Received: 27 November 2014
IOTC-2014-WPM05-07

Working/Info Paper for the IOTC Working Party on Methods Dec2014

Computational Framework to Support Indian Ocean Bigeye and Yellowfin Management Strategy Evaluation: A review of software requirements and options

Dale Kolody, Rich Hillary, Anne Preece

E: dale.kolody@csiro.au

CSIRO Ocean & Atmosphere Flagship. GPO Box 1538, Hoabart, TAS, 7001, Australia

Abstract

This paper describes a number of considerations for the computational framework to support Management Strategy Evaluation (MSE) for Indian Ocean yellowfin (YFT) and bigeye (BET) tunas, including choice of programming languages, operating model structure and conditioning options.

Programming languages are evaluated with respect to:

- Price and familiarity within the fisheries science community

Operating Model structural options include:

- Single or multiple species evaluated simultaneously
- Age-structure, multiple fleets with independent selectivity
- Spatial structure, including:
    - Stock structure (independent spawning populations for an individual species)
    - Seasonal and interannual variability in movement/recruitment distribution
- Non-stationary production dynamics (recruitment, M and/or growth)
- Data available for Harvest Control Rules (type and error structures):
    - Catch by fleet
    - CPUE-based abundance indices
    - Catch-at-Length
    - Tag dynamics: conventional and/or genetic options (potentially including close kin)
- Management decision options:
    - Catch controls
    - Effort Controls
    - Time/Area Closures

Conditioning Options include:

- Build likelihood-based estimators into the operating model to ensure structural consistency
- Use the assessment to parameterize the operating model with ad hoc adaptations for structural incompatibilities

Comparisons with some existing MSE software is made in the interest of code re-usability and the potential for shared development. These items were initially discussed at the informal Working Party on Methods (Ispra Mar2014) and have been updated to reflect those discussions. We hope for additional feedback prior to full commencement of the IOTC BET/YFT MSE technical support project in 2015.

Introduction

In 2011, the Indian Ocean Tuna Commission (IOTC) endorsed a plan to pursue Management Strategy Evaluation (MSE, largely pioneered in a fisheries context by the International Whaling Commission, e.g. de la Mare 1986) as an approach for achieving responsible fishery management (IOTC 2011) and it has been recognized that MSE is the best tool for addressing IOTC Resolution 12/01 (IOTC 2012):

[The Commission] AGREES, in accordance with paragraph 1 of Article IX of the IOTC Agreement, to the following:

- *To apply the precautionary approach, in accordance with relevant internationally agreed standards, in particular with the guidelines set forth in the UNFSA, and to ensure the sustainable utilization of fisheries resources as set forth in Article V of the IOTC Agreement.*
- *In applying the precautionary approach, the Commission shall adopt, after due consideration of the advice supplied by the Scientific Committee, stock-specific reference points (including, but not necessarily limited to, target and limit reference points), relative to fishing mortality and biomass, and associated harvest control rules, that is, management actions to be taken as the reference points for stock status are approached or if they are breached*

The IOTC MSE process has begun for Albacore tuna under the auspices of the IOTC WPM, adopted as a priority because of concerns about recent stock status. A parallel process has begun for skipjack (SKJ) tuna, motivated by the conditions imposed by the Marine Stewardship Certification (MSC) for the Maldives pole and line fishery (Adam et al. 2013). To date, there has been limited progress on similar initiatives for the other major target species of the IOTC, with yellowfin and bigeye tuna being the prime candidates. Though we note that some work exploring MSE for BET was pursued independently of the IOTC process (e.g. Tong et al, 2011).

The implementation of an MSE approach typically involves the following components:

1) Specification of Management Objectives – what fisheries management hopes to achieve, and the quantitative criteria against which harvest strategies can be evaluated. Interim reference points are defined for IOTC stocks in resolution 13/10 (IOTC 2013: target RPs of B_{MSY}, F_{MSY} and limit RPs of 0.4-0.5 B_{MSY} and 1.3 – 1.5 F_{MSY} depending on the species).

2) Candidate Harvest Strategies – decision rules that specify (in advance) how data will be collected, interpreted and translated into a management action (the Harvest Control Rule – HCR is the decision rule that converts data into the management action).

3) Operating Models – simulation models that represent the key features and uncertainties of the fish population, fishing fleets, data collection and management systems in forward projections.

4) Simulation Testing – the process of using the operating models to evaluate the candidate harvest strategies in stochastic dynamic projections

5) Harvest Strategy Selection – choosing the Harvest Strategy that has the most desirable performance trade-offs with respect to the management objectives

6) HS Implementation – adoption of the selected harvest stratey to provide the management advice for the fishery.

This list does not necessarily describe a sequential process, as some elements can be pursued in parallel and several components are typically revisited as understanding of the system and management priorities tend to iteratively improve. The overall process requires the engagement of the IOTC member nations and stakeholders to participate in informed decision making at several points. The IOTC agreed to initiate such a consultative process (IOTC 2012), and relevant meetings were begun in earnest in 2014 (e.g. WWF/GEF – Areas Beyond National Jurisdiction workshops in Sri Lanka in Apr 2014, MSE consultation day associated with the 2014 Commission Meeting, and Introductory Stock Assessment and Science/Management Workshop South Africa Sep 2014).

This document pertains to components 2-4 above, outlining some of the key technical decisions and developments that need to be made to move the MSE agenda forward for Indian Ocean yellowfin (YFT) and bigeye (BET) tunas. Computer software and analytical requirements include:

i) The operating model (OM) – the simulator for the fish population, fishery and data collection, which attempts to provide a realistic description of the main uncertainties in the system.
ii) Conditioning – The process used to specify the parameters (and uncertainty) for the operating model. This software could be distinct from the simulator, or could use the same structure as a statistical estimator. Conditioning should be reasonably consistent with and informed by the stock assessment process.
iii) Harvest Control Rule – the algorithm that uses the data to prescribe the management action within the simulator.
iv) The scenario controller. Running an MSE typically requires a large number of complicated scenario specifications that are best handled with automated file processing to the extent possible.
v) Results summarization. A high level language is desirable for generating automated summary graphics and tables of the simulator output files.

