# Updates on development of MSE analyses for Indian Ocean swordfish

Thomas BRUNEL, Iago MOSQUEIRA (WMR) [1]

Management Strategy Evaluation Task Force of the Working Party on Methods Meeting. 28-31 March 2023

A brief summary of current status and recent developments on the work for an MSE analysis for Indian ocean swordfish is presented here. An updated uncertainty grid for the OM construction, new proposal for a model free MP and the implementation of a surplus production model for a model-based MP need to be discussed by MSE task force of WPM to guide the next steps of work for this species.

## Contents



[1] Wageningen Marine Research. Haringkade 1, Postbus 68, 1976CP, IJmuiden, The Netherlands, thomas.brunel@wur.nl , iago.mosqueira@wur.nl

## Operating model development

The status of the current swordfish OM was presented to the recent WPB session (Brunel & Mosqueira, 2022). The document presented a revision of the OM grid that decrease the factors and levels considered, by identifying those not affecting the estimated variability in stock status and productivity. This exercise was partly motivated by the large uncertainty observed in the figures presenting the performance of various MPs, especially on the performance statistic used for tuning: the probability of the stock being in the green quadrant of the Kobe plot in the 2034-2039 period. Two factors in the grid were eliminated as they seem to have no influence on stock status and productivity: choice of selectivity function shape for the CPUE fleets (previously set as double normal or logistic) and the scaling factor applied to the biomass by area estimates (previously based on surface area, biomass or catch). A change in the range of steepness values was also suggested, to 0.6-0.8 from the previous 0.6-0.9.

The WPB supported the proposal to remove from the grid the parameters that are less influential on stock status and productivity, but did not endorse the proposal regarding lower steepness values. The group noted that values between 0.6 and 0.9 were used for swordfish in other oceans (e.g. at ICCAT). The updated structural uncertainty grid after WPB 2022 and the original grid are given in table 1 and 2 respectively.

**Table 1 : Reference OM structural uncertainty grid**

| Variable | Values | | |
|---|---|---|---|
| Selectivity | Double Normal | Logistic | |
| Steepness | 0.6 | 0.75 | 0.9 |
| Growth + Maturity | Slow growth, late maturity (Wang et al.,2010) | Fast growth, early maturity (Farley et al., 2016, otoliths) | |
| M | Low = 0.2 | High = 0.3 | Sex-specific Lorenzen M (Farley et al. (2016), otoliths) |
| Sigma R | 0.2 | 0.6 | |
| ESS | 2 | 20 | |
| CPUE scaling schemes | Area effect x Surface | Catch | Biomass |
| CPUEs | JPN late + EU.PRT | JPN late | TWN + EU.PRT |
| Catchability increase | 0% | 1% / year | |

**Table 2 : Proposal for a new OM structural uncertainty grid (difference highlighted in bold)**

| Variable | Values | | |
|---|---|---|---|
| Selectivity | Double Normal | | |
| Steepness | 0.6 | 0.75 | 0.90 |
| Growth + Maturity | Slow growth, late maturity (Wang et al.,2010) | Fast growth, early maturity (Farley et al., 2016, otoliths) | |

| | | | |
|---|---|---|---|
| M | Low = 0.2 | High = 0.3 | Sex-specific Lorenzen M (Farley et al. (2016), otoliths) |
| Sigma R | 0.2 | **0.4** | 0.6 |
| ESS | 2 | 20 | |
| CPUE scaling schemes | | | Biomass |
| CPUEs | JPN late + EU.PRT | JPN late | TWN + EU.PRT |
| Catchability increase | 0% | 1% / year | |

The decisions made at WPB were implemented, which resulted in a new grid containing 648 combinations, of which 175 were selected by factorial design optimization (vs 2592 and 108 respectively for the original OM). The SS3 stock assessment was run for these 175 parameter combinations, and 130 runs were ultimately considered acceptable, based on their index of abundance prediction skill, and used as a basis for the OM (vs 67 for the original OM).

The SS3 runs based on the latest available stock assessment data for the Indian ocean swordfish cover the development of the stock until the year 2018. In order to conduct simulations starting with a stock status as close as possible to the current status, the OM was projected forward over the years 2019-2022 using the IOTC catches estimates for the years 2019 to 2021, and assuming a status quo fishing mortality for 2022 ($F_{2022}$=$F_{2021}$).

The updated OM presents a slightly different distribution of stock status at the end of the assessment period from the previous OM (figure 1), with broader distribution of values, and, overall, slightly higher values. The updated OM gives the perception of a stock that was unfished at the start of the assessment period, for which exploitation started in the 1990s which caused stock size to decrease until the mid-2000s after which it stabilized (figure 2).

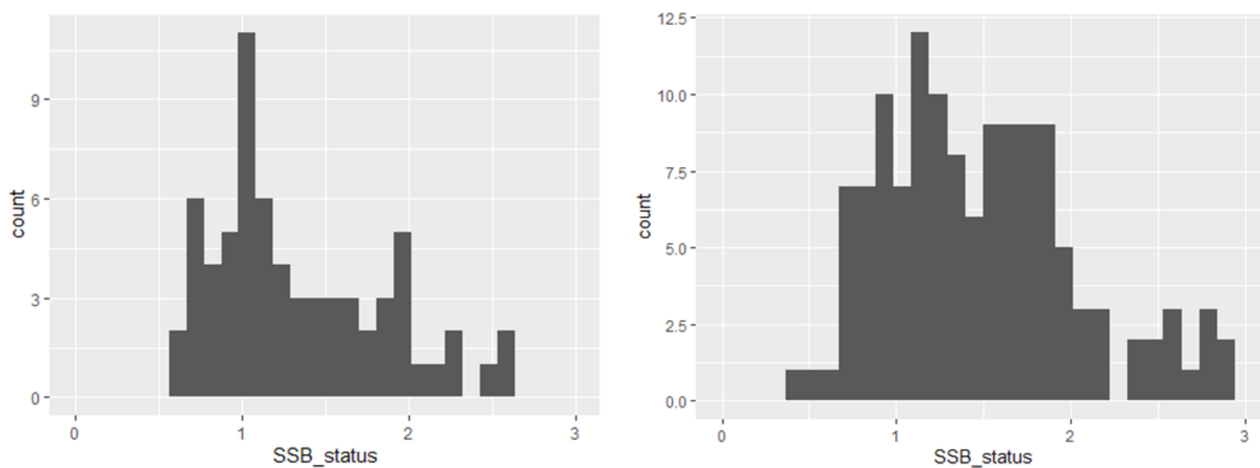


*Figure 1 : Distribution of the estimated stock status in 2018 (SB/$SB_{MSY}$) on the previous OM grid (left) and new one (right).*

