**Technical Report of the 2018 ICCAT Bluefin Tuna and North Atlantic Swordfish MSE Meeting**

Tom Carruthers (bluemattersci@gmail.com)

2nd July 2018

# Introduction

The Atlantic Bluefin Tuna Management Strategy Evaluation (ABT-MSE) process has entered a phase where various SCRS scientists and managers are developing candidate management procedures (CMPs) based on a proposed set of relative abundance indices. The 2018 ICCAT bluefin tuna and North Atlantic swordfish MSE meeting was convened to facilitate this process by providing a venue to engage with developers of CMPs and receive feedback on the current MSE software package.

The meeting covered various subject areas relating to these goals including: technical feedback on the current MSE software; a review of the current suite of CMPs; agreement on tuning specifications to facilitate comparison of future results; input from stakeholders on performance metrics and CMPs; suggested changes to the software package and operating models; and the future work program.

Feedback from SCRS scientists and stakeholders included a number of concerns, for example: that the rate of stock mixing was being overestimated by the operating models; the Western scenarios do not encompass the predictions of the Western stock assessment in terms of stock size and trajectory; and that the current operating model reports do not include enough detail to fully understand their behavior and estimation. In addressing these comments a number of changes to operating model structure, assumptions and interpretation of data were suggested that are detailed below.

A full meeting report was written by Dr Butterworth the BFT MSE chair and is included as Appendix A to this document. Rather than reiterate that general and comprehensive report, this document distils the technical aspects as they relate to the design and features of the ABT-MSE operating models and software package. These issues are rephrased as recommended changes and suggested additions to the MSE package (v2.8). This report also highlights any relevant changes in the current package (v3.1) and the corresponding Trial Specifications document (Appendix B).

# Interpretation of data

## Stock of origin data

An important area of feedback related to the apparently high degree of stock mixing estimated by the operating model. There are two principal causes. The first relates to the stock assignment data that often provide inference of Eastern Fish in Western areas (including the GOM) and Western fish in Eastern areas (including the Mediterranean) (see ‘Raw’ rows of Figure 1). The suggested approach of ignoring all assignments in the interval 30-70% Eastern (Western < 30%, Eastern > 70%) does not fully address this issue (Table 1 & Figure 1). A solution was developed that treated the assignments in the Natal spawning areas to be 100% stock of origin and then calculated the stock composition of assignments in mixed areas based on these stock-specific signatures (Mixture distribution). This approach recognizes the GOM and Mediterranean as stock exclusive and greatly reduced the inference of mixing, particularly of western fish in eastern areas (see draft SCRS paper Appendix C).

Table 1.The assignment scores (% Eastern) by ocean area for methods using just the otolith microchemistry data among the raw, the 30-70 cut-off and the new mixture distribution approach. The % raw data is lower than 100% for the raw approach since genetics assignment data are not included. The mixture distribution is lower than this again as it excludes any strata with fewer than 5 observations (too few to characterize a distribution). The mixture distribution approach assumes that GOM and MED fish are 0% and 100% Eastern respectively.



