Agenda Item G.7.a
Attachment 2
September 2013

Management Strategy Evaluation for Rebuilding Revision Rules: A Proof of Concept

André E. Punt
School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195 USA

Summary

Rebuilding Revision Rules are decision rules which relate to how to change management actions in response to changes in the outcomes of Rebuilding Analyses. It is possible to develop many alternative Rebuilding Revision Rules (and they can be combined, with various components perhaps interacting in unexpected ways). This document outlines a management strategy evaluation which could be used to compare alternative Rebuilding Revision Rules. The document is primarily a "proof of concept" to show that it is possible to develop the MSE. However, illustrative results are shown for two example strategies.

Introduction

Rebuilding Plans are required for stocks which are assessed to be below the Minimum Stock Size Threshold (MSST). Rebuilding Plans involve identifying a target year for recovery to the proxy for B_{MSY} (denoted T_{TARGET}), and the adjusting fishing mortality (usually expressed as a Spawning Potential Ratio, SPR, for US west coast groundfish stocks) so that recovery occurs at or before T_{TARGET}. T_{TARGET} is constrained to be less than T_{MAX}, which is 10 years after the stock was declared overfished or one mean generation time plus the time to recover to the B_{MSY} proxy if there were no future catches (i.e. T_{MIN} plus one mean generation).

The biological information on which to select T_{TARGET} (T_{MAX} is a biological concept) relates to projections of rates of recovery to the proxy for B_{MSY} under different harvest strategies (usually levels of constant SPR, but perhaps also with a phase-in). The results of projections under different harvest strategies (and sometimes alternative states of nature) are referred to as a "Rebuilding Analysis".

Rebuilding Plans need to be reviewed every two years. For stocks with no new information, or for which no assessment is conducted, this may involve comparing the actual catches with those expected under the Rebuilding Plan. However, the information on a rebuilding stock does generally change over time. The changes can be "mild", such as the addition of new survey results and fishery length and age data, or "severe" such as a change to the assumed rate of natural mortality, the steepness of the stock recruitment relationship or the time-series of historical catches. Consequently, the results from projections will change even if management has followed the currently-adopted Rebuilding Plan[1]. Rebuilding Revision Rules involve (1) assessing adequacy of progress toward rebuilding and (2) altering Rebuilding Plans, given a change in stock status (Punt and Ralston, 2007).

There are, however, many possible Rebuilding Revision Rules, and it is not clear how they are likely to perform. Management Strategy Evaluation (MSE, Smith, 1994) involves using simulation testing to evaluate the performance of candidate management rules. In this context, management rules are the combination of the assessment method used to estimate stock status and productivity, and the control rules which translate these estimates into allowable catch levels (the "Rebuilding Revision Rules").

This document describes a MSE framework which could be used to evaluate alternative Rebuilding Revision Rules. As such, it provides a "proof of concept" to allow the Council

[1] Note that there will always be changes to projection outcomes because projections are conducted under the assumption that future recruitment will be randomly distributed about some average value. A new assessment will provide information on (recent) recruitments which would in the past have assumed to have been random.

and its advisory bodies to decide whether they wish to use MSE to compare alternative Rebuilding Rules. The document outlines a (simple) operating model, how assessments are conducted and two example Rebuilding Revision Rules. It also provides some example diagnostic plots and tables which could be used to quantify the performance of Rebuilding Revision Rules relative to Council goals and legal mandates.

Methods

Overview

The Management Strategy Evaluation involves three key components: (a) the operating model, which reflects the "truth" for the simulations, (b) the assessment method, and (c) the harvest control rules. An assessment is conducted every 4th year, and the subsequent application of the harvest control rules lead to catch limits (which are assumed to be taken exactly) for the next four years.

Operating model

The operating model is an age- and sex-structured population dynamics model (Appendix A). The historical level of fishing effort is pre-specified while the catchability coefficient is selected so that the spawning biomass when assessments are first conducted is 10% of B_0. The data available for assessment purposes are the catches, a survey index of relative abundance, and fishery and survey age-composition data. The data on which the illustrative example is based are assumed to be highly informative. It is assumed that steepness is known exactly as is natural mortality and weight- and fecundity-at-age.

Assessment method

The assessment is based on Stock Synthesis 3 (Methot and Wetzel, 2013). The parameters which are estimated each time an assessment is conducted are B_0, the parameters which define logistic selectivity functions for the fishery and the survey, and the annual deviations about the stock-recruitment relationship.

Harvest control rule

Two harvest control rules are evaluated. The two harvest control rules: "50%-rule" and "flexible" have several features in common, as outlined below, but differ in terms of how they deal with cases in which a Rebuilding Plan is implemented and has not "failed". The harvest control rule for setting harvest specifications for year *y* (based on assessment with data up to year *y*-1) is:

A. If $S_y / S_0 \geq 0.4$, the stock is rebuilt so set catch limits to the ABCs. Stop (denoted Case X).
B. If $0.25 \leq S_y / S_0 < 0.4$ and the stock is not currently under a Rebuilding Plan, set the catch limits based on the 40-10 harvest control rule. Stop (denoted Case -X).
C. If $S_y / S_0 < 0.25$ and the stock is not currently under a Rebuilding Plan:
    a. Conduct a rebuilding analysis to determine T_{MAX}.
    b. Determine T_{TARGET} so that there is a 0.6 probability of rebuilding to $0.4B_0$ by T_{MAX}.
    c. Set $SPR_{current}$ so that the probability of rebuilding by T_{TARGET} is 0.5.
    d. Calculate the catch limits for the next four years based on $SPR_{current}$.
    e. Stop (denoted Case 0).

D. If S_y / S_0 < 0.4, the stock is currently under a Rebuilding Plan, and the year is beyond T_{MAX}, "Reboot", i.e. start a new Rebuilding Plan (reset T_{MAX} and T_{TARGET}), but do not let SPR increase as a result of this. Stop (denoted Case 3).

E. If S_y / S_0 < 0.4, the stock is currently under a Rebuilding Plan, and the year is beyond T_{TARGET} but not T_{MAX}:
   a. Set T_{TARGET} equal to T_{MAX} and try to find the SPR so that the probability of rebuilding to $0.4B_0$ is 0.5 by the new T_{TARGET}.
   b. If this SPR exists then set $SPR_{current}$ and compute catch limits for the next four years. Stop.
   c. If this SPR does not exist then "reboot", i.e. start a new Rebuilding Plan (reset T_{MAX} and T_{TARGET}), but do not let SPR increase as a result of this. Stop.

F. If S_y / S_0 < 0.4, the stock is currently under a Rebuilding Plan and the year is not yet T_{TARGET}:
   a. Strategy "flexible"
      i. Project forward under $SPR_{current}$.
      ii. If the probability of rebuilding to $0.4B_0$ at least 0.4, set the catch limits for the next four years based on $SPR_{current}$. Stop (denoted Case 1).
      iii. If there is an SPR which corresponds to a 0.5 probability of rebuilding to $0.4B_0$ by T_{MAX}, set $SPR_{current}$ to this SPR, reset T_{TARGET} to T_{MAX}, and compute catch limits for the next four years. Stop (denoted Case 2).
      iv. "Reboot", i.e. start a new Rebuilding Plan (reset T_{MAX} and T_{TARGET}), but do not let SPR increase as a result of this. Stop.
   b. Strategy "50% rule"
      i. If there is an SPR which corresponds to a 0.5 probability of rebuilding to $0.4B_0$ by T_{TARGET}, set $SPR_{current}$ to this SPR, compute catch limits for the next four years. Stop (denoted Case 5).
      ii. If there is an SPR which corresponds to a 0.5 probability of rebuilding to $0.4B_0$ by T_{MAX}, set $SPR_{current}$ to this SPR, reset T_{TARGET} to T_{MAX}, and compute catch limits for the next four years. Stop (denoted Case 2).
      iii. "Reboot", i.e. start a new Rebuilding Plan (reset T_{MAX} and T_{TARGET}) but do not let SPR increase as a result of this. Stop.

Performance metrics

The performance metrics can be divided into those which are graphical and those which are numerical. An example graphical summary is provided and discussed below.

Punt and Ralston (2007) note that there are many statistics that could be used to summarize the performance of a management strategy. They focused on five principal management goals: (a) a high probability of the stock recovering by the T_{MAX} selected when the Rebuilding Plan was originally developed[2], (b) high catches during rebuilding, (c) low inter-annual variation in catches, (d) stability in the Rebuilding Plan (i.e., minimizing changes to the value of T_{MAX}[2]), and (e) simplicity. Punt and Ralston (2007) noted that the first three of these five goals are typical of those commonly selected when conducting an MSE. The fourth goal is included because it measures the "administrative cost" of a management strategy; changing the SPR used to set the harvest guideline and changing harvest guidelines themselves is relatively straightforward administratively. In contrast, a reboot may require an amendment to the Fishery Management Plan. Punt and Ralston (2007) argue that the

[2] Given clarity in what is required in a Rebuilding Plan since Punt and Ralston (2007) was published, T_{MAX} would need to be replaced by T_{TARGET} if a metric along these lines was considered, and this is what was done for the example application.

importance of the goal of simplicity cannot be overstated. It is likely that the PFMC would select a simple set of Rebuilding Revision Rules over a more complicated set even if the performance of the more complicated set was marginally better than that of the simple set, purely because of the need for the public to know how decisions are made regarding the management of overfished stocks.

The (example / illustrative) performance measures used to quantify these five goals are:

1. The "rebuilding ratio," the ratio of the number of years before the stock was assessed to be rebuilt divided by the number of years that it was expected that rebuilding would take based on the original Rebuilding Plan, i.e., if the rebuilding ratio exceeds unity then rebuilding is perceived to have taken longer than originally expected.
2. A measure of the variability of the catches (abbreviation AAV), defined as:

$$AAV = 100 \frac{\sum_y \left| C_y - C_{y+1} \right|}{\sum_y C_y} \qquad (1)$$

   where C_y is the catch during year y.
3. The average catch during the years when the resource was under a Rebuilding Plan.
4. The average catch during the first ten years of the rebuilding period.
5. The number of times it was necessary to "reboot".

Punt and Ralston (2007) argue that the rebuilding ratio should be based on the perception that the stock has rebuilt, rather than the stock having actually rebuilt. This is because this performance measure relates to what the decision makers would actually see. The performance measures should include both short- and long-term catches because the short-term catch reflects the likely immediate impact of the fishery.

