SCRS/2015/179 Collect. Vol. Sci. Pap. ICCAT, 72(7): 1782-1795 (2016)

STRUCTURE AND ESTIMATION FRAMEWORK FOR ATLANTIC BLUEFIN TUNA OPERATING MODELS

Tom Carruthers[1], Ai Kimoto[2], Joseph E. Powers[3], Laurence Kell[4],
Doug S. Butterworth[5], Matthew V. Lauretta[6], Toshihide Kitakado[7]

SUMMARY

A preliminary spatial, multi-stock statistical catch-at-length assessment model is developed as a basis for defining operating models for Atlantic bluefin tuna. The modifiable multi-stock model (M3) aims to improve upon previous multi-stock models such as MAST (Taylor et al. 2011) in three core areas. The first iteration of the model: (1) makes use of indices of abundance specific to time-area strata (e.g. for a given ocean area and month of the year), (2) does not use conventional tagging data to inform exploitation rates, (3) is fitted to samples of length composition data and therefore avoids established problems related to ageing individuals based on a growth curve and length data only.

In this paper we provide a full account of preliminary M3 model equations and discuss the results of simulation evaluations of model estimation performance. Limitations of the current approach and future areas for model development are also discussed.

RÉSUMÉ

Un modèle d'évaluation préliminaire, spatial, statistique, multi-stock, de prise par taille est développé comme base pour définir des modèles opérationnels pour le thon rouge de l'Atlantique. Le modèle multi-stocks modifiable (M3) vise à perfectionner les modèles multi-stocks précédents, tels que MAST (Taylor et al. 2011) dans trois domaines principaux. La première itération du modèle: (1) utilise des indices d'abondance spécifiques aux strates spatiotemporelles (par exemple pour une zone océanique donnée et un mois de l'année), (2) n'utilise pas les données du marquage conventionnel pour apporter des informations aux taux d'exploitation, (3) est ajustée à des échantillons de données de composition par taille et évite donc les problèmes établis liés au vieillissement des spécimens sur la base uniquement d'une courbe de croissance et de données de taille.

Dans ce document, nous fournissons un compte rendu complet des équations préliminaires du modèle M3 et discutons des résultats des évaluations de simulation des performances d'estimation du modèle. Sont également discutées les limitations de l'approche actuelle et les futurs domaines de développement du modèle.

RESUMEN

Se desarrolla un modelo de evaluación preliminar espacial, multistock, estadístico de captura por talla como base para definir modelos operativos para el atún rojo del Atlántico. El modelo multistock modificable (M3) tiene como objetivo mejorar modelos previos multistock como el MAST (Taylor et al. 2011) en tres aspectos principales. La primera iteración del modelo: (1) utiliza índices de abundancia específicos de los estratos espacio-temporales (por ejemplo, para una determinada zona oceánica y mes del año), (2) no utiliza datos de marcado convencional para aportar información a las tasas de explotación, (3) está ajustado a muestras de datos de composición por tallas y por tanto evita problemas establecidos relacionados con la determinación de la edad de ejemplares basándose solo en una curva de crecimiento y en datos de tallas.

[1] IOF, 2202 Main Mall, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z4. t.carruthers@fisheries.ubc.ca
[2] NRIFSF. 5-7-1, Orido, Shimizu, Shizuoka, 424-8633. Japan. aikimoto@affrc.go.jp
[3] 8465 SW 141 Street, Palmetto Bay, Fl 33149, United States. jepowers@lsu.edu
[4] International Commission for the Conservation of Atlantic Tunas, Calle Corazón de María, 8, 28002 Madrid, Spain
[5] Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa. Doug.butterworth@uct.ac.za
[6] U.S. National Marine Fisheries Service, Southeast Fisheries Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL, 33149-1099, USA. E-mail: matthew.lauretta@noaa.gov
[7] Tokyo University of Marine Science and Technology, 5-7, Konan 4, Minato-ku, Tokyo 108-8477 Japan. kitakado@kaiyodai.ac.jp

En este documento se proporciona una descripción total de ecuaciones preliminares del modelo M3 y se discuten los resultados de las evaluaciones de simulación del rendimiento de estimación del modelo. Se discuten también las limitaciones del enfoque actual y campos futuros de desarrollo del modelo.

KEYWORDS

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

1 Introduction

The Atlantic-Wide Research Programme on Bluefin Tuna (GBYP) aims to develop a new scientific management framework by improving data collection, knowledge of key biological and ecological processes, assessment models and management. A critical component of the GBYP is the construction of a robust advice framework consistent with the precautionary approach (GBYP 2014). A Management Strategy Evaluation (MSE, Cochrane 1998, Butterworth 1999, Kell *et al.* 2014, Punt *et al.* 2014) approach has been proposed to address this goal (Anon. 2014b). MSE establishes operating models that represent credible hypotheses for population and fishery dynamics which are used to quantify the efficacy of various management procedures. These management procedures may encompass a wide range of complexity from conventional stock assessments linked to harvest control rules (Hilborn 2003) through to simple empirical management procedures that calculate catch limits directly from resource monitoring data indices (Geromont and Butterworth 2014a;b, Kell *et al.* 2015).

MSE applications generally develop operating models from stock assessments that are fitted to data in order to ensure that model assumptions and estimated parameters are empirically credible (Punt *et al.* 2014, e.g. CCSBT 2011). In the case of Atlantic bluefin tuna, such a model requires enough complexity to capture the core uncertainties regarding Atlantic bluefin tuna dynamics (Fromentin *et al.* 2014, Leach *et al.* 2014). These include stock structure (Kell *et al.* 2012), stock mixing, migration (Fromentin and Lopuszanski 2014) and biases in observed data (e.g. annual catch data). Additionally the model should be able to accommodate the wide range of data that have been collected for Atlantic bluefin tuna including catch rate indices (Abid *et al.* 2015, Hanke *et al.* 2015, Kimoto *et al.* 2015, Lauretta and Brown 2015, Santiago *et al.* 2015, and Walter 2015), aerial surveys (Bonhommeau *et al.* 2010), length composition data, larval surveys (Ingram *et al.* 2015), electronic tagging data (Block *et al.* 2005) and stock of origin data (Rooker *et al.* 2014).

