SIMULATION TESTING A MULTI-STOCK MODEL WITH AGE-BASED MOVEMENT

Tom Carruthers[1] and Laurence Kell[2]

SUMMARY

Data were simulated from a multi-stock, multi-fleet, spatial, seasonal population dynamics to evaluate estimation performance of an operating model with age-based movement. Preliminary analysis suggests potential serious estimation biases that may be addressed by model restructuring, addition of new data, use of alternative likelihood functions or weighting schemes.

KEYWORDS

Stock assessment, simulation, migrations, population dynamics, seasonal variations, tuna fisheries, tagging, fishery management

[1] IOF, 2202 Main Mall, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z4. t.carruthers@fisheries.ubc.ca
[2] International Commission for the Conservation of Atlantic Tunas, Calle Corazón de María, 8, 28002 Madrid, Spain

1 Introduction

A Management Strategy Evaluation (MSE, Butterworth 1999, Cochrane 1998, Punt et al. 2014) approach has been proposed for Atlantic bluefin tuna (SCRS 2013) as a suitable framework for providing robust management advice consistent with the precautionary approach (GBYP 2014). A principal task in the construction of an MSE framework is the development of operating models which represent credible hypotheses for population and fishery dynamics. Operating models are typically fitted to data to ensure that model assumptions and estimated parameters are empirically credible (Punt et al. 2014, e.g. CCSBT 2011).

Prior to fitting an operating model to real data it is necessary to validate the model by simulation testing. Validation provides a check for coding errors, appropriate weighting of various likelihood functions and also reveals model instability and estimation bias / precision (Deroba et al. 2014). In this paper we generated data from a simulated multi-stock, multi-fleet, spatial, seasonal population dynamics model to which we fitted the latest version of the M3 (v0.18) operating model.

2 Methods

The model closely follows the equations described in Carruthers et al. 2015a (M3 operating model v0.15). Several modifications to the operating model were suggested following a meeting of the Core Modelling Group in Monterey (Jan 2016). The most important changes were:

(1) a shift to age-based movement in which individuals of three discrete different age classes (e.g. ages 0-3, 4-8, 9+) exhibit important differences in migration;

(2) initialization of the model at equilibrium fishing levels consistent with the average estimated over the first five model years;

(3) modelling of a ‘plus group’ whereby individuals over a particular age (e.g. 25) aggregated in a single age class.

(4) implement a 1 year lag between spawning biomass and recruitment (ie recruitment is predicted from spawning biomass in spawning areas in the spawning season of the previous year).

This new operating model (v.0.18) was programmed in the non-linear optimization software ADMB. A test unit was also developed in the statistical environment R that uses identical equations and dimensions (Table 1) and simple observation error models for the most common data types (Table 2).

The core differences between the test unit and the operating model relate to simplification that are necessary to rapidly fit operating models to potentially sparse data: (i) the operating model uses a movement model parameterized as a gravity model; (ii) the operating model estimates 5-year blocks of recruitment deviations. The simplified movement model aims to generalize spatial distribution rather than individually estimate all the movements from areas to areas which is how the data were simulated. Since the model is a statistical catch-at-length model, information about annual recruitment is smeared through the inverse age-length key.

The parameter ranges for the test unit are described in Table 3. A total of 64 simulations were undertaken in which a different level of each parameter was sampled from these distributions. Data were then simulated subject to an observation error model (also described in Table 3).

The age based movement was simulated to encapsulate a plausible hypothesis for Atlantic bluefin tuna dynamics (Figure 1, represents the equilibrium population distribution for each of these age classes). Age class 1 refers to a relatively sedentary juvenile stage (ages 0-3) that largely remain in a dedicated spawning area. Age class 2 are highly mobile adult fish (ages 4-8) that mix in two central areas during the second subyear but remain in their dedicated spawning area in subyear 1. The oldest age class (ages 9+) have intermediate mobility that is less seasonal.