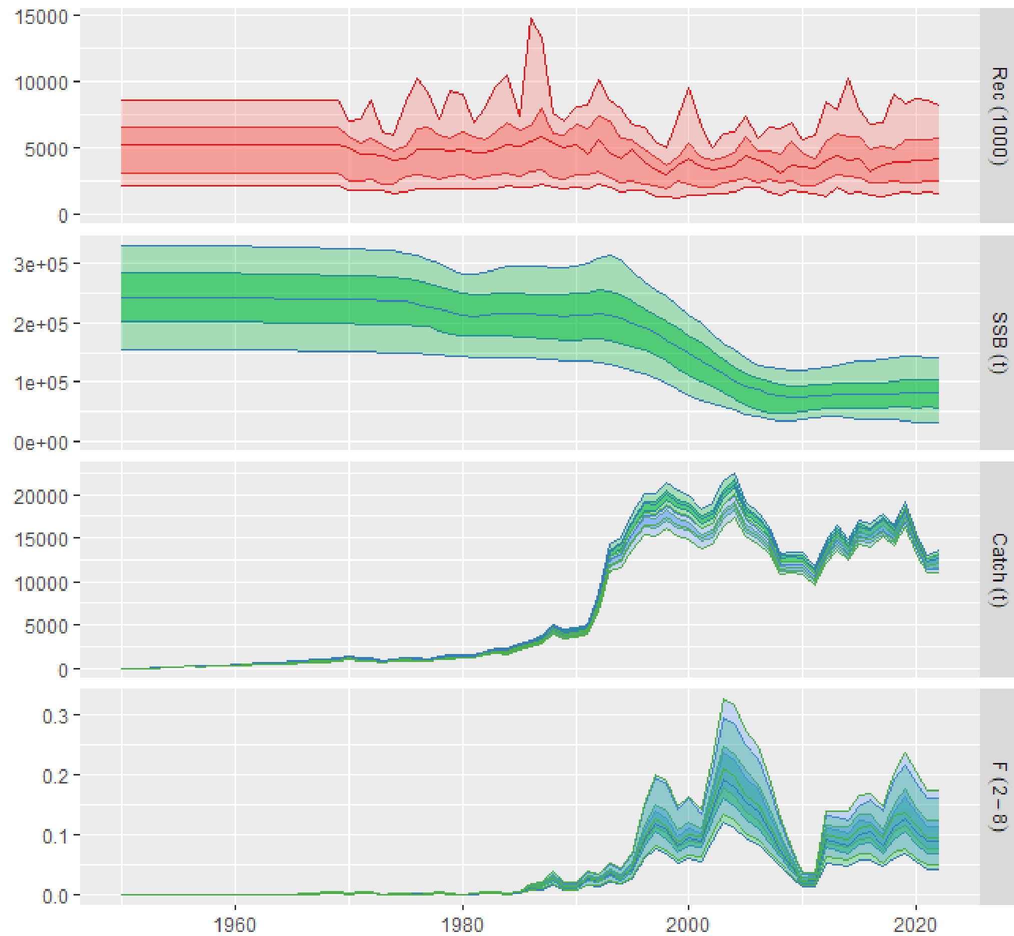


*Figure 2 : historical stock development in the updated swordfish OM*

# Management plan testing

## Model free MPs

### Relative harvest rate

*Description of the MP*

The concept of *relative* harvest rate, the ratio of total catch to a relative indicator of stock size (such as a survey index), has been proposed and tested recently as a useful driver for management on stocks for which a stock assessment is unavailable (Fischer et al. 2022). The relative harvest rate estimator is used in a hockey-stick shaped HCR, implemented as proposed by Fischer et al. (2022):

$$TAC_{y+1} = I_y \times H_{target} \times BSG \times \gamma$$

with

$$BSG = min\left(1\,;\, {}^{I_y}\!/_{I_{trigger}}\right)$$

The TAC for the coming year is defined as the product of the survey index $I_y$ and the target harvest rate value, $H_{target}$. In addition, a biomass safeguard, BSG, is applied, which reduces the target harvest rate when the index falls below an index trigger value, $I_{trigger}$.The biomass safeguard BSG essentially imposes a hockey-stick functional form on the control rule, similar to the one employed by IOTC in model-based MPs. The parameter $\gamma$ can be used to rescale the entire harvest control rule, and adjust its performance to any desired criteria.

The two control points in the HCR can be defined empirically, from the data. The proxy proposed by Fischer et al. (2022) for $I_{trigger}$ is the lowest observed stock index, multiplied by an uncertainty buffer of 1.4 in the absence of better knowledge. The reference level for harvest rate, $H_{target}$, can be defined as the mean of the past relative harvest rate values for years in which the stock was considered exploited at $F < F_{MSY}$.

*Implementation*

For the implementation of the MP to the Indian Ocean swordfish stock, the relative harvest rate was calculated using the Japanese longline CPUE (figure 3), the index for which the SS3 model appeared to have the highest predictability (lowest MASE). The trigger value for this index was set to 0.79 following the approach proposed by Fisher et al. (2022), based on the lowest observed values (0.57) multiply the 1.4 buffer.

The target harvest rate can be set empirically (as proposed by Fisher et al. 2022) by looking at the years for which the stock was in the green quadrant of the Kobe plot according to the base case stock assessment, and taking the average of this value (around 30 000 on figure 3).

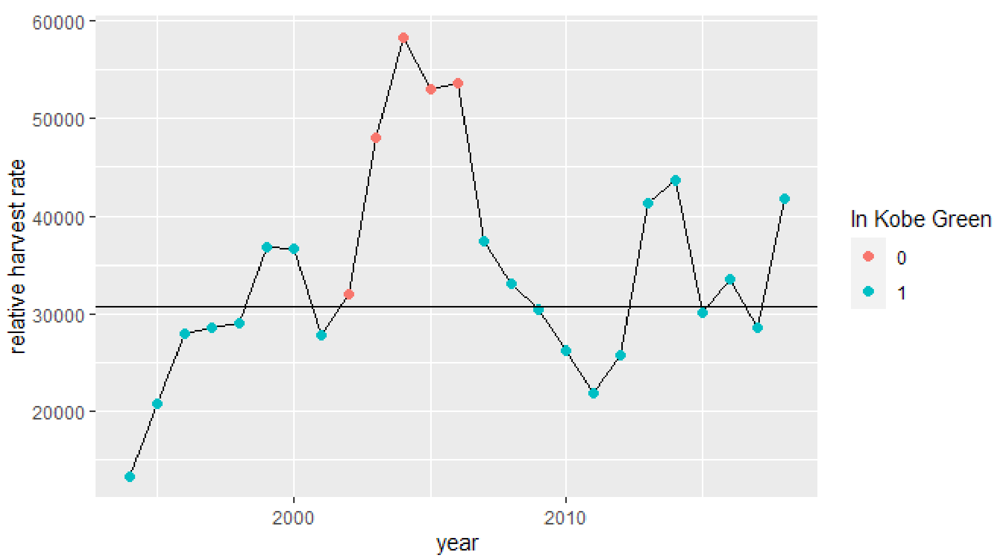


*Figure 3 : relative harvest rate for Indian Ocean swordfish (catches divided by Japanese longline CPUE). Years when the stock was in the green part of the Kobe plot are represented in red. The*

*mean relative harvest rate of these years is used as a target harvest rate in the MP (horizontal line).*

## CPUE-based rule

The CPUE rule (IOTC 2018) bases the decision on future TAC on the recent trend in a stock size index, combined with the distance between the current index value and a tunable target (Figure 4). As for the harvest rate, the Japanese longline CPUE index was used in the MP tested here. Future TAC is calculated as a proportion, $TAC_{mult}$, of the current TAC, which is defined as

$$TAC_{mult} = 1 + k_a\, Sl + k_b\, D$$

with

$$k_a = k_1 \; if \; Sl > 0 \; or \; k_a = k_2 \; if \; Sl \leq 0$$

And

$$k_b = k_3 \; if \; D > 0 \; or \; k_b = k_4 \; if \; D \leq 0$$

Where $Sl$ is the slope of the log CPUE over the last 5 years, $D$ is the difference between recent CPUE value (average over the last 3 years) and the target CPUE value, and $k_a$ and $k_b$ are parameters of the relative weight assigned to the previous two quantities (Figure 4), controlling the responsiveness of the MP. By setting, $k_1 \neq k_2$ and $k_3 \neq k_4$ the responsiveness can be different when the CPUE is decreasing (or when the CPUE is under the target) than when it is increase (or over the target).

