

A SUMMARY OF DATA TO INFORM OPERATING MODELS IN MANAGEMENT STRATEGY EVALUATION OF ATLANTIC BLUEFIN TUNA

Tom Carruthers¹, Joseph E. Powers², Matthew V. Lauretta³, Antonio Di Natale⁴, Laurence Kell⁵

SUMMARY

We provide a summary of various data that have been collected which may be used to describe alternative hypotheses for fishing and population dynamics of Atlantic bluefin tuna. When formalized in operating models these alternative hypotheses may be used in Management Strategy Evaluation to provide management advice. In this paper we discuss the roles of these data in informing operating models and priorities for obtaining data.

RÉSUMÉ

Nous fournissons un résumé des différentes données qui ont été collectées et qui peuvent être utilisées pour décrire des hypothèses alternatives pour la dynamique de la pêcherie et de la population de thon rouge de l'Atlantique. Une fois formalisées dans des modèles opérationnels, ces hypothèses alternatives peuvent servir dans l'évaluation de la stratégie de gestion pour fournir des avis de gestion. Nous discutons dans ce document des rôles de ces données pour apporter des éléments informatifs aux modèles opérationnels ainsi que des priorités pour l'obtention de données.

RESUMEN

Proporciamos un resumen de diversos datos que han sido recopilados y que podrían utilizarse para describir hipótesis alternativas para la dinámica pesquera y de población del atún rojo. Cuando se formalicen en modelos operativos, estas hipótesis alternativas podrían usarse en evaluaciones de la estrategia de ordenación con el fin de proporcionar asesoramiento en materia de ordenación. En este documento se discuten los papeles de estos datos a la hora de aportar información a los modelos operativos y prioridades para obtener datos.

KEYWORDS

Data collections, age composition, biological sampling, aerial surveys, catch statistics, fishery statistics, fishing effort, size composition, tagging

1 Introduction

A Management Strategy Evaluation (MSE, Butterworth 1999, Cochrane 1998, Punt *et al.* 2014) approach has been proposed for Atlantic bluefin tuna (Anon. 2014) 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). Data for Atlantic bluefin tuna are numerous, vary widely in their information content, may be interpreted in various ways (e.g. exploitation rate, growth,

¹ IOF, 2202 Main Mall, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z4. t.carruthers@fisheries.ubc.ca

² 8465 SW 141 Street, Palmetto Bay, FL 33149, United States. jepowers@lsu.edu

³ 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

⁴ International Commission for the Conservation of Atlantic Tunas, ICCAT, Calle Corazón de María, 8, 28002 Madrid, Spain. antonio.dinatale@iccat.int

⁵ ICCAT. laurie.kell@iccat.int

movement) and often various data types provide contrasting information about fleet and population dynamics. For example, size composition data and catch rate indices may provide different inferences regarding stock depletion. A range of operating models may be developed based on the types of data used to fit the model and the way in which the model interprets these data. A precursor to discussions about how data should be used by operating models is a metadata summary that outlines the types of data that have been collected for Atlantic bluefin tuna. In this paper we summarize these data and describe their potential role in the development of operating models. We also describe an interim solution to operating model development in which a preliminary operating model is established rapidly from data that are freely available, after which the model may be refined as additional data become available.

Since operating models for Atlantic bluefin tuna must be able to represent hypotheses regarding spatio-temporal distribution and stock mixing (Kell *et al.* 2011, Arrizabalaga *et al.* 2014, Fromentin and Lopuszanski 2014) a suitable operating model must include spatial and seasonal structure. The 2015 bluefin tuna preparatory meeting identified eight large marine areas for disaggregation of bluefin tuna data that may be used to model spatial dynamics (**Figure 1**, ICCAT 2015a). Concerns have been expressed about the appropriateness of these spatial definitions for calculating standardized CPUE indices for the Japanese longline fleet however (Kimoto *et al.* 2015). We also identify four subyear temporal definitions that may approximate the spatial distribution of bluefin tuna within model years:

(subyear 1) 1st January – 15th March,
(subyear 2) 16th March – 15th May,
(subyear 3) 16th May – 15th July,
(subyear 4) 16th July – 31st December.

In the remainder of this document we refer to this spatial and temporal structure by ‘area’ and ‘subyear’. A preliminary operating model (Modifiable Multi-stock Model, or ‘M3’) is under development (presented in Carruthers *et al.* (2016, *in press*) which is to be fitted to the various data discussed in this paper. From herein we refer to this as the ‘operating model’.

2 Data and their role in preliminary operating models

A coarse description of Metadata to support bluefin tuna operating model development is presented in **Table 1**.

2.1 CPUE indices

The operating model for Atlantic bluefin tuna is intended to be fitted to standardized catch rate data at the resolution of area and subyear. Providing relative abundance data at this resolution anchors the estimation of movement to only those scenarios that maintain a credible spatial distribution of vulnerable biomass. Models that include only annual indices have no such constraint and can predict movement and spatial distribution from tagging data that are not credible given fishery catch rate data (for example an absence of fish where observed seasonal catch rates are substantial).

It follows that when fitting spatial models, the value of relative abundance indices at the same resolution as estimated movements can be very high. Previous simulation evaluations have shown that reliable estimation of spatial distribution can be obtained from only spatial catch rate data and characterizing this reliably is a primary concern for the estimation of variables of management interest such as stock depletion and MSY reference points (Carruthers *et al.* 2011b). The same spatial catch rate data can also provide additional information about age- or size-specific movement if fleets have varying size selectivities. Since these indices constrain movement estimation to observed spatial distributions, the information content of electronic tagging can be used to explore additional characteristics of movement such as temporal variability or different movement of juvenile and mature fish. Since commercial catch rate data are generally reported at a sufficient temporal and spatial resolution to produce these indices at the scale of area and subyear, a relatively large quantity of information about spatial distribution is provided to the operating model with no additional data collection requirements.