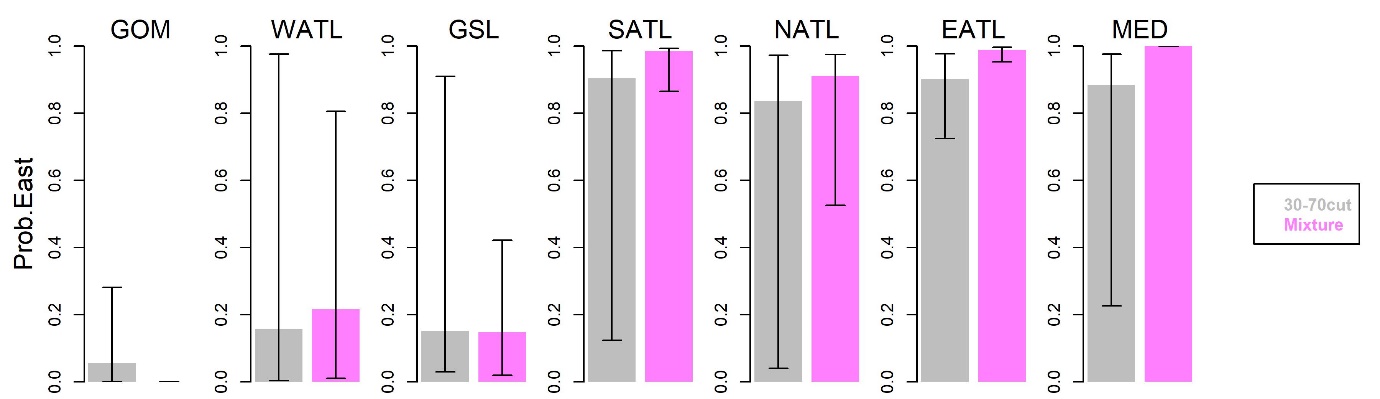
****

Figure 1. The assignment scores (of Table 1) by oceans areas for the 30-70 cut-off approach and the mixture model. Black bars represent the 5th and 95th percentiles. For the mixture model these are approximate, and based on the assumption that the square root transformation used in equation (2) provides distribution normality.

The second source of the apparent overestimation of stock mixing appears to arise due to the patchy observation of stock assignment data and the default assumption of the age-based movement model. In version 2.8, the model estimates fully mixing (mimicking a spatially aggregated assessment model) and estimates age-based movement where data are available. The new approach for v3.1 is to model movement of all age classes to be the same, borrowing information from other age classes that have data to impute the movement of age classes with no data. In this configuration, the model does not default to full mixing.

## Auto-correlation accounted for in the likelihood functions for relative abundance indices

In many cases the operating model fits indices acceptably but there are runs in residuals where several years of predicted indices are lower than, or higher than those observed. While this property is already used to generate future index data with similar attributes, this statistical phenomenon has not be accounted for properly in the operating model fits to data (the equations of the likelihood functions).

## Indices of abundance

Following revisions to the Western assessment, the US rod and reel CPUE index for fish over 177cm should be included.

Also to be included is the Japanese longline CPUE index in the Gulf of Mexico that ran from 1974-1981, over which time there was inference of stock declines.

The group also noted that the French Aerial Survey should be divided into two time periods 2000-2003 and 2009-2016.

The Canadian Combined index should be removed and replaced with two separate indices for the South West Nova Scotia (West Atlantic) and the Gulf of St Lawrence (GSL).

# Changes to operating model structure and assumptions

## Start the SCA before 1983

The M3 operating model has two time periods, the later time period is similar to a traditional statistical catch-at-age (SCA) stock assessment and aims to fit catch observations, catch composition data and relative abundance indices. The early time period assumes the same selectivity as the early years of the SCA but given the lack of catch composition data, conducts a stock-reduction analysis (SRA) where catches are removed exactly assuming selectivities are the same as the early SCA period. Version 2.8 ran the SRA from 1864-1982 and the SCA from 1983-2015. The proposed changes would start the SCA earlier to encompass indices (e.g. GOM Japanese longline indices) and data that may indicate a decline in stock levels during the period from 1970 onwards (v3.1 ends the SRA in 1964 and starts the SCA in 1965).

## Steepness of 0.6 for Western stock

The intention of the various recruitment scenarios for the western stock was to encompass the so-called high and low recruitment scenarios. It was considered that the current Beverton-Holt fits (level 2 of the recruitment factor) did not represent this particular uncertainty (steepness was estimated to be high and similar to the other scenario of 0.98) and instead it was suggested that a lower level of recruitment compensation should be user-specified (steepness = 0.6, included in v3.1) for level 2 – the ‘high recruitment’ scenario.

## Stock-recruitment relationships from '87 onwards

It was noted that the latter period of stock-recruitment started too late and should start from 1987 to include years where there were significant shifts in recruitment strength (included in v3.1)

## R0 historical estimation

Version 2.8 estimates mean recruitment over the whole SCA time period including biannual (2 year blocks) deviations from this mean recruitment. Stock-recruitment relationships were then estimated post-hoc (similarly to Virtual Population Analysis). This configuration led to difficulties in estimating the relevant MSY reference points and did not allow users to fix historical levels of recruitment compensation (steepness could not be user-specified). To address this version 3.1 includes a reconfiguration of the stock-recruitment model in M3 that estimates unfished recruitment R0 (for multiple time periods) given user specified steepness.

## Hockey stick model with fixed hinge points

Version 2.8 of the package did not correctly recreate the ‘hockey stick’ stock recruitment assumptions of previous assessments (stock recruitment factor levels 1 and 3). The hinge point (spawning biomass below which recruitment declines) is now fixed in the model to the mean of model estimated spawning biomass for years 1990-1995.

## Movement defaults to age-invariant

As discussed above, the model should be restructured to borrow movement information from age classes where there are stock of origin or electronic tagging data so that the model does not default to a fully mixed scenario. One approach to this is to estimate the movement of all age classes and then model age-class-specific deviations from this (a hierarchical modelling approach).

## Move to a simplified spatial model 7-area

An alternative, simpler spatial model was suggested as a sensitivity analysis that combines the Gulf of Mexico and Caribbean areas, two North Atlantic and two South Atlantic regions into single Gulf of Mexico, North Atlantic and South Atlantic regions (Figure 2). This spatial arrangement is far simpler and greatly reduces the sparseness of data improving model stability.

### 

Figure 2. Spatial definitions tabled by the 2015 ICCAT data preparatory meeting (left panel) with simplification to a single Mediterranean area. An alternative seven-area model is also presented that merges Gulf of Mexico and Caribbean areas, the central and south sub areas of the North and South Atlantic.

Note that the justification for explicit spatial structure is that there are different trends in total biomass or stock composition between areas over time. If this is not the case an aggregate area provides a good approximation. It follows that any justifications for a more complex arrangement than the 7 area model should originate from data demonstrating such trends.

## Include the correct time-lags for the various types of simulated data

Currently all data are lagged by 2 years: catches and indices are available up to two years before the current year that an MP is applied. This is to be revised such that catch data are lagged by 3 years and for the most recent two years the TAC will be assumed to be have taken exactly.

## Electronic tags released in the Bay of Biscay

The M3 model can only estimate movement based on electronic tags of known stock of origin. In version 2.8 these were limited to those tags that were in the natal spawning areas of the Gulf of Mexico or the Mediterranean. Following feedback from Spanish scientists it was considered that any tags released in the Bay of Biscay (the AZTI East Atlantic electronic tags) can be assigned Eastern Stock of Origin.

# Performance metrics

## Calculation of quantities relating to Maximum Sustainable Yield

The calculation of BMSY (and BMSY relative to unfished) for version 2.8 used the mean recruitment estimated over the time period of the Statistical Catch at Age (SCA) model to infer unfished stock size. This is to be revised in the move towards a stock-recruitment model where recruitment compensation (steepness) is fixed at user-defined values and unfished recruitment R0, is estimated.

## Dynamic B0 / BMSY calculation

The reference operating models include changes in stock productivity that entail changes in stock depletion (unfished biomass changes) and MSY quantities (MSY, BMSY, FMSY). It is necessary to be able to phrase performance of management procedures in terms of a representation of what could be achieved given current stock dynamics. A suitable means of doing this was suggested in which an unfished stock is calculated as changes occur in productivity – essentially maintaining a track of stock size with zero catches – so called ‘dynamic B0’. The same approach could be used to track dynamic BMSY by assuming that BMSY is a constant fraction of unfished biomass.

## Average annual variation in yield – negative TAC adjustments only

It was pointed out that stakeholders are probably less concerned with increases in TACs than decreases and hence a suitable performance metric could be the variability in downward TAC adjustments rather than the average annual variability in catches (added to v3.1)

# Presentation and Reporting

## Better characterization of spatial model predictions

Users of the ABT-MSE package requested further detail on the spatial predictions of the operating models such as the fraction of the Western Stock in Eastern areas and vice versa. These additional plots are necessary to diagnose changes in model predictions following the reparameterization of the movement model and the new stock of origin inferences arising from the mixture model analysis (Appendix C). Any such plots should also communicate among-simulation variability within the same operating model.

## Cryptic biomass

Spatial-seasonal multistock models differ from spatially aggregated annual models in that it is possible for the modelled fish population to exist in time-area strata where there is limited fishing. Participants noted that it would be useful to be able to quantify the extent of this cryptic biomass when reporting operating model behavior and estimates. It was thought that this might partly explain why operating model estimates of Western stock biomass are higher than the current assessment models.

An additional form of cryptic biomass occurs due to the declining size selectivity of fishing gears. This is potentially important as VPA stock assessments have shown considerable sensitivity to the assumed value of the ‘Fratio’, a parameter that controls the relative fishing mortality on older individuals in the plus group.