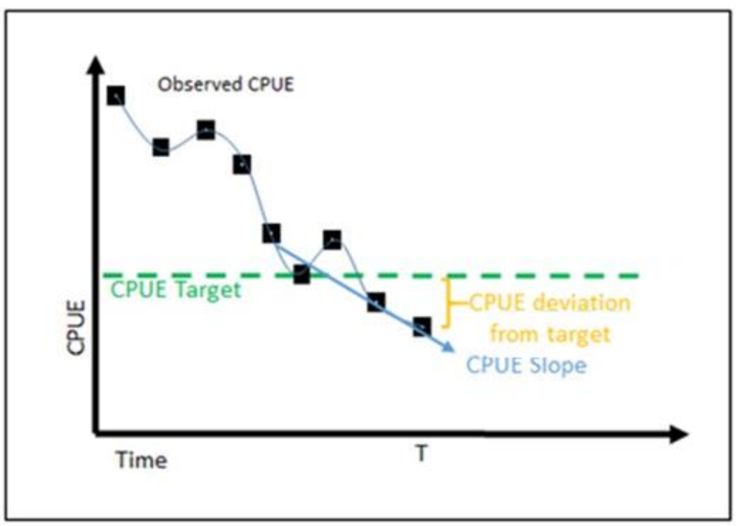


*Figure 4 : The CPUE rule is based on the recent slope in the CPUE index and the distance to the target index value.*

# Model based MP

## Stock assessment methods

The Bayesian surplus production model JABBA (Winker, 2018) has been applied in the model-based MP for swordfish. A formulation of JABBA has been adopted that attempts to speed up the model runs without a measurable degradation in fit and robustness. The production function shape is set to Schaefer, regardless of any possible assumption about the actual shape for this stock. Model estimates of depletion, which form the input to the HCR, should be less sensitive to the choice of curve than MSY-based quantities, and this choice introduces a level of safety in the estimation of productivity.

Both process and observation error have been kept at minimum levels, with the process error standard deviation limited to 0.2 in the log scale, to 0.1 for the index observation error, and a 10% CV in the error on the catch observations. An alternative setup could be attempted that concentrates on process error. Assuming that observation error is likely to lead to biases that can be corrected by the MP tuning process, a good estimate of process error could increase model robustness under certain circumstances.

<u>Data</u>

Total annual catch and relative abundance data for the years 1951 to 2018, were obtained from the 2019 swordfish stock assessment inputs. Two indices of abundance have been chosen as input to the model, based on their prediction skill on the base case stock assessment, as measured by their MASE value: JAP LL NW and TWN LL NW.

<u>Priors</u>

Loosely informative priors were defined, without any stock or species-specific information: r=LN(0.2, 0.5), dep=LN(0.9, 0.25), K=LN(log(2.5e6), 0.85). The performance of the model might be increased if some of the priors are made more informative. This could be based on the results of the base case stock assessment, which then becomes part of the MP, for r and K. In the case of initial depletion, the assumption of $B_{1950}$ = K could be explicitly incorporated.

<u>Model exploration</u>

The model setup presented above was briefly tested using a simulated dataset in which a stock was generated with swordfish life-history characteristics. The main objective of this test was to assess the robustness of the software and model formulation. A one-way-trip trajectory was forced on the virgin population to levels below $B_{MSY}$. Catch and indices were extracted and used to fit the JABBA setup under test. The overall trend appeared to be well captured (figure 5), and all runs converged.

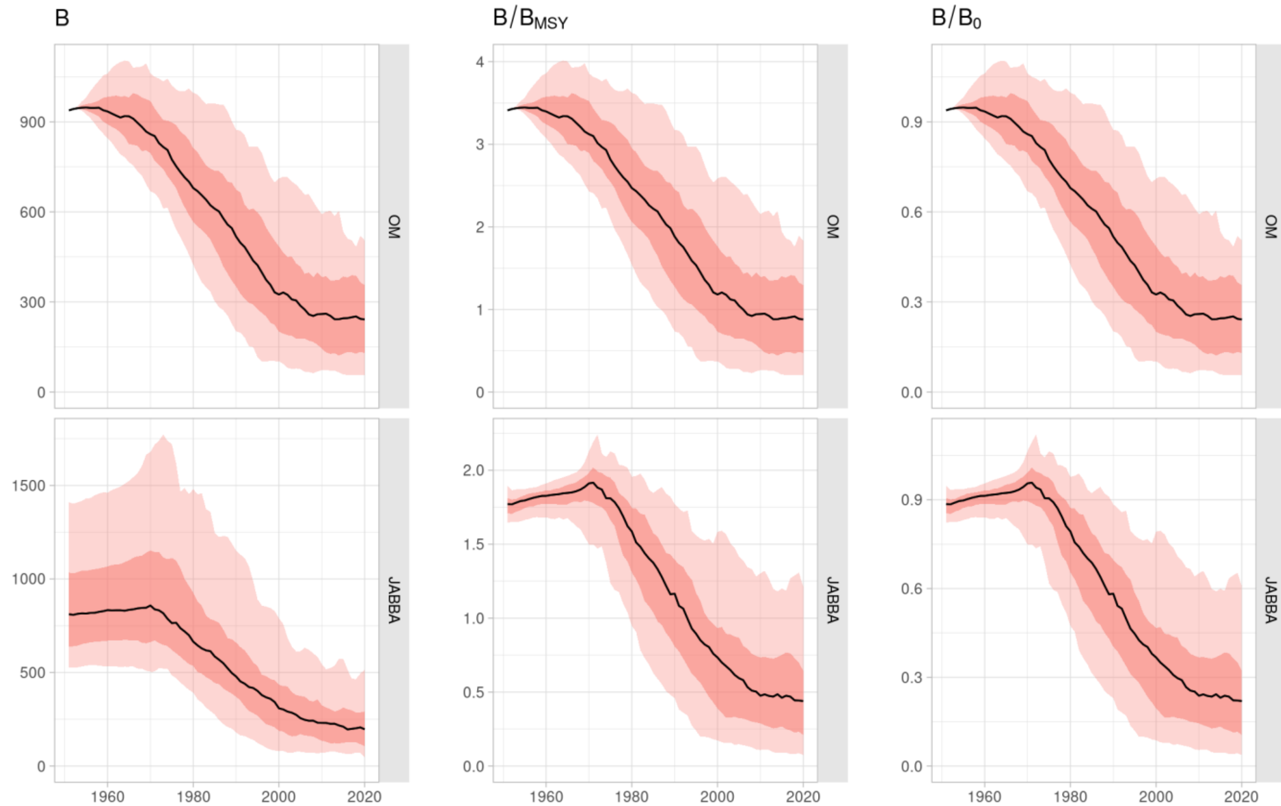


*Figure 5: JABBA model fits to a swordfish life history simulated dataset, and employing the settings employed in the IOTC SWO MP runs. Top panels present the OM value for vulnerable biomass (B) and its ratios to $B_{MSY}$ and B0, while bottom panels show the JABBA estimates for the same quantities*

## Harvest control rule

The model-based MPs (figure 6) involve two steps: 1) fitting a surplus production model to estimate current depletion rate, and 2) applying a Harvest Control Rule (HCR) to the model estimates of current depletion. The shape of the HCR (hockey-stick) is defined by three control parameters:

- CP1: minimum stock level below which no fishing (or the least possible) should take place, which is by convention set at SB/SB0= 0.10
- CP2: trigger stock level below which Catch advice should be decreased proportionally to current depletion, which is by convention set at SB/SB0= 0.40
- CP3: maximum catch that can be taken when the stock is estimated to be above the trigger level, which each approximated by tunning the MP.