In this paper we focus on key discussion points to get the process started, particularly the operating model and conditioning. Harvest Control Rule development is only discussed superficially at this time. This document is presented as a strawman for discussion and feedback from the IOTC Working Party on Methods and other interested parties.

Software options

Programming Languages

MSE in an international RFMO context needs to be an open and transparent process, and toward this end, efforts should be made to ensure that software and computing platforms do not represent unnecessary hurdles for CPC scientists (e.g. through the use of esoteric or expensive proprietary software). We propose that the high level statistical and graphical R software should be used to control the MSE evaluation process (i.e. including automated batch handling of a large number of combinations of operating models and harvest control rules, and summarizing simulation results). If appropriate, this may involve the R package FLR - Fisheries Library in R (Kell et al. 2007), which has many useful features designed for MSE. The computationally demanding work should be executed using an efficient language that can be compiled into machine code with C++ (GNU g++ version) suggested as a likely candidate. Open source stock assessment packages (or purpose-built C++ models, using ADMB – AD Model Builder, Fournier et al. 2012) represent a good option for operating model conditioning (options discussed below). R, g++, and ADMB are all freely available and supported in several computing environments.

Structural needs for YFT and BET Operating models

There are a number of trade-offs in developing a fishery simulator. One can always argue that more detail is required to increase realism, but there are usually shortcuts that capture the essence of the real-world feature with a minimum of computational overhead. In considering the workplan over the next 2 years, we recognize that some IOTC working party participants might desire additional features that cannot be delivered in the initial timeframe. We consider the highest priority output to be an MSE framework that may be simplistic in some respects, but which can be used to simulate fishery dynamics and management at a sufficient level to begin evaluating and comparing alternative HCRs. Foresight for future expansion is appropriate, but innovative complexity should initially be frugal and added incrementally after the framework is functioning, in relation to the recommendations of the IOTC Working Parties on Methods and Tropical Tunas, Scientific Committee, and relevant cross tuna-RFMO initiatives.

The majority of fisheries MSE applications to date have involved relatively simple structural assumptions that are largely consistent with the stock assessment modelling tradition that has prevailed over the past decades:

- Single species
- Single stock
- Spatially-aggregated
- Age-structured
- Sex-aggregated
- Stationary growth, mortality and stock recruitment relationship
- Multiple fleets with stationary selectivity
- Data for conditioning includes:
    - Total catches
    - Relative abundance indices (survey or standardised CPUE)

  - Size (or age) composition
  - Possibly tag releases and recoveries
- Data for HCR includes:
  - Total catches
  - Relative abundance indices (survey or standardised CPUE)
  - Size (or age) composition
- Management through total catch controls

We highlight that the CCSBT operating model is a good example with most of these features, and represents the first simulation-tested and adopted international tuna harvest strategy. CCSBT is also a much simpler system than the IO tropical tunas, because i) it is largely a single species fishery with minimal targeting confounding in the core of the fishery, ii) there is a simple single stock population structure and (seemingly) well understood movement patterns that justify a spatially-aggregated assessment, iii) it is relatively data rich (including fisheries data plus a juvenile aerial survey, conventional tagging studies and a close-kin genetic mark recapture estimator for spawning biomass), and iv) there are a small number of fleets. While it is sobering to note that it took ~10 years of iterative MSE development to achieve adoption of a management procedure, the technical requirements were not the limiting factor. Below, we briefly discuss each of the bulleted points in relation to Indian Ocean YFT and BET, identifying a range of options from simple to complex, and propose a set of preliminary operating model features in Table 1.

Single or Multiple Species

Multi-species models are typically categorized in terms of trophic interactions or technical interactions. Despite the growing trend in ecosystem-based fishery management, for BET and YFT, the need to represent trophic interactions is probably a low priority because observations and understanding of prey and predator interactions is very limited. Given the speculative way in which one would probably have to specify the dynamics of lower (or higher) trophic levels, it seems more productive to speculate about how the trophic interactions might have a net effect on BET/YFT variability in growth, mortality, recruitment and/or movement.

Technical interactions are potentially more important for BET and YFT because they are typically caught together in mixed species fisheries, though many fleets have considerable capacity for selective targeting. If the main species are to be managed by simultaneous catch limits on the individual species, there may be a need to quantify the degree to which limits on one species will affect catches of the other. If the IOTC intends to manage on the basis of overall effort, or time-area closures, it may be appropriate to simulate the effect of different management decisions on multiple species simultaneously. Inescapably, skipjack (SKJ) fisheries will be an important part of multispecies considerations, but SKJ is outside the scope of the current project.

Suggested priorities:

1) Single species
2) Multi-species YFT, BET (and SKJ?) will only be simulated simultaneously if management options are pursued which do not aim for in-season monitoring and control of individual species catches.

Stock structure and spatial connectivity

For the purposes of this document, we define stock structure as the spatial (and/or temporal) distribution of individuals derived from different reproductively isolated spawning populations. Stocks may overlap spatially (and temporally) to a greater or lesser extent. Mixed stock fisheries can represent a serious management problem (e.g. the IWC and Pacific salmon fisheries spend a lot of effort on this issue), largely due to questions of "ownership", and different stock-specific management objectives (e.g. if two stocks are harvested together but have different productivity, there will be a trade-off between over-harvesting the less productive stock and foregoing lost economic opportunity on the more productive stock). We note that some authors describe spatially-disaggregated population models as including stock structure despite aggregating spawning biomass calculations across regions, as long as the different regions have individual recruitment estimates (even though these estimates are not independent). In this document we use the term stock structure to refer to somewhat reproductively isolated populations, which potentially have independent stock recruitment relationships and other biological characteristics (e.g. growth rates, migration patterns, etc.)

Most of the recent tropical tuna assessments in the Indian Ocean (and elsewhere) have assumed a single spawning stock at the scale of the RFMO management unit (e.g. Figure 1, Figure 2 left panels). In some cases, only the western half of the Indian Ocean has been assessed, under the speculative assumption that it may represent a single discrete population reasonably separated from the eastern IO. As discussed in Kolody et al. (2013), stock structure speculation in the Indian Ocean has been motivated by i) observations of somewhat different data characteristics on the two sides (size composition and CPUE), ii) observation of relatively few tags released on one side of the ocean recovered on the other (though this inference is complicated by poor reporting rates in most fleets outside of the western area), and iii) there have now been at least 5 studies suggesting that there may be genetic population structure within the tropical tuna species of the Indian Ocean (including YFT, BET and SKJ), identified within the study areas roughly shown in Figure 3. Furthermore, high resolution analyses of tagging studies (e.g. in the Pacific Ocean) have suggested that average movement rates are much more restricted than has historically been assumed on the basis of relatively few rapid long-range tag displacements (e.g. Sibert and Hampton 2003, Kolody and Hoyle, in press). While there is currently no consensus on Indian Ocean stock structure, it will hopefully be examined with a concerted effort over the next few years, with EU funds identified for the task.

The spatially-disaggregated assessments that have been applied to the IO tropical tuna stocks (MULTIFAN-CL and Stock Synthesis) have all assumed a single stock. The models do estimate recruitment by sub-region (i.e. recruitment deviates subject to constraints that render them somewhat independent of one another spatially), and many of the assessments have suggested that some areas have very limited mixing with other areas (e.g. Figure 1, right panel), but this may be largely an artefact (e.g. due to inaccurate and conflicting CPUE trends by region and poor tag recovery reporting rates outside of the western purse seine / pole and line fishing areas).

One can easily contrive situations in which the failure to resolve stock structure would be critical, but it is not clear that those situation are realistic at this time. To fully describe multiple stocks (e.g. Figure 2, right panel), would require multiple spawning populations, migration characteristics that link most fish to their natal regions, and understanding of the degree of overlap in the fisheries (and possibly meta-population structure if there is some mixing among spawning populations). Reference points should be calculated independently for different stocks. However, its worth emphasizing that it may be possible to devise HCRs which are robust to the stock structure complexity, and which would not require extensive and continuous stock structure monitoring.

Indian Ocean YFT, BET and SKJ have all been assessed with spatially-structured and spatially-aggregated models. Spatial structure is intended to partition the population into more homogeneous units that may exhibit some degree of independence. However, for BET (and SKJ), the decision was made to place greater emphasis on the spatially–aggregated assessments. Reasons for this include (*inter alia*): i) the unbalanced tag release design and poor reporting rates outside of the core purse seine fishery limit the ability to estimate movement, ii) data in many regions is of poor or unknown quality, and/or iii) the overall stock status estimates from aggregated and disaggregated models did not necessarily indicate the need for different management advice.

The YFT assessment spatial structure has remained relatively stable in recent years (Figure 1). The structure was derived from a (largely qualitative) attempt to identify reasonably homogenous units, including consideration of i) distinct fisheries (in terms of gear type, flag and/or seasonality), ii) Longhurst bio-geographical provinces, iii) the need to keep the estimators numerically tractable.

For BET and YFT operating models, it would be difficult to argue that spatial structure is completely irrelevant (e.g. harvesting off the east coast of Africa probably has a negligible effect on the Indonesian fishery). The stock structure is uncertain, but plausibly exists at a finer scale than has been generally appreciated. There is a further interaction with the choice of management measures. If management options are to be implemented via a range of high resolution space-time closures, one needs to understand how those measures will affect the population at the scale of interest.

Option priorities:

1) 1-2 stocks, 2 regions (west and east). Given the limited data available with which to reliably estimate stock structure and movement, conditioning would probably require forced input of a range of different migration assumptions, potentially including:
    - Stationary, region-specific, seasonal, age-dependent proportional migration (i.e. standard bulk transfer co-efficients)
    - Stock-specific migration with site fidelity to natal spawning regions
    - Interannual variability in migration (e.g. environment- or density- dependent)
2) More than 2 stocks, spatially-disaggregated, contingent on the parallel IOTC stock structure project.

We note that if small time-area closures were to be revisited as the main tool for IOTC management actions, fundamentally different tools might need to be explored. High spatial/temporal resolution models (e.g. SEAPODYM, APECOSM, or single-species equivalents) may represent the best method to usefully represent movement dynamics. These models could not be easily embedded within a traditional fisheries MSE system, but could provide the supporting analyses to estimate how time-area closures would scale up to the low resolution population model.

Age Structure

Age-structured population models are fairly standard these days, and are strongly justified in assessment and operating models for tuna which have highly variable size/age dependent selectivity with different gear types. Aside from an aggregated population, the only likely alternative would be a length-structured model. We are not aware of any strong reason to pursue a length-based model for these species, noting that that key length-based processes can probably be described within an age-based model (e.g. using growth morphs as described below, or pseudo-length-based selectivity)

Sex Structure / Growth morphs

There is compelling evidence that mature YFT and BET exhibit sex dimorphism. Growth morphs in a population model refer to partitioning an age-class into sub-units that exhibit different length-at-age, which are potentially differentially vulnerable to harvesting (and can capture some of the long-term effects of size-based selectivity on the size structure of a population). Sex disaggregation could simply involve the use of two growth morphs, but could involve other sex-specific characteristics including selectivity, migration patterns or natural mortality (none of which are well understood for tuna). It is not clear that the management implications of ignoring sex differences would be sufficient to justify sex-disaggregation in the operating model, and we certainly expect that other factors would be more critical for assessment and management. We would expect that including a sufficiently broad range of uncertainty in the basic tuna life history parameters of a sex-aggregated model would encompass the key HCR performance characteristics arising from sex disaggregation.

Option priorities:

1) Sex-aggregated
2) 2 sexes
3) 2 sexes plus growth morphs to capture size selective fishery effects

Fleet structure and behaviour

We would propose to model each nation / gear type / spatial region combination as an individual fishery, with some shared characteristics (e.g. because of poor data). Each fishery will be designated as either a participant or non-participant in the HCR. OM scenarios for non-participant fleet behaviour would default to constant recent catch or effort levels, with a provision for an explicit time series of future catch or effort (i.e. consistent with fleet development plans). If some of the smaller fleets are not likely to be regulated by the HCR and/or do not require output of individual performance measures, they would be pooled.

Stationarity in key OM characteristics

Stock assessment models generally have a number of stationarity assumptions required to make computationally tractable estimators. e.g. Nobody really believes that M is constant over time, but since there is unlikely to be evidence to the contrary, and not much chance of directly estimating time-varying M, this is an assumption of convenience. For many of these characteristics, a substantial degree of stochastic variability can make surprisingly little difference to the performance of a Harvest Strategy, however, sudden sustained shifts, or long term trends can be very problematic, particularly if they are not recognized or are difficult to estimate. E.g. If growth changes annually, but there is a good time series of observed length-at-age, this is not likely to cause a big problem for an assessment (e.g. CCSBT uses a varying length-at-age time series derived from multiple tagging programmes). However a gradual shift in selectivity toward smaller ages, or increasing catchability of fleets that are assumed to have CPUE proportional to abundance could be very misleading for an assessment and HCR. While non-stationarity can be very difficult to estimate, it is easy to simulate speculative scenarios. We propose to simulate a form of non-stationarity using auto-correlated random deviates for key model features, potentially including:

- Recruitment deviations from the stock-recruitment relationship
- CPUE observation errors (and/or catchability)
- Selectivity

The auto-correlation can be thought of as an approximation for a number of processes (that are challenging for the HCR) such as gradual changes to the distribution of fishing fleets and fish populations, or long-term trophic interaction effects on larval survival, etc. The time series of deviations between predictions and observations in the conditioning model provide an option for estimating the (minimum) level of auto-correlation to use in the OM (but this is discussed further below in the context of CPUE observation errors).

Simulated data collection

Any capacity for an HCR to make feedback-based management decisions will depend on the quality and information content of the data that is collected and used within the HCR. Unfortunately the data quality is not always well understood and probably over-estimated in many cases. In particular, the accuracy and precision of standardized CPUE as a relative abundance index is questionable.

The CCSBT has had a long-running debate about the utility of longline CPUE in assessments and the Harvest Control Rule, in large part due to i) a prolonged contraction of the effort distribution (creating uncertainty about fish density in unfished times and areas), ii) shifting targeting practices and iii) uncertainty about the origin and implications of unreported over-catches. The first two issues are clearly an issue for Indian Ocean yellowfin and bigeye tunas as well, with targeting shifts being even more prominent. Figure 4 illustrates the recent changes to the Indian Ocean longline effort distribution largely driven by Somali piracy. CPUE variance and autocorrelation for the OM projections can be estimated from the residual characteristics (annual observations and conditioned OM predictions). This ensures that the projected errors should be consistent with the conditioning model fit to the data. However, this is a minimum error magnitude that does not account for systematic CPUE trend biases that result from the limitations of the CPUE standardization. Without a reliable independent abundance index it is difficult to speculate about the expected error structure of CPUE indices. Figure 5 illustrates standardized Indian Ocean BET longline CPUE time series estimated from 3 longline fleets. The three series are very different, despite similar analyses to standardize the data. There is a tendency in tuna RFMOs to assume that the Japanese CPUE series is preferable, but this is the most important information in the assessment (along with total catches), there are lots of plausible mechanisms suggesting that there may be problems, and there really is no way of knowing how reliable these series are.

Despite considerable effort spent attempting to understand SBT CPUE, there has long been a strong desire to use fisheries independent research data to overcome the known and suspected CPUE problems. This has included conventional tagging, aerial surveys, transect acoustic surveys, and close-kin genetic mark-recapture studies. The close-kin mark-recapture studies appears to have finally provided a reliable fisheries independent index of spawning biomass which can be used in stock assessment and harvest control rules (Bravington et al 2012). This method (along with the genetic analogue of conventional tagging) bypass many problems of conventional tagging including unknown reporting rates and tag shedding. The transferability and cost-benefit analyses for the approach have not yet been undertaken for many species, but we note that this may prove feasible for yellowfin and bigeye populations.

Proposed OM simulated data options:

1) Essential
    - Total catches with lognormal observation errors
    - Relative abundance indices (standardised CPUE) with auto-correlated lognormal observation errors
    - Size composition – sampled with an appropriate effective sample size with size- and time-dependent correlation structure

2) Possible – if the uncertainties and limitations around CPUE-based abundance indices are recognized as debilitating to the HCR, tag-based HCRs should be explored:
    - Continuous or episodic conventional tagging
    - Genetics-based mark-recapture methods (conventional or close-kin based)

Conditioning the operating model

Conditioning (or estimating initial states and parameters with a statistical model fit to real data) is not essential in MSE, but is desirable. Conditioning ensures that the fishery simulator is plausibly consistent with the available data. If the OM is not conditioned, it may be difficult to defend a recommended management action, because it may be predicated on an OM that is demonstrably not consistent with the data. There is generally a recognition that the OM should encompass a broader suite of uncertainties than a traditional stock assessment model (though we would argue that most stock assessments tend to understate the key uncertainties). Stock assessment problems generally have more unknowns than data, and tractable estimators can only be formulated by imposing constraints (e.g. Schnute and Richards 2001). There are typically a number of models that would fit the data equally well, but they are never formulated, or they are dismissed in the interest of parsimony (and ease of formulating management advice on the basis of a single model). However, these other models may be more realistic and could have very different implications for management than the selected model(s). We propose the following general process for conditioning (roughly modelled on the CCSBT process), in consultation with the appropriate IOTC working parties:

1) Identify the key "inestimable" uncertainties that need to be represented in the fishery simulator, those that are likely to have a non-trivial effect on management performance, but which generally cannot be reliably estimated from the available data (e.g. M, stock-recruit steepness, stock structure and migration rates, etc.).
2) Specify a plausible range of point values for each option (e.g. stock-recruit steepness = 0.65, 0.8, 0.95)

3) For each combination of the quantities from 2, fit an assessment model to the available data, and estimate the remaining parameters (i.e. those that are considered to be estimable conditional on the fixed inputs, including numbers-at-age, selectivity, catchability, recruitment time series, etc.). For computational reasons, the number of options needs to be carefully restrained (e.g. 5 factors each with 3 values results in 243 models)
4) Identify a suite of models (or relative weighting scheme for all models) that adequately represents the plausible uncertainty. In some cases, the likelihood associated with the models from 3 may provide guidance on relative weightings (but we would be cautious in interpreting likelihoods literally given that we know some assumptions are unrealistic, data are uninformative about some parameters, and likelihoods are not always directly comparable). In some cases, a large number of models may suggest very similar dynamics, and it may not be worthwhile including all options in the MSE, because it is the more extreme OM options that provide interesting contrast to help identify robust management procedures.
5) More extreme, but less likely, OM options might be retained as a "robustness set" of options that are used to identify key uncertainties that might require additional information to resolve, or demonstrate performance differences among management procedures that are otherwise similar.

