Update of the Indian Ocean swordfish operating model

12th Working Party on Methods

Daniela Rosa[*] Dan Fu[†] Rui Coelho[*] Iago Mosqueira[‡]

Online, 18th - 20th October 2021.

Abstract

This document presents the current status of development of an Operating Model for the Indian Ocean swordfish (*Xiphias gladius*) stock, with the re-condition of the models to the most recent swor sfish stock assessment, conducted in 2020. It explores the role of 9 axis of structural uncertainty, each with 2 to 3 levels. The current grid results in 2592 alternative population trajectories and productivity estimates. To reduce the number of model runs, a partial factorial design of 108 model runs is proposed. A set of diagnostics are then applied, and currently a set of 70 models is considered to compose the OM.

1 Introduction

The Indian Ocean Tuna Commission (IOTC) has committed to a path of using Management Strategy Evaluation (MSE) to meet its obligations for adopting the precautionary approach. A species-specific workplan was adopted at the 2017 IOTC meeting (IOTC 2017a), outlining the steps required to adopt simulation-tested Management Procedures for the highest priority species, among them the Indian Ocean swordfish stock. The 2017 session of the IOTC Working Party on Methods (WPM) (IOTC 2017b) discussed and proposed an initial set of elements likely to be responsible for most of the model uncertainty, both in past dynamics and current stock status. The development of the operating model (OM) started with the 2017 assessment and several updates have been presented as well as an initial testing for candidate management procedures (Mosqueira et al. (2017), Rosa et al. (2018), Rosa et al. (2019)). The objective of the present working document is to update the OM conditioning to the most recent IOTC swordfish stock assessment, conducted in 2020, including diagnostics applied to OMs.

2 2020 Stock Assessment

The most recent Indian Ocean swordfish stock assessment (Fu 2020) was presented at the 2020 session of the Working Party on Billfish (WPB). The Stock Synthesis 3 (Methot and Wetzel 2013) population model is age-based (with ages 0-30), separated by sex, and partitioned into four areas. Information from 15 fisheries, defined by fleet and region, was used, including length composition data for fourteen of them.

The stock assessment explored the uncertainty with respect to various assumptions through a grid of 24 model runs (Figure 1), based around two alternative values for growth, three values of stock-recruit steepness, two of recruitment variability, and two of effective sample size of the length composition data. All of these elements have been incorporated in the grid developed by WPM. Sensitivity runs included changes to CPUEs, length composition and growth information. For complete details of the models please refer to Fu (2020) and IOTC (2020).

[*]Instituto Português do Mar e da Atmosfera (IPMA), Av. 5 de Outubro, 8700-305 Olhão, Portugal./Centre of Marine Sciences (CCMAR), University of Algarve, Campus de Gambelas Ed 7. 8005-139 Faro, Portugal

[†]Indian Ocean Tuna Commission Secretariat, Le Chantier Mall (2nd floor), PO Box 1011, Victoria Mahe, Seychelles

[‡]Wageningen Marine Research, Ijmuiden, Netherlands

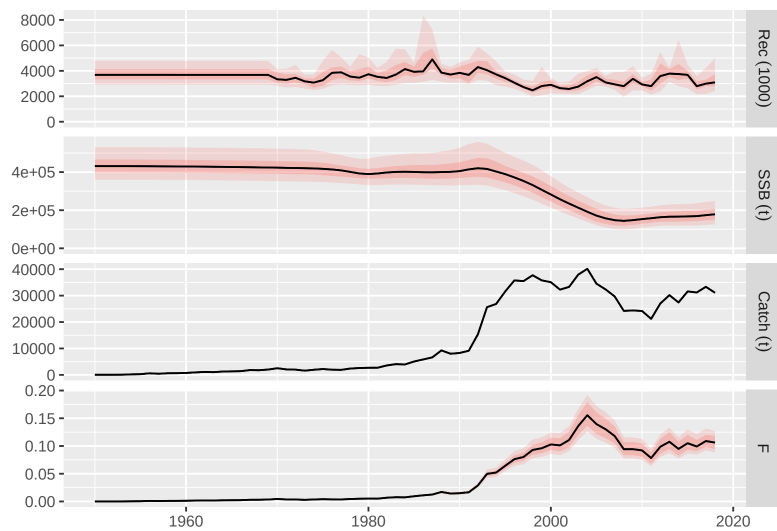

Figure 1: Population trajectories (recruitment, SSB, catch and F) estimated by the 2020 Stock Synthesis stock assessment of Indian Ocean swordfish.

The operating model (OM) was built around one of the stock assessment models `io4_h80_GoMf_r2_CL005`. This is a four area model, with both the Japanese (all areas), Portuguese and South African (southwest) CPUEs, steepness of 0.80, fast growth, recruitment sigmaR=0.2, effective length frequency sample size capped at 5. For simplicity of building the OM grid, as both the Portuguese and South African CPUE presented similar trends, the last was dropped from the model fit. This model will be hereafter refered to as "base case."

2.1 Model diagnostics

As in albacore MSE (Mosqueira 2020), an initial set of model diagnostics, following Carvalho et al. (2020), was applied to the base case. Diagnostics were based on convergence, retrospective analysis, runs test and hindcast cross validation. These diagnostics were run using the R package *ss3diags* (Winker et al. 2021).

2.1.1 Convergence level

Convergence is assessed by the inversion of the Hessian matrix, and the value of the gradient at the solution. Carvalho et al. (2020) suggest the value of the gradient at the solution should be smaller than 1e-4. In the first iteration the convergence value of the base case model was higher than the proposed value, initial parameters were jittered and a convergence lower than 1e-4 was obtained.

2.1.2 Retrospective analysis

The retrospective pattern for the estimated SSB from the base case model is shown in Figure 2. This plot also includes a one-step ahead projection of SSB based on the known total catches. Retrospective analysis is a form of hindcasting commonly used to assess the stability of a model formulation to updates in the data. Mohn's rho (Mohn 1999) is then used to quantify the strength of the retrospective pattern. (Hurtado-Ferro et al. 2014) proposed that values outside the -0.15 to 0.20 range, should indicate an undesirable retrospective pattern for longer lived species. In the case of the base case model Monh's rho is within the proposed range.