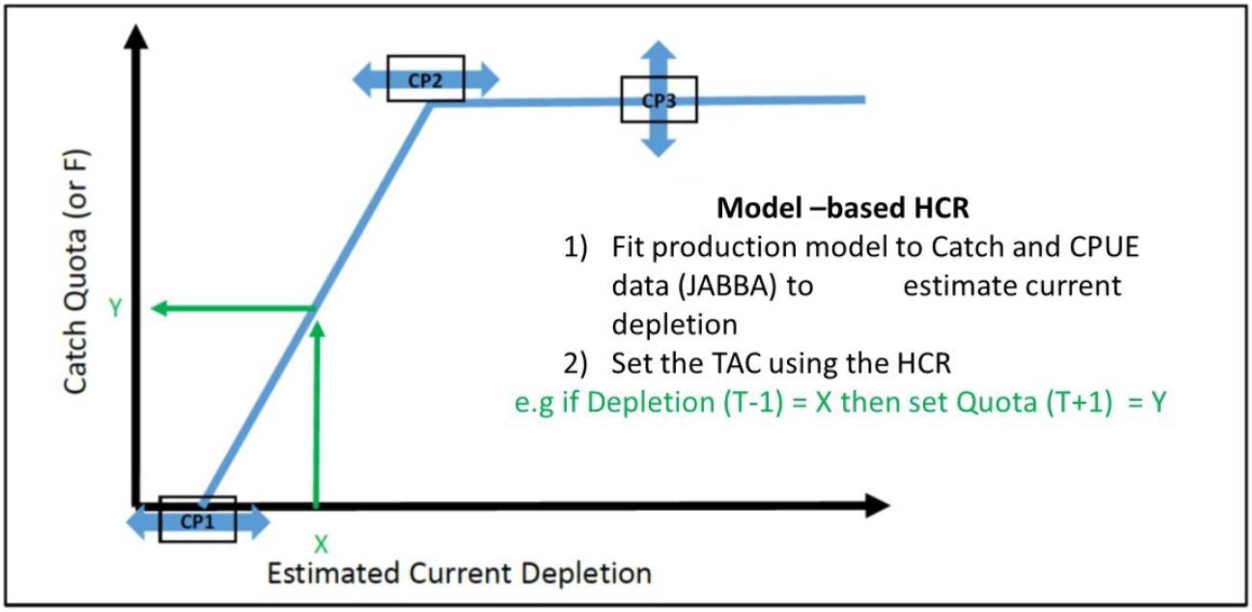


*Figure 6 : the hockey stick harvest control rule sets the future TAC based on the JABBA estimate of current depletion*.

## Results

### Exploration of the impact of the responsiveness parameters $k_{1,2,3}$ and $_4$ and CPUE target value on MP performance

The CPUE MP has multiple control parameters, and it is not straightforward to decide which parameters should have values chosen a priori and which parameters should have values obtained by tunning. In order to examine the impact of the different parameters on the MP performance, simulations were run for a grid of parameters values and key performance indicators were calculated over the period for which tunning should be done (2034-2039). The grids used were the following :

- Grid 1 (symmetric responsiveness) :
    - k1=k2 varying between 0.1 and 3 and
    - k3=k4 varying between 0.1 and 1.2
    - target = mean of historical CPUE index values (1.00)
- Grid 2 (asymmetric responsiveness) :
    - k1 varying between 0.1 and 3
    - k2 = 2 x k1
    - k3 varying between 0.1 and 1.2
    - k4 = 2 x k3
    - target = mean of historical CPUE index values (1.00)
- grid 3 (target range ):
    - target values between 0.1 and 2

  - k parameters chosen based on results from grid 1 and 2 that correspond to low and to high responsiveness of the MP to the CPUE index, and to symmetric (grid1) and asymmetric (grid2) responsiveness.

Simulations were run on a subset of the OM (iters = 50) to make the computing time acceptable.

For the grid 1, the MP performance indicators were more influenced by the responsiveness to the distance to the target CPUE (k3 and k4, figure 7) rather than to the slope (k1 and k2). The mean catch is rather stable along the isolines for p(Kobe green), showing that these two performance metrics are linked to a large extent. Along the p(Kobe green) isolines, management strategies that are more reactive to the CPUE index (higher values of k parameters) lead to higher interannual catch variability and increased biological risk (note : risk metrics based on simulations with only 50 iterations are not well estimated and are likely to be underestimates).

For the grid 2, the influence of the parameters related to the CPUE slope (k1 and k2) on MP performance is larger (figure 8), but parameters related to the distance to CPUE target remain the most influential. Along the isolines for p(Kobe green), lower reactivity scenarios seem to yield the highest catches, while catches are overall lower than for grid 1. As for grid 1, increasing responsiveness lead to increased catch variability and biological risk.

For the grid 3, increasing the CPUE target value logically increases the probability of being in the green part of the Kobe plot, and decreases biological risk (figure 9). In general, the asymmetric MPs and the faster reacting MPs appear more precautionary (higher p(Kobe green) and lower risk) than the symmetric MPs and the slow reacting MPs. A dome shape relationship is observed between CPUE target and the resulting catch, with an optimal CPUE target value depending on MP reactiveness. The slow reacting MPs lead to overall higher catches than the faster reacting one, and to less variable catches. The asymmetry in the MP reaction to the sign of the slope and difference with target CPUE lead lower catch (but less so for the slow reacting MPs), and increased catch variability for the fast reacting MP.

The options that lead to (values as close as possible to) 50% and 60% probability of being in the green part of the Kobe plot were identified (dots and triangles respectively on figure 9). Scenarios with the best performance (for both 50% and 60% p(Kobe green)) were for the slow reactivity option and the symmetric reaction, leading to the highest catches and the lowest variability. All scenarios achieving these levels of p(Kobe green) had low biological risk associated (most times lower that 5%).

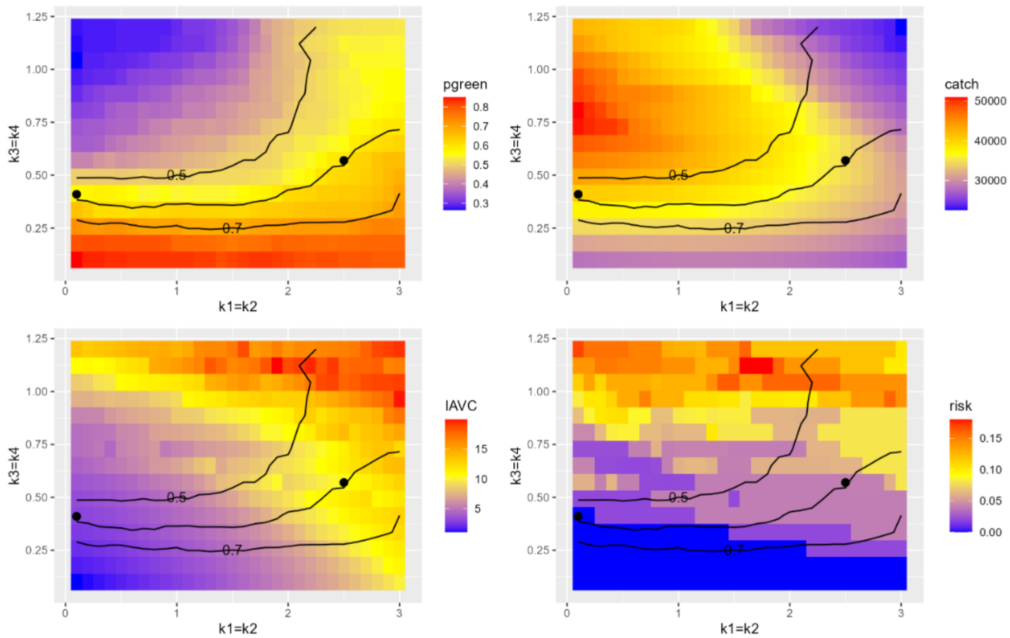


*Figure 7 : values of 4 MP performance indicators (pgreen, mean catch, interannual catch variability and risk3 computed over the period 2034-2039) for a grid of k responsiveness parameters with symmetric response for positive and negative CPUE slope and distance to target (isolines show the 50, 60 and 70% probability of being in the green zone of the Lobe plot).*