In this paper we document a preliminary version of a spatial, multi-stock, statistical catch-at-length assessment model which we refer to as the modifiable multi-stock model or M3 (v1.03)(referred to as 'the model' herein). The most notable multi-stock model previously applied to Atlantic bluefin tuna was MAST (Taylor *et al.* 2011). In the development of M3 we aimed to address a number of central weaknesses of MAST and produce a credible, robust and faster assessment tool that can be applied to evaluate alternative hypotheses for Atlantic bluefin tuna rapidly. The purpose of this paper is to provide a full description of an early version of the model in order to illicit feedback from a wider group of scientists and stakeholders. This model version provides the basic model structure and anticipated data inputs in order to demonstrate a proof of concept and serve as a basis for discussing possible model improvements and features.

2 Methods

The model is based on conventional age-structured accounting (e.g. Quinn and Deriso 1999, Chapter 8) which is common to stock assessment models such as Stock Synthesis 3 (Methot and Wetzel 2013), CASAL (Bull *et al.* 2012), Multifan-CL (Fournier *et al.* 1998) and iSCAM (Martell 2015). Similar to these assessment packages, M3 is developed using ADMB (Fournier *et al.* 2012) for its rapid and robust non-linear estimation performance for problems with relatively large numbers of parameters (i.e. more than 100 parameters). The more challenging aspects of developing a multi-stock spatial model relate to the modelling of movement and initializing the model.

2.1 Estimated parameters

The majority of parameters estimated by the model relate to movement probabilities and annual recruitment deviations (**Table 1**). The number of estimated parameters can be reduced substantially by limiting estimation to only those movements that have been recorded or are considered credible. For example, given a quarterly time step (e.g. Jan-Mar, April-Jun etc.) and the spatial definitions of the 2015 data preparatory meeting (Anon. 2015, **Figure 1**), an evaluation of conventional tagging for Atlantic bluefin tuna data reveals that less than 80 parameters of the 224 possible are required to characterize all of the possible movements recorded by these tagging data.

2.2 Transition equations

The standard age-structured equations are complicated somewhat by the subyear temporal structure in which ageing and recruitment occur in a particular subyear. In this version of the model, spawning occurs for all stocks in a subyear *ms*, after subyear 1 (this is also likely to be the case in any final model fitted to bluefin tuna data since spawning in the Mediterranean and Gulf of Mexico is thought to occur after a period of movement early in the year).

Numbers of individuals *N*, for stock *s*, in a model year *y*, in the first subyear *m*=1, age class *a*, and area *r* are calculated from individuals that have moved $\vec{N}$, in the previous year, final subyear n_m, of the same age class subject to combined natural and fishing mortality rate *Z*:

1) $$N_{s,y,m=1,a,r} = \vec{N}_{s,y-1,n_m,a,r} \cdot e^{-Z_{s,y-1,n_m,a,r}}$$

where total mortality rate is calculated from annual natural mortality rate *M*, divided by the fraction of the year represented by the subyear t_m, and fishing mortality rate *F*, summed over all fleets *f*.

2) $$Z_{s,y,m,a,r} = \frac{M_{s,a}}{t_m} \sum_f F_{y,m,a,r,f}$$

Fishing mortality rate at age is derived from fishing mortality rate by length class *FL* and the conditional probability of fish being in length class *l*, given age *a* (an inverse age-length key, LAK).:

3) $$F_{y,m,a,r,f} = \sum_l FL_{y,m,l,r,f} \cdot LAK_{s,y,a,l}$$

The fishing mortality rate at length is calculated from an index of fishing mortality rate *I*, an estimated catchability coefficient *q* and a length selectivity ogive *s*, by fleet:

4) $$FL_{y,m,l,r,f} = q_f \cdot I_{y,f} \cdot s_{f,l}$$

Selectivity is calculated by the Thompson (1994) ogive and an estimate of mean length *L* of an age class *l*:

5) $$s_{f,l} = \frac{1}{1-s_{dome}} \cdot \left(\frac{(1-s_{dome})}{s_{dome}}\right)^{s_{dome}} \cdot e^{s_{prec} \cdot s_{dome} \cdot (s_{mode}-L_l)} \cdot \frac{1}{1+e^{s_{prec} \cdot (s_{mode}-L_l)}}$$

In the spawning subyear *ms*, aging and recruitment occur:

6) $$N_{s,y,ms,a,r} = \vec{N}_{s,y,ms-1,a-1,r} \cdot e^{-Z_{s,y,ms-1,a-1,r}}$$

Recruitment is currently assumed to occur in user-specified spawning area for each stock *rs*. Recruitment is assumed to follow a Beverton-Holt function of spawning stock biomass *SSB* in the defined spawning areas *rs* relative to unfished spawning stock biomass *SSB0* and is subject to annual recruitment deviations *R*, for each stock.

7) $$N_{s,y,ms,1,rs} = R_{s,y} \cdot \frac{0.8 \cdot R0_s \cdot h_s \cdot SSB_{s,y}}{0.2 \cdot SSB0_{s,y} \cdot (1-h_s) + (h_s - 0.2) \cdot SSB_{s,y}}$$

where *h* is the steepness parameter (fraction of unfished recruitment at 1/5 unfished spawning stock biomass) and spawning stock biomass is calculated from moved stock numbers in the subyear prior to spawning subyear *ms*, in spawning area *rs*, weight of individuals at age *w*, and the fraction of individuals mature at age *mat*:

8) $$SSB_{s,y} = \sum_a \sum_r \vec{N}_{s,y,ms-1,a,rs} \cdot e^{-Z_{s,y,ms-1,a,rs}} \cdot w_{s,a} \cdot mat_{s,a}$$

where weight is calculated from length at age *l*:

9) $$w_{s,a} = \alpha_s \cdot l_{s,a}^{\beta_s}$$

and fraction mature at age is assumed to be a logistic function of age with parameters for the age at 50% maturity γ, and slope ϑ:

10) $$mat_{s,a} = 1/\left(1 + e^{(\gamma_s - a)/\vartheta_s}\right)$$

Stock numbers for subyears that are not the first subyear of the year and are not the spawning subyear are calculated:

11) $$N_{s,y,m,a,r} = \vec{N}_{s,y,m-1,a,r} \cdot e^{-Z_{s,y,m-1,a,r}}$$

In each subyear, after mortality and recruitment, fish are moved according to a Markov transition matrix *mov* that represents the probability of a fish moving from area *k* to area *r* at the end of the subyear *m*:

12) $$\vec{N}_{s,y,m,a,r} = \sum_k N_{s,y,m,a,k} \cdot mov_{s,m,k,r}$$

The movement matrix is calculated from a log-space matrix *lnmov* and a logit model to ensure each row *k*, sums to 1:

13) $$mov_{s,m,k,r} = e^{lnmov_{s,m,k,r}} / \sum_r e^{lnmov_{s,m,k,r}}$$

Movements from an area *k* to an area *r* that are considered not to be credible (e.g. from the Eastern Mediterranean to the Gulf of Mexico) are assigned a large negative number (essentially zero movement). For each area *k*, from which individuals can move, the first possible value is assigned a value of zero; subsequent possible movements are assigned an estimated parameter ψ (since rows must sum to 1 there is one less degree of freedom):

14) $$lnmov_{s,m,k,r} = \begin{cases} -1E10 & no\ movement\ from\ k\ to\ r \\ 0 & first\ possible\ movement\ from\ k\ to\ r \\ \Psi_{s,m,k,r} & other\ possible\ movements\ from\ k\ to\ r \end{cases}$$

This movement formulation limits estimation to only those movements that are possible given the data (e.g. consistent with observed tagging data).

