Updates on development of MSE analyses for Indian Ocean albacore tuna and swordfish

Iago MOSQUEIRA, Thomas BRUNEL (WMR) *

13th Session IOTC Working Party on Methods. 19-21 October 2022

Abstract

A brief summary of current status and recent developments on the work for an MSE analysis for Indian ocean albacore tuna and swordfish is presented here. A new stock assessment for albacore tuna, and an updated grid for the swordfish OM grid, need both to be discussed by WPM to guide the next steps of work for these species.

Albacore tuna

The past session of WPTmT has updated the stock assessment for the stock, employing a similar formulation of the Stock Synthesis platform (Methot, XX). Two changes are worth noticing. First, the separation by season, with independent selectivity and catchability in each quarter, of the longline fisheries. The second is that the combined CPUE series for all longline fisheries is not based this time on operational-level data (Kitakado et al, 2022).

A final set of two model runs were chosen by WPTmT (IOTC, 2022) to provide advice on stock status for albacore. The basic configuration of both models is based on the diagnostic model presented to the meeting (Rice, 2022). They differ in the choice of longline CPUE index to be fitted: the NW longline aggregated index (LLCPUE1) for one, and the SW index (LLCPUE3) for the other.

Given the extended disaggregation of the various abundance indices in the new model, a number of diagnostics were examined to ascertain their suitability for running any management procedure that employs them. This choice would also determine which prediction skill statistics should be used to select and resample models in the OM grid. The prediction skill of each quarterly LL CPUE index was computed across the two model runs. Table 1 presents those area and quarter combinations where MASE < 1. Only the model fitted to the SW (LLCPUE3) series shows an index with acceptable MASE in the SW index in any quarter. Indices for the NW area in quarters 2 and 3, selected in bold, appear informative in both models. These two indices could form the basis for a management procedure, either combined in the case of the CPUE trend-based MP, or input separately in the case of the model-based one.

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

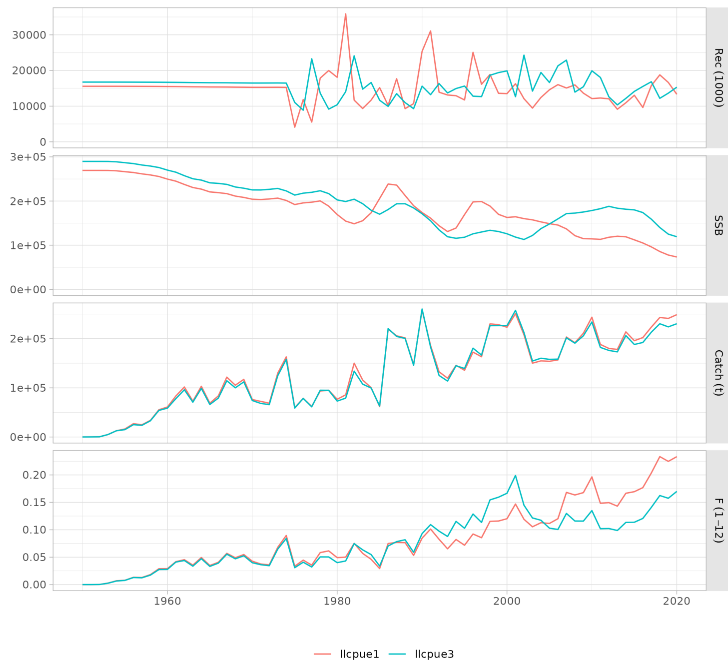

Figure 1: Summary trajectories for the two stock assessment model runs chosen by WPTmT in 2022.

Table 1: MASE values for those combinations of model run, area and season for which MASE < 1.

Model	Area	Season	MASE
LLCPUE1	**NW**	**2**	0.7465
LLCPUE1	**NW**	**3**	0.8837
LLCPUE3	**NW**	**2**	0.6955
LLCPUE3	**NW**	**3**	0.9190
LLCPUE3	NE	2	0.9546
LLCPUE3	SW	3	0.8970

The runs test for the runs employing the NW and SW indices are presented in Figures 2 and 3 respectively. Results for the LLCPUE1 index are shown by quarter in the top half, and those for LLCPUE3 in the bottom half. While for the NW model some series in both areas 1 and 3 present reasonable residual patterns (Figure 2), only two out of four indices in area 3 pass the test even when the model, SW, is set to fit to them (Figure 3).