The Japanese longline and US longline are important standardized indices for fitting operating models due to their relatively wide spatial and temporal coverage in mixing areas of the North Atlantic. A proposed US-Canadian combined longline index may provide additional continuity over a wider spatial range that covers the range of the population in the West Atlantic Ocean (Lauretta *et al.* 2015, SCRS-2015-171). In the Mediterranean, Moroccan, Portuguese, Spanish and Italian trap catch rate data may be used to provide indices for eastern and western areas. While these indices are a priority for operating model development, the calculation

and testing of standardizing catch rates can be time intensive. We propose an interim solution in which freely available data, requested and organized by ICCAT, are used in preliminary operating models which are later updated with standardized indices. The most relevant data in this regard are the nominal (non-standardized) Task II catch and effort data, much of which is available at the temporal resolution of day and spatial resolution of 5x5 degree ocean square (ICCAT 2015b) consistent with the defined areas. These data do not include covariates that can be used to account for shifts in fishing practices that are typically included in catch rate standardization. Effects of spatial expansion and contraction may still be accounted for (e.g. Ahrens and Walters 2005; Carruthers *et al.* 2011a).

For a number of reasons, catch rate indices may not be linearly related to true changes in vulnerable biomass and may decline faster (hyperdepletion) or slower (hyperstability) than vulnerable biomass. Scenarios that incorporate hyperstability and hyperdepletion may be important for fitting operating models to historical relative abundance indices and also proposing observation models for MSE.

2.2 Larval indices

Recent papers have suggested that indices derived from larval surveys may be closely correlated with stock assessment predictions of spawning stock biomass (SCRS/2015/035). Larval surveys have been carried out in the Gulf of Mexico (1977-2013) and more recently the western Mediterranean (2001-2005, 2012-2013). It is possible that the operating model could be fitted to these indices under the assumption they are representative of trends in spawning stock biomass in spawning areas. If used in this way, these indices can be expected to strongly influence model predictions of trends in spawning stock biomass (since their interpretation is not contingent on estimates of selectivity and movement). Adding indices of spawning stock biomass may also indirectly inform age-based movement, maturity and selectivity of fishing in spawning areas.

2.3 Catches

The operating model predicts exploitation rate at the scale of subyear and area. It follows that catches are required at this scale and a fleet must be assigned to all catches (this may be an umbrella 'other' fleet of general size selectivity). In order to meet custom subyear definitions above it may be necessary to manually raise Task II catches for subyears and areas to sum to total reported landings.

Catch data are generally the only source of scale in stock assessments: if catches are reported in kilograms the estimates of MSY, catch recommendations and the estimate of current biomass will be a number 1000 times larger than the same catches reported in metric tonnes. It follows that consistent bias in catches across a time series (for example consistent 25% underreporting of catches) has no effect on estimates of stock depletion and fishing mortality rate and therefore is ignorable when specifying operating models (for which scale is not important unless a specific catch recommendation is to be considered). However temporal patterns in bias in catches can be significant (for example significant reductions in illegal, unreported and unregulated fishing of the eastern stock after 2008, ICCAT 2014b) and may provide contrasting view of historical stock trends and exploitation rate (e.g. Carruthers *et al.* 2015a). It follows that scenarios for catch underreporting (and related quantities such as dead discarding) could be used as alternative hypotheses that may be represented by operating models.

In terms of MSE results and selection of management procedures, a potentially more important issue is the level of bias in future observations of catches that is simulated in the management strategy evaluation. Carruthers *et al.* (2015b) found bias in catches to be amongst the most influential of observation processes determining the performance of management procedures. In their study, catch observation bias was often much stronger determinant of MP performance than current stock depletion or historical trend in exploitation rate (estimated by the fitted operating models potentially arising from historical biases in catches). It is therefore important to carefully consider plausible scenarios for biases in reported catches in the future.

2.4 Catch compositions

Currently the operating model is intended to be fitted to length sample data provided to ICCAT (Task II size). The majority of these data are reported at a sufficiently fine spatial and temporal resolution to be aggregated to the subyear and area definitions of the operating model. These data are the primary source of information regarding the size-selectivity of fishing. The representativeness of these data, their spatial-temporal coverage and the contribution to exploitation of the associated fleets is likely to determine how fleets are defined in the operating model.

A fleet in the operating model represents a distinct size selectivity of substantial exploitation rate. The purpose of including fleet heterogeneity in the operating model is to accurately characterize patterns in the exploitation of size classes in order to best estimate quantities of interest such as stock status and fishing mortality rate at maximum sustainable yield. Each fleet in the operating model must have paired observations of catch and an index of fishing mortality rate I (catch divided by a standardized index of abundance) by area and subyear. Each fleet must also have some size composition data. The catches of fleets that do not have both composition data and a standardized index of fishing mortality rate must be aggregated with catches of a fleet of similar size selectivity (that does have both composition data and an index of fishing mortality rate). This relatively strict constraint ensures that the model does not require the estimation of a very large number of fishing mortality rate parameters (e.g. up to 12,480 free F parameters estimated for an eight area model, with four subyears, 65 historical years of exploitation and 6 fleets). By including an index of fishing mortality rate, just one catchability parameter q is estimated per fleet: $F=qI$ (e.g. 6 q parameters as opposed to 12,480 F parameters)

Due to the computational overhead, it is not possible to model the large number of flag and gear combinations recorded by ICCAT which have caught bluefin tuna historically. However it is relatively straightforward to aggregate fleets of similar size selectivity for model fitting and then predict expected individual catch rates *post-hoc*. This allows for parsimonious operating models (avoid including many fleets) while retaining predictive capacity for catch rates of individual fleets which may be a critical component of stakeholder utility.