2.3 Initializing the model

Compared with spatially aggregated models, initialization is more complex for spatial models, particularly those that may need to accommodate movement by age and include regional spawning and recruitment. The solution used here is to iterate the transition equations above (Equations 1, 6, 7, 11, 12) given zero fishing mortality until the spatial distribution of stock numbers converges for each of the subyears.

Prior to this iterative process an initial guess at the spatial and age structure of stock numbers $\hat{N}$ is made using the estimated movement matrix and natural mortality rate at age *M*:

15) $$\hat{N}_{s,m,a,r} = R0_s \cdot \mathrm{e}^{-\sum_1^a M_{s,a}} \cdot \sum_k \frac{1}{n_r} \cdot mov_{s,m,k,r}$$

It typically takes between 50 and 100 iteration years of unfished conditions for stock numbers to converge to within 1/10 of a percent of the previous iteration. To ensure stability of the estimation, a fixed number of iterations is defined by the user.

2.4 Predicting data

For each fleet *f*, total predicted catches in weight $\hat{C}$, are calculated from the Baranov equation:

16) $$\hat{C}_{y,m,r,f} = \sum_s \sum_a w_{s,a} \cdot N_{s,y,m,a,r} \cdot (1 - e^{-Z_{s,y,m,a,r}}) \cdot \left(\frac{F_{y,m,a,r,f}}{Z_{s,y,m,a,r}}\right)$$

Similarly predicted catches in numbers at age *CAA*, is given by:

17) $$\widehat{CAA}_{s,y,m,a,r,f} = N_{s,y,m,a,r} \cdot (1 - e^{-Z_{s,y,m,a,r}}) \cdot \left(\frac{F_{y,m,a,r,f}}{Z_{s,y,m,a,r}}\right)$$

This can be converted to a prediction of total catches in numbers by length class *CAL* using a stock specific inverse age-length key, *LAK*:

18) $$\widehat{CAL}_{y,m,l,r,f} = \sum_s \sum_a \widehat{CAA}_{s,y,m,a,r,f} \cdot LAK_{s,y,a,l}$$

The model predicts spawning stock biomass indices $\widehat{Issb}$, that are standardized to have a mean of 1 for each stock over the total number of years *ny*:

19) $$\widehat{Issb}_{s,y} = n_y \cdot SSB_{s,y} / \sum_y SSB_{s,y}$$

The model predicts vulnerable biomass indices $\hat{I}$, by fleet that are standardized to have a mean of 1 for each fleet:

20) $$\hat{I}_{y,m,r,f} = n_y \cdot n_m \cdot n_r \cdot V_{y,m,r,f} / \sum_y \sum_m \sum_r V_{y,m,r,f}$$

Where vulnerable biomass *V* is calculated:

21) $$V_{y,m,r,f} = \sum_l \left(s_{f,l} \cdot \sum_s \sum_a \left(N_{s,y,m,a,r,f} \cdot ALK_{s,y,a,l} \cdot w_{s,a}\right)\right)$$

The model predicts stock of origin composition of catches $\widehat{SOO}$, from predicted catch numbers at age:

22) $$\widehat{SOO}_{s,y,m,r,f} = \sum_a \widehat{CAA}_{s,y,m,a,r,f} / \sum_s \sum_a \widehat{CAA}_{s,y,m,a,r,f}$$

2.5 Likelihood functions, priors and the global objective function

Table 2, summarizes the likelihood functions for the various data types. A log-normal likelihood function was assumed for total catches by fleet. The log-likelihood was calculated:

23) $$OBJ_c = \sum_y \sum_m \sum_r \sum_f \frac{log(\sigma_{catch}) + \left(\log(\hat{C}_{y,m,r,f}) - \log(C_{y,m,r,f})\right)^2}{2 \cdot \sigma_{catch}^2}$$

Similarly the log-likelihood component for indices of vulnerable biomass and spawning stock biomass were calculated:

24) $$OBJ_i = \sum_y \sum_m \sum_r \sum_f \frac{log(\sigma_{index}) + \left(log(\hat{I}_{y,m,r,f}) - log(I_{y,m,r,f})\right)^2}{2 \cdot \sigma_{index}^2}$$

25) $$OBJ_{SSB} = \sum_s \sum_y \frac{log(\sigma_{SSB}) + \left(log(\widehat{Issb}_{s,y}) - log(Issb_{s,y})\right)^2}{2 \cdot \sigma_{SSB}^2}$$

The length composition data are assumed to be distributed multinomially. In traditional stock assessment settings catch composition data may often dominate the likelihood function due to the large number of observations. This is exacerbated by a failure to account for non-independence in size composition samples. There are two possible solutions: (1) manually specify the effective sample size (ESS) of length-composition samples or (2) use a multinomial likelihood function that includes the conditional maximum likelihood estimate of the ESS (perhaps even a freely estimated ESS, S. Martell personal communication). In this version of the code, ESS is user-specified.