The operating model was fitted to data using two weighting schemes for the various likelihood components of the global objective function (Table 4). Since the same random seed is preserved between runs so while the number of simulations is relatively low (64 in this preliminary simulation test) the results are directly comparable among likelihood weighting schemes.

Bias was evaluated in estimates of five variables and parameters of interest: (i) the fraction of spawning stock biomass in the spawning area during the spawning season (to check estimation of movement); (ii) current stock depletion (spawning stock biomass relative to unfished); (iii) current population biomass; (iv) current exploitation rate and; (v) unfished spawning biomass.

3 Results

Bias in estimates of spawning stock biomass (bottom left plot, Figure 2) and spawning biomass distribution (top left panel, Figure 2) were underestimated by 5% for stock 1 (less spawning biomass in spawning area than simulated) and over estimated by 10% for stock 2 (more spawning biomass in spawning area than simulated). In general the range of biases for these estimated variables was much lower than the other quantities (depletion, current biomass and current fishing mortality rate). In general the capacity of the model to provide biased estimates is relatively high and indicates the need for further model development or the inclusion of other data sources. The solution may be unrelated to data weighting: the alternative weighting scheme provided very similar estimation biases.

4 Discussion

Simulation testing is an invaluable tool in both checking that an estimation model is coded correctly but also for tuning the various aspects of the model to improve estimation performance. In this preliminary simulation test of the latest operating model (M3 v0.18) it is clear that estimation performance is relatively poor and there is a need for further exploration of model structure, weighting of data, prescription of likelihood functions, parameterization of movement processes. Simulation evaluation provides a principled and transparent means of developing more robust and accurate models.

An important avenue for development are additional data sources. For example there is interest in incorporating indices of spawning biomass such as larval surveys (Ingram et al. 2015). These may strongly constrain model estimation and improve the robustness of the operating model. Various movement models are available including a detailed Markov movement model (a probability of moving from each area to each area in each time step), a gravity model (fractions in each area with a viscosity parameter to limit stock mixing) and a fractional model (a fully mixed stock in which a fixed fraction of individuals are found in each area in each time step). It is important to simulation test operating models using a test unit that generates these types of dynamics to understand what estimation scheme is robust to uncertainties (ie a 3 x 3 factorial simulation – estimation test).

When occasionally catches are simulated that are occasionally very small, the log-normal observation error model of the operating model is essentially scale-less and can lead to strong negative bias in estimates of stock size. This is alleviated by using a normal (or least squares) likelihood component for catch observations. It is important therefore to evaluate the correct error structure for the real catch data to avoid this problem. Similarly a number of observation processes are assumed to follow multinomial model such as the length composition data and the electronic tag track data. It may be beneficial to investigate alternative likelihood functions such as a multivariate logistic function that are less likely to dominate the global objective function and overly strongly determine model fitting (particularly considering the likely quality of the catch composition data for bluefin tuna).

The simulation framework developed here allows for variability in a wide range of biological, fishery and observation processes (Table 3). Many other processes were not varied in this preliminary simulation test such as trends in growth, mortality, fishery selectivity, hyperstability in indices and patchiness in fishery data. This complexity allows for post-hoc evaluation of the processes most strongly linked to estimation bias. Perhaps biases were most prevalent where data were patchy or fishery size selectivity varied strongly between years. In such a case alternative model structures and likelihood weighting schemes may be considered.

It is essential to develop simulation for any unique estimation framework as the interaction of data and model structure are likely to be specific and not easily generalized. For example the findings of accurate estimation of stock trend but not absolute magnitude (Deroba et al. 2014) were not consistent in an age-based movement population dynamics model for grouper (Carruthers et al. 2015). As data for fitting operating models become available, the simulator should be adjusted to best reflect the quality and quantity of the data in order to make model adaptations that are most appropriate. This simulation exercise was relatively compact and included 45 years, 2 seasons, 4 areas, 2 stocks and 2 fleet types. The final operating model for Atlantic bluefin tuna is likely to be closer to 55 years, 4 seasons, 11 areas, 2 (or 3) stocks and 6+ fleets and therefore proposes a substantially different estimation problem. The primary role of an early simulation test such as this is to check for programming errors and map out the various axes for model development.

The simulation exercise described here is preliminary and should be improved by focusing on quantities that are most appropriate to management. For example accurate estimates of unfished stock size may be of less interest than estimates of current exploitation rate and biomass relative to MSY (i.e. an accurate Kobe plot).

5 Acknowledgements

This work was carried out by TC 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.

6 References

Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56, 985-998.

Carruthers, T.R., Walters, J.F., McAllister, M.K., Bryan, M.D. 2015. Modelling age-dependent movement: an application to red and gag groupers in the Gulf of Mexico. 72(8): 1159-1176. doi: 10.1139/cjfas-2014-0471

CCSBT. 2011. Report of the Sixteenth Meeting of the Scientific Committee. Bali, Indonesia. Commission for the Conservation of Southern Bluefin Tuna. Available at: http://www.ccsbt.org/userfiles/file/docs_english/meetings/meeting_reports/ccsbt_18/report_of_SC16.pdf

Cochrane, K L., Butterworth, D.S., De Oliveira, J.A.A., Roel, B.A., 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: experiences in the South African pelagic fishery. Rev. Fish. Biol. Fisher. 8, 177-214.

Deroba, J.J. et al. 2014. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES. J. Mar. Sci. doi: 10.1093/icesjms/fst237.

Ingram, G.W., Jr., D. Alvarez-Berastegui, P. Reglero, R. Balbín, A. García5, and F. Alemany. Indices of larval bluefin tuna (Thunnus thynnus) in the Western Mediterranean Sea (2001-2013). SCRS/2015/035

Punt, A.E., Butterworth, D.S., de Moore, C. L., De Oliveira, J. A. A., Haddon, M. 2014. Management strategy evaluation: best practices. Fish and Fisheries. doi: 10.1111/faf.12104

Table 1. The dimensions and estimated parameters of the test unit simulator and operating model

Model dimension		
Number of fleets	n_f	2
Number of areas	n_r	4
Maximum age (plus group)	n_a	25
Number of years	n_y	45
Number of subyears (seasons)	n_m	2
Number of stocks	n_s	2

Estimated parameter		**Number estimated**
Unfished recruitment	n_s	2
Length a modal selectivity	n_f	2
Precision of selectivity	n_f	2
Dome-shape of selectivity	n_f	1
Recruitment deviations	$n_s * n_y /5$	18
Fleet catchability	n_f	2
Movement	Up to: $(n_r\text{-}1) \cdot (n_r) \cdot n_m$	48
	Total	**74**

Table 2. Observation error models used in both simulation and estimation.

Type of data	**Disaggregation**	**Likelihood function**
Total catches (weight)	year, subyear, area, fleet	Log-normal
Index of vulnerable biomass (e.g. a CPUE index)	year, subyear, area, fleet	Log-normal
Length composition	year, subyear, area	Multinomial
Electronic tag (known stock of origin)	stock, year, subyear, area	Multinomial
Stock of origin	year, subyear, area	Multinomial

Table 3. Specification of simulations.

	Parameter / variable	Fixed value	Uniform LB	Uniform UB	Normal Mean	Normal CV
Biological	Depletion stock 1		0.05	0.25		
	Depletion stock 2		0.3	0.4		
	Inverse logit movement (inter simulation variability)				1	0.2
	Unfished recruitment stock 1		232k	449k		
	Unfished recruitment stock 2		38k	75k		
	Steepness stocks 1 and 2		0.354	0.65		
	Growth rate κ stock 1		0.087	0.091		
	Growth rate κ stock 2		0.091	0.095		
	Maximum length stock 1		313	317		
	Maximum length stock 2		318	341		
	Age at maturity stock 1		3.5	4.5		
	Age at maturity stock 2		8.5	9.5		
	Autocorrelation in recruitment deviations stocks 1 and 2		0	0.5		
	Variability in recruitment deviations (normal CV)		0.1	0.3		
Exploitation	Trajectory in current fishing mortality (fleet 1 and 2) % y^{-1}		-2	5		
	Ratio of current fishing mortality rate (F fleet 1 / F fleet 2)	1				
	Minimum length at maximum selectivity (fleet 1 and 2)		160	200		
	Selectivity at maximum length fleet 1		1	1		
	Selectivity at maximum length fleet 2		0.7	1		
	Number of spool up initialization years	20				
	Years of 'representative' exploitation used in spool up	5				
Obs. model	Bias in catches across simulation			none		
	Imprecision in catches (year, subyear, area, fleet)				1	0.2
	Number of annual catch at length observations		2000	5000		
	Number of electronic tagging transitions		1000	2000		
	Number of stock-of-origin observations		1000	2000		
	Hyperstability in master index			none		

