An Initial Illustrative Example of the Application of MSE to Gulf Menhaden to Address Issues related to MSC Certification of Ecosystem-related Reference Points

Doug S. Butterworth and Rebecca A. Rademeyer

January 2019

Summary

An illustration of the use of an MSE process to assist decide a future management approach to the Gulf menhaden fishery is set out. The intent is for such an approach to provide a very low probability that menhaden abundance would fall below its historically lowest level, which should safeguard against adverse ecosystem impacts arising from the fishery. MSC certification will also require adoption of some management rule that does ensure that catches are reduced in response if the resource abundance drops too low. Under the BAM Base assessment model and maintenance of landings levels in the range that has occurred since 2000, it seems that no further regulations are necessary. However, the possibility of other dynamics (e.g. a period of poor recruitment) needs to be taken into account, and it is shown that in these circumstances application of a simple rule to control catches would result in abundance not falling as low as would otherwise be the case. Issues that would need to be considered in taking this approach further are discussed.

Introduction

This document seeks to show how the simulation-testing process of Management Strategy Evaluation (MSE) might be used to address two major issues which are soon to arise in considering refinements to the current management approach for the Gulf menhaden fishery.

The first is MSC certification. The MSC will insist (as their rules require) on some formally adopted "Harvest Control Rule" being used to set catches. This does NOT, however, mean that such a rule need apply ALL the time. The MSC's primary motivation for this requirement is: "if something is going wrong with the resource, there must be already-agreed measures in place to correct this". The current situation of the absence of such catch limits could still apply **provided** that this would be overridden if evidence of problems with the resource became available, most likely from the two primary abundance indices used for the recent SEDAR assessment. This document addresses the matter of how analyses would need to be carried out to provide the necessary assurances (*inter alia* to the MSC) that such an "override" rule would indeed achieve the desired effect.

The second is a move to include ecosystem considerations in a management system for the resource. The rationale suggested here is that this consideration can be combined in solution-terms with that advocated above. There is no strong evidence that this menhaden fishery has in the past led to deleterious ecosystem effects. Presumably therefore, if abundance can be maintained above its lowest level historically, ecosystem problems should not arise. If, then, the process above is pursued along the lines of developing a management rule for menhaden that indicates a very low probability of abundance falling below its historically lowest level, conceivably "two birds are killed with the same stone".

NB: Please note that this document is intended to be illustrative, not comprehensive. The latter would likely be more appropriate later in the process.

Methods

The Operating Model (OM) taken forward here to reflect the dynamics of the Gulf menhaden population mimics the BAM Base Model developed for the assessment of this resource (SEDAR, 2018).

The projections

Key aspects of the 20-year projections conducted are as follows, with full details (including some exceptions to the broad statements made below) set out in Appendix A.

- Unless otherwise specified, future dynamics are the same as for the BAM Base Model assessment.
- Future annual landings are drawn at random, with replacement, from the 2000-2017 values. The landings in 2018 are taken to be 525 635 mt. A maximum full fishing mortality (Fmax) is imposed to avoid unrealistic values, i.e. instances where the low size of the resource makes it unlikely the future intended catch could be taken, so that this is overridden by a value corresponding to Fmax.
- A hockey-stick is assumed for the stock-recruitment curve, with the break taken as SSB=1.8×10^6 (in billions of eggs) – see Figure 1.
- Future recruitment residuals are drawn at random, with replacement, from the 1978-2017 model estimated residuals.
- Future survey results are computed assuming log-normal observation error, with standard deviation computed from past (2005+) model estimated error. The selectivity and catchability values are taken as estimated for the BAM Base Model.

Robustness tests

Only one OM is used here (the BAM Base Model) and the robustness tests differ from the Base Case in the projections only.

- Three values of Fmax have been used (1.8, 2.5 and 4.0).
- Five (2020-2024) or ten (2020-2029) years of bad recruitments are assumed, for which the recruitment that would otherwise have been generated is halved.

Importantly this is NOT intended as a comprehensive analysis (which would need to include many more such tests). The purpose of this test here is purely illustrative – to show that there are instances where application of the Management Rule (see below) to reduce catches is needed to prevent deleterious further depletion of the resource when it is at lower levels of abundance than customary.

The Management Rule

The management rule is empirical. It only overrides and reduces a landing drawn from the historical set if the value of a combined abundance index falls below a threshold level specified for that index. If the threshold is breached, a TAC is set proportional to the value of the index. The index suggested is a weighted average of the gill net and seine indices. Figure 2 illustrates the rule for a particular choice of tuning parameter values, and also plots historical values of the combined index. Details are given in Appendix B.

Results and Discussion

Figure 3 plots historical estimates and 20-year projections for three possible trajectories for landings, full F, SSB, recruitment, the gill net and seine indices and the combined index for the Base Case with Fmax=1.8 and projecting without the management rule.