The log-likelihood component for length composition data is calculated:

26) $$OBJ_{CAL} = -\sum_y \sum_m \sum_l \sum_r \sum_f CAL_{y,m,l,r,f} \cdot log(\hat{p}_{y,m,l,r,f})/ESS_f$$

Where the model predicted fraction of catch numbers in each length class *p*, is calculated:

27) $$\hat{p}_{y,m,l,r,f} = \widehat{CAL}_{y,m,l,r,f} / \sum_l \widehat{CAL}_{y,m,l,r,f}$$

Similarly the log-likelihood component for PSAT tagging data of known stock of origin (SOO), released in year *y*, subyear *m*, area *r* and recaptured in year *y2*, subyear *m2*, and area *k* is calculated:

28) $$OBJ_{PSAT} = -\sum_s \sum_y \sum_m \sum_{y2} \sum_{m2} \sum_r \sum_k PSAT_{s,y,m,y2,m2,k} \cdot log(\hat{\theta}_{s,y,m,y2,m2,r,k})$$

where recapture probabilities θ, are calculated by repeatedly multiplying a distribution vector *d*, by the movement probability matrix *mov*. For example for a tag released on a fish of stock 1 in year 2, subyear 3, and area 4, the probability of detecting the tag in year 3, subyear 2 for the various areas is calculated:

29) $$\hat{\theta}_{s=1,y=2,m=3,y2=3,m2=2,r=4,1:n_r} = \left(\left(d \cdot mov_{s,m=3}\right) \cdot mov_{s,m=4}\right) mov_{s,m=1}$$

where

30) $$d_k = \begin{cases} 0 & k \neq r \\ 1 & k = r \end{cases}$$

The log-likelihood component for PSAT tagging data of unknown stock of origin PSATu, is currently weighted according to the compound probability that a fish is of a particular stock given the track history for that tag. For example for a tag *t*, tracked in series of years y_i, subyears m_i, and regions r_i, the weight *w*, of that tag for a specific stock is calculated:

31) $$w_{t,s} = \frac{\prod_i[(\sum_a N_{si,yi,mi,ai,ri})/(\sum_s \sum_a N_{si,yi,mi,ai,ri})]}{\prod_i[1-(\sum_a N_{si,yi,mi,ai,ri})/(\sum_s \sum_a N_{si,yi,mi,ai,ri})]}$$

This is simply the product of fractions of that stock in those time-area strata divided by the product of the fractions of other stocks in those time-area strata. An alternative approach would be to compare the relative probabilities of the observed movements among the stocks although it is unclear whether this circularity (PSAT data are a primary source of information regarding movement) could lead to estimation problems.

The weighted likelihood function is similar to that of the stocks of known origin but includes the appropriate weighting term for each tag

32) $$OBJ_{PSATu} = -\sum_t \sum_s \sum_y \sum_m \sum_{y2} \sum_{m2} \sum_r \sum_k PSATu_{t,s,y,m,y2,m2,k} \cdot log(\hat{\theta}_{s,y,m,y2,m2,r,k}) \cdot w_{t,s}$$

The log-likelihood component for stock of origin data *SOO* was also calculated assuming a multinomial distribution:

33) $$OBJ_{PSATu} = -\sum_{\boldsymbol{s}} \sum_y \sum_m \sum_r \sum_{\boldsymbol{f}} SOO_{s,y,\boldsymbol{m},\boldsymbol{r},\boldsymbol{f}} \cdot log(\widehat{SOO}_{s,y,m,r,f})$$

In addition to these likelihood functions for observed data, priors may be placed on the steepness parameter *h*, of the stock recruitment relationship and a factor *Mfac*, multiplied by the user specified natural mortality rate-at-age schedule *Minit*.

34) $$M_{s,a} = Minit_{s,a} \cdot Mfac_s$$

The factor applied to the natural mortality rate-at-age schedule is assumed to be lognormally distributed according to user specified mean and standard deviation parameters.

35) $$OBJ_M = \sum_s \frac{log(\sigma M_s) + (Mfac_s - \mu M_s)^2}{2 \cdot \sigma M_s^2}$$

Steepness is parameterized by a logit model constrained between 0.2 and 1:

36) $$h_s = 0.2 + 0.8 \cdot e^{\hat{h}_s} / \left(1 + e^{\hat{h}_s}\right)$$

In the logit^{-1} space, a normal prior is adopted for this transformed steepness $\hat{h}$, parameter that includes user specified mean $\widehat{\mu h}$, and standard deviation $\widehat{\sigma h}$, parameters. The corresponding log-likelihood component is:

37) $$OBJ_h = \sum_s \frac{log(\sigma h_s) + \left(\hat{h}_s - \mu h_s\right)^2}{2 \cdot \sigma h_s^2}$$

The global objective function OBJ_T, to be minimized is the summation of the weighted , likelihood components:

38) $$OBJ_T = \omega_c \cdot OBJ_c + \omega_i \cdot OBJ_i + \omega_{SSB} \cdot OBJ_{SSB} + \omega_{CAL} \cdot OBJ_{CAL} + \omega_{PSAT} \cdot OBJ_{PSAT} + \omega_{PSATu} \cdot OBJ_{PSATu} + \omega_M \cdot OBJ_M + \omega_h \cdot OBJ_h$$

3 Simulation evaluations

The demonstration MSE framework previously presented to the GBYP Core Modelling Group (December 2014, Anon. 2014) was used to simulate data to determine whether the model could estimate quantities of interest such as stock depletion, current stock size and spatial distribution reliably. In this preliminary simulation evaluation, 200 datasets were simulated with varying stock depletion, exploitation history, gear selectivity, movement and spatial distribution. The simulation was kept relatively simple and included only two fleet types, 4 areas, 2 sub-years and 40 historical years of exploitation (**Figure 2** illustrates the simplified mixing and spatial structure that borrows four areas from the spatial definitions of Anon. 2015, Figure 1). A summary of inputs and parameter ranges for the simulations is included in **Table 3**.

This simulation evaluation was intended as a proof of model concept and consequently did not simulate biases in observed data (e.g. catch reporting, non-independence in length composition data) or evaluate model misspecification (e.g. incorrect aggregation of fleets, misspecification of selectivity, incorrect natural mortality rate at age, functional form of the stock recruitment relationship). The simulation results presented here are for a former version of the M3 (v1.02) model which does not include priors for steepness and natural mortality rate. In this simulation test these were assumed to be known perfectly without error.

Simulation testing reveals that the model provides estimates of stock depletion, stock size and spatial distribution that are not strongly biased (**Figure 3**). For example, biases in estimates of stock depletion were on average within 2% of unbiased for both simulated stocks. The range of biases was also relatively low with standard deviations among simulations of 5.3% and 8.3% for stocks 1 and 2, respectively. Estimates of current stock size were somewhat negatively biased (around -5% for both stocks) but not strongly so. Among stocks, biases in estimates of current stock size were negatively correlated. This is to be expected since an overestimate of the size of stock 1 is likely to be paired with an underestimate of stock size 2, as the model aims to generate a similar total vulnerable biomass to that simulated. Among the simulations the biases in estimates of current stock size could be larger for stock 2 (a standard deviation of 10%) than stock 1 (a standard deviation of 6%) which should be anticipated since stock 2 is estimated to be smaller than stock 1 (unfished recruitment was 1/6 that of stock 1).