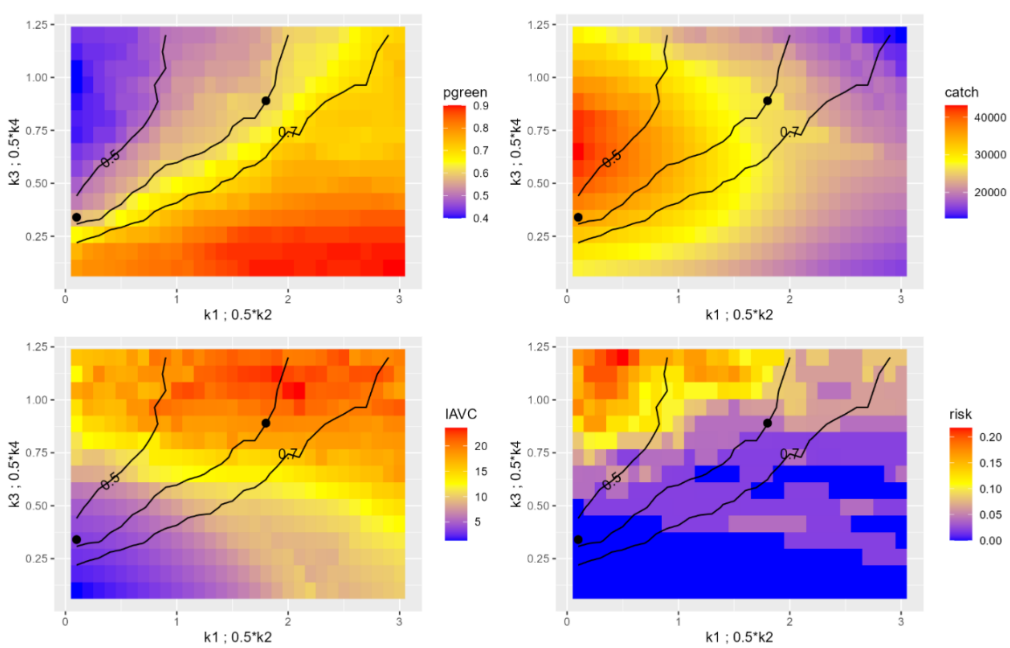


*Figure 8 : values of 4 MP performance indicators (pgreen, mean catch, interannual catch variability and risk3 computed over the period 2034-2039) for a grid of k responsiveness parameters with responsiveness twice high for negative compared to positive CPUE slope and distance to target (isolines show the 50, 60 and 70% probability of being in the green zone of the Lobe plot).*

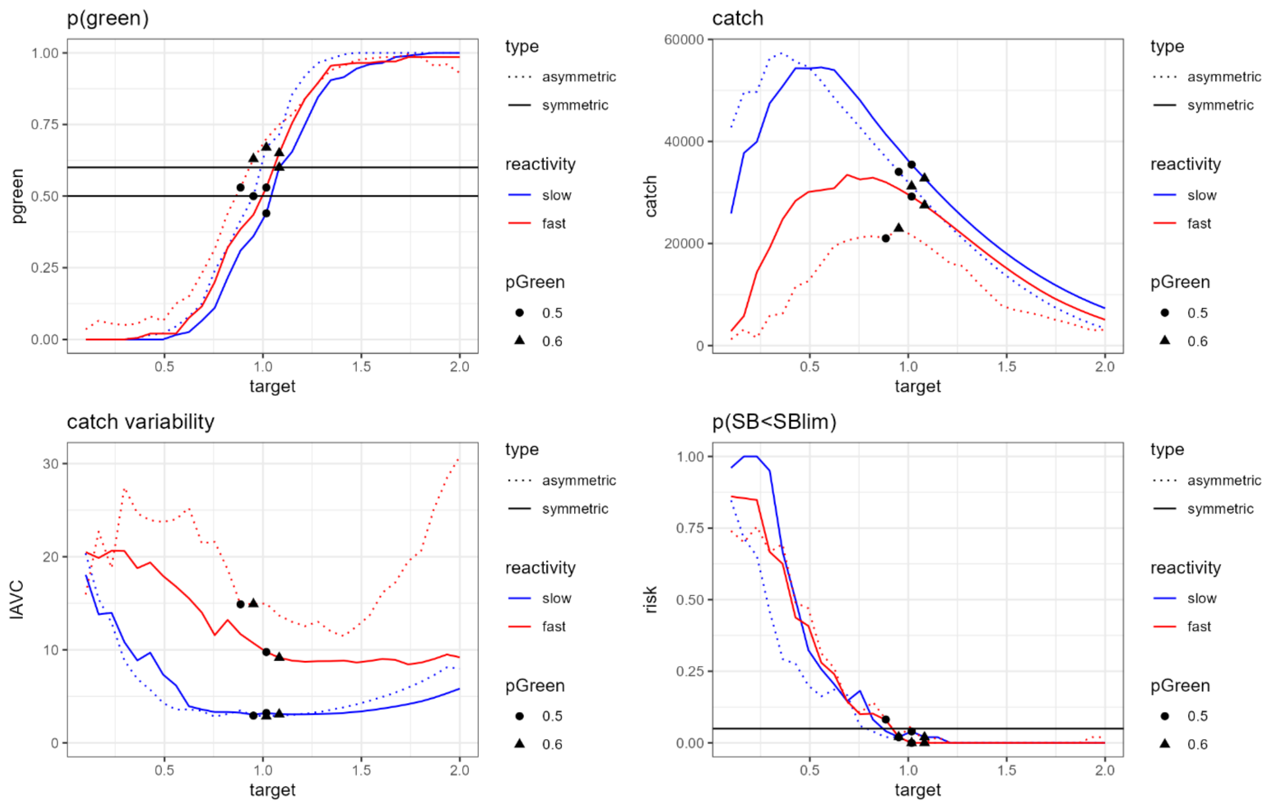


*Figure 9 : distribution of the values of 4 MP performance indicators (pgreen, mean catch, interannual catch variability and risk3 computed over the period 2034-2037) for a range of CPUE target values for a selection of 4 the combinations of k parameter values (identified by the dots on figure 6 and 7, with low and high responsiveness, symmetric or asymmetric responsiveness depending on sign of the slope and difference to target). For each combination, CPUE target value closest to leading to 50 and 60% probability of being in the Kobe green are identified by the dots and triangles respectively*.

## Robustness test

The conclusion on the CPUE rule, that low and symmetric *k* values may be optimal, needs to be challenged by a robustness test. Indeed, given the current good state of the stock, such MPs could lead to the best and most stable catches, without resulting in a high risk, but they might not be MPs that would protect the stock or ensure its recovery if it was put at risk by a particular event (e.g. recruitment failure).

To test that, a series of poor recruitments were imposed at the start of the simulation, by setting all deviations from the stock-recruitment model to 0.1 for the period 2022-2025. This purely fictive scenario was chosen in order to produce a substantial decrease in stock size.

Simulations were run for the slow and fast reacting CPUE MP (that were each tuned beforehand). The figure 10 shows that yearly values of p(Kobe green) and risk3 for both MP run both with the base case OM and with the OM with recruitment failure. Values of p(Kobe green) and risk3 are similar with the base case OM, with p(Kobe green) decreasing from current levels to about 0.5 in 2045 (tunned for 60% over the period 2034-2039). In both cases the risk increase over the period but remains very low.

When a recruitment failure is introduced, the risk increases more quickly, especially for the slow reacting MP, that seems to be less able to protect the stock from declining. However, even in this case, the risk is remain low, below 5%.

This seems to indicate that given the current status of the stock, both MPs, even the slow reacting one are robust to an event as extreme as 5 consecutive years of low recruitment.