## Dynamic B0 plots

The concept of dynamic B0 can be hard to communicate and it is also difficult to understand how stock productivity changes impact the existing spawning biomass (e.g. how quickly a large unfished stock would declined given a reduction in stock productivity). These plots should show stock-specific shifts in dynamic B0 for Eastern, Western and the Atlantic-wide population.

## Redo OM reports

The existing automatic operating model reports should be overhauled to include the new spatial diagnostics, and dynamic B0.

## Standardized automatic MSE / CMP reporting

New automatic reports for multi-MP MSE analyses and also one for a specific CMP should be developed so that all developers can compare standardized outputs. These should include the new figures, diagnostics and performance metrics identified in this report arising from the Bluefin Tuna and North Atlantic Swordfish MSE Meeting.

## Spatial / stock-specific reporting

The group agreed that catches should be reported by area (East / West management areas) as this is how management is organized but depletion should be reported by stock.

## Statistical properties of simulated indices

In projection years, the MSE framework generates simulated relative abundance indices subject to error, auto-correlation and non-linearity (hyperdepletion / hyperstability). A more comprehensive presentation of these values and their derivation should be provided in future operating model reports so that the group can better understand whether these are credible.

# Alternative operating models (uncertainties)

## Robustness test for increasing gravity in GSL

It has been hypothesized that the inability of the model (and Western assessment models) to fit the gradual and pronounced increase in catch rates in the Gulf of St Lawrence may be due to a spatial shift northwards. To recreate such a phenomenon one option is to create an operating model where the gravity weight of the GSL increases over time. It was concluded that this could be discussed as a possible robustness operating model during a ‘second round’.

## New catchability reference case

The existing catchability scenarios of the robustness model include a possible 1% annual increase in fishing efficiency that could lead to CPUE indices that are overly optimistic. The group decided that a 2% increase was more appropriate (roughly doubling catch rates per fishing mortality rate every 50 years for the same vulnerable biomass).

## Split the Mediterranean Larval index

A possible discontinuity in the Mediterranean Larval index such that earlier observations (much lower) should be considered to have a different catchability than the most recent observations. The result will likely be the inference of an Eastern stock that is increasing less strongly over the last two decades.

## Recalculation of the Master Index

The Master Index (Carruthers 2017) is used to calculate effective effort for all spatial strata for all fleets by dividing observed catches by a complete prediction of spatio-temporal abundance. Although this prior inference is weak it initialized the model in a credible spatio-temporal distribution, greatly speeding up model convergence and stability. This index is calculated using a range of fleet-specific CPUE data and the standardized CPUE indices / relative abundance indices that the operating model is fitted to. This index must be recalculated including the changes to indices listed above.

## Temporally constant movement

The group discussed the assumption of temporally constant movement as estimated by the operating model (Carruthers et al. 2016). While stock-specific movement is fixed in the model over time, mixing is not and fluctuates according to the magnitude of the two stocks and the component age-classes. Rather than increase the dimensionality of the reference set of operating models it was decided to add additional movement scenarios to the robustness set including a scenario with halved mixing (e.g. where 40% of eastern fish are found in the west, this is reduced to 20%) and the increasing GSL gravity scenario described above.

# Trial specifications

## Details on the observation model

Greater details on how data are generated should be included, in this case the annual catch observations and the relative abundance indices.

## Details on the gravity model

The movement modelling is parameterized as a simple gravity model that has been subject to simulation testing that is now referenced (e.g. Carruthers et al. 2011).

# Other discussion points

## Operating model predictions versus assessments

The group acknowledged that a spatial, seasonal, multistock operating model is likely to provide varying estimates of stock size, productivity and depletion compared with single area, discrete East-West management area assessments. However it was considered that the second axis abundance should have three levels (of report by Dr Butterworth): “

1. Best estimate OM fit. If this implies large differences with the accepted assessments, the reason(s) for the differences should be identified.
2. The trends and scales in SSB resulting from OM conditioning for both East and West are simultaneously forced to follow the results of the 2017 stock assessments closely in terms of both absolute magnitude and trend (the final assessments agreed by SCRS in 2017 should be used for this). This should help identify the reasons for any possible differences identified in A.
3. This is similar to scenario A but including some broad constraints to prevent the results of the OM conditioning from diverging from the current general knowledge of past stocks dynamics. The Group considered that it would be appropriate to require that the results of the OM for both East and West BFT show that they were overfished at some point in the past. This means not just spawning stock biomass being lower than SSBmsy, but also a low relative SSB level in certain time periods. Preventing SSB increases during past periods of high catches may also be considered and should be clearly explained if it is included in the scenario. These ideas are meant to reflect public perception of BFT being at low level (particularly in the east) around the turn of the century. The GBYP modelling expert was given flexibility here, depending on outcomes found from various explorations.”