Table 4. Weightings for likelihood components.

Data type	Catch	Length comp.	Stock of origin	Electronic tag tracks	Recruit. devs	Mov. par. penalty	Sel. Par. penalty
Default weighting	1	1/1000	1/10	1/10	1	1	1
Alternative weighting	10	1/100	1/10	1/100	1	1	1
Typical unweighted contribution max. posterior density.	-2k	680k	440	6k	30	30	30

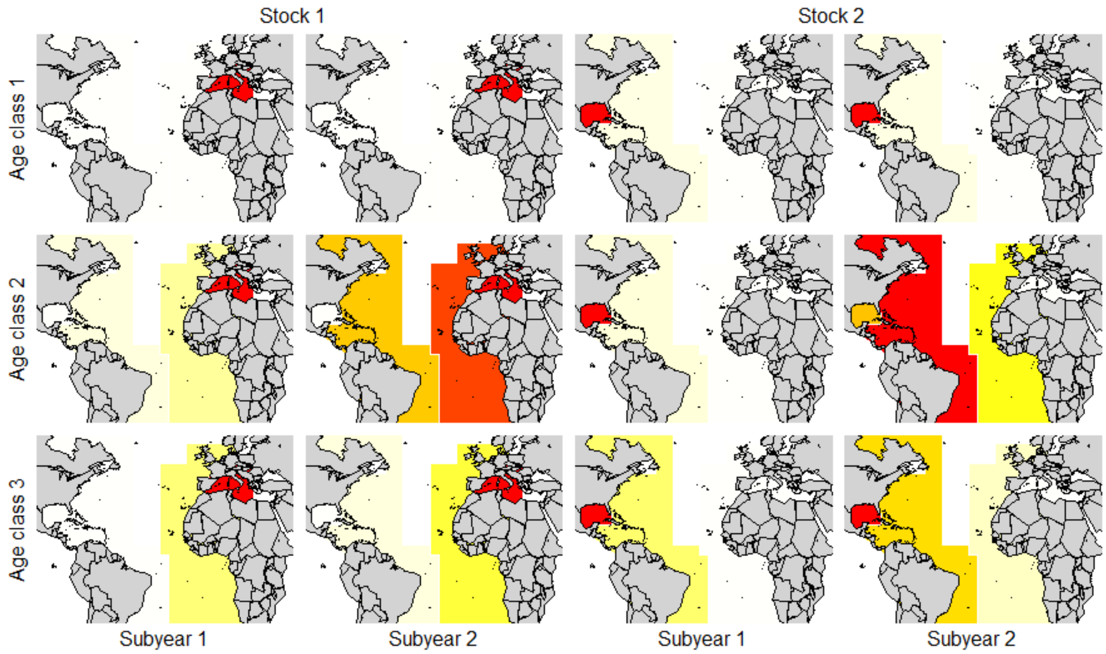

Figure 1. The hypothetic stock structure simulated in the test unit and operating model. Areas shaded red are those with largest stock numbers. Age class 1 refers to a relatively sedentary juvenile stage (ages 0-3) that largely remain in a dedicated spawning area. Age class 2 are highly mobile adult fish (ages 4-8) that mix in two central areas during the second subyear but remain in their dedicated spawning area in subyear 1. The oldest age class (ages 9+) have intermediate mobility that is less seasonal.

