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], Campbell R. Davies[5].

1 Fisheries Centre, AERL, 2202 Main Mall, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z4

2 International Commission for the Conservation of Atlantic Tunas, Calle Corazón de María, 8, 28002 Madrid, 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, Australia, 7000.

6 Centre for Environmental Policy, Imperial College London, Silwood Park, Buckhurst Road, Ascot, Berkshire SL5 7PY, U.K.

7 Tokyo University of Marine Science and Technology, 5-7, Konan 4, Minato-ku, Tokyo 108-8477 Japan

Corresponding Author

Thomas R Carruthers

Tel: (+1) 604 822 6903

Fax: (+1) 604 822 8934

Email: t.carruthers@fisheries.ubc.ca

Abstract

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 auto-correlation in recruitment strength. We quantify the robustness of each management procedure 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. Management procedure performance is most sensitive to biases in catch data. Our results indicate that in many cases 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

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

Introduction

Management advice for the majority of economically important fish stocks, and increasingly also for by-catch 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 amongst 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 on an annual basis (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, aka, 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 paper 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*. Southern Bluefin Tuna - CCSBT, 2011; South African hake - Rademeyer *et al*., 2008; Namibian Hake – Butterworth and Geromont 2001; 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[1]. 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*. Geromont and Butterworth, 2015b; Maunder, 2014).

In this paper 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

[1] 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 more than 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.

also include new approaches that operate on alternative fishery information. A number of the MPs of this paper 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 paper 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 paper 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 management procedures

In this paper 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, 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 paper (Rcontrol, Rcontrol2, Gcontrol, SPmod, SPslope, Fadapt, and DynF) rely on the same calculation of surplus production *S*:

(1) $$S_y = B_{y+1} - B_y + C_y$$

The derivative of surplus production with respect to biomass *G* (*i.e.* d*S*/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 the stock is below that size, while *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*.

Ten additional generic rules are included that were proposed by Geromont and Butterworth (2015b; CC1, CC4, LstepCC1, LstepCC4, Ltarget1, Ltarget4, Islope1, Islope4, Itarget1, 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 tuned-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 an historical average CPUE.

All of 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 2014b describe four constant catch MPs: CC1, CC2, CC3 and CC4, which set 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 paper; the equations of the generic MPs are presented in Table 2.

Tuned management procedures

In many MSE applications MPs are tuned through simulation testing to achieve the pre-specified 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 paper by Geromont and Butterworth (2015a) who identified MPs for data-rich fisheries.

In this paper 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). In order 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

In order 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 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}.

Depletion Corrected Average Catch (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 the Appendix A.1.). DCAC is currently used by the Pacific Fishery Management Council to set catch-limits for 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 all of the generic MPs, the simpler Fratio MP provides a useful comparison.

Operating model structure and simulation design

The operating model 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 Appendix B).

We constructed operating models using a factorial design encompassing 36 sets of operating model assumptions. The four factors were (1) life history with three levels, (2) temporal autocorrelation in recruitment with two levels, (3) starting stock depletion with two levels and (4) data quality with three levels (Table 4). 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 management procedure. 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 assessed level of stock depletion was not simulated 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 2.5-15% of unfished levels (less than $B_{MSY}/2$) and an overexploited scenario in which spawning biomass is between 15-35% of unfished levels (between $B_{MSY}/2$ and B_{MSY}).

Autocorrelation in recruitment was simulated in order to evaluate the performance of the MPs in situations where stock productivity varies over time (an AR1 process, App. Eqn.B.1). This may not fully reflect step-changes in recruitment that have been 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 fundamental 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 of the inputs to the MPs (*e.g.*, natural mortality rate, observations of absolute biomass) (see App. Table C.1 which includes a summary of the observation error model). We simulated three different 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 but 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 4). For example, a value of 0.82 could be drawn for the first simulation and be used as consistent 18% downward bias for all observed catches in that time-series. In addition to this bias we superimpose imprecision that is log-normal 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 in order to simulate the poorer quality of data. Biases were generated from log-normal 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 3).

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 of the operating models and management procedures evaluated in this paper (computer code for reproducing our methods is available online at https://github.com/tcarruth/Carruthers-et-at-2015-MP-MSE).

Performance criteria

Performance was summarized by five metrics: yield, average annual variability in yield, fishing mortality rate, spawning stock biomass and spawning stock biomass 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 Average Annual Variability in Yield (AAVY). AAVY is the mean difference in the yield of adjacent projected years (starting from the last historical year) divided by the mean yield over the same time period.

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

where n_p is the number of projected years, n_h is the number of historical years, 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 spawning stock biomass over 50% *MSY* levels and yields greater than 50% *MSY* levels etc., Table 5).

In specific management contexts, a set of performance metrics for a management procedure 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 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, observation error in catches σ_C) was sampled from a range of values that were considered to be credible *a priori* (Table 4 and Table App.B.1). 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, Figure 2). In the case of 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 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[2].

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] 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.

3 – 5). 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 given the herring and bluefin tuna life histories (Figures 3 and 5). In contrast the MPs DD, DD4010, DynF, Islope4, Itarget1, Itarget4, LstepCC4 and Ltarget4 were the most likely to lead to stock rebuilding for these two life history types. In most cases, simulating better quality data quality improved the probability of reaching B_{MSY}. Exceptions include the Fadapt and CC1 rules. In the case of 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, that sets the TAC to mean historical catches, generally led to a low probability of rebuilding. In many cases 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 around 40% less likely to rebuild over 15 years given poor quality data.

Figure 6 illustrates the trade-off between the probability of achieving more than half *MSY* catches over the last ten 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 under-exploitation. 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, 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 60-80% probability of returning yields in excess of half *MSY* (Figure 6) and these rebuilt stocks are achieved in less than 70% of simulations (Figure 4).

Overexploited scenario

When starting from a stock depletion of between 50% and 100% B_{MSY}, many MPs satisfied the spawning stock biomass and fishing mortality rate performance targets but failed to meet the yield target, for example CC4, LstepCC4, Islope4 and Itarget4 (Figures 7 – 9). 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, Fratio). Some MPs which performed amongst the best in the overexploited scenario, such as DCAC, performed amongst 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 versus SSB) was similar regardless of life-history type (Figures 7 - 9). 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 in the case of herring and bluefin tuna (Figures 7 and 8, respectively) but not rockfish (Figure 9). 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 by a much larger margin for a greater number of operating models in the case of the rockfish life-history. In almost all instances the simple MPs achieved the AAVY target (greater than a 60% chance of AAVY less than 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 in comparison to stock assessments. 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 to management under traditional stock assessment paradigm.

Value of information and sensitivity analysis

For most MPs the observation process that most strongly determined long-term yield (average yield over the final 10 years of the projection) was bias in observation of annual catches (Table 7). The sensitivity to bias in catch observations varied amongst MPs. For example given negative bias in annual catch observations for herring ('catch –', Table 7) the SBT2 MP led to a 90% reduction in yields from the reference level compared with a 45% reduction in the case of Islope4. The generic MPs were sensitive to other observation model variables that impact calculations of surplus production. 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 generic MPs that are tested here for the first time such as DynF are sensitive to high imprecision in catches ('catch high', Table 7) and the relative abundance index ('index high') arising from the calculation of surplus production (Eqn. 1).

With the exception of methods such as DynF and Fadapt that do not rely on annual catch observations, positive bias in catches led to much more depleted stocks than compatible with the target reference levels. In many instances catches that were inflated to between 1.5 and 2 times their true value lead to chronic stock declines (*i.e*. a 100% decline in depletion relative to unbiased reference levels, Table 7). Even though higher levels of bias were simulated for observations of absolute biomass (Figure 2) methods using these data such as DynF and Fadapt led to less pronounced stock declines (mean depletion declines between 50% and 80%).

Discussion

The simulations indicate that the absolute performance of the MPs considered can vary widely among the three life-history types. Performance rankings however, were relatively constant and the five best and worst performing methods remained largely the same across life-history types. This apparent 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. While methods such as 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 for 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 auto-correlation, is not appropriate for mimicking regime shifts that are characterised by abrupt and 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 MPs were generic (*e.g.* DynF, Fadapt) or reference MPs (*e.g.* DD, Fratio) that require inputs for 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 five-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*, Eqn. 1) to update the TAC, generally performed poorly (Gcontrol, Rcontrol, Rcontrol2, SPslope, SPmod, 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 (a) data-rich approaches using time series of catch and effort data and (b) 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 could often out-perform the more data intensive delay-difference assessment over a range of data-qualities, depletion levels and life-history types. This suggests that if an index of current absolute biomass 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 to 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, 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 less than 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). It is possible that 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 (*e.g*. Table 7) 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 amongst 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 more than 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 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.

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 paper) was one of a suite of candidates – biomass dynamic, empirical, 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. Secondly, because all of 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 to initial depletion levels. 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 even though 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 focussing 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 of 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 paper, and it is important 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 organisations are now expected to provide catch-limits for all fisheries in a fishery management plan with the exception of some short-lived stocks (*e.g.* USA, Australia, New Zealand). For this reason we have focused in this paper 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.

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 paper 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., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science, 56: 985-998.

Butterworth, D.S., 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(1): 357-374.

Carruthers, T.R. 2015. DLMtool: Data-Limited Methods Toolkit. v1.35. http://cran.r-project.org/web/packages/DLMtool/index.html [accessed 5/4/2015]

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

CCSBT. 2011. Report of the Sixteenth Meeting of the Scientific Committee. Bali, Indonesia. Commission for the Conservation of Southern Bluefin Tuna. Available at: http://www.ccsbt.org/userfiles/file/docs_english/meetings/meeting_reports/ccsbt_18/report_of_SC16.pdf [accessed 6/5/2015]

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. Reviews in Fish Biology and Fisheries, 8: 177-214.