We note that the approach described above places an emphasis on what we tend to call model uncertainty (though some of the factors are more appropriately described as parameter uncertainty). The emphasis on point estimates ignores parameter estimation uncertainty (e.g. which could be estimated using MCMC for a single or few models). We propose this approach because in our experience,

i) Parameter estimation uncertainty associated with an individual model is generally much narrower than uncertainty associated with a range of model point estimates. Assessment model simulations often suggest that, even in models with very good assumptions, estimated confidence intervals are often too narrow.
ii) For highly parameterized models such as those proposed for the tropical tunas, MCMC chains would probably be computationally prohibitive and prone to poor mixing.

Depending on the management options pursued, there may be a requirement for additional parallel work to estimate how high resolution processes scale up to the units of the assessment models. At this time, we have serious concerns about the incompatibility between high resolution tropical tuna tag dynamics and the low resolution assessment models, which may result in sizeable estimation biases within an assessment (e.g. Langley and Million 2012, Kolody and Hoyle, in press). In the first instance, we propose to evaluate a broad range of movement rate options as inestimable in the scheme proposed above, and conduct independent conditioning for each movement rate assumption. In this case tags would probably not be explicitly included in the conditioning, but external tag analyses are recognized as a method of assigning relative credibility to different migration scenarios (particularly if this proves important for HCR performance).

Existing Fisheries Models that could be extended for the Indian Ocean YFT/BET Operating Model

The following provides a brief description of potentially relevant models, noting that the descriptions represent the authors' current understanding which may be outdated, or otherwise inaccurate.

Stock Synthesis / Multifan-CL – These are two general purpose stock assessment packages frequently applied to tuna populations, including BET and YFT in the IOTC. They have a range of feature options (contrasted with the CCSBT OM in Table 2). They could be used to condition the YFT and BET operating models, but neither is designed to function as an operating model in an iterative feedback-based management simulation context (though it is possible that they can be used in this way with some ad hoc file manipulations). These models would not be the ideal starting point if new features were required.

CCSBT (Southern Bluefin Tuna) OM (e.g. CCSBT 2013)- A key decision for specifying an operating model relates to whether one wants to use the same dynamics for the conditioning and future projections. As far as we are aware, the "designed by committee" CCSBT operating model represents the most ambitious attempt to use the same model for both objectives, at least for any tuna-like species. This purpose-built software (which underpins the only active simulation-tested RFMO Management Procedure) is designed to include a number of structural uncertainties specific to the needs of the CCSBT. It is freely available, implemented in ADMB (with supporting R summary routines) and has exposure among several nations that are involved with both CCSBT and IOTC. This is potentially a good starting platform to adapt for IOTC needs.

Indian Ocean SKJ OM (Bentley and Adam 2014) – The OM under development for SKJ has some obvious synergies with YFT and BET requirements. The model represents a single stock, with migration among three regions (western, Maldives EEZ, eastern). There is no explicit statistical fitting facility for conditioning to the available data and two conditioning options have been proposed. The first option involves stochastic sampling of model parameters from priors, with parameter combinations retained if the model predictions meet a number of selection criteria, e.g. CPUE(year=X) < CPUE(year=Y). In principle, the selection criteria could be statistically-based, but to date the criteria explored have been more ad hoc. This approach should lead to a broader representation of uncertainty than a statistical approach, which is probably a good thing for SKJ (which have poorly understood relative abundance indices, and are thought to have good stock status currently). This is an interesting idea, however, it is also largely untested (and we note that overstatement of uncertainty has the risk of overly precautionary management for fully-exploited stocks). The second conditioning option involves adopting parameters from the WPTT SKJ stock assessment, which has the added challenge of reconciling the structural inconsistencies between the two models (i.e. notably in terms of spatial structure and migration).

Indian Ocean Albacore (ALB) OM (Iago Mosqueira, JRC, Italy, pers. Comm.) – The model under development conforms to fairly standard population dynamics assumptions (e.g. single-stock, spatially aggregated). Conditioning involves adopting the parameters from Stock Synthesis (based on, but likely expanded from the WPTmT stock assessment).

ICCAT Northern Bluefin Tuna OM (Tom Carruthers, UBC, pers. comm.)– There is a plan underway to develop an NBT OM which is expected to include considerable spatial complexity (including multiple stocks and migration between areas), with the capacity to condition and conduct simulation projections from the same software. While development has not progressed very far to date, there are potential synergies between the NBT and YFT/BET requirements.

Harvest Control Rules

The key feature of the MSE system that we want to establish is the means to test alternative HCRs. This does not require the HCR to be embedded in the OM software, as long as the OM platform can export the simulated data files, call an external subroutine and import the final management recommendation. As part of the MSE process, this could be achieved with different levels of detail:

HCR Option priorities:

1) Synthetic HCR - use the known stock status from the OM, with a specified error distribution, to calculate a management action. These simulation results would be meaningful if assessment errors were small or one truly understood the error characteristics of the assessment process.
2) Empirical decision rule - convert the simulated fisheries data directly into a management recommendation. It has often been shown that relatively simple HCRs can produce very good management performance when tailored to a specific MSE problem (particularly over the relatively short time periods over which an HCR is likely to be operating without review). These have the advantage of being easily interpretable by stakeholders.
3) Simple model-based decision rule – involves fitting an assessment model of some sort and using the model estimates to set the management action. This approach is more likely to "learn" about the system than an empirical rule, but has downsides in that i) it is more computationally demanding to evaluate the HCR, and ii) automating the model fitting process (including testing for implausible results) is not always straightforward.