Figure 4 compares the projection results for the Base Case (Fmax=1.8) with and without the management rule by using projection probability envelopes, and by showing medians with 90%iles. In addition, a pie chart is shown to illustrate the levels of false positives and false negatives. The first is where the management rule is brought into play unnecessarily (observation error has led to the resource index giving too negative an impression of the true state of the resource which is actually not any problem); the second is the converse.

The projections in Figure 4 seem reasonable in that they broadly mimic the post-millennium values evident in the fishery. Importantly, even without the management rule, annual egg production (at the lower 5%ile) never falls below the historical lowest level. Hence essentially no further “rules” seem needed to ensure respect of the reference point suggested.

However, one cannot be sure that resource dynamics in the future will continue as in the recent past, or indeed that the BAM Base Model reflects the resource’s situation exactly. Hence it is important to repeat these computations for other plausible scenarios, to check whether a management rule would be needed for them.

To this end, Figure 5 is an illustrative (though not necessarily realistic) example intended to assist understand this point. It compares the projection results for the robustness test with 10 years of bad recruitment (Fmax=4.0) first without and then with the management rule. Here the poor recruitments lead to a reduction in the abundance of the resource and hence in its egg production. Without the management rule, there is a high probability that the resource is virtually wiped out; with the rule, although the resource is still reduced in size, the reduction in catch is sufficient to secure recovery in due course. (More technical details about this, including an explanation of the role of the Fmax value, may be found in Appendix C.)

Next steps

The intent of this document is purely illustrative. Although it indicates that if the dynamics of the BAM Base Model continue to apply, and catches remain within their range over the 2000-2017 period, there is only a small probability that resource abundance fall below its lowest level historically, this is not sufficient to justify continuation of management in the absence of further potential restrictions. A number of further aspects would first need to be investigated to provide the requisite advice. A non-exhaustive list of such aspects is provided below, with examples given under a number of sub-heads.

<u>Different OMs</u>

Evaluations thus far consider the actual historical dynamics of the resource and fishery to be reflected exactly by the BAM Base Model. These evaluations need to be repeated for a number of plausible alternatives which could also reflect the actual situation. Discussion amongst an appropriate group is likely the best approach to select these alternatives, which would probably for the most part be drawn from sensitivities examined in the assessment (SEDAR, 2018).

Different projection assumptions

For the same reasons as given immediately above, consideration of alternative assumptions about the future would also be needed, for example deviations from the proportionality assumptions made for the relationships between the survey indices and the underlying associated component of the resource abundance. Another important factor to consider would be alternatives to the hockey-stick stock-recruitment model used to generate future recruitments.

Interpretation of robustness test results

The Base Model and associated projection specifications reflect the current best (most "plausible") perception of menhaden dynamics. Alternative "robustness" tests, in which either or both of these are varied in some way, are consequently considered to be less "plausible" as reflections of the underlying reality. Hence whatever probability is decided to be the maximum acceptable for falling below the lowest historical abundance for the Base Model, presumably some slightly higher probability should be considered acceptable for a robustness test. But how much such increase is acceptable? Sensibly this would be larger if the test is considered less plausible as a reflection of reality, but in what relationship? In the extreme, what potential tests should be considered so implausible that they can be ignored? These are not straightforward issues, and thus will require some discussions and ideally pre-specification by an appropriate group.

Furthermore, if for example recruitment drops in future, some abundance drop is inevitable, but then how (if at all) should this be taken into account in assessing whether a criterion of abundance not falling below its lowest level historically has been met? For example, often a "dynamic B0" concept is used for such circumstances, where abundance is scaled to the abundance that would have been present had there never been a fishery.

Alternative Management rules

Trade-off considerations always come into play when selecting management rules. Figure 6 illustrates a typical range available, within which different choices are possible. If a threshold is crossed, in principle any option between closing the fishery to continuing as normal is possible, and the Figure shows the respective smallest and largest abundance declines which will result in the short-to-medium term given a 10-year period of poor recruitments, together with results for the rule illustrated in this document. Clearly one can obtain results differently situated between these two extremes by amending the rule.

Other possibilities that merit attention include:

- Different relative weightings of the two abundance indices in computing the composite index.
- Including a maximum "cap" on annual landings.
- Restricting the extent to which a catch limit (if imposed) can change from year-to-year.

Process

Decisions on issues raised above are best made by some small group with representatives from the various stakeholders involved in Gulf menhaden assessment and management, in particular also as a way to ensure improved broad buy-in to the outcome from the process. Some discussion of how such a group might best be appointed and operate would be desirable.