A persistent issue in the fitting of integrated stock assessment models is the correct weighting of various sources of data (Candy 2008, Francis 2011, Maunder and Punt 2013). This is particularly applicable to size composition data that can often dominate the global objective function due to the number of observations and likelihood functions that are typically assumed for these data (e.g. a multinomial observation model). By fitting to length-composition data we bypass many of the issues associated with deriving age of fish from length data but do not avoid problems associated with overweighting composition data due to incorrectly accounting for non-independence among observations. For example bluefin tuna are observed to school in size mono-specific groups and hence multiple observations of length from a fishing set may in fact represent a single independent length observation. It may be the case that operating model results are sensitive to specification of this ‘effective sample size’. It follows that various scenarios for fleet aggregation and effective sample size may be considered as alternative hypotheses to be represented by operating models in future MSEs.

2.5 Conventional tagging data

Conventional tagging data provide valuable information about the range of movements that are possible for Atlantic bluefin tuna and provide a basis for formulating alternative hypotheses about movement. Conventional tagging data have been used in previous bluefin tuna assessment models to estimate fishing mortality rates (and hence abundance) by quarter and area (MAST, Taylor *et al.* 2011). These approaches have assumed a prior for reporting rate that is constant over time, space and among fleets. There is however evidence that reporting rates vary in time, space and among fleets by factors as great as 500 (Carruthers *et al.* 2011, Carruthers and McAllister 2010, Hillary *et al.* 2008). Similarly large disparities between predicted catches and tag recapture probabilities can be expected if these variable reporting rates are not accounted for.

A fleet specific prior for reporting rate may be prescribed (e.g. Carruthers and McAllister 2010). The principal problem with this approach is that these reporting rate estimates are generally sensitive to alternative assumptions and can be as much as 1/3 or 3 times base-case estimates. While the difference between a reporting rate prior mean of 1/1000 or 3/1000 may not seem large, in relative terms this is highly uncertain and consistent with predictions of catches and abundance that differ by a factor of 3.

It is possible to estimate reporting rates by fleet inside the operating model but since these are confounded with recapture probability these data may only weakly inform abundance. This would not be a serious problem if the computational overhead associated with conventional tagging data was small. However it can be as large as 10 times the total number of other calculations slowing estimation speed by a similar margin (more so if reporting rates are estimated). An operating model of 4 subyears, 65 historical years, 8 areas and 6 fleets requires the calculation of approximately ~14M recapture probabilities in addition to the likelihood calculation for recaptured tags.

Many tags have been recaptured by observer programs for which it may be assumed that reporting rates are 100%. In order to retain information regarding recapture probability (and hence abundance) it is possible to fit to only observer tag recapture data and provide post-hoc estimates of reporting rates (also observed versus predicted recaptures) for the various commercial fleets.

The preliminary version of the operating model is designed to use conventional tagging data to identify the range of possible bluefin tuna movements and also provide information about possible shifts in growth rate of bluefin tuna. The latter is a novel use of conventional tagging data in the context of integrated stock assessment. It has been hypothesized that as exploitation rates increase there should be a shift in the size composition of the modelled population due to the attrition of faster growing individuals (Walters and Martell 2004). Traditional length-based assessments (e.g. MULTIFAN-CL, Fournier *et al.* 1998) have not accounted for this phenomenon and instead, renew the distribution of growth types in each cohort as it ages, irrespective of the level of exploitation. In contrast, a cohort may start with a normal distribution of maximum length among individuals which becomes strongly positively skewed as larger fish are preferentially removed from a cohort (perhaps at high exploitation rates, older fish would belong to the very lower tail of the unfished distribution of maximum length). Simulation studies indicate that estimates of fishing mortality rates from models that do not account for this phenomenon could be strongly positively biased (as much as 3 times that of true fishing mortality rate, C. Walters and A. Hordyk, personal communication).

The operating models under development include a dynamic inverse age-length key that accounts for attrition of faster/larger growth type groups in order to investigate this potentially important source of model misspecification (SCRS/2015/179). Conventional tagging data could provide a valuable empirical validation of shifts in growth rate predicted by the model since each observation of a release and recapture provides an estimate of the growth parameters (e.g. Walters and Essington 2010). Given it may strongly impact estimates of stock status and productivity, growth type attrition may be considered an important alternative hypothesis to be captured in Atlantic bluefin tuna operating models.

2.6 Surgically implanted archival tags

Surgically implanted archival tags (SI tags, e.g. Walli *et al.* 2009, LPRC 2015) provide detailed tag track information and have been included in previous spatial population models to predict both movement and exploitation rates (Taylor *et al.* 2011). While reporting rates of archival tags are generally much higher than conventional tags (due to larger rewards for reporting archival tags), uncertainty over reporting rates and post-release mortality rate still complicates their use in the estimation of exploitation rate. SI tags could be used (1) to estimate exploitation rate and abundance with some assumption about reporting rate, (2) to estimate only recapture probabilities of tags reported by fleets of known reporting rate (e.g. observer fleets), (3) to estimate only movement assuming that releases and recaptures are independent of fishing (similarly to PSAT tags).

2.7 Pop-off Satellite Archival Tags (PSATs)

The primary source of information regarding movements from model areas to model areas among subyears comes from PSAT tags (Block *et al.* 2005, Lutcavage *et al.* 2012, Cermeno *et al.* 2015). The operating model includes these data formatted into separate subyear segments (i.e. has the fields 'year-from', 'area-from', 'area-to', 'subyear-from', 'age-from').

