# A Management Strategy Evaluation

# for orange roughy

**ISL Client Report**
**for**
**Deepwater Group Ltd**

**P.L. Cordue**
**August 2014**

## Executive summary

A management strategy evaluation was performed with a generic orange roughy stock to determine an appropriate limit reference point, target biomass range, and harvest control rule (HCR) for use in managing orange roughy stocks. The proposed management strategy was designed to be consistent with New Zealand's Harvest Strategy Standard (MFish, 2008), and the Marine Stewardship Council's certification requirements (MSC, 2013).

The first step in the work was to estimate stock recruitment steepness (the percentage of virgin recruitment, on average, when at 20% of virgin biomass, $B_0$) by performing extra stock assessment runs for the Mid-East Coast stock. Of the four stocks assessed in 2014, this was the only stock which had adequate age data on cohorts spawned at low stock size (and hence information on how average recruitment changes at low stock size). Assessment runs were done for a Beverton-Holt and a Ricker stock-recruitment relationship. The results were similar for both relationships with median steepness (for the combined posterior) equal to 60% with a 95% CI of 31-95%.

The large level of uncertainty in steepness, as well as the form of the stock-recruitment relationship, created a high degree of uncertainty in the estimates of $B_{MSY}$. For Beverton-Holt, the median estimate and 95% CI of $B_{MSY}$ were 26% $B_0$ and 12–39% $B_0$; and for Ricker they were 42% $B_0$ and 37-47% $B_0$. As there was no basis for choosing between the Beverton-Holt and Ricker stock-recruitment relationships, it was concluded that the mid-point of the target range needed to be at about 40% $B_0$, as a trade-off between the $B_{MSY}$ estimates from the two relationships.

The limit reference point was defined to be the greater of 20% $B_0$ and 50% $B_{MSY}$. Under this definition, the Bayesian estimate of the limit reference point was 20% $B_0$ with a very high level of certainty.

Experimentation with various HCRs showed that (orange roughy) mature biomass, even when managed with a (perfectly) constant $F$, was prone to large *long-term* fluctuations (because of the low natural mortality and gradual recruitment). A fairly wide target range was needed to accommodate these *long-term* fluctuations and a breadth of 20% $B_0$ was proposed. Taken with the mid-point of 40% $B_0$, this gave a target biomass range of 30-50% $B_0$ (with a limit reference point of 20% $B_0$). With narrower biomass target ranges, the natural variability in the stock made it difficult to keep the stock within the range for a sufficiently high proportion of the time.

HCRs, based on the above range and limit reference point, were tested in long-term simulations (generally run over 16,000 years and with the first 1000 years ignored) to check that they performed adequately with regard to maintaining the biomass within the target range with little possibility of ever being below the limit reference point. The proposed rule was found to be robust to the uncertainty in steepness and natural mortality as well as one-off and multiple violations in major assumptions. The rule should perform well unless there is a severe violation of the assumptions made in the evaluation *and* steepness and/or natural mortality are low.

The proposed rule was also used in projections from the 2014 stock assessment results. The projections showed an increasing level of total allowable commercial catch (TACC) for each stock and a flat or increasing stock-status trajectory. For the North West Chatham Rise and ORH 7A stocks, the stock-status trajectory was within the target biomass range (30–50% $B_0$) throughout the five year projection period. The East and South Chatham Rise stock had an estimated stock status of just 30% $B_0$ in 2014. However, the projected stock-status, under the proposed HCR, was firmly within the target biomass range within 3–4 years (and under a "worst case" scenario, the median projected stock status reached 30% $B_0$ within 10–11 years).

The proposed rule should be re-evaluated about every five years as new data become available on stock-recruitment steepness and natural mortality in particular.

## Introduction

This document describes a management strategy evaluation (MSE) for orange roughy. The objective of the work was to determine a limit reference point, a biomass target range, and a harvest control rule (HCR) which are compatible with both the New Zealand's Harvest Strategy Standard (MFish, 2008) and the Marine Stewardship Council's certification requirements (MSC, 2013). In particular, the reference points and HCR developed in this document aim to be consistent with: PI 1.1.2 Reference Points, PI 1.2.1 Harvest Strategy, and PI 1.2.2 Harvest Control Rules and Tools.

A single generic orange roughy stock was modelled in the MSE. It was based on the three fisheries which are being evaluated against the MSC Standard in 2014: East and South Chatham Rise (ESCR), North West Chatham Rise (NWCR), and ORH7A (Challenger Plateau). The greatest uncertainties with regard to orange roughy population parameters are the stock-recruitment (SR) relationship and the value of natural mortality ($M$). The MSE focused on ensuring that the proposed harvest strategy is robust to these uncertainties. In terms of $B_{MSY}$, the SR relationship is critical, with both the form of the relationship (Beverton-Holt and Ricker were considered) and the level of steepness ($h$, being the proportion of virgin recruitment at 20% $B_0$) being important. The fourth orange roughy stock assessed in 2014, Mid-East Coast (MEC), which is not currently being evaluated against the MSC Standard because of its low stock status (< 20% $B_0$), is a crucial source of information on SR steepness.  New MEC assessment runs were performed to estimate $h$ for use in the MSE. Uncertainty in $M$ was quantified using the four existing assessments (MPI 2014).

A HCR with excellent long-term performance was determined by simulation using the generic orange roughy stock. The final step was checking that it also provided good short-term performance for the specific stocks under consideration. This was done by applying the HCR to projections from the 2014 stock assessment results.

## Methods

An MSE requires a number of components. There must be a population model which keeps track of the true state of the population and incorporates an appropriate level of "reality". In this MSE the model was an age-structured model that was very similar to those used in the 2014 stock assessments. However, the MSE model had some extra features which allow for the specification of some parameters that are not usually estimated during a stock assessment (e.g., correlation between annual year class strengths).

The other major component of an MSE is a method whereby the total allowable catch (TAC) is updated. In reality, for orange roughy, this will be by Bayesian stock assessment from time to time in conjunction with a HCR (and projections). It is not possible to model such assessments realistically as the calculation of estimates can take several days for a single assessment (therefore doing thousands of simulated Bayesian assessments is not possible in a reasonable timeframe). In this MSE, the stock assessment approach was approximated by using estimators based on the true values from the operating model. Two types of estimates were made: current stock status (current mid-season mature biomass divided by virgin mid-season mature biomass) and current vulnerable biomass (the beginning of year biomass available to the fishery). The two estimators are highly correlated to reflect that, in reality, they are products of the same stock assessment. Also, the estimators across years are highly correlated to reflect that, in reality, data sets used in successive assessments are cumulative (i.e., new data are added to an existing data set each year).

The population model and the method of updating TACs are part of what can be termed the "operating model". It represents "reality" at any time during a simulation. The testing of various HCRs requires that the properties of each HCR are determined by very long-term simulations. The question that needs addressing is, how well does a control rule perform on "average" and, for orange roughy in particular? This requires that thousands of years are simulated to accurately calculate the average performance.

The objective of the MSE is to find a HCR that maintains the mid-season mature biomass within a biomass range that is consistent with $B_{MSY}$ and allows little possibility of recruitment overfishing. The HCR must perform well over a wide range of assumptions; it should perform very well when the operating model is consistent with the assumptions under which the HCR was defined, but it must also be robust to errors in a wide range of assumptions (e.g., the form of the SR relationship, different values of natural mortality (*M*) and SR steepness).

The MSE was performed using purpose written code in the statistical package R.

**The operating model**

Full details of the operating model equations are given in Appendix A. A summary is given below.

The population model keeps track of fish in a single stock according to age (1-200 years with no plus group) and maturity (i.e., mature or immature). Therefore, the model is single-sex, single-area, and age-structured. The annual cycle was: ageing, recruitment (into age class 1), maturation, and then a full-year of mortality using the Baranov catch equation. The SR relationship was either Beverton-Holt or Ricker and average recruitment is calculated according to mid-season (i.e., after half the mortality) stock status. Year class strengths (YCS) were log-normally distributed with a specified recruitment variability (sigmaR = s.d. of log YCS) and correlation (rho = 1-year lag correlation of log YCS).

Maturation was constant from year-to-year and was logistic producing (i.e., the proportion mature at age in the virgin population was logistic). Fishing was either logistic by age (i.e., independent of maturity state) or only on mature fish (i.e., no fishing mortality on immature fish and full selection, independent of age, on mature fish).

The population was initialised in deterministic equilibrium and virgin mid-season mature biomass is denoted as $B_0$. The average unfished mid-season mature biomass was calculated from a long-term simulation with no fishing. If this was not equal to $B_0$ then a correction factor was applied (this was generally very small except when SR steepness was low and sigmaR and/or rho were high; see Cordue, 2001).

Two estimates were produced for each year of a simulation: stock status and vulnerable biomass. As already described, the estimators were highly correlated within year and across years. The estimate of stock status was available to be used in a HCR to calculate the $F$ to be applied to the estimate of current vulnerable biomass: $\text{TACC} = F\hat{B}_{vul}$. A HCR need not update the TACC each year, but the estimates are available to be used if needed. In the base model the control rule updates the TACC every third year. (Note, for orange roughy, the TAC is equal to the TACC plus 5% to allow for incidental catch).

**Ground-truthing of the operating model**

The results from four stock assessments in 2014 were used for specifying parameter values and/or ranges used in the operating model. The only missing piece from the assessments was guidance for SR steepness as this was specified to equal 0.75 in a Beverton-Holt SR relationship in each assessment. For the MSE, steepness was estimated for the MEC assessment as it had the most data from year classes that were spawned from low stock status (see Appendix B and a summary in the next section).

The four 2014 stock assessments all included MCMC runs where $M$ was estimated:

| **Stock** | ***M* (median)** | **95% CI** |
|---|---|---|
| NWCR | 0.041 | 0.033–0.051 |
| ESCR | 0.037 | 0.027–0.048 |
| MEC | 0.032 | 0.028–0.037 |
| ORH7A | 0.038 | 0.031–0.047 |
| Combined | 0.037 | 0.029–0.049 |

To represent the uncertainty in $M$, the four posterior distributions were combined (with equal weight) and a random sample of 5000 was taken from the combined distribution for use in the MSE.

Logistic-producing maturation had also been estimated in each of the assessments (posterior medians):

| **Stock** | **$a_{50}$ (years)** | **$a_{to95}$ (years)** |
|---|---|---|
| NWCR | 37 | 13 |
| ESCR | 41 | 12 |
| MEC | 35 | 10 |
| ORH7A | 32 | 10 |

For the operating model the median of the median estimates (above) was used: $a_{50}$ = 36 years, $a_{to95}$ = 11 years. These were fixed in the operating model in all simulations as it was similar across all four stocks and was reasonably well determined in each assessment. The more important issue was where the fishing selectivity was relative to maturity. For example, if fishing was on an older-age subset of the mature fish, then very high *F*s are sustainable because a part of the mature biomass would never be threatened; alternatively, if full selection was at an age well below maturity, then high *F*s would be very detrimental to stock status.

In the stock assessments, fishing was assumed to be on mature/spawning fish for ORH7A and NWCR. Also, for ESCR, the fishing selectivities (of the multiple fisheries) were centered at about the age of maturity. It is only for MEC that the fishing selectivity was at a much lower age than maturation. Since MEC is not being considered for MSC certification the fishing selectivity was based on the other three stocks. Within the operating model, fishing selectivity was assumed to be logistic with $s_{50} = a_{50}$ and $s_{to95} = a_{to95}$. This was the base assumption and, as an alternative, fishing was assumed to be just on the mature fish. Uncertainty in fishing selection relative to the age of maturity was not considered (except in one robustness run) as current fishing selectivity (relative to maturity) was well determined in the stock assessments (i.e., we know whether the fishery is currently on mature/spawning fish or not).

The East and Northwest Chatham Rise growth parameters and length-weight parameters from the 2014 Plenary report were used in the operating model (MPI, 2014).

Estimates of recruitment variability (sigmaR) and (1-year lag) correlation (rho) were available from the base model MCMC results:

| | SigmaR | | Rho | |
|---|---|---|---|---|
| **Stock** | **median** | **95% CI** | **median** | **95% CI** |
| NWCR | 0.94 | 0.80–1.07 | –0.04 | –0.29 to 0.28 |
| ESCR | 0.92 | 0.73–1.11 | –0.01 | –0.22 to 0.24 |
| MEC | 0.93 | 0.41–1.18 | 0.17 | –0.05 to 0.35 |
| ORH7A | 0.94 | 0.68–1.11 | 0.09 | –0.15 to 0.35 |

The consistency of the median sigmaR estimates is because the assessments contained very little information on sigmaR and the medians have moved little from the medians of the implied priors (from the Haist parameterisation and the nearly uniform prior on YCS: prior medians ≈ 0.96 and 95% CI ≈ 0.7–1.3). The major change from the priors is the elimination of the right-hand tail (i.e., values above approximately 1.1). In the base model, sigmaR = 0.9 was assumed and a robustness test was done for sigmaR = 1.1 (which is the usual value assumed for orange roughy but which appears to be a bit high given the stock assessment results). For the base model, rho = 0 was assumed and a robustness test was done at rho = 0.4.

The stock assessments produced fairly precise estimates of current mature biomass and stock status:

| Stock | CV($B_{2014}$)(%) | CV($B_{2014}/B_0$)(%) |
|---|---|---|
| NWCR | 17 | 11 |
| ESCR | 9 | 8 |
| MEC | 23 | 20 |
| ORH7A | 12 | 9 |

For the MSE base model, a CV of 15% was assumed for the vulnerable biomass estimator and the stock status estimator. A robustness test was done with CVs of 25%. The within-year correlation between the vulnerable biomass and stock status estimators was close to 1 as the same random errors were used in their construction. The 1-year lag correlation for each estimator was also close to 1 (see Appendix A).

**Estimation of steepness**

The MEC stock assessment model provided information on steepness because there were age frequency data for a number of cohorts that were spawned when stock status was low. The MEC base stock assessment model was re-run with $h$ and $M$ estimated. Informed priors were used for both parameters and the runs were taken through to full MCMC estimation (Appendix B). Estimation of steepness was made for both the Beverton-Holt and Ricker SR relationships. A random sample of 5000 was taken from the posterior distributions of each run for use in the MSE. For the MSE simulations, the base model assumed Beverton-Holt and in a robustness test the Ricker relationship was used in the operating model. Both forms were given equal weight when considering $B_{MSY}$ and the limit reference point (LRP).

**Estimation of $B_{MSY}$ and the LRP**

Bayesian estimation of $B_{MSY}$ and the LRP was performed to account for uncertainty in $h$ and $M$. This was achieved by calculating $B_{MSY}$ and the LRP as a function of $h$ and $M$ over a two-dimensional grid of values and then obtaining a posterior distribution by using the given posterior samples of $h$ and $M$ (see above). For each pair of posterior samples ($h$, $M$) the value of $B_{MSY}$ or the LRP was calculated by interpolation using the corresponding "grid function". The "spline" and "splinefun" functions in R were used to provide the interpolated values (these are cubic splines). Hence, the 5000 samples from the joint posterior of $h$ and $M$ provided 5000 samples from the posteriors of $B_{MSY}$ and the LRP.

For given values of $h$ and $M$, $B_{MSY}$ was calculated by running the base model (or the Ricker model) with deterministic recruitment and constant $F$ over a range of $F$ values to determine the yield curve. The model was run for 3000 years at each value of $F$ (to ensure deterministic equilibrium was reached) and $F_{MSY}$ was determined to two significant figures. The LRP was defined to be the greater of 20% $B_0$ or 50% $B_{MSY}$.

## Estimation of performance indicators for harvest control rules

Four performance indicators (for a given HCR) were estimated in each run: mean annual mid-season mature biomass; mean annual yield; the probability of the mid-season mature biomass being above the LRP (denoted $LRP_p$); and the probability of the mid-season mature biomass being above the lower bound of the biomass target range (denoted $LB_p$).

Bayesian estimation was used to account for the uncertainty in $h$ and $M$. This was achieved in the same fashion as for $B_{MSY}$ and the LRP, using interpolation via cubic splines over the grid of calculated values for fixed $h$ and $M$. For each fixed pair ($h$, $M$) the HCR was applied for 16,000 years. The first 1000 years were ignored and the statistics were calculated from the remaining 15,000 years. As a check on stochastic equilibrium, the median biomass was calculated for each 5000 year segment after the initial 1000 years were discarded. The CV of the three medians was required to be less than 5% otherwise a warning was issued. Warnings were rare except for the run where sigmaR = 1.1 and rho = 0.4. For this run, the 15,000 year period was doubled to 30,000 years (which eliminated warnings except for one or two ($h$, $M$) pairs with very low values).

The Bayesian posteriors of $LRP_p$ and $LB_p$ were used to derive two summary measures:

- **LRP risk**: the probability that the HCR will allow mid-season mature biomass to be below the LRP more than 5% of the time
- **depletion risk**: the probability that the HCR will allow mid-season mature biomass to be below the lower bound of the biomass target range, LB, more than 30% of the time

The probabilities are the proportion of $h$, $M$ pairs where the HCR allows the poor performance with respect to the LRP or LB. The choices of 5% and 30% were not arbitrary.

The mid-season mature biomass should rarely go below the LRP in the long-term; the choice is perhaps between 1% and 5%. The rarer the event the harder it is to estimate the actual probability so 5% was chosen rather than 1% (if 1% had been chosen, the simulation period of 15,000 years would probably had to have been increased).

The 30% level for the LB was chosen to reflect the MSC definition of a depleted stock which is one which is “consistently” below the LB of the target biomass range. Certainly if a stock was below the LB only 10-20% of the time it could not reasonably be argued that it was “consistently” below. At 30% the argument is somewhat moot, but this level was chosen to be conservative.