Reference

SEDAR. 2018. SEDAR 63 – Gulf Menhaden Stock Assessment Report. SEDAR, North Charleston SC. 352 pp. available online at: http://sedarweb.org/sedar-63

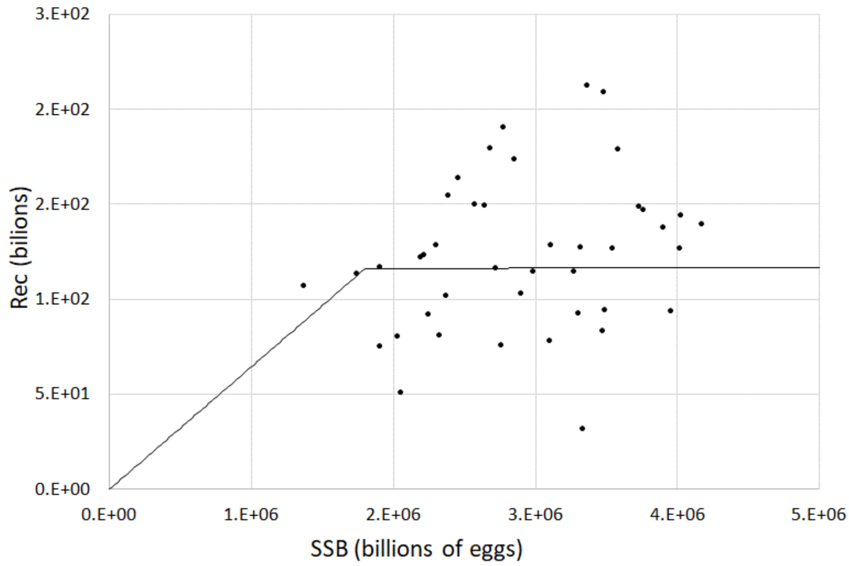
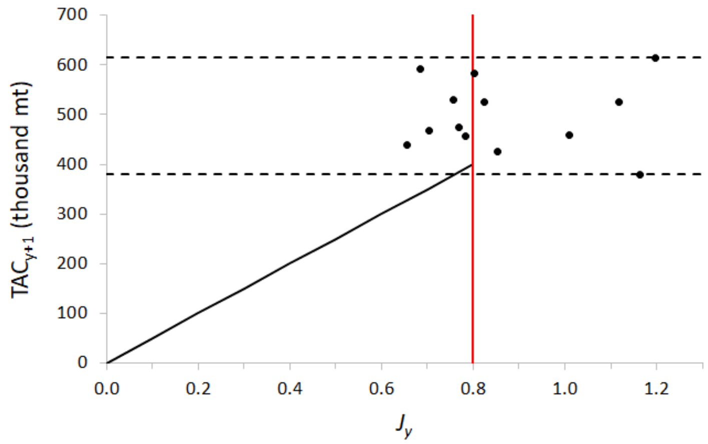


Figure 1: Hockey-stick stock recruitment curve for Gulf Menhaden which is used to compute projected recruitment. The data points are those estimated in the BAM Base Model.



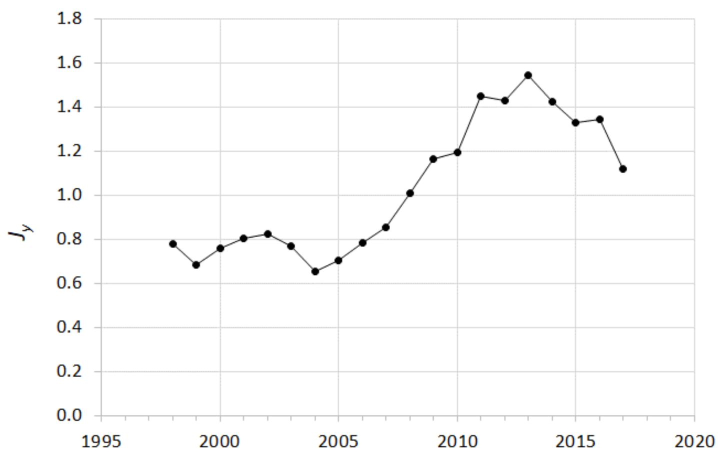


Figure 2: Top plot: Illustration of the management rule for set tuning values considered in the example for which results are reported. The horizontal dash lines show the 2000-2017 minimum and maximum landing values. The historical (1999-2017) J_y vs. TAC_{y+1} are shown as black dots. Bottom plot: Historical combined index J_y values.

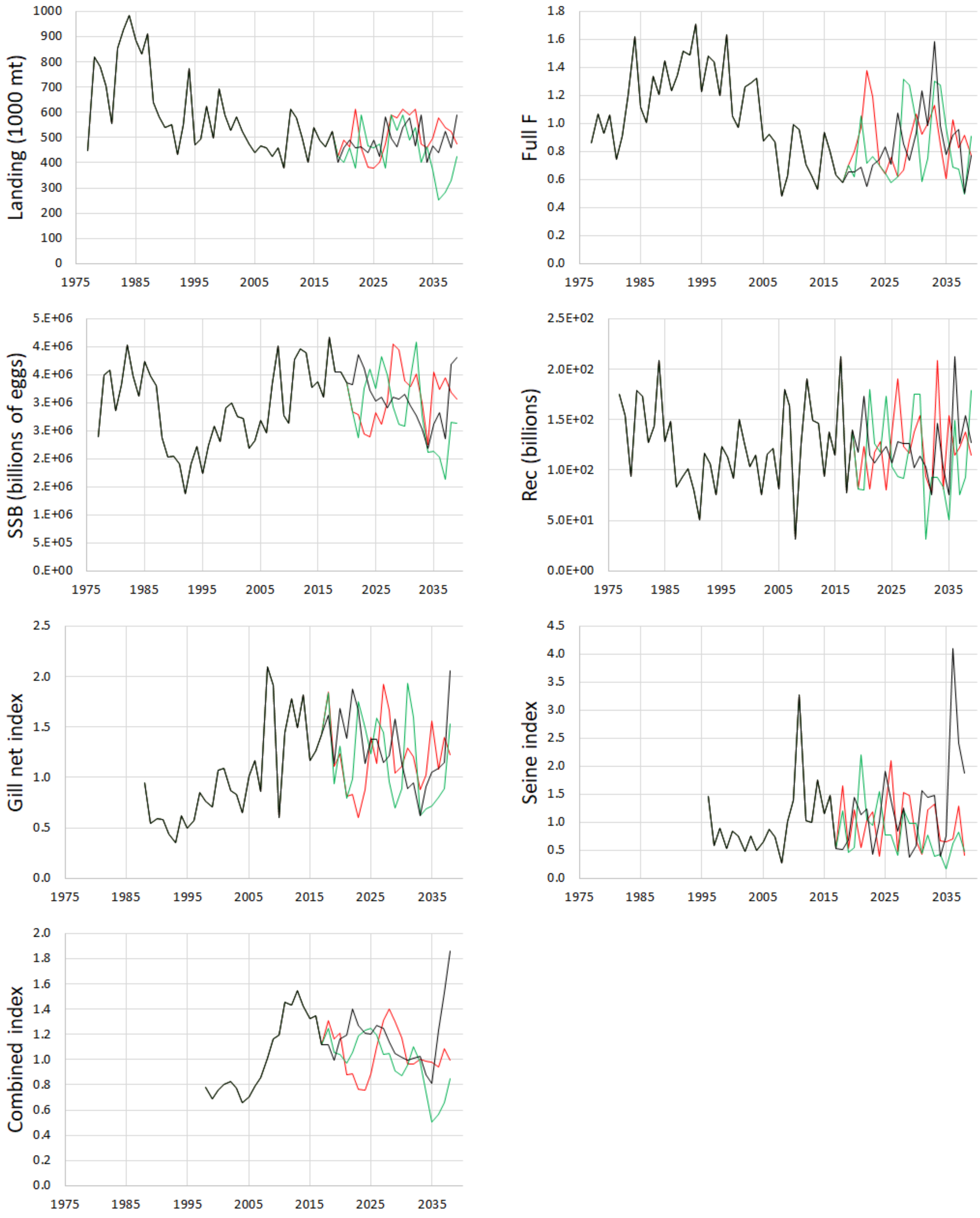


Figure 3: "Worm" plots showing historical estimates and 20-year projections for three individual trajectories for landings, full F, SSB, recruitment, the gill net and seine indices and the combined index for the **Base Case** with Fmax=1.8, **without** the management rule.

Figure 4: Historical estimates and projected 20-year median and 90%iles for a series of quantities for the **Base Case** with Fmax=1.8, without (red lines) and with (black lines) the management rule.