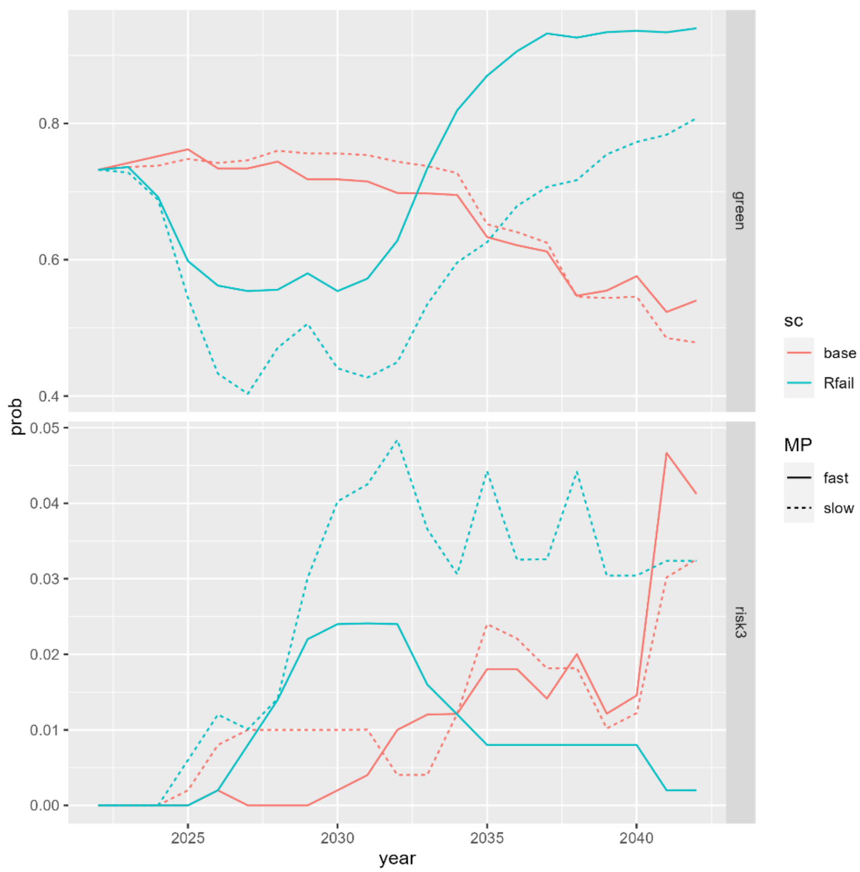


*Figure 10 : probability of being in the green part of the Kobe plot (top pannel) and probability of SB<SB$_{lim}$ (bottom pannel) over time for a slow and a fast reacting CPUE MP, for the base case OM and for an OM with 5 years of recruitment failure at the start of the simulations.*

## Tunning of the MPs

The three MPs were tuned for an objective of 50%, 60% and 70% probability of being in the green part of the Kobe plot. Performance indicators for the 9 tunned MPs are presented on figure 10:

- The tunning criteria is achieved in all cases (mean of p(Kobe green) at 0.5, 0.6 or 0.7 depending on MP), but there is a large variability in the distribution of the values between iterations (25% quartile and 75% quartile at 0 and 1 p(Kobe green) respectively). Analysis of the results of earlier simulations carried out on the earlier version of the SWO OM showed that the distribution of p(Kobe green) was bimodal, with most iterations having a value at either 0 or 1, and very few in between (Brunel and Mosqueira, 2022). This was interpreted as the consequence of two causes : first, the large variability in the OM, second, the fact that none of the MP need to lead to any substantial change in stock size to achieve the tuning criteria (p(Kobe green)$_{2022}$ = 73%) and therefore most of the iterations being in the green part of the Kobe plot at the start of the simulation will remain there for the whole simulation period (and similarly of red).
- the harvest rate (hr) MP leads, overall to a small stock size (SB/SB[MSY]) and the cpue MP to the largest. The hockey stick MP (hcst) leads to stock sizes intermediary between the two others, but with much larger variability.
- All MPs lead to a high probability of the stock remaining above Blim (p(SB>SBlim) > 95% in all cases).
- The mean catch is substantially higher for the cpue MP (33 to 37 thousand tonnes), but the variation amongst iterations is large. Catches are lower (31 thousand tonnes) for the hr MP, with little difference for the different the tunning objectives. Catches are the lowest for the hcst MP (between 26 and 29 thousand tonnes depending on tuning criteria), with very low variability across iterations.
- Interannual variation in catches is highest for the hr MP, lower for the cpue MP, and the lowest for the hcst MP.

Future trajectories of Fbar and SB (figure 11 and 12) are very stable (for the annual median across iterations as well as for individual iterations), and with little contrast across the MPs. Only in the longer term (after 2040) an increasing trend is observed for Fbar in the cpue MPs (and decreasing for SB). Catches are generally increasing with the cpue MP, while for the hr MP, they have an initial increase (at the first year the MP is used to set advice), followed by a slow increase and then a decrease. For the hcst MP catches are almost always equal to the maximum catch in the HCR, except when the MP is tuned for 50% probability for p(Kobe green), where catches are reduced (sliding slope of the HCR) for an increasing number of iterations as simulation time progresses.

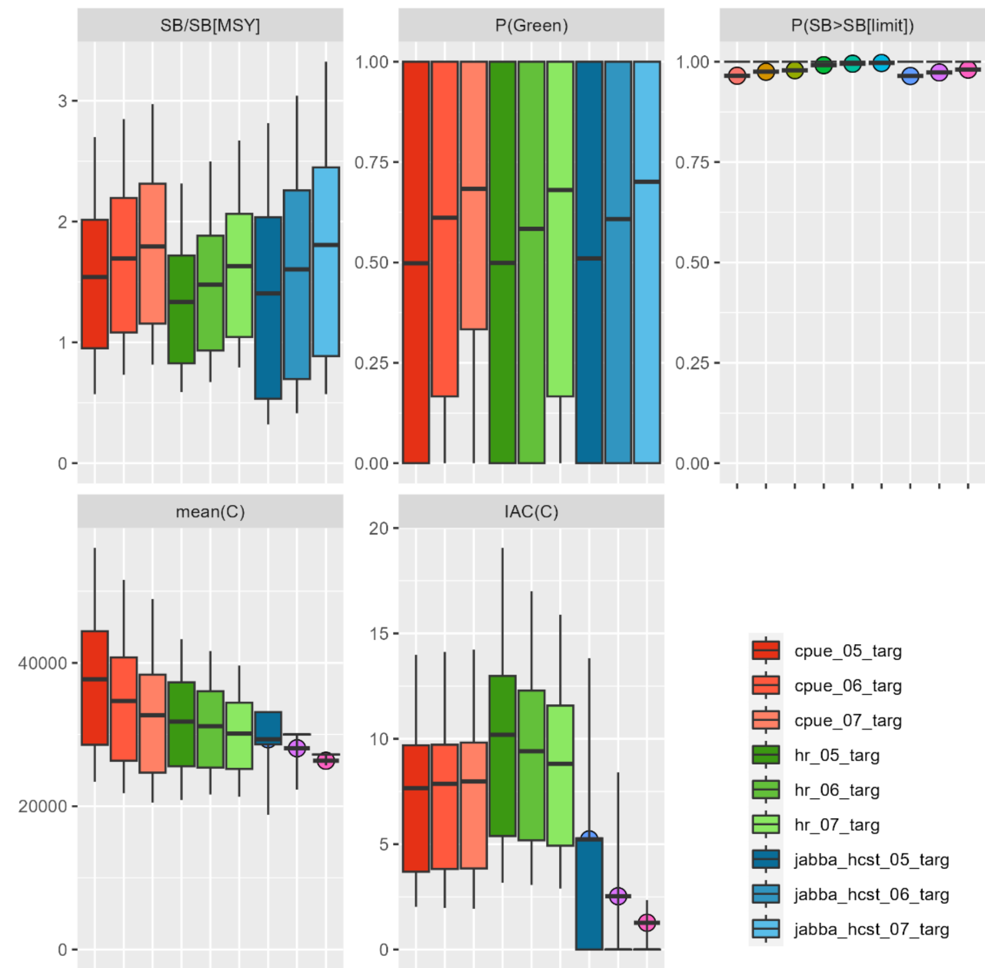


*Figure 11 : performance indicators for the three MPs proposed for SWO, each tunned for three levels of probability on being in the green part of the Kobe plot over the years 2034:2039.*