The majority of PSAT tags do not have a definite stock of origin (SOO) (they were not tagged in a spawning area or ocean area specific to a single stock). In such cases it is not clear how to weight these data such that information about movement can be correctly attributed to stock. Determining SOO (or SOO weights) may be undertaken prior to model fitting based on other SOO data. Alternatively this calculation can occur inside the model using the relative likelihood of SOO based on either (1) the model predicted composition of stocks in the areas of a tag track (as in MAST, Taylor *et al.* 2011) or (2) the predicted likelihood of tag track movements given estimated movement parameters of stocks. Both methods that occur inside the model can be expected to be unreliable since they are likely to lead to estimated movements that best divide the tags of unknown SOO (i.e. they can be expected to underestimate stock mixing simply due to the likelihood weighting). This potential source of bias should be simulation tested. The way in which SOO is assigned to PSAT tags may impact estimates of spatial distribution and movement and hence management reference points, and could be considered as alternative scenarios for operating models.

It may be possible to estimate spatial distribution of bluefin tuna in the absence of PSAT data using simplified spatial models that do not attempt to characterize the full range of movement among areas (e.g. the gravity models of Carruthers *et al.* (2011) and MAST Taylor *et al.* (2011)). However PSAT data contain vital information for evaluating alternative movement hypotheses such as age-based movement and temporally variable movement and their collection is a priority for operating model development.

2.8 *Otolith microchemistry*

It is generally not possible to estimate stock size and mixing using a multi-stock model without external data regarding stock of origin: without these data the magnitudes of the various stocks are confounded. Previous multi-stock models such as MAST have used otolith microchemistry data as the primary source of information regarding stock of origin (e.g. Rooker *et al.* 2008, Rooker *et al.* 2014). The challenge is processing microchemistry data at the resolution of subyear and area. While these data have been gathered in all of the areas of the operating model, they have not been processed at this resolution and only exist from 2005 onwards (i.e. Rooker *et al.* 2008)

Given that ICCAT data are already freely available that can be used to characterize trends in abundance, catch compositions and spatial distribution, data that provide information about stock of origin are arguably the most important for fitting a preliminary bluefin tuna operating model (after which other data sources can be refined).

2.9 *Otolith shape analysis*

Otolith shape analysis is an alternative approach for characterizing stock mixing which is both cheaper than micro-chemical analysis and also non-destructive. The method has been compared with traditional micro-chemical analysis and shows promise, achieving similar accuracy (SCRS/P/2015/004). Further detail on the approach is summarized in the report of the 2015 data preparatory meeting (ICCAT 2015a). Similar to other data that provide information regarding stock of origin, these data are particularly valuable in the early stage of operating model development.

2.10 *Single nucleotide polymorphism (SNP)*

SNP data provide another basis for assigning individuals to stock of origin and have been collected for both mixing areas in the Atlantic and spawning areas in the Gulf of Mexico and Mediterranean. The most recent SNP analyses (e.g. SCRS/2015/048) have made use of a larger number of genetic markers and have demonstrated an ability to accurately determine sub-stock structure such as spawning area (e.g. Gulf of Mexico, Balearic Sea, Strait of Sicily and Levantine Sea). It follows that SNP data could provide the necessary empirical basis for formalizing hypotheses that have been proposed for sub stock structure (e.g. Kell *et al.* 2011, Fromentin and Lopuszanski 2014), particularly for the Eastern stock that is thought to have spawning areas in western, central and eastern Mediterranean.

An ongoing challenge for operating model development is generating model datasets that can inform a model with more than one Mediterranean sub-stock. In principle this is possible if SNP data can be used to quantify fractions of individuals in each sub-stock in sub-areas of the Mediterranean and there are sufficient electronic tags (that can be assigned to each Mediterranean sub-stock) to estimate movement. Such data would still have to be provided at the resolution of subyear and area after which they may be too sparse to be considered a reliable characterization of sub-stock structure.

2.11 *Other genetics data*

Coupled with electronic tagging data, older genetic studies based on microsatellites (Carlsson *et al.* 2007) and mitochondrial DNA (Boustany *et al.* 2008) provided early confirmation of the broad east-west stock structure. It is not clear whether these data could be processed at the level of subyear and area in order to inform operating models. However this may be worth pursuing since these data broaden the temporal range of data on SOO.

Another potential for obtaining abundance information is close-kin analysis to estimate the spawning stock abundance (Bravington *et al.* 2013). This provides the potential for genetic mark recapture experiments to estimate absolute abundance, mortality rates or migration, addressing directly some of the key uncertainties in BFT assessments. In particular the close-kin analysis could provide a fishery-independent estimate of spawning stock numbers, particularly for Western Atlantic Bluefin.

2.12 *Fishery independent surveys*

In addition to the fishery-independent larval surveys discussed above, aerial surveys have been suggested as potential source of index information. Aerial surveys (Bonhommeau *et al.* 2010) have been conducted in the Mediterranean since 2010 and may represent the longest running index of spawning biomass for the eastern stock. Similarly to larval indices, the aerial survey may strongly influence operating model estimates of recent trends in spawning biomass. Since the aerial surveys in the Mediterranean have covered areas that encapsulate eastern sub-stock structure, these data may also be used to inform alternative sub-stock structure hypotheses.

Similar aerial surveys have been proposed for the Western stock in addition to a western hydro acoustic survey and a hydro acoustic curtain in the Strait of Gibraltar. While such proposals may not necessarily support operating model development, these data may support management procedures in the future. It follows that it is important to characterize metadata for these types of future data collection programs.