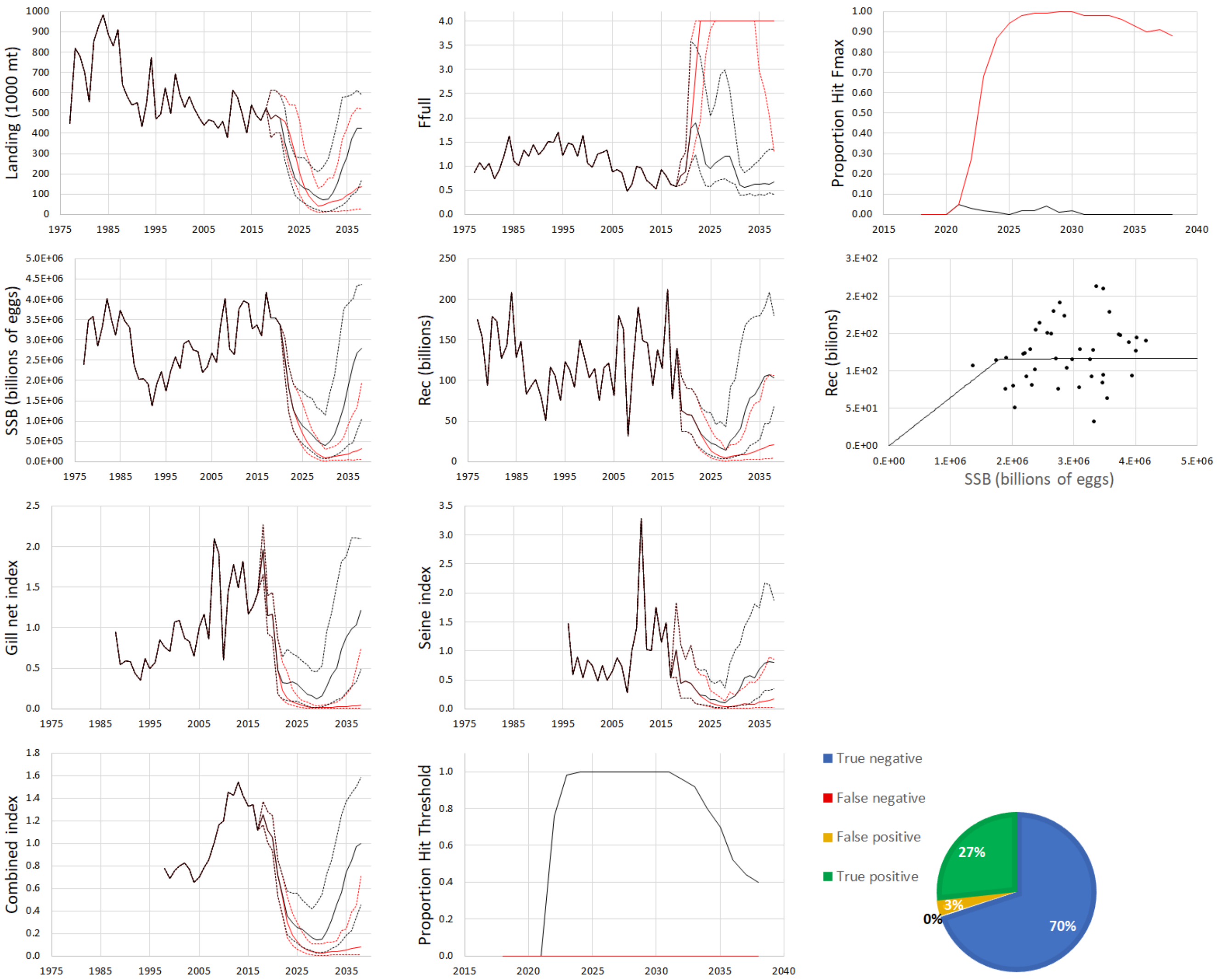


Figure 5: Historical estimates and projected 20-year median and 90%iles for a series of quantities for the **robustness test** with **10 years bad recruitment** with **Fmax=4.0**, without (red lines) and with (black lines) the management rule.

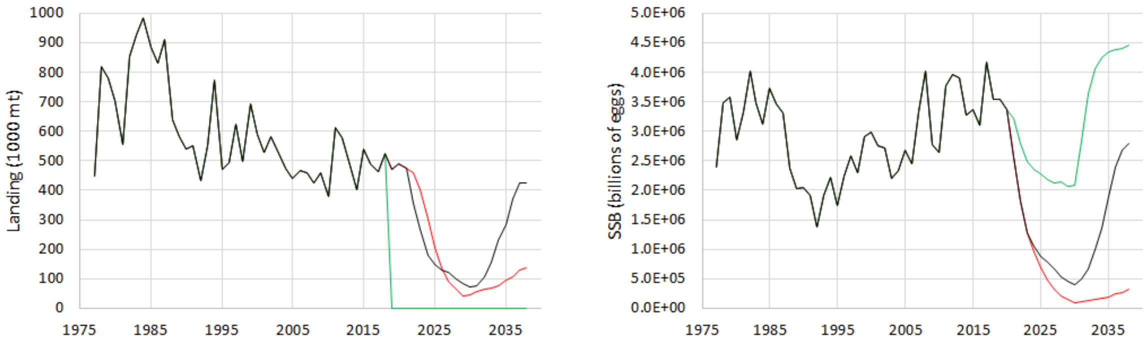


Figure 6: Historical estimates and projected 20-year medians for landings and SSB for the **robustness test** with 10 years bad recruitment with Fmax=4.0, with zero catch from 2019 onwards (green lines), and without (red lines) and with (black lines) the management rule.

Appendix A – Projection methodology details

Projections into the future under a specific management rule are performed using the following steps.

Step 1: Begin-year numbers at age

The components of the numbers-at-age vector at the start of 2018 ($N_{2018,a}$: *a* =1,…, *m* – where *m* is a plus-group) are obtained from the MLEs for an assessment of the resource. The assessment used here is the BAM Base model.

Step 2: Annual landings

For 2018, $L_{2018} = 525\ 635 \text{ mt.}$ (A.1)

From 2019 onwards:

L_y is drawn at random, with replacement, from the observed 2000-2017 landings.