## Robustness testing of control rules

The main robustness testing focused on the proposed control rule but the approach was the same for any rule. The main focus of the testing was robustness to uncertainty in $h$ and $M$ but in addition to this, various assumption violations were laid on top of the uncertainty testing. For example, the proposed HCR was tested over the whole plausible range of $h$ and $M$ with a

Beverton-Holt SR relationship and alternatively with a Ricker SR relationship. Also, the base model assumed unbiased estimators of stock status and vulnerable biomass. As this is unlikely to be the case, the robustness of the proposed HCR was tested against a 20% positive bias in one or other or both of the estimators. A higher assumed CV for the stock assessment estimators was also tested; as were higher values of sigmaR and rho.

## Results

### Bayesian estimates of steepness and natural mortality

The MEC assessment runs in which $h$ was estimated gave similar results for Beverton-Holt and Ricker stock-recruitment relationships (Table 1, Appendix B). The combined results (giving both runs equal weight) gave a median steepness of 60% (i.e., an average of 60% of virgin recruitment when at 20% $B_0$) with a 95% CI of about 30-90% (Table 1). The combined posterior distribution had a single mode with a very broad range (Figure 1 – note, values above 100% are from the Ricker posterior). For both runs the estimated median correlation between $h$ and $M$ was close to zero.

**Table 1: Bayesian estimates of steepness for the MEC assessment models that assumed a Beverton-Holt or a Ricker stock-recruitment relationship. The median and 95% CIs are given as a percentage of virgin recruitment ($R_0$).**

| | Steepness ($h$) (% $R_0$) | |
|---|---|---|
| | **Median** | **95% CI** |
| **Beverton-Holt** | 68 | 39–93 |
| **Ricker** | 53 | 28–99 |
| **Combined (equal weight)** | 60 | 31–95 |

The posterior distribution for $M$ was calculated as the (equal-weight) combined posterior distribution of the four assessed stocks in 2014 (see "Ground truthing" above). The distribution had a median of 0.037 with a 95% CI: 0.029–0.049 (Figure 2).

### Bayesian estimates of $B_{MSY}$

For the base model, estimates of $B_{MSY}$ were highly dependent on the form of the SR relationship and the level of $h$ (Tables 2 & 3). For both SR relationships, increasing $M$ reduces $B_{MSY}$ slightly when $h$ is high. Also, in both cases, increasing $h$ reduced $B_{MSY}$, but the reduction was far less for the Ricker relationship than the Beverton-Holt relationship (Tables 2 & 3).

**Table 2: $B_{MSY}$ (%$B_0$) as a function of $h$ and $M$ when a Beverton-Holt stock recruitment relationship was assumed for the base model. "–" denotes that $B_{MSY}$ was not defined (i.e., the yield curve did not have a maximum).**

| | Natural mortality (*M*) | | | | | | |
|---|---|---|---|---|---|---|---|
| **Steepness (*h*)** | **0.02** | **0.025** | **0.03** | **0.035** | **0.045** | **0.05** | **0.06** |
| **0.25** | 45 | 45 | 45 | 45 | 45 | 45 | 45 |
| **0.30** | 45 | 44 | 44 | 44 | 43 | 43 | 43 |
| **0.35** | 42 | 42 | 41 | 41 | 41 | 41 | 40 |
| **0.40** | 40 | 40 | 39 | 39 | 39 | 38 | 38 |
| **0.50** | 35 | 34 | 34 | 34 | 34 | 34 | 33 |
| **0.60** | 31 | 30 | 30 | 30 | 29 | 29 | 29 |
| **0.75** | 25 | 24 | 24 | 23 | 23 | 22 | 22 |
| **0.90** | 17 | 17 | 16 | 16 | 15 | 14 | 13 |
| **1.00** | 6 | 4 | 3 | – | – | – | – |

**Table 3: $B_{MSY}$ (%$B_0$) as a function of $h$ and $M$ when a Ricker stock-recruitment relationship was assumed for the base model.**

| | Natural mortality (*M*) | | | | | | |
|---|---|---|---|---|---|---|---|
| **Steepness (*h*)** | **0.02** | **0.025** | **0.03** | **0.035** | **0.045** | **0.05** | **0.06** |
| **0.25** | 48 | 48 | 48 | 48 | 48 | 48 | 48 |
| **0.30** | 48 | 47 | 47 | 47 | 47 | 47 | 47 |
| **0.35** | 46 | 46 | 46 | 46 | 46 | 46 | 46 |
| **0.40** | 45 | 45 | 45 | 45 | 44 | 44 | 44 |
| **0.50** | 43 | 43 | 43 | 42 | 42 | 42 | 42 |
| **0.60** | 42 | 42 | 41 | 41 | 41 | 41 | 40 |
| **0.75** | 40 | 40 | 39 | 39 | 39 | 39 | 38 |
| **0.90** | 38 | 38 | 38 | 38 | 37 | 37 | 36 |
| **1.00** | 37 | 37 | 37 | 37 | 36 | 36 | 35 |
| **1.20** | 36 | 36 | 36 | 35 | 34 | 34 | 34 |

The Bayesian estimates of $B_{MSY}$ were obtained from the implicit functions in Tables 2 and 3 weighted by the $h$ and $M$ posterior distributions. The steepness estimates for Beverton-Holt were much lower than from Ricker with the 95% CIs barely overlapping (Table 4). The combined posterior covers a very broad range with the median just below the commonly used $B_{MSY}$ proxy of 40% $B_0$ (Table 4).

**Table 4: Bayesian estimates of $B_{MSY}$ for the base model assuming a Beverton-Holt or a Ricker stock-recruitment relationship. The median and 95% CIs are given as a percentage of virgin mid-season mature biomass ($B_0$).**

| | $B_{MSY}$ (%$B_0$) | |
|---|---|---|
| | **Median** | **95% CI** |
| **Beverton-Holt** | 26 | 12–39 |
| **Ricker** | 42 | 37–47 |
| **Combined (equal weight)** | 38 | 15–47 |

## Bayesian estimates of the limit reference point

The limit reference point was defined to be the greater of 20% $B_0$ or 50% $B_{MSY}$. The rationale for this definition is that the stock should not be too far below $B_{MSY}$ and ideally never be below 20% $B_0$. Allowing the stock to be below 20% $B_0$ could impair recruitment, alter the role of the fish in the ecosystem and possibly result in a regime shift to a much lower carrying capacity (i.e., niche replacement).

The estimates of the limit reference point (LRP) for the two SR relationships were very similar and the median of the combined posterior distribution was 20% $B_0$ (Table 5).

**Table 5: Bayesian estimates of the LRP for the base model assuming a Beverton-Holt or a Ricker stock-recruitment relationship. The median and 95% CIs are given as a percentage of virgin mid-season mature biomass ($B_0$).**

| | LRP (% $B_0$) | |
|---|---|---|
| | **Median** | **95% CI** |
| **Beverton-Holt** | 20 | 20–20 |
| **Ricker** | 21 | 20–24 |
| **Combined (equal weight)** | 20 | 20–23 |

The current knowledge of steepness for orange roughy suggests that 20% $B_0$ is appropriate as a LRP irrespective of whether a Beverton-Holt or Ricker relationship is assumed. When the two relationships are given equal weight, the median estimate of the percentage of virgin recruitment at 20% $B_0$ is 60% (i.e., the median steepness as given in Table 1). This seems appropriate for a LRP; less than 60% (on average) of virgin recruitment does constitute some impairment of recruitment that should be avoided to ensure long-term sustainability of the stock and dependent fishery.

## Determination of a suitable target range

A suitable biomass target range has to ensure that biomass will be maintained well above the LRP for the majority of the time and that it is consistent with $B_{MSY}$.

The existing target range of 30–40% $B_0$ would be suitable in the first regard if biomass could be maintained within the range (or at least generally above 30% $B_0$). However, if the mid-point of the range (35% $B_0$) is targeted, it is unlikely that biomass can generally be maintained above 30% $B_0$. For example, if the base-model stock is fished at a constant fishing mortality of $F_{35\%B0}$, then 95% of the time the stock status is in the range 24–48% $B_0$ (Table 6, Figure 3). If a "bent" control rule is applied where $F$ reduces linearly from $F_{35\%B0}$ at 30% $B_0$ down to 0 at 20% $B_0$ the "95%-range" is still too broad (27–48% $B_0$) with a 19% probability of being below 30% $B_0$ (Table 6). This all assumes that stock status and current vulnerable biomass are known exactly every year; if more realistic assumptions are made (with the same HCRs) then stock status will be more variable (and therefore will more often be at less than 30% $B_0$).

In terms of $B_{MSY}$, the mid-point of the target range at 35% $B_0$ seems low, as the median estimate of $B_{MSY}$ is 38% $B_0$ and the 95% CI on the Ricker $B_{MSY}$ is 37–47% $B_0$ (Table 4). As a compromise between potentially very low $B_{MSY}$ from Beverton-Holt (95% CI: 12–39 % $B_0$) and the higher Ricker range, it is appropriate to set a mid-point for the biomass target range at about the median of the combined posterior distribution (38% $B_0$). Since the commonly used $B_{MSY}$ proxy of 40% $B_0$ is slightly above the median estimate it is convenient to use 40% as the mid-point of the target range.

**Table 6: Stock status statistics for the base model when fishing at constant $F_{35\%B0}$ or with a harvest control rule that declines linearly from $F_{35\%B0}$ at 30% $B_0$ down to 0 at 20% $B_0$. These results assume no error in stock status or fishing mortality.**

| | Stock status (% $B_0$) | | | |
|---|---|---|---|---|
| | **Median** | **95% CI** | **P(B < 20% $B_0$)** | **P(B < 30% $B_0$)** |
| $F_{35\%B0}$ | 34 | 24–48 | 0.00 | 0.24 |
| $F_{35\%B0}$ **+ bend** | 34 | 27–48 | 0.00 | 0.19 |

The target range should be broad enough to accommodate the sustained trends in stock status that can occur due to good or poor recruitment (e.g., see Figure 3). It is clear from the base-model results, when fishing at a constant $F$ (or using the bent rule), that allowing only 10% $B_0$ for the breadth of the range is inadequate (e.g., the 95% CI for the bent rule spans 21% $B_0$). A breadth of 20% $B_0$ is the obvious candidate. Combined with the mid-point of 40% $B_0$ this gives a target range of 30–50% $B_0$.

The lower bound of the target range at 30% $B_0$ is well above the LRP (20% $B_0$) and the target range, in terms of the median estimates, is expected to provide 75–90% of virgin recruitment (Table 7).

**Table 7: Bayesian estimates of average recruitment at 30% $B_0$ and 50% $B_0$ for the base model assuming a Beverton-Holt or a Ricker stock-recruitment relationship. The median and 95% CIs are given as a percentage of virgin recruitment ($R_0$).**

| | Average recruitment at 30% $B_0$ (%$R_0$) | | Average recruitment at 50% $B_0$ (%$R_0$) | |
|---|---|---|---|---|
| | **Median** | **95% CI** | **Median** | **95% CI** |
| **Beverton-Holt** | 78 | 52–96 | 89 | 72–98 |
| **Ricker** | 69 | 41–122 | 91 | 62–136 |
| **Combined (equal weight)** | 75 | 44–110 | 90 | 66–127 |

## The proposed harvest control rule (HCR)

The proposed HCR is based on the existing HCR which has $F$ ramping up from 0 at 10% $B_0$ to 0.045 at 30% $B_0$ and remaining constant thereafter ($M$ = 0.045 is the assumed value for stock assessments, so the existing HCR is an "$F = M$" strategy) . However, the proposed HCR is "dynamic" and additionally has $F$ ramping up within the target biomass range (Figure 4). The dynamic aspect of the HCR (which is explained below) is crucial in enabling long-term performance which is robust to uncertainty in parameter estimates (e.g., $h$ and $M$, recruitment variability and correlation) and errors in assumptions (e.g., the form of the SR relationship, bias in the estimators of stock status and/or current biomass).

A "static" HCR primarily consists of a functional relationship between estimated stock status and $F$. That is, there is a function $g$, which for a given estimated stock status $s$, will return $F = g(s)$ which is the fishing mortality to be applied to the estimate of vulnerable biomass to obtain the recommended catch. In a static HCR, the functional relationship $g$ does not change over time.

In a dynamic HCR, there is an initial functional relationship and a rule by which that relationship can change over time. For the proposed rule, the initial functional relationship is essentially the existing HCR except that it has $F$ ramping up within the target biomass range of 30–50% $B_0$ (Figure 4). There are two ways in which $g$ can change. If estimated stock status is below the lower bound of the target biomass range (LB) then $g$ is scaled down by a proportion $p < 1$ (i.e., the new relationship is such that the $F$s are equal to the old $F$s multiplied by $p$). The scaling down of the relationship $g$ occurs every time that an assessment estimates stock status below the LB (although the proportion $p$ also changes as a function of stock status, and the cumulative scaling down cannot exceed a specified limit – see Appendix 1). If estimated stock status is found to be greater than 60% $B_0$ and $g$ is less than the initial $g$ then it is rescaled upwards (i.e., the new $F$s are equal to the old $F$s divided by a proportion $p$ $< 1$). The $p$ values are determined as a function of stock status but they are always between 0.9 and 1 (see Appendix 1, the recommended HCR has $l$ = LB = 30% $B_0$, $r$ = 60% $B_0$, $k = 0.9$, $m = 10$, and $p_{limit} = 0.3$).

The rescaling of the functional relationship essentially allows the HCR to "learn" over time about the average production of the stock. If estimated stock status is below the LB, then the HCR becomes progressively more conservative. Similarly, if the stock status becomes very large (after scaling down of the functional relationship) then the HCR becomes progressively more aggressive. If the HCR is properly designed, stock status should remain with the target biomass range most of the time (when measured over the long term).

The full specification of the proposed HCR, which we will denote as "dynamic HCR10", is: LRP = 20% $B_0$, target biomass range = 30–50% $B_0$, initial $F_{mid} = 0.045$, slope within the target range: $p$ = 25%; ramps down to zero at 10% $B_0$; rescaling limit points: $l$ = 30% $B_0$, $r$ = 60% $B_0$; $k = 0.9$, $m = 10$, $p_{limit} = 0.3$. Assessments were specified to occur every 3 years in line with MPI's draft 10-year deepwater plan for orange roughy.

**Performance of the harvest control rule**

Four performance indicators were evaluated: mean annual mid-season mature biomass; mean annual yield; the probability of the mid-season mature biomass being above the LRP (20% $B_0$); and the probability of the mid-season mature biomass being above the lower bound of the target range (30% $B_0$). Also, two risks were estimated from the Bayesian posteriors: LRP risk and depletion risk (see above).

The average mid-season mature biomass maintained by dynamic HCR10 (for the base model) depends strongly on $h$ and $M$ with higher values of these parameters giving higher biomass (Figure 5). The same is true for average yield with higher yield obtained for higher values of $h$ and $M$ (Figure 6). There is little chance of the biomass falling below the LRP except when $h$ and $M$ are very low (Figure 7). The probability of the rule maintaining biomass above the LB also depends strongly on $h$ and $M$ (Figure 8). For very low values of $h$ and $M$ there is little chance of being above the LB.

When we look at how the functional relationship (between stock status and $F$) has been cumulatively rescaled over the full simulated time period (16,000 years) we see that the HCR "learnt" much about the mean production of the stock (Figure 9). For the lowest values of $h$ and $M$ the functional relationship was scaled down as much as possible to $p_{limit} = 0.3$ (Figure 9). For the grid of $h$ and $M$ values there is progressively less rescaling as $h$ and $M$ increase. For the three highest values of $M$, the cumulative rescaling value is somewhat "bumpy" as $h$ increases (Figure 9). This is because it is a matter of luck as to how much the relationship is rescaled after estimated stock status dips below 30% $B_0$ (because of a sequence of poor YCS) and then ultimately increases above 60% $B_0$ (because the rescaled functional relationship removes less than the average stock production).

Application of the $h$ and $M$ posteriors to the grid functions provides Bayesian posteriors for each of the four performance indicators. The spread in the posterior for each indicator is caused by the variation in $h$ and $M$ across the respective distributions. The mean biomass under dynamic HCR10 has a mode at about 42% $B_0$ and very little weight anywhere else; the mean yield has a mode at about 1.5% $B_0$; the probability of being above the LRP is very tightly distributed at or just below 1; as is the probability of being above the LB (Figure 10). That is to say, there are very few combinations of $h$ and $M$ for which the rule does not have excellent long-term performance (see Table 8). For the base model, dynamic HCR10 has zero LRP and depletion risk (see Table 12).

**Table 8: Base model: Bayesian estimates of the four performance indicators for the proposed harvest control rule: dynamic HCR10.**

| Mean B(% $B_0$) | | Mean yield (% $B_0$) | | P(B > 20% ($B_0$)(%) | | P(B > 30% $B_0$)(%) | |
|---|---|---|---|---|---|---|---|
| **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** |
| 42 | 41–43 | 1.4 | 0.8–2.1 | 100 | 100–100 | 97 | 96–98 |

A careful interpretation of the median estimate of mean yield is needed. The median mean yield is the "best" estimate available given the current knowledge with regard to $h$ and $M$. However, in practice, the mean long-term yield will depend strongly on the actual values of $h$ and $M$; depending on what they are, mean yield could be less than 1% $B_0$ or as high as 2% $B_0$ (see the 95% CIs in Table 8).

Each HCR contains a specification for the frequency of TAC updates. The base assumption is for updates every 3 years. For orange roughy, low natural mortality implies that a large number of cohorts are in the mature biomass when it is maintained in a range of 30-50% $B_0$. The mature biomass changes little from year to year and the frequency of TAC updates is not important in terms of the theoretical performance of the control rules (see Table 9). However, regular monitoring of biomass and age composition and continued research on generic issues (e.g., acoustic target strength) are important to improve the stock assessments.