Results and Discussion

Illustrative results for single simulations

Figures 1 and 2 explore the performance of the two candidate Rebuilding Revision Rules. The plots can be interpreted as follows

A. Upper Left Panel. This panel shows the true (i.e. operating model) time trajectory of spawning biomass relative to B_0 (solid line), along with the target biomass ($0.4B_0$) and the threshold which defines an overfished groundfish stock ($0.25B_0$). It also shows the estimates of spawning biomass relative to B_0 from each assessment[3]. For the simulation in Figure 1, spawning biomass is underestimated in the first assessment then overestimated for next eight assessments. In contrast, spawning biomass is underestimated for many years in the simulation in Figure 2. One consequence of the latter outcome is that the stock assessment does not detect that the stock has rebuilt for three assessment cycles after it actually rebuilds to B_{MSY}. The bars at the top of this panel illustrate the changing nature of the Rebuilding Plan. The horizontal line ranges from when the stock was declared overfished to the current $\mathrm{T}_{\mathrm{MAX}}$ (the open circle). $\mathrm{T}_{\mathrm{TARGET}}$ is indicated by the closed circle. A "reboot" to the Rebuilding Plan is indicated by a change to the start of the horizontal line. The closed and open circles are the same for the fourth line in this panel in Figure 3; this reflects a case when $\mathrm{T}_{\mathrm{TARGET}}$ was increased to $\mathrm{T}_{\mathrm{MAX}}$ given it was found that rebuilding to $\mathrm{T}_{\mathrm{TARGET}}$ could not occur with 50% probability (Case 2 above). There are periods in which there is no horizontal line in Figure 2 because the stock is assessed to have been rebuilt to B_{MSY} (so catches are based on the 40-10 rule).

[3] The assessment produces a time-series of estimates of biomass but the estimate for the last year in most critical to management so only this value is shown.

B. Upper Right Panel. This panel shows the time-trajectories of catches (the vertical dotted line indicates when the stock was first declared overfished). As expected, there can be major changes in catch every four or so years. The very low catches at the end of the projection period in Figure 1 reflect the situation where rebuilding can occur, but only if the fishery is effectively closed.
C. Lower Left Panel. This panel shows the SPR used to set the catch limit as a function of time. The symbols pertain to the case under consideration. All simulations start with a "0", indicating the start of a Rebuilding Plan. A "1" (or a "5") means that no major changes to the Rebuilding Plan were needed (the SPR was kept its current value ("1"; "flexible" strategy) or changed so that the probability of rebuilding was 0.5 ("5"; "50-rule" strategy)), whereas a "2" means that T_{TARGET} was changed to T_{MAX}. A "3" in this panel indicates that a reboot took place.
D. Lower Right Panel. This panel shows the time-trajectory of the probability of rebuilding. As expected, this probability is (generally) 0.5 for the "50%-rule" strategy (Figure 1), but can be as low as 0.4 for the "flexible" strategy (Figure 2).

The results in Figures 1 and 2 are meant to be illustrative (the results for the two strategies are not even for the same simulations). However, they are meant to help the Council and its Advisory Bodies to better understand how the various components of a Rebuilding Revision Rule interact.

Illustrative results for 100 simulations

Figures 3 and 4 summarize the results of 100 simulations in terms of histograms of the five performance metrics outlined above. The ideal set of Rebuilding Revision Rules should have a rebuilding ratio of 100%, i.e. the stock rebuilt in the year it was predicted to rebuild in when the Rebuilding Plan was developed (or before this year, reflecting recovery "as soon as possible"), a low value for the AAR statistic, which measures the extent to which catches vary from one year to the next (a value for this statistic of 20%, means that catches change on average 20% from one to the next), and a value of zero for number of times it was necessary to reboot the Rebuilding Plan. The catches should be as high as possible (and preferably not vary much among simulations). It would also be expected that the average catch during the first ten years of the Rebuilding Plan would be less than those over the entire rebuilding period given the focus on harvest strategies with constant fishing mortality.

Less than 50% of simulations rebuilt to B_{MSY} by the T_{TARGET} selected when the Rebuilding Plan was developed for the "50% rule" strategy (Figure 3). The results for the 'flexible" strategy suggest that rebuilding to B_{MSY} occurs close to when it was expected to occur on average, but there is considerable uncertainty in this (Figure 4). In addition, ~75% of Rebuilding Plans had no "reboots", but ~20% had one reboot and ~5% two reboots for the "flexible" strategy whereas only ~60% of Rebuilding Plans were not "rebooted" for the "50%-rule" strategy and three reboots occurred in one simulation. Catch variability was generally higher for the "50%-rule" strategy than for the "flexible" strategy (contrast the distributions for the AAV statistic in Figures 3 and 4).

Next steps and overall discussion

The framework outlined in this document represents an ideal situation which will over-estimate the performance of candidate Rebuilding Revision Rules. In particular, the stock assessment is structurally correct (no errors in pre-specified parameters such as steepness and *M*) while the data are high quality and available for all years. More realistic scenarios will have higher rates for the need for reboots, for example. However, these simulations provide a basis for estimating a baseline rate of, for example, reboots against which more sophisticated rebuilding revision rules can be compared.

There are three main next steps if the Council decides to pursue use of MSE to evaluate Rebuilding Revision Rules:

- Specification (or confirmation) of the performance metrics and plots. Performance metrics should be selected to capture performance relative to Council goals and objectives. The performance metrics and plots on which the analyses of this document were based were taken from Punt and Ralston (2007) and do not reflect Council (or Advisory Body) deliberations.
- Specification of additional Rebuilding Revision Rules. While it is not technically feasible to evaluate hundreds of Rebuilding Revision Rules (100 simulations for 80 years for one Rebuilding Revision Rule takes ~3-4 days on a fast desktop computer), more sets should be considered. Possible factors to consider in candidate Rebuilding Revision rules are: (a) not changing the SPR during the last xx years of the Rebuilding period, (b) only allowing the probability of rebuilding to lag behind 0.5 for one (or two) assessment cycles, etc. Martin Dorn has suggested that track should be kept of the projected year of rebuilding and "adequate progress towards rebuilding" defined as the projected time to rebuild to B_{MSY} under the current SPR (T_{REBUILD}) being between 25% and 75% of (the current) T_{TARGET}.
- Specification of operating models. The operating model of this document is unrealistic for several reasons, including that it assumes unrealistically good data, but also because the biological parameters are fairly unrealistic (relatively high M with very low steepness), having been chosen for illustrative purposes. The operating model does not include "black swan" events such as that a major change impacting the assessment (e.g., the pre-specified values assumed for steepness or M change). Such events are known to impact actual Rebuilding Plans and will be necessary to evaluate.

Acknowledgements

Chantel Wetzel (NWFSC, UW) and Martin Dorn (AFSC) are thanked for comments on an early draft of this document.

References

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

Punt, A.E. and S.V. Ralston. 2007. A Management Strategy Evaluation of rebuilding revision rules for overfished rockfish species. p. 329-351. In: J. Heifetz, J. DiCosimo, A.J. Gharrett, M.S. Love, V.M. O'Connell, and R.D. Stankey [Ed.] *Biology, Assessment and Management of North Pacific Rockfishes***.** Alaska Sea Grant College Program, University of Alaska Fairbanks.

Smith, A.D.M. 1994. Management strategy evaluation: The light on the hill. In: D.A. Hancock (ed.), Population dynamics for fisheries management. Australian Society for Fish Biology, Perth, Western Australia, pp. 249-253.