4 Discussion

4.1 Limitations

In this paper we document the first version of the M3 model which is designed primarily to outline the basic framework in order to initiate a dialogue with the GBYP core modelling group and wider SCRS regarding priorities for operating model development. In order to get a version of the model working and simulation tested, a number of features were omitted. One of the most important was the conventional tagging data which are not used to estimate exploitation rates in this preliminary M3 model.

It is relatively simple to include the model code to predict the dynamics of a tagged population of bluefin tuna and predict capture probabilities. However there are concerns that uncertain and variable reporting rates serve to contaminate these data which therefore could provide a misleading picture of movement and exploitation rate (and hence stock size). The confounding of reporting rate and fishing mortality rate estimates is a known problem in the use of conventional tagging data when no information exists to inform reporting rate estimates, as is currently the case for bluefin tuna. Previous spatial, multi-stock models of bluefin tuna such as MAST (Taylor *et al.* 2010) have assumed similar reporting rates among fleets. There is however evidence that reporting rates may vary widely among fleets, for example between 1/1000 and 1/2 (Carruthers and McAllister 2010). Thus, the model may be confronted with observed recapture rates and observed catch rates that differ by a factor of 500 for some fleets. The likely result is a poorly defined objective function, model outputs that are sensitive to initial values and a failure to satisfy convergence criteria.

While the benefits of adding conventional tagging data are uncertain, the additional computations are likely to be greater than 200% for a model such as M3. It may be possible to include conventional tagging data and estimate the fleet-specific reporting rates but the benefits would be weak additional information regarding movements and exploitation rates at the cost of a much more computationally intensive model. Despite these potential drawbacks and the omission of these data in this first version of the model, conventional tags should still be considered as a source of information about exploitation rate and this should be a primary subject of discussion. There are a number of other ways in which conventional tagging data can be incorporated into the model such as characterising growth, informing estimates of fleet selectivity and vulnerability-at-age, and defining the full range of possible movements. Recaptures of conventional tags from observer-based programs provide one potential source of unbiased estimates as 100% reporting can be assumed. For example, the pelagic longline observer programs in the West Atlantic may provide unbiased estimates of rates of recapture.

Currently the model does not attempt to model movement by age or length. Other approaches have aimed to estimate a separate set of movement parameters for juvenile and mature fish (e.g. MAST). This is a priority for model development. An alternative to estimating a separate movement model for certain ages or length classes would be to model continuous models of regional gravity or mixing rate with age (e.g. Carruthers *et al.* 2015). This has the benefit of adding the same number of parameters but having smooth transitions in movement among age classes rather than an abrupt shift to an alternative movement model at a particular life-stage.

Similar to other statistical catch-at-age and catch-at-length models which approximate historical fishing dynamics according to a finite number of fleets, the M3 model requires at least one selectivity curve to be either user-specified or assigned a ‘flat-topped’ (e.g. logistic) selectivity curve. The aggregation of fleets should also be considered carefully to avoid overly complex models with redundant fleets or those that have not contributed substantially to the exploitation of one or more of the stocks. Alternative selectivity assumptions, aggregations of fleets or temporal definitions of fleets may serve as alternative hypotheses to be taken into account in the operating models of a future MSE. Conventional tagging data may provide information on fleet-specific selectivity for some fishing fleets that operate in areas where a large number of tags have been released, e.g., the bait boat fisheries in the Bay of Biscay or the rod-and-reel fisheries off of North America.

4.2 Strengths and opportunities

Since all of the catch composition data for bluefin tuna are in the form of length samples, statistical catch-at-age models require these sampled lengths to be converted to ages. The problems associated with this practice are well established (Allioud *et al.* 2014, Hilborn and Walters 1992, Kell and Kell 2011). A central strength of the model proposed here is that it is fitted to these length samples directly. A statistical catch-at-length approach also provides avenues for accounting for varying growth rates among individuals which may be informed by conventional tagging data.

Once a range of plausible hypotheses has been identified for bluefin tuna and suitable operating models have been developed to represent these, there are potentially a large number of research questions that may also be addressed in addition to MSE. For example: to what extent is movement estimation biased by release location of PSAT tags? How should data be weighted in a spatial, integrated assessment model? What is a suitable experimental design of a genetic tagging program? When assessing populations with complex stock structure and highly migratory dynamics, what assessment model complexity is particularly important: space, age, both or neither? If explicit performance metrics are available, can suitable harvest control rules be derived?

An advantage of not integrating conventional tagging into the model is that it runs relatively quickly (e.g., in less than a few minutes). This may allow a simplified version of the approach to be included within a management procedure in future MSE analyses, broadening the range of complexity in management procedures. Rapid model fitting also allows for extensive simulation testing which confirms that the model has been programmed correctly and reveals potential areas for re-parameterization or simplification.

A number of model features are in development to allow for alternative hypotheses for bluefin tuna dynamics. Once such extension is an approximation to a Growth-Type Group model (GTG) via calculation of the inverse age-length key. A common oversight in fisheries stock assessment models is the inability to account for downward shifts in the expected length composition at age which occur as larger, faster growing individuals experience higher exploitation rates. Preliminary simulation evaluations using models that simulate many (300+) growth type groups, indicates that fishing mortality rate estimates can be highly inaccurate (as much as 200% biased) when this phenomenon is not taken into account. The conventional approach (also considered for the MAST model, Taylor *et al.* 2011) is to model discrete groups of individual of varying growth parameters. In Stock Synthesis these are referred to as 'platoons' (Methot and Wetzel 2013).

The principal problem with these approaches is that it can take many extra growth type groups (300+) or several platoons to generate suitably smooth predicted length compositions. This is a major problem for models such as M3 that are already computationally intensive as each GTG or platoon is an additional dimension, and hence the number of calculations is increased by a factor equal to the number of GTGs or platoons. A solution under investigation here is an approximation that takes percentiles of the distribution of GTGs and uses linear interpolation to predict the shift in the length structure given historical fishing mortality rates. This is a much more tractable approach as it adds no additional dimensionality to the transition calculations: additional calculation is limited to the construction of the inverse age-at-length key. Furthermore, conventional tagging data may be used as an empirical source of growth data for fitting the inverse age-length key. This additional feature would add around 10% more computation time per iteration.