2.13 Growth and aging

The operating model is fitted to catch-at-length data and predicts fishing mortality rate by length class for each fleet. The operating model therefore requires an inverse age-length key (probability of an individual being in a length class given it is of an age class) to convert estimates of fishing mortality rate at length to fishing mortality rate at age. The operating model either (1) applies a static, user-specified inverse age-length key or (2) attempts to generate a dynamic inverse age-length key from model predicted fishing mortality rate and user-specified variability in individual growths. In either case, growth must be characterized in order to develop these keys and also predict the weight of individuals of a given age class. Since the inverse age-length keys are a requirement of the operating model and are not yet available, their derivation is priority for developing a preliminary operating model.

2.14 Maturity

There have been numerous studies focused on bluefin tuna reproductive biology with which to characterize maturity at age and fecundity in general (e.g. Corriero *et al.* 2005, Medina *et al.* 2007, Diaz 2010, Aranda *et al.* 2013, Knapp *et al.* 2014). However participants of the recent data-preparatory meeting (ICCAT 2015) concluded that the body of work on bluefin tuna reproduction should be reconsidered prior to the next assessment. It may be possible to identify a various hypotheses for bluefin tuna maturity and fecundity that are represented by operating models. Participants of the same meeting (ICCAT 2015) also reiterated an ongoing concern that the age-at-maturity estimated from fish on the spawning grounds may not be representative of reproductive contribution at age of the wider population. Again this source of uncertainty may be formalized in operating models.

3 Discussion

3.1 Strategy for operating model development and data priorities

In settings where MSE has been successful it has been an iterative process (Punt *et al.* 2014). It is generally recognised that stakeholders require prior exposure to MSE concepts in order to identify operating models and performance metrics. In recognition of this, it is a priority to develop a working MSE framework based on operating models fitted to data in order to advance the GBYP MSE process. A fitted operating model also allows data providers to see the benefit of their contributions and allows further refinement of the broader MSE framework.

A possible strategy for operating model development is to fit the model to ICCAT data that are freely available and admit more rigorous data to the model as they become available (e.g. standardized indices, PSAT tagging data). Catch data, relative abundance indices and length composition data are currently available as are the parameters to produce preliminary inverse age-length keys and maturity schedules. It follows that the most significant data gaps for developing a preliminary operating model are data regarding stock of origin (e.g. otolith microchemistry, SNP) and movement (e.g. PSAT data). If simplified gravity movement models are assumed tagging data are not necessary to estimate seasonal spatial distribution. As more detailed PSAT data become available alternative movement models can be considered such as age-based movement.

3.2 Possible alternative operating models arising from data assumptions

In this document we discuss various ways in which data can be interpreted by operating models. These may be considered for alternative operating model hypotheses. For example:

- (1) (a) inclusion / (b) exclusion of larval indices
- (2) (a) inclusion / (b) exclusion of PSAT tags of unknown stock of origin (SOO)
- (3) (a) *a priori* assignment of SOO to PSATs / (b) within model assignment of SOO based on stock ratios in tag track areas / (c) within model assignment of SOO based on predicted movements of stocks.
- (4) (a) inclusion of all conventional tags / (b) observer recaptures only / (c) exclusion of conventional tags
- (5) fleet definitions (e.g. all fleets of gear types are aggregated)

- (6) (a) movement by discrete age groups / (b) movement continuously changing with age / (c) age-invariant movement
- (7) weighting of various data regarding stock of origin (Otolith microchemistry, otolith shape, SNP, mtDNA)
- (8) Mediterranean sub-stock structure: (a) aggregated / (b) western-eastern / (c) western-central-eastern
- (9) historical biases in catch reporting
- (10) future biases in catch reporting
- (11) weighting of catch-at-length data (scenarios for effective sample size).
- (12) hyperstability / hyperdepletion in relative abundance indices
- (13) recruitment based on spawning stock biomass (a) in spawning area (b) stock-wide

3.3 Opportunities

The operating model has been subject to simulation testing (SCRS/2015/179) to evaluate estimation performance. These simulations can be broadened to establish which data are most critical in determining model predictions of stock status and productivity. This type of value of information analysis may provide a more rigorous basis for prioritising data gathering to support operating model development.

4 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.

References

- Ahrens, R., Walters, C. 2005. Why are there still large pelagic predators in the oceans? Evidence of severe hyper-depletion in longline catch-per-effort. WCPFC-SC1 ME-WP-3. Available at: https://www.wcpfc.int/system/files/SC1_ME_WP_3.pdf [accessed September 2015].
- Aranda, G., Medina, A., Santos, A., Abascal, F.J., Galaz, T. 2013. Evaluation of Atlantic bluefin tuna reproductive potential from ovarian histology and its use in stock assessments. *Journal of Sea Research*. 76: 154-160.
- Block, B., A., Teo, S.L.H., Walli, A., Boustany, A., Stokesbury, M.J.W., Farwell, C.J., Weng, K.C., Heidi, D., Williams, T.D. 2005. Electronic tagging and population structure of Atlantic bluefin tuna. *Nature*. 434: 1121-1127.
- Bonhommeau, S., Farrugio, H., Poisson, F., Fromentin, J-M. 2010. Aerial surveys of bluefin tuna in the western Mediterranean sea: retrospective, prospective, perspective. *Collect. Vol. Sci. Pap. ICCAT*, 65(3): 801-811 (2010).
- Boustany, A. M., Reeb C. A., Block B. A. 2008. Mitochondrial DNA and electronic tracking reveal population structure of Atlantic bluefin tuna (*Thunnus thynnus*). *Marine Biology* 156:13–24.
- Bravington, M. V., Grewe, P.G., Davies, C.R. 2013. Fishery-independent estimate of spawning biomass of Southern Bluefin Tuna through identification of close-kin using genetic markers. FRDC Report 2007/034. CSIRO, Australia.
- Butterworth, D.S., Punt, A.E., 1999. Experiences in the evaluation and implementation of management procedures. *ICES J. Mar. Sci.* 56, 985-998.
- Candy, S.G. 2008. Estimation of effective sample size for catch-at-age and catch-at-length data using simulated data from the Dirichlet-multinomial distribution. *CCAMLR Science*. 15: 115-138.
- Carruthers, T.R., McAllister, M. 2010. Estimating reporting rates using coincidental tag recaptures. *Journal of Aquatic Living Resources*. 23: 343-352.
- Carruthers, T.R., Ahrens, R.N.M., McAllister, M.K., Walters, C. 2011. Integrating imputation and standardization of catch rate data in the calculation of relative abundance indices. *Fish. Res.* 109(1): 157-167.
- Carruthers, T.R., Fonteneau, A., Hallier, J-P. 2014. Estimating tag reporting rates for tropical tuna fleets of the Indian Ocean. 155: 20-32.