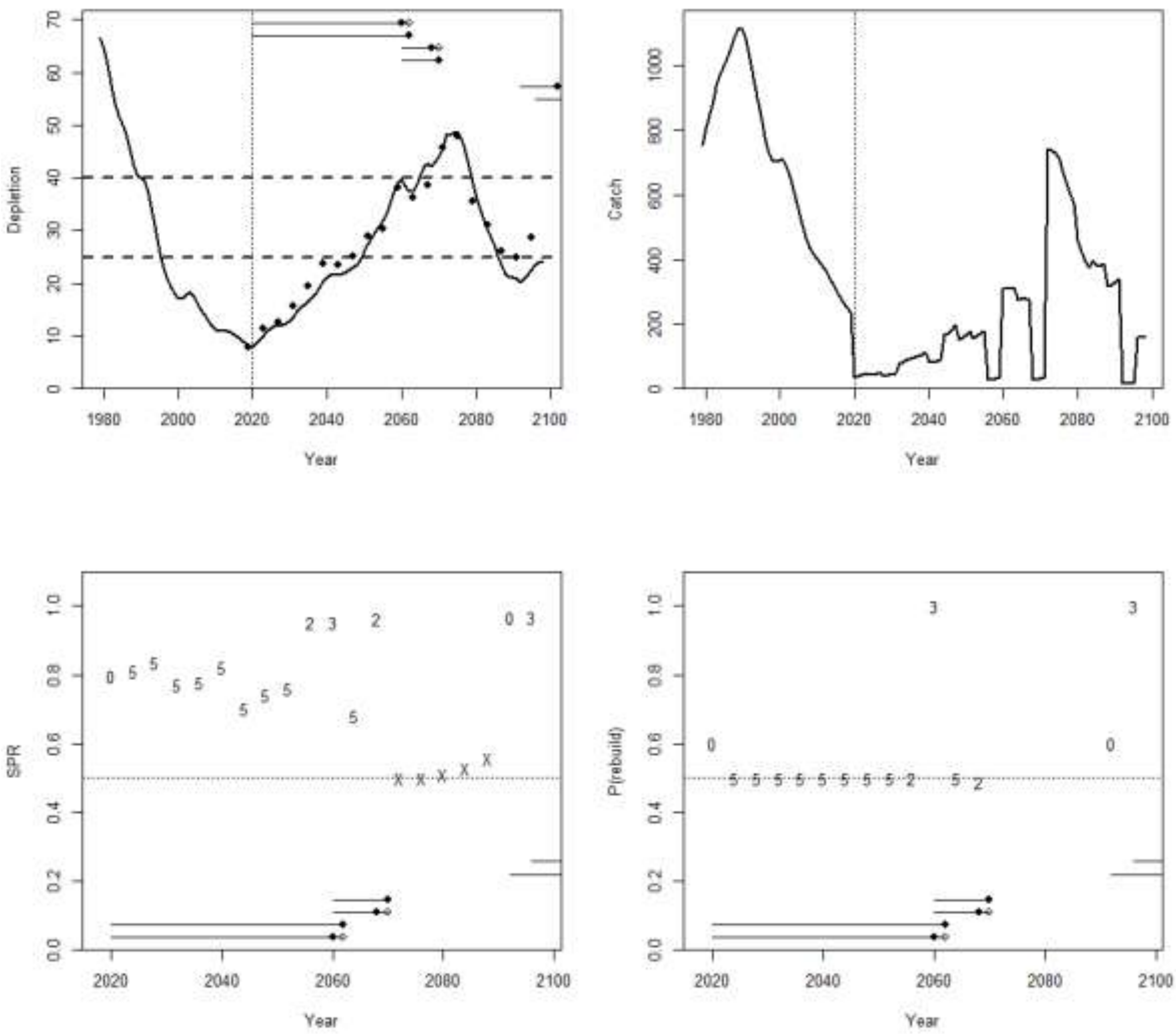

Figure 1. Plot showing overall performance for a single simulation. The Rebuilding Revision Rule is "50% rule" for this analysis. The numbers in the lower two panels indicate the outcomes of applying the Rebuilding Revision Rules every fourth year from 2020. The "3" in the lower panels indicates a year in which the Rebuilding Plan was "rebooted".

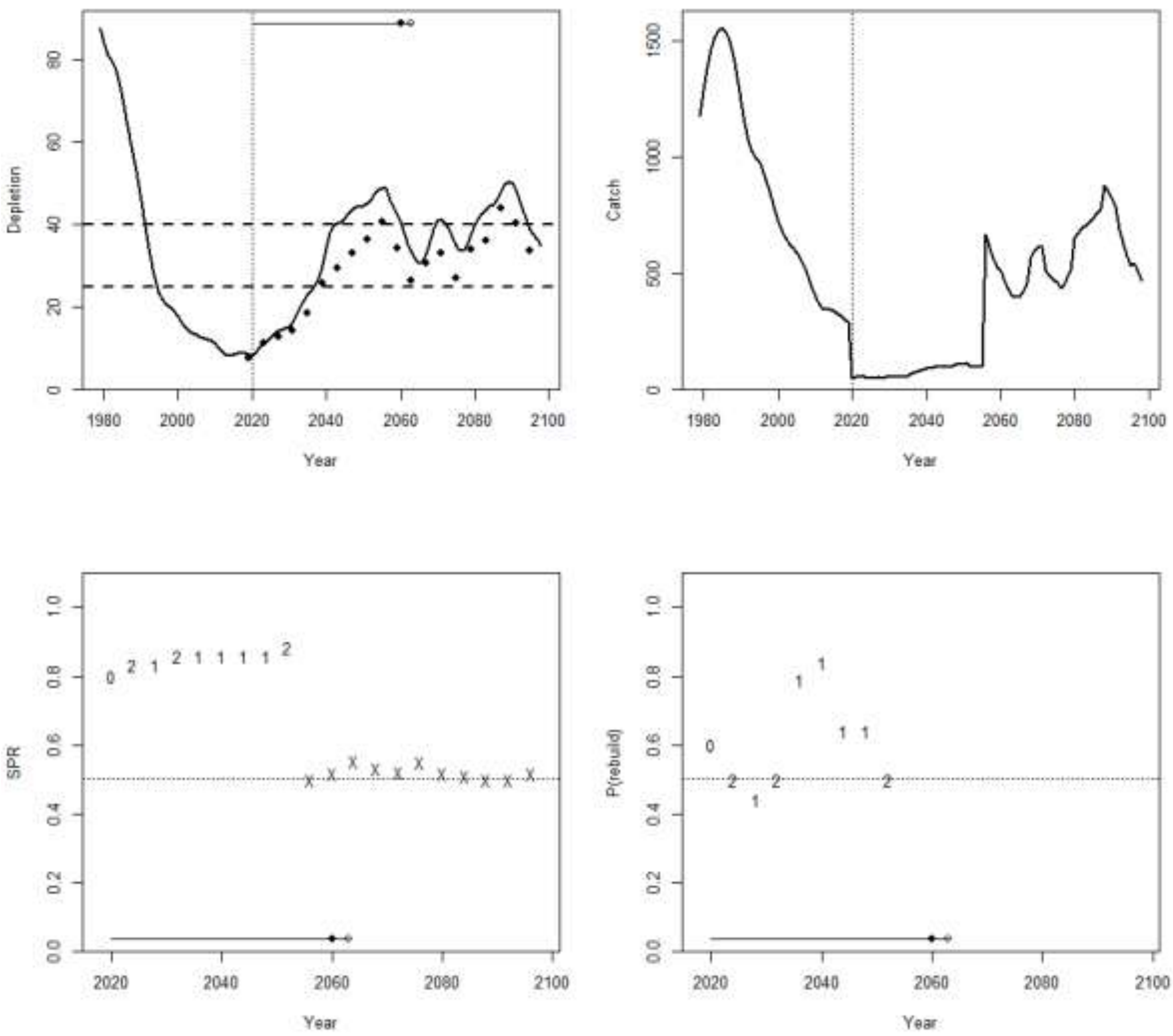

Figure 2. Plot showing overall performance for a single simulation. The Rebuilding Revision Rule is “flexible” for this analysis.