If the combined abundance index (see equation B2 of the main text) for year *y*-1 is below the threshold value, then a TAC applies to year *y* is computed using the management rule (see main text and Appendix B).

Step 3: Landings-at-age (by number)

The $L_{y,a}$ values are obtained under the assumption that the commercial selectivity function (S_a) estimated for the most recent period in the BAM Base Model (1996+) continues in the future. The full fishing mortality F_y is solved iteratively to achieve the annual landing by mass:

$$L_y = \sum_{a=1}^{m} w_a^{mid} N_{y,a} S_a F_y \left(1 - e^{-Z_{y,a}}\right) / Z_{y,a} \qquad \text{(A.2)}$$

where

w_a^{mid} is the time invariant weight-at-age in the middle of the year,

$N_{y,a}$ is the number-at-age vector for age *a* at the start of year *y* (with *m* the plus group),

and

$Z_{y,a} = F_y S_a + M_a$ is the total mortality-at-age vector for age *a* and year *y*.

M_a is the natural mortality-at-age *a* (input).

The numbers-at-age can then be computed for the beginning of the following year (*y*+1):

$$N_{y+1,1} = R_{y+1} \qquad \text{(A.3)}$$

$$N_{y+1,a+1} = N_{y,a} e^{-Z_{y,a}} \quad \text{for } 1 \le a \le m-2 \qquad \text{(A.4)}$$

$$N_{y+1,m} = N_{y,m-1} e^{-Z_{y,m-1}} + N_{y,m} e^{-Z_{y,m}} \qquad \text{(A.5)}$$

If the intended landing is such that the apical fishing mortality (that at the age at which selectivity is 1) exceeds Fmax, then the landings are instead limited to those corresponding to Fmax. The initial selection for Fmax was based on the output from the BAM Base Model assessment, which indicated 1.8 as a round number slightly in excess of the maximum that had occurred historically.

Step 4: Recruitment

Expected values (in log space) for future recruitments (R_y) are provided by a hockey-stick stock-recruitment relationship:

$$R_y = \begin{cases} R & if\ SSB_y \geq SSB_{threshold} \\ \frac{R}{SSB_{threshold}} SSB_y & if\ SSB_y < SSB_{threshold} \end{cases} \tag{A.6}$$

where

R is the geometric average of the model estimated past (1977-2017) values

$SSB_{threshold}$ is a fixed value (180 000 billion eggs produced),

and

$$SSB_y = \sum_{a=2}^{m} f_a\, N_{y,a} \tag{A.7}$$

with

$f_a = \rho_a mat_a fec_a$ the reproductive output of a female fish of age *a*,

ρ_a is the proportion of female at age *a*,

mat_a is the proportion mature at age *a*, and

fec_a is the fecundity at age *a*.

When projecting, error is added to this expected value, so that for simulation replicate *s*, if

$S = \{\varepsilon_y = lnR_y - lnR: y = 1977, \ldots, 2017\}$, then when projecting:

$R_y^s = Re^{\varepsilon^*}$

where ε^* is drawn at random with replacement from the set I of ε_y values

Although the Recruitment vs Eggs produced plot from the BAM Base Model assessment shows no obvious relationship between the two, clearly there must eventually be some reduction in the number of recruits to be expected as egg production falls. We have taken the fairly standard approach here of assuming a hockey stick relationship whether the hinge-point occurs at the lowest historical annual egg production estimated, though for robustness and precaution a slightly higher value of 180 000 billion eggs was chosen so as to avoid undue influence from the lowest two historical values.

Step 5:

The projected values for numbers-at-age are used to generate values of the abundance indices I_{y+1}^i (in terms of numbers), and similarly for following years. Indices of abundance in future years will not be exactly proportional to true abundance, as they are subject to observation error. Log-normal observation error is therefore added to the expected value of the abundance index in question (in log space), i.e.:

$$I_y^i = q^i B_y^i e^{\varepsilon_y^i} \tag{A.8}$$

with

$$\varepsilon_y^i \text{ from } N\left(0, \left(\sigma^i\right)^2\right) \tag{A.9}$$

where

B_y^i is the abundance available to and indexed by the survey:

$$B_y^i = \sum_{a=1}^{m} S_a^i N_{y,a} e^{-Z_{y,a} T^i/12} \tag{A.10}$$

T^i is the timing of the survey (in month) ($T^i = 6$ for the gill net index and 3 for the seine index).

The survey selectivities are assumed to remain unchanged. The catchabilities are taken to be those estimated in the OM (the BAM Base Model assessment).

The residual standard deviations σ^i are estimated from the model fit. Since residuals seem to have increased in recent years, the residuals from 2005 onwards have been used for their computation:

$$\sigma^i = \sqrt{\frac{1}{\sum_{y=2005}^{2017} 1} \sum_{y=2005}^{2017} \left(lnI_y^i - ln\hat{I}_y^i\right)^2} \tag{A.11}$$

where I_y^i is the observed index value in year *y* for survey *i* and $\hat{I}_y^i$ is the corresponding model estimated value this yields σ^i =0.11 for the gill net index and 0.41 for the seine index.