- Carruthers, T.R., Kell, L.T., Butterworth, D.S., Maunder, M.N., Geromont, H.F., Walters, C., McAllister, M.K., Hillary, R., Levontin, P., Kitakado, T., Davies, C.R. 2015a. Performance review of simple management procedures. ICES J. Mar. Sci. In press.
- 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. 2015b. 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
- Cermeno, P., Quilez-Badia, G., Ospina-Alvarez, A., Sainz-Trapaga, S., Boustany, A., Seitz, A.C., Tudela, S., Block, B.A. 2015. Electronic tagging of Atlantic bluefin tuna (*Thunnus thynnus*, L.) reveals habitat use and behaviors in the Mediterranean Sea. *PLoS ONE* 10 (2): e0116638. doi: 0.1371/journal.pone.0116638
- 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.
- Corriero, A., Karakulak, S., Santamaria, N., Deflorio, M., Spedicato, D., Addis, P., Desantis, S., Cirillo, F., FenechFarrugia, A., Vassallo-Agius, R., de la Serna, J. M., Oray, Y., Cau, A., Megalofonou, P., Metrio, G., 2005. Size and age at sexual maturity of female bluefin tuna (*Thunnus thynnus* L. 1758) from the Mediterranean Sea. *Journal of Applied Ichthyology* 21, 483–486.
- Diaz, G. 2010. A revision of western Atlantic bluefin tuna age of maturity derived from size samples collected by the Japanese longline fleet in the Gulf of Mexico (1975-1980). *ICCAT, Collect. Vol. Sci. Pap.* 66 (3): 1216-1226.
- 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.
- Francis, R,I,C,C. 2011. Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aqua. Sci.* 68(6): 1124-1138.
- 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., 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.
- 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.
- GBYP. 2014. ICCAT Atlantic wide research programme for Bluefin Tuna. Available online at: <http://www.iccat.int/GBYP/en/index.htm> [accessed October 2014]
- 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.
- Hillary, R.M., Million, J., Anganuzzi, A., Areso, J.J. 2008. Tag shedding and reporting rate estimates for Indian Ocean tuna using double-tagging and tag-seeding experiments. Working paper of the Indian Ocean Tuna Commission. IOTC-2008-WPTDA-04. Available at: <http://www.iotc.org/files/proceedings/2008/wptda/IOTC-2008-WPTDA-04.pdf> [accessed September 2015]
- ICCAT 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]
- ICCAT 2014b. Executive summary 2014, status of Eastern stock of Atlantic bluefin tuna. Available at: https://www.iccat.int/Documents/SCRS/ExecSum/BFT_EN.pdf [accessed September 2015]

- ICCAT 2015a. Report of the 2015 ICCAT bluefin data preparatory meeting. Available at: https://www.iccat.int/Documents/Meetings/Docs/BFT_DATA_PREP_2015_eng.pdf [accessed September 2015].
- ICCAT 2015b. Information published on TaskII – Catch and Effort (T2CE). Available at: <https://www.iccat.int/Data/t2ce-ENG.pdf> [accessed September 2015]
- ICCAT 2015c. CATDIS readme file. Available at: http://www.iccat.int/Data/Catdis/Readme_catdis.docx [accessed September 2015]
- Ingram, G.W., Jr., D. Alvarez-Berastegui, P. Reglero, R. Balbín, A. García5, and F. Alemany. 2016. In press. SCRS/2015/035. Indices of larval bluefin tuna (*Thunnus thynnus*) in the Western Mediterranean Sea (2001-2013).
- 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.
- Knapp, J.M, Aranda, G., Medina, A., Lutcavage, M. 2014. Comparative assessment of the reproductive status of female Atlantic bluefin tuna from the Gulf of Mexico and the Mediterranean Sea. PLOS One. doi: 10.1371/journal.pone.0098233.
- 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*). Marine Policy. 48: 84-92.
- Levontin, P., Leach, A.W., Holt, J., Mumford, J.D., and Kell.L. 2014. Specifying and weighting scenarios for MSE robustness trials. Collect. Vol. Sci. Pap. ICCAT, 71(3): 1326-1343 (2015).
- LPRC. 2015. Large pelagics research centre. Available at: <http://www.tunalab.org/> [accessed September 2015]
- Lutcavage, M.E., Galuardi, B., Lam, T.C.H. 2012. Predicting potential Atlantic spawning grounds of western Atlantic bluefin tuna based on electronic tagging results, 2002-2011. Col. Vol. Sci. Pap. ICCAT. 69(2): 955-961.
- Maunder M.N., Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research. 142: 61-74.
- Medina, A., Abascal, F.J., Aragon, L, Mourente G, Aranda, G., *et al.* 2007. Influence of sampling gear in assessment of reproductive parameters for bluefin tuna in the western Mediterranean. Mar. Ecol. Prog. Ser. 337: 221-230.
- 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
- 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.
- Rooker JR, Secor DH, De Metrio G, Schloesser R, Block BA, *et al.* 2008. Natal homing and connectivity in Atlantic bluefin tuna populations. Science 322: 742–744.
- Anon. 2014. 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.
- 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.
- Walli A, Teo SLH, Boustany A, Farwell CJ, Williams T, Dewar, H., Prince, E., Block, B.A. 2009. Seasonal movements, aggregations and diving behavior of Atlantic bluefin tuna (*Thunnus thynnus*) revealed with archival tags. PLoS ONE 4: e6151. doi: 10.1371/journal.pone.0006151
- Walters, C., Essington, T. 2010. Recovery of Bioenergetics Parameters from Information on Growth: Overview of an Approach Based on Statistical Analysis of Tagging and Size-At-Age Data. The Open Fish Science Journal, 2010, 3, 52-68