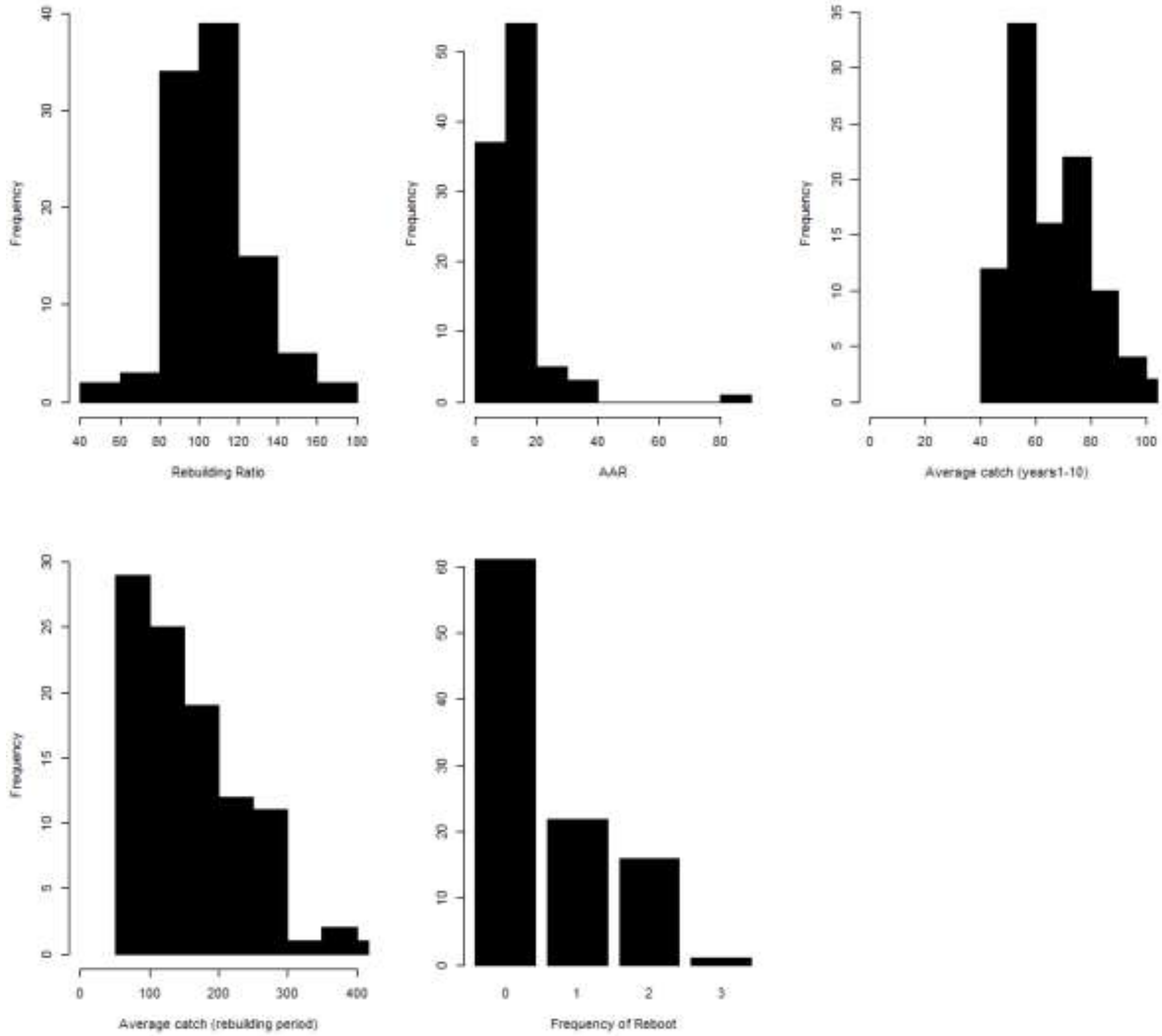

Figure 3. Histograms of performance metrics for the "50% rule" strategy based on 100 simulations.