<u>Step 6</u>:

Steps 1-5 are repeated for each future year in turn for as long a period as desired.

Appendix B – The Management Rule

The management rule is empirical. It only overrides and reduces a landing drawn from the historical set if the value of a combined abundance index (see below) falls below a threshold level specified for that index. The basis for the associated computations is set out below:

If $J_y < J_{threshold}$:

$$TAC_{y+1} = \gamma J_y \tag{B.1}$$

where

TAC_y is the catch limit that applies for year *y*,

$J_{threshold}$ and γ are tuning values (in this example, $J_{threshold} = 0.8$ and $\gamma = 500\,000$); Figure 2 illustrates the rule for these choices for these tuning values, and

J_y is a measure of the immediate past level in the abundance indices that are available to use for calculations for year *y*:

$$J_y = \frac{1}{p}\sum_{y'=y-p+1}^{y}\left[\left(w_{gill}\frac{I_{y'}^{gill}}{I_{2017}^{gill}} + w_{seine}\frac{I_{y'}^{seine}}{I_{2017}^{seine}}\right)\Big/\left(w_{gill} + w_{seine}\right)\right] \tag{B.2}$$

with

I_y^{gill} and I_y^{seine} being the observed gill net and seine indices, respectively, in year *y*,

w_{gill} and w_{seine} being the weights given to each index (here $w_{gill} = 4$ and $w_{seine} = 1$, which corresponds roughly to inverse variance weighting given the standard deviations of the residuals in the BAM Base Model fit),

and *p* being a tuning parameter (here *p*=3); this parameter is used to smooth away some of the noise in the index by averaging over a few years rather than consider only the most recent year.

Note the assumption has been made that when a TAC is set in year *y* for year *y*+1, values of these abundance indices will be available for the current year *y*.

Appendix C

In the interests of simplicity, the main text (Figure 5) includes results for only one of the robustness tests specified, viz. 10 years of poor recruitment and an Fmax value increased from 1.8 for the Base Case to 4.0. These choices may be somewhat extreme, but are intended only to provide an illustration which shows that there are scenarios where the introduction of the management rule can improve resource performance by reducing the extent to which abundance declines.

As indicated by the plots below, which show sensitivities to the period of low recruitment and the Fmax value, the higher Fmax value is necessary to produce this effect; otherwise a lower Fmax value alone is sufficient to safeguard the menhaden from undue depletion. Even though an Fmax value as high as 4.0 has not occurred historically (see Figure 3), it is not implausible that it could occur in circumstances where there was pressure to maintain catches even though abundance had fallen – in reality Fmax at the apical age would not increase to this extent, but rather the selectivity curve would increase in width as fishing changed to cover a wider geographical area to maintain catches by taking a greater proportion of the catch at higher and lower ages.