Coelho, M., Filipe, J.A., Ferreira, M.A.M. 2013. Modelling enforcement and compliance in fisheries: a survey. International Journal of Latest Trends in Finance and Economic Sciences, 2(3): 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(1): 19-30.

DFO. 2012. Stock assessment report on Pacific herring in British Columbia in 2012. Fisheries and Oceans Canada. Science Advisory Report 2012/062. Available at: http://www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2012/2012_062-eng.pdf [accessed 30/9/2014]

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., 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. Available at: 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. Available at: http://www.iattc.org/Meetings/Meetings2014/MAYSAC/PDFs/SAC-05-10b-Management-Strategy-Evaluation.pdf [accessed 3/10/2014]

McAllister, M.K., Pikitch, E.K., 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(9): 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 [accessed 6/5/2015]

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., 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., 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. Url: http://www.R-project.org.

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., S.R. Sloan. 2013. Implementing harvest strategies in Australia: 5 years on. ICES Journal of Marine Science. doi:10.1093/icesjms/fst158.

Vert-pre, K.A., Amoroso, R.O., Jensen, O.P., Hilborn, R. 2013. Frequency and intensity of productivity regime shifts in marine fish stocks. Proceeding of the National Academy of Sciences, U.S.A., 110(5): 1779-1784.

Wallace, J.R. and Cope, J.M. 2011. Status update of the U.S. canary rockfish resource in 2011. Report of the National Marine Fisheries Service. Available at: http:// www.pcouncil.org/ wp-content/ uploads/ Canary_2011_Assessment_Update.pdf [accessed 3/10/2014]

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, NJ: Princeton University Press.

Wetzel, C.R., 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. doi: 10.1016/j.fishres.2011.04.024

Zeller D., Booth, S., Davis, G., Pauly, D. 2007. Re-estimation of small-scale fishery catches for U.S. flag-associated island areas in the western Pacific: the last 50 years. Fisheries Bulletin, 105(2): 266-277.

Zeller D., Harper, S., Zylich, K., Pauly, D. 2015. Synthesis of under-reported small-scale fisheries catch in Pacific-island waters. Coral Reefs, 34(1): 25-39.

Tables

Table 1. Overview of the management procedures and their data requirements.

Type	Name	Description	Data requirements / inputs	References
Generic MPs	Gcontrol	Historical trend in inferred surplus production with 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 one year of biomass and catch to update the TAC		
	DynF	The TAC is set according to a fixed target fishing rate and varies according to the recent trend in surplus production with biomass	Recent catch, recent absolute biomass, natural mortality rate, ratio of F_{MSY} to natural mortality rate	
	Fadapt	The TAC is set according to a walk in target fishing rate that is modified according to recent trend in surplus production with biomass.		
	CC1	The TAC is recent catch levels	Recent catch	Geromont and Butterworth 2015b
	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 in order 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 in order 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 2015a
	GB_slope	Incremental changes in TAC are made in order to maintain a constant relative abundance index	Recent catch, recent relative abundance index	
	GB_target	Incremental changes in TAC are made in order 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 downward 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 $_{FMSY}$ 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 management procedures. *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).

MP Name	Parameter values	TAC calculation
Gcontrol	$g^U = 0.5$ $g^L = 2$	$TAC_y = \begin{cases} TAC_y^{try} & g^L < TAC_y^{try} < g^U \\ g^L \bar{C}_y & TAC_y^{try} < g^L \\ g^U \bar{C}_y & g^U < TAC_y^{try} \end{cases} \quad , \quad TAC_y^{try} = S_y\left(1 - 2G_y\right)$ G_y is the slope in surplus production S with biomass over the last ten 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 ten years
Rcontrol		As Gcontrol except: $G_y = r\left(1 - 2D_y\right)$ where r is the demographically derived prior for 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		$TAC_y = \begin{cases} \left[1 - \left(\bar{B}_{y-4} - \bar{B}_y\right)/\bar{B}_{y-4}\right] C_y^{ave} & \Delta^B < 9/10 \\ \frac{9}{10} S_{y-1} & \Delta^B > 11/10 \\ TAC_{y-1} & 9/10 < \Delta^B < 11/10 \end{cases}$ $\Delta^B = \bar{B}_y/\bar{B}_{y-4} \quad , \quad S_y = \bar{B}_y - \bar{B}_{y-1} + \bar{C}_{y-1} \quad , \quad C_y^{ave} = 1/4 \sum_{t=y-3}^{y} C_t$
SPmod		$TAC_y = \begin{cases} 4C_y/5 & \Delta^B < 4/5 \\ 6S_{y-1}/5 & \Delta^B > 6/5 \\ C_{y-1} & 4/5 < \Delta^B < 6/5 \end{cases} \quad , \quad \Delta^B = \frac{B_y}{B_{y-1}} \ , \quad S_y = B_y - B_{y-1} + C_{y-1}$
DynF	$F^L = \frac{FMSY}{2}$	$TAC_y = \bar{F}_y \bar{B}_y \quad , \quad \bar{F}_y = \begin{cases} F_y^{try} & F^L < F_y^{try} < F^U \\ F^L & F_y^{try} < F^L \\ F^U & F^U < F_y^{try} \end{cases} \quad , \quad F_y^{try} = F_{y-1} e^{-2G_y}$ As Gcontrol, G_y is the derivative of S with respect to biomass over the last 7 years.
Fadapt	$F^U = 2 \cdot FMSY$ $FMSY = M\left(\frac{FMSY}{M}\right)$	$TAC_y = \bar{F}_y \bar{B}_y \quad , \quad \bar{F} = F^L + logit^{-1}\left(W_y - G_y\right)\left(F^U - F^L\right)$ $W_y = \begin{cases} logit\left(\frac{F_y^{ave} - F^L}{F^U - F^L}\right) & F^L < F_y^{ave} < F^U \\ -2 & F_y^{ave} < F^L \\ 2 & F^U < F_y^{ave} \end{cases}$ As Gcontrol, G_y is the derivative of S with respect to biomass over the last 7 years.

Table 2 continued.

MP Name	Parameter values	TAC calculation
CC1	$x = 1$	$TAC_{y+1} = x/5\sum_{t=y-4}^{y} TAC_t$
CC4	$x = 0.7$	
LstepCC1	$x = 1$	$TAC_{y+1} = \begin{cases} TAC_y + \frac{1}{20}\bar{C}_{\,y+1} & 1.05 < L_y^{rec}/L_y^{ave} \\ TAC_y & 0.98 < L_y^{rec}/L_y^{ave} < 1.05 \\ TAC_y - \frac{1}{20}\bar{C}_{\,y+1} & 0.96 < L_y^{rec}/L_y^{ave} < 0.98 \\ TAC_y - \frac{1}{10}\bar{C}_{\,y+1} & L_y^{rec}/L_y^{ave} < 0.96 \end{cases}$
LstepCC4	$x = 0.7$	$TAC_{y*} = x\bar{C}_{y*}$, $\bar{C}_{y+1} = 1/5\sum_{t=y-4}^{y} C_t$
Ltarget1	$x = 1$ $v = 1.05$	$TAC_{y+1} = \begin{cases} \frac{TAC_{y*}}{2}\left[1 + \left(\frac{L_y^{rec} - L^0}{L^{targ} - L^0}\right)\right] & L_y^{rec} \geq L^0 \\ \frac{TAC_{y*}}{2}\left[\frac{L_y^{rec}}{L^0}\right]^2 & L_y^{rec} < L^0 \end{cases}$, $\overline{TAC}_{y*} = x/5\sum_{t=y*-4}^{y*} TAC_t$
Ltarget4	$x = 0.8$ $v = 1.15$	$L^0 = \frac{9}{10}L_{y*}^{ave}$ $L_y^{rec} = \frac{1}{5}\sum_{t=y-4}^{y}\sum_{i=1}^{n_t} L_{t,i}^{obs}/n_t$, $L_y^{ave} = \frac{1}{10}\sum_{t=y-9}^{y}\sum_{i=1}^{n_t} L_{t,i}^{obs}/n_t$ $L^{targ} = vL_y^{ave}$
Islope1	$x = 0.8$ $\lambda = 0.4$	$\overline{TAC}_{y*} = x/5\sum_{t=y*-4}^{y*} TAC_t$, $TAC_{y+1} = TAC_y(1 + \lambda s_y)$
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	$x = 1$ $v = 1.5$	$TAC_{y+1} = \begin{cases} \frac{\overline{TAC}_{y*}}{2}\left[1 + \left(\frac{I_y^{rec} - I^0}{I_y^{targ} - I^0}\right)\right] & I_y^{rec} \geq I^0 \\ \frac{\overline{TAC}_{y*}}{2}\left[\frac{I_y^{rec}}{I^0}\right]^2 & I_y^{rec} < I^0 \end{cases}$, $\overline{TAC}_{y*} =$ $x/5\sum_{t=y*-4}^{y*} TAC_t$,
Itarget4	$x = 0.7$ $v = 2.5$	$I_y^{rec} = \frac{1}{5}\sum_{t=y-4}^{y} I_t$, $I^{ave} = \frac{1}{10}\sum_{t=y*-9}^{y*} I_t$, $I^0 = \frac{4}{5}I_y^{ave}$, $I_y^{targ} = v\, I_y^{ave}$

Table 3. The equations of the tuned management procedures. *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. *Robs:* an estimate of recruitment strength. *MSY:* Maximum Sustainable Yield (though this is not known perfectly).

MP Name	TAC calculation
SBT1	$TAC_y = TAC_{y-1} \cdot \begin{cases} 1 + 2G_y/3 & G_y < 0 \\ 1 + 3G_y & G_y \geq 0 \end{cases}$ where G_y is the derivative of log CPUE with respect to time over the last 10 years
SBT2	$TAC_y = \frac{1}{2}TAC_{y-1} + \delta MSY \quad , \quad \delta = \begin{cases} \Delta^{7/4} & \Delta < 1 \\ \Delta^{1/4} & \Delta > 1 \end{cases} \quad , \quad \Delta = R^{ave}/R^{hist}$ $R_y^{ave} = \frac{1}{5}\sum_{t=y-4}^{y} Robs_t \quad , \quad R_y^{hist} = \frac{1}{10}\sum_{t=y-9}^{y} Robs_t$
GB_CC	$TAC_y = \begin{cases} 4C^{ave}/5 & MSY < 4C^{ave}/5 \\ C^{ave} & 4C^{ave}/5 < MSY < 6C^{ave}/5 \\ 6C^{ave}/5 & 6C^{ave}/5 < MSY \end{cases}$ where C^{ave} is the mean historical annual catch
GB_slope	$TAC_y = \begin{cases} 4C^{ave}/5 & TAC_y^{try} < 4C^{ave}/5 \\ TAC_y^{try} & 4C^{ave}/5 < TAC_y^{try} < 6C^{ave}/5 \\ 6C^{ave}/5 & 6C^{ave}/5 < TAC_y^{try} \end{cases} \quad , \quad TAC_y^{try} = C^{ave}(1 + G_y)$ 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	$TAC_y = \begin{cases} 4C^{ave}/5 & TAC_y^{try} < 4C^{ave}/5 \\ TAC_y^{try} & 4C^{ave}/5 < TAC_y^{try} < 6C^{ave}/5 \\ 6C^{ave}/5 & 6C^{ave}/5 < TAC_y^{try} \end{cases}$ $TAC_y^{try} = \begin{cases} MSY\left(\frac{1}{2} + \frac{I_y^{rec} - I^0}{2(I^{targ} - I^0)}\right) & I^0 \leq I_y^{rec} \\ \frac{MSY}{2}\left(\frac{I_y^{rec}}{I^0}\right)^2 & I_y^{rec} < I^0 \end{cases} \quad , \; I^0 = \frac{1}{25}\sum_{t=y*-4}^{y*} I_t \; , \; I_y^{rec} = \frac{1}{4}\sum_{t=y-3}^{y} I_t$ 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

Table 4. Overview of simulation design. Parameter ranges represent the lower and upper bounds of a uniform random variable. SD refers to the standard deviation. In the case of 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, 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.

MSE attribute	**Symbol**	**Case 1**	**Case 2**	**Case 3**
Life history		**Herring**	**Bluefin tuna**	**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 log-normal 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 (log-normal SD)	Υ_D	none	0.2	0.5
Bias in annual catches (log-normal SD)	Υ_C	none	0.2	0.5
Observation error in annual catches (log-normal SD)	σ_C	none	0.2 - 0.4	0.3 - 0.6
Observation error in relative abundance index (log-normal SD)	σ_I	none	0.1 - 0.3	0.2 - 0.6
Observation error in recruitment (log-normal SD)	σ_{Robs}	none	0.05 - 0.1	0.1 - 0.3
Hyperstablity / hyperdepletion in index	β	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 (log-normal SD)	σ_B	none	0.1 - 0.5	0.5 - 1
Bias in natural mortality rate (log-normal SD)	Υ_M	none	0.25	0.5
Bias in von Bertalanffy K parameter (log-normal SD)	Υ_K	none	0.05	0.1
Bias in B_{msy} relative to unfished (log-normal SD)	Υ_{Bpeak}	none	0.1	0.2
Bias in age at maturity (log-normal SD)	Υ_{Am}	none	0.1	0.2
Number of annual age / length observations	n_{CAA}	10000-15000	200 - 500	50 -100
Other control rule inputs				
Bias in ratio of F_{MSY} to M (log-normal SD)	Υ_{FMSY_M}	none	0.25	0.5
Bias in target CPUE (B_{MSY})(log-normal SD)	Υ_{Bmsy}	none	0.3	1
Bias in target catch (MSY)(log-normal SD)	Υ_{MSY}	none	0.2	0.3

Table 5. Performance metrics of this simulation evaluation. Targets are probabilities of achieving a reference level. Note that MP performance is considered acceptable if all targets are met in combination.

Performance metric	**Reference level**	**Target**
Spawning Stock Biomass (SSB)	SSB is above 50% *MSY* levels evaluated over all projected years	> 60%
Long-term Spawning Stock Biomass (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%
Average Annual Variability in Yield (AAVY)	AAVY is less than 30%	> 70%

Table 6. Specification of robustness tests. Reference regions are 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 (Figure 2).

Parameter controlling data quality	**Symbol**	**Ranges for observation error level / bias**		
		Reference	**Low / -**	**High / +**
Bias in biomass relative to unfished (log-normal SD)	D	3/4 - 4/3	1/3 - 1/2	2 - 3
Bias in annual catches (log-normal SD)	Υ_C	3/4 - 4/3	1/2 - 2/3	3/2 - 2
Observation error in annual catches (log-normal SD)	σ_C	0.35 - 0.45	0.2 - 0.3	0.5 - 0.6
Observation error in relative abundance index (log-normal SD)	σ_I	0.35 - 0.45	0.2 - 0.3	0.5 - 0.6
Hyperstablity / hyperdepletion in index	β	3/4 - 4/3	1/2 - 2/3	3/2 - 2
Bias in target CPUE (BMSY)(log-normal SD)	Υ_{CPUE}	3/4 - 4/3	1/2 - 2/3	3/2 - 2
Bias in target catch (MSY)(log-normal 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 (log-normal SD)	σ_B	0.6 - 0.7	0.5 - 0.6	0.7 - 0.9
Bias in ratio of FMSY to M (log-normal SD)	Υ_{FMSY_M}	3/4 - 4/3	1/2 - 2/3	3/2 - 2
Bias in natural mortality rate (log-normal SD)	Υ_M	3/4 - 4/3	3/4 - 4/3	3/2 - 2

Table 7. The robustness of yields to bias and imprecision in the inputs to MPs for the data-poor, overexploited scenario with low auto-correlation in recruitment. For each MP and life-history type results for the three most critical sources of imperfect information are reported. Numbers represent relative yields and are accompanied by text description of the associated observation and the direction of bias / level of precision. Relative yields are the percentage change in the reference yield. 'Catch', 'Biomass', 'Index' and 'Depln.' refer to annual observations of catch, absolute biomass, a relative abundance index and stock depletion, respectively. 'von B κ' and 'M' refer to the von Bertalanffy growth rate parameter and natural mortality rate. 'F_{MSY}/M' and 'B_{MSY}/B_0' are inputs for the ratios of F_{MSY} to natural mortality rate and B_{MSY} relative to unfished levels. The symbols '-' and '+' refer to negative and positive bias in observations. The text 'high' and 'low' refer to high and low observation error. For example 'Catch –' results refer to a set of contingency simulations for which negative bias in catches was simulated, 'Index low' are contingency simulations where low observation error in relative abundance indices was simulated (Table 6). For each MP and life-history the three most critical data issues are arranged in order of descending sensitivity (left to right) and color-coded with respect to reductions in yield with red representing the greatest loss of yields. In instances where an MP did not operate, the percentage change in yield is replaced with 'DNO'. Increases in yield are shaded green, reductions in yield of 80% or more are shaded red.

MP	Herring			Bluefin			Rockfish		
Gcontrol	Catch -	Catch +	Biomass -	Catch -	Catch +	Biomass +	Catch +	Catch -	Index high
	-89	-79	-27	-92	-79	69	-100	-85	-42
Rcontrol	Catch +	Catch -	Depln. +	von B k +	Catch -	Catch +	Catch +	von B k -	Catch -
	-85	-82	-60	-99	-89	-65	-100	-89	-84
Rcontrol2	von B k +	Catch +	Depln. +	von B k +	von B k -	Catch -	Catch +	von B k +	Biomass -
	-95	-90	-70	-100	-83	-79	-100	2318	-82
SPslope	Catch +	Index high	Catch -	Catch +	Biomass -	Catch low	Biomass +	Catch -	Catch +
	DNO	-71	-49	-82	153	-48	DNO	DNO	-79
SPmod	Catch +	Catch high	Catch -	Catch +	Catch -	Index low	Catch +	Biomass -	Biomass +
	-76	154	128	-79	164	91	-100	1208	-90
DynF	Biomass +	Catch +	Catch high	Biomass +	FMSY/M +	Index high	Biomass +	FMSY/M -	FMSY/M +
	-42	45	36	-44	-28	-23	-41	28	-16
Fadapt	Biomass +	Catch high	Catch +	Biomass +	FMSY/M -	M +	Biomass +	Biomass -	FMSY/M -
	-55	22	19	-47	23	-17	-43	-27	18
CC1	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high
	-89	-73	83	-92	-85	10	-100	-89	-30
CC4	Catch -	Catch +	Catch high	Catch -	Catch high	Catch +	Catch +	Catch -	Catch high
	-99	-85	-21	-99	-25	-15	-100	-99	-32
LstepCC1	Catch +	Catch -	Catch low	Catch +	Catch low	Catch high	Catch +	Catch -	Catch high
	-70	36	-15	-85	20	6	-100	69	-9
LstepCC4	Catch +	Catch -	Catch high	Catch +	Catch -	Catch low	Catch +	Catch -	Catch low
	-47	-34	6	-57	-47	-20	-58	-31	26
Ltarget1	Catch -	Catch +	Catch low	Catch -	Catch +	Catch low	Catch +	Catch -	Catch low
	-94	-87	49	-98	-90	8	-100	-98	41
Ltarget4	Catch -	Catch +	Catch high	Catch -	Catch +	Catch low	Catch -	Catch +	Catch low
	-100	189	39	-100	505	90	-99	1345	100
Islope1	Catch +	Catch high	Catch -	Catch +	Catch -	Index low	Catch +	Index high	Catch high
	-46	10	-9	-75	-40	-14	-85	-16	5
Islope4	Catch -	Catch +	Index high	Catch -	Index low	Catch +	Catch -	Catch +	Index high
	-45	-27	-12	-52	-22	-17	-40	-36	-13
Itarget1	Catch -	Index low	Catch high	Catch -	Catch +	Index low	Catch -	Catch +	Index high
	-99	30	-19	-99	121	-25	-99	342	-31
Itarget4	Catch -	Catch +	Catch low	Catch -	Catch +	Catch high	Catch -	Catch +	Catch high
	-99	4308	47	-99	5791	-82	-99	6069	-84
SBT1	Catch +	Catch -	Catch high	Catch -	Catch +	Catch low	Catch +	Catch -	Index low
	-89	-59	39	-91	-89	-15	-100	-99	193
SBT2	Catch +	Catch high	Catch -	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high
	-89	67	48	-92	64	-21	-75	29	-15
GB_CC	Catch +	Catch -	Catch low	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high
	-91	-78	28	-93	-85	-24	-100	-95	-33
GB_slope	Catch +	Catch -	Catch high	Catch -	Catch +	Catch low	Catch +	Catch -	Index low
	-86	-85	47	-93	-90	-11	-100	-99	109
GB_target	Catch +	Catch -	Catch low	Catch +	Catch -	Index high	Catch -	Catch +	Index low
	-91	-65	34	-92	-92	-24	-95	-95	38
DD	von B k +	Catch +	von B k -	von B k +	Catch +	Catch high	von B k +	Catch +	M -
	98	-46	-27	-57	-37	-29	-60	-53	31
DD4010	von B k +	Catch +	Catch low	Catch -	Index low	Index high	von B k +	von B k -	Catch -
	80	-30	30	50	-27	-19	125	-50	-42
DCAC	BMSY/B0 +	BMSY/B0 -	FMSY/M +	FMSY/M -	BMSY/B0 -	BMSY/B0 +	M +	FMSY/M -	BMSY/B0 +
	-29	14	-12	22	17	14	-19	15	-13
Fratio	Biomass +	Biomass -	FMSY/M +	Biomass +	Biomass -	M +	Biomass -	Biomass +	FMSY/M +
	-34	-22	-4	-34	-18	-9	-34	-25	-4

Table 8. The robustness of projected stock biomass to bias and imprecision in the inputs to MPs for the data-poor, overexploited scenario with low auto-correlation in recruitment. As for Table 7 but figures relate to percentage changes in biomass at the end of the projection relative to reference simulations. Increases in biomass are color coded green, reductions in biomass of 80% or more are color coded red.

MP		**Herring**			**Bluefin**			**Rockfish**	
Gcontrol	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high	Catch +	Catch -	Biomass +
	-92	147	79	-94	201	69	-100	173	-27
Rcontrol	Catch +	Depln. +	Catch -	von B k +	Catch +	von B k -	Catch +	von B k -	Catch -
	-75	-68	172	-100	-90	-62	-100	-77	173
Rcontrol2	Catch +	von B κ +	Biomass -	Catch +	Catch -	von B k -	Catch +	von B k +	von B k -
	-91	146	-59	-99	86	50	-100	-80	-69
SPslope	Catch +	Index high	Index low	Catch +	Catch -	Biomass -	Biomass +	Catch -	Catch +
	DNO	-77	-63	-92	265	241	DNO	DNO	-82
SPmod	Catch +	Catch high	Biomass +	Catch +	Catch -	Index low	Catch +	Biomass -	Catch -
	-76	249	104	-88	462	282	-100	199	150
DynF	Biomass +	Catch +	FMSY/M +	Biomass +	FMSY/M +	Catch high	Biomass +	Biomass -	M +
	-55	-41	-36	-63	-41	39	-78	54	-31
Fadapt	Biomass +	Catch +	M +	Biomass +	FMSY/M +	Biomass -	Biomass +	Biomass -	M +
	-63	-55	-42	-74	-37	42	-75	52	-32
CC1	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high	Catch +	Catch -	Catch low
	-93	425	85	-99	480	34	-100	290	-10
CC4	Catch +	Catch high	Catch -	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high
	-98	39	22	-97	25	10	-100	38	-5
LstepCC1	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high	Catch +	Catch -	Catch low
	-74	64	37	-93	115	31	-100	181	-18
LstepCC4	Catch +	Catch high	Catch low	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high
	-66	27	10	-78	20	14	-91	63	-8
Ltarget1	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high	Catch +	Catch -	Catch low
	-96	123	56	-99	104	29	-100	93	-8
Ltarget4	Catch +	Catch high	Catch -	Catch +	Catch high	Catch low	Catch +	Catch -	Catch low
	-83	18	-8	-84	15	7	-86	12	4
Islope1	Catch +	Catch -	Catch high	Catch +	Catch -	Index low	Catch +	Catch -	Index high
	-68	33	28	-90	51	36	-98	108	29
Islope4	Catch +	Catch high	Catch low	Catch +	Index low	Catch high	Catch +	Catch -	Index high
	-61	23	10	-65	22	17	-83	47	13
Itarget1	Catch +	Catch high	Catch -	Catch +	Index low	Catch -	Catch +	Catch -	Index high
	-86	24	9	-88	21	14	-84	19	8
Itarget4	Catch +	Catch -	Catch high	Catch +	Catch high	Index low	Catch +	Catch -	Catch high
	-25	-14	16	-23	14	14	-24	5	-4
SBT1	Catch +	Catch -	Catch high	Catch +	Catch -	Index low	Catch +	Catch -	Index low
	-95	289	128	-99	231	65	-100	189	-22
SBT2	Catch +	Catch high	Catch -	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high
	-94	117	95	-99	110	22	-92	92	-12
GB_CC	Catch +	Catch -	Catch high	Catch +	Catch -	Catch high	Catch +	Catch -	Catch low
	-95	217	119	-99	270	40	-100	198	-14
GB_slope	Catch +	Catch -	Catch high	Catch +	Catch -	Index low	Catch +	Catch -	Index low
	-96	252	104	-99	277	58	-100	180	-27
GB_target	Catch +	Catch -	Catch high	Catch +	Catch -	Index low	Catch +	Catch -	Index low
	-94	252	124	-98	309	65	-97	189	-25
DD	Catch +	von B k -	Catch -	von B κ +	Catch +	von B k -	von B k -	Catch +	Catch -
	-81	119	62	-90	-80	-46	-84	-55	58
DD4010	Catch +	Catch -	Index low	Catch +	von B k +	Index low	Catch +	Catch -	von B k +
	-38	28	-21	-58	-34	40	-30	29	-22
DCAC	BMSY/B0 +	M +	M -	M -	FMSY/M -	FMSY/M +	M +	M -	FMSY/M -
	34	-21	18	24	21	-16	-25	33	26
Fratio	Biomass +	FMSY/M +	M +	Biomass +	FMSY/M +	M +	Biomass +	M +	Biomass -
	-56	-34	-33	-67	-35	-29	-70	-33	37

Figures

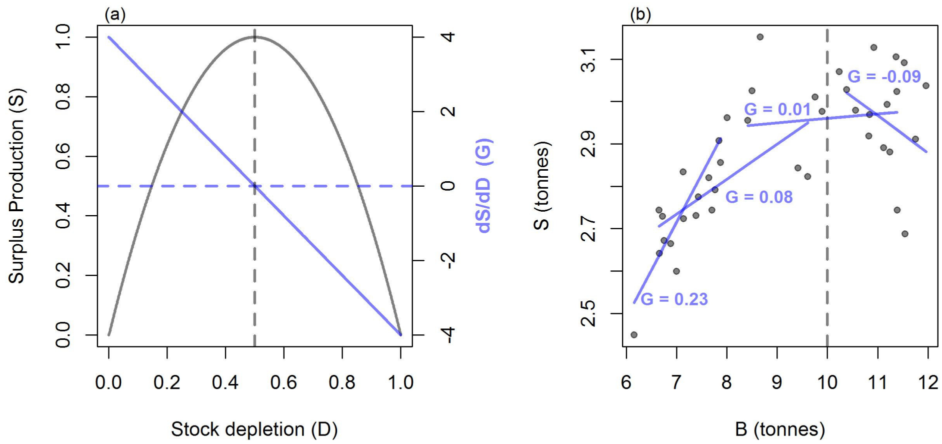


Figure 1. The theoretical model of derivative (*G*) in surplus production (*S*) with respect to stock depletion (*D,* biomass relative to the unfished level) according to the Schaefer production model (panel a). The vertical dashed line represents the simulated level of biomass at the most productive stock size. Observations of catch and biomass (*B*) may be used to infer surplus production (*S*): $S_y=B_y-B_{y-1}+C_{y-1}$ (panel 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). Panel b illustrates estimates of *G* (grey lines) for four simulated time periods.

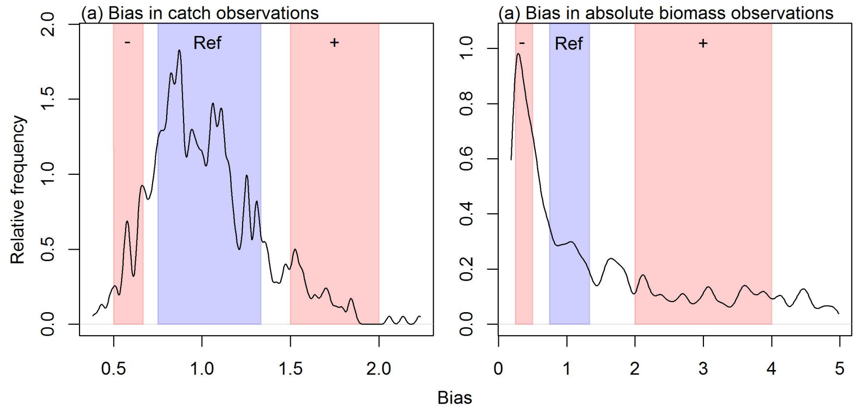


Figure 2. Relative frequency of simulated biases that were used to generate imperfect observations of catch and absolute biomass. Blue shaded areas represent the reference region (Ref) in which relative unbiased data were simulated. Red shaded areas represent contingency regions of low (-) and high (+) bias. The robustness of MPs was evaluated by comparing their performance among these simulated regions (Table 6).

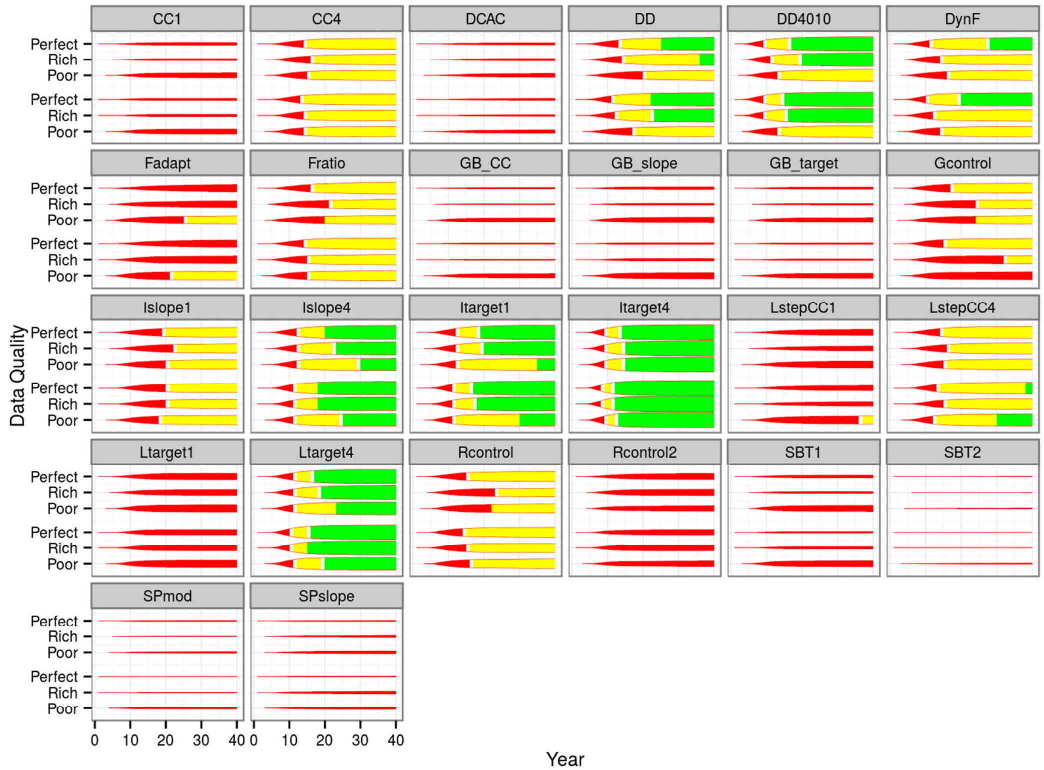
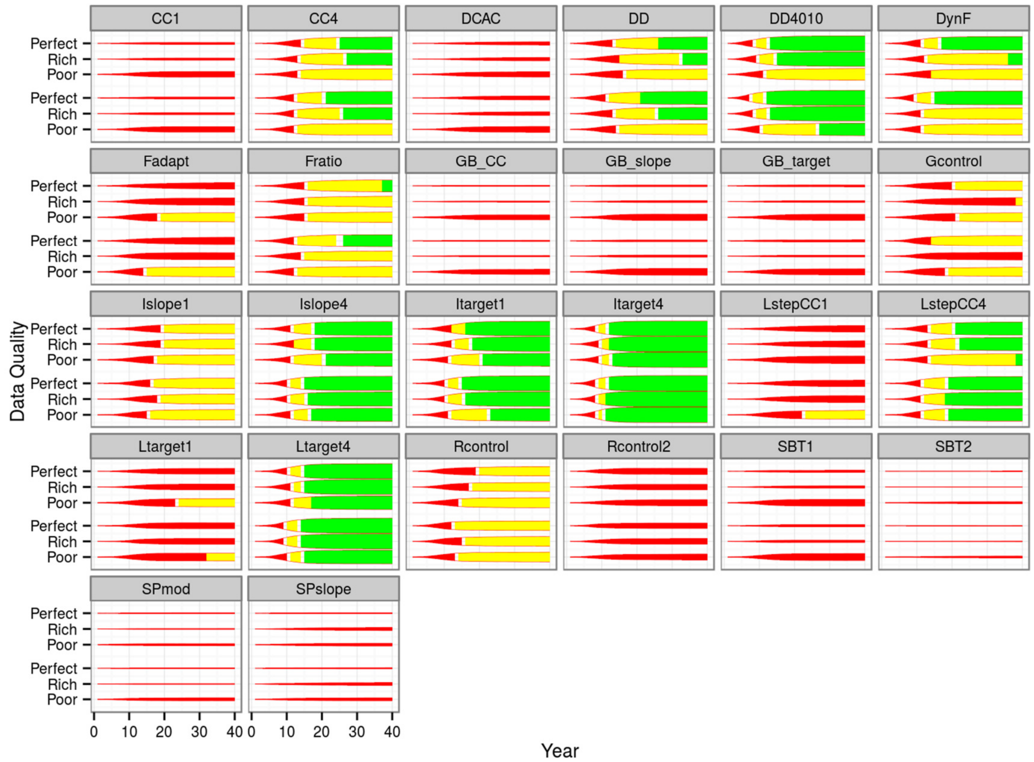


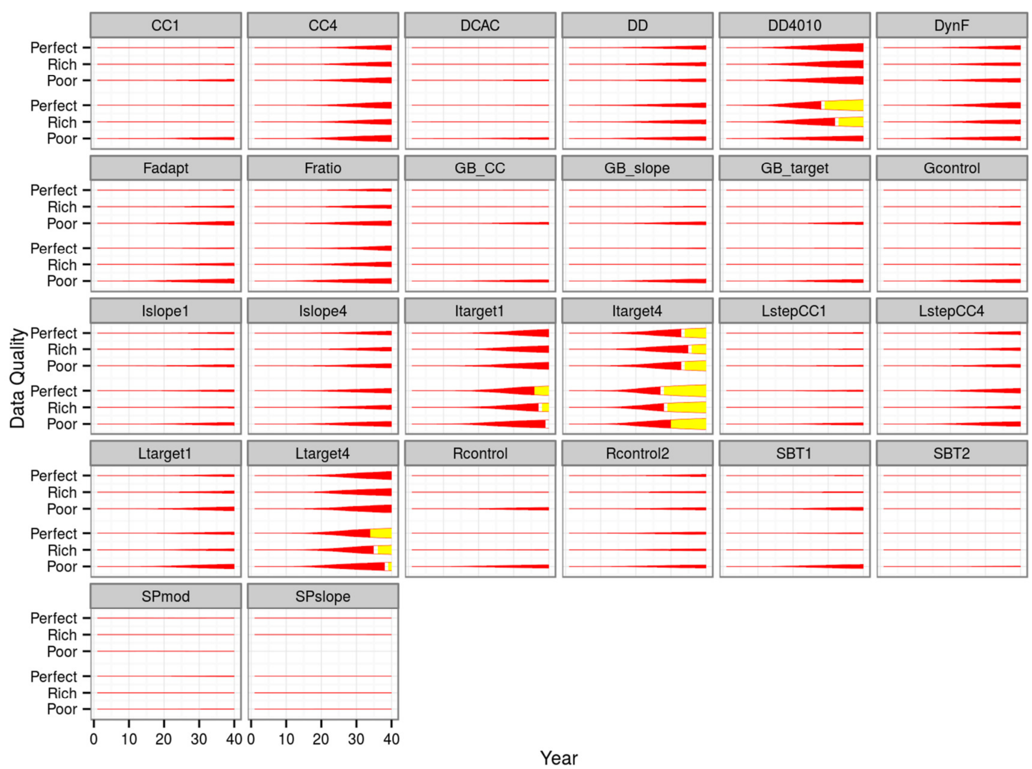
Figure 3. Duration of recovery of spawning biomass to *MSY* levels for the herring life-history. The thickness of the horizontal bars represents the cumulative frequency of simulations that have recovered to B_{MSY} levels. The bars are colored red until 50% of simulations have recovered to B_{MSY} levels, then colored yellow until 75% of simulations have recovered to B_{MSY} levels, after which they are colored green. The top three bars in each panel represent high-autocorrelation in recruitment, the bottom three bars represent low autocorrelation in recruitment.

747



748 Figure 4. As Figure 3 but for the bluefin tuna life-history.

749



750 Figure 5. As Figure 3 but for the rockfish life-history.

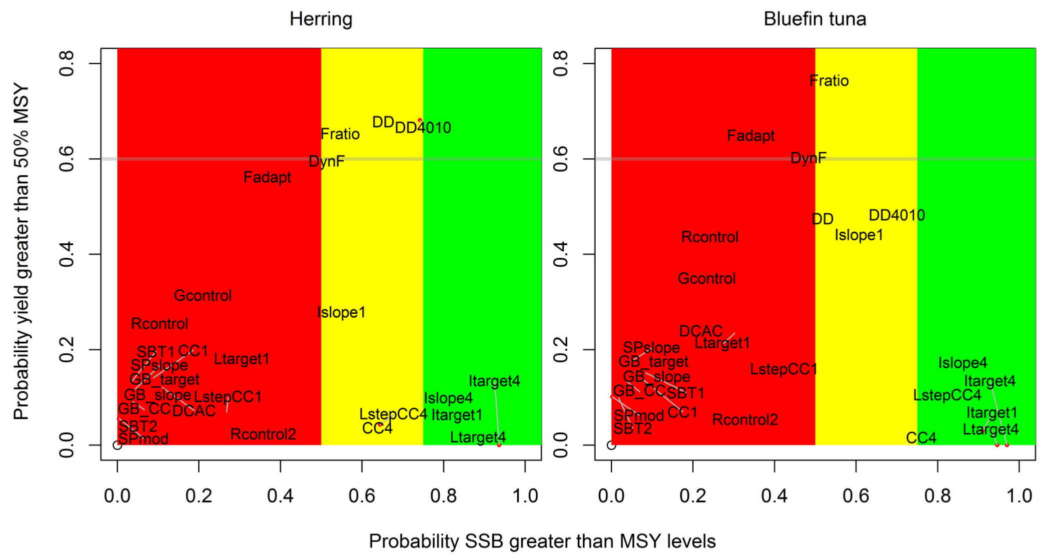


Figure 6. The trade-off between probability of rebuilding to B_{MSY} levels and the probability of meeting a yield target of 50% *MSY*. Plotted are the results of the data-rich simulations with low autocorrelation in recruitment. Colored shading reflects the probability of rebuilding the stock to above SSB after 40 years. Red plot regions correspond to less than a 50% probability of rebuilding, yellow regions less than a 75% probability of rebuilding and green regions greater than 75% probability of rebuilding. Each panel includes a horizontal line at 60% probability of achieving a yield of greater than 50% *MSY*, which corresponds to the target.

Figure 7. 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 percent of simulations succeeded in keeping spawning stock biomass above the target level of 50% of *MSY* levels. Green areas represent regions that exceed the target performance level. Black and red points represent simulations with low and high recruitment autocorrelation, respectively.

768

769 Figure 8. As Figure 7 but for the bluefin tuna life-history.

Figure 9. As Figure 7 but for the rockfish life-history.

Appendix A: Reference methods

DCAC

In circumstances where the information available is insufficient to derive a catch-limit from stock assessment Depletion Corrected Average Catch has been applied (DCAC, MacCall 2009). DCAC attempts to calculate average catch accounting for the removal of "windfall harvest" of less productive biomass that may have occurred as the stock became depleted. DCAC requires inputs for M, F_{MSY}/M (or c), B_{MSY}/B_0 (stock biomass at MSY relative to unfished, B_{peak}) and B_{cur}/B_0 (current stock depletion, D). A number of samples are drawn from the following distributions:

App.A.1a) $M_{DCAC} \sim \text{lognormal}(\mu=M, SD=0.5)$

App.A.1b) $c_{DCAC} \sim \text{lognormal}(\mu=c, SD=0.2)$

App.A.1c) $D_{DCAC} \sim \text{lognormal}(\mu=D, SD=0.2)$

where, in keeping with MacCall's (2009) approach, the SDs for M and c are set to 0.5 and 0.2, respectively.. MacCall (2009) states that "unlike the other parameters, the precision of [depletion D] is entirely dependent on the data and method used in its estimation, and there is no clear value of precision that can serve as a default". Subsequently, Dick and MacCall (2011) assume a default distribution with a CV of 0.25. We adopt the same beta distribution for depletion to remain consistent with the assumptions made in simulating DB-SRA (detailed above in management scenario M1), *i.e.*:

App.A.2a) $D_{DCAC} \sim \text{beta}(\mu=D_{obs}, CV = 0.25)$ where $D_{obs} < 0.5$

App.A.2b) $1\text{-}D_{DCAC} \sim \text{beta}(\mu=1\text{-}D_{obs}, CV = 0.25)$ where $D_{obs} > 0.5$

For each sample of these parameters, sustainable yield (SY) is calculated by:

App.A.3)

$$SY_{DCAC} = \frac{\sum C_{obs}}{n + (1 - D_{DCAC}) / (Bpeak_{DCAC} c_{DCAC} M_{DCAC})} = \frac{\sum C_{obs}}{n + (1 - D_{DCAC}) / (0.4 c_{DCAC} M_{DCAC})}$$

where the C_{obs} are annual historical catches and n is the number of years of historical catches.

This stochastic approach produces numerous samples of the derived sustainable yield (SY) of which a percentile (typically the median) is used as the TAC.

F_{MSY}/M ratio 'Fratio'

It has been suggested that ratios of F_{MSY}/M (c) may be robust to broad life-history types and fisheries exploitation scenarios. Gulland (1971) proposed a simple method of setting maximum sustainable yield $MSY = 0.5M \cdot B_0$, in doing so assuming that B_{MSY}/B_0 =0.5 and F_{MSY}/M = 1. Subsequent publications have revised this F_{MSY} recommendation downwards. The Fratio MP is simulated by generating imperfect knowledge regarding M, current absolute biomass and the ratio of F_{MSY}/M.

Delay-difference stock assessment (DD)

The performance of a delay-difference model (Deriso, 1980; Schnute, 1985) fitted to catch and effort data is evaluated to provide a reference for the performance of the other MPs. The delay-difference model requires additional auxiliary (independent) information regarding the form of the stock-recruit function, the fraction mature at age, somatic growth, M, and the selectivity-at-age curve. The delay-difference stock assessment method provides estimates of B_{cur} and F_{MSY} and therefore direct estimates of an appropriate catch limit.

The delay-difference model is fitted to annual total catch and effort data. The model is parameterized according to: maximum sustainable yield, MSY_{DD} and harvest rate at maximum sustainable yield, $Umsy_{DD}$. The catchability coefficient scaling effort to fishing mortality rate is also estimated. The growth parameters α and ρ of the Ford-Brody growth model ($W_{a+1}=\alpha+\rho W_a$) are approximated from the known weight at age W, for each simulation:

App.A.4) $$\alpha = W_\infty(1-\rho); \qquad \rho = \frac{W_{V_{obs}+2} - W_\infty}{W_{V_{obs}+1} - W_\infty}$$

where W_∞ is the maximum weight of an individual. Selectivity at age is assumed to follow the maturity schedule and AM_{obs} is the observed age at 50% maturity selectivity determined from the ascending limb of the selectivity curve ω (Eqn. App.A.12). Since bias in the age at 50% maturity may strongly affect the delay-difference model, AM_{obs} is simulated subject to imperfect knowledge (Table App.C.1). Survival rate at maximum sustainable yield is given by $Smsy = \exp(-M_{obs})(1-Umsy_{DD})$ so that the number of spawners per recruit, SPR is given by:

App.A.5) $$SPR = \frac{(\alpha \cdot Smsy)/(1-Smsy) + W_{AM_{obs}}}{1-\rho \cdot Smsy}$$

The Beverton-Holt parameter α_{rec}, the maximum recruits per spawner as spawner biomass approaches zero, is calculated:

App.A.6) $$\alpha_{rec} = 1/\left((1-Umsy_{DD})^2 (SPR + Umsy_{DD} \cdot \Delta_{SPR})\right)$$

The derivative of yield with respect to harvest rate Δ_{SPR}, evaluated at $Umsy_{DD}$ is given by:

App.A.7) $$\Delta_{SPR} = -S_0 \frac{p}{1-\rho \cdot Smsy} \frac{SPR+1}{1-\rho \cdot Smsy} \frac{\alpha}{(1-Smsy)} + \frac{Smsy \cdot \alpha}{(1-Smsy)^2}$$

where S_0 is unfished survival rate $S_0 = \exp(-M)$. The Beverton-Holt parameter β_{rec} is calculated as:

App.A.8) $$\beta_{rec} = \frac{Umsy_{DD} \cdot (\alpha_{rec} \cdot SPR - 1/(1-Umsy_{DD}))}{MSY_{DD}}$$

Unfished recruitment R_0 is allocated to recruitments up to and including the age at recruitment to the fishery Am_{obs} and is given by:

App.A.9) $$R_0 = \frac{\alpha_{rec} \cdot SPR_0 - 1}{\beta_{rec} \cdot SPR_0}$$

where unfished spawners per recruit SPR_0 is calculated using Eqn. App.A.5 when S_{msy} is replaced by S_0. It follows that initial biomass B_1 is given by: $B_1 = R_0 \cdot SPR_0$ and initial numbers N_1 is given by $N_1 = R_0 / (1 - S_0)$. From this initialization, biomass dynamics are calculated by:

App.A.10) $$B_{y+1} = S_y \left(\alpha \cdot N_y + \rho \cdot B_y\right) + W_V \cdot R_{y+1} \quad ; \quad N_{y+1} = S_y \cdot N_y + R_{y+1}$$

where $S_y = \exp(-E_y q_{DD} - M)$ is the survival rate in year y, N represents stock numbers, B is the biomass, W_k is the weight of an individual at the age at 50% selectivity k, M is the natural mortality rate (assumed to be known exactly), q_{DD} is the estimated catchability, E_{y} is the observed fishing effort during year y, and R_{y} represents the number of recruits during year y:

App.A.11) $$R_{y+k} = \frac{\alpha_{rec}\left(B_y - C_y\right)}{1 + \beta_{rec}\left(B_y - C_y\right)}$$

where catches C, are given by: $C_y = B_y\left(1 - \exp\left(-q_{DD} E_y\right)\right)$.