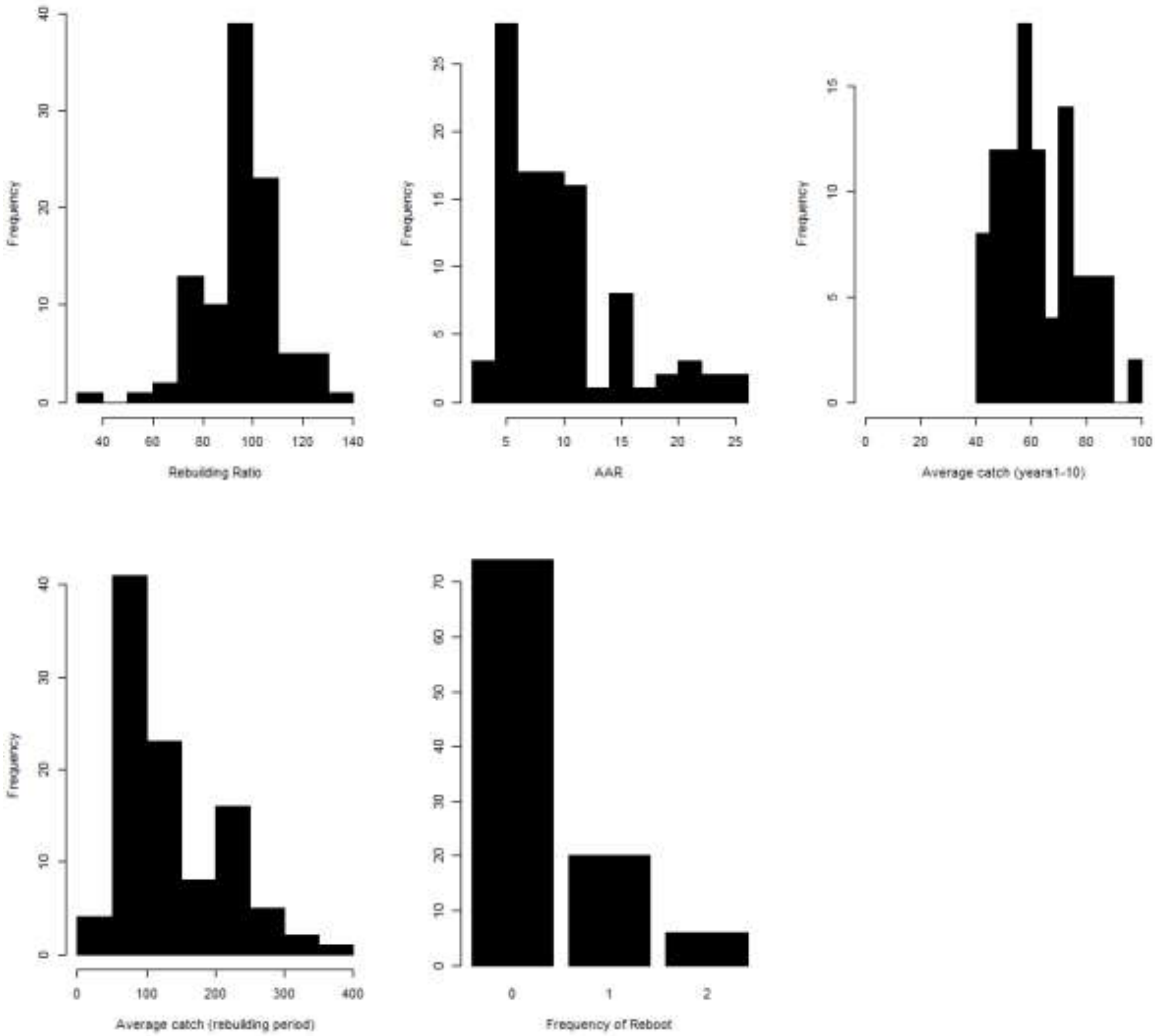

Figure 4. Histograms of performance metrics for the “flexible” strategy based on 100 simulations.

Appendix A: The Operating Model

Basic Dynamics

The dynamics of the simulated population are governed by the equation:

$$N^s_{y+1,a} = \begin{cases} 0.5\,R_y & \text{if } a = 0 \\ N^s_{y,a-1}\, e^{-Z^s_{y,a-1}} & \text{if } 1 \le a \le x-1 \\ N^s_{y,x-1}\, e^{-Z^s_{y,x-1}} + N^s_{y,x}\, e^{-Z^s_{y,x}} & \text{if } a = x \end{cases} \quad \text{(A.1)}$$

where $N^s_{y,a}$ is the number of animals of sex *s* and age *a* at the start of year *y*, $Z^s_{y,a}$ is total mortality for animals of sex *s* and age *a* during year *y*:

$$Z^s_{y,a} = M^s + S_a F_y \quad \text{(A.2)}$$

M^s is the instantaneous rate of natural mortality for animals of sex *s*, S_a is the selectivity on animals of age *a*, F_y is the fully-selected fishing mortality, R_y is the recruitment during year *y*, and *x* is the plus group age.

The annual recruitment is governed by a Beverton-Holt stock-recruitment relationship, parameterized in terms of steepness (*h*) and average unfished recruitment (R_0), i.e.:

$$R_y = \frac{4hR_0(B^S_y / B^S_0)}{(1-h)+(5h-1)(B^S_y / B^S_0)} e^{\varepsilon_y - \sigma^2_R/2} \quad \text{(A.3)}$$

where B^S_y is the spawning stock biomass at the start of year *y*:

$$B^S_y = \sum_{a=1}^{x} f_a\, w_a\, N^{\text{fem}}_{y,a} \quad \text{(A.4)}$$

f_a is the proportion of females which are mature at age *a*, and w_a is the weight of an animal of age *a* at the start of the year.

The catch (in weight) during year *y*, $\tilde{C}_y$, is given by:

$$\tilde{C}_y = \sum_s \sum_{a=0}^{x} w_{a+1/2}\, C^s_{y,a} = \sum_s \sum_{a=0}^{x} w_{a+1/2} \tfrac{S_a F_y}{Z^s_{y,a}} N^s_{y,a}(1-e^{-Z^s_{y,a}}) \quad \text{(A.5)}$$

where $w_{a+1/2}$ is the weight of animals of age *a* in the middle of the year.

The stock is assumed to be in an unfished state at the start of the first year with catches. The operating mortality is projected for 50 years. The time-series of effort, E_y, is given in Figure A.1. Fully-selected fishing mortality for these 50 years is given by $F_y = qE_y$ where the catchability coefficient *q* is selected so that the relative spawning at the start of year 51, B^S_{51} / B^S_0, equals a pre-specified value.

Data generation

The data available for assessment purposes are the annual catches in weight (assumed to be known exactly), a survey index of biomass, and the age-composition of the fishery and survey catches. For the purposes of this example application, survey data are available for all years, and catch age-composition data are available for all years in which the catch is non-zero.