**Table 9: Base model: Bayesian estimates of the four performance indicators for the proposed HCR (dynamic HCR10, update every 3 years, results in bold) and the same HCR with more frequent or less frequent TAC updates.**

| | **Mean B (%$B_0$)** | | **Mean yield (%$B_0$)** | | **P(B > 20% $B_0$)(%)** | | **P(B > 30% $B_0$)(%)** | |
|---|---|---|---|---|---|---|---|---|
| **Years** | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** |
| ***n* = 1** | 42 | 41–43 | 1.4 | 0.8–2.1 | 100 | 100–100 | 96 | 94–98 |
| ***n* = 3** | **42** | **41–43** | **1.4** | **0.8–2.1** | **100** | **100–100** | **97** | **96–98** |
| ***n* = 5** | 42 | 41–43 | 1.4 | 0.8–2.1 | 100 | 100–100 | 97 | 97–98 |

## Robustness of the proposed HCR

Many assumptions are made in the base model and the robustness of the proposed HCR to most of these assumptions was tested. If the Ricker SR relationship and the associated posterior distribution are used then the mean biomass and yield have similar median values as for the base model but they have wider 95% CIs (Table 10). There is still very low LRP and depletion risk with the estimated values respectively 2% and 3%. This was the only sensitivity where LRP risk was greater than 0 and depletion risk was greater than 1% (see Table 12).

Additional recruitment variability (sigmaR = 1.1 instead of 0.9) does not compromise the HCR, nor does the presence of moderate correlation in recruitment strengths (rho = 0.4 instead of 0; see Table 10). Even when there is increased variability and moderate correlation, the HCR still has excellent long-term performance as it does when there is a major error in the selectivity (6 years younger than assumed in the *F* calculations) (Table 10).

**Table 10: Bayesian estimates of the four performance indicators for dynamic HCR10 with the base model (results in bold), and variations of the base model: Ricker SR relationship; fishing only on mature fish (Mature); sigmaR = 1.1 (instead of 0.9); rho = 0.4 (instead of 0); sigmaR = 1.1 & rho =0.4 (1.1 & 0.4); $s_{50}$ = 30 years (instead of 36 years).**

| | **Mean B (%$B_0$)** | | **Mean yield (%$B_0$)** | | **P(B > 20% $B_0$)(%)** | | **P(B > 30% $B_0$)(%)** | |
|---|---|---|---|---|---|---|---|---|
| | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** |
| **Base** | **42** | **41–43** | **1.4** | **0.8–2.1** | **100** | **100–100** | **97** | **96–98** |
| **Ricker** | 42 | 32–50 | 1.4 | 0.4–2.9 | 100 | 97–100 | 96 | 61–100 |
| **Mature** | 42 | 41–43 | 1.4 | 0.8–2.0 | 100 | 100–100 | 97 | 95–98 |
| **Sig. 1.1** | 42 | 40–43 | 1.4 | 0.7–2.1 | 100 | 100–100 | 95 | 93–97 |
| **Rho 0.4** | 43 | 41–45 | 1.4 | 0.7–2.1 | 100 | 100–100 | 93 | 89–96 |
| **1.1 & 0.4** | 43 | 39–45 | 1.4 | 0.7–2.1 | 100 | 99–100 | 89 | 80–92 |
| **$s_{50}$ = 30** | 42 | 40–42 | 1.4 | 0.7–2.2 | 100 | 100–100 | 97 | 96–97 |

The performance of the HCR is only slightly affected by a 20% positive bias in the estimation of stock status or current vulnerable biomass (Table 11). When both biases are present there are some lower values of *h* and *M* that could lead to the biomass spending a significant amount of time below the LB (Table 11). Considerable care is taken during the stock assessment process to ensure that such systematic biases are eliminated (e.g., in the way acoustic biomasses are developed from survey data and treated as relative indices). However, there is no LRP risk associated with both biases and only a 1% depletion risk (Table 12). The HCR is robust to a higher level of imprecision in the estimation of stock status and vulnerable biomass (CV = 25% instead of 15%)(see Table 11).

**Table 11: Bayesian estimates of the four performance indicators for dynamic HCR10 with the base model (results in bold) and base model variations: a 20% bias in estimates of vulnerable biomass (Vul. 20%) or stock status (S.S. 20%) or both (Both 20%); and a CV for the estimates of 25% (instead of 15%).**

| | Mean B (% $B_0$) | | Mean yield (% $B_0$) | | P(B > 20% $B_0$)(%) | | P(B > 30% $B_0$)(%) | |
|---|---|---|---|---|---|---|---|---|
| | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** | **Median** | **95% CI** |
| **Base** | **42** | **41–43** | **1.4** | **0.8–2.1** | **100** | **100–100** | **97** | **96–98** |
| **S.S. 20%** | 35 | 35–41 | 1.5 | 0.8–2.2 | 100 | 100–100 | 80 | 77–95 |
| **Vul. 20%** | 42 | 41–42 | 1.4 | 0.8–2.2 | 100 | 100–100 | 97 | 96–97 |
| **Both 20%** | 35 | 34–37 | 1.5 | 0.8–2.3 | 100 | 100–100 | 78 | 76–87 |
| **CV 25%** | 42 | 41–44 | 1.4 | 0.7–2.1 | 100 | 99–100 | 92 | 92–95 |

**Table 12: Estimated LRP risk and depletion risk for dynamic HCR10 with the base model and all sensitivities tested.**

| | LRP risk (%) | Depletion risk (%) |
|---|---|---|
| **Base** | 0 | 0 |
| **Ricker** | 2 | 3 |
| **Mature** | 0 | 0 |
| **Sig. 1.1** | 0 | 0 |
| **Rho 0.4** | 0 | 0 |
| **1.1 & 0.4** | 0 | 1 |
| **$s_{50}$ = 30** | 0 | 0 |
| **S.S. 20%** | 0 | 1 |
| **Vul. 20%** | 0 | 0 |
| **Both 20%** | 0 | 1 |
| **CV 25%** | 0 | 0 |

## Application of the proposed HCR to the 2014 stock assessments

The proposed HCR (dynamic HCR10) is designed to maintain biomass, over the long term, within a target biomass range of 30–50% $B_0$ for a wide range of values of $h$ and $M$. The 2014 stock assessment calculation of yields used $U_{35\%B0}$ based on $h$ = 0.75 and $M$ = 0.045 (Cordue, 2014). A constant $U_{35\%B0}$ HCR is a more aggressive rule than dynamic HCR10, as it targets a mid-point of 35% $B_0$ rather than 40% $B_0$. Also, the proposed HCR has a slope within the target biomass range which reduces $F$ from its midpoint value when estimated stock status is less than 40% $B_0$ (this affects ESCR as its estimated stock status was 30% $B_0$; for NWCR and ORH7A there is little effect from the slope as they were estimated at 37% and 42% $B_0$ respectively).

The yield estimates from dynamic HCR10 are less than the $U_{35\%B0}$ yields but similar to the highest options given by MPI in the discussion documents (IPPs) for the 2014-15 TAC review (Table 13).

**Table 13: Yield estimates for the 2014-15 fishing year from the stock assessments, the MSE and MPI consulted options for catch limits.**

| Stock | $U_{35\%B0}$ yield (t) | HCR10 yield (t) | MPI TACC/catch-limit options (t) |
|---|---|---|---|
| NWCR | 1414 | 1043 | 750, 900, 1250 |
| ORH7A | 2128 | 1748 | 500, 900, 1600 |
| ESCR | 6400 | 3772 | 3100 |

Dynamic HCR10 was used in projections from the 2014 stock assessment results for NWCR, ESCR, and ORH7A (sampling from the last 10 estimated YCS as in the stock assessment – see Cordue, 2014). That is, the HCR was applied in 2014 (using the base model and the medians of the posterior distributions for stock status and vulnerable biomass) to set the TACC/catch-limit in 2015 and subsequent years, until the next scheduled assessment; then the HCR was again applied (to the medians of the projected distributions for stock status and vulnerable biomass) to set the TACC/catch-limit for the next set of years; and so on. The projections were first done for the base models (which determined the TACCs/catch-limits under the base models) and then repeated for the lowM-highq models (using the TACCs/catch-limits from the base models). Projections were done up to 2019 for NWCR and ORH7A but were extended to 2025 for ESCR (to check on whether the stock rebuilt to 30% $B_0$ for the "worst-case" lowM-highq model). Also, for ESCR, projections for the base model and lowM-highq were repeated with $h = 0.6$ (which is 20% lower than the assumed value of 0.75 in the stock assessments).

The years in which to do future assessments were taken from the draft 10-Year Deepwater Fisheries Research and Monitoring Programme which has the assessments for ESCR, NWCR, and ORH7A scheduled respectively for 2018, 2018, and 2016; and then every 3 years for each stock. The projected TACCs/catch-limits are shown in Table 14.

**Table 14: Projected catch limits for the three orange roughy fisheries from the MSE.**

| | ESCR catch limit (t) | NWCR catch limit (t) | ORH7A (TACC) |
|---|---|---|---|
| 2015 | 3772 | 1043 | 1748 |
| 2016 | 3772 | 1043 | 1748 |
| 2017 | 3772 | 1043 | 1799 |
| 2018 | 3772 | 1043 | 1799 |
| 2019 | 4965 | 1332 | 1799 |
| 2020 | 4965 | | |
| 2021 | 4965 | | |
| 2022 | 5768 | | |
| 2023 | 5768 | | |
| 2024 | 5768 | | |
| 2025 | 6317 | | |

For the NWCR, the HCR allowed the median stock status to increase each year for the base model (Figure 11) and for the "worst-case" lowM-highq model (Figure 12). The main difference between the two models was that the base started from a higher stock status than the lowM-highq model. Both models finish in 2019 with the median stock status within the target biomass range (Figures 11 and 12).

The results are similar for ORH7A in that the shape of the trajectory is flat showing a slight decline in the later years for both the base model and the lowM-highq model (Figures 13 and 14). Also, in 2019 both models finish with the median stock status within the target biomass range.

For the ESCR base model, which starts with a stock status of 30% $B_0$ in 2014, a steady increase in stock status is shown with the 95% CIs wholly within the target biomass range from 2020 onwards (Figure 15). For the lowM-highq model, a steady rebuild is also shown but the median stock status does not get into the target biomass range until 2023 (Figure 16). The results are similar for the base and lowM-highq models when $h$ = 0.75 is replaced with $h$ = 0.6 when doing the projections (Figures 17 and 18). The results are only slightly more pessimistic since the spawning biomass, during much of the projection period, is receiving recruitment spawned from biomass that was not substantially depleted (i.e., which means that the average recruitment is very similar for $h$ = 0.6 and $h$ = 0.75).

## Discussion and conclusion

The current target biomass range for orange roughy is 30–40% $B_0$. The lower bound of this target range is from early work following the recommendations of Francis (1992) to use a default steepness of 0.75 and calculate $B_{MAY}$ (biomass that gives the Maximum Average Yield under a constant $F$) subject to the constraint that biomass should not be below 20% $B_0$ more than 10% of the time. His methods, when applied to orange roughy stocks, gave $B_{MAY}$ = 30% $B_0$ for all of the stocks that had been assessed (MPI 2014). This value is lower than that calculated, using the method of Francis (1992), for many other species because the low natural mortality of orange roughy ensures that there are many cohorts in the mature biomass (when it is not at a low level) and hence the biomass is more stable than for other species with higher natural mortality.

The current target biomass range of 30–40% $B_0$ appears to be too narrow to meet the criteria set in the Marine Stewardship Council Standard. It does not appear to be consistent with $B_{MSY}$ as the current best estimate has a 95% CI of 15–47% $B_0$ with a median of 38% $B_0$. If only a Beverton-Holt SR relationship is considered then 30–40% $B_0$ would be consistent with $B_{MSY}$ which is then estimated with a median of 26% $B_0$ and a 95% CI of 12–39% $B_0$. However, in that case there is still the problem that biomass will often be below 30% $B_0$ unless a higher mid-point is set within the target range (i.e., if biomass fluctuates around the mid-point of 35% $B_0$ it will often be below 30% $B_0$).

A wider biomass target range is needed to be consistent with $B_{MSY}$ and to ensure that biomass will almost always be above the lower bound of the target range. The proposed range of 30–50% $B_0$ allows for the relatively large *long-term* fluctuations which occur naturally in orange roughy stocks. The target biomass range recommended is consistent with the commonly used proxy for $B_{MSY}$ of 40% $B_0$. It is also consistent with the finding from Punt et al. (2014) that a target of 35–40% $B_0$ will minimize the potential loss of yield relative to that which would be achieved when $B_{MSY}$ is known (Punt et al. considered Beverton-Holt and Ricker SR relationships).

The proposed HCR (dynamic HCR10) will maintain the biomass within the target biomass range, in the long-term, under most circumstances. Because the rule is dynamic, and adjusts the functional relationship between stock status and $F$ over time, the rule is robust to the large uncertainties in SR steepness and natural mortality. Further, it is robust to one-off and some multiple violations of the assumptions made in the study. It will take a severe violation of one or more assumptions before the rule will perform badly (and only then if steepness and/or natural mortality are low).

In the short term, dynamic HCR10 was shown to be very safe for the three stocks being considered. In the base models, stock status was maintained (NWCR, ORH7A) or rebuilt (ESCR) into the target biomass range. In the "worst case" scenarios (lowM-highq), the TACCs under dynamic HCR10 from the base model created similar stock-status trajectories to those from the base model. That is, if dynamic HCR10 is applied under the base assumptions, but it happens that the stock status and productivity are lower than expected, the resultant stock status is still more than adequate.

For the ESCR, dynamic HCR10 allows the stock to rebuild to be firmly within the target biomass range within 3–4 years. Under the lowM-highq model the rebuild takes longer but median projected stock status is within the target biomass range within 10–11 years. For both the base and lowM-highq models the increasing stock-status trajectory and rebuild times are not unduly affected by an $h$ which is 20% lower than assumed in the base model.

The dynamic nature (i.e., rescaling of the functional relationship) of the recommended HCR is not relevant in the short-term for two of the three stocks under consideration because NWCR and ORH7A are at about 40% $B_0$. The ESCR is rapidly rebuilding into the target biomass range (since the 2014-15 TACC will be set well below the current surplus production – assuming that $h = 0.75$ and $M = 0.045$) but a change in stock status resulting from an updated stock assessment could result in the need to rescale. The HCR is not relevant to other orange roughy stocks, such as MEC, which are currently estimated to be below the target biomass range. The rule has been designed for stocks which are within or above the target biomass range. The purpose of the rescaling, below 30% $B_0$, is to allow for low values of $h$ and $M$ which are a possibility if a stock, managed under the HCR, becomes depleted. It is not for stocks that have been depleted due to historical levels of exploitation. It can be applied to stocks, such as MEC, once they are rebuilt to within the target biomass range.

The MSE should be repeated in 4 to 5 years from now when more information will be available on $h$ and $M$ from new stock assessments.

## References


Cordue, P.L. (2001). Short communication: A note on incorporating stochastic recruitment into deterministic age structured population models. *ICES J. Mar. Sci. 58*: 794–798.

Cordue, P.L. (2014). The 2014 stock assessments of orange roughy. Draft New Zealand Fisheries Assessment Report.

DFO. (2010). Stock assessment update for British Columbia canary rockfish. DFO Can. Sci. Advis. Sec. Sci. Resp. 2009/019.

Forrest R.E.; McAllister M.K.; Dorn M.W.; Martell S.J.D.; Stanley R.D. (2010). Hierarchical Bayesian estimation of recruitment parameters and reference points for Pacific rockfishes (*Sebastes* spp.) under alternative assumptions about the stock-recruit function. *Can. J. Fish. Aquat. Sci. 67*:1611–1634.

Francis, R.I.C.C. (1992). Recommendations concerning the calculation of maximum constant yield (*MCY*) and current annual yield (*CAY*). New Zealand Fisheries Assessment Research Document 1992/8. 27 p.

MFish (2008). Harvest Strategy Standard for New Zealand Fisheries. Ministry of Fisheries, Wellington, New Zealand. 30 p.

MPI. (2014). Fisheries Assessment Plenary, May 2014: stock assessments and stock status. Compiled by the Fisheries Science Group, Ministry for Primary Industries, Wellington, New Zealand. 1381 p.

MSC. (2013). MSC Certification Requirements Version 1.3. Marine Stewardship Council,14 January 2013. 355 p.

Punt, A. E.; Smith, A. D. M.; Smith, D. C.; Tuck, G. N.; and Klaer, N. L. (2014). Selecting relative abundance proxies for $B_{MSY}$ and $B_{MEY}$. *ICES J. Mar. Sci. 71*: 469-483.

Shertzer, K.W.; Conn, P.B. (2012). Spawner-recruit relationships of demersal marine fishes: prior distribution of steepness. *Bulletin of Marine Science 88*: 39-50.

Thorson, J. (2013). Preliminary summary of replicating 2011 steepness results and estimating a 2013 steepness prior. Draft document, NWFSC. 6 p.