## Need for a streamlined MP development guide

Guidelines for CMP development are current included in the ABT-MSE package manual but these are included in Section 7 of that document and it may be better to create a shorter, more focused guide to MP development in a separate document.

## Discontinuities in harvest control rules

Harvest control rules often include abrupt changes in management advice with declining estimates of spawning biomass. For example, the 40-10 control rule maintains FMSY fishing until spawning biomass reaches 40% of unfished levels at which point fishing linearly decreases as a fraction of FMSY to zero at 10% of unfished biomass. Such rules are problematic as they can lead to high variability in TAC advice in response to natural fluctuations in stock level and observation errors. CMPs should therefore avoid such control rules where possible.

## Imposed limits on TAC variability

Version 2.8 of the package imposed maximum rates of TAC change of +/-20%. This guideline of the trial specifications was intended for CMP developers and should be removed as a restriction codified into the MSE framework.

## Tuning specifications for CMP comparison

In order to provide a level playing field for comparing the behavior among CMPs, a standard approach in other MSE settings is to tune CMPs so that they provide comparable performance in projections. Proposed here is the tuning of CMPs so that median (across simulations) depletion after 30 projected years is the same. It may be possible to include an optimization function in the package to achieve this.

## Mimicking the F0.1 fishing of Conventional stock assessments

Where possible it is useful to compare CMPs to ‘status quo’ performance. However this requires that status-quo approaches (e.g. a Stock Synthesis assessment with fishing at F0.1. Methot and Wetzel 2013; Anon 2018) can be made sufficiently fast, robust and reproducible, and that the data they require can be credibly simulated by the MSE. The former has proven very difficult and the latter is challenging in this context as existing F0.1 interpretations of assessments would require the simulation of age and length composition data. It follows that this has been noted as an area for future development but is of lower priority. A possible solution to the former problem would be to characterize the possible bias and error in estimates of stock size and F0.1 by doing retrospective analysis of stock assessments. Then a pseudo-F0.1 scenario could be simulated with similar biases / error, but would have to assume the latest assessment gave unbiased predictions of stock size and F0.1 for previous years.

# Acknowledgements

This report was developed under the provision of the ICCAT Atlantic Wide Research Programme for Bluefin Tuna (GBYP), funded by the European Union, several ICCAT CPCs, the ICCAT Secretariat and by other entities (see: http://www.iccat.int/GBYP/en/Budget.htm). The contents of this paper do not necessarily reflect the point of view of ICCAT or other funders and in no ways anticipate ICCAT future policy in this area.

References

Anon. 2018. Report of the 2017 ICCAT bluefin stock assessment meeting. Col. Vol. Sci. Pap. ICCAT, 74(6): 2372-2535.

Carruthers, T.R. 2017. Calculating population-wide spatial and seasonal relative abundance indices for Atlantic Bluefin tuna for use in operational modelling. Col. Vol. Sci. Pap. ICCAT, 74(6): 2586-2595.

Carruthers, T.R., Kimoto, A., Powers, J., Kell, L., Butterworth, D., Lauretta, M. and Kitakado, T. 2015b. Structure and estimation framework for Atlantic bluefin tuna operating models. ICCAT SCRS/2015/179.

Carruthers, T.R., McAllister, M.K., Taylor, N.G. 2011. Spatial surplus production modelling of Atlantic tunas and billfish. Ecological Applications. 21(7), 2734-2755.

Methot, R.D., Wetzel, C.R. 2013. Stock Synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142: 86-99.

Appendix A: Report of the 2018 ICCAT Bluefin Tuna and North Atlantic Swordfish MSE Meeting by D.S. Butterworth

Attached.

Appendix B: Current Trial Specifications document

Attached.

Appendix C. Draft SCRS paper on revaluation of stock assignment data using a mixture model

Attached.