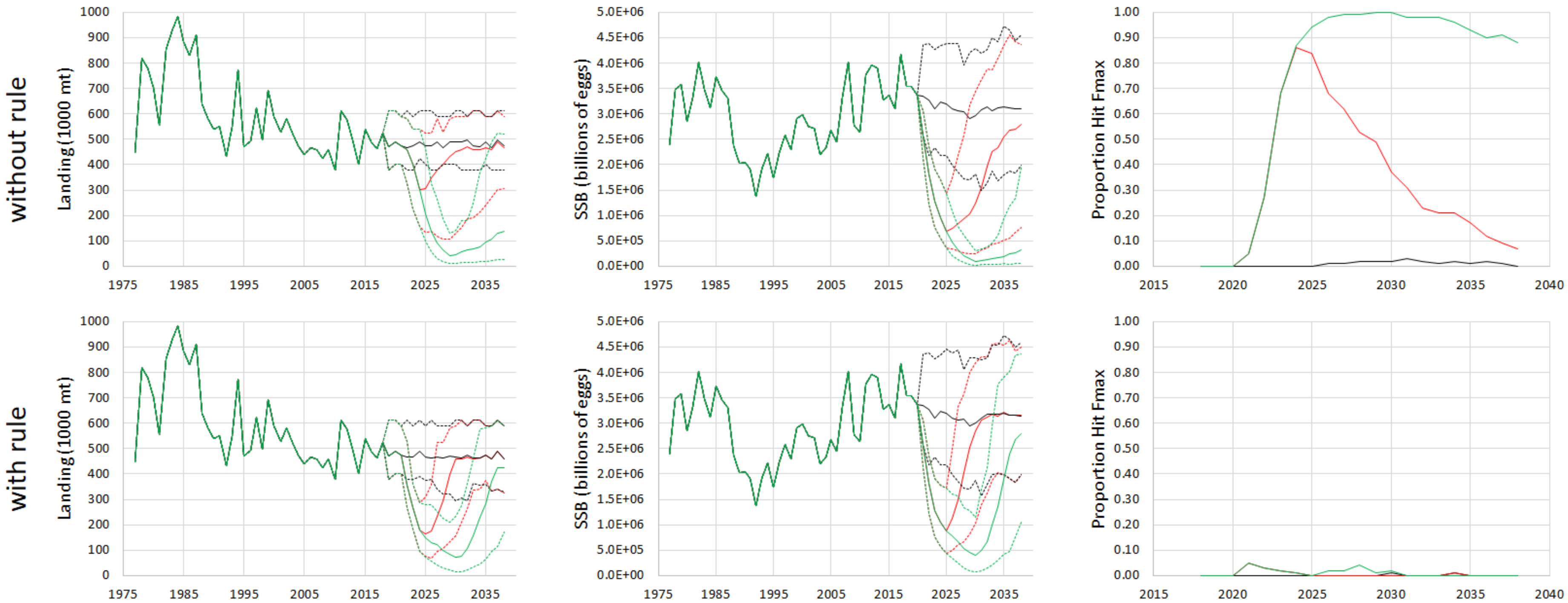


Figure C.1: Landings and SSB time trajectories (median and 90% PI) for the Base Case (black lines), and the robustness tests with 5 years (red lines) and 10 years (green lines) bad recruitments, without (top row) and with (bottom row) the management rule, all with Fmax=4.0. The proportion of time Fmax is hit each year is also shown for each case.

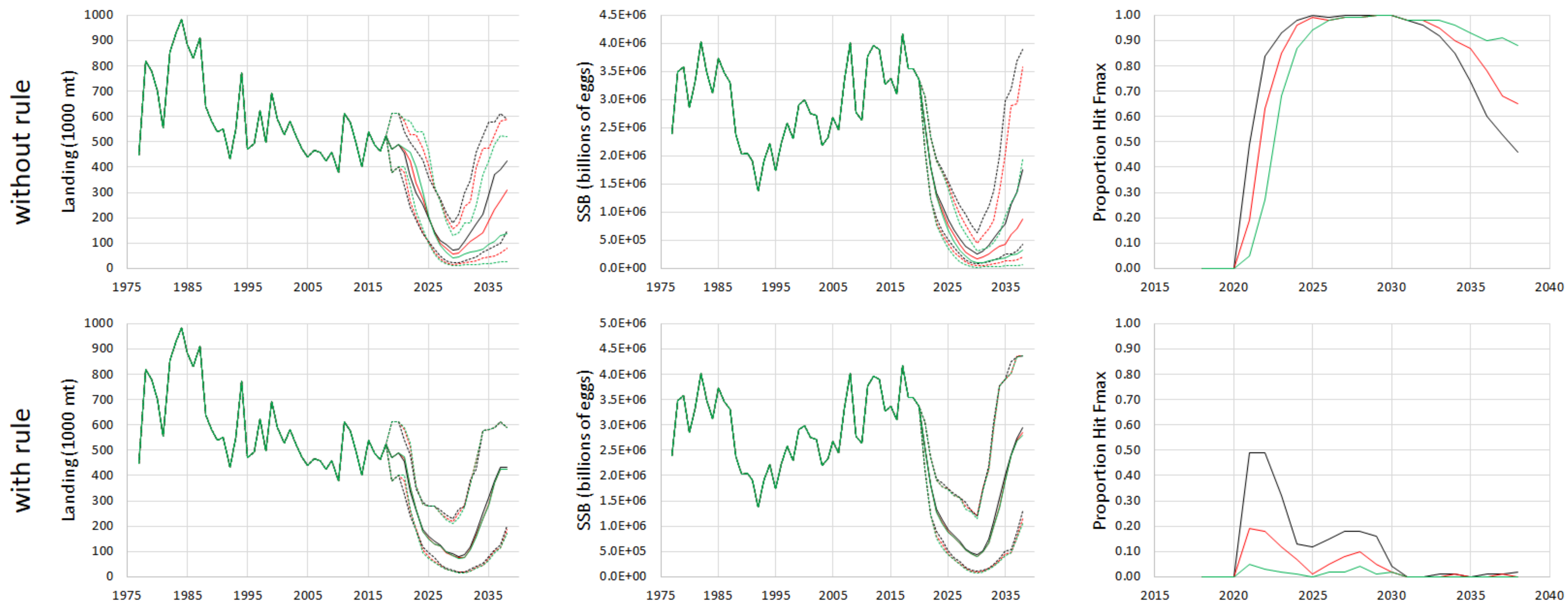


Figure C.2: Landings and SSB time trajectories (median and 90% PI) for Fmax=1.8 (black lines), Fmax=2.5 (red lines) and Fmax=4.0 (green lines) for the robustness tests with 10 years bad recruitments, without (top row) and with (bottom row) the management rule. The proportion of time Fmax is hit each year is also shown for each case.