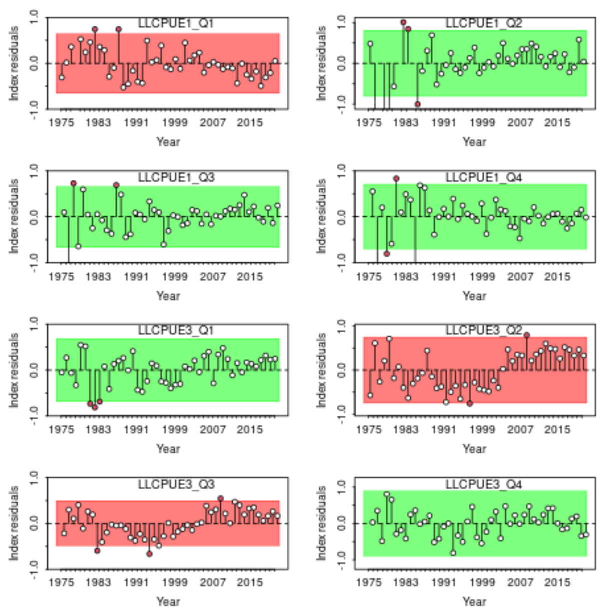

Figure 2: Runs tests of the NW (LLCPUE1) and SW (LLCPUE3) quarterly abundance indices for model fitting to the NW series.

The same test has been applied to the length-frequency data from the same two areas, NW and SW, and for the two models (Figures 4 and 5, respectively).

A test has been carried out of an updated operating model grid based on the new stock assessment SS3 model runs, was carried out. A partial factorial design was employed at this point to limit the computational workload to 120 model runs. As in the previous iteration, three selection criteria were applied to the model runs, which removed from the set any SS3 run that:

- Estimated virgin biomass (B_0) to be larger than 3.5 kt (three times the base case estimate), or current stock status relative to B_{MSY} to be greater than three, for a total of 11 runs.
- Convergence was deemed not to have been achieved (convergence level > 1e-4), which no model failed this time.
- MASE of both chosen indices of abundance (NW LLCPUE1 seasons 2 and 3) was larger than 1, which applied to 22 runs.

Of the 120 models in the partial factorial design grid, 99 passed the three tests and could be employed in the final OM. Figure 6 shows a comparison of the distributions of three estimated values meant to convey the scale, initial simulation status and overall variability of both grids. It appears mostly extreme values are removed from the grid while the distributions do not vary greatly.

The effect of some grid factors and levels on the determined stock status is presented in Figure 7.

The OM obtained from selecting grid runs as presented above, then resampled with weights based on the

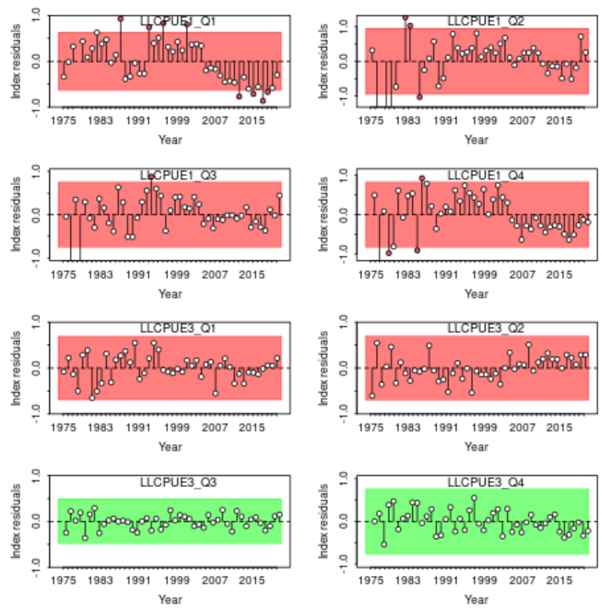

Figure 3: Runs tests of the NW (LLCPUE1) and SW (LLCPUE3) quarterly abundance indices for model fitting to the SW series.

p-value of the Diebold-Mariano test (Diebold and Mariano, 1995), is compared in Figure 8 with the trajectories from the two stock assessments used by WPTmT for this stock.