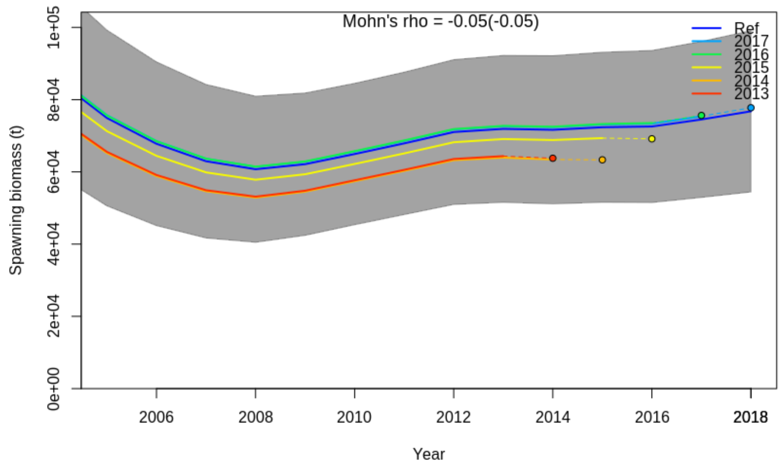

Figure 2: Five year restrospective runs with a one step ahead forecast of spawning stock biomass according to total catch for Indian Ocean swordfish.

2.1.3 Runs test

Runs tests on the CPUE and length-frequency data sources are common diagnostics of goodness of fit (Carvalho et al. 2017). The Wald-Wolfowitz runs test, a non-parametric statistical test that checks the randomness hypothesis for a data sequence, can be used to identify residuals patterns that should not be considered random. The runs test for the CPUEs and for the length-frequency data sources in the base case (Figure 3 and Figure 4) indicate that the model fit to most CPUE data series (those in green) do not present significant patterns, while for length half the series present significant patterns in the residuals.

2.1.4 Hindcasting cross validation

A proposed model-free hindcasting technique (HCXval) uses cross-validation to compare observations to their predicted future values (Kell, Kimoto, and Kitakado 2016). The prediction skill of a model is then computed from the prediction residuals. A robust statistic for evaluation the prediction skill can be constructed using the mean absolute scaled error (MASE) of Hyndman et al. (2006). The R package *ss3diags* contains functions that simplify the calculation of prediction skill and the computation of their MASE for both CPUE and length frequencies. The prediction skills of the indices for each area are presented in Figure 5. A MASE score larger than 1 indicates the model does only as well as a random walk at predicting the quantity, while a value of 0.5 indicates the model is twice as good as a random walk. The model prediction skill appears to be better for the Northwest area.

3 Operating model

The OM being developed here is a second iteration of conditioning of the operating model. It is based on the population and fishery models used for the assessment of the stock status of Indian Ocean swordfish in 2020 (see above). Work on this MSE exercise is being carried out around a public source code repository, available at https://github.com/iotcwpm/SWO.

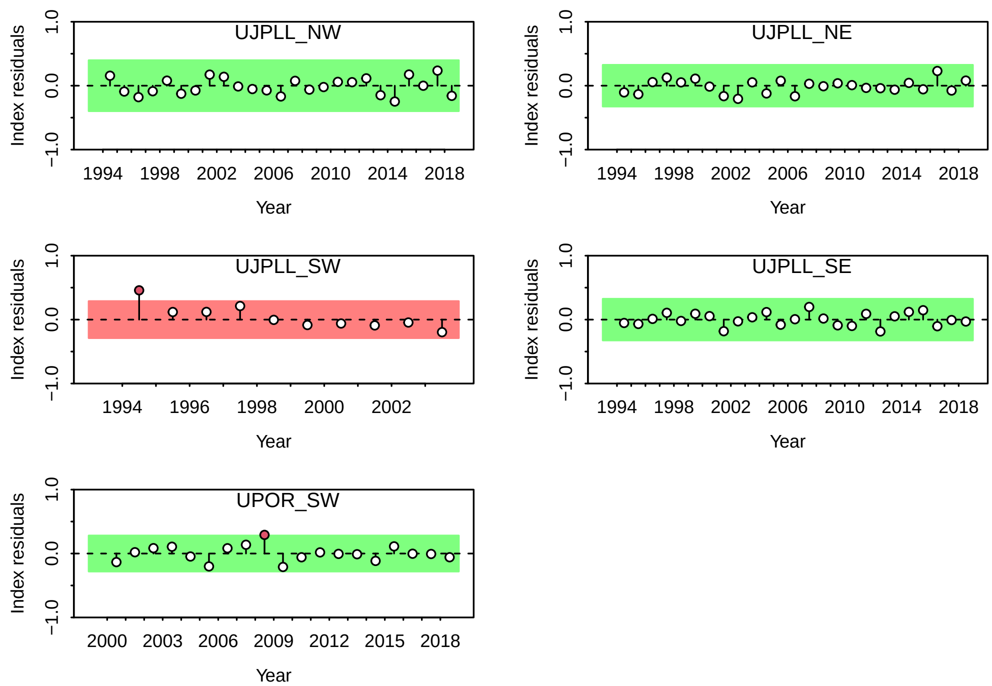

Figure 3: Runs tests of the CPUE series in the base case model for Indian Ocean swordfish.

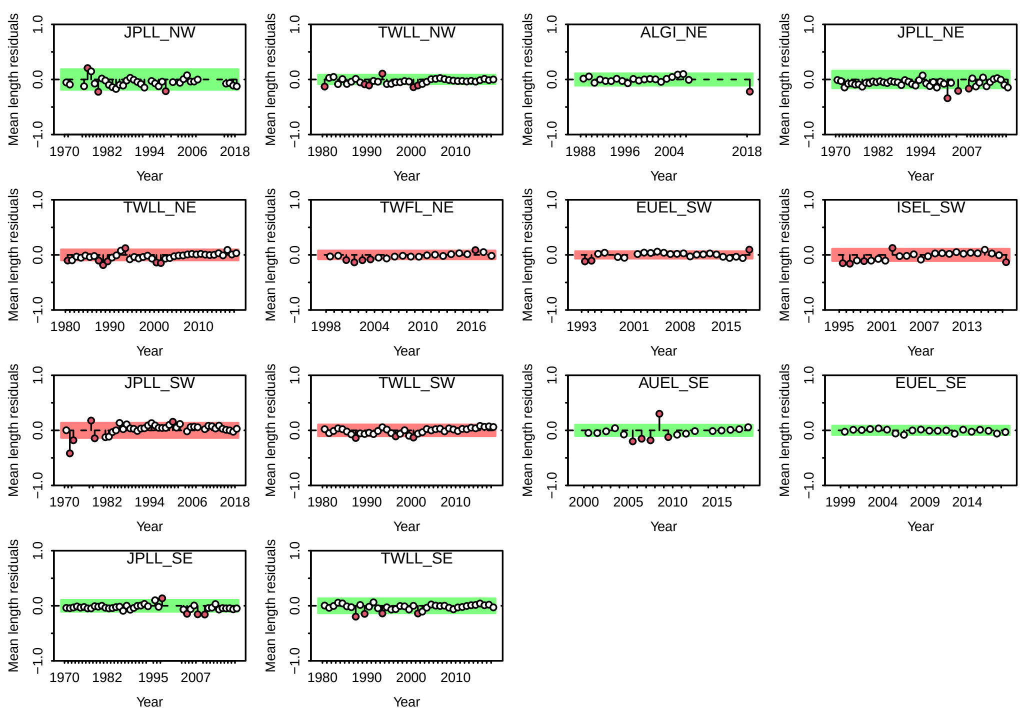

Figure 4: Runs tests for the various sources of length frequency data in the Indian Ocean swordfish base case.

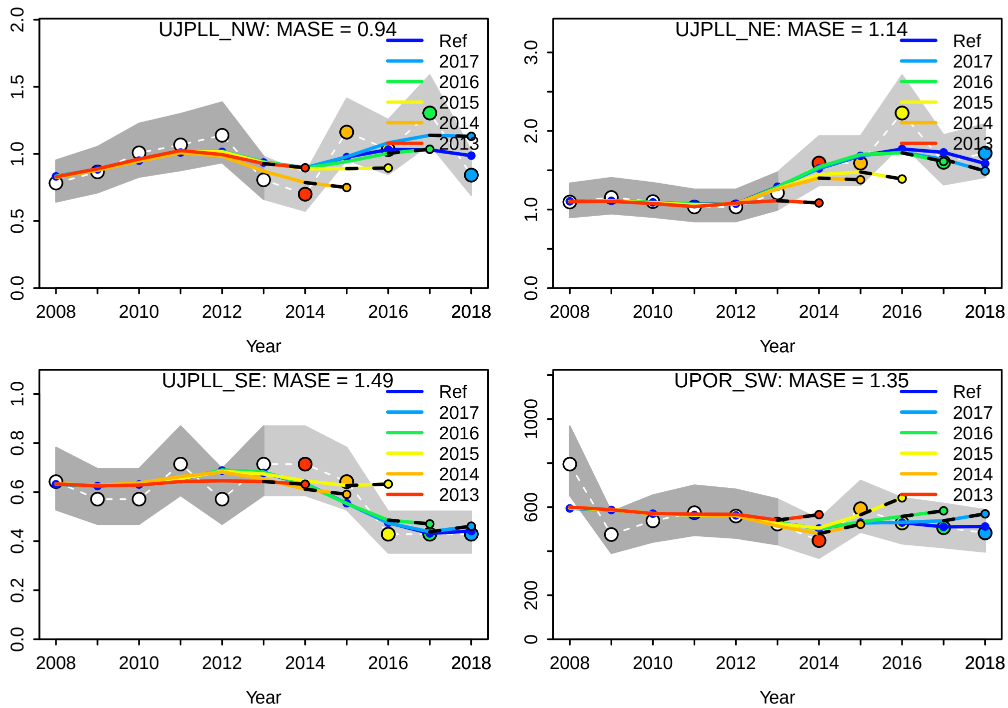

Figure 5: Hindcasting cross-validation results for CPUEs by area for Indian Ocean swordfish base case.

3.1 Structural uncertainty grid

The 2017 session of the WPM proposed an initial set of options for characterizing the structure of the uncertainty grid for generating the OM, based on a set of SS3 model runs (IOTC 2017b). Those options were further discussed during the workshop meeting of the authors to start the conditioning of the OM. The decision was to construct a grid of model runs built around those suggested by the WPB on feasible, or at least not too extreme, values for a number of assumptions and fixed parameters in the population model. The impact of some of these elements in the model were already explored in some detail by the researchers carrying out past stock assessments (Secretariat (2017), Fu (2020)).

The nine factors currently considered in the structural uncertainty grid for the swordfish OM are the following:

3.1.1 Selectivity

Two functions were considered for the selectivity-at-length of the longline fleets: the current *double normal*, in which selectivity decreases in the older ages, and a *logistic* function, in which selectivity remains flat after reaching its asymptote.

3.1.2 Steepness

Steepness (h) from Beverton and Holt stock-recruitment function is often a very influential parameter which is difficult to estimate in most stock assessments. The base case SA models used *0.80*, to reflect plausible lower and higher values, three values are considered in the uncertainty grid (*0.6*, *0.75* and *0.9*).

3.1.3 Growth & Maturity

Growth and maturity are very important parameters in stock assessments. Swordfish exhibit a marked difference in growth between male and female, therefore sex-specific growth and maturity estimates are used in all cases. There are concerns in the age estimation of swordfish, with differences being found in the results depending on what structure is used to estimate age (fin rays or otoliths). This uncertainty also undermines the maturity by age relationship. Two growth curves and maturity estimates are considered for the OM (Figure 6):

- Slow growth and late maturity (Wang, Lin, and Chiang (2010))
- Faster growth and earlier maturity (Farley et al. (2016), from otoliths)

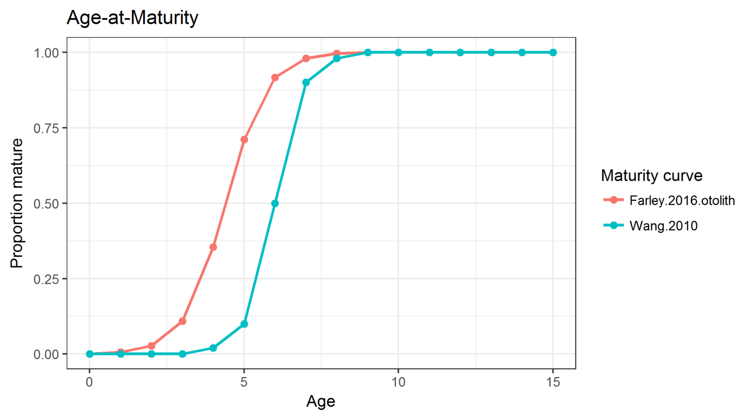

Figure 6: Maturity at age scenarios used for conditioning the Swordfish operating model.

3.1.4 Natural Mortality (M)

Natural mortality is a common unknown in most stock assessment models. The base case considered in the stock assessment model was 0.2 constant for all ages, which was supplemented with an alternative value of 0.4 also constant for all ages as suggested by the WPM. After initial exploration of OM results it was clear that setting M at 0.4 would not produce plausible estimates of biomass, therefore, based on these results, the authors decided to set natural mortality to 0.3 instead of 0.4. A 3rd possibility using age-specific M values, based on the the Lorenzen equation was also included in the grid. The age specific mortality was scaled so that M at age at maturity (age 6) was 0.25. A total of 3 possibilities were therefore considered for M in this grid (Figure 7):

- 0.2, constant for all ages
- 0.3, constant for all ages
- Age and sex specific values based on the Lorenzen equation

3.1.5 Effective Sampling Size (ESS)

Two values were used for the relative weight of length sampling data in the total likelihood, through changes in the effective sampling size parameter, of 2 and 20. This alters the relative weighting of length samples and CPUE series in informing the model about stock dynamics and the effects of fishing at length.

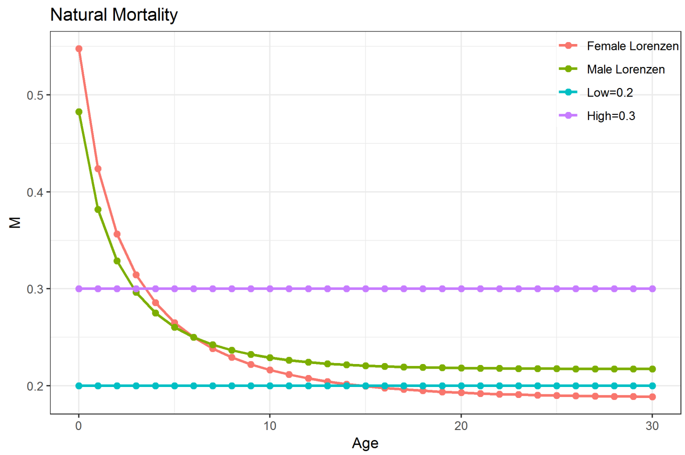

Figure 7: Natural Mortality options used for conditioning the swordfish operating model.

3.1.6 CPUE series

CPUE series presented to the WPB showed conflicting trends, specially in the final years of the series. The base case considered in the assessment used the Japanese late (1994-2018) CPUEs, with the Portuguese indices from 2000-2015 being used in the Southwest area. An alternative view could be generated by using the Taiwanese CPUEs, again in combination with those from the Portuguese fleet for the SW. A total of 3 possibilities are thus being considered for CPUE series in this model grid:

- *JPNlate + EU.PRT*: Japanese CPUE (1994-2018) with indices 2000-2018 in the SW replaced by the Portuguese index (with an overlap in the years 2000-2003),
- *JPNlate*: Japanese CPUE (1994-2018) for all areas,
- *TWN + EU.PRT*: Taiwanese CPUE (1994-2018) with indices 2000-2018 in the SW replaced by the Portuguese index (with an overlap in the years 2000-2003).

3.1.7 CPUE scaling

The stock assessment assumed a stock residing in four areas and CPUE scaling was used as a mean to determine regional biomass distribution. Three alternatives for scaling the CPUES were considered:

- *area* effect * surface
- *catch* by area
- *biomass* by area, as estimated from a four area model with no scaling

3.1.8 Catchability increase

Two scenarios were considered for the effective catchability of the CPUE fleet. On the first one it was assumed that the fleets have not improved their ability to fish for swordfish over time, or that any increase had been captured by the CPUE standardization process (0% increase). An alternative scenario considered a 1%/year increase in catchability by correcting the CPUE index to reflect this.

3.1.9 SigmaR

Two values were considered for the true variability of recruitment in the population (*sigmaR*), specifically 0.2 and 0.6. The WPM discussed that both lower and higher options should be considered, but that a further middle value could also be added in the future (0.4) (IOTC 2017b). At this stage, and in order to not increase too much the grid of model runs, only the two extremes (0.2 and 0.6) were considered for recruitment variability.

3.1.10 Summary of the OM grid of uncertanties

Table 1 below summarizes the grid of uncertanties considered for the conditioning of the OM. This grid results in a total of 2,592 model runs.

Table 1: Summary view of the swordfish operating model 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)
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	

3.2 Main effects

A first exploration has been carried out of the individual effect on model output of adopting an alternative value for each variable, one at a time, using the base case model. The estimates of virgin spawning biomass (SSB_0) are shown in Figure 8, while trend (SSB_{2018}/SSB_{MSY}) is presented in Figure 9.

3.3 Fractional factorial design

In the first iteration of the OM the 2,592 models were run, which is computationally demanding, specially if running also the diagnostics that are in need of removing years of data. The MSE WPM (IOTC 2021a) suggested the fractional factorial design could be attempted. The fractional factorial design consists of a subset of the models in the full factorial design, considering that 3-way (and higher) interactions are rare and that models in the full factorial design are redundant, not providing any new information. In this case, 108 model runs were chosen to represent the uncertainties in stock dynamics.

3.3.1 Model selection

Model selection followed the suggestions by Mosqueira (2020) for albacore MSE. Initial selection of models was based on convergence level. Based on the distribution of the convergence level of the models (figure 10), models with a convergence level of 0.001 were excluded from further analysis (14 models).

Another selection was based on the CPUE MASE scores, in order to keep models with some prediction skill. Considering this, model runs with a MASE score>1 for the CPUE series in the NW region were also excluded from the operating model (26 models).

Given these two selection objectives, from the 108 models, 70 models are considered to constitute the operating model.

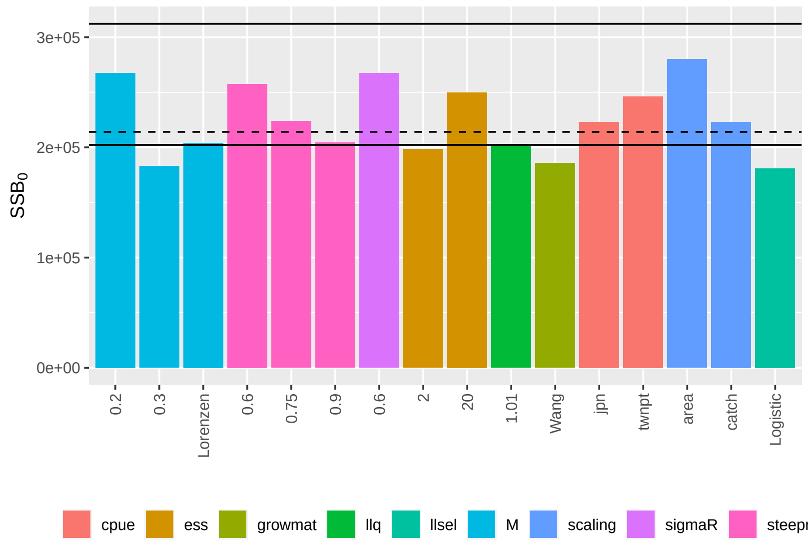

Figure 8: Changes in estimated virgin spawning stock biomass (SSB_0) for each level within each factor. Horizontal lines represent the minimum and maximum estimates of the stock assessment models, dashed line represents the model estimates used to construct the base case.

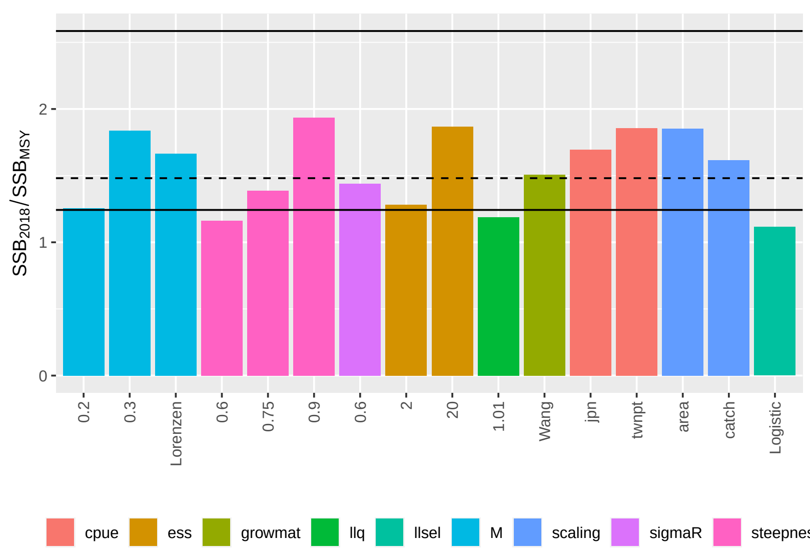

Figure 9: Changes in estimated spawning stock biomass in the last year (SSB_{2018}) over the SSB at MSY (SSB_{MSY}) for each level within each factor. Horizontal lines represent the minimum and maximum estimates of the stock assessment models, dashed line represents the model estimates used to construct the base case.

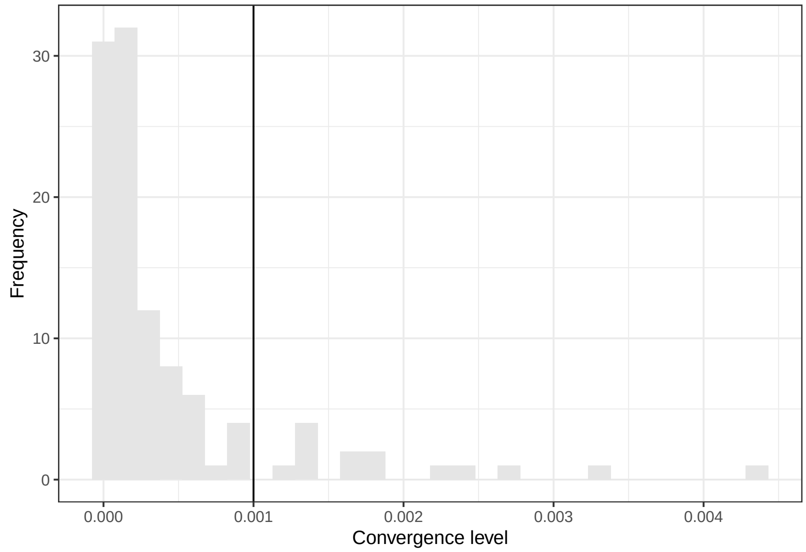

Figure 10: Frequency distribution of the convergence level for the 108 model runs. The black line represents the applied cut-off value (1e-3).

Further selection would be based on unrealistic estimates of SSB_0 (Figure 11) or SSB_{2018}/SSB_{MSY} (Figure 12).

In albacore models that cannot explain current catches are also being excluded from the OM, in the case of swordfish updating to recent catches (2019 and 2020-provisional) does not lead to the exclusion of any model.

A comparison of key quantities by each level of each factor is presented in Figures 13 and 14. Additionally, the overall time series plot of the OM (70 runs), shows values for abundance and fishing mortality to be widely distributed around the stock assessment (Figure 15), although somewhat more pessimistic.

4 Next steps

This document presents development of the work that has been carried out regarding to the management strategy evaluation of swordfish in the Indian Ocean, focus has been given to the re-conditioning of the OM. Progress on the MSE work for swordfish will continue on the development of the OM, testing of candidate management procedures with model weighting based on prediction skill (through MASE) using the p-value from Diebold-Mariano test and incorporating the feedback from TCPM04 (IOTC (2021b)). Potential robustness tests are also proposed to carried out as bellow:

- Continued low recruitment
- CPUE overcompensation bias
- Reported overcatch
- Not reported over catch
- Tracking the Southwest area abundance.

5 Acknowledgements

D. Rosa is supported by an FCT Doctoral grant (Ref: SFRH/BD/136074/2018).

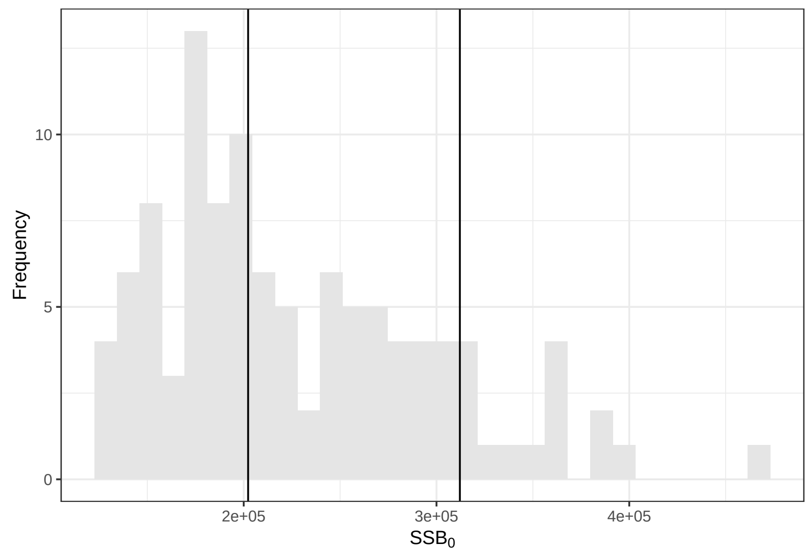

Figure 11: Distribution of the 108 models estimated values of spawning stock virgin biomass (SSB_0). The black lines show the minimum and maximum values of SSB_0 returned by the stock assessment grid.

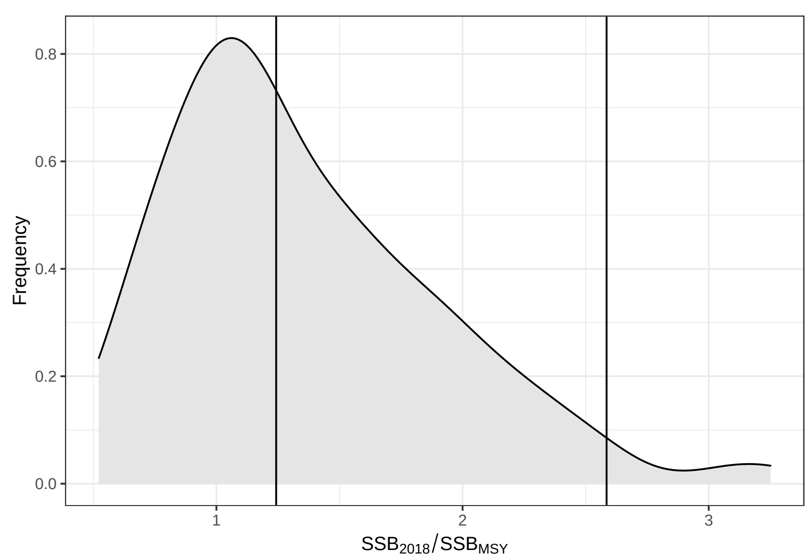

Figure 12: Distribution of the 108 models estimated values of spawning stock biomass in the last year (SSB_{2018}) over the SSB at MSY (SSB_{MSY}). The black lines show the minimum and maximum values of SSB_0/SSB_{MSY} returned by the stock assessment grid.

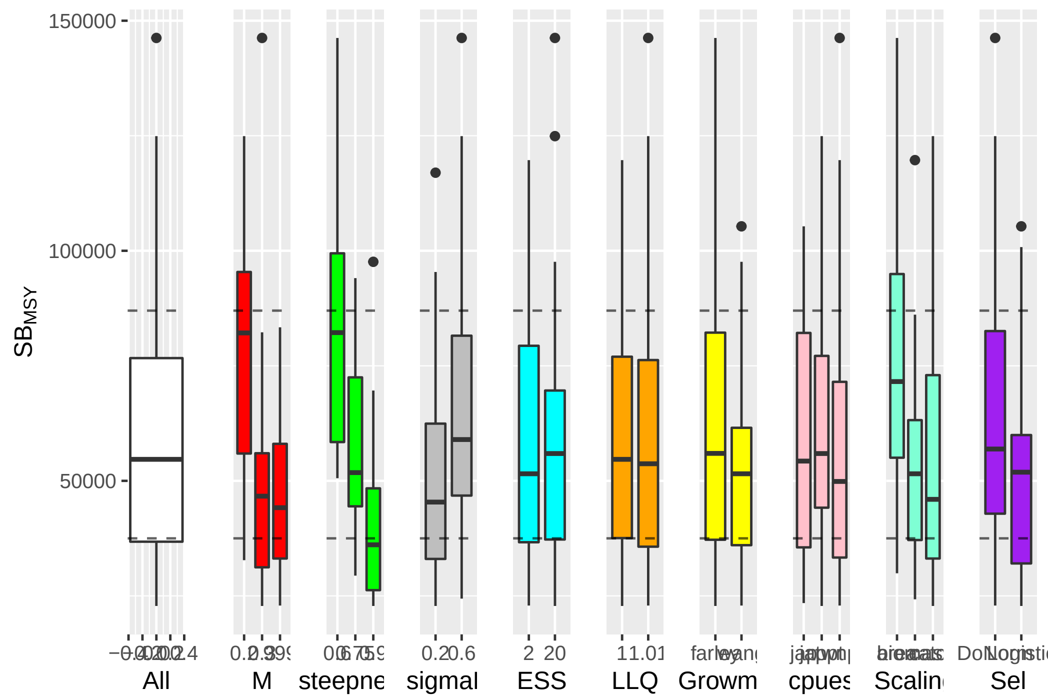

Figure 13: Comparison of spawning stock biomass at MSY (SSB_{MSY}) by level of each uncertainty factor. The dashed lines show the minimum and maximum values of SSB_{MSY} returned by the stock assessment grid.

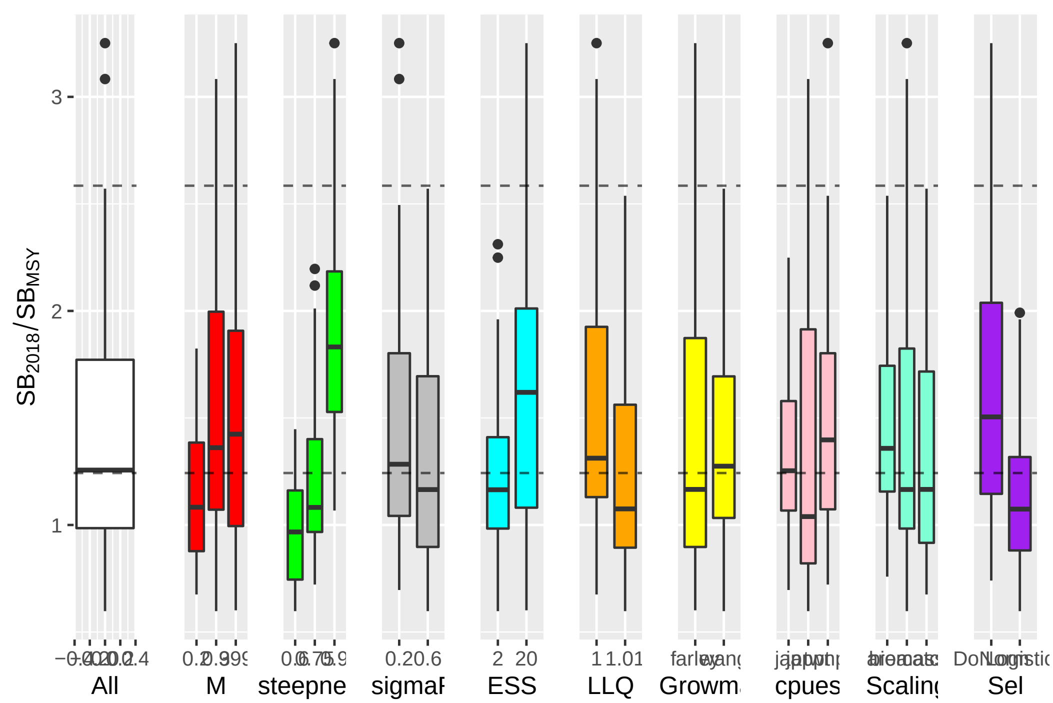

Figure 14: Comparison of spawning stock biomass 2018 over spawning stock at MSY (SSB_{2018}/SSB_{MSY}) by level of each uncertainty factor. The dashed lines show the minimum and maximum values of SSB_{MSY}] returned by the stock assessment grid.

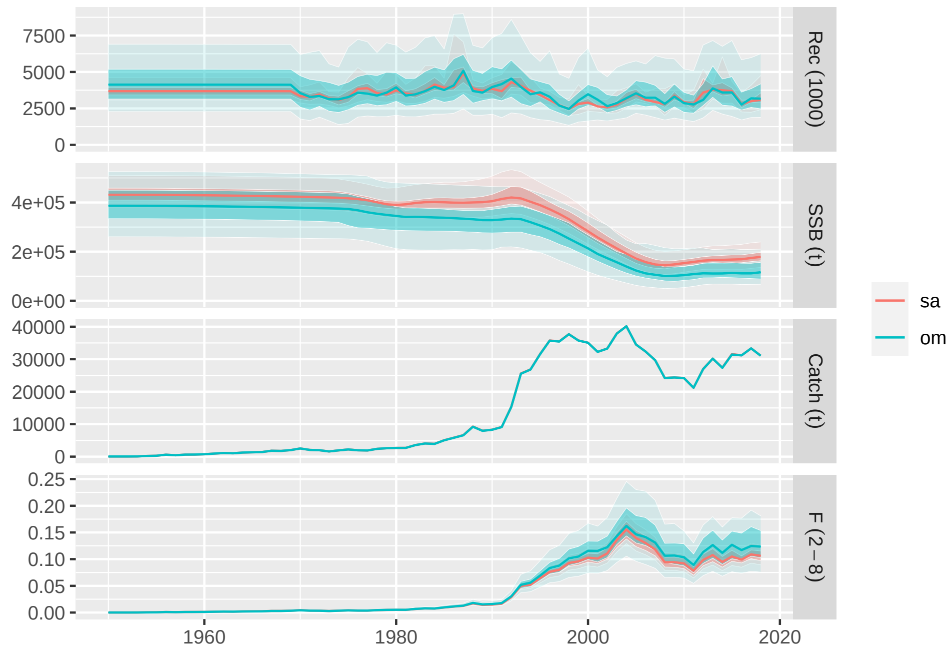

Figure 15: Population trajectories (recruitment, SSB, catch and F) estimated by the OM grid (in blue) and the stock assessment (in red) for Indian Ocean swordfish. The line shows the median value, while the darker and lighter ribbons show the 50% and 90% quantiles, respectively.