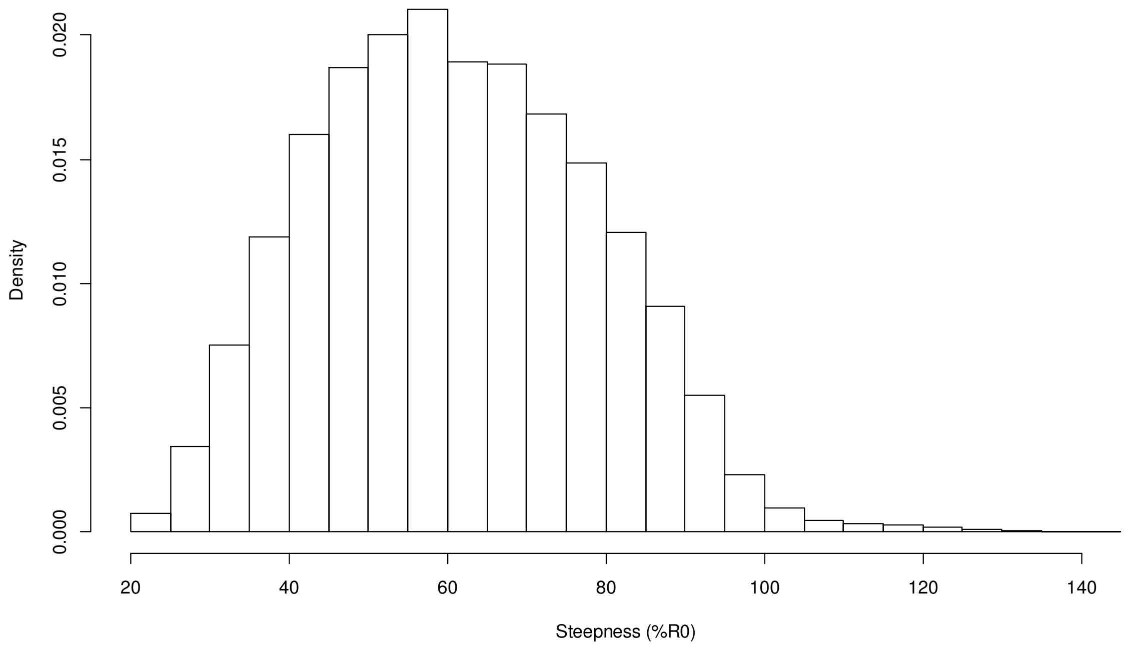


**Figure 1: Combined posterior distribution for steepness from the two MEC assessment runs which assumed a Beverton-Holt or a Ricker stock-recruitment relationship. The two runs were given equal weight.**

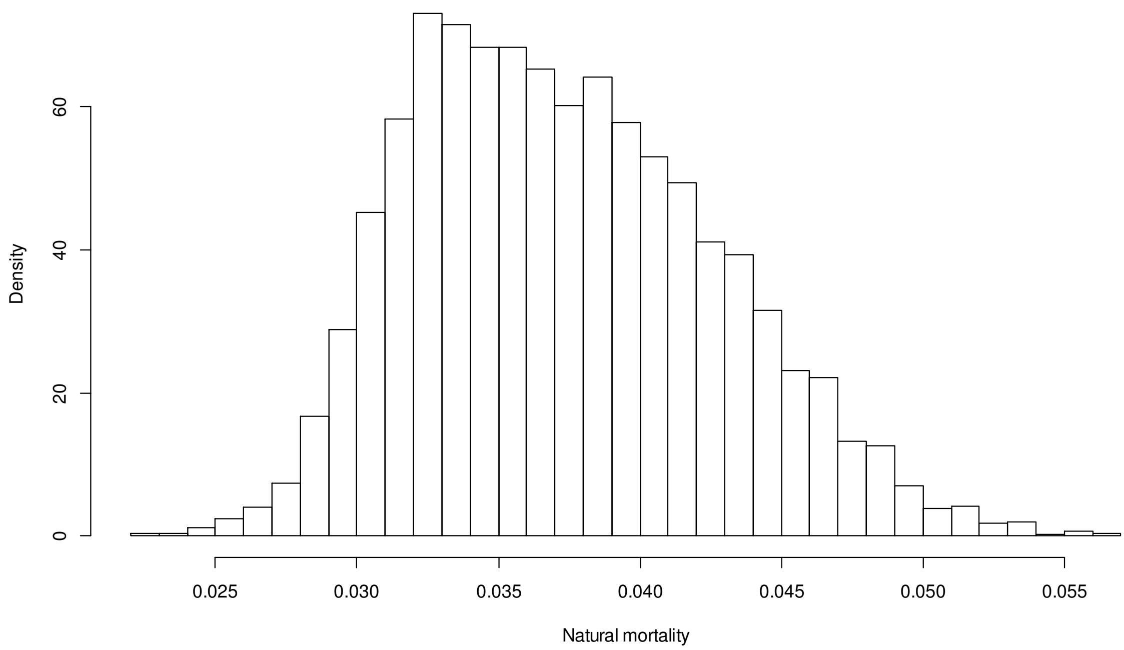


**Figure 2: Combined posterior distribution for natural mortality from the four assessed stocks in 2014. Each stock was given equal weight.**

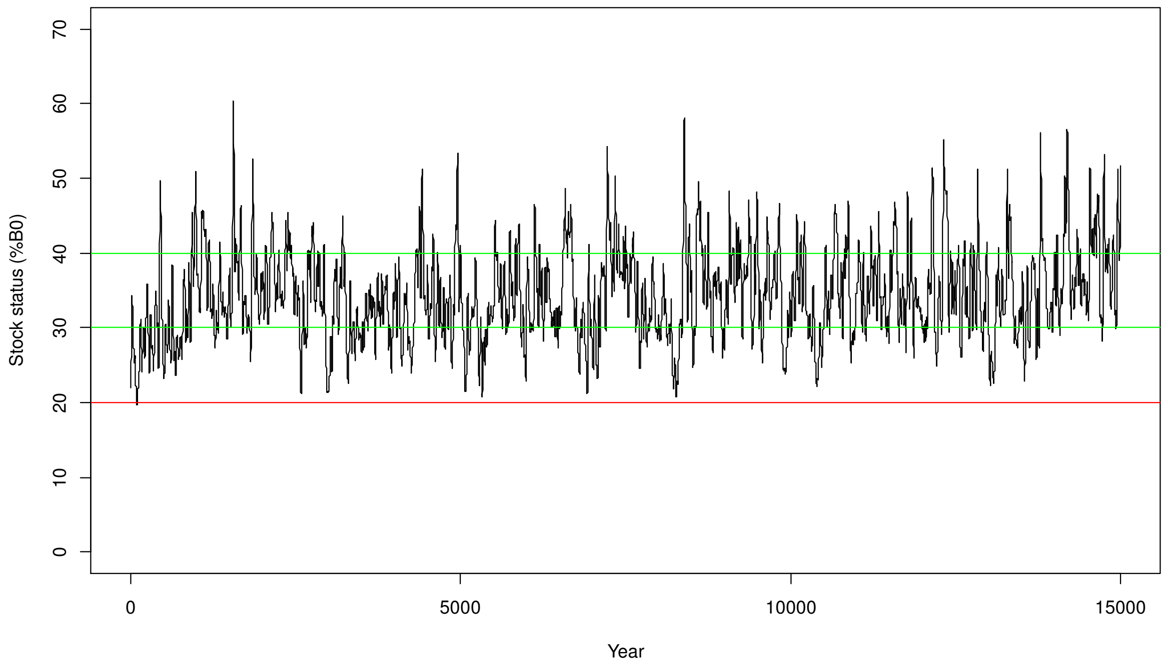


**Figure 3: An example time series of stock-status (mean mid-season spawning biomass as $\%B_0$) for the base model when fishing at a constant $F_{35\%B0}$.**

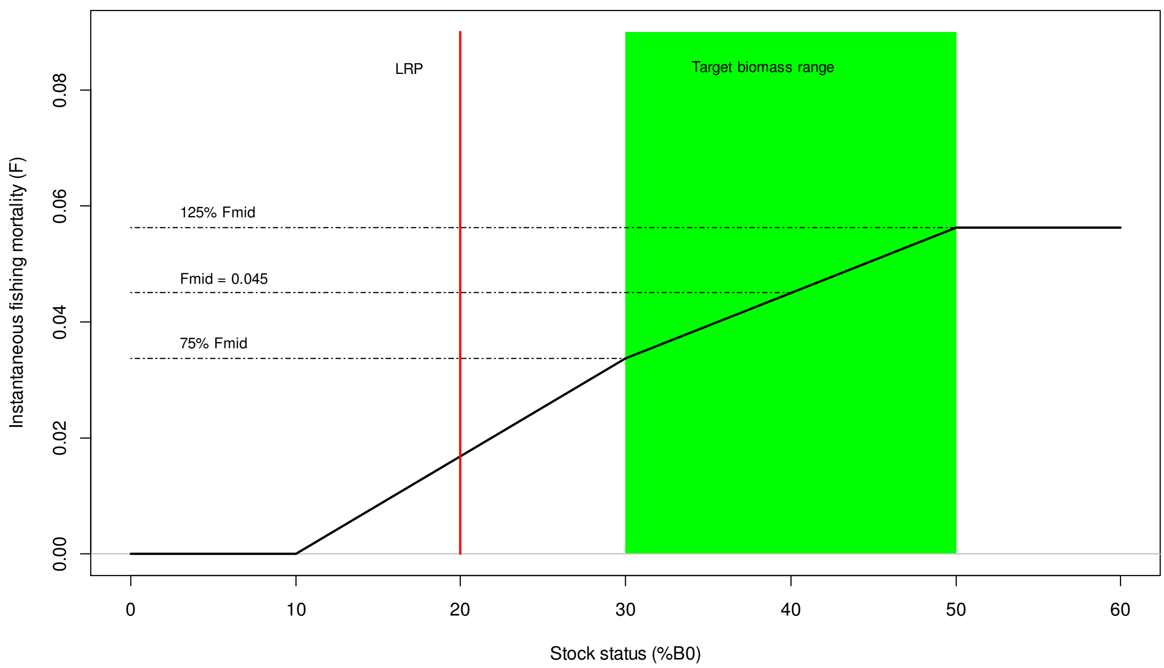


**Figure 4: Proposed harvest control rule, dynamic HCR10: LRP = 20% $B_0$, target biomass range = 30-50% $B_0$, initial $F_{mid}$ = 0.045, slope within the target range: $p$ = 25%; ramps down to zero at 10% $B_0$; rescaling limit points: $l$ = 30% $B_0$, $r$ = 60% $B_0$; $k$ = 0.9, $m$ = 10, $p_{limit}$ = 0.3.**

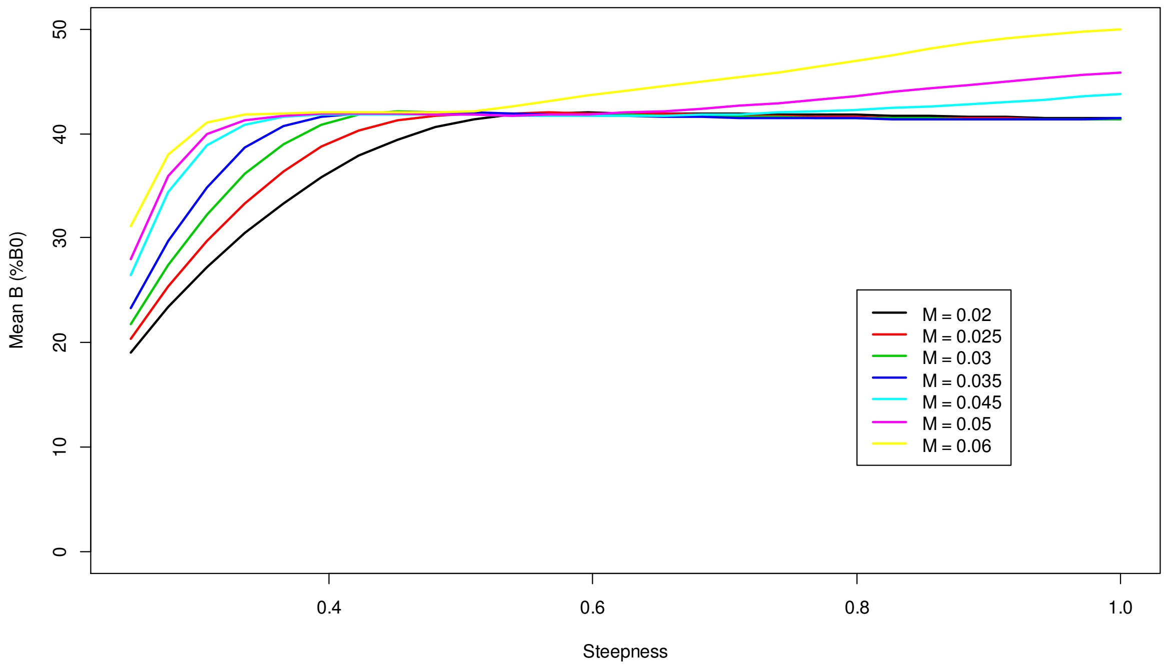


**Figure 5: Base model: average mid-season mature biomass (% $B_0$) maintained by dynamic HCR10 as a function of *h* and *M*. The cubic splines across *h* are plotted for each value of *M* in the grid over which mean biomass was calculated.**

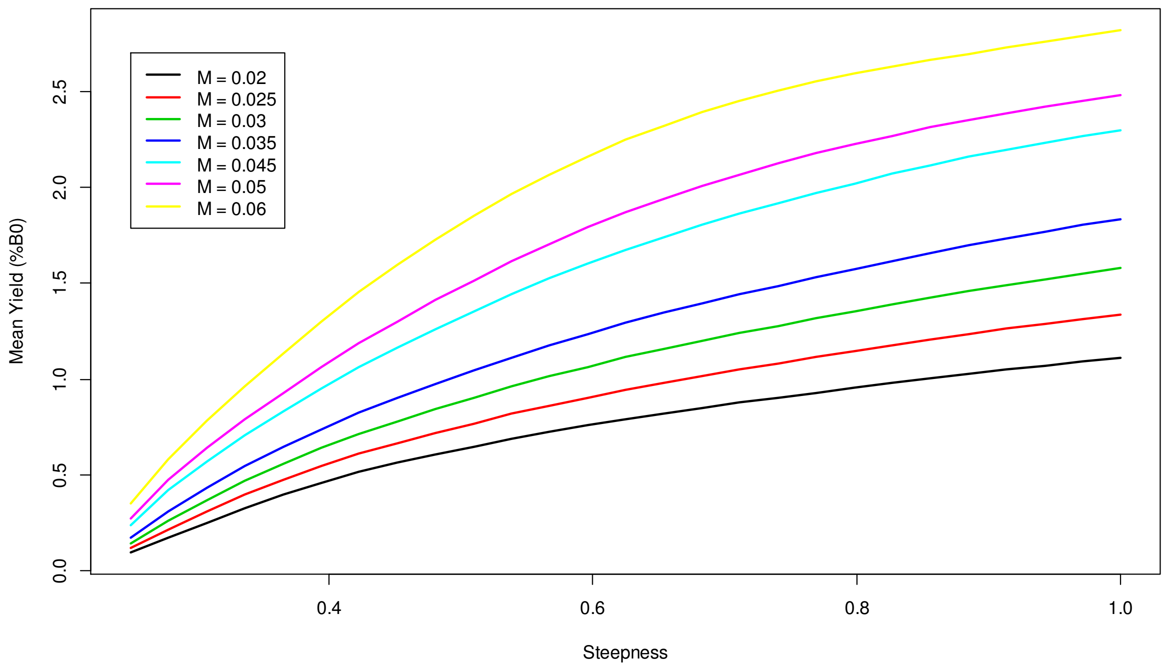


**Figure 6: Base model: average yield (% $B_0$) achieved by dynamic HCR10 as a function of *h* and *M*. The cubic splines across *h* are plotted for each value of *M* in the grid over which mean yield was calculated.**

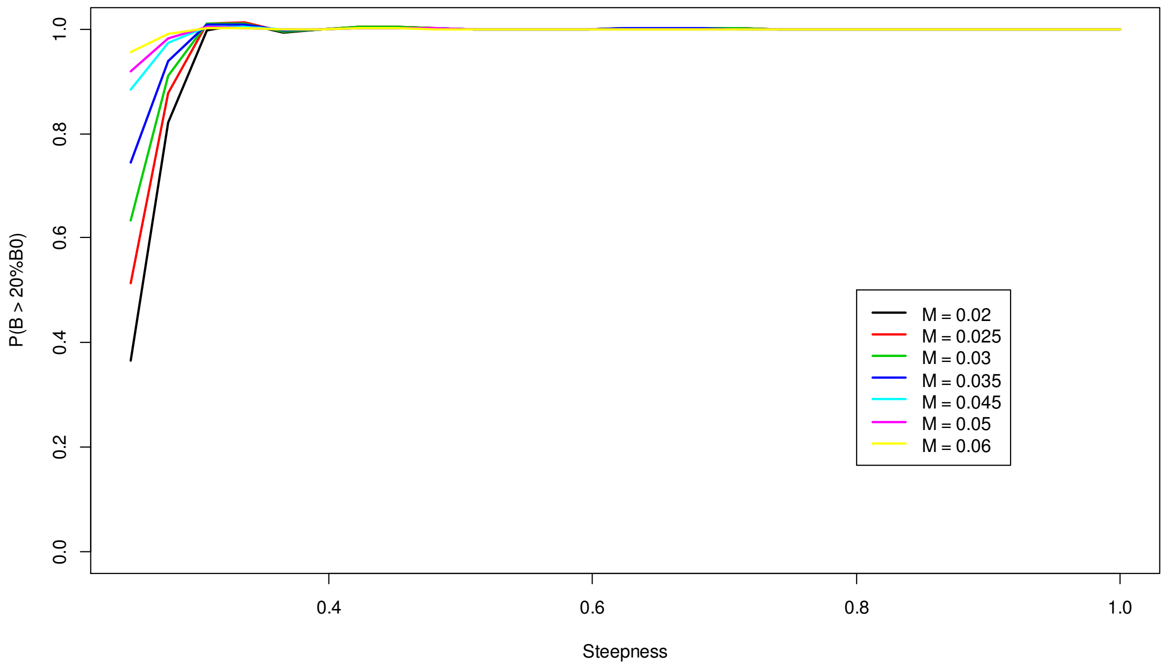


**Figure 7: Base model: the probability that mid-season mature biomass was maintained above the LRP (20% $B_0$) by dynamic HCR10, as a function of *h* and *M*. The cubic splines across *h* are plotted for each value of *M* in the grid over which the probability was calculated.**