The first approach is obviously the simplest and fastest, but it not likely to realistically reflect the errors that would arise from the data and assessment methods (particularly the time series structure), unless the approach was carefully calibrated with simulations. The latter two approaches should more realistically capture the error structure of the real HCR, and enable one to meaningfully compare alternative HCRs, including evaluating the information content of different data types in a cost-benefit analysis. A fourth option, of including a full stock assessment as the basis of the HCR, is not considered realistic in this context, because Indian Ocean tuna assessments are rapidly evolving, with changing inputs, supporting analyses and assumptions that cannot be meaningfully simulated at this time.

We note that the specific form of the HCR should not matter, i.e. it does not necessarily need to emulate an assessment, or directly estimate stock status relative to any reference points. An effective HCR should extract useful information from the data (e.g. learn how to distinguish productive and unproductive stocks, or identify poor recruitment), but how this is achieved is not important as long as it performs well with respect to the management objectives, and is robust to the uncertainties in the system.

Conclusions

We propose this list of software options, potential specifications for Indian Ocean YFT and BET operating models and conditioning options, for the consideration of the WPM and other interested parties for feedback.

References

Adam MA, Sharma R, Bentley N. 2013. Progress and Arrangement for Management Strategy. Evaluation of Indian Ocean Skipjack Tuna. IOTC–2013–WPTT15–33 Rev_1

Bentley, N, and Adam, MS. Management strategy evaluation for Indian ocean skipjack tuna : first steps (Bentley N & Adam MS). IOTC–2014–WPTT16–39

Bravington, M., Grewe, P., Davies, C., 2012. Report of the close-kin project: estimating the absolute spawning stock size of SBT using genetics. Commission for the Conservation of Southern Bluefin Tuna working paper CCSBT-ESC/1208/19.

CCSBT 2013. Report of the eighteenth meeting of the Scientific Committee, 7 Sep 2013, Canberra, Australia.

de la Mare WK. 1986. Simulation studies on whale management procedures. Thirty sixth Report of the International Whaling Commission. Document SC/37/O14.

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

IOTC 2011. Report of the fifteenth session of the Indian Ocean Tuna Commission. Colombo, Sri Lanka 18-22 March 2011.

IOTC 2012. Report of the sixteenth session of the Indian Ocean Tuna Commission. Fremantle, Australia 22-26 April 2012.

IOTC 2013. Report of the seventeenth session of the Indian Ocean Tuna Commission. Mauritius, 6-10 May 2013.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M. A., Poos, J. J., Scott, F., and Scott, R. D. 2007. FLR: an open-source framework for the evaluation and development of management strategies. – ICES Journal of Marine Science, 64: 640–646.

Kolody D, Grewe, P, Davies, C, Proctor, P. 2013. Are Indian Ocean tuna populations assessed and managed at the appropriate spatial scale? A brief review of the evidence and implications. IOTC-2013-WPTT15-13.

Kolody, D. and S. Hoyle. In press. Evaluation of tag mixing assumptions in western Pacific Ocean skipjack tuna stock assessment models. Fish Res http://dx.doi.org/10.1016/j.fishres.2014.05.008

Langley, A., Million, J., 2012. Determining An Appropriate Tag Mixing Period for theIndian Ocean Yellowfin Tuna Stock Assessment. Indian Ocean Tuna Commission14th Working Party on Tropical Tuna Working Paper IOTC-2012-WPTT14-31.

Langley, A, Herrera, M, Million, J. 2012. Stock assessment of yellowfin tuna in the Indian Ocean using Multifan-CL. Indian Ocean Tuna Commission Working Party on Tropical Tuna 2012 Working Paper: IOTC-2012-WPTT14-38 Rev_1.

Schnute, J.T., Richards, L. 2001. Use and abuse of fishery models. Canadian Journal of Fisheries and Aquatic Science 58: 10–17.

Sibert, J., Hampton, J., 2003. Mobility of tropical tunas and the implications for fisheries management. Mar. Pol. 27, 87–95.

Tong, Y., D. Kolody, X. Chen, and Y. Chen. 2011. Evaluation of three harvest control rules for bigeye tuna (*Thunnus obesus*) fisheries in the Indian Ocean. IOTC–2011–WPTT13–INF15.

Table 1. Proposed set of key features for IOTC YFT and BET operating models

Feature	Initial Proposal
Space-time structure	
Spatial-structure	Disaggregated (west and east)
Time-step	quarterly (annual reporting)
Biology	
Age-structure	Yes
Sex-structure/growth morphs	No
Stock-structure*	1-2 stocks
Stock-recruit function	Beverton-Holt
M	Stationary, age-dependent
Recruitment deviates	Lognormal auto-correlated
Length-at-age	Stationary, best estimate(s) from the WPTT
Migration	bulk transfer (seasonal if necessary)
Maturity	Stationary
Fishery	
Fleets	Minimum number to describe HCR-regulated fleets
Selectivity	Stationary age-based
Catchability	Stationary with time-dependent option
Simulated Data for HCR	
Catch	Fleet-specific observation error
CPUE	Observation errors specified by variance and temporal correlation
Size composition	Effective sample size with temporally correlated bias
Age composition	None
Tag dynamics	Not to be included unless/until such time as tagging is accepted by IOTC as a viable basis for implementing a HCR (noting that close-kin genetic tags may be the preferred option)
Management Controls	
Catch controls	TACs removed with lognormal implementation error
Unregulated Fleets	Effort remains constant at current levels (or specified time series)

*Stock structure to be reviewed in parallel with Indian Ocean Stock Structure Project

Table 2. Assessment package features available for conditioning tuna operating models. We note that CASAL shares many of these features as well, but has a limited history with complicated tuna configurations.

Feature	CCSBT-OM	Stock Synthesis	Multifan-CL
Generic flexible design	N	Y	Y
Age-structure	Y	Y	Y
Sex-structure	N	Y	In testing
Spatial-Structure	N	Y (single spawning stock, but multiple morphs with different migration)	Y (single spawning stock, no migration morphs)
Recruitment spatial options	N	?	Deviations constrained among regions
Stock-Structure	N	N (see spatial)	N (see spatial)
Length-at-age	Temporally variable	Temporally variable option (?)	stationary
Selectivity	Age-based, temporally variable	Age, pseudo-length–based, or length-based with platoons; temporally variable option	Age or pseudo-length–based (stationary?)
Relative abundance indices, size composition data	Y	Y	Y
Conventional tag dynamics	Y (Non-spatial)	Y (spatial)	Y (spatial)
Close-kin genetic tag dynamics	Y (Non-spatial)	N	N

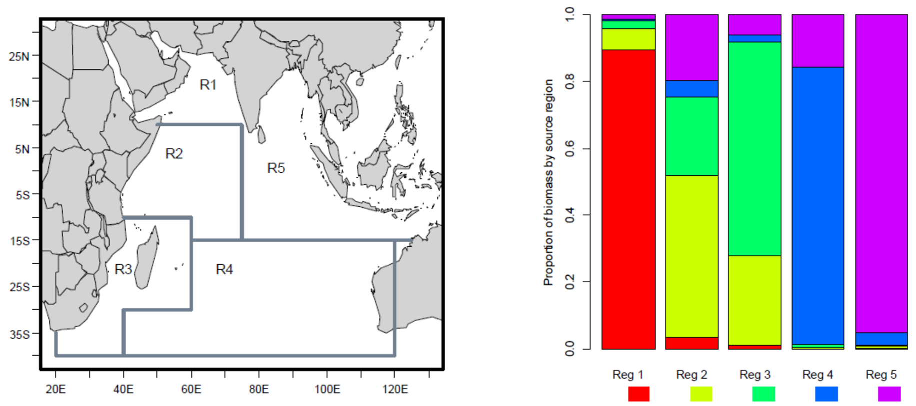

Figure 1. Left panel: IOTC YFT stock assessment spatial structure. Right panel: YFT movement summary estimated from the reference case assessment (i.e. . (from Langley et al. 2012).

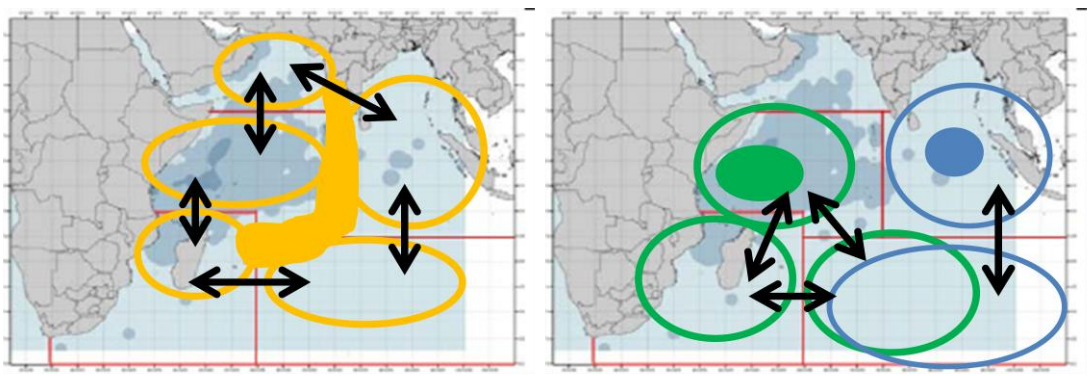

Figure 2. Schematic representation of a single stock, spatially-disaggregated YFT assessment model (left panel), and a hypothetical 2 stock model (right panel). The solid blobs represent spawning areas, and the ellipses represent internally homogeneous foraging areas for the corresponding spawning areas, arrows indicate migration links.

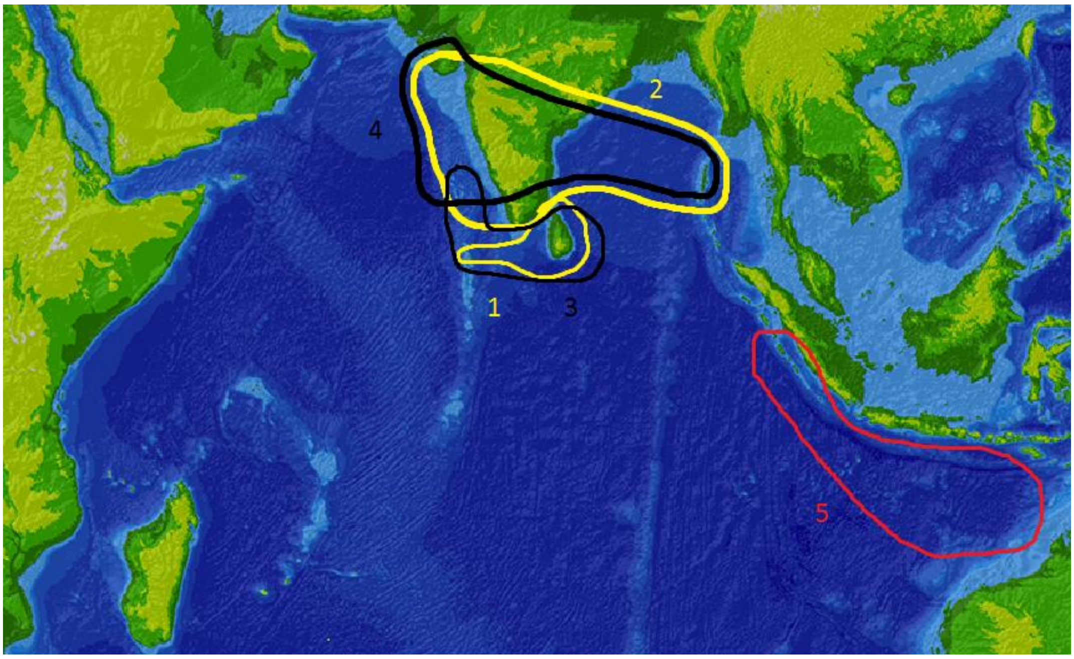

Figure 3. Genetic evidence for structured tuna populations has been reported for studies within each of the regions approximately outlined above (1-5=Dammanogado....). SKJ = black, YFT = yellow, BET = red. (from Kolody et al 2013)