4.3 Priorities

In this paper, simulation testing was cursory and should be much more thorough for future releases of the model. For example future tests should examine various types of model misspecification such as ignoring differences in juvenile/mature movement and misspecification of maturity. Problematic observation processes may also be simulated: for example persistent biases in annual catches, spatially biased length sampling and relative abundance indices that are non-linearly related to abundance. Model diagnostics in future simulation tests could also include other checks such as of the biases in MSY-related reference points.

The M3 model differs from previous multi-stock model such as MAST in that it requires indices of abundance (e.g. standardized CPUE indices) for fleets by time-area strata (e.g. for a given ocean area and subyear). The core advantage of this is that the movement estimation is constrained to combinations of parameters that are consistent with other spatial data. Previous spatial modelling has demonstrated that these data alone are sufficient to estimate spatial heterogeneity reliably. This is more important than estimating specific movement transitions. For example to estimate MSY-related reference points reliably, it is more important to know the spatial distribution of the stock at a given time of year than to know exactly what fraction of individuals moved to and from the various areas (Carruthers *et al.* 2011). This means that the development of fleet-specific indices at the resolution of subyear and spatial area is a central priority for operating model development.

In this paper we reference the spatial definitions of the 2015 data preparatory meeting (Anon. 2015, Figure 1). Kimoto *et al.* (2015) recommend that the ICCAT Bluefin Working Group should carefully examine spatial stratifications particularly those of the northeast Atlantic, a region that is currently a main fishing area for Japanese longline vessels. Historical abundance indices for the Japanese longline fleet are perhaps the most important data for the stock assessment of Atlantic bluefin tuna and are likely to be pivotal in fitting operating models. **Figure 1** originates from the bluefin mixing workshop in 2001 (Anon. 2002) which apart from the separation of the Gulf of Mexico and the Mediterranean, was not fully agreed upon by the group. Furthermore the reasoning behind the boundaries identified are not fully supported by knowledge accumulated over the last 15 years. It was recommended to revise the area stratification with new and updated information collected, in the light of both biological and fisheries aspects.

Version 1.03 of the M3 model includes user-specified priors for steepness (recruitment compensation) and the natural mortality rate factor *Mfac* (a multiplier of the user-specified natural mortality rate at age schedule). It is important to account for uncertainty in these parameters in operating models since these are among the least well known and most influential in the estimation of Atlantic bluefin tuna status and productivity. It follows that moving to a Bayesian version of the model using the Metropolis Hastings algorithm of ADMB is a future priority.

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 way anticipates future ICCAT policy in this area.

References

Abid, N., Faraj, A., De la Serna, J.M., Macías, D., Saber, S., Ortiz de Urbina J. 2015. Updated standardized joint CPUE index for bluefin tuna (*Thunnus thynnus*) caught by Moroccan and Spanish traps for the period 1981-2013. ICCAT, Col. Vol. Sci. Pap., 71(3): 1308-1312.

Ailloud, L.E., Smith, M.W., Then, A.Y, Omori, K.L., Ralph, G.M., Hoenig, J.M. 2014. Properties of age compositions and mortality estimates derived from cohort slicing of length data. ICES. J. Mar. Sci. 72(1): 1-6.

Anon. 2014a. Report of the 1st Meeting of ICCAT GBYP Core Modelling Group. Available at: http://www.iccat.int/GBYP/Documents/MODELLING/PHASE%204/BFTMSE_CMG1_Report.pdf [accessed September 2015]

Anon. 2002. Report of the 2010 ICCAT Workshop on Bluefin Mixing (Madrid, Spain, September 3 to 7, 2001). ICCAT, Col. Vol. Sci. Pap., 54(2): 261-352.

Anon. 2014b. Report of the 2013 Meeting on bluefin stock assessment methods (Gloucester, Massachusetts, United States – July 20 to 22, 2013). ICCAT, Col. Vol. Sci. Pap., 70(1): 160-189.

Anon. 2015b. Report of the 2015 ICCAT bluefin data preparatory meeting (Madrid, Spain - March 2-6, 2015).

Bonhommeau, S., Farrugio, H., Poisson, F, Fromentin, J-M. 2010. Aerial surveys of bluefin tuna in the western Mediterranean sea: retrospective, prospective, perspective. ICCAT, Col. Vol. Sci. Pap., 65(3): 801-811

Bull, B., Francis, R.I.C.C, Dunn, A., McKensie, A., Gilbert, D.J., Smith, M.H., Bain, R., Fu, D. 2012. CASAL (C++ algorithmic stock assessment library): CASAL user manual. V2.30-2012/03/21. NIWA Technical Report. 135. 280pp.

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., McAllister, M. 2010. Estimating reporting rates using coincidental tag recaptures. Journal of Aquatic Living Resources. 23: 343-352.

Carruthers, T.R., McAllister, M.K. Taylor, N.G. 2011. Spatial surplus production modelling of Atlantic Tunas and Billfish. Ecological Applications. 21: 2734-2755.

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.

Fournier, D.A., Hampton, J., Sibert, J.R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Can. J. Fish. Aqua. Sci. 55: 2105-2116.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A. and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Meth. Soft. 27 (2): 233-249.

Fromentin, J.M. 2009, Lessons from the past: investigating historical data from bluefin tuna fisheries}, Fish and Fisheries, 10, 2, 197--216, Wiley Online Library.

Fromentin, J-M., Lopuszanski, D. 2014. Migration, residency and homing of Bluefin tuna in the western Mediterranean Sea. ICES J. Mar. Sci. 71(3): 510-518.

Fromentin, J-M., Bonhommeau, S., Arrizabalaga, H., Kell, L.T. 2014. The spectre of uncertainty in management of exploited fish stocks: the illustrative case of Atlantic bluefin tuna. Mar. Pol. 47:8-14.

Geromont, H.F., Butterworth, D.S. 2014a. Generic management procedures for data-poor fisheries; forecasting with few data. ICES J. Mar. Sci. doi:10.1093/icesjms/fst232.

Geromont, H.F., Butterworth, D.S. 2014b. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. ICES J. Mar. Sci. doi:10.1093/icesjms/fsu017.

GBYP. 2014. ICCAT Atlantic wide research programme for Bluefin Tuna. Available online at: http://www.iccat.int/GBYP/en/index.htm [accessed October 2014]

Hanke, A.R., Andrushchenko, I., Whelan, C. 2015. Indices of stock status from the Canadian bluefin tuna fishery:1981 to 2013. ICCAT, Col. Vol. Sci. Pap., 71(2): 983-1017.