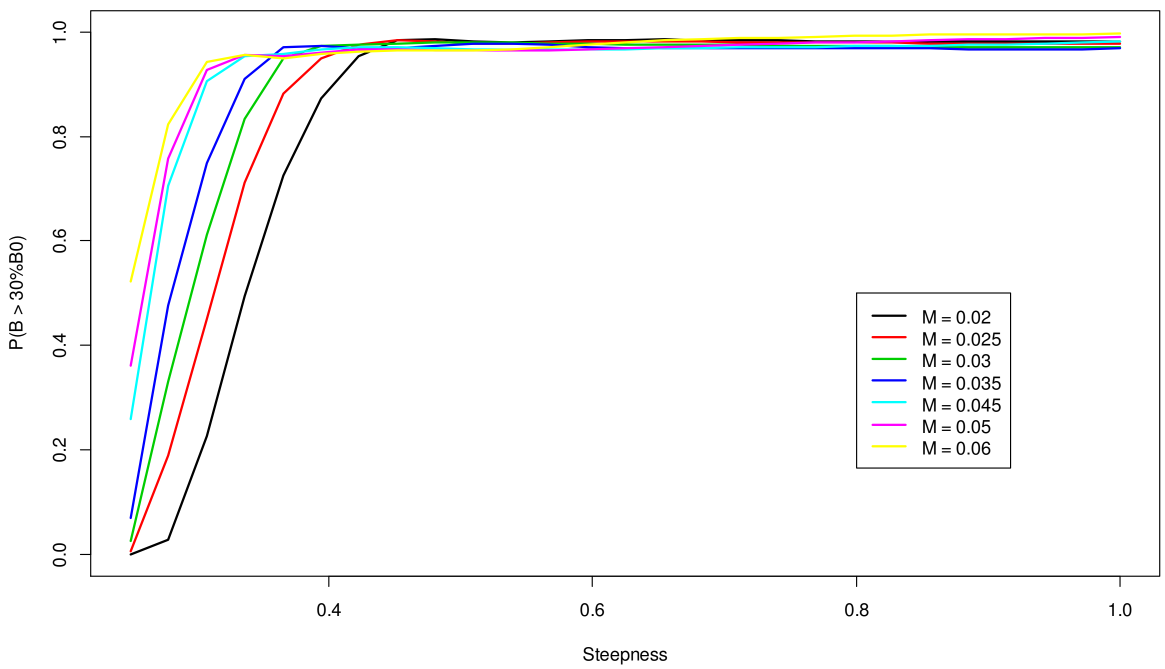


**Figure 8: Base model: the probability that mid-season mature biomass was maintained above the lower bound of the target range (30% $B_0$) by dynamic HCR10, as a function of *h* and *M*. The cubic splines across *h* are plotted for each value of *M* in the grid over which the probability was calculated.**

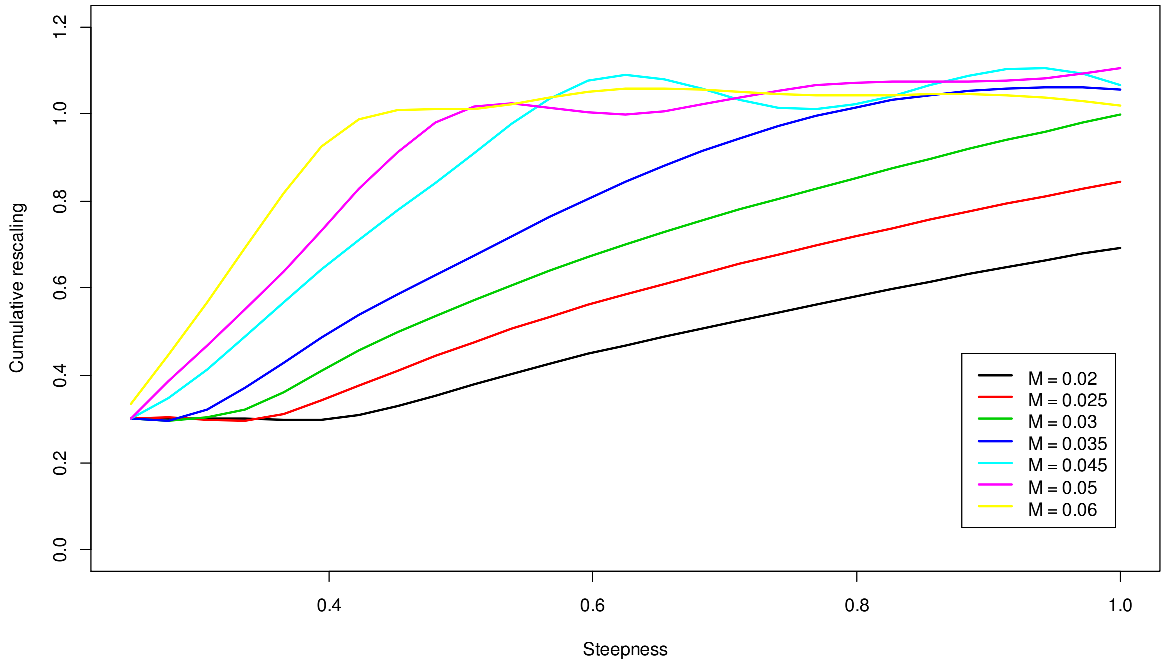


**Figure 9: Base model: the cumulative scaling of the functional relationship (between stock status and *F*) that occurs for dynamic HCR10, as a function of *h* and *M*. The cubic splines across *h* are plotted for each value of *M* in the grid over which the cumulative scaling was calculated.**

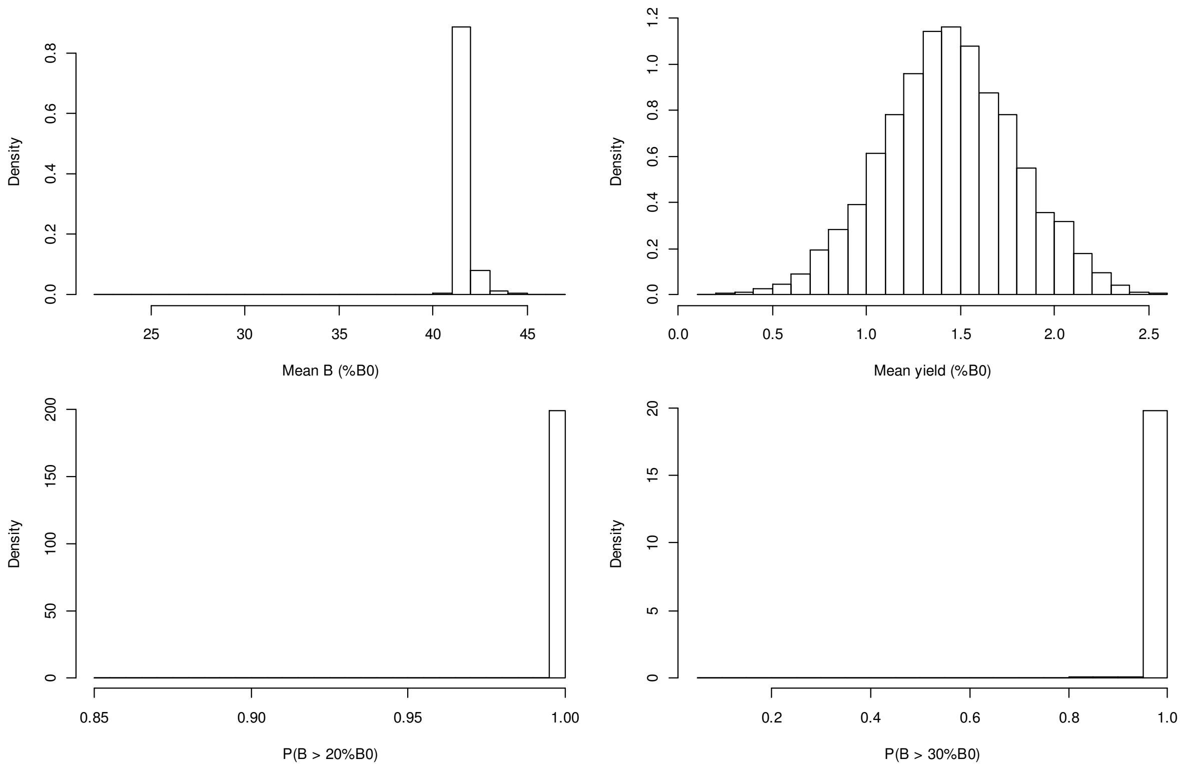


**Figure 10: Base model, dynamic HCR10: Bayesian posteriors for mean mid-season mature biomass (% $B_0$), mean yield (% $B_0$), and the probabilities of mid-season mature biomass being above the LRP (20% $B_0$) or the lower bound of the target range (30% $B_0$).**

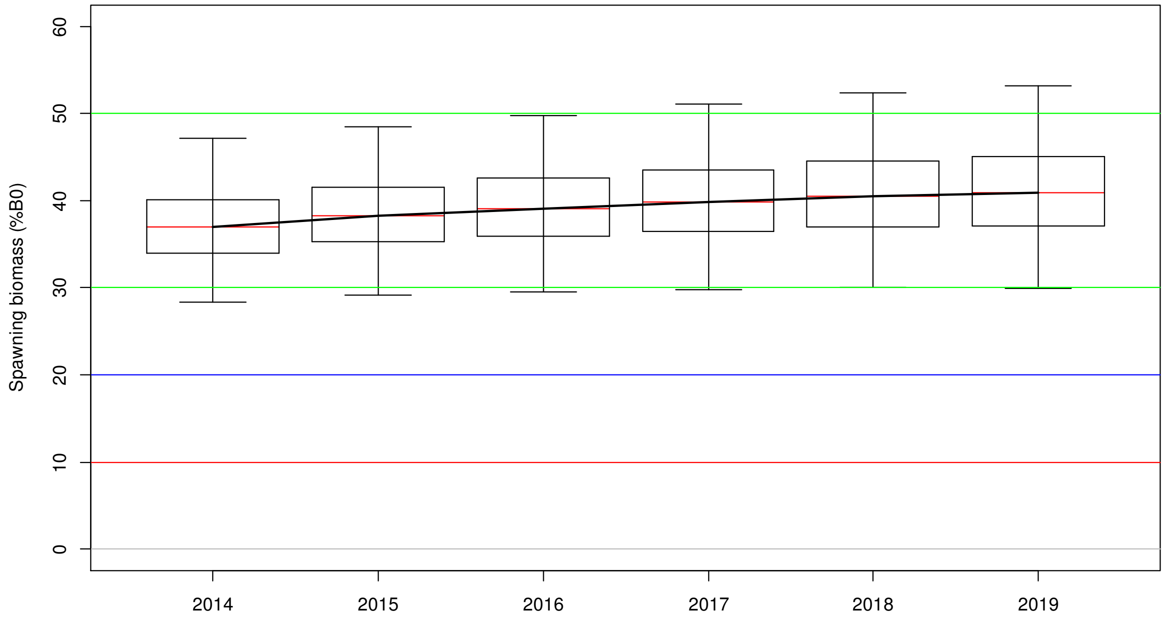


**Figure 11: NWCR, base model: projections under dynamic HCR10 (catch limit: 1043 t for 2015–2018 inclusive; 1332 t in 2019, see Table 14). The box and whiskers plots are of projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

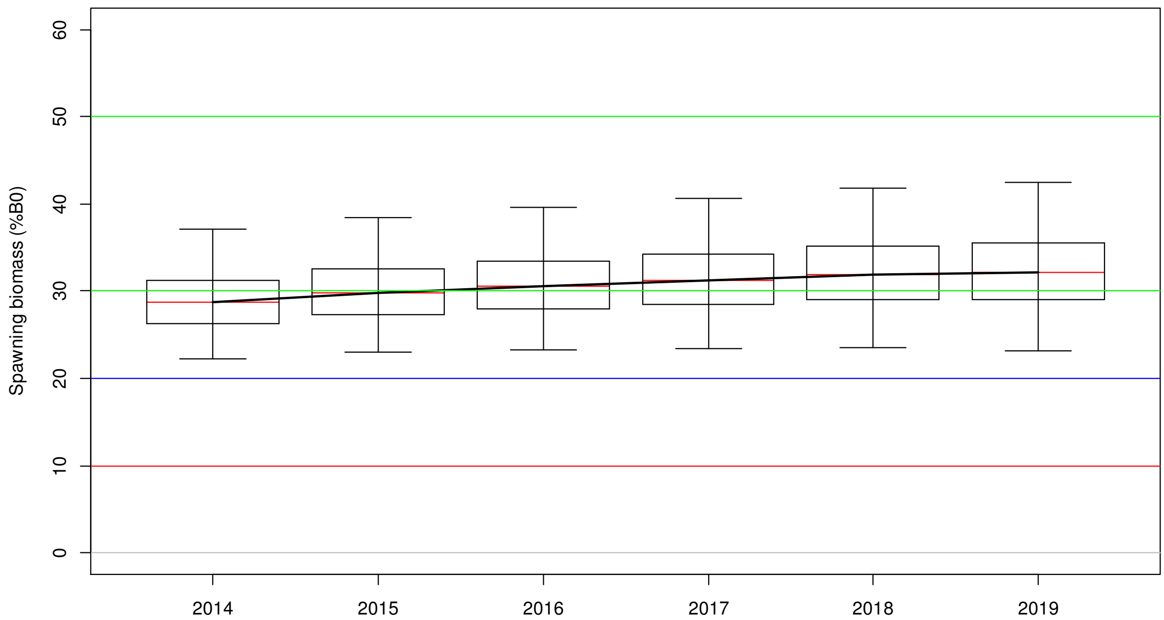


**Figure 12: NWCR, "worst-case" lowM-highq model: projections under the catch limits from dynamic HCR10 applied to the base model (1043 t for 2015–2018 inclusive; 1332 t in 2019, see Table 14). The box and whiskers plots are for projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

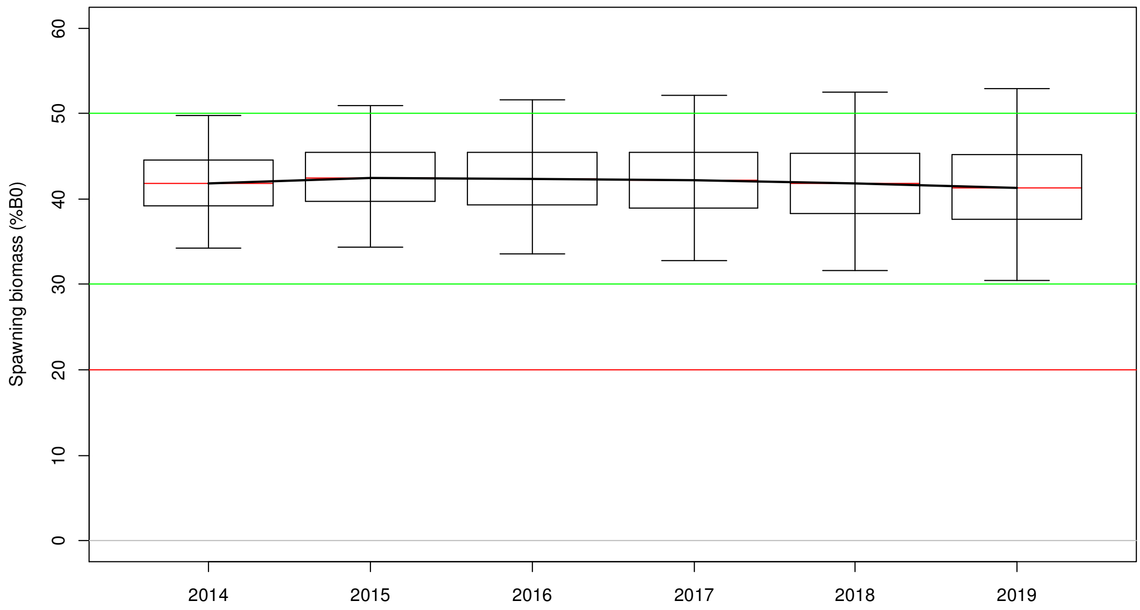


**Figure 13: ORH7A, base model: projections under dynamic HCR10 (catch limit: 1748 t for 2015–2016 inclusive; 1799 t for 2017–2019 inclusive, see Table 14). The box and whiskers plots are of projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

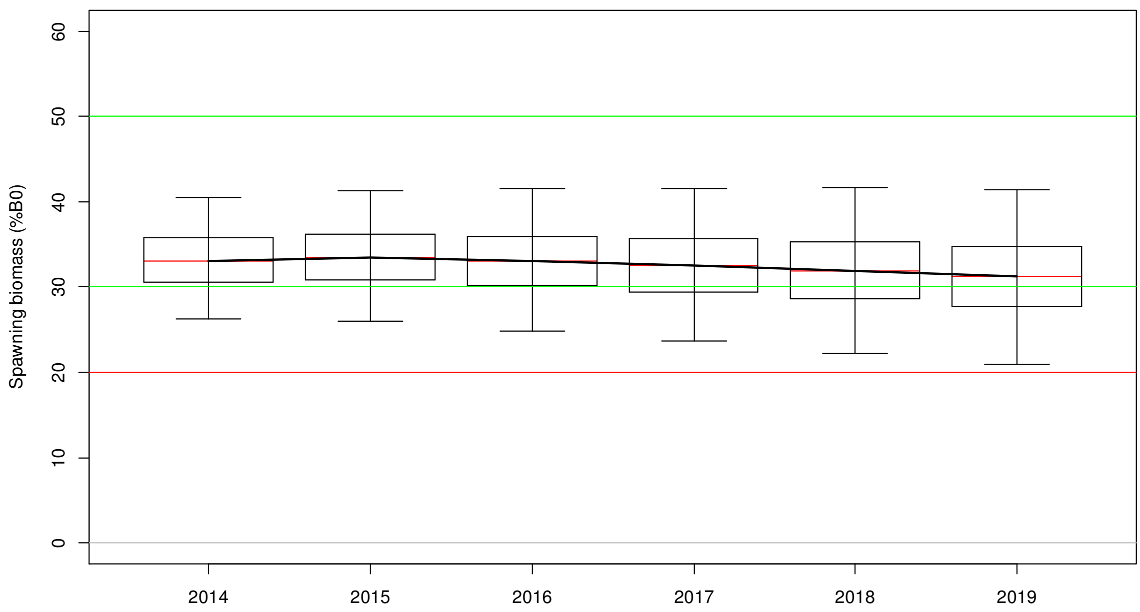


**Figure 14: ORH7A, "worst-case" lowM-highq model: projections under the catch limits from dynamic HCR10 applied to the base model (1748 t for 2015–2016 inclusive; 1799 t for 2017–2019 inclusive, see Table 14). The box and whiskers plots are for projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

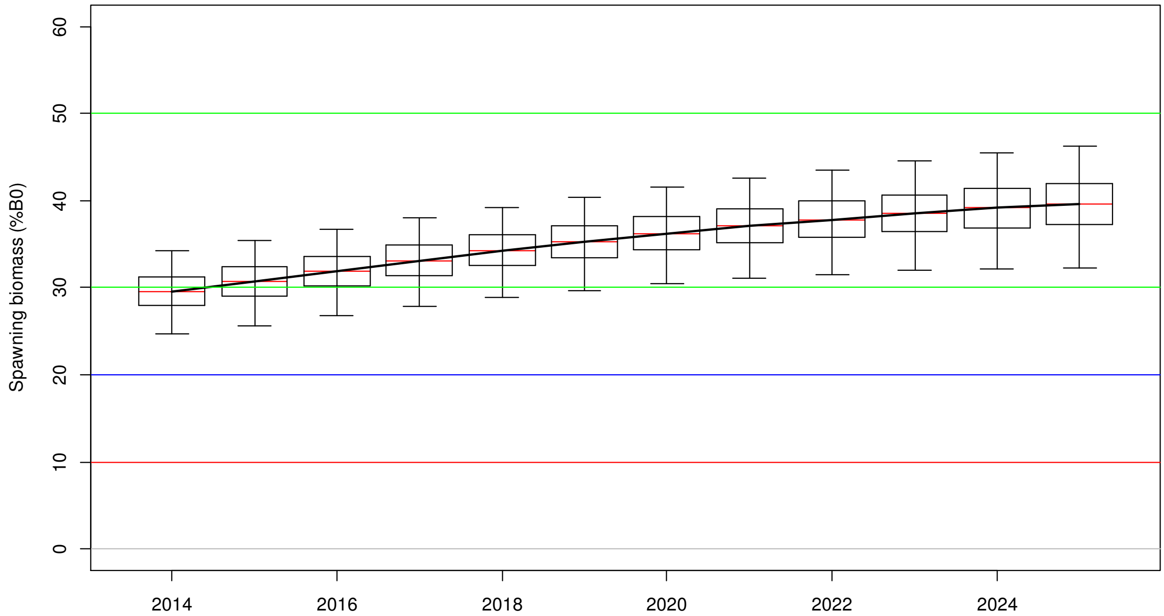


**Figure 15: ESCR, base model: projections under dynamic HCR10 (catch limit: 3772 t for 2015–2018 inclusive; 4965 t for 2019–2021 inclusive; 5768 t for 2022–2024 inclusive; 6317 t in 2025, see Table 14). The box and whiskers plots are of projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

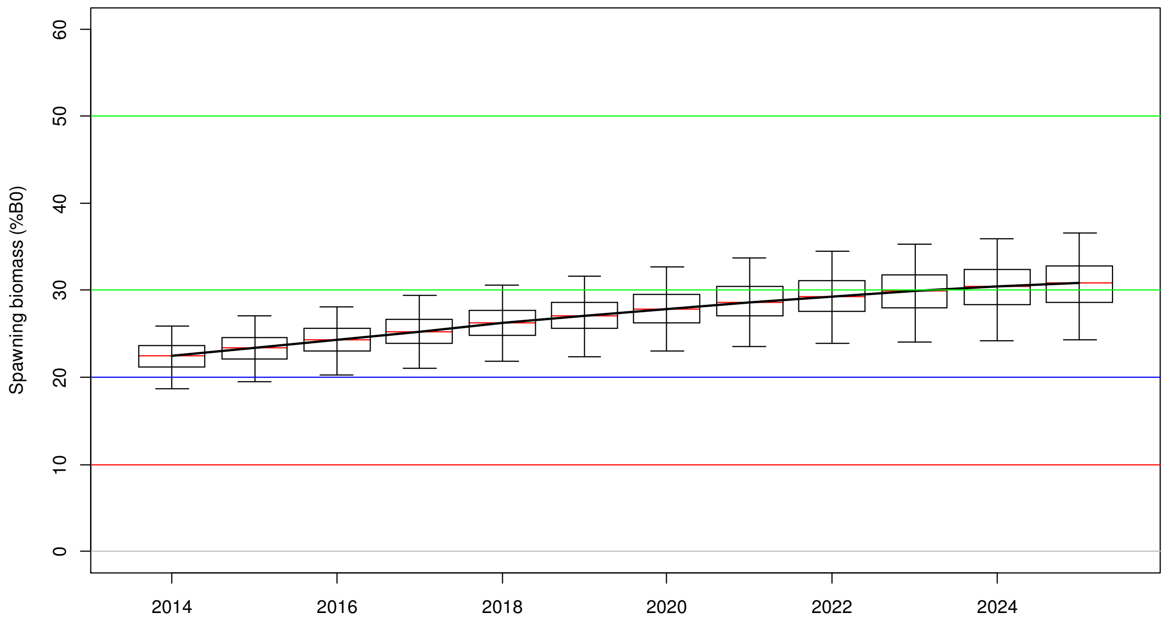


**Figure 16: ESCR, "worst case" lowM-highq model: projections under the catch limits from dynamic HCR10 applied to the base model (3772 t for 2015–2018 inclusive; 4965 t for 2019–2021 inclusive; 5768 t for 2022–2024 inclusive; 6317 t in 2025, see Table 14). The box and whiskers plots are for projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

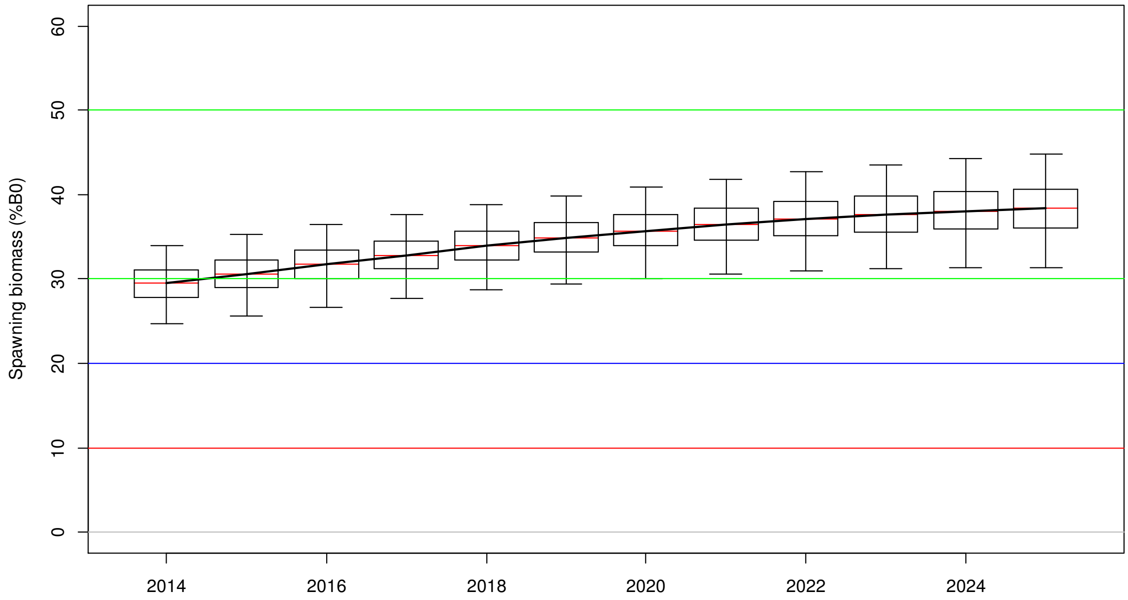


**Figure 17: ESCR, base with $h$ = 0.6: projections under the catch limits from dynamic HCR10 applied to the base model (3772 t for 2015–2018 inclusive; 4965 t for 2019–2021 inclusive; 5768 t for 2022–2024 inclusive; 6317 t in 2025, see Table 14). The box and whiskers plots are for projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

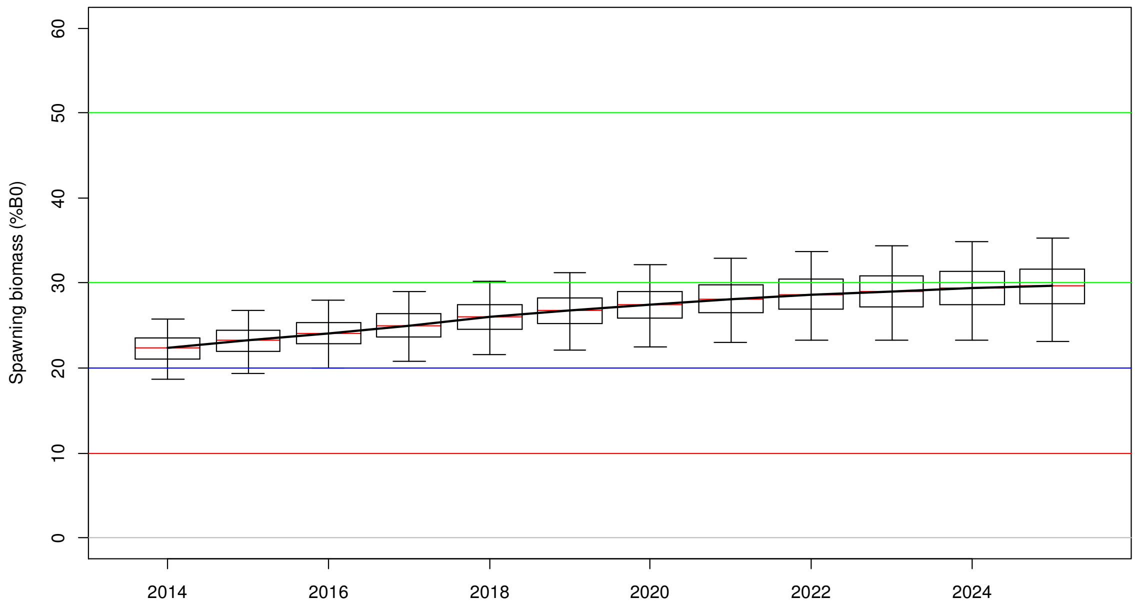


**Figure 18: ESCR, lowM-highq with $h$ = 0.6: projections under the catch limits from dynamic HCR10 applied to the base model (3772 t for 2015–2018 inclusive; 4965 t for 2019–2021 inclusive; 5768 t for 2022–2024 inclusive; 6317 t in 2025, see Table 14). The box and whiskers plots are for projected mid-season spawning biomass. The medians are shown by the horizontal red lines; the boxes cover the middle 50%; and the whiskers extend to the 95% CI.**

# Appendix A: Model equations

## Population dynamics

A standard age-structured population dynamics model was used in the simulations: single sex, single area, Baranov catch equation, with fish numbers tracked by age and maturity state (mature: "*mat*", or immature: "*imm*"). The model was started in deterministic equilibrium with the end-of-year total numbers at age $a$ = 1, …, 200 years, $N_{0,a}$:

$$N_{0,a} = R_0 e^{-aM}$$

where $R_0$ is the number of recruits at age 1 in the virgin population (an arbitrary value of $R_0$ = 100 fish was used). The proportion mature at age $a$ (in the virgin population) was defined to be logistic with given parameters $a_{50}$ and $a_{\mathrm{to}95}$:

$$p_{mat,a} = \frac{1}{1 + 19^{(a50-a)/ato95}}$$

The above equation was modified slightly by specifying that all proportions were 0 below age 10 years and 1 above age 60 years.

The annual cycle consisted of ageing, recruitment, maturation, and mortality (a full year of natural and fishing mortality assuming the Baranov catch equation) in that order. The total number of fish in year $y$+1 at age $a$+1 were obtained from the previous end-of-year numbers:

Ageing: a = 1,…,199 years $N_{y+1,a+1} = N_{y,a}$

The recruitment at age 1, in year $y$+1, was the product of virgin recruitment ($R_0$), the response from the stock-recruitment relationship ($p_{SR}(B_y)$, where $B_y$ is the mid-season mature biomass in year $y$ (see below)) and the "year class strength" ($Y_y$) of the cohort:

Recruitment: $N_{y+1,1} = Y_y p_{SR}(B_y) R_0$

A fixed proportion of immature fish were matured at each age in each year. The fixed maturation ogive was calculated from the logistic proportions mature-at-age in the virgin population:

Maturation: a = 10,…,60 years

$$N_{new,mat,a} = \left(\frac{p_{mat,a} - p_{mat,a-1}}{1 - p_{mat,a-1}}\right) N_{y+1,imm,a}$$
$$N_{y+1,mat,a} = N_{y+1,mat,a} + N_{new,mat,a}$$
$$N_{y+1,imm,a} = N_{y+1,imm,a} - N_{new,mat,a}$$

with no fish matured below 10 years of age and all fish matured above 60 years of age. This formulation ensures that the proportions mature-at-age are in deterministic equilibrium in the virgin population (i.e., do not change when there is no fishing and all YCS are equal to 1).

Mortality was modelled with the Baranov catch equation with either a logistic fishing selectivity at age or a non-age-selective fishery on just the mature fish.

Mortality: $$N_{y,a,end} = e^{-(M+s_aF_y)}N_{y,a,begin}$$

where $M$ is natural mortality (independent of age or maturity), $F_y$ is the fishing mortality in year $y$, and $s_a$ is the selectivity at age $a$ years. The "$N$" terms refer to mature or immature numbers at the beginning and end of the mortality period for a non-age-selective fishery (in which case each $s_a = 1$ for mature fish and $s_a = 0$ for immature fish) or to the total number of fish for an age-selective fishery.

The catch was calculated in the usual way:

Catch: $$C_{y,a} = \frac{s_aF_y}{M+s_aF_y}\left(N_{y,a,begin} - N_{y,a,end}\right)$$
$$C_y = \sum_a w_a C_{y,a}$$

where $w_a$ is the mean fish weight at age $a$ years (calculated from given von Bertalanffy growth and length-weight relationships which are independent of maturity).

Stock status or depletion in year $y$, $D_y$, is defined to be the mid-season mature biomass divided by the mid-season unfished mature biomass: $D_y = B_y/B_{unfished}$. Mid-season occurs when half of the total mortality has been applied. The unfished biomass is the average mid-season mature biomass in the virgin population which is almost equal to the deterministic mid-season virgin mature biomass ($B_0$):

$$B_0 = \sum_a w_a\, p_{mat,a} R_0 e^{-(a-0.5)M}$$
$$B_{unfished} = cB_0.$$

$c$ is a correction factor which depends on many of the parameters in the population model (particularly the variability and correlation driving the year class strengths, $M$, and steepness, $h$, in the stock-recruitment relationship). The correction factors were calculated, as needed, by running the virgin population over 50,000-150,000 years (depending on what was required to make the result independent of the random number seed). The correction factors for the base model ranged from 0.95–1 (over the grid of $h$ and $M$ values used)(Table A1). In the main text, "$B_0$" is used to denote "$B_{unfished}$" as the distinction is obscure for the general reader.

**Table A1: Correction factors required in the base model to scale deterministic mid-season virgin mature biomass to the average mid-season virgin mature biomass.**

| | Natural mortality (*M*) | | | | | | |
|---|---|---|---|---|---|---|---|
| **Steepness (*h*)** | **0.02** | **0.025** | **0.03** | **0.035** | **0.045** | **0.05** | **0.06** |
| **0.25** | 0.97 | 0.97 | 0.96 | 0.96 | 0.96 | 0.96 | 0.95 |
| **0.30** | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.97 |
| **0.35** | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 |
| **0.40** | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| **0.50** | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| **0.60** | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| **0.75** | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| **0.90** | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| **1.00** | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

In the stock-recruitment relationship the uncorrected depletion level ($B_y/B_0$) was used because that is what gives rise to $B_{unfished}$. In the base model the Beverton-Holt relationship was used:

$$p_{SR}(B_y) = \frac{B_y/B_0}{\left[1 - \frac{(5h-1)}{4h}(1 - B_y/B_0)\right]}$$

and in some sensitivities the Ricker relationship was used:

$$p_{SR}(B_y) = \frac{B_y}{B_0}(5h)^{\frac{5}{4}(1-B_y/B_0)} \; .$$

The year class strengths, $Y_y$, were assumed to follow an AR(1) process in log space:

$$Y_y \sim \text{LN}(\mu, \sigma_R),\; Y_y = e^{X_y} \text{ where } X_y = d + \rho X_{y-1} + \epsilon_y,\; \epsilon_y \sim \text{N}(0, \sigma^2),\; \text{and } X_0 = 0.$$

It follows that: $\mu = d/(1-\rho)$ and $\sigma_R = \sigma/\text{sqrt}(1-\rho^2)$. The constant $d$ is defined by the requirement that $\text{E}(Y_y) = 1$. There appears to be no analytical solution for $d$ but a good approximation can be found by solving the following equation iteratively:

$$d = \log\left[\frac{1-\rho}{1-\rho+\rho d+\frac{\rho^2}{2}\left(\frac{\sigma^2}{1+\rho}+\frac{d^2}{1-\rho}\right)}\right] - \frac{\sigma^2}{2}$$

The above equation is derived by noting that

$$E(Y_y) = E(e^{X_y}) = E(e^d e^{\rho X_{y-1}} e^{\epsilon_y}) = e^d E(e^{\rho X_{y-1}}) e^{\frac{\sigma^2}{2}} = 1$$

and approximating $\text{E}(e^{\rho X_{y-1}})$ with a second order Taylor approximation:

$$\mathrm{E}(e^{\rho X_{y-1}}) \cong 1 + \rho\left(\frac{d}{1-\rho}\right) + \frac{\rho^2}{2}\left(\frac{\sigma^2}{1+\rho} + \frac{d^2}{1-\rho}\right)$$

Note that $\rho$ is the correlation coefficient for successive YCS and that when $\rho = 0$ we have $\sigma_R = \sigma$ and the familiar $d = -\sigma^2/2$.