The model is fitted to observed (simulated) catches by minimizing a global objective O that is calculated by the sum of the negative log likelihood of the catches (excluding constant terms):

App.A.12) $$O = \sum_y \left[\frac{\log(2\pi)}{2} + \log(\sigma_c) + \frac{\left(\log\left(C_y^{obs}\right) - \log\left(C_y\right)\right)^2}{2\sigma_c^{\,2}}\right]$$

where σ_c is the assumed standard deviation (in log space) of the observation error.

Appendix B. Operating model

Simulating stock dynamics

A standard age-structured, spatial model identical to that of Carruthers *et al*. (2014) was used to simulate population and fishery dynamics. Ranges of parameters and variables allowed variation among simulations for a given stock (*e.g*., natural mortality rate M, slope in recent fishing effort, targeting). All parameters that vary as random variables across simulations are denoted with a tilde (*e.g.* $\tilde{\sigma}$). The probability distributions from which these parameters are sampled are detailed in Table App.B.1. Hence, each parameter or variable denoted with a tilde represents a sample from a distribution specific to each stock. This convention alleviates the need for a simulation and stock subscript for every parameter or variable described below. For example, the symbol $\tilde{\sigma}$ represents $\tilde{\sigma}_{s,i} \sim f(\theta_s)$ which is the sample of the parameter $\tilde{\sigma}$ corresponding with the i^{th} simulation for stock s, drawn from a distribution function $f()$, which has stock specific parameters θ_s.

The numbers of individuals recruited to the first age group $N_{y,a=1,r}$ in each year y, and area r is calculated using a Beverton-Holt stock-recruitment relationship with log-normal recruitment deviations:

App. B.1) $$N_{y,a=1,r} = \exp\left(P_{y,r} - \frac{\tilde{\sigma}_{proc}^{\;2}}{2}\right)\frac{0.8R_0\tilde{h}\,SSB_{y,r}}{0.2SSB_0(1-\tilde{h})+(\tilde{h}-0.2)SSB_{y,r}}$$

where h is the steepness parameter, R_0 is the mean recruitment given unfished conditions, $SSB_{y,r}$ is spawning stock biomass in the previous year and SSB_0 is the mean spawning stock biomass under unfished conditions. The process error term P, is an autocorrelated random variable:

App.B.2) $$P_{y,r} = \varsigma \cdot P_{y-1,r} + \phi_{y,r} \cdot \sqrt{1-\varsigma^2}$$

where ς controls the level of autocorrelation in recruitment deviations and ϕ is a normally distributed random variable with mean zero and standard deviation $\tilde{\sigma}_{proc}$:

App.B.3) $$\phi_{y,r} \sim \text{normal}\left(0, \tilde{\sigma}_{proc}\right)$$

The spawning stock biomass *SSB*, is given by:

App.B.4) $$SSB_{y,r} = \sum_{a=1}^{n_a} m_a W_a N_{y,a,r}$$

where m_a is the maturity-at-age *a,* and the maximum age n_a is specific to each stock. Maturity-at-age is assumed to follow a logistic relationship with age; the slope of the transition from immature to mature is determined by the precision parameter σ_A, and the inflection point $\tilde{A}_m$ that is the age where 50% of individuals are mature (sampled form a random uniform distribution):

App.B.5) $$m_a = \frac{1}{1 + \exp\left(\left(\tilde{A}_m - a\right) / \sigma_A\right)}$$

Numbers at age are converted to biomass using the von Bertalanffy growth equation:

App.B.6) $$L_a = \tilde{L}_{\inf}\left(1 - e^{-\tilde{\kappa}(a - t_0)}\right)$$

where L_a is the length of an individual of age *a*, the asymptotic length is L_{inf}, and κ is the slope at the theoretical age at zero length t_0. Simulated L_{inf} and κ are sampled independently and assumed to be time-varying with mean percentage slope α_{Linf} and α_{κ} (Table App.B.1.). Parameters α_{Linf} and α_{κ} were sampled independently from uniform distributions between -0.25 and 0.25 percent per year to investigate whether small temporal changes in growth could affect MP performance. Inter-annual variability in L_{inf} and κ were simulated from log-normal distributions with mean 1, and standard deviations sd_{Linf} and sd_K.

Weight at age W_a is assumed to be related to length by:

App.B.7) $$W_a = \beta L_a^{\alpha}$$

For ages greater than 1, fishing mortality is assumed to occur before natural mortality and the numbers-at-age are calculated by:

App.B.8) $$N_{y,a,r} = \left(N_{y-1,a-1,r} - C_{y-1,a-1,r}\right)\exp\left(-\tilde{M}\right)$$

Similarly to L_{inf} and κ, inter-annual variability in natural mortality rate was generated by sampling from a lognormal distribution with mean 1 and standard deviation sd_M. The underlying trend (per cent per year) in natural mortality was sampled from a random uniform distribution (Table App.B.1.). No "plus group" is modelled; instead the maximum age is set sufficiently high that survival to the maximum age is less than 1% under unfished conditions.

Movement and spatial targeting dynamics were not the focus of this simulation evaluation. The generic two-area model of the simulation framework was parameterized to mimic a fully diffuse stock that was not subject to spatial targeting.

Simulating fishery dynamics

The selectivity at age ω_a, was calculated using a double normal curve with age at maximum selectivity m_{sel}, an ascending limb standard deviation of σ_{sel1} and a descending limb standard deviation σ_{sel2}. These standard deviations were determined for each simulation by numerically solving for two user-specified quantities that are more intuitive: (1) the minimum age at 5% maximum selectivity $\tilde{a}_{0.05}$, and (2) the selectivity of the oldest age class $\tilde{\omega}_{old}$. In order to sample a wide range of selectivity dynamics $\tilde{a}_{0.05}$

was sampled from a uniform distribution between 20% and 50% age at maturity. Additionally $\tilde{\omega}_{old}$ could range from a 0-100%, representing dome-shaped selectivity curve where older fish are not fished to a ‘flatted – topped’ selectivity where older fish are fished at the same rate as younger fish.

The ascending limb age selectivity A_a (before normalization to a maximum value of 1) is given by:

App.B.9) $$A_a = \frac{1}{\sqrt{2\pi\tilde{\sigma}_{sel1}^{2}}} \exp\left(-\frac{(a - m_{sel})^2}{\tilde{\sigma}_{sel1}^{2}}\right)$$

The descending limb selectivity D_a is given by:

App.B.10) $$D_a = \frac{1}{\sqrt{2\pi\tilde{\sigma}_{sel2}^{2}}} \exp\left(-\frac{(a - m_{sel})^2}{\tilde{\sigma}_{sel2}^{2}}\right)$$

The selectivity at age is given by:

App.B.11) $$\omega_a = \begin{cases} A_a / \max(A_j) & j \le m_{sel} \\ D_a / \max(D_j) & j > m_{sel} \end{cases}$$

Catch in numbers is calculated by:

App.B.12) $$C_{y,a,r} = N_{y,a,r}\left(1 - \exp\left(-\omega_a\, p_{y,r}\, F_{y,a}\right)\right)$$

where F is the fishing mortality rate.

Observed catch is calculated by multiplying simulated catch in numbers-at-age by weight-at-age and adding observation error:

App.B.13) $$C_y^{obs} = \exp\left(\varepsilon_{y,a,r} - \frac{\tilde{\sigma}_{obs}^{\ 2}}{2}\right)\sum_a \sum_r C_{y,a,r} W_a$$

The error term ε, is drawn from a standard normal distribution whose standard deviation σ_{obs} is sampled at random in each simulation:

App.B.14) $$\varepsilon_{y,a,r} \sim \text{dnorm}\ (0, \tilde{\sigma}_{obs})$$

Fishing mortality rate F was assumed to be proportional to effort according to the constant $\tilde{q}$ which was determined by numerical optimizing for sampled current depletion D (Table App.B.1. below).

App.B.15) $$F_y = \tilde{q} E_y$$

Total effort is not related to biomass levels, and in historical and future projections can remain high even at very low biomass levels. The maximum instantaneous fishing mortality rate is limited to 90% to prevent the simulation of large declines in stock biomass in any year due to TAC recommendations that are occasionally very high.