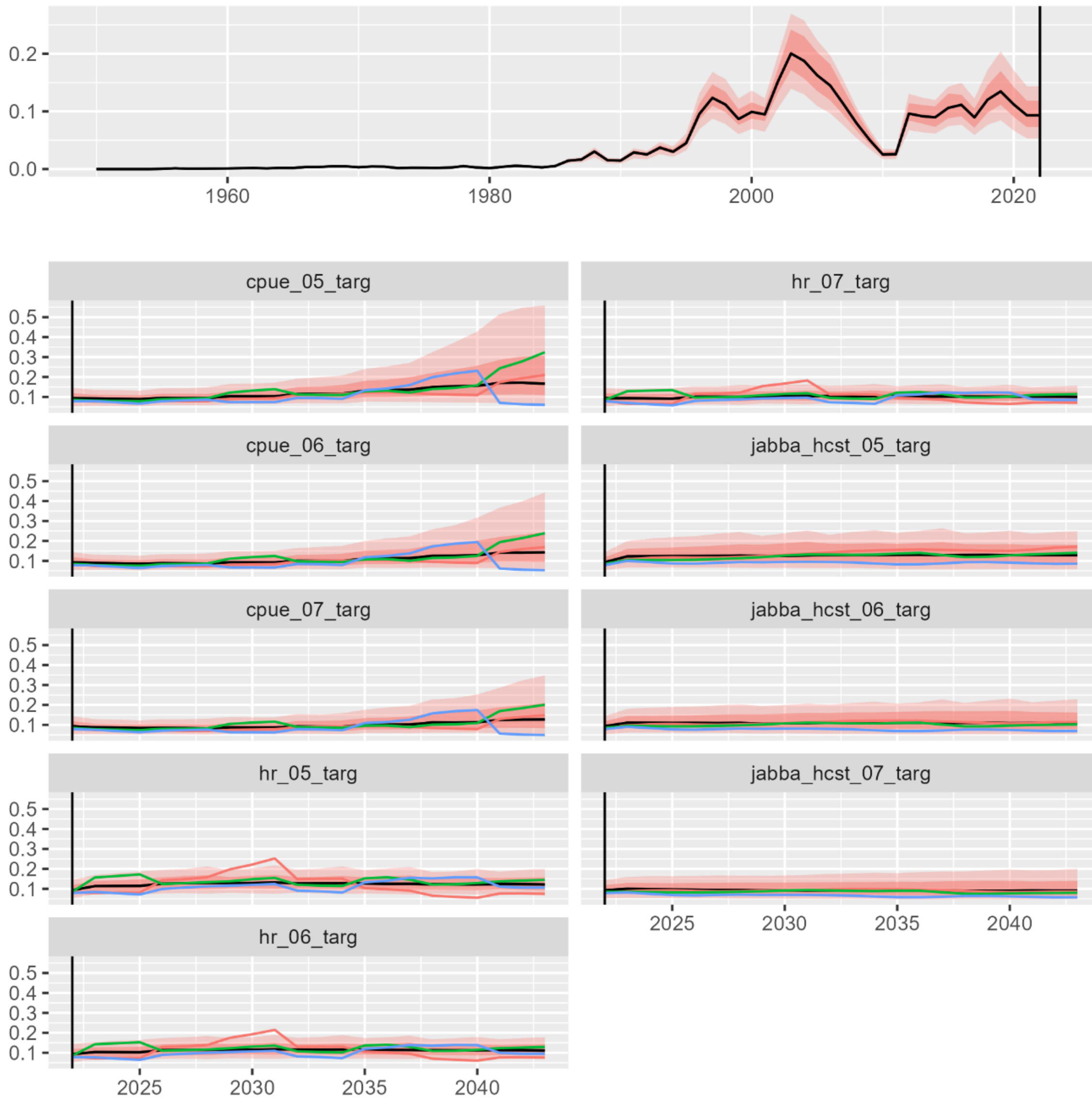


*Figure 12 : past (OM) and future developments of Fbar of SWO for the 9 tunned MP tested*

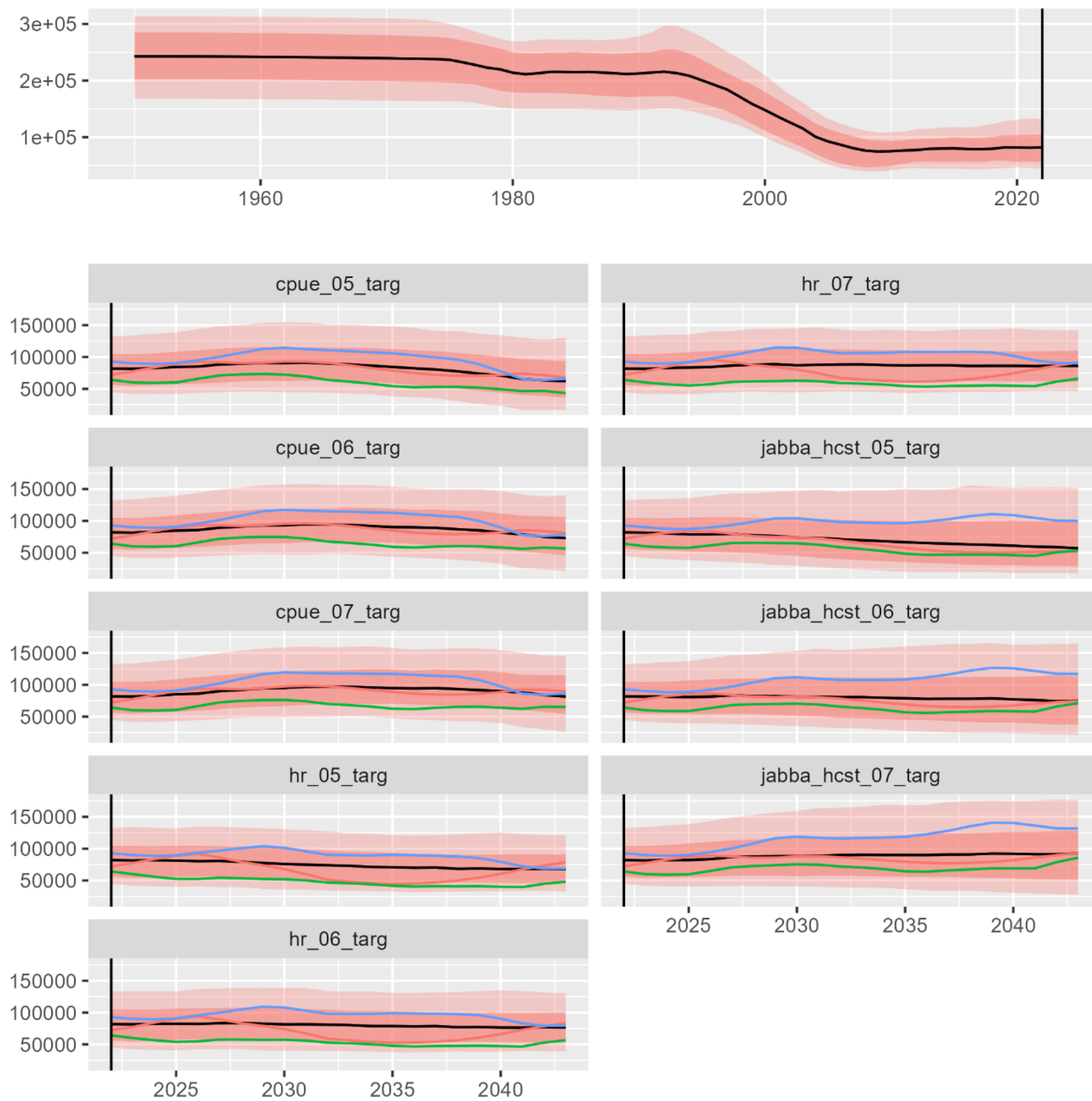


*Figure 13 : past (OM) and future developments of SB of SWO for the 9 tunned MP tested*