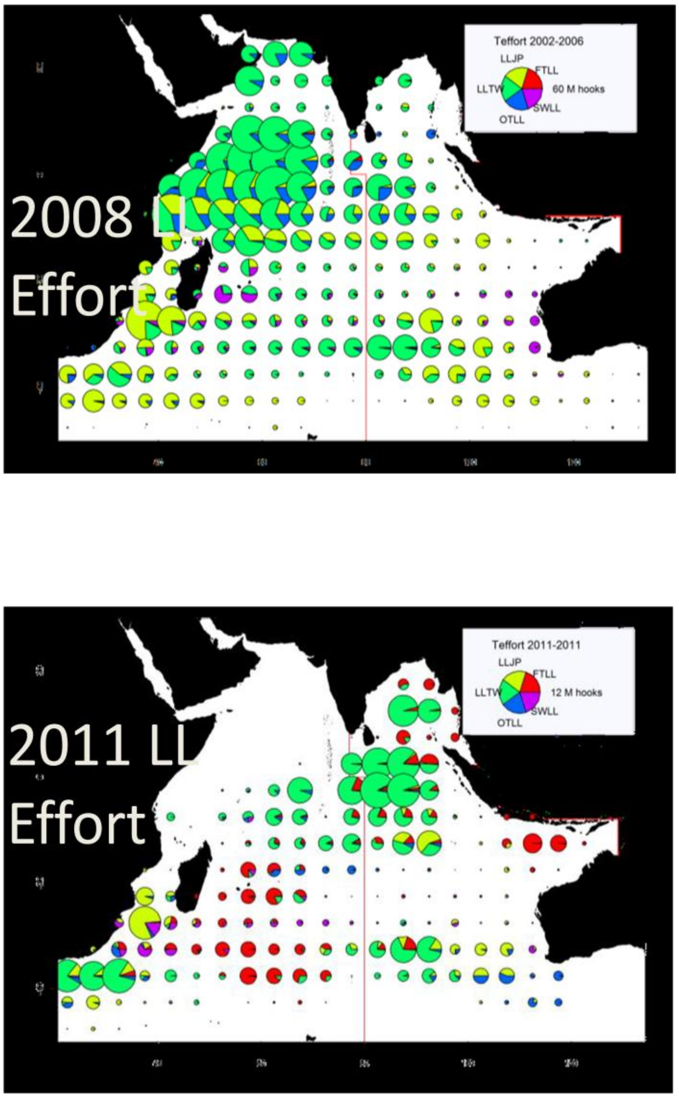

Figure 4. Changes in the longline effort distribution in recent years, largely due to Somalian piracy (from IOTC WPTT 2012).

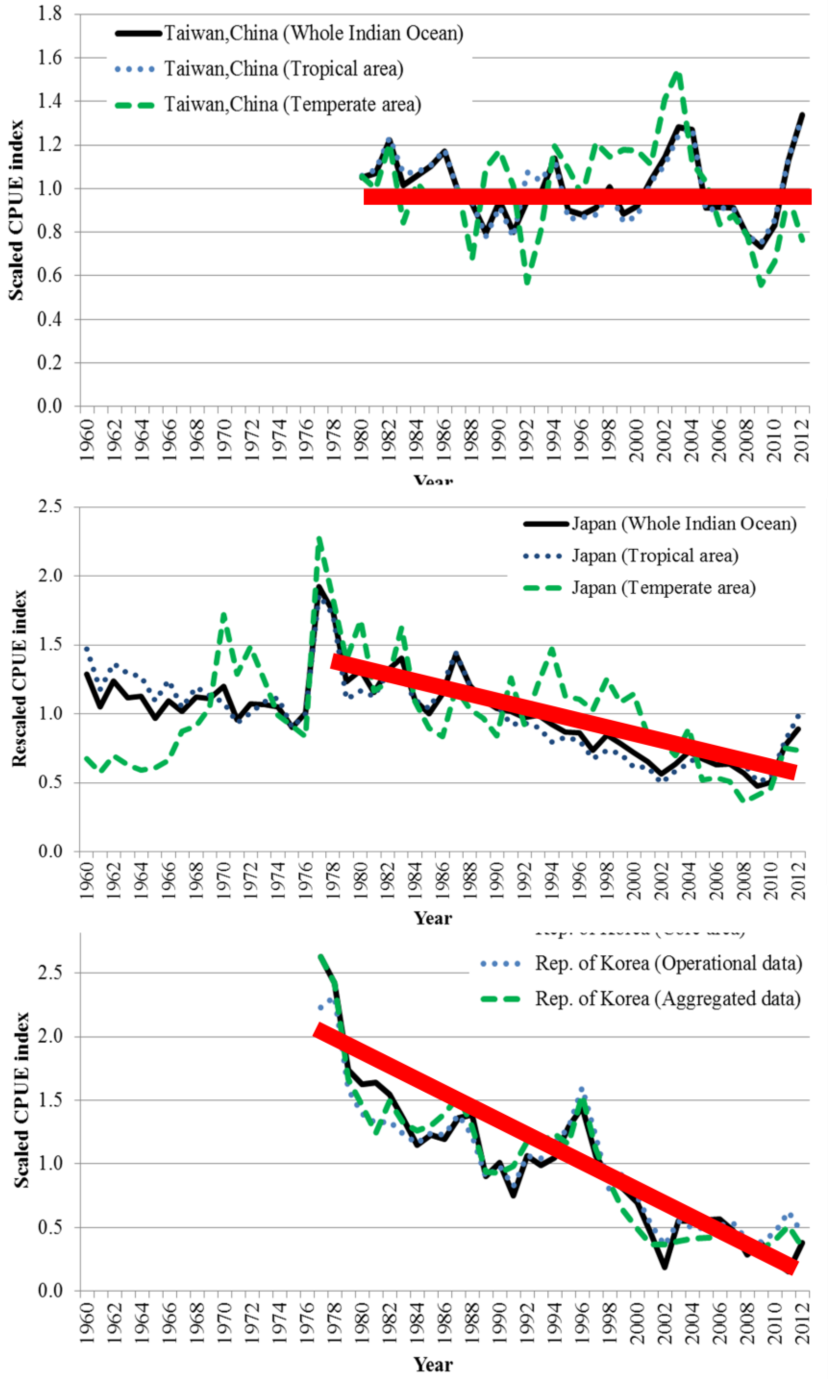

Figure 5. Comparison of Taiwanese, Japanese and Korean bigeye tuna longline CPUE trends in the Indian Ocean (from IOTC WPTT 2012).