Table 1. Simplified summary of data to inform operating models for Atlantic bluefin tuna. Spatial range either refers to the areas of **Figure 1** or stock (Es = eastern stock, Ws = western stock).

Type of data (<i>Informs</i>)	Year range	Til	Spatial range	Can be by season?	Contact	Collab	Available to:				Used in OM?
							TC	CMG	ICCAT	ALL	
1. CPUE indices (relative abundance, movement, performance at stakeholder level)											
1.1. ICCAT task II CPUE	1950-2014	∞	All	Y	Carlos Palma (ICCAT)	Y	Y	Y	Y	Y	Yes
1.2. Japanese LL standardized spatial	1976-2013	∞	E, NE, W, C	Y	Ai Kimoto	Y	N	N	N	N	Not yet
	1990-2013	∞		Y		Y	N	N	N	N	Not yet
	1992-2014	∞	W	Y		Y	N	N	N	N	Not yet
1.3. USA LL standardized spatial	1992-2004	∞	GOM	Y		Y	N	N	N	N	Not yet
	2005-2014	∞	GOM	Y	Matt Lauretta (NOAA)	Y	N	N	N	N	Not yet
1.4. USA HL standardized spatial	1980-2014	∞	W	Y		Y	N	N	N	N	Not yet
1.5. USA RR standardized spatial	1992-2014	∞	W	Y							
1.6. USA-CAN LL standardized spatial	1992-2014	∞	W, C	Y	M. Lauretta (NOAA) /	Y	N	N	N	N	Not yet
1.7. USA-CAN HL standardized spatial	1993-2014	∞	W, C	Y	A. Hanke (DFO)	Y	N	N	N	N	Not yet
1.8. CAN LL standardized		∞	W, GSL	Y		Y	N	N	N	N	Not yet
	1981-2014	∞	GSL	Y	Alex Hanke (DFO)	Y	N	N	N	N	Not yet
1.9. CAN HL standardized	1988-2014	∞	W	Y		Y	N	N	N	N	Not yet
1.10. TWN LL standardized	1960-2004	2004	W, NE, E	Y	Julia Huang (NTOU)		N	N	N	N	Not yet
1.11. MOR TRAP standardized	1981-2014	∞	WM	Y	N. Abid		N	N	N	N	Not yet
1.12. POR TRAP standardized			W, WM	Y			N	N	N	N	Not yet
1.13. ESP TRAP standardized			W, WM	Y	Jose Miguel de la Serna		N	N	N	N	Not yet
1.14. ITA TRAP standardised			CM	Y	Pierantonio Addis	Y		N	N	N	Not yet
2. Larval indices (SSB, movement)											
2.1. USA	1977-2013	∞	GOM	Y	Walter Ingram (NOAA)	Y	N	N	N	N	Not yet
2.2. ESP	01-'05 '12-'13	2018	W Med	Y	Franciso Alemany (IEO)	Y	N	N	N	N	Not yet
3. Catches (stock size, harvest rate)											
3.1. ICCAT task I	1950-2015	∞	non-spatial	N	Carlos Palma (ICCAT)	Y	Y	Y	Y	Y	Yes
3.2. ICCAT task II			All	Y		Y	Y	Y	Y	Y	Yes
3.3. GBYP	1512-1950		E, M	Y	Carlos Palma (ICCAT)	Y	Y	Y	Y	Y	Not yet
4. Catch composition (selectivity, depletion)											
4.1. ICCAT catch-at-size	1950-2015	∞	All	Y	Carlos Palma (ICCAT)	Y	Y	Y	Y	Y	Yes
4.2. Stereo video caging	2014	ended	WM, EM	Y	Mauricio Ortiz (ICCAT)		N	N	N	N	Not yet
4.3. Canadian fisheries				N	Alex Hanke (DFO)						
4.4. GBYP Historical catches	1910-1950	=	E, M	Y	Carlos Palma (ICCAT)	Y	N	Y	Y	Y	Not yet
5. Conventional tags (feasible movement, growth, GTG heterogeneity)											
5.1. ICCAT	1954-2014	2015	All	Y	Carlos Palma (ICCAT)	Y	Y	Y	Y	Y	Yes
6. SI archival tags (feasible movement)											
6.1. LPRC (n=4000)	2011-2015		W	Y	Molly Lutcavage		N	N	N	N	Not yet
7. PSAT tags (movement)											
7.1. LPRC (n=423)	2005-2009	ended	W	Y	Molly Lutcavage		N	N	N	N	Not yet
7.2. DFO (n=135)	2013-2015	∞	GSL,W,GOM	Y	Alex Hanke (DFO)	Y	N	N	N	N	Not yet
7.3. Stanford (n=1783)	1996-2010	∞	W	Y	Barbara Block		N	N	N	N	Not yet
7.4. GBYP (n = 103)	2012-2014	2015	E,MED	Y	Antonio Di Natale	Y	Y	N	N	N	Not yet
7.5. WWF (n = 100)				Y	Pablo Cermenon		N	N	N	N	Not yet
7.6. SEFSC (NOAA)	2011-2013		GOM,W,GSL	Y	Craig Brown		N	N	N	N	Not yet
7.7. Acadia (NS)			GSL	Y	Mike Stokesbury		N	N	N	N	Not yet
7.8. UCA	2011	ended	W, C, WM	Y	Antonio Medina	Y	Y	N	N	N	Not yet