## Simulation of assessments

To apply a HCR the current stock status and vulnerable biomass must be estimated. It is clear, in reality, that successive biomass estimates will be highly correlated as each new estimate uses only a small amount of extra data. For each simulation run, time series of correlated and potentially biased biomass estimates were constructed for beginning-of-year vulnerable biomass and mid-season mature biomass. Estimates were available in every year but only those estimates that were needed by the control rule were used. The estimate of stock status was the estimate of mid-season mature biomass divided by $B_{unfished}$.

Let $B_1, \ldots, B_y$ be a sequence of true biomasses from the model (i.e., vulnerable or mature). The estimated biomass series $\hat{B}_1, \ldots, \hat{B}_y$ was formed:

$$\hat{B}_1 = qB_1\epsilon_1$$

$$\hat{B}_y = p\left[q\left(B_y - B_{y-1}\right) + \hat{B}_{y-1}\right] + q(1-p)B_y\epsilon_y$$

where $p$ is a non-zero proportion (providing the correlation between estimates), $q$ provides for a potential bias, and $\mathrm{E}\left(\epsilon_y\right) = 1,\ \mathrm{Var}\left(\epsilon_y\right) = \sigma_y{}^2$. It is easy to prove, by induction, that $\mathrm{E}\left(\hat{B}_y\right) = qB_y$. Further, it was assumed that the $\epsilon_y$ were lognormal random variables and that the biomass estimators had a constant CV. It follows that:

$$\sigma_1 = c \ \text{ and } \ \sigma_y = \frac{c\left[1 - p^2\left(\frac{B_{y-1}}{B_y}\right)^2\right]^{1/2}}{1-p}$$

where $\mathrm{CV}\left(\hat{B}_y\right) = c$. Also, $\epsilon_y \sim \mathrm{LN}\left(\frac{-\gamma_y{}^2}{2}, \gamma_y\right)$ where $\gamma_y{}^2 = \log\left(\sigma_y{}^2 + 1\right)$. A requirement for the constant CV is that $p\,B_{y-1}/B_y < 1$ (so that $\sigma_y$ is defined). For some runs this condition was not always satisfied (especially early in the time series when the fish-down was occurring) in which case the previous $\sigma_y$ was used (i.e., $\sigma_y = \sigma_{y-1}$). Note, the correlation coefficient for successive biomass estimators is $p\,B_{y-1}/B_y$.

The same $\epsilon_y$ were applied to the stock status and vulnerable biomass estimates (as they are both the product of the same stock assessment).

## Application of harvest control rules (HCRs)

Each HCR specified an assessment frequency $n$. In a simulation run, with a given HCR, an assessment was performed in the first year and then every $n$ years after that. In a non-assessment year, the TAC was unchanged. In an assessment year, the TAC and TACC were calculated from the HCR using the *estimates* of stock status and vulnerable biomass (see above) and the associated $F$ from the HCR: $\mathrm{TACC}_y = F\hat{B}_{vul,y}$. The TAC was derived from the TACC by adding an allowance of 5% for incidental catch: TAC = 1.05 × TACC. The TAC, in each year, was removed from the stock by calculating the actual fishing mortality required to remove the TAC (i.e., $F_y = \mathrm{TAC}_y/B_{vul,y}$; note this approximation is very accurate given the low total mortalities involved).

## Dynamic HCRs

HCRs are normally "static" in the sense that for any given estimate of stock status a specific $F$ is specified and this relationship *never* changes over time. A "dynamic" HCR is one where the functional relationship between estimated stock status and $F$ can be changed over time as part of the HCR. A dynamic HCR consists of a rule for changing the functional relationship and an initial functional relationship.

In this study, the dynamic HCRs used a mechanism for changing the functional relationship ($g$) over time based on two limit points ($l$, $r$) and a scaling value which also depended on the estimated stock status ($p(s) < 1$). At each assessment, if the estimated stock status ($s$) was below the limit point, $l$, then $g$ was scaled down by $p(s)$:

$$\text{If } (s < l) \text{ then } g_{new} = p(s)\ g_{old}$$

Similarly, if the estimated stock status was above the other (much higher) limit point, $r$, and the cumulative scaling on $g$ was less than 1 , then $g$ was scaled up:

$$\text{If } (s > r \text{ and "cumulative scaling on } g\text{"} < 1) \text{ then } g_{new} = g_{old} / p(s)$$

The scaling up or down of the functional relationship means that each $F$ is scaled up or down. For example,

$$g_{new} = p(s)\ g_{old}$$

means that for every stock status $x$,

$$g_{new}(x) = p(s)\ g_{old}(x).$$

The maximum cumulative scaling down of the initial $g$ was limited to $p_{limit}$.

The scaling function $p$ was piecewise defined:

$$p(s) = k + (1 - k)\left(\frac{s}{l}\right)^m \quad \text{for } 0 \le s < l$$

$$p(s) = k + (1 - k)\left(\frac{r + l - s}{l}\right)^m \quad \text{for } r < s \le r + l$$

$$p(s) = k \quad \text{for s} > r + l.$$

The function is complicated but, for large $m$, it is essentially equal to $k$ between 0 and $l$ and above $r$ (e.g., see Figure A1 for the scaling function used in the recommended HCR). The purpose of having this complicated scaling function, rather than a simple step function, is to avoid the discontinuity at $l$ (e.g., if $l$ = 30% $B_0$ and there is a step function then $s$ = 29.9% $B_0$ would require a scaling down of the functional relationship; with the continuous scaling function it is still scaled down but by a factor very close to 1).

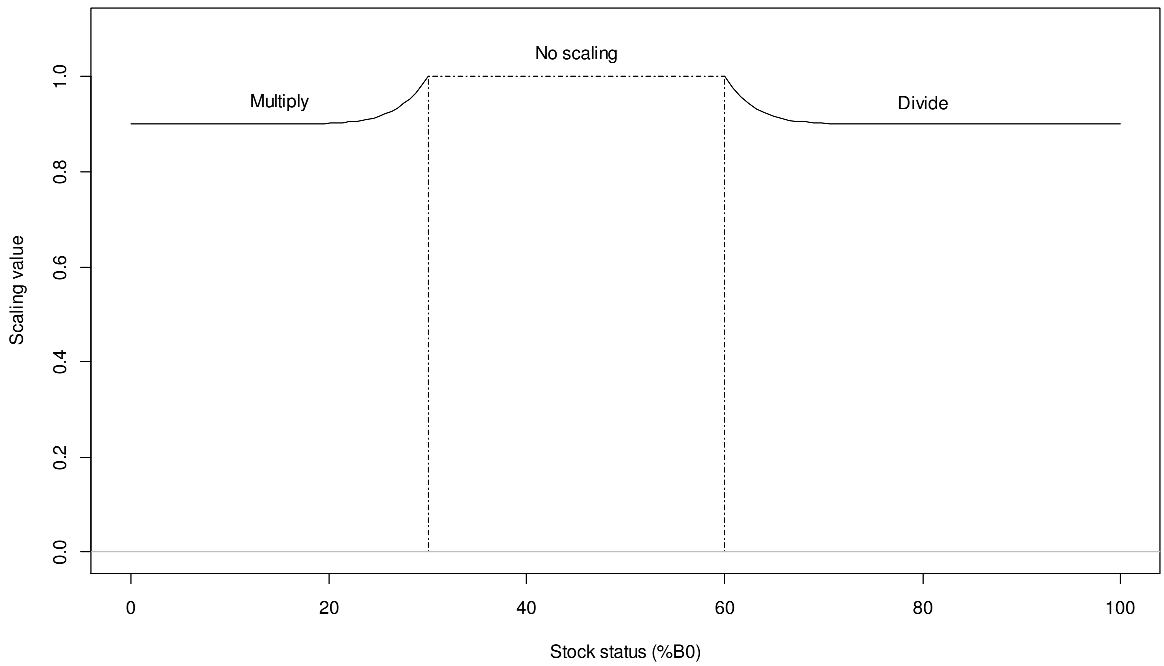


**Figure A1: The scaling function used in dynamic HCR10: $l$ = 30% $B_0$, $r$ = 60% $B_0$, $k$ = 0.9, $m$ = 10. There is no scaling when estimated stock status is from 30–60% $B_0$; below 30% $B_0$ the functional relationship is multiplied by the scaling value; above 60% $B_0$ the functional relationship is divided by the scaling value (provided the current cumulative scaling is less than 1).**

## Appendix B: MEC MCMC runs to estimate stock recruitment steepness

This appendix describes the two MEC assessment runs that were done to estimate steepness ($h$) in the stock-recruitment relationship. The runs differed in the assumed form of the relationship (Beverton-Holt or Ricker) and the informed prior on $h$ (although the priors were equivalent in terms of the prior assumptions with regard to the slope at the origin of the stock-recruit relationship). Apart from those differences the two runs were the "estM" run of the 2014 MEC stock assessment (MPI, 2014) – that is the runs were the MEC base model except that $h$ and $M$ were estimated (with informed priors).

The Beverton-Holt informed prior for $h$ was taken from an update of a Canadian canary rockfish assessment (DFO, 2010) which used a prior developed from west coast USA rockfish assessment results (Forrest et al. 2010). This was the most conservative of three priors that were considered (Figure B1). Drawing on work for a group of rockfish species is appropriate because of the biological similarity between the rockfish and orange roughy, notably in relation to their longevity.

The Ricker informed prior for $h$ was constructed from the Beverton-Holt prior assuming that the distribution of the slopes at the origin were identical for the Beverton-Holt and Ricker relationships. This was achieved by transforming a large random sample from the Beverton-Holt prior using the relationship:

$$h_R = \frac{1}{5}\left(\frac{4h}{1-h}\right)^{\frac{4}{5}}$$

where $h$ is the Beverton-Holt steepness and $h_R$ is the Ricker steepness (the relationship follows from differentiating the two relationships and equating the slopes at the origin). A random sample of 10,000 from the Beta(5,2) prior was truncated at 0.95 (to avoid very high values of $h_R$) and transformed as above. This gave an approximately lognormal distribution with mean = 1.66 and CV = 0.69 (the parameterisation required by CASAL; or, for R: LN(0.27, 0.62)).

For both runs, MCMC estimates were obtained by running 3 independent chains starting at a random jump from the MPD estimate. Chains were run out to 15,000,000 samples with every 1000$^{th}$ sample retained. A burn-in of 2000 samples was applied based on a plot of the objective function values from each of the three chains (e.g., Figure B2). Convergence was judged to be adequate on the basis of two diagnostics: a plot comparing the posterior distributions of the three chains; and a plot examining the stability of the median across all three chains combined. For both runs, the three chains had very similar posteriors (Figures B3 and B4) and, for the three chains combined, there was a high degree of stability in the estimated steepness (being the median of the posterior, see Figures B5 and B6).

A formal procedure was developed to obtain a measure of the precision with which the median was determined for each run. For each run, the effective sample size for an autoregressive (AR-1) sampling process, and a CV and 95% CI on the median were determined by bootstrapping. The results indicate that the CVs on the estimates of the posterior medians were about 2.5% and hence the 95% CIs were fairly tight (plus or minus about 5%)(Table B1).

**Table B1: To do with the precision with which the median of the posterior distribution for steepness was estimated. The estimate of median steepness and CVs and 95% CIs for median steepness from the Beverton-Holt and Ricker MCMC runs are given. See the text for a description of the correlation coefficient ($p$), the CV of the segment medians ($CV_{segments}$), and the effective sample size of the AR-1 process (AR1-$N_{eff}$).**

| | Median steepness | | | | | |
|---|---|---|---|---|---|---|
| | **CV (%)** | **Estimate** | **95% CI** | $p$ | $CV_{segments}$ **(%)** | **AR1-$N_{eff}$** |
| **Beverton-Holt** | 2.4 | 0.68 | 0.64-0.71 | 0.89 | 13 | 507 |
| **Ricker** | 2.7 | 0.53 | 0.50-0.55 | 0.73 | 16 | 390 |

The formal procedure is described below.

A distribution $D(\mu,\sigma)$ can be "sampled" by an AR-1 process giving a time series of random variables $X_1, \ldots, X_n$:

$$X_1 \sim D(\mu, \sigma)$$

$$X_i = pX_{i-1} + (1-p)\epsilon_i$$

where

$$\epsilon_i = \sqrt{\frac{1+p}{1-p}} R_i + \left(1 - \sqrt{\frac{1+p}{1-p}}\right)\mu \quad \text{and} \quad R_i \sim D(\mu, \sigma).$$

It is easy to show by induction that, for every $i$, $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. Also, the correlation between consecutive samples is $p$. When $p = 0$, we have simple random sampling from the distribution.

The MCMC samples from a single variable (or derived variable) are conceptually similar to the samples from AR-1 sampling of the marginal posterior of the variable. The MCMC sampling will not be exactly the same but a comparable AR-1 process can be used to estimate the likely precision that has been achieved after sampling by multiple chains up to some given length (after a suitable burn-in). The following procedure was used for both runs (each of which had 3 chains of 13,000 retained samples or 39,000 samples when the chains are concatenated).

The samples (after burn-in) from the combined chains were used to represent the distribution $D$ (i.e. in the AR-1 process, when random samples from $D$ were needed, the concatenated MCMC samples were sampled at random with replacement). The correlation, $p$, of the AR-1 process was taken to equal the 1-lag correlation coefficient for the concatenated samples and $\mu$ was set equal to the sample mean. An effective sample size for the AR-1 process was determined by equating variance between the AR-1 process and the concatenated samples. The concatenated samples were split into 39 segments each of 1000 samples and the CV of the segment medians was calculated. The effective $n$ was then determined by bootstrapping (5000 simulations) at different values of $n$ until the median CV of the 39 bootstrapped segments was equal to the observed CV of the segment medians. When equality was achieved, the CV of the distribution of the medians from the bootstrapped concatenated samples was calculated. Also, the spread of the bootstrap medians was applied to the median of the actual concatenated samples to give a 95% CI on the median. (See Table B1.)

The posterior distribution for steepness in the Beverton-Holt run was slightly to the left and slightly tighter than the prior distribution (Figure B7). However, for the Ricker run the posterior distribution was well to the left of the prior (eliminating most values over 1) and had a much reduced spread (Figure B8).

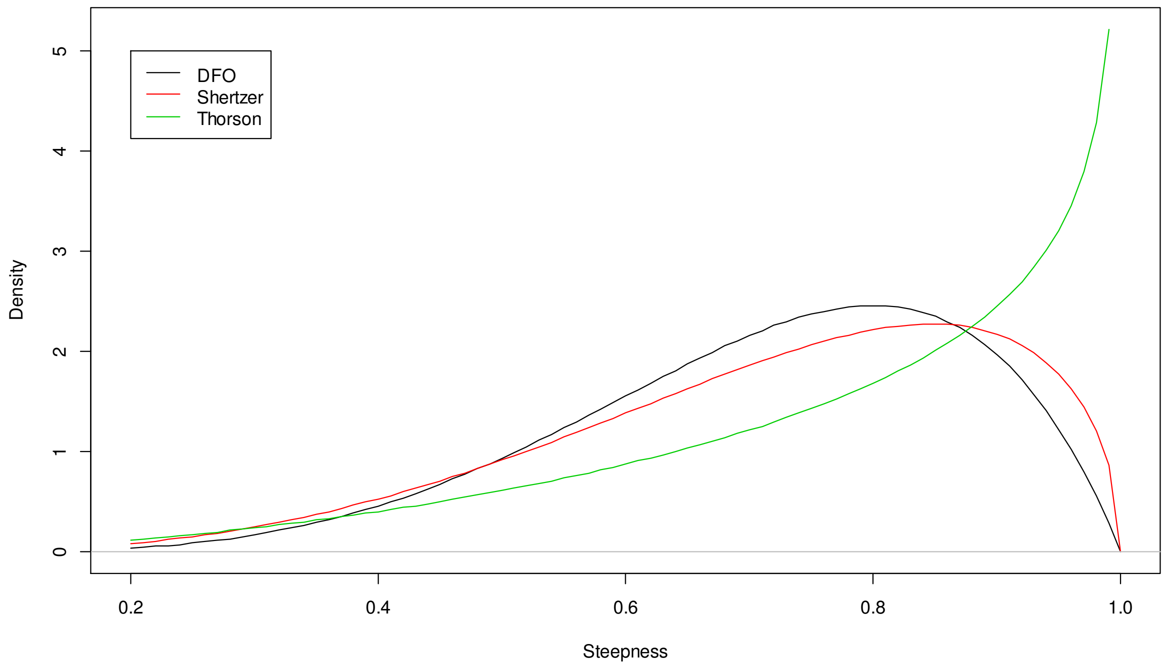


**Figure B1: Three alternative prior distributions for Beverton-Holt steepness from DFO (2010), Shertzer & Conn (2012) and Thorson (2013). The distributions are respectively: beta(5,2), beta(3.89, 1.52), and beta(2.65, 0.74) (reported as a distribution with mean = 0.782, sd = 0.197; the beta parameters were derived by equating the beta distribution mean and sd to those given; comparing the densities visually confirmed this was adequate).**

