ICES Journal of Marine Science Advance Access published February 26, 2014

ICES Journal of Marine Science

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

ICES Journal of Marine Science; doi:10.1093/icesjms/fsu017

Complex assessments or simple management procedures for efficient fisheries management: a comparative study

H. F. Geromont* and D. S. Butterworth

Marine Resource Assessment and Management Group, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa

Corresponding author. e-mail: geromont@telkomsa.net

Geromont, H. F., and Butterworth, D. S. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsu017.

Received 22 September 2013; accepted 18 January 2014.

Complex stock assessment methods are data- and expertise-hungry, with the annual updates of catch-at-age data and models typically seen as an essential requirement for sound management. But are the heavy commitments of resources required for this level of annual intervention really necessary to achieve efficient long-term fishery management? This question is addressed through a retrospective analysis of management performance over the last 20 years for four North Atlantic fish stocks. The assessments for two of these stocks have exhibited fairly strong retrospective patterns. The actual assessment advice for these stocks was provided based on complex assessment methods making use of age data. The outcomes are compared with what could have been achieved with much simpler catch control rules based upon age-aggregated survey indices alone. Even for the stocks whose assessments exhibit retrospective patterns, these simple rules can achieve virtually equivalent catch and risk performance, with much less interannual TAC variability, compared with what actually occurred over the past 20 years.

Keywords: complex assessments, empirical management procedures, hindsight/forecast projections, management performance, retrospective comparison.

Introduction

Scientific management advice for more valuable fish resources has generally been based on regular (often yearly) assessments. These stock assessments usually utilize age-structured population models such as Virtual Population Analysis (VPA), Statistical Catch-At-Age (SCAA), or Integrated Analyses Models (IA) which rely heavily on annual updates of fishery catch-at-age or catch-at-length data. They are often complex, requiring substantial scientific expertise to implement (ICES, 2012). The assessment process is further complicated by decisions regarding which data to include in the analyses, and how these data are to be incorporated in the objective function being minimized to fit the population model assumed. But is this complexity really necessary for sound long-term management?

With the thrust of current fisheries management primarily focused on rebuilding of overfished stocks, the three key questions that need to be addressed are: "where are we now?", "where do we want to go?", and "how do we get there?". For data-rich stocks for which adequate reliable data are available, the first question is answered by performing a traditional stock assessment to estimate current resource status in terms of biological reference points. The second question, which relates to the choice of target stock size and the period over which to achieve it, is generally a policy decision, partly informed by the stock assessment (Brodziak *et al.*, 2008). [In the United States, this policy is reflected in the Magnuson–Stevens Fishery Conservation and Management Act (MSA), amended in 1996 to include specifications in terms of maximum sustainable yield (MSY) targets and thresholds for purposes of stock rebuilding, is collectively known as the Sustainable Fisheries Act (SFA). The International Commission for the Exploration of the Seas (ICES) base management advice on satisfying biological limits and target fishing mortality rates.] However, it is the third question that is the focus of this paper: how best to achieve a target biomass in a chosen time frame.

While complex age-structured assessments are necessary to obtain the estimates of current stock status (in answer to the "where are we now" question), they may not provide the best basis for optimal long-term fisheries management advice (in answer to the "how do we get there" question). Indeed, detailed annual stock assessments may not be necessary to achieve management goals and may constitute an inappropriate use of limited resources. Are there simpler, more efficient ways of providing reliable ongoing management advice?

The management procedure (MP) approach has established itself as a powerful fisheries management tool which helps to meet multiple management objectives in a manner that checks robustness to uncertainty for compatibility with the Precautionary Approach (De Oliveira *et al*., 2009). For this reason, *inter alia*, this approach is favoured over that of annual stock assessments as basis for the provision of management advice for the most valuable data-rich fish stocks in South Africa (Geromont *et al*., 1999). For these data-rich stocks, complex statistical assessments are typically performed only at multiyear intervals to answer the "where are we now" question, while simple empirical MPs (harvest control rules based on relatively few data) are employed to move high-value stocks to preselected target levels in answer to the "how do we get there" question. Both approaches are therefore important in their own right, but they answer different questions and fulfil different roles: the goal of the MP approach is to simplify and automate annual fisheries management advice (for example, TAC recommendations) to achieve a target stock size over a chosen period.

Unfortunately, the MP approach is frequently misunderstood and criticized for the simplicity of its harvest control rules by scientists who advocate that all available data should be used in conjunction with the "best" current assessment model to best inform ongoing management of a resource. The aim of this paper is to show that simple harvest control rules might perform as well as, or perhaps even better, than complex annual assessments in achieving long-term management goals (i.e. in addressing the "how do we get there" question), even when confronted with typical assessment problems such as retrospective patterns.

Retrospective comparison

The performances of the two management paradigms are compared in a simple and self-consistent manner by conducting a retrospective study of four data-rich fisheries in the North Atlantic, specifically North Sea sole and plaice and New England witch flounder and plaice. The two North Sea stocks were selected for this retrospective comparative study because of the data-intensive nature of the age-structured models currently applied to provide annual TAC advice (ICES, 2010). The two New England stocks have been included because of the persistent retrospective patterns (in these instances, systematic overestimation of recent spawning biomass and underestimation of recent fishing mortality) present in the VPA assessments on which catch advice has been based (NEFSC, 2012); the advice provided for these stocks has had to adjust for these patterns to avoid TAC recommendations being too high.

Recent assessments [2010 ICES assessments for North Sea sole and plaice, and 2012 assessments for Gulf of Maine witch flounder and plaice (see Supplementary material for details)] of these stocks are used as the basis to compare the two management approaches. The MPs are selected from simulation results based on resource information available up to 1989/1990 (depending on the stock investigated). Their performances are then compared with what actually happened over the subsequent 20-year period under advice arising from regular (in some cases, annual) "best" assessments. (This is an admitted simplification in the interests of drawing the overall comparison. In reality, TACs were hardly ever exactly equal to the output from the assessments, but they were nevertheless closely informed by them.) In all cases, the empirical MPs simulation tested here rely on annual survey abundance estimates only, compared with the annual assessments of the resources which were typically based on a full set of age data, as well as on one or more indices of abundance.

The comparison consists of four steps.

1. Deterministic "hindsight" projections

To start, deterministic projections are performed where perfect knowledge (i.e. hindsight) of parameter values and residuals over the projection period from 1990/1991 to 2009/2010 is assumed. Three simple empirical MPs are each tuned to achieve over these 20 years the same final (2009/2010) spawning biomass as estimated by the recent assessments in question. These MPs with their associated tunings are referred to as "hindsight" MPs, as they have the benefit of hindsight in "knowing" what will happen in the next 20 years in terms of uncertainties (the residuals related to recruitments and survey sampling errors, as well as future selectivity- and weight-at-age vectors).

2. Stochastic "forecast" projections of "hindsight" MPs.

If one had been choosing an MP 20 years ago, one would not have had the benefit of the "hindsight" above at that time, and projections would have been based on information available up to 1989/1990 only. For these stochastic projections then, 1000 simulations are performed to incorporate the key sources of future uncertainty: process error and observation error.

When the "hindsight" MPs of step (1) are applied under these stochastic "forecast" conditions, their performance deteriorates in the absence of exact knowledge of the future, in particular in often yielding final biomasses after 20 years which are considerably below those achieved under the catches that actually occurred. The purpose of this step is to check whether the performance of these "hindsight" MPs would have been considered sufficiently acceptable to have led to their implementation 20 years ago.

3. Use of stochastic "forecast" projections to tune MPs

This step involves selecting alternative tunings of the empirical MPs considered in step (1) that might have led to their being considered acceptable 20 years ago. The stochastic projections are used to select control parameters for these MPs that achieve a spawning biomass distribution in 20 years' time which at the lower 2.5 percentile is at least as large as the biomass estimated by the recent assessment under the actual catches that were subsequently taken. These MPs are thus deliberately more conservative in making allowance for uncertainty in the spirit of the Precautionary Principle, and are termed "forecast" MPs.

4. Deterministic projections of "forecast" MPs

In this final step, deterministic projections are again performed, but this time with the more conservative "forecast" MPs selected at step (3). This provides a self-consistent basis to determine how well simple MPs would have managed the stocks under consideration compared with complex assessments. In particular: how do the resultant average catches and fishing mortalities, their interannual variability, and the final spawning biomass at the end of the projection period compare with what actually occurred (or is estimated to have occurred) under management based on the use of advice arising primarily from regularly, sometimes annually, updated assessments?

Note that, for the sake of simplicity, annual TACs generated by the MPs are assumed to be caught exactly with no allowance made for implementation error.

The basic question to be addressed is: could simple empirical MPs, based solely on a survey index of abundance, have been used to generate appropriate annual TACs for the data-rich resources under consideration?

Projections

The projection period spans the last 20 years of recent assessments for each of the stocks considered, i.e. from 1990 to 2009, or 1991 to 2010, depending on the stock under investigation. Therefore, projections start in 1990/1991, and are moved forward year by year by first obtaining the TAC according to a particular MP based on the latest survey abundance index, from which the corresponding fishing mortality, F_y, rate can be computed for that year, given the selectivity- and natural mortality-at-age vector selected. The population numbers for the start of the next year can then be computed. Once the population numbers for the next year are available, the spawning biomass is computed and, subsequently, the number of recruits for the following year according to a stock–recruit relationship. In addition, the survey abundance index for the next year, required by the MP that calculates the yearly TAC, is generated. By repeating this cycle, population numbers-at-age are projected forward from 1990/1991 to 2009/2010.

Input to projections

Recent assessment results for North Sea sole and plaice (ICES, 2010) and New England witch flounder and plaice (NEFSC, 2012) are used as basis for the retrospective analysis.

For purposes of this exercise, the management period for each stock is split in 1990/1991 into a "historic" and "future" period. For the "historic" pre-1990/1991 period, population numbers, fishing mortalities-at-age, natural mortality rate, and number of recruits are taken directly from the corresponding VPA assessments. These are assumed to be known exactly, and are used to provide the operating model (OM) on which the harvesting strategies are tested for each stock. Deterministic and stochastic projections are then performed over a 20-year post-1990/1991 projection period under various catch control rules. When performing a deterministic "hindsight" projection under the actual catches, the same final spawning biomass is achieved as in the VPA assessment.

Input data (observed annual catches, estimated number of recruits, and associated spawning biomasses as estimated by the VPA assessments) and pertinent parameter values are given in Supplementary Appendix SA1–SA4 for each of the stocks investigated. Tables and plots of input data and model parameters are also given in the Supplementary material.

The operating model (OM)

The OM that forms the basis of this retrospective comparison is an age-structured production model. Fishing is assumed to be continuous throughout the year, so that the population dynamics are as described by Supplementary Equations (S.1)–(S.8). A simple two-line "hockey-stick" stock–recruitment function is used for the projections reported in the paper.

Perfect knowledge of the past (pre-1990) is assumed, as provided by the VPA assessment that is used to parameterize the OM. While this might appear a gross oversimplification of an MP approach, where alternative catch control rules are typically simulation tested over a suite of OMs (Butterworth and Punt, 1999; Punt, 2006; Butterworth, 2007, 2008; Punt and Donovan, 2007; De Oliveira *et al*., 2009), the use here of a single OM achieves the primary intent of these analyses which is to compare the outputs of different simple control rules with the catches which actually occurred in a clear-cut and self-consistent manner.

For the deterministic projections (steps 1 and 4 in the "Retrospective comparison" section), some assumptions need to be made: the same multiplicative residuals as for the fits of the actual assessment to the stock–recruitment and survey abundance indices are taken to apply [Supplementary Equations (S.22) and (S.25), respectively], and the assessment selectivities-at-age together with population and catch weights-at-age and natural mortality rate values are maintained unchanged.

For the stochastic "forecast" projections (steps 2 and 3 in the "Retrospective comparison" section), rather than knowing exactly what recruitment residual will apply each year in the future, projections have to assume that these will be drawn at random each year from distributions estimated from fits of stock–recruitment relationships to the assessment results available at that time (i.e. results up to 1989/1990 as indicated by the VPA assessment). Similar assumptions need to be made about the abundance estimates forthcoming from future surveys. In addition, "future" (post-1989/1990) population model parameters are based on "past" (pre-1990/1991) estimates.

In particular, stochastic components are incorporated as follows:

(i) observation error: lognormal error distribution about the generated survey abundance values with a standard deviation of $\sigma = 0.2$;

(ii) process error: stock–recruitment residuals lognormally distributed about the stock–recruitment function with a standard deviation of $\sigma^R = 0.8$ for North Sea sole and $\sigma^R = 0.5$ for the other stocks;

(iii) process error: random re-sampling of "future" fishing selectivity vectors from historical vectors from 1980/1981 to 1989/1990; "future" population and catch weights-at-age averaged over historical weights-at-age for the last 3 years of the pre-management period.

Technical details are given in Supplementary Section S.1.

Candidate MPs

Three types of MPs with different harvest control rules are simulated to compare performance with assessment-based management: a constant catch rule and two very simple empirical MPs, based on changes in survey indices of abundance. These simple MPs are easy to understand, test, and apply, and have shown comparable performance and robustness to uncertainty as their model-based counterparts [for example, model-based MPs used in the 1990s (Geromont *et al*., 1999) have been replaced in the 2000s by empirical rules (Rademeyer, 2012) to manage the South African hake resource]. In particular, these simple empirical MPs are chosen for this retrospective analysis to better compare the data-hungry assessment approach with a data-sparse MP approach.

Note that in implementation for relatively data-rich stocks, such as the four cases considered here, a simple MP approach like this would remain underpinned by a full resource assessment: the former provides ongoing yearly management advice to move the resource to some target stock size ("how to get there?"), while the latter is re-considered at multiyear intervals to check the appropriateness of the MP and if necessary to adjust some of its control parameters by

Table 1. The three candidate MPs considered for retrospective study of the four data-rich stocks.

Empirical MPs	Control parameters
MP constant catch	$TAC_{y+1} = TAC^{\text{target}}$, where TAC^{target} is the annual catch required to reach the target spawning biomass
MP slope: TAC adjusted up or down if the trend in recent survey index values is positive or negative	$TAC_{y+1} = TAC_y\,(1 + \lambda S_y)$, where λ is the smoothing parameter, and S_y the average survey slope over the most recent p years
MP target: TAC adjusted up or down if average of recent survey index values is above or below the target index value.	$TAC_{y+1} = TAC^{\text{target}}[w + (1-w)((I_y^{\text{recent}} - I^0)/(I^{\text{target}} - I^0))]$ *if* $I_y^{\text{recent}} \geq I^0$ $TAC_{y+1} = wTAC^{\text{target}}(I_y^{\text{recent}}/I^0)^2$ if $I_y^{\text{recent}} < I^0$, where I^{target} is the target reference point for survey, $I^0 = 0.2I^{\text{ave}}$ is the limit reference point for survey, I^{ave} the average survey abundance index over past 5 years, I_y^{recent} the average survey of most recent 4 years, TAC^{target} the equilibrium catch, and w a smoothing parameter

Full specifications of these MPs are given in the Supplementary material. A 20% limitation on the extent of interannual change in TAC has been imposed for most cases.

determining whether resource behaviour has remained within the range assumed when last testing the MP (to address "where are we now?").

Constant catch harvesting strategy

This is an extreme and the simplest of all empirical MPs; it requires no data to set the annual TAC. The future TAC is given by the first equation in Table 1.

A constant catch harvesting strategy is not recommended as there is no self-correcting feedback-control mechanism built into this type of MP. It does, however, give a ball-park indication of the average yield that can be expected during the projection period given a chosen target abundance, which is useful for later comparisons among the different candidate MPs.

Survey slope-based harvesting strategy

Slope-based MPs utilize the trend in a limited subset of data (typically the most recent 4 or 5 years of survey biomass estimates) for input. The annual TAC is simply moved up or down from where it was the previous year, depending on whether the estimated trend is positive or negative. The TAC for the next year is given by the second equation in Table 1. Butterworth and Geromont (2001) provides an example of an application of this strategy.

Survey target-based harvesting strategy

A target-type MP moves the resource abundance towards a prespecified target level for some abundance index. The future TAC is adjusted up or down each year depending on whether the average of the most recent surveys is above or below the target survey level. The TAC for the next year is given by the third set of equations in Table 1 (similar to the Tier 4 rule of Wayte, 2009).

Full technical details are provided in Supplementary Section S.3. While the constant catch MP has only a single control parameter, the slope and target MPs have a number of tuneable parameters. Control parameters corresponding to the best-performing "hindsight" and "forecast" MPs for each stock are given in Supplementary Tables S.1–S.16 of the Supplementary Results section for each of the stocks considered. These include specifications of limitations placed on interannual TAC changes, which for the most part are restricted not to exceed 20%.

Results

To aid comparisons among the MP and assessment-based performances, the MPs are first tuned to reach the same final spawning biomass as indicated by the assessment under the actual catches for the deterministic projections of step 1. For the stochastic projections of step 3, the MPs are tuned, so that the 2.5 percentile of the simulated final spawning biomass distribution is the same as, or close to, the final biomass estimated by that assessment.

The statistics reported for comparison of performance of the MPs over the projection period are:

(i) average total annual catch over the 20-year projection period, $\overline{TAC}$,

(ii) average annual variation of catch (variation given by the absolute value of the annual change in catch as a proportion of the catch in the previous year) over the projection period, AAV,

(iii) average annual fishing mortality rate over the projection period, $\overline{F}$,

(iv) average annual variation (given by the absolute value of the change) in fishing mortality, $\overline{\Delta F}$,

(v) final spawning biomass at the end of the projection period as a fraction of the target biomass, $B_{\text{final}}^{\text{sp}}/B^{\text{target}}$, where B^{target} corresponds to the final (2009 or 2010) spawning biomass as estimated by the original assessment, and

(vi) minimum spawning biomass over the projection period, expressed as a fraction of the target biomass, min $B_y^{\text{sp}}/B^{\text{target}}$.

In most cases considered the reasons for choosing the statistics are self-evident. The last is included because, given the uncertainty always associated with a stock–recruitment relationship, one wants to avoid decreases in abundance to levels where recruitment success could be compromised. Thus, biological risk here is defined in terms of a lower percentile of the spawning biomass distributions where recruitment success may be impaired for biomasses below that level.

Four sets of results corresponding to each of the four steps are shown for North Sea sole (Figures 1–4, respectively), with only summary statistics shown for North Sea plaice (Figure 5), New England witch flounder (Figure 6), and American plaice (Figure 7).

North Sea sole

North Sea sole is caught together with North Sea plaice in the demersal flatfish fishery in ICES Subarea IV. Annual TAC advice is based on harvest rules that depend on the state of the stock relative to a biomass reference point and a target fishing mortality. However, to satisfy socio-economic objectives, annual TAC advice has been aligned with limit (precautionary) rather than target reference points, resulting in excessive fishing pressure over the projection

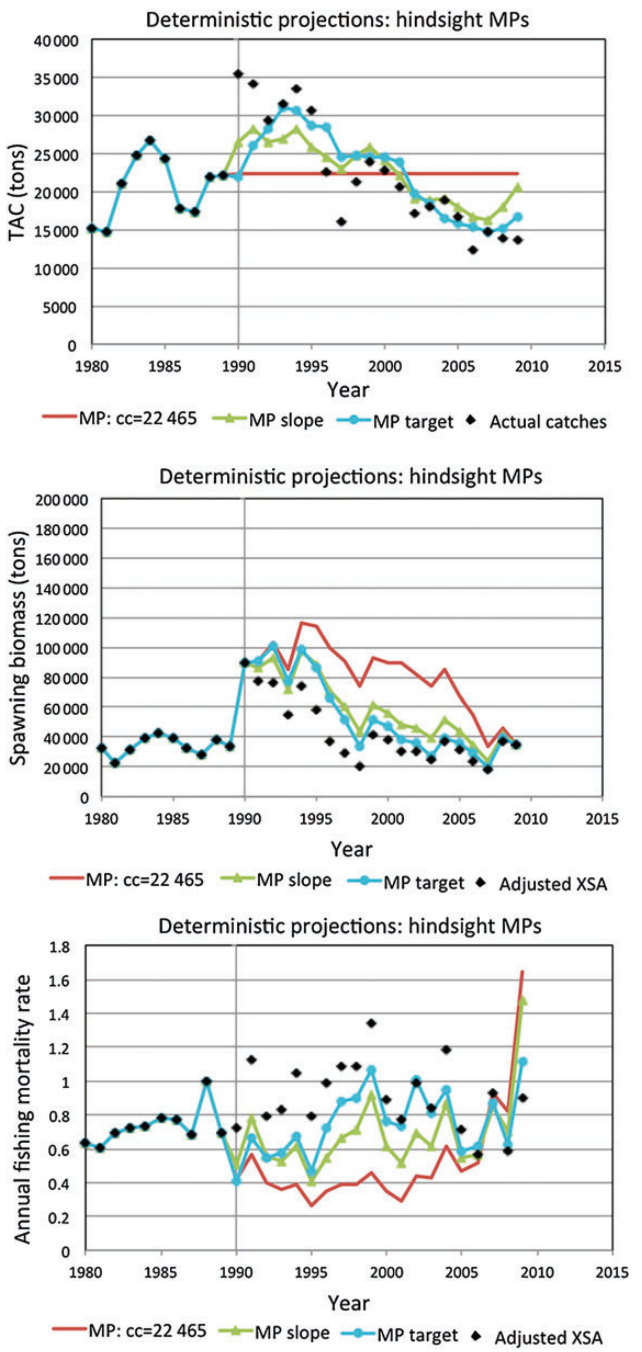

Figure 1. Deterministic projections of constant catch (red line), slope (green triangles), and target (blue dots) "hindsight" MPs, tuned to hit the target biomass in 2009 exactly. TAC (top), spawning biomass (middle), and fishing mortality rate (bottom) trajectories under these three different "hindsight" MPs are compared with biomass and fishing mortality rates under catches which actually occurred (black diamonds) for North Sea sole. (Colours are visible in the online version of this paper.)

period considered (J.J. Poos, pers. comm.). With the aim to rebuild overfished stocks by gradually reducing fishing mortality, the European Union (EU) adopted a multi-annual management plan in 2007, based on MSY (Miller and Poos, 2010). The current precautionary reference points for North Sea sole in terms of spawning biomass and fishing mortality rate are 35 000 t and 0.4 year^{-1}, respectively (ICES, 2010).

Retrospective projections under three simple empirical MPs are compared with actual catches from 1990 to 2009. For purposes of this exercise, the VPA (Extended Survivor Analysis, XSA) stock assessment results reported in the 2010 ICES WGNSSK Report (ICES, 2010) are used to parameterize the OM from which projections are developed. The survey slope and target-based MPs input indices of abundance based on the BTS Isis survey.

Input data and parameter values used for the retrospective projections for North Sea sole are given in Supplementary Tables SA1.1–SA1.6. Plots of input data and pertinent parameter values are given in Supplementary Figures SA1.1–SA1.12.

Biomass and catch trajectories

Deterministic "hindsight" projections for the three candidate MPs, each tuned so that the spawning biomass in 2009 hits the same final spawning biomass for that year as estimated by the 2010 VPA assessment, are compared with projections under actual catches in the plots in Figure 1 (step 1).

The assessment-based approach shows very high catches in the first half of the projection period followed by a rather large drop in annual catches later. The slope and target MPs achieve the same spawning biomass in 2009 by keeping the initial increase in catch smaller than actually occurred, which allows for larger catches later. The constant catch MP achieves the same final spawning biomass in 2009 by leaving the TAC unchanged at the pre-management level.

In Figure 2 (step 3), the three "hindsight" MPs are re-tuned to yield final spawning biomass distributions whose lower 2.5 percentiles are at least as large as occurred under the actual catches. To ensure this level of safety, the constant catch strategy requires an immediate large drop in catch in 1990, followed by an additional decrease in 1991 to reach the level of catch which would ensure adequate spawning biomass recovery at the 2.5 percentile level (top plots). To achieve comparable spawning biomass levels at the end of the 20-year projection period, the slope and target MPs result in greater potential yield as is evident from Figure 2 (middle and bottom plots), with the median future catch (indicated by solid black lines in the plots) under these harvesting strategies greater than the constant catch required to achieve the same level of biological risk. High median catches in the first few years of the projection period under the slope and target harvesting strategies are caused by the relatively high BTS-Isis survey estimates in the most recent pre-management period (see Supplementary Figure SA1.4 of the Supplementary Appendix). However, both MPs are able to self-correct, with subsequent decreases in the median catch necessary to achieve the spawning biomass target in 2009 at the 2.5 percentile.

Deterministic projections for the three "forecast" MPs are compared in Figure 3 (step 4). The slope and target MPs lead to slightly less average yield than was taken in reality, with the final spawning biomass almost twice that achieved under actual catches. The constant catch strategy needs to be overly conservative as it lacks feedback control. Of note is that annual fishing mortality rates for the three "forecast" MPs are much lower than those associated with actual catches, with values in the region of (and sometimes below) the precautionary reference point of 0.4 year^{-1}, compared with much higher values when projecting under actual catches.

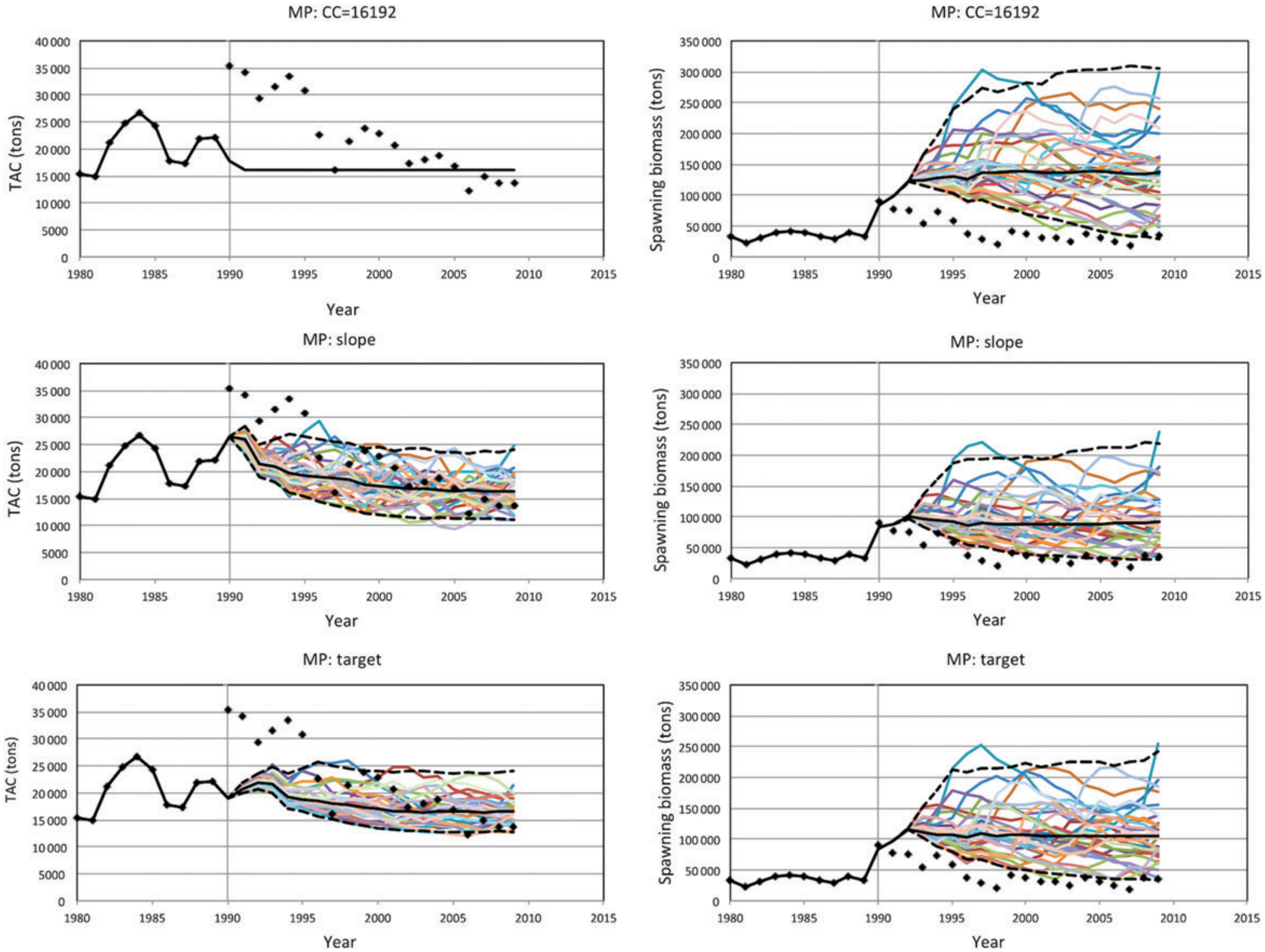

Figure 2. Stochastic TAC (left) and spawning biomass (right) trajectories for the three candidate "forecast" MPs tuned so that the lower 2.5 percentile reaches the target biomass in 2009 (30 of the 1000 simulations conducted are shown here). Top: constant catch MP, middle: slope MP, and bottom: target MP. Spawning biomass trajectories corresponding to the catches of North Sea sole which were actually made are indicated by the black diamonds. The medians and 95% probability interval envelopes are indicated by the solid and dashed lines, respectively.

Summary statistics

Figure 4 provides summaries of performance statistics for both the stochastic and deterministic "forecast" MPs in a graphical form that is helpful when making comparisons.

Due to future uncertainty, these "forecast" MPs need to be rather conservative and take ~20% less catch in the median terms. The main advantage provided by these MPs is that they result in less interannual fluctuation in TAC, compared with formal TAC advice given over the projection period, and therefore likely also less interannual variation in catch. Given hindsight information, the "forecast" MPs are seen to overshoot the 2009 target biomass, with the final spawning biomass about double that target, as is evident from Figure 3. This is at the cost of ~10% less overall catch for the target and slope MPs. The constant catch strategy reflects the smallest total catches on average and would not be a candidate in reality because of its lack of feedback features.

Of note is that the spawning biomass is maintained above the target level over the entire projection period for all three forecast MPs, compared with a drop in abundance to half the target level under the actual catches. As a consequence of the tuning parameters chosen, yearly fluctuations in TAC are kept low under MP management: 7% for the forecast target-type MP compared with 15% interannual fluctuations in actual catch and 14% interannual fluctuation in actual TACs set over this period (see Supplementary Figure SA1.1 for a plot of the actual TACs).

Further retrospective comparisons

Summary statistics are shown in Figures 5–7 for North Sea plaice and New England witch flounder and plaice, respectively. Descriptions of these three fisheries and their assessments can be found in the Supplementary material, together with figures equivalent to Figures 1–3 for North Sea sole.

North Sea plaice

Figure 5 gives a summary of performance statistics for the stochastic and deterministic projections under the three candidate "forecast" MPs for North Sea plaice. Taking roughly the same total catch on average leads to the same or somewhat higher final spawning biomass estimates than occurred under actual catches. This outcome is achieved with far less interannual variation in catch (below 10% compared with the almost 17% that occurred in reality), although these values are not exactly comparable as no allowance has been made for implementation error. [The interannual variation in actual TACs (based on annual assessments) over the

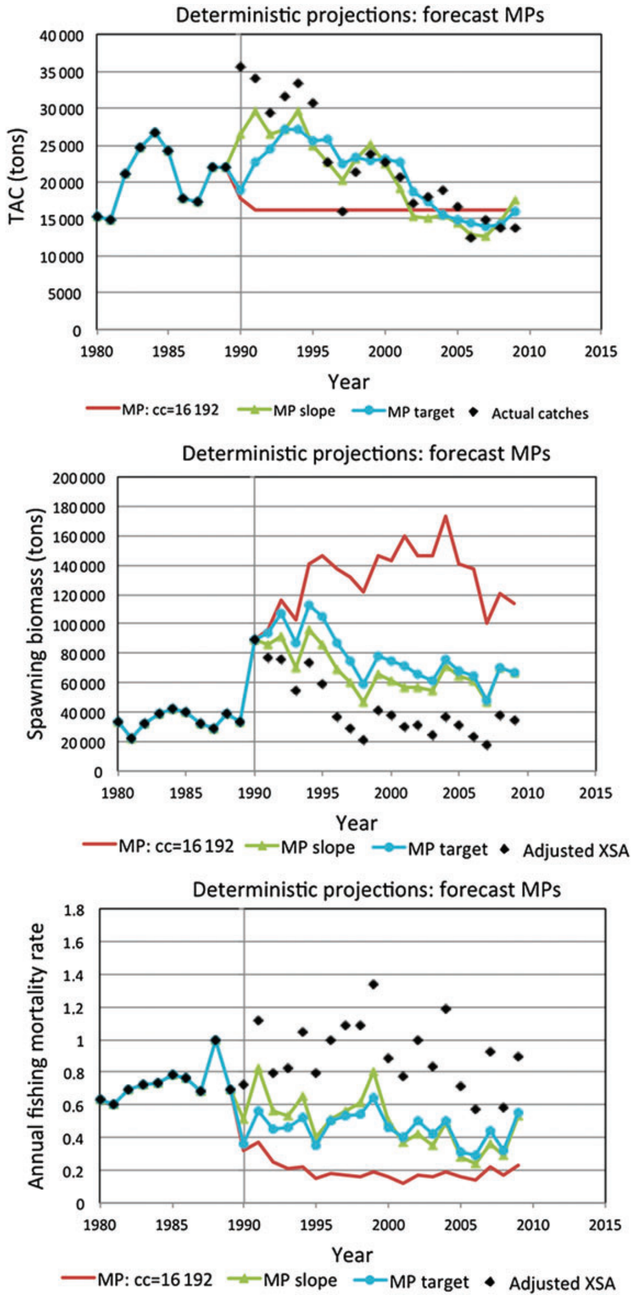

Figure 3. Deterministic projections for the constant catch, slope, and target MPs of Figure 2, which have been tuned, so that the lower 2.5 percentile reaches the target biomass in 2009. TAC (top), spawning biomass (middle), and fishing mortality rate (bottom) trajectories under these three different "forecast" MPs are compared with biomass and fishing mortality rates under the catches which actually occurred (black diamonds) for North Sea sole.

projection period is roughly 10% compared with the 17% in actual catches (landings plus discards), with this difference being mainly due to the large fluctuations in discards (see Supplementary Figure SA2.1).]

For the summary statistics shown in these plots, the performance of the target MP dominates that of the slope MP in terms of every performance statistic. Furthermore, the target MP outperforms the data-hungry assessment-based approach, ensuring potentially less interannual fluctuation in catch.

New England witch flounder

A comparison of summary statistics for different MPs applied to New England witch flounder is given in Figure 6. For deterministic hindsight projections, these MPs result in approximately the same average catch over the projection period as actually occurred and also achieve the same final spawning biomass. The advantage of the going the MP route is the potential drop in the interannual fluctuation in catches (5% compared with the more than 15% that occurred in reality [this high AAV is partly caused by almost annual changes to management regulations for New England groundfish species since the 1970s (Wigley *et al.*, 2003)] and maintaining higher biomass levels throughout the management period.

American plaice

Summary statistics for American plaice are given in Figure 7. These are somewhat less optimistic in terms of the deterministic "hindsight" comparison on the right: For approximately the same average annual catch as actually occurred (top plot), the MPs fail to achieve the biomass target in 2010 (third plot). The target MP performs rather poorly with the final spawning biomass ~25% below target.

The underlying cause of this deterioration in performance from the stochastic "forecast" to the deterministic "hindsight" projection results lies in the systematic decrease in population weights of adult fish (ages 6 and older) over the projection period (see Supplementary Figure SA4.12): the deterministic "hindsight" projections incorporate this decrease explicitly, but the stochastic "forecast" projections are performed using an average historical population weight-at-age vector with the average being taken over the 3 years preceding the management period when the weights of adult fish were high. It is these higher population weights associated with the "forecast" projections which lead to the systematic underachievement of the spawning biomass target, not unlike the retrospective patterns exhibited in the VPA assessments.

In an attempt to improve performance, the target MP was re-tuned to allow for a mid-cycle update in 2000 to better incorporate the systematic changes in model parameters/variables over the 20-year projection period ("MP target+"). Updating pertinent model parameters halfway through the projection period leads to improved performance of the target MP, with a decrease in the median average annual catch and an associated increase in the final spawning biomass in 2010 to close to the target level.

An additional safeguard against such temporal changes in population weights might be desirable: the empirical MPs of Table 1 could readily be modified to accommodate any potential decrease by simply scaling the annual TAC by a factor ($\bar{w}^S_y/\bar{w}^S$), where $\bar{w}^S_y$ is the average weight for year y and $\bar{w}^S$ the average stock weight over the pre-management period.

Summary

These initial results highlight some of the advantages of following the MP approach, which include a potential drop in unnecessary interannual fluctuations in catch and maintaining biomass at higher levels while relying on far less data to do so, i.e. more efficient management than achieved under the historical data-hungry assessment approach. These simple harvest control rules are able to achieve almost the same catch and risk performance, even in the presence of relatively strong retrospective patterns: the slope and

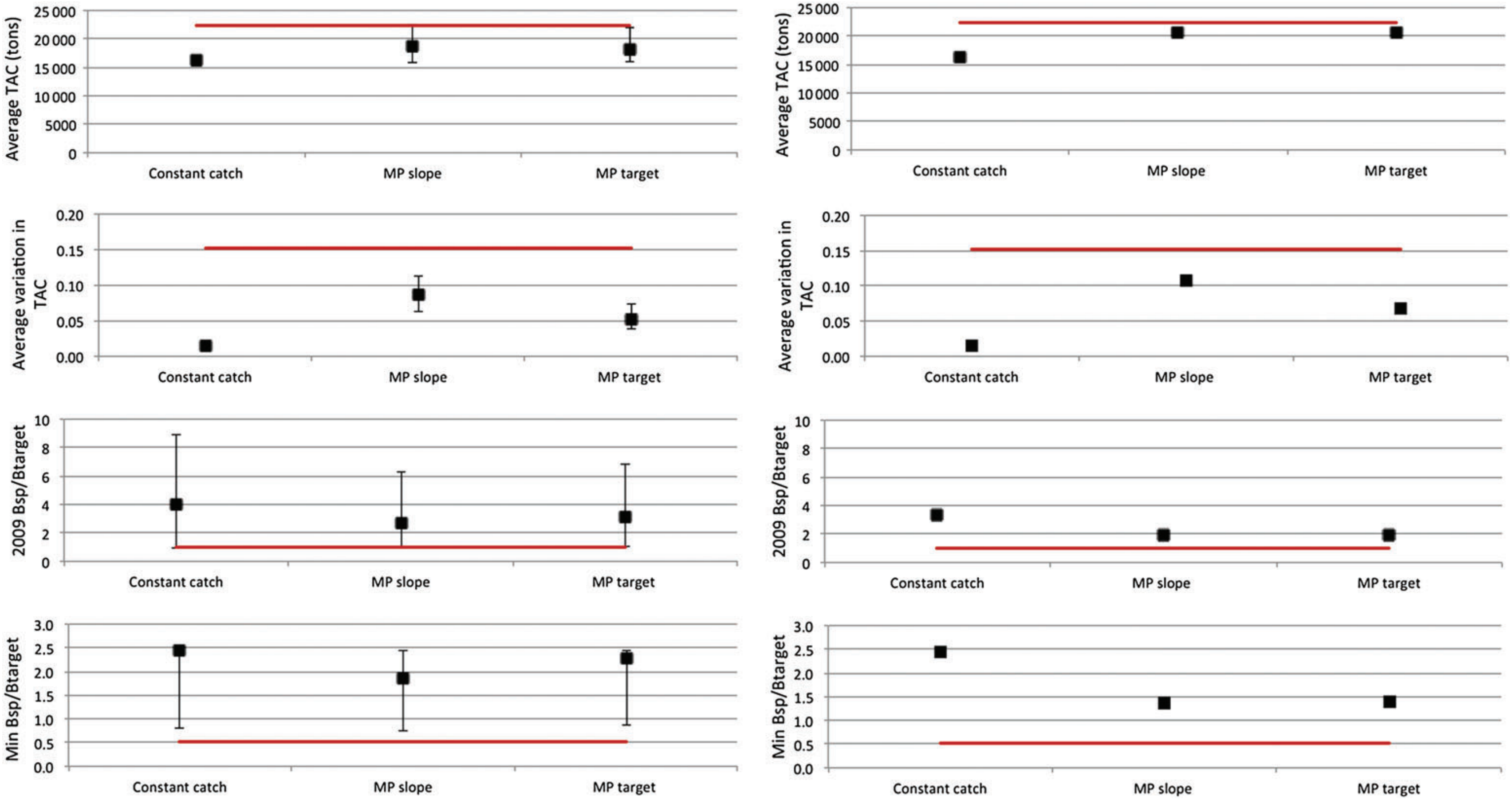

Figure 4. Summary performance statistics for stochastic (left) and deterministic (right) projections of the three candidate "forecast" MPs for North Sea sole. The medians and 95% probability intervals for a thousand simulations are shown on the plots on the left. From top to bottom: average annual TAC, average interannual variation in TAC, final spawning biomass as a fraction of the target level, and minimum future spawning biomass as a fraction of the target level. The thick horizontal lines indicate values when projecting under the catches which actually occurred.

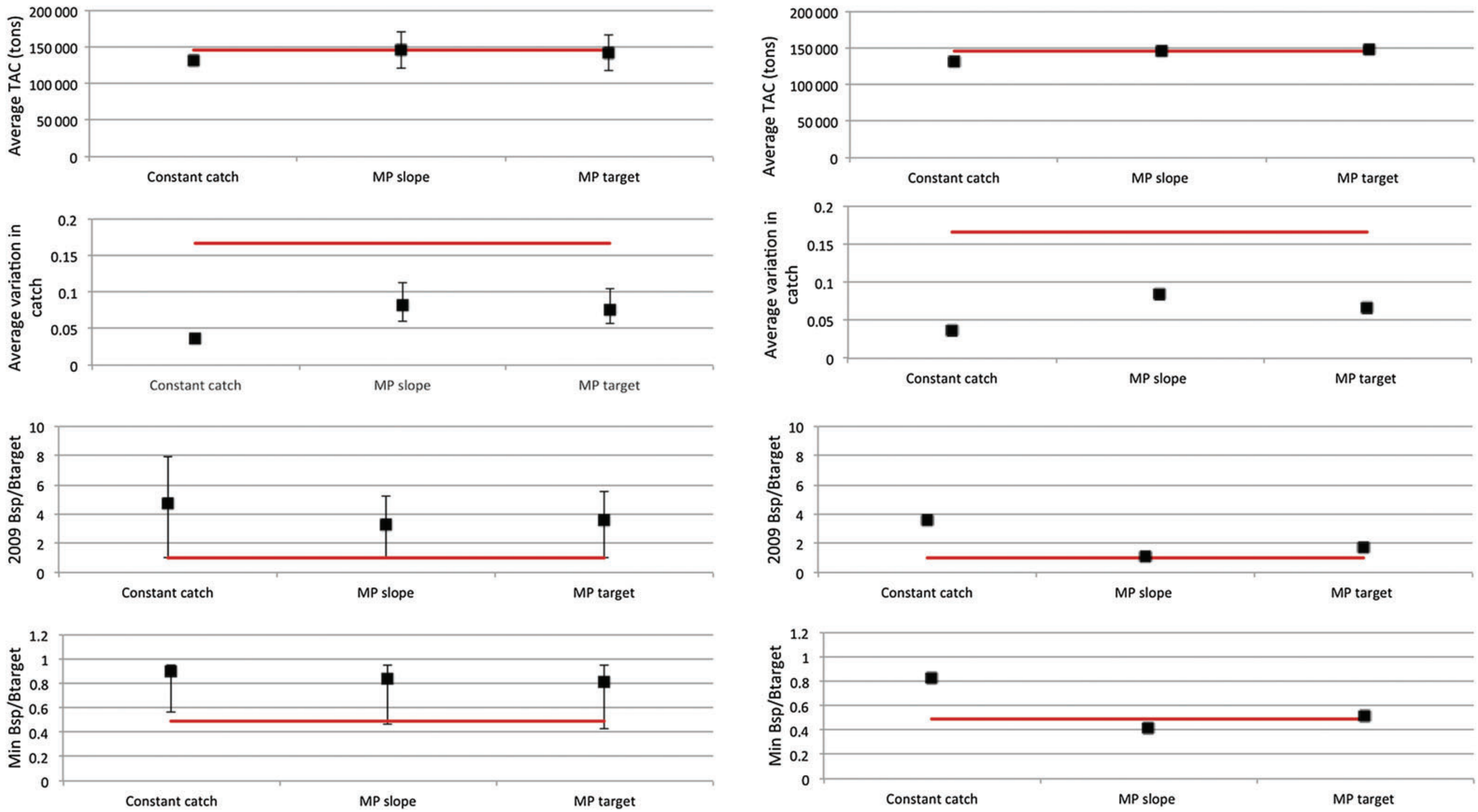

Figure 5. Summary performance statistics for stochastic (left) and deterministic (right) projections of the three candidate "forecast" MPs for North Sea plaice. The notation is as for Figure 4.

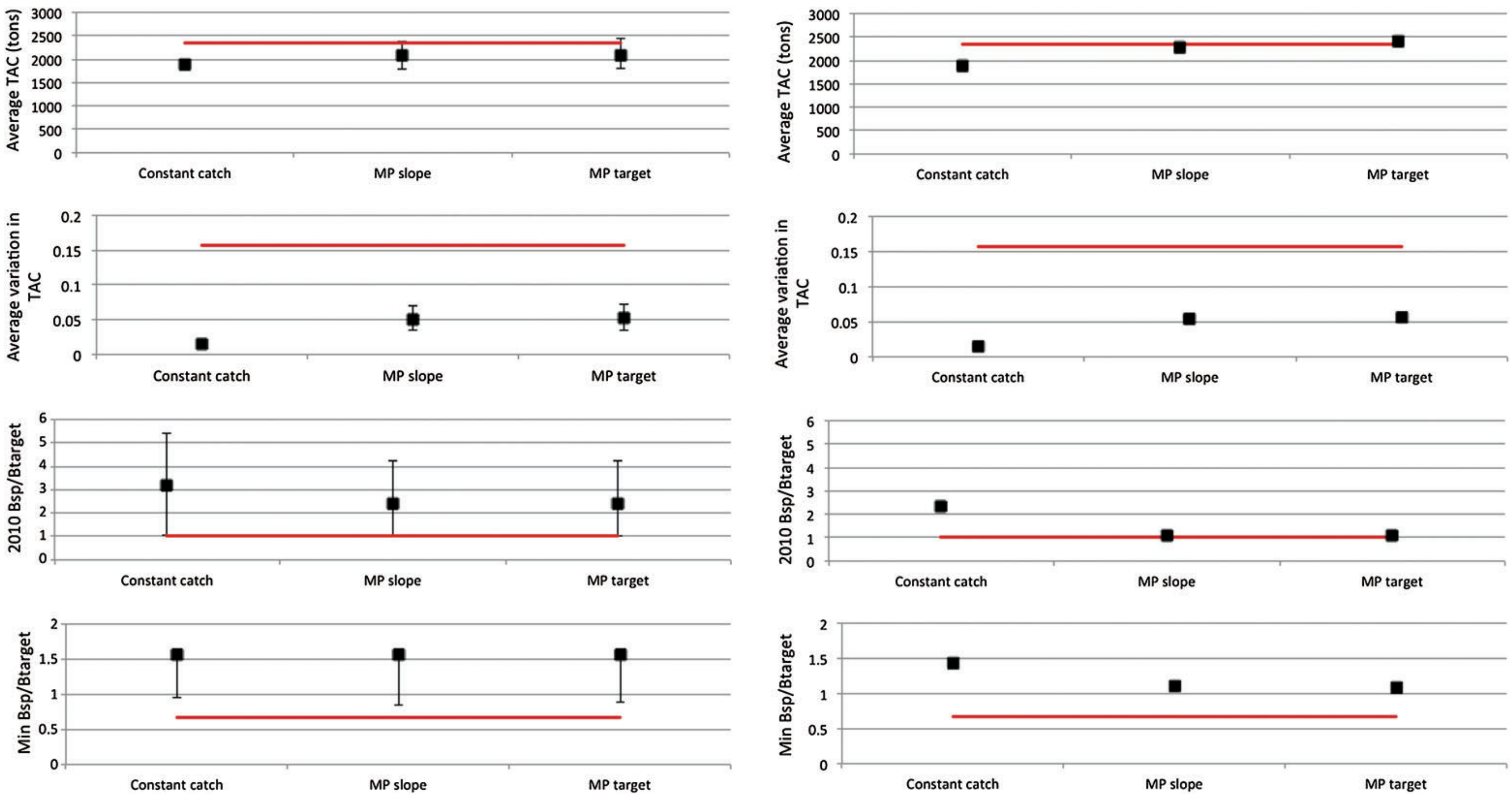

Figure 6. Summary performance statistics for stochastic (left) and deterministic (right) projections of the three candidate "forecast" MPs for New England witch flounder. The notation is as for Figure 4.

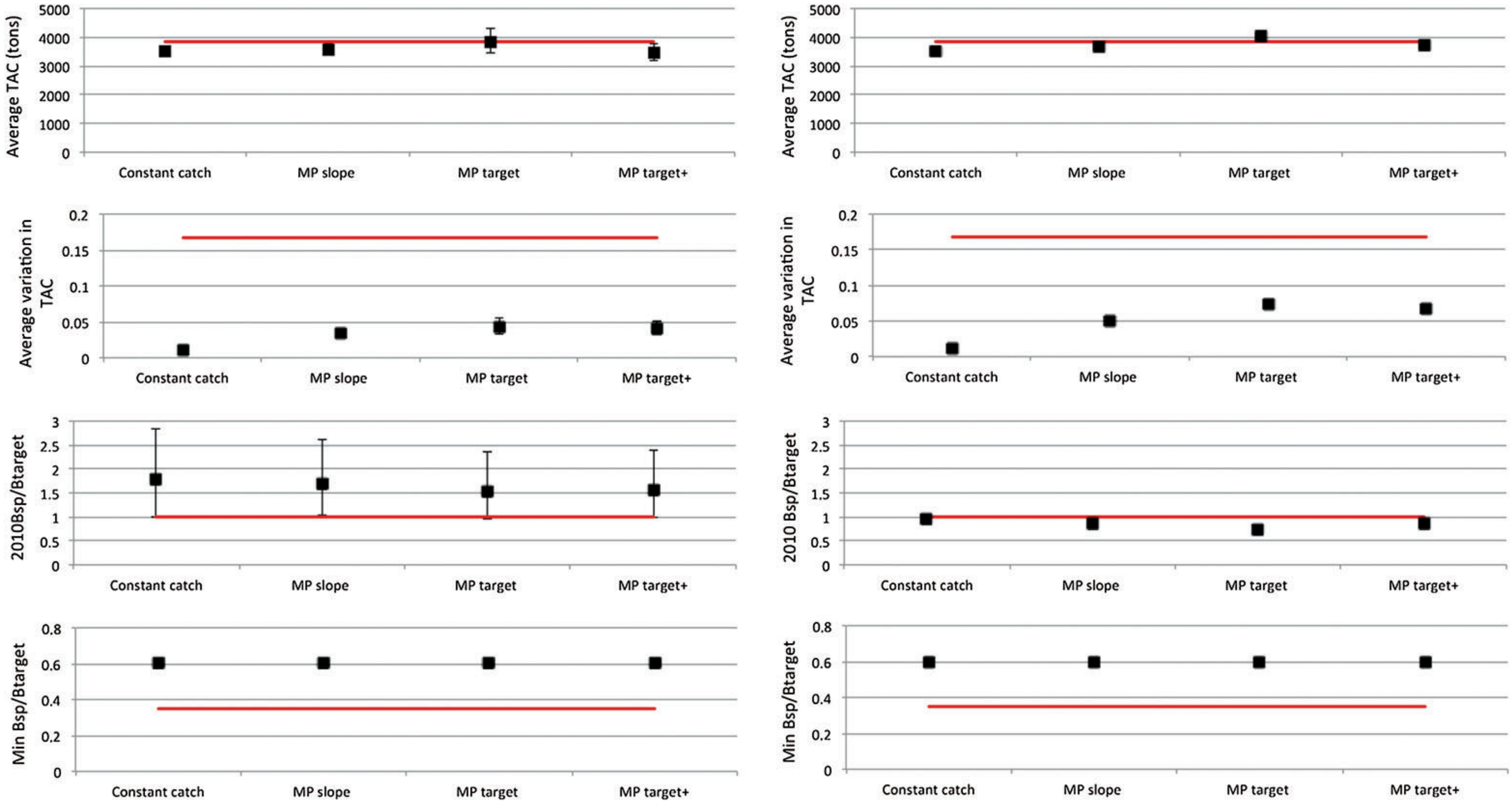

Figure 7. Summary performance statistics for stochastic (left) and deterministic (right) projections of the three candidate "forecast" MPs for American plaice. The notation is as for Figure 4. The "MP target+" candidate is retuned in mid-cycle (i.e. in 2000)—see text.

target MPs are able to adjust for these retrospective patterns to a large (though not complete) extent through feedback. The American plaice example does however also point to the need for expanded analyses which include robustness tests which incorporate a wider suite of uncertainties.

Furthermore, these results suggest that a 20-year MP cycle would probably be too long, as might have been expected, given that here a more realistic MP update period of 10 years leads to improved performance. Based on the South African experience, where MPs have been implemented with some success over the past two decades, the optimal MP cycle is ~4–5 years: a shorter cycle will lead to more workload without associated benefits and with increased interannual fluctuations in TAC. The IWC Scientific Committee uses a 5-year cycle.

Stock rebuilding: one MP or several assessments?

These simple retrospective comparisons suggest that a simple empirical MP, implemented over a 20-year period, could have given comparable, or better, performance than the potentially 20 annual VPA assessments actually conducted, even in the presence of strong retrospective patterns in the assessments.

This study has considered only empirical MPs. While model-based MPs might yield similar results, there are a number of reasons why the authors consider these less preferable in most circumstances, the chief of which is transparency of understanding and buy-in by the stake-holders (Butterworth, 2008)—see also Supplementary material, Section S.4 for a comprehensive discussion.

This study does not include implementation error in its retrospective comparisons (in terms of the interannual variation statistics), and therefore does not paint the full picture in this regard. However, a plot of annual catches and TACs (see Supplementary Figures SA1.1 and SA2.1) highlights typical volatility, not only in annual catches, but also in annual TACs when ongoing management advice is based on regular assessments. This would suggest that the assessment-based management typically leads to large seemingly unwarranted changes in TAC, with fluctuations caused by factors unrelated to the trend in resource biomass, but rather by policy decisions that differ in their basis from the underlying scientific advice. An MP approach goes a long way to repair such discontinuities by automating the decision-making process thereby sidestepping political interference that ultimately conflicts with the attainment of longer term management objectives.

In general, these results are similar to those obtained in simulation studies by Punt (1993), which showed that compared with simpler MP approaches for South African hake based on production models, attempts to take age-structure into account through VPA in recommending catch limits led to greater variability in those limits without any corresponding enhancement of performance in terms of reduced conservation risk. In line with these early indications, management of the South African data-rich hake resource has since moved away from age-aggregated model-based to empirical MPs (Rademeyer, 2012), for which it has been relatively simple to demonstrate that the latter achieve satisfactory performance under a wide range of robustness tests to key uncertainties (see also Supplementary material, Section S.5).

Furthermore, because of diminishing resources for research, there may well be difficulties in sustaining the level of input data (such as the annual ageing of the catch as required by VPA methods) and expertise required for complex annual assessments. This raises the question of whether such complex assessments can continue to serve as the primary basis to provide scientific advice on catch limits, so that there is a need to explore alternative possibilities as has been done here.

However, this is certainly not to suggest that complex assessments must be abandoned. Rather they still need to be conducted from time to time to provide reliable estimates of stocks, as well as to provide updated OMs for the MP review processes which are typically conducted every 4/5 years. Nevertheless, when the goal is to aim for a target biomass within a specified time frame, the MP approach seems best suited, particularly in terms of achieving greater stability in annual TAC advice.

To conclude, these initial results strongly suggest that simple empirical MPs may provide a defensible, simpler, and less costly alternative approach to the provision of scientific management advice based on frequent (possibly annual) updates of complex age- or length-based assessments.

Where next?

These restrospective comparisons were intentionally simple to highlight some of the key features of an MP approach. A comprehensive MP evaluation approach tailored to each of the stocks considered, and which takes account of the full range of uncertainty that is present in reality, would be a complex and time-consuming exercise which is beyond the scope of this paper. However, even without approaching such a fully comprehensive MP study, there are nevertheless some aspects of this work which might be usefully extended further:

(i) In the calculations above, the TAC specified by the MP is assumed to be exactly equal to the total removals for the year concerned. Realistic levels of implementation error need to be incorporated into projections. In particular, a distinction between landings and discards needs to be introduced. Furthermore, different components of implementation error (e.g. policy decisions that differ from scientific advice) need to be identified and simulated to render these retrospective studies more comparable with what actually occurred.

(ii) The stochastic projection trial exercise should be extended to incorporate more checks of robustness. Aspects to be considered for inclusion in such an extension include estimation and model structure uncertainty in the numbers-at-age vector that initiates the projections, variability in natural mortality over time, systematic changes in population weights, and changes in survey catchability.

(iii) In comparing performances above, "forecast" MPs were tuned to achieve the same final spawning biomass level at some low percentile (2.5%) of the distribution of this statistic. Trade-off comparisons might be more meaningful to stakeholders if instead tuning was effected to achieve the same total catch over the period.

(iv) At a later stage, if this approach finds wider favour, rather than demonstrations of adequacy based on history, the analyses will need to move on to consider simulations projecting forward from the present time, so as to develop MPs that could be seriously considered for implementation. This could involve extension beyond the very simple types of MPs considered here, and would require a wide range of robustness testing.

Supplementary material

Supplementary material is available at the ICESJMS online version of this manuscript. Section S.1 provides technical details of the OM, which is followed by Section S.2 which gives the input specifications and Section S.3 which provides the technical details of the MPs. Section S.4 presents a discussion comparing model-based and empirical MPs. The supplementary material concludes with Section S.5 which presents a complete set of results for each of the stocks considered. The Appendix of the supplementary material lists OM parameters and the assessment outputs that form the basis of the retrospective analyses.

Acknowledgements

We thank José De Oliveira, Charlie Edwards, and Laurie Kell for assistance in providing the assessment data which we have used. David Miller is thanked for useful comments and constructive criticism on an earlier version of this work, as are José De Oliveira and an anonymous reviewer. Financial assistance of the National Research Foundation (NRF) of South Africa towards this research is hereby acknowledged.

References

Brodziak, J., Cadrin, S. X., Legault, C. M., and Murawski, S. A. 2008. Goals and strategies for rebuilding New England groundfish stocks. Fisheries Research, 94: 355–366. doi:10.1016/j.fishres

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

Butterworth, D. S. 2008. Some lessons from implementing management procedures. Fisheries for Global Welfare and Environment, 5th World Fisheries Congress 2008, pp. 381–397.

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

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

De Oliveira, J. A. A., Kell, L. T., Punt, A. E., Roel, B. A., and Butterworth, D. S. 2009. Managing without best predictions: the Management Strategy Evaluation framework. *In* Advances in Fisheries Science. 50 Years on from Beverton and Holt, pp. 104–134. Blackwell Publishing, Oxford. 2008.

Geromont, H. F., de Oliveira, J. A. A., Johnston, S. J., and Cunningham, C. L. 1999. Development and application of management procedures for fisheries in southern Africa. ICES Journal of Marine Science, 56: 952–966.

ICES. 2010. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK), 5–11 May 2010, Ices Headquarters, Copenhagen. ICES CM 2010/ACOM: 13. 1058 pp.

ICES. 2012. Report on the Classification of Stock Assessment methods developed by SISAM. ICES CM 2012/ACOM/SCICOM: 01. 15 pp.

Miller, D. C. M., and Poos, J. J. 2010. Combined ex post and ex ante evaluation of the long term management plan for sole and plaice in the North Sea, including responses to ICES re-view. ICES Document CM 2010/ACOM: 62. 109 pp.

NEFSC (Northeast Fisheries Science Center). 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Department of Commerce, Northeast Fisheries Science Center Reference Document, 12-06. 789 pp. http://www.nefsc.noaa.gov/nefsc/publications/ (last accessed August 2013).

Punt, A. E. 1993. The comparative performance of production-model and ad hoc tuned VPA based feedback-control management procedures for the stock of Cape hake off the west coast of South Africa. *In* Risk Evaluation and Biological Reference Points for Fisheries Management, pp. 283–299. Ed. By S. J. Smith, J. J. Hunt, and D. Rivard. Canadian Special Publication of Fisheries and Aquatic Sciences, 120.

Punt, A. E. 2006. The FAO precautionary approach after almost 10 years: have we progressed towards implementing simulation-tested feedback-control management systems for fisheries management? Natural Resource Modeling, 19: 441–464.

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.

Rademeyer, R. A. 2012. The evolution of management procedures for the South African hake resource in the 2000s. PhD thesis, University of Cape Town. 381 pp. http://www.mth.uct.ac.za/maram/.

Wayte, S. E. (Ed.) 2009. Evaluation of new harvest strategies for SESSF species. CSIRO Marine and Atmospheric Research, Hobart and Australian Fisheries Management Authority, Canberra. 137 pp.

Wigley, S. E., Brodziak, J. K. T., and Col, L. 2003. Assessment of the Gulf of Maine and Georges Bank witch flounder stock for 2003. Northeast Fisheries Science Center Reference Document, 03-14. 186 pp.

Handling editor: Ernesto Jardim