Hilborn, R. 2003. The state of the art in stock assessment: where we are and where we are going. Scientia Marina 67 (supplement 1): 15-20.

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.Kell, L.T., Kell, A.A. 2011. A comparison of age slicing and statistical age estimation for Mediterranean swordfish (*Xiphias gladius).* ICCAT, Col. Vol. Sci. Pap., 66(4): 1522-1534.

Kell, L.T. T., Fromentin, J.-M., Bonhommeau, S. 2012. An Evaluation of the Implications of Population Structure on the Current Bluefin Tuna Advice Framework. ICCAT, Col. Vol. Sci. Pap., 68(1): 370-378

Kell, L.T., Merino, G., De Bruyn, P., Ortiz de Urbinaz, J, Arrizabalaga., H., Santiago, J., Murua H. 2014. An Example Management Strategy Evaluation of a Harvest Control Rule. ICCAT, Col. Vol. Sci. Pap., 70(5): 2096-2110.

Kell, L.T., Hillary, R., Fromentin, J.-M., Bonhommeau, S. 2015. An example management strategy evaluation of a model free harvest control rule. ICCAT, Col. Vol. Sci. Pap., 71(6): 2790-2797.

Kimoto, A., Takeuchi, Y., Itoh, T. 2015. Updated standardized bluefin CPUE from the Japanese longline fishery in the Atlantic to 2014 fishing year. ICCAT, Col. Vol. Sci. Pap., 71(2): 1059-1097.

Lauretta, M.V., Brown, C.A. 2015. Standardized catch rates of bluefin tuna (*Thunnus thynnus*) from the rod and reel/handline fishery off the northeast United States during 1993-2013. ICCAT, Col. Vol. Sci. Pap., 71(3): 1223-1237.

Leach, A.W., Levontin, P., Holt, J., Kell, L.T., Mumford, J.D. 2014. Identification and prioritization of uncertainties for management of Eastern Atlantic bluefin tuna (Thunnus thynnus). Mar. Pol. 48: 84-92.

Levontin, P., Leach, A.W., Holt, J., Mumford, J.D. Specifying and weighting scenarios for MSE robustness trials. ICCAT, Col. Vol. Sci. Pap., 71(3): 1326-1343.

Martell, S. 2015. The iSCAM project. Available at: https://code.google.com/p/iscam-project/ [accessed 9/9/2015]

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

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

Quinn, T.J., Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Rooker, J. R., Arrizabalaga, H., Fraile, I., Secor, D.H., Dettman, D.L., Abid, N., Addis, P., Deguara, S., Saadet Karakulak, F., Kimoto, A., Sakai, O., Macias, D., Santos, M.N. Crossing the line: migratory and homing behaviors of Atlantic bluefin tuna. Mar. Ecol. Prog. Ser. 504: 265-276.

Santiago, J., Arrizabalaga, H., Ortiz, M., Goñi, N. 2015. Updated standardized bluefin tuna CPUE index of the Bay of Biscay baitboat fishery (1952-2013). SCRS/2014/054.

Taylor, N.G., McAllister, M.K., Lawson, G. L. Carruthers, T.R., Block, B.A. 2011. Atlantic bluefin tuna: a novel multistock spatial model for assessment of population biomass. PLoS ONE. 6(12):1:10.

Thompson, G.G. 1994. Confounding of gear selectivity and the natural mortality rate in cases where the former is a non-monotone function of age. Can. J. Fish. Aqua. Sci. 51:2654-2664.

Walter J. 2015. Update of standardized catch rates of large bluefin tuna (*Thunnus thynnus*) from the U.S. pelagic longline fishery in the Gulf of Mexico 1987-2013 with correction for weak hook effects. ICCAT, Col. Vol. Sci. Pap., 71(3): 1260-1278.

Table 1. The parameters estimated by the model. The example is for a possible bluefin tuna operating model of 8 areas (**Figure 1**), 4 subyears, 5 fleets, 65 years and 25 age classes.

Parameter	Number of parameters	Example	Symbol
Unfished recruitment	n_s	2	R_0
Length a modal selectivity	n_f	5	s_{mode}
Precision of selectivity	n_f	5	s_{prec}
Dome-shape of selectivity	n_f	5	s_{dome}
Recruitment deviations	$(n_{y} + n_a - 1) \cdot n_s$	178	r
Fleet catchability	n_f	5	q
Movement	Up to: $(n_r\text{-}1) \cdot (n_r) \cdot n_m$	224	ψ
Steepness (recruitment compensation)	n_s	2	h
Natural mortality rate modifier	n_s	2	M_{fac}
	Total	428	

Table 2. Summary of the likelihood function for various data.

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
Index of spawning stock biomass (e.g. a larval survey)	year, stock	Log-normal
Length composition	year, subyear, area	Multinomial
PSAT tag (known stock of origin)	stock, year, subyear, area	Multinomial
PSAT tag (unknown stock of origin)	year, subyear, area	Multinomial
Stock of origin	Year, subyear, area	Multinomial

Table 3. Simulation model specification.

Parameter/variable	**Symbol**	**Description**	**Value (range of simulated values)**
Number of stocks	n_s		2
Number of fleets	n_f		2
Number of areas	n_r		4
Number of years	n_y	Historical years of exploitation	40
Number of subyears	n_m		2
Spawning subyear	*ms*		2
Unfished recruitment	*R0*		Stock 1: 225-450 Stock 2: 37 - 75
Natural mortality rate at age	*Minit*	User-specified natural mortality rate at age schedule	Stock 1: 0.49, 0.24, 0.24, 0.24, 0.24, 0.2, 0.175, 0.15, 0.125, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1. Stock 2: 0.1 (all ages)
Natural mortality rate factor	*Mfac*	Multiplier of the Natural mortality rate at age schedule	0.75 – 1.5
Steepness	*h*	Recruitment compensation	0.35 – 0.65
Inter-annual recruitment variation	σ_R	Log-normal standard deviation of recruitment deviations	0.1 - 0.3
Recruitment autocorrelation	Ψ_R		0.5 – 0.9
von Bertalanffy maximum growth rate parameter	κ		Stock 1: 0.087- 0.091 Stock 2: 0.091 - 0.095
Age at maturity	γ	Age when 50% of individuals are mature	Stock 1: 3.5 - 4.5 Stock 2: 8.5 - 9.5
Stock depletion	$SSB_{ny}/SSB0$	Spawning stock biomass relative to unfished	Stock 1: 0.035 - 0.2625 Stock 2: 0.3 - 0.4
Age at 100 selectivity	s_{mode}		5 – 8 (all fleets, all stocks)
Age at 5% selectivity		Ascending limb of sel. curve	2 – 3 (all fleets, all stocks
Sel. of oldest age			Fleet 1: 1 (all stocks) Fleet 2: 0.5 – 1 (all stocks)
Slope in recent exploitation rates			(-5) – 5 % per year (all fleets, all stocks)