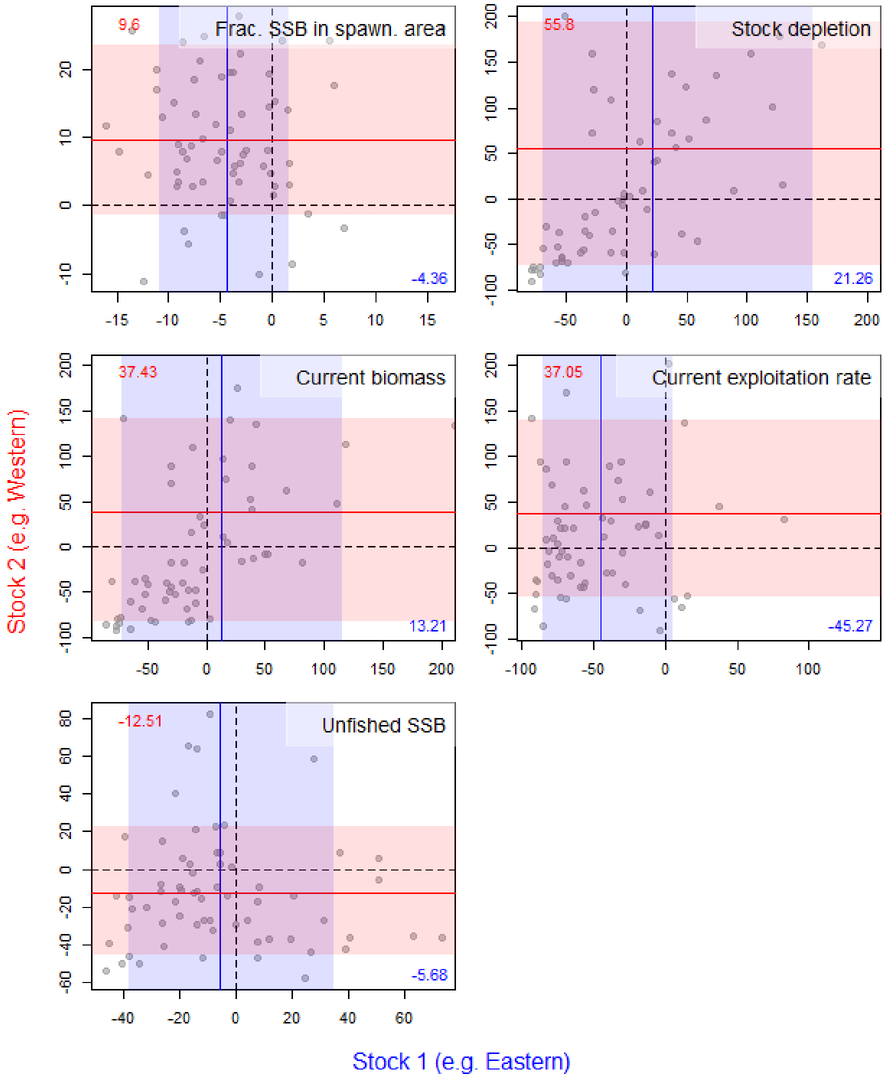

Figure 2. Estimation bias in model variables/parameters ((estimated-simulated)/simulated, expressed as a %) for data-weighting scheme 1 (Table XX). Dashed vertical and horizontal lines represent unbiased estimation. Vertical blue lines and numbers represent mean bias of stock 1. Horizontal red lines and numbers represent mean bias of stock 2. Shaded horizontal and vertical areas represent the range of bias among 10th and 90th percentiles.