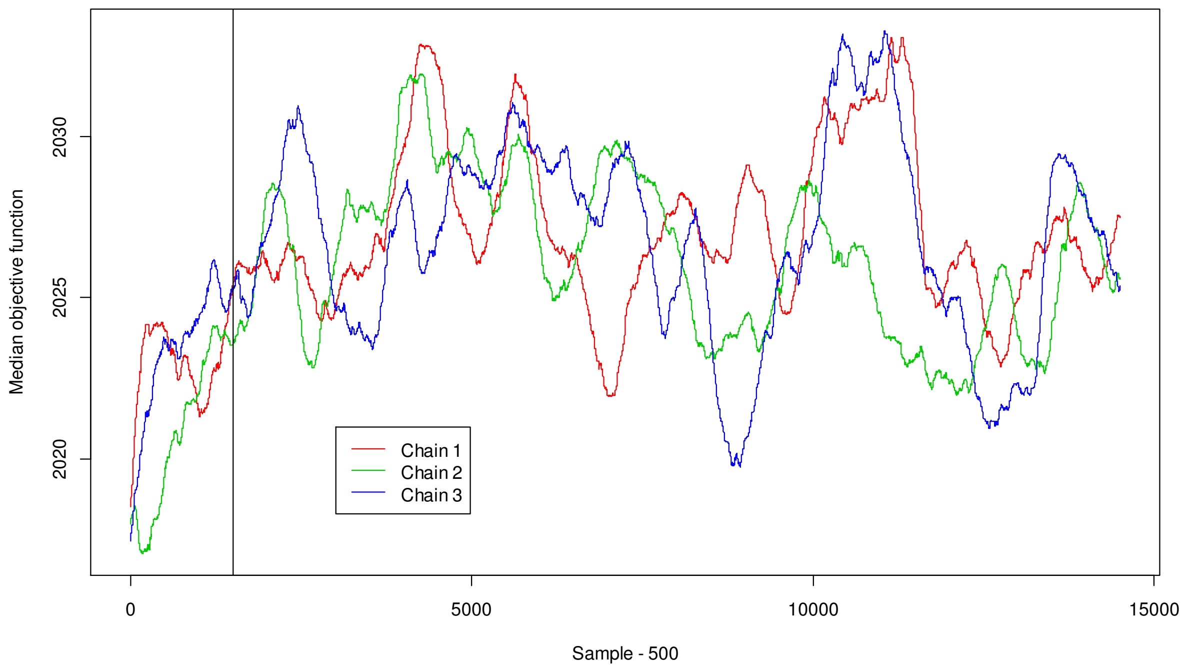


**Figure B2: The moving median (length 500 samples) of the objective function for the three chains in the MCMC run estimating Beverton-Holt steepness. The vertical line is at sample 1500 corresponding to a burn-in of 2000 retained samples. (Note, the three chains have an increasing median objective function during the burn-in as they move away from the MPD before fluctuating over a range of higher values.)**

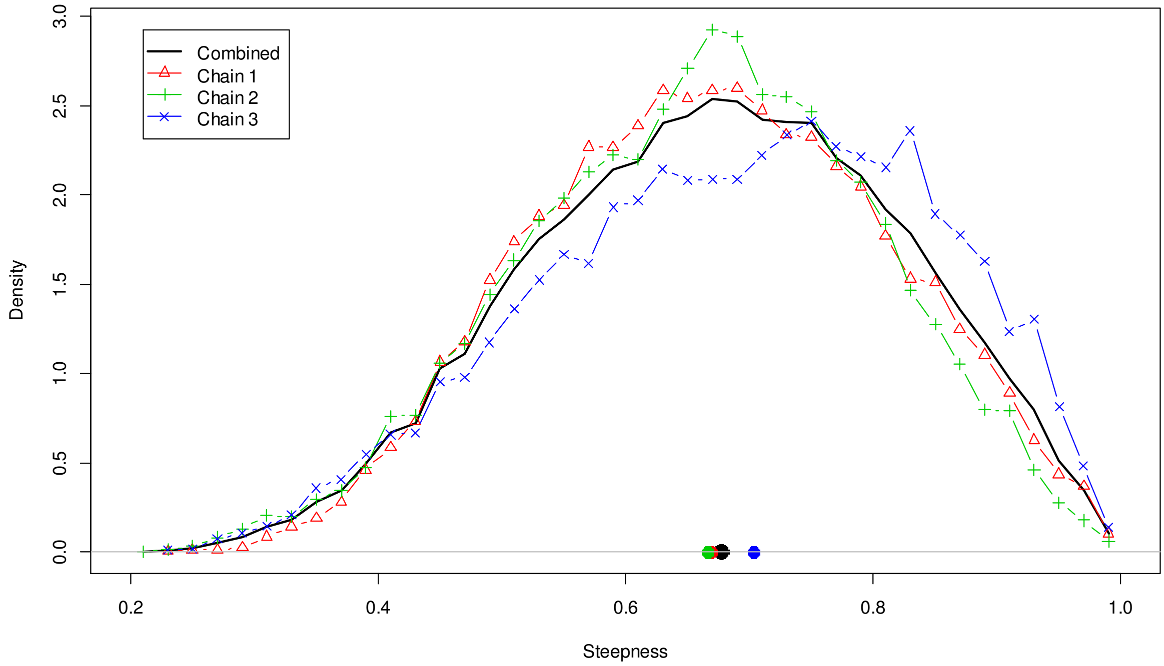


**Figure B3: The posterior distributions of steepness for the three chains in the Beverton-Holt MCMC run and the combined distribution across the chains. The medians are shown by the solid circles on the x-axis.**

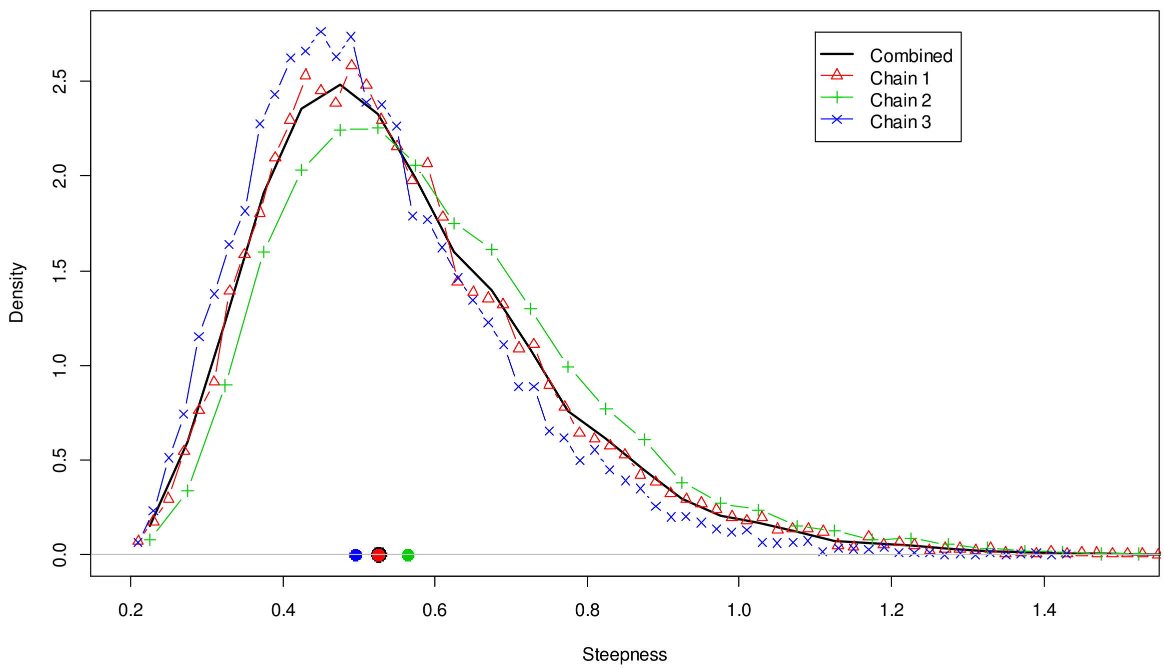


**Figure B4: The posterior distributions of steepness for the three chains in the Ricker MCMC run and the combined distribution across the chains. The medians are shown by the solid circles on the x-axis.**

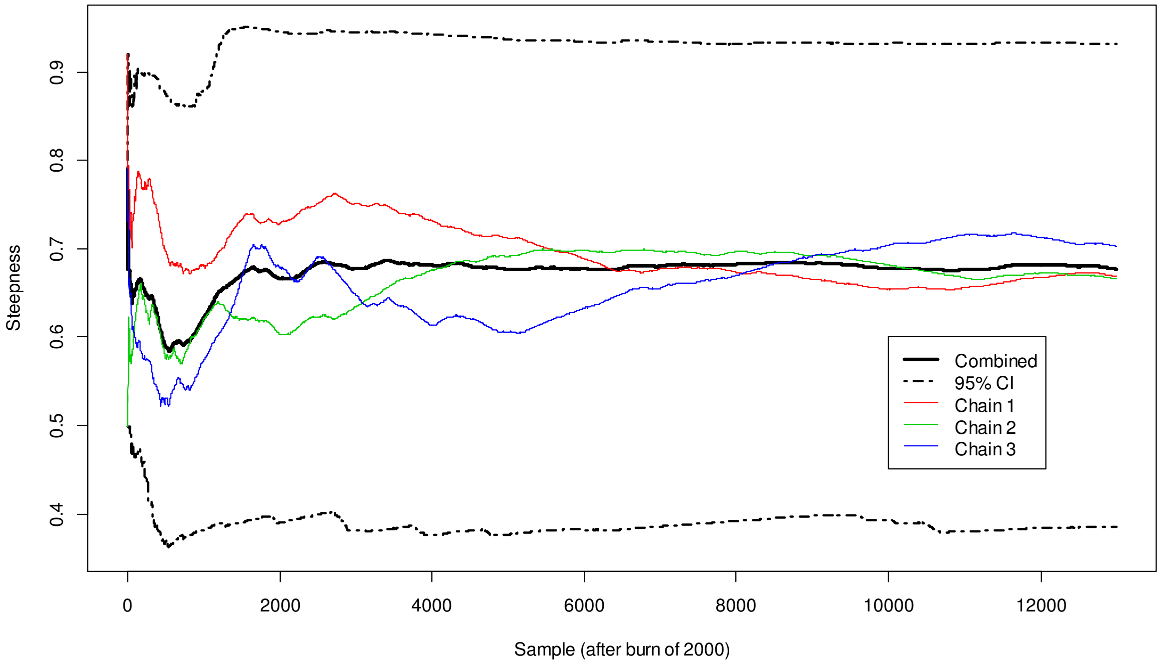


**Figure B5: The median steepness for the three chains in the Beverton-Holt MCMC run and the median and 95% CI for the combined distribution across the chains (as a function of the cumulative number of retained samples after the burn-in of 2000).**

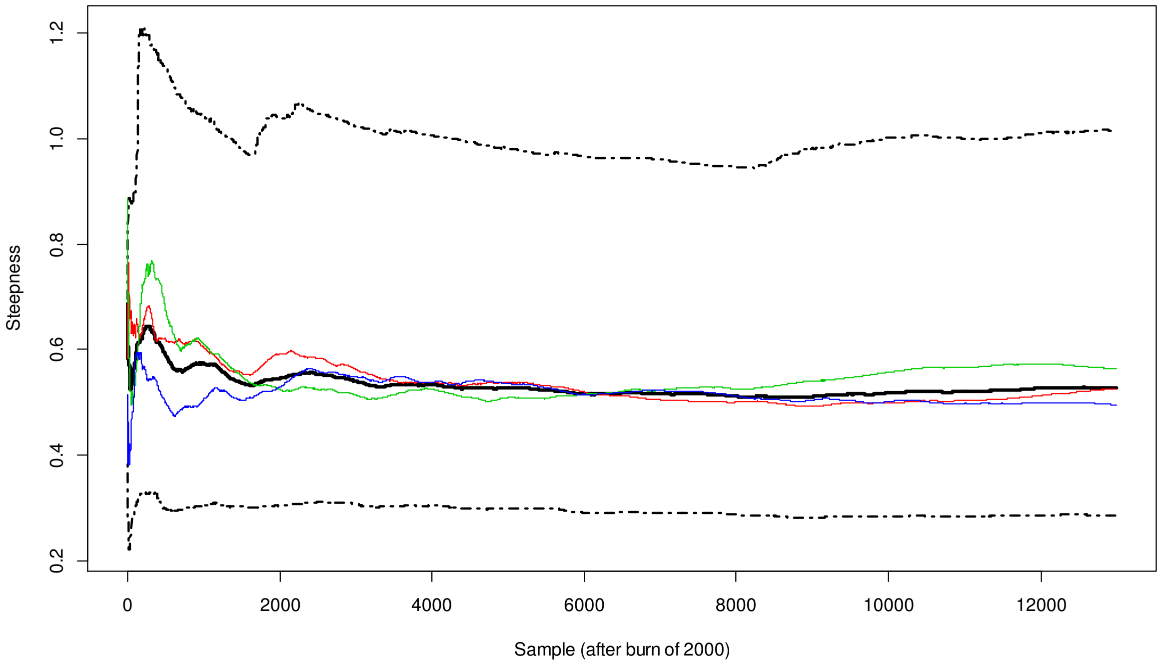


**Figure B6: The median steepness for the three chains in the Ricker MCMC run and the median and 95% CI for the combined distribution across the chains (as a function of the cumulative number of retained samples after the burn-in of 2000).**

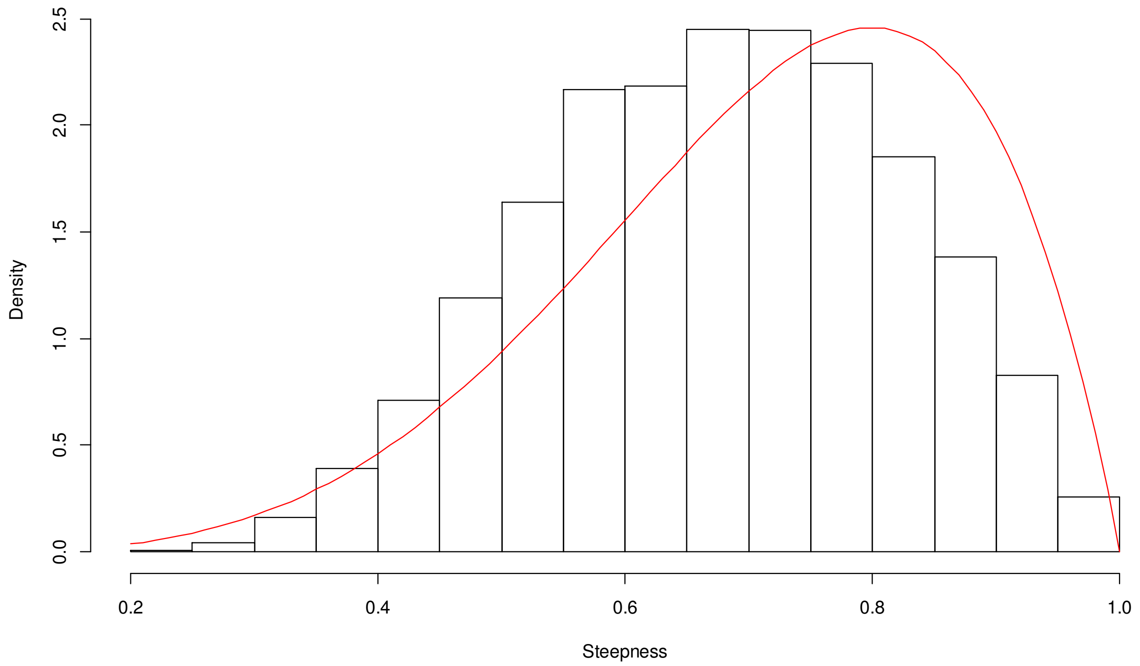


**Figure B7: The prior (red line) and posterior (histogram) for steepness from the Beverton-Holt MCMC run.**

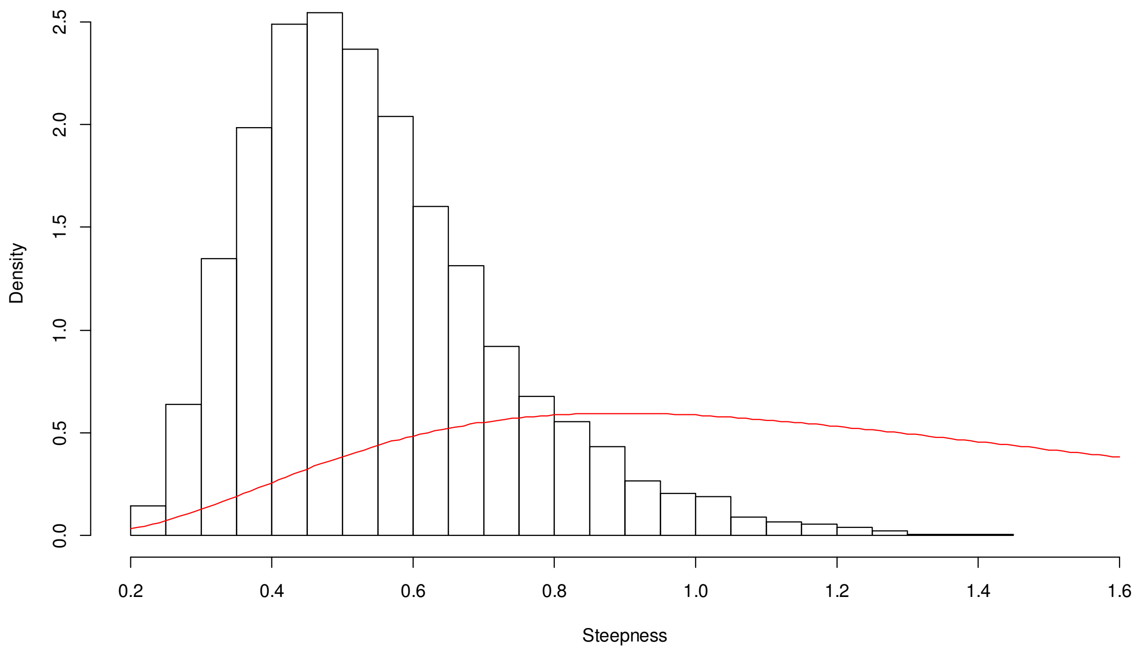


**Figure B8: The prior (red line) and posterior (histogram) for steepness from the Ricker MCMC run.**