The survey index (and the expected survey age-compositions) is given by:

$$I_y = \sum_s \sum_{a=0}^{x} w_{a+1/2} \tilde{C}^s_{y,a} e^{\eta_y - \sigma_\eta^2/2} = \sum_s \sum_{a=0}^{x} w_{a+1/2} \tilde{S}_a N^s_{y,a} e^{-Z^s_{y,a}/2} e^{\eta_y - \sigma_\eta^2/2} \quad \text{(A.6)}$$

where $\tilde{S}_a$ is the survey selectivity at age *a*.

The survey age-compositions are assumed to be multinomially distributed about the true survey age-proportions, while the fishery age-compositions are assumed to be multinomially distributed about the true fishery age-compositions (Equation A.5).

Table A.1. Values for the parameters of the operating model

Parameter	Values
Biological Parameters	
Natural mortality: females	0.15yr^{-1}
Natural mortality: males	0.2yr^{-1}
Fecundity-at-age	Figure A.2
Weight-at-age	Figure A.2
Selectivity-at-age	Figure A.2
Stock-recruitment relationship	
Steepness, *h*	0.4
Extent of variation in recruitment, σ_R	0.6
Other	
Number of age-classes	20
Initial depletion	0.1
Data collection	
Survey CV	0.2
Effective sample size for the survey age data	100
Effective sample size for the fishery age data	100

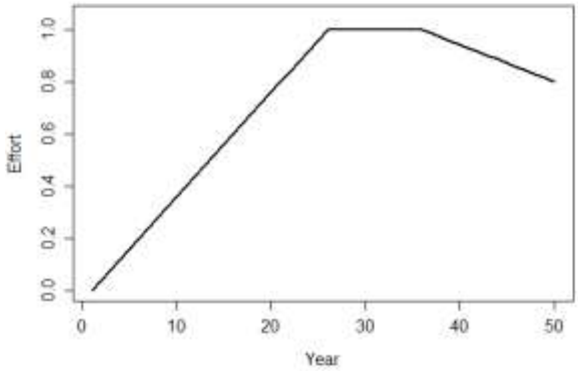

Figure A.1. Time-trajectory of effort.

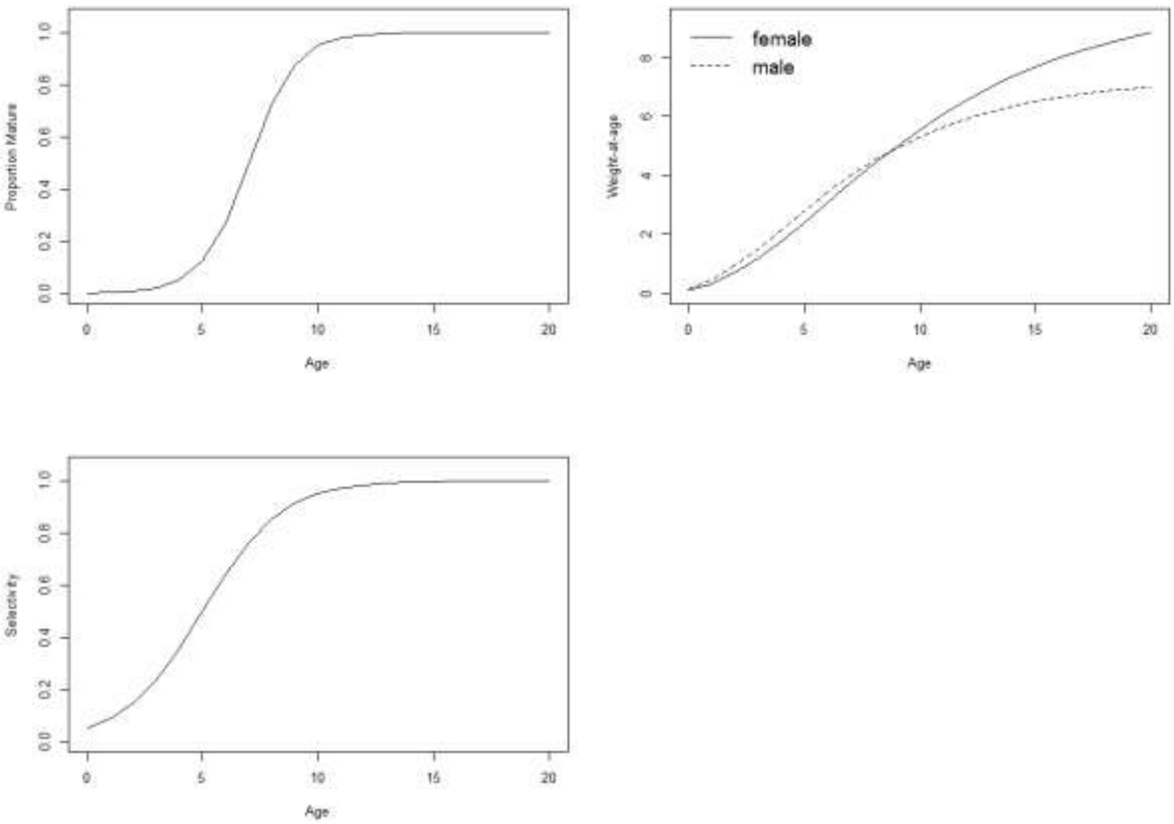

Figure A.2 Maturity, weight-at-age (mid-year), and selectivity as a function of age.