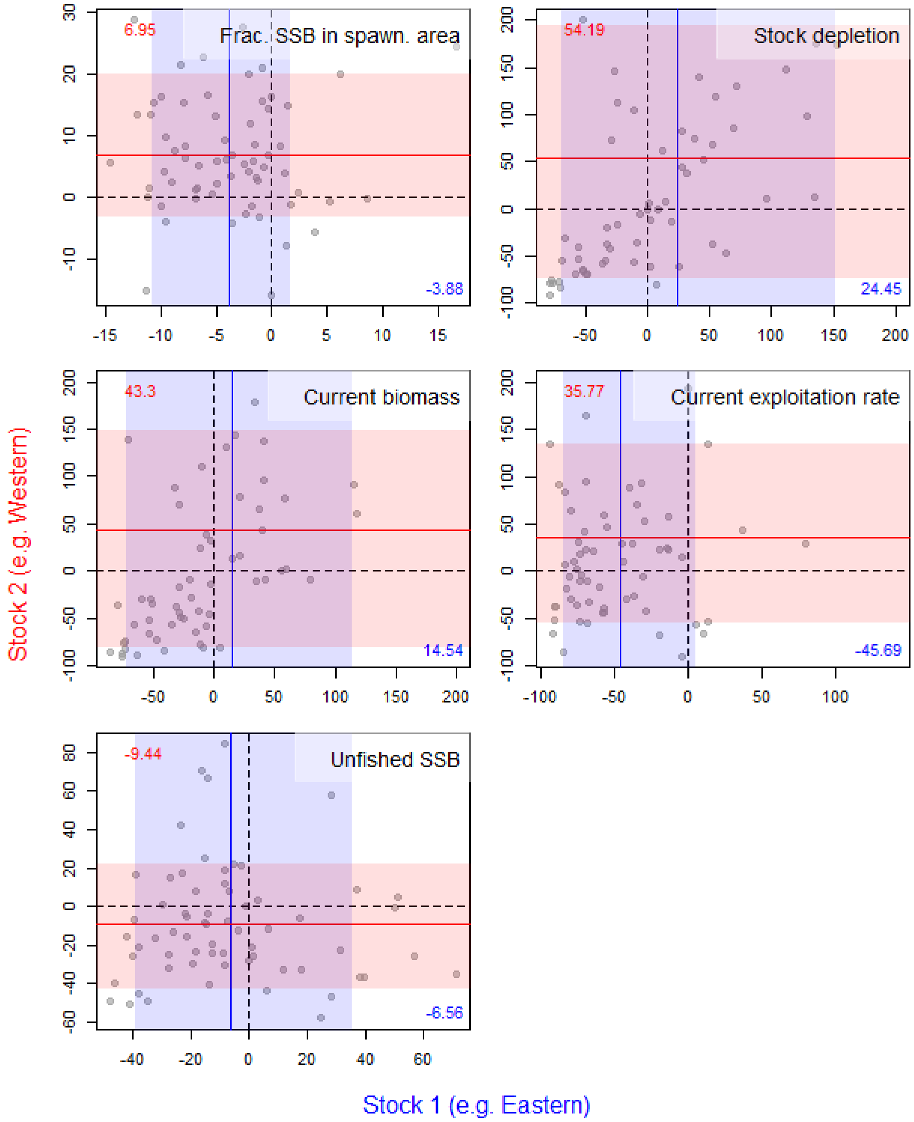

Figure 3. Estimation bias in model variables/parameters ((estimated-simulated)/simulated, expressed as a %) for data weighting scheme 2 (Table XX). Dashed vertical and horizontal lines represent unbiased estimation. Vertical blue lines and numbers represent mean bias of stock 1. Horizontal red lines and numbers represent mean bias of stock 2. Shaded horizontal and vertical areas represent the range of bias among 10th and 90th percentiles.

Figure 4. Estimation bias in model variables/parameters ((estimated-simulated)/simulated, expressed as a %) for data weighting scheme 3 (Table XX). Dashed vertical and horizontal lines represent unbiased estimation. Vertical blue lines and numbers represent mean bias of stock 1. Horizontal red lines and numbers represent mean bias of stock 2. Shaded horizontal and vertical areas represent the range of bias among 10th and 90th percentiles.