Two documents were presented to WTPmT this year relevant to the OM conditioning for albacore. One covered the status, problems and lessons learned from the previous OM grid (Mosqueira, 2022). The second (Hillary & Mosqueira, 2022) presented a first draft implementation of the conditioning procedure based on an Approximate Bayesian Computation algorithm. After some discussion on the merits and possible problems of the suggested approach, the WPTmT “AGREED to support the proposed methodology as the basis for the update of the albacore Operating Model and REQUESTED the WPM to further discuss the technical details of this method” (IOTC, 2022).

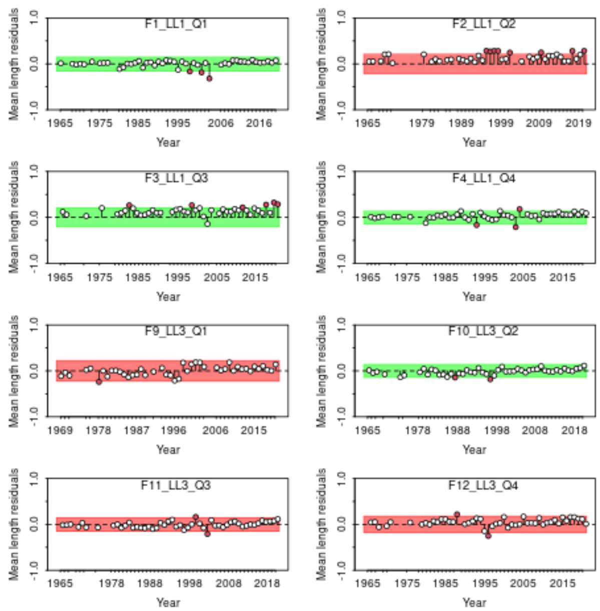

Figure 4: Runs tests of the length-frequency data for the NW (LLCPUE1) and SW (LLCPUE3) fleets fitting to the NW series.

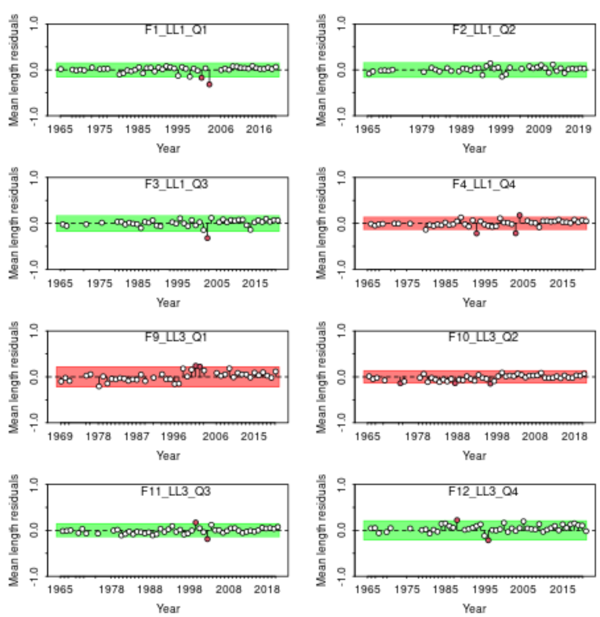

Figure 5: Runs tests of the length-frequency data for the NW (LLCPUE1) and SW (LLCPUE3) fleets fitting to the SW series.