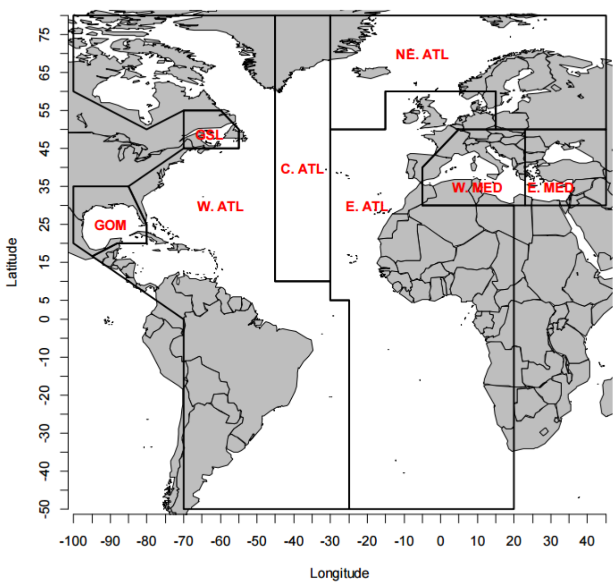

Figure 1. Spatial definitions of the 2015 ICCAT bluefin tuna data preparatory meeting (Anon. 2015).

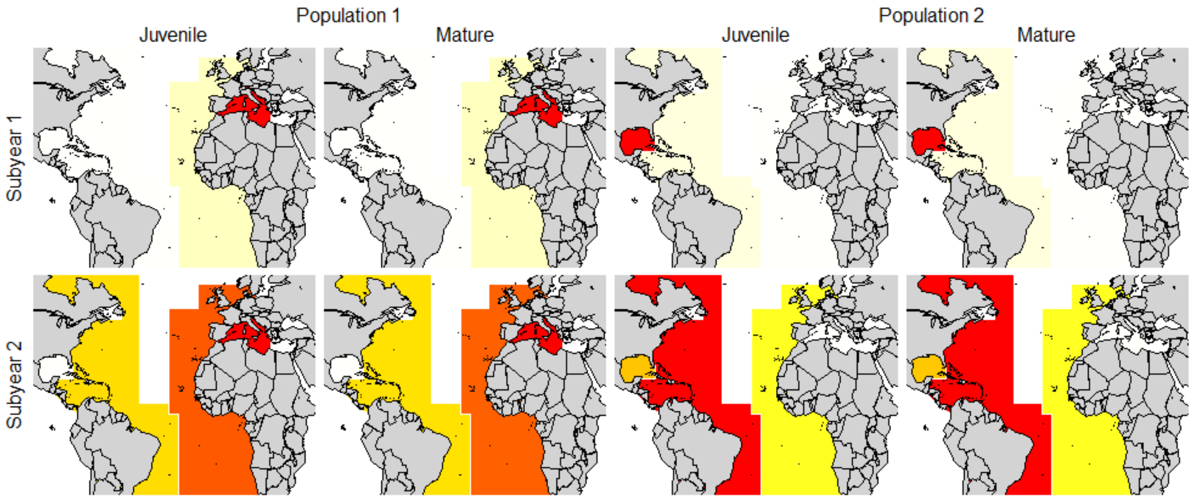

Figure 2. Spatio-temporal distribution of simulated population dynamics. In order to conduct a simplified test of estimation performance, only two subyears and four areas were simulated of the eight identified at the 2015 bluefin data preparatory meeting (ICCAT 2015, Figure 1).

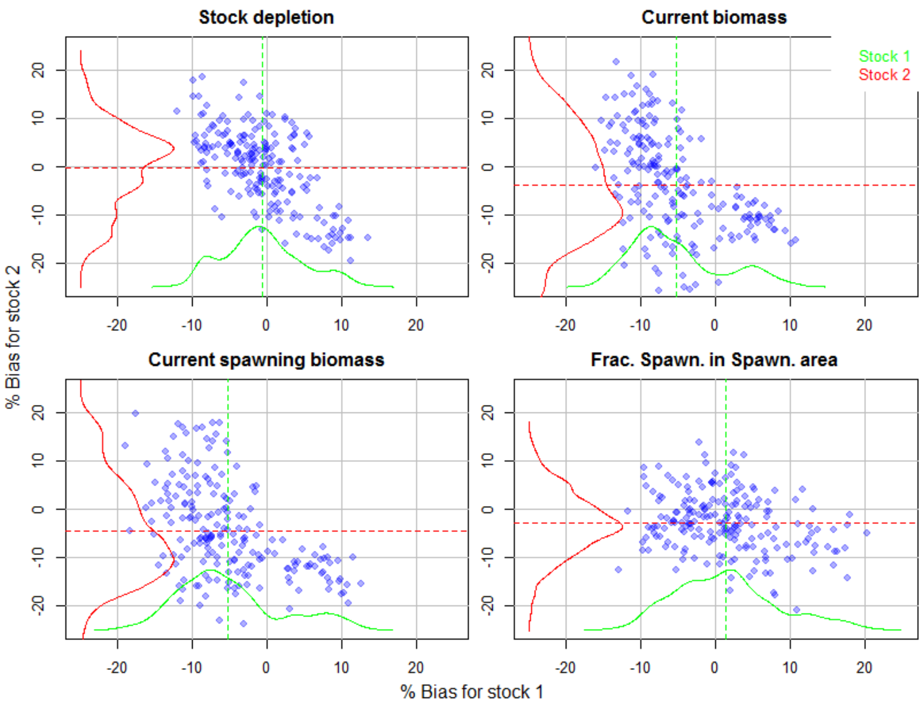

Figure 3. Bias in estimates of simulation model quantities ((estimated value – simulated value)/simulated value). Current refers to the final year of the simulation (ie most recent). Frac. Spawn. in Spawn. area is the fraction of the spawning stock biomass predicted in the spawning area for the final year of the simulation and is included here to examine the ability of the model to estimate spatial distribution reliably. The vertical and horizontal dashed lines represent the mean bias for stocks 1 and 2 respectively.