Table 1 continued.

8. Otolith microchemistry (stock of origin)											
8.1. UMCES, TAMU	2012-2013			Y	David Secor	Y	N	N	N	N	Not yet
8.2. NOAA							N	N	N	N	Not yet
8.3. EU (AZTI)	2009-2011	ended	E	Y	Igaratza Fraile	Y	N	N	N	N	Not yet
8.4. DFO / UMCES	2011-2013	∞	W, GSL	Y	Alex Hanke (DFO)		N	N	N	N	Not yet
8.5 GBYP	2011-2015		All	Y	GBYP	Y		Y	Y	Y	Not yet
9. Otolith shape analysis (stock of origin)											
9.1. GBYP GMIT (n=718)	2013	2015	E, W, C, WM	Y	Deirdre Brophy		N	N	N	N	Not yet
10. SNP (population structure, genetic structure)											
10.1. Med HCMR					Gianpaolo Zmpicinini		N	N	N	N	Not yet
10.2. GBYP UB	2011-2015		All		Gregory Puncher	Y	N	N	N	N	Not yet
					Naiara Rodriguez						
10.3. AZTI (n=130)					Espeleta	Y	N	N	N	N	Not yet
10.4 NOAA/VIMS/CSIRO	2015		GOM/MED	N	John Walter		N	N	N	N	Not yet
10.5 GBYP Historical UB	200 BC - 1927		E, M	Y	Alessia Cariani	Y	N	N	N	N	Not yet
11. Other genetics on population structure (population structure, genetic structure)											
11.1. mtDNA					Barbara Block		N	N	N	N	Not yet
11.2. Micro Sat/ mtDNA (n=320 / 147)	2003	ended	GOM, WM	Y	Carlsson		N	N	N	N	Not yet
12. Fish. Ind. surveys (relative abundance, movement)											
12.1. ICCAT Aerial	2010-2015		M	Y	Antonio Di Natale	Y	N	N	N	N	Not yet
12.2. USA Aerial	2015-		W	Y	Molly Lutcavage		N	N	N	N	Not yet
12.3. USA Acoustic	2015-		W	Y	Molly Lutcavage		N	N	N	N	Not yet
12.4. SOG Hydro acoustic curtain (OTN)	proposed		W, WM	Y	Mike Stokesbury		N	N	N	N	Not yet
13. Growth, aging (age-length keys, length-age keys)											
13.1. Age-length keys (NOAA)				Y	John Walter		N	N	N	N	Not yet
					Enrique Rodriguez-						
13.2. Age-length keys (IEO)	2010-2012	ended	E, WM	Y	Marin		N	N	N	N	Not yet
13.3. Age-length keys (DFO)	2010-2013	ended	GSL, W	Y	Alex Hanke (DFO)		N	N	N	N	Not yet
13.4. Derived from tagging	1963-2012	ended	Es, W s	Y	Lisa Allouid		N	N	N	N	Not yet
13.5 Age-length keys (GBYP)	2011-2015		E, M	Y	Antonio Di Natale	Y	N	Y	Y	Y	Not yet
13.5 Ageing calibration (GBYP)	2014		E, M	Y	Antonio Di Natale	Y	N	Y	Y	Y	Not yet
14. Maturity (Spawning biomass)											
14.1. Western (NOAA)	1975-1981	ended	GOM	Y	Guillermo Diaz (NOAA)		N	N	N	N	Not yet
14.2 Mediterranean		rew	M	Y	GBYP	Y	N	Y	Y	Y	Not yet
15. Other ecological data (spatial distribution, covariates for CPUE standardization, steepness, natural mortality rate, spawning locations etc.)											
15.1. Larval ecology (IEO)				Y	Diego Alvarez		N	N	N	N	Not yet
15.2. Habitat model				Y	Berastegui		N	N	N	N	Not yet
					Jean-Noel Druon		N	N	N	N	Not yet

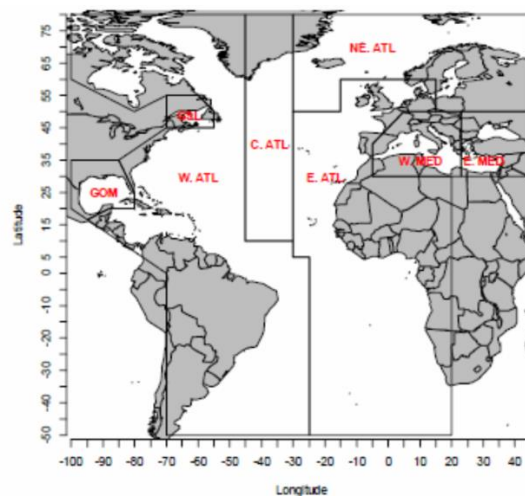


Figure 1. Spatial definitions of the 2015 ICCAT bluefin tuna data preparatory meeting (ICCAT 2015).