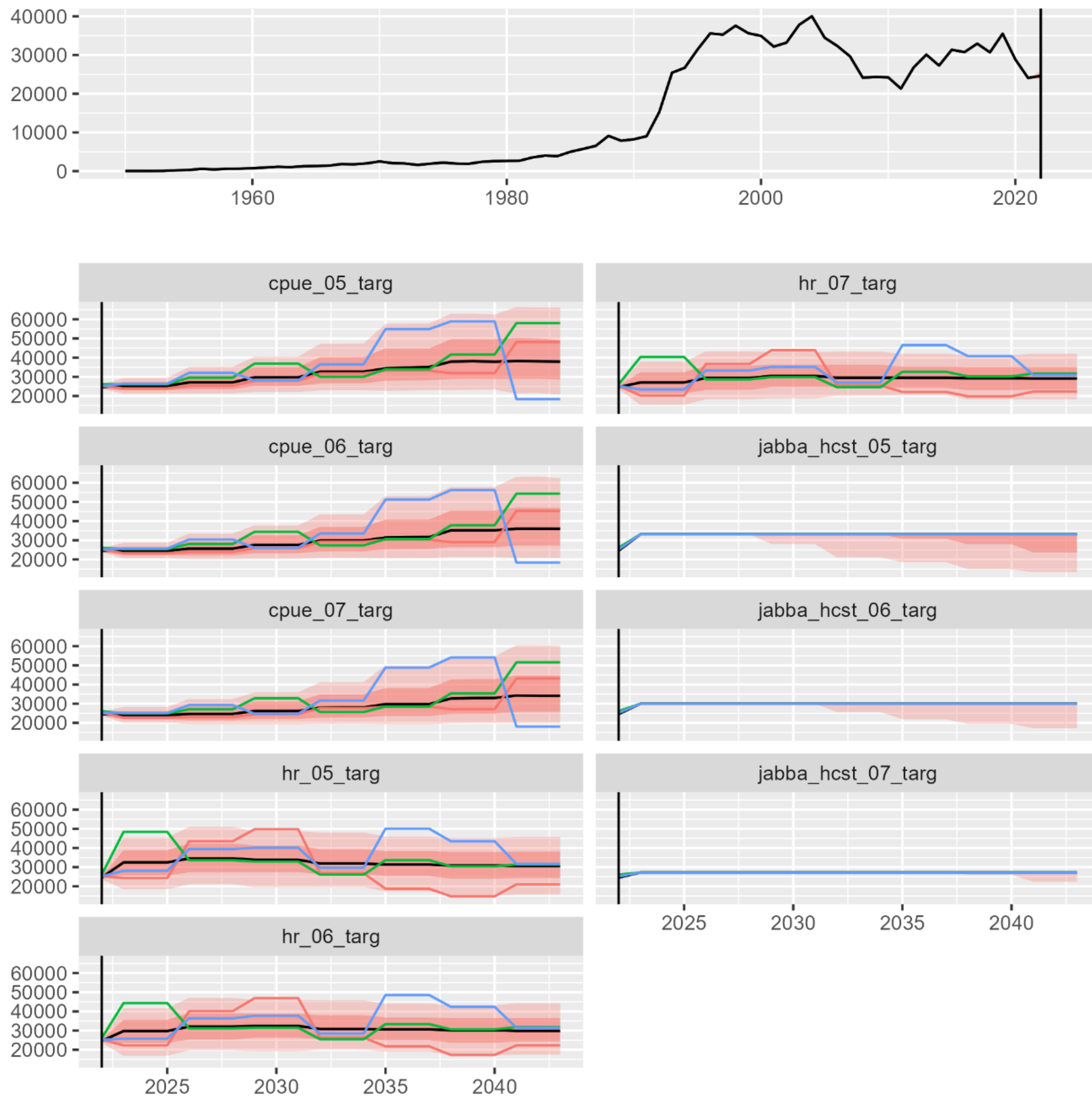


*Figure 14 : past (OM) and future developments of the catch of SWO for the 9 tunned MP tested*

## Discussion

### OM development

The proposed changes in the uncertainty grid lead to a greater number of model runs that can be used to build the OM, which can be viewed as an improvement. Concerns were raised by the model developer at the WPB in 2023 that the OM based on the old grid had a very large variability in stock status at the start of the simulation period, which lead to a very wide distribution of some of the performance indicators, amongst which the one used for tuning (p (Kobe green ), figure 10). This triggered the proposal from the model developers to try and simplify the uncertainty grid. The proposal made at WPB2023 included a revision of the range of recruitment steepness (0.6 to 0.8 proposed instead of 0.6 to 0.9). This proposal did not result in a noticeable reduction of the variability in the OM, and removing high steepness values from the OM grid was not judged appropriate by the WPB (as values up to 0.9 are used in other tuna commissions for swordfish).

The stock assessment of Indian Ocean swordfish is due to be updated by the WPB in 2023. The model developers will compare the new perception of the stock provided by this update assessment with the current OM to decide whether the current OM is still a suitable basis to conduct an MSE, or if it has to be updated, using as basis the 2023 assessment. Comparison will be based on the distribution of stock status at the start of the simulations, as well as on the population dynamics parameters.

## Implementation of JABBA in the MSE

Implementing a surplus production model in the simulations has proved difficult, mainly due to the fact that it requires to define a model configuration that would be suitable for a wide range of situations, and remains suitable in the future, with longer simulated input data time series. Here, a version of JABBA designed for implementation in MSE was used. Improved computation time and more stable model output are obtained by fixing the process error variance and observation error variances and let the model estimate only the population dynamics parameters. Assumptions on these variances are likely to have consequences on the estimated stock development, and therefore on the current depletion estimate, which is used in the HCR. An alternative approach could be to use in JABBA estimates of observation error obtained independently from the model, and let the model estimate process error variance. In addition, the model could be made more robust by using more informative priors on population dynamics parameters (e.g. K) that would be derived from the stock synthesis assessments on which the OM is based.

## Preliminary results with tunned MP

The CPUE MP can be parameterized with various degrees of responsiveness, and this aspect of the MP affects the performance of the MP as much the value of the parameter that is estimated by tuning (target CPUE index). The investigation done here using a grid of values of k parameters suggests that in this case, low k values provide advantages in terms of higher and more stable catches. This slow reacting CPUE MP is also robust to a recruitment failure, even with the pretty pessimistic scenario tested here (0.10 times the stock-recruitment model predictions for 4 years in a raw, while the lowest recruitment log-residual in the historical part of the OM is 0.31).

In the simulations conducted in 2022, it appeared that the hockey-stick MP resulted in a narrower distribution of the future stock status (i.e. performance indicator SB/SB[MSY]) than the data based MP. This year's simulations lead to the opposite conclusion, with much wider distribution of SB/SB[MSY] for the hockey-stick MP (figure 10). Last year, due to technical issues with the implementation of JABBA, this MP was applied assuming a perfect assessment (depletion status taken from the OM, not estimated by JABBA). The increase in variability is likely an effect of introducing assessment error in the system (i.e. managing on an stock assessment estimate and not the OM value). Therefore the variability in stock status estimates in this year's simulations are more realistic than last year

Overall, all MPs appear to results in precautionary management, successfully maintaining the stock above Blim. This would suggest that using 50% probability of being in the green part of the Kobe plot would be a reasonable tuning criteria. Based on the preliminary results presented here, there are two main management options. Choosing for the data-based CPUE MP would lead to the highest catch, but with some moderate interannual variations in the TAC (around 7% change from year to year on average). As an alternative, managers could choose the hockey-stick, which provides very stable catch, but at the cost of a lower catch level than with the CPUE MP (29 vs 38 thousand tonnes).

Finally, it should be noted that by tuning for an objective in the medium term and calculated over a short number of year, applying those MP would result in the long term in performance that could differ from the one reported here for the tuning period. For example, the CPUE MP clearly sets the stock on a decline trend over the long term, and performance indicators, if they were calculated on the long term, would be not as good as those reported on figure 10.

## References

Brunel, T., Mosqueira I. 2022. Indian Ocean swordfish MSE: investigating an alternative Operating model. IOTC-2022-WPB20-18.

Fischer, Simon H, José A A De Oliveira, John D Mumford, and Laurence T Kell. 2022. "Exploring a Relative Harvest Rate Strategy for Moderately Data-Limited Fisheries Management." Edited by M S M Siddeek. *ICES Journal of Marine Science* 79 (6): 1730–41