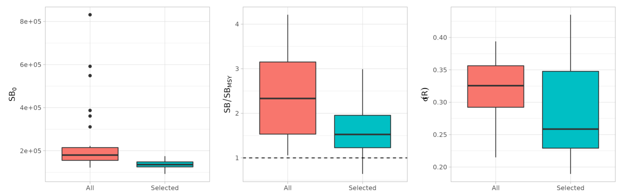

Figure 6: Distribution of values for three estimated values (virgin biomass, SB_0, biomass status, SB/SB_{MSY}, and variability, $\sigma(R)$, between the complete partial grid and the runs selected according to the three criteria.)

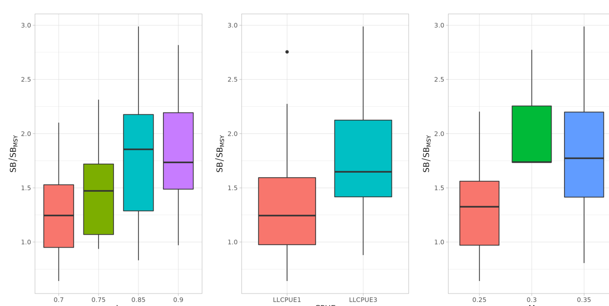

Figure 7: Distribution of values for current biomass status, SB_{2020}/SB_{MSY}, by level across three factors in the grid, SRR steeness (h), choice of index of abundance (CPUE), and natural mortality level (M).

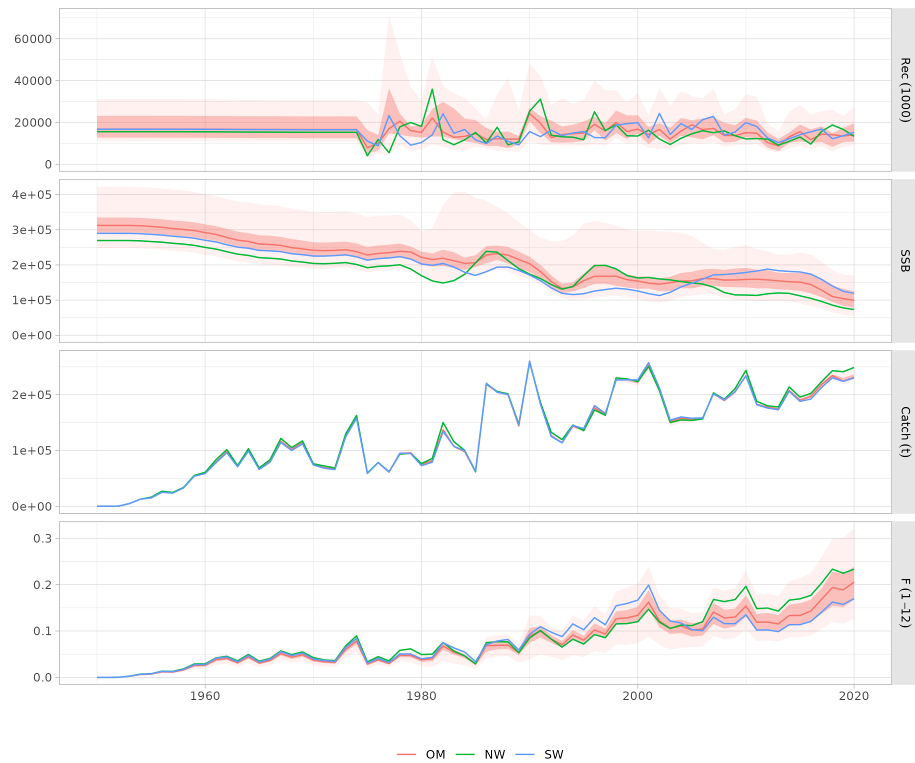

Figure 8: Time series of the main metrics (recruitment, SSB, catch and mean fishing mortality) for the full partial-factorial OM grid (in red with 50 and 80% credibility intervals)

Swordfish

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 perfomance of various MPs, specially on the performance statistic used for tuning: the probability of the stock falling in the green quadrant of the Kobe plot in the 2034-2039 period (Figure 9, top central panel).

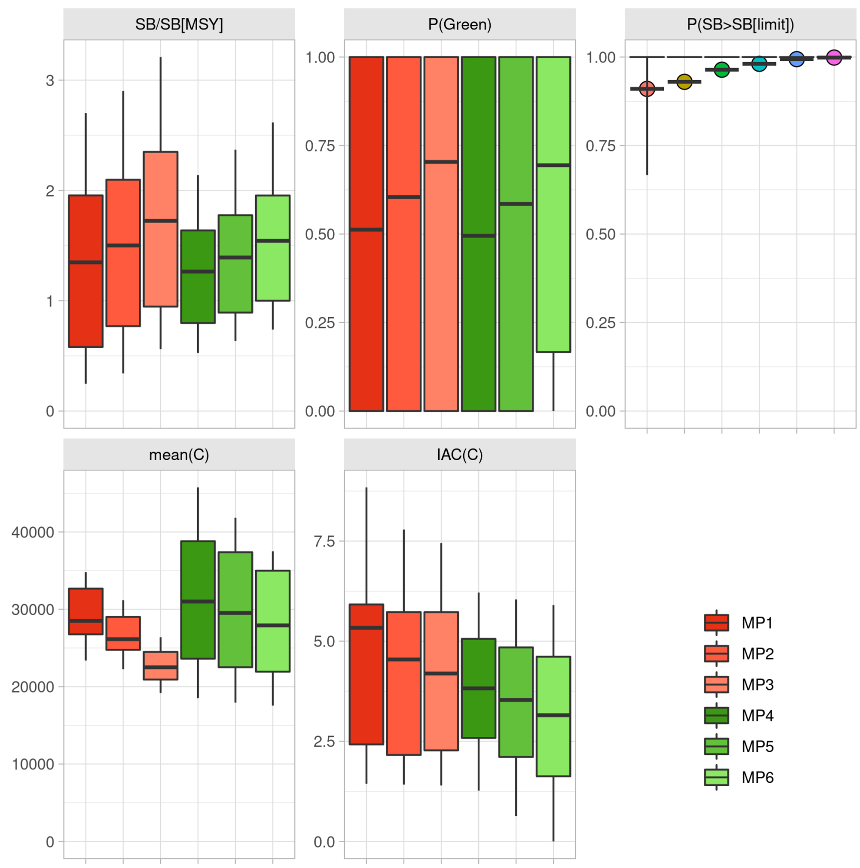

Figure 9: Boxplots of the five main performance statistics for a range of MPS applied to the swordfish OM.

Although the mean values, showed as a horizontal black line, show that the tuning criteria have been met, the 80% quantiles extend over the full range. To explore if a reduction in the spread of the OM grid could lead to a reduced uncertainty coverage in stock status, and have an effect in this statistic plot, an alternative OM grid was proposed and tested. 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 estimated virgin biomass (B0) in the new grid does not differ markedly from the previous one (Figure 10), and has no effect in the distribution of the Kobe green performance statistic.

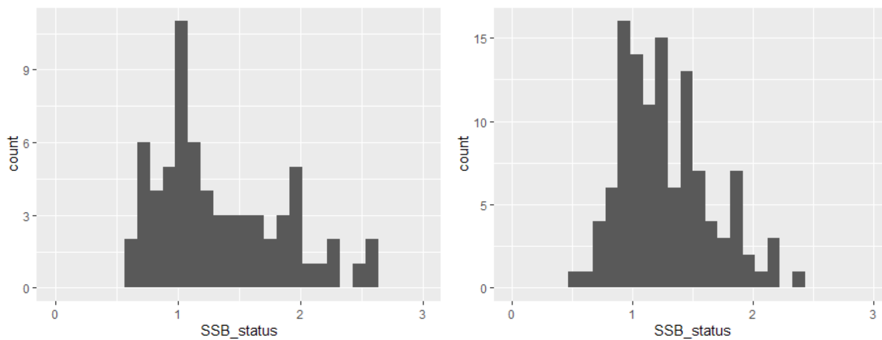

Figure 10: Distribution of the estimated virgin biomass (B0) on the previous OM grid (left) and new one (right).