Log-normal variability in effort is added to a general effort trend V:

App.B.16) $$E_y = \exp\left(\varphi_y - \frac{\tilde{\sigma}_{eff}^{\ 2}}{2}\right) V_y$$

The effort variability term φ_y is randomly sampled from a standard normal distribution that has a standard deviation, $\tilde{\sigma}_{eff}$ drawn at random for each simulation from a uniform distribution ranging from 0.1 to 0.4:

App.B.17) $$\varphi_y \sim \text{dnorm}\ (0, \tilde{\sigma}_{eff})$$

A range of effort variability is sampled to assess how the degree of auto-correlation affected the performance of stock status classification methods. The general trend in effort is determined by a linear model of change in effort over time with slope a_E, and intercept of 0.5:

App.B.18) $$\frac{dV_y}{dy} = a_E y + 0.5$$

This functional form allows effort to increase, decrease or remain unchanged over time. This effort model is constrained by sampling positive values for initial changes in effort (effort is increasing at the start of the time series). The final annual change in effort $\tilde{\Delta}_E$, was sampled from a uniform distribution between -1 and 1 to simulate a range of final effort trajectories including strongly decreasing and increasing effort:

App.B.19) $$\tilde{\Delta}_E = \frac{dV_{final}}{dy}$$

For any simulated effort time series, the slope a_E, can then be calculated from the total number of years in the time series n_y, and the sampled intercept of 0.5:

App.B.20) $$a_E = \left(\tilde{\Delta}_E - 0.5\right) / n_y$$

Simulated effort time series that included negative values were discarded. All of the stocks experienced the same effort dynamics.

In any given year, spatial fishing effort is assumed to be proportional to the distribution of the vulnerable biomass in the previous year, modified by a targeting parameter λ, that controls how strongly fishing effort will be distributed in relation to vulnerable biomass. The fraction of fishing effort P, allocated to each region r, in a given year y, is calculated:

App.B.21) $$p_{y,r} = \left(\sum_a \omega_a W_a N_{y,a,r} \right)^{\lambda} \Bigg/ \sum_r \left(\sum_a \omega_a W_a N_{y,a,r} \right)^{\lambda}$$

The values for p sum to 1 in any year so they can be used to distribute total effort E_y across areas in each year such that mean F among areas is the same as total annual F. Fishing is distributed evenly regardless of the vulnerable biomass in the previous year when the targeting parameter λ is zero. Spatial fishing will be distributed in favour of areas of high vulnerable biomass when λ is positive and distributed away from such areas when λ is negative. For all stocks a range of the targeting parameter was sampled from a random uniform distribution between -0.5 and 1 to evaluate the impact on MPs of the distribution of fishing relative to the population.

Parameterization of stock dynamics

Given the availability of full stock assessments with which to characterize their stock dynamics, we chose Pacific herring (DFO, 2012), Atlantic bluefin tuna (ICCAT, 2012), and canary rockfish (Wallace and Cope, 2011) as case-studies that span a range of longevity. The values of input parameters and the sources of these inputs are detailed in Table App.B.1.

Table App.B.1. Summary of the variables/parameters that define each of the stock simulations, including values and/or the range over which they are sampled. The values for simulations were taken from recent stock assessments for Pacific herring (DFO, 2012), eastern Atlantic bluefin tuna (ICCAT, 2012) and canary rockfish (Wallace and Cope, 2011). Where two values are provided, variables are sampled from a uniform distribution with the lower and upper bounds listed.

Name		**Pacific herring**		**Eastern Atlantic bluefin tuna**		**Canary rockfish**	
Maximum age	n_a	10		32		64	
Steepness	h	0.4	0.6	0.6	0.9	0.35	0.7
Mean natural mortality rate	μ_M	0.28	0.38	0.12	0.16	0.04	0.08
Interannual variability in natural mortality rate	sd_M	0	0.1	0	0.1	0	0.05
Gradient in natural mortality rate (per cent y^{-1})	α_M	-0.5	0.5	-0.5	0.5	-0.5	0.5
Theoretical age at length zero	t0	-0.025		-0.97		-0.04	
Mean maximum length	μ_{Linf}	25	29	315	325	62	68
Interannual variability in maximum length	sd_{Linf}	0	0.025	0	0.025	0	0.025
Gradient in maximum length (per cent y^{-1})	α_{Linf}	-0.25	0.25	-0.25	0.25	-0.25	0.25
Mean von Bertalanffy growth coefficient	μ_K	0.43	0.53	0.08	0.1	0.122	0.128
Interannual variability in the growth coefficient κ	sd_K	0	0.025	0	0.025	0	0.025
Gradient in the growth coefficient κ (per cent y^{-1})	α_K	-0.25	0.25	-0.25	0.25	-0.25	0.25
Weight-length parameter a ($W=aL^b$)	α_{WL}	4.50E-06		1.96E-05		1.55E-05	
Weight-length parameter b ($W=aL^b$)	b_{WL}	3.127		3.009		3.03	
Stock depletion, biomass relative to unfished	D	0.025	0.6	0.025	0.6	0.025	0.6
Age at 50% maturity	A_m	1.7	2.3	3.5	5	6.5	9.5
Log-normal recruitment variation	σ_R	0.2	0.4	0.1	0.3	0.2	0.5

Appendix C: Simulating imperfect information

Table App.C.1. 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. The log-normal 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: $\text{dnorm}\left(-0.5\log\left(1+\sigma^2/\mu^2\right),\sqrt{\log\left(1+\sigma^2/\mu^2\right)}\right)$.

Variable	Symbol	Related functions
The standard deviation of the log-normally distributed bias in natural mortality rate M (μ_M varies among simulations)	Υ_M	$M_{obs} = M \times \mu_M$ μ_M~lognormal(μ=1,Υ_M)
The standard deviation of the log-normally distributed bias in von Bertalanffy growth rate parameter K (μ_K varies among simulations)	Υ_K	$K_{obs} = K \times \mu_K$ μ_K~lognormal(μ=1,Υ_K)
The standard deviation of the log-normally distributed bias in biomass at maximum sustainble yield B_{MSY} (μ_{Bmsy} varies among simulations)	Υ_{Bmsy}	$Bmsy_{obs} = Bmsy \times \mu_{Bmsy}$ μ_{Bmsy}~lognormal(μ=1,Υ_{Bmsy})
The standard deviation of the log-normally distributed bias in biomass at maximum sustainable yield relative to unfished Bpeak (B_{MSY}/B_0, μ_{Bpeak} varies among simulations)	Υ_{Bpeak}	$Bpeak_{obs} = Bpeak \times \mu_{Bpeak}$ μ_{Bpeak}~lognormal(μ=1,Υ_{Bpeak})
The standard deviation of the log-normally distributed bias in the ratio of maximum sustainable fishing mortality rate to natural mortality rate F_{MSY}_M (μ_{FMSY_M} varies among simulations)	Υ_{FMSY_M}	$c_{obs} = c \times \mu_{FMSY\_M}$ μ_{FMSY_M}~lognormal(μ=1,Υ_{FMSY_M})
The standard deviation of the log-normally distributed bias in MSY (μ_{MSY} varies among simulations)	Υ_{MSY}	$MSY_{obs} = c \times \mu_{MSY}$ μ_{MSY}~lognormal(μ=1,Υ_{MSY})
The standard deviation of the log-normally distributed bias in the age at first maturity Am (μ_{Am} varies among simulations)	Υ_{Am}	$Am_{obs} = Am \times \mu_{Am}$ μ_{Am}~lognormal(μ=1,Υ_{Am})
Uniformly distributed observation error in recruitment (R_{obs} varies among years and simulations, σ_{Robs} varies among simulations)	σ_{Robs}	R_{obs} = lognormal(μ=R,σ_{Robs}) σ_{Robs}~$U(L_{Robs},U_{Robs})$
The standard deviation of the log-normally distributed bias in the current level of stock depletion D (B/B_0; D_{obs} and Υ_D vary among projected years and simulations; μ_D and σ_D vary among simulations)	Υ_D	$D_{obs} = D \times \Upsilon_D$ Υ_D~lognormal(μ_D,σ_D) μ_D~lognormal(μ=1,Υ_D)
Uniformly distributed observation error in current stock depletion μ_D for projected years	σ_D	σ_D~$U(L_D,U_D)$
The standard deviation of the log-normally distributed bias in the current catches C (C_{obs} and Υ_C vary among projected years and simulations; μ_C and σ_C vary among simulations)	Υ_C	$C_{obs} = C \times \Upsilon_C$ Υ_C~lognormal(μ_C,σ_C) μ_C~lognormal(μ=1,Υ_C)
Uniformly distributed observation error in current catches μ_C for projected years	σ_C	σ_C~$U(L_C,U_C)$
Standard deviation in log-normal 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	$I=B^{\beta} \times \Upsilon_B$ Υ_B~lognormal(1, σ_I) σ_I~$U(L_I,U_I)$
The beta parameter controlling hyperstability / hyperdepletion in the abundance index (β varies among simulations)	β	LN(β)~U(LN(β_{min}),LN(β_{max}))
Loguniform bias in current biomass (B_{obs} and Υ_B vary among years and simulations, μ_B and σ_B vary among simulations)	min_B max_B	$B_{obs} = B \times \Upsilon_B$ Υ_B~lognormal(μ_B,σ_B)
The maximum standard deviation for log-normal error in current biomass for projected years	σ_B	LN(μ_B)~U(min_B,max_B) σ_B~$U(L_B,U_B)$