References

Carvalho, F., A. E Punt, Y.-J. Chang, M. N. Maunder, and K. R. and Piner. 2017. "Can Diagnostic Tests Help Identify Model Misspecification in Integrated Stock Assessments?" *Fisheries Research* 192: 28–40. https://doi.org/http://dx.doi.org/10.1016/j.fishres.2016.09.018.

Carvalho, F., H. Winker, D. Courtney, M Kapur, L. Kell, M. Cardinale, M. Schirripa, et al. 2020. "A Cookbook for Using Model Diagnostics in Integrated Stock Assessments." *Fisheries Research* 240: 105959. https://doi.org/http://dx.doi.org/10.1016/j.fishres.2016.09.018.

Farley, J., N. Clear, D. Kolody, K. Krusic-Golub, P. Eveson, and J. Young. 2016. "Determination of Swordfish Growth and Maturity Relevant to the Southwest Pacific Stock." R 2014/0821.

Fu, D. 2020. "Preliminary Indian Ocean Swordfish Stock Assessment 1950-2018 (Stock Synthesis)." IOTC–2020–WPB18–16. Indian Ocean Tuna Commission.

Hurtado-Ferro, F., C. S. Szuwalski, J. L Valero, S. C. Anderson, C. J. Cunningham, K. F. Johnson, R. Licandeo, et al. 2014. "Looking in the Rear-View Mirror: Bias and Retrospective Patterns in Integrated, Age-Structured Stock Assessment Models." *ICES Journal of Marine Science* 72 (1): 99–110. https://doi.org/https://doi.org/10.1093/icesjms/fsu198.

Hyndman, R. J., A. B. Koehler, R. D. Methot, and C. R. Wetzel. 2006. "Another Look at Measures of Forecast Accuracy." *International Journal of Forecasting* 22 (4): 679–88. https://doi.org/https://doi.org/10.1016/j.ijforecast.2006.03.001.

IOTC. 2017a. "Report of the 21st Session of the Indian Ocean Tuna Commission." IOTC–2017–S21–R[E]. Indian Ocean Tuna Commission.

———. 2017b. "Report of the 8th Session of the IOTC Working Party on Methods." IOTC–2017–WPM08–R[E]. Indian Ocean Tuna Commission.

———. 2020. “Report of the 18th Session of the IOTC Working Party on Billfish).” IOTC-2020-WPM18-R[E]. Indian Ocean Tuna Commission.

———. 2021a. “Report of the 12th Session of the IOTC Working Party on Methods (Management Strategy Evaluation Task Force).” IOTC–2021–TCMP04–R[E]. Indian Ocean Tuna Commission.

———. 2021b. “Report of the 4th IOTC Technical Committee on Management Procedures.” IOTC–2021–TCMP04–R[E]. Indian Ocean Tuna Commission.

Kell, L. T., A. Kimoto, and T. Kitakado. 2016. “Evaluation of the Prediction Skill of Stock Assessment Using Hindcasting.” *Fisheries Research* 183: 119–27. https://doi.org/https://doi.org/10.1016/j.fishres.2016.05.017.

Methot, R. D., and C. R. Wetzel. 2013. “Stock Synthesis: A Biological and Statistical Framework for Fish Stock Assessment and Fishery Management.” *Fisheries Research* 142 (Supplement C): 86–99. https://doi.org/https://doi.org/10.1016/j.fishres.2012.10.012.

Mohn, Robert. 1999. “The Retrospective Problem in Sequential Population Analysis: An Investigation Using Cod Fishery and Simulated Data.” *ICES Journal of Marine Science* 56 (4): 473–78. https://doi.org/https://doi.org/10.1006/jmsc.1999.0481.

Mosqueira, I. 2020. “Update of the Indian Ocean Albacore Operating Model.” IOTC-2020-WPM11-08. Indian Ocean Tuna Commission.

Mosqueira, I., D. Rosa, D. Fu, and R. Coelho. 2017. “Update on the Conditioning of an Operating Model for the Indian Ocean Swordfish Stock.” IOTC-2017-SC20-11. Indian Ocean Tuna Commission.

Rosa, D., I. Mosqueira, D. Fu, and R. Coelho. 2018. “Update on the Conditioning of an Operating Model for the Indian Ocean Swordfish Stock.” IOTC-2018-SC21-12. Indian Ocean Tuna Commission.

———. 2019. “Updates on the Indian Ocean Swordfish Management Strategy Evaluation: Initial Testing of Candidate Management Procedures.” IOTC-2019-WPM10-12. Indian Ocean Tuna Commission.

Secretariat, IOTC. 2017. “An Age, Sex and Spatially-Structured Stock Assessment of the Indian Ocean Swordfish Fishery 1950-2015, Using Stock Synthesis.” Indian Ocean Tuna Commission.

Wang, S.-P., C.-H. Lin, and W.-C. Chiang. 2010. “Age and Growth Analysis of Swordfish (Xiphias Gladius) in the Indian Ocean Based on the Specimens Collected by Taiwanese Observer Program,” no. IOTC-2019-WPM-08.

Winker, H., F. Carvalho, M. Cardinale, and L. Kell. 2021. *Ss3diags: Cookbook for Model Diagnostics in Stock Synthesis*.