In the case of swordfish, the starting point in the simulation period is for a stock very close (62%) to the range of probability values chosen for tuning (50, 60 and 70%). A reduced number of OM trajectories are thus required to change exploitation status level for those desired probabilities to be achieved. The calculation of the performance statistic over a five year period also limits the range of possible values over which the average is computed. Combined with the relatively large uncertainty in stock status presented by this OM, this appears to explain the unexpected performance plot for this statistic. A question could be raised on whether the boxplot for the probability of being in the Kobe green should be included in the standard set. The main result of interest is whether the tuning procedure has succeeded in achieving the requested levels, and this can done more briefly by reporting the mean probability only. This would also avoid the current discrepancy, as the boxplots for the other four chosen statistics are centered around the median rather than the mean.

The proposed alternative OM was discussed by the experts at WPB. There was a general agreement that the two factors of the uncertainty grid that had little influence on the OM starting condition could be removed from the grid. However there was no consensus about the choice of an appropriate range for the recruitment steepness value. For a part, it was felt that values as high as 0.9 were potentially appropriate for the more productive tropical tunas, but likely not for swordfish. On the other hand, it was argued that, due to swordfish females high fecundity, a high steepness value was also plausible. Considering the decision made at ICCAT, where a maximum steepness of 0.9 is used for swordfish, the decision was made to keep using the same range of 0.6-0.9 as previously done.

Discussion

The open questions presented above and that affect the operating models for both albacore tuna and swordfish, should be discussed and settled by WPM to allow the developers to progress on the evaluation of management procedures.

The first one is on the choice of methodology for the next version of the albacore OM. The recent WPTmT showed its support to decoupling the OM conditioning from stock assessment fitting by adopting an alternative methodology, like the presented ABC (Hillary & Mosqueira, 2022). Although it is not clear at this point if the same issues of instability and problems in forecasting encountered in the previous OM will still apply to a grid based on the 2022 stock assessment, the general interest on separating both processes remains. Albacore is likely to prove an easier stock in which to develop this methodology for any possible wider application to IOTC stocks. Possible refinements to the current grid could also be considered if the group identifies any issue that could improve its robustness.

The WPTmT assessment session has requested WPm to consider the use of prediction skill, and particularly the MASE statistic on the indices of abundance to filter and weight model runs on the grid. Reservations have been raised on removing runs from the OM when the MASE for the selected index or indices is above one. The logic of this criteria has been discussed before, bt we would welcome a firm decision from WPM on the matter.

Regarding the suggested changes to the swordfish OM grid, we would request WPM to discuss, amend and/or endorse the proposal made to WPB, which simplifies slightly the grid without apparent changes in estimated uncertainty.

Finally, the suitability of the current graphical presentation of the probability of being in the Kobe green could be considered, as it does not seem to add substantial information. An alternative performance statistic, that provides a more informative comparison among management procedures, could be proposed to substitute it.

References

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

Diebold, F.X. and R.S. Mariano. (1995). Comparing predictive accuracy. Journal of Business and Economic Statistics, 13: 253-63.

Hillary, R., Mosqueira, I. 2022. Exploring a wider approach to OM conditioning in IOTC MSE work. IOTC-2022-WPTmT08-10.

Kitakado, T., Wang, S-P., Matsumoto, T., Lee, S.I., Satoh, K., Yokoi, H., Okamoto, K., Lee, M.K., Lim, J.-H., Kwon, Y., Tsai, W.P., Su, N.J., Chang, S.T., Chang, F.-C. 2022. Joint CPUE indices for the albacore Thunnus alalunga in the Indian Ocean based on Japanese, Korean and Taiwanese longline fisheries data. IOTC–2022-WPTmT08(DP)-15.

Mosqueira, I. 2022. Conditioning an operating model for Indian Ocean albacore. IOTC-2022-WPTmT08-05.

Rice, J. 2022. Stock assessment of albacore tuna (Thunnus alalunga) in the Indian Ocean using Stock Synthesis. IOTC-2022-WPTmT08-09.