Received: 5 October 2015
IOTC-2015-WPM06-10

Indian Ocean Tuna Commission Working Parties on Methods

Indian Ocean Yellowfin and Bigeye Tuna Management Strategy Evaluation Development Framework – Draft Progress Update

Oct 2015

Dale Kolody[1], Paavo Jumppanen[1], Tom Carruthers[2], Adam Langley[3]

Contact: Dale.Kolody@csiro.au

[1]CSIRO Oceans & Atmosphere, Australia

[2]UBC Fisheries Centre, Canada

[3]Trophia, New Zealand

Abstract

Recent progress on the development of a Management Strategy Evaluation (MSE, or Management Procedure Evaluation) technical framework for Indian Ocean yellowfin (YFT) and bigeye (BET) tunas is described. This includes i) an outline of the key software features implemented to date, ii) an exploration of YFT Operating Model (OM) options (conditioned using Stock Synthesis software in association with the draft 2015 assessment), and iii) an outline of the software development plan through to mid-2016.

We emphasize that this technical project is only one part of a much larger MSE process that requires the engagement and exchange of ideas among many parties, including technical experts that will need to contribute to the review and development of operating models and management procedures, and various stakeholders (including fisheries managers and IOTC Commissioners) that will need to articulate their expectations about management objectives and options. This specific component of the project is scheduled for completion mid-2016, so this presentation represents the primary opportunity to solicit feedback from the general participants of the IOTC WP Methods, WP Tropical Tunas, and Scientific Committee. We welcome feedback about the defined feature set for the projection model, and the approach to Operating Model conditioning.

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 implementation of an MSE approach typically involves the following components (though generally achieved through an iterative process that involves revisiting some steps in an iterative cycle):

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) 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.

3) Candidate Management Procedures (harvest strategies) – the specification of 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).

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 document pertains to a technical project to advance the software required to address components 2-4 in the above process for bigeye and yellowfin tunas and represents an update to Kolody et al. (2014), which outlined the original workplan. This software is only one part of the much larger process that requires the engagement and exchange of ideas among many parties, including technical experts that will need to contribute to the review and development of operating models and management procedures, and various stakeholders (including fisheries managers and IOTC Commissioners) that will need to articulate their expectations about management objectives and options. The software developed in this project is designed to be suitably flexible to describe a number of potentially important fisheries characteristics. However, these are very complex systems, and the example operating models and management procedures are intended primarily to demonstrate the available feature set, rather than provide specific results upon which management recommendations might be formulated. It is expected that scientists, managers and stakeholders will request additional modifications to the fisheries dynamics, uncertainties, and performance indicators, to reflect the diversity of opinions and experiences within the IOTC. It is hoped that the operating models and management procedures can be updated with minimal modification to the core computational framework, such that this might be achievable by knowledgeable fisheries scientists that do not need to understand all of the MSE code.

The remainder of this document is presented in the following sections:

1. Outline of software for the simulation framework
2. YFT operating model demonstration case
3. Considerations for the BET operating model
4. YFT interim MSE results
5. Candidate Management Procedures
6. Development timeline and priorities
7. Request for feedback from the IOTC WPM, WPTT and SC

1. Outline of software for simulation Framework

Following a review of the current MSE needs and existing software options, we opted to build on pre-existing initiatives. The intent in doing this was to save development time, increase code reliability (i.e. through independent testing by different teams), and hopefully contribute to tools that can be used more broadly across tuna-RFMOs.

MSE Projection code – this includes a number of sub-routines for representing uncertainty, projecting the population, calling management procedures, simulating the fishery, simulating future data collection, and summarizing results. We opted to build on the platform described in Carruthers et al. (2014), designed for Atlantic Bluefin Tuna (ABT). The main advantages of this platform include:

- It is completely self-contained within R, has built in parallelization for distributing the workload across a cluster, it is very portable across platforms, and is accessible to anyone without investment in proprietary software or lower level language programming skills.
- The platform is designed to include many complications of interest for tunas (some of which have not received much attention in most MSE applications to date). The main features include (but are not obligatory):
    - Age-structured
    - Spatial-disaggregation
    - Multiple populations with independent biological characteristics
        - M, Length-at-age, weight-at-age, migration
    - Multiple fishing fleets with age-dependent selectivity
    - Seasonally-structured parameters (migration)
    - Stochastic variability in parameters
        - including realization-specific deviations
        - autocorrelation and trends

    - Standard fisheries data simulation
        - Catch-at-age, catch-at-length, abundance indices
    - A number of MPs have already been coded for the ABT evaluations
    - Implementation error in quotas

However, the platform is not ideal for the IOTC YFT-BET MSE system, and requires several modifications:

- The ABT implementation was designed for MPs that recommend a single annual quota only. The modified software will be able to evaluate an arbitrary mix of fleet-specific catch quota and effort recommendations (which may include scenarios of some fleets not participating).
- The ABT approach was designed for a relatively long-lived species with annual recruitment, while the tropical tunas need to represent faster biological processes (including multiple spawning events within years and seasonal growth, maturity, etc.).
- The ABT implementation cannot evaluate multiple species management simultaneously. We hope to extend this capacity after the basic single species dynamics are implemented and tested. However, Management Procedures for multi-species situations are not well developed, and we do not expect to progress much beyond a demonstration of the technical feasibility in this project.

The ABT platform is also reliant on a certain fluency in multi-dimensional array manipulations to maintain computational efficiency, which are not intuitive to R novices. Furthermore, some of the implementation details differ from the standard assumptions that are likely to be used in stock assessment models. Notably, the ABT implementation does not use the Baranov catch equations, which reduces the computational overhead by avoiding iterative solutions. Solving the Baranov catch equations within R is slow owing to the use of numerical approximations to the gradient of the minimization objective function, or worse still minimising without a gradient function. By re-writing the projection model component as an R callable C++ dynamic library with an automatically differentiated objective (machine precision gradient rather than an approximation) we should be able to construct a workable solution using ADT and ADlib. ADlib is a C++ library that provides a mechanism for constructing dynamic multi-dimensional arrays that can be dereferenced in the same manner as statically defined ones. This makes for the ability of writing more readable and maintainable code without sacrificing performance. The arrays themselves are constructed as contiguous blocks of memory as with all arrays within R internally and ADlib provides the means to make accessing R created arrays within the C++ code equally simple. The command line tool ADT provides, in conjunction with the automatic differentiation tool Tapenade (developed by INRIA) the means to create differentials of functions in our C++ code, thereby providing the means to implementing a gradient to the objective function. In addition to the automatic differentiation support, ADT also provides support for automatically creating interface code between R and C++ thereby relieving the programmer of that onerous task. This makes for faster and less buggy code development.

Appendix 1 compares the ABT approach with the usual Baranov approach, and concludes that i) a modified version of the ABT approach should be adequate to maintain reasonable consistency with the Baranov approach (i.e. <2% inconsistency in calculated values of N, M and F). However, the numerically-efficient external sub-routine should allow for greater speed, an independent check on the integrity of the code, and greater consistency with standard assessment models.

Software Development Version Control – Given that this software is to be freely available to the broader IOTC community (and presumably other tuna-RFMOs) for use and further development, it will be distributed and managed through a formal version control system. It is initially being implemented using git/stash on an internal CSIRO repository, and will be made public (possibly migrated to gitHub) when the basic feature set is complete and the software is stable.

2. Yellowfin Operating Model demonstration case

In the following, we use the term Operating Model (OM) to refer to the structural assumptions and parameter estimates that describe the populations of interest, while the specific software that implements the dynamics is referred to as the projection code. There are a number of ways in which the parameters of the operating models can be defined. "Conditioning" is the process of estimating the parameters (e.g. with a statistical stock assessment model), which ensures that the simulation dynamics are reasonably consistent with the historical data, and perceptions of stock status derived from the stock assessment process. Accordingly, we are using Stock Synthesis (Methot and Wetzel, 2013) software for the conditioning, and attempting to be consistent with the main insights from the stock assessments prepared for the WPTT. However, we also recognize that a key objective of the MSE should be a demonstration that Management Procedures are robust to assessment assumptions and uncertainties that are plausible, but not necessarily described in the assessment. We expect the conditioning undertaken as part of this project will differ from recent stock assessments in a number of ways, including:

- The YFT and BET models will be harmonized to the same spatial, temporal and fleet structure. This is required in a multi-species (technical interaction) context, because the two species are often caught together, and some by-catch is inevitable.
- Scenarios with potentially different stock structures will be explored (recognizing that further insight into this issue should arise as a consequence of a number of recent stock structure research initiatives (including the recent IOTC call for EOIs for a comprehensive stock structure quantification project).
- Non-stationary recruitment dynamics (and potentially other biological parameters) will be simulated, in the first instance including strongly auto-correlated recruitment deviations (where the correlation is derived to be consistent with the point estimates from the SS model).
- Simulated relative abundance indices (CPUE) will also have the capacity to represent non-stationary patterns (e.g. strongly auto-correlated patterns, including trends in catchability).

Describing uncertainty in the context of a comprehensive stock assessment and MSE is a challenging process. Understatement of the uncertainty leads to an elevated level of risk (whether or not it is recognized), while overstatement may lead to lost economic opportunities from overly precautionary management. The approach we have proposed to use emphasizes the importance of "model uncertainty" (including different structural assumptions, and a range of fixed values for key parameters that are known to be difficult to estimate). This approach is expected to provide a broader representation of uncertainty than the statistical uncertainty conditional on any individual model. Furthermore, as Langley (2015) discusses, it does not appear to be technically practical to attempt to describe the statistical uncertainty of models with this level of complexity using Bayesian approaches.

Some exploratory model ensembles for YFT Operating Model have been derived from the current Stock Synthesis assessment (Langley 2015), including the same time/area strata and fleet definitions and most core structural assumptions. However, this work was pursued in parallel with and diverged from the assessment in some important ways (comparable models may not be identical to the assessment report because of minor changes, including minimization controls). These ensembles are presented primarily to illustrate the process being undertaken and to seek feedback from the Working Parties about conditioning priorities.

The preliminary YFT OMs are identical to the reference case assessment with respect to the following general structural characteristics:

- Single species, single stock
- Age-structured
- Sex-aggregated
- Stationary growth, mortality and stock recruitment relationship
- Multiple fleets, each with stationary selectivity (in forward projections, some fleets have historical temporal variability)
- Conditioning consists of estimating the parameters which result in the best fit to the following data:
    - Total catch in mass or numbers (depending on the fleet)
    - CPUE (fishery-selected relative abundance indices)
    - Catch-at-size composition
    - Tag recaptures (not all scenarios)
- The estimated states/parameters include:
    - mean unfished recruitment
    - recruitment deviation time series (including spatial deviations in some cases)
    - fishery selectivity
    - fishery catchability
    - Numbers-at-age (last year in the assessment determines the first year for the MSE projection code)
    - tag reporting rates (not relevant for the projections)
- A number of biological parameters are fixed (possibly at multiple plausible values in different OMs), including:
    - M
    - Beverton-Holt stock recruit steepness
    - length-at-age
    - mass-length relationship
    - maturity schedule

The reference case assessment included four spatial regions (Figure 1), while we have considered 4 area and 2 area options.

OM ensembles 1-4 are described in Tables 1-4. Model ensemble 1 is the most similar to the base case assessment, and emphasizes the critical importance of uncertainty in M and steepness.

Ensembles 2-4 explored some options that are potentially of interest to the MSE, but did not receive a strong emphasis in the assessment (i.e. forcing alternative migration scenarios, reducing the influence of the tags, and reducing spatial complexity to two regions). These and other OM considerations are discussed under the headings below.

Figure 2- Figure 5 show the stock status summaries of the MPD estimates of OM ensembles 1-4. All of these ensembles demonstrate a reasonable consistency with the reference case assessment, with the range of models bracketing the reference case situation (though the reference case is not necessarily near the middle of the ensemble). As a model ensemble typically includes a large number of models, they are generally not scrutinized with the level of detail attributed to individual assessment outputs. However, as it is essential to ensure that the models are not completely absurd, we tend to look at a number of aggregate summary diagnostics as shown in Figure 6, which can be used to identify model convergence, quality of fit, and stock status characteristics in relation to the input assumptions.

As part of the current project, we will not be attempting to identify the definitive operating models for MSE decisions, but seek to ensure that the demonstration case is plausible, takes on board the key insights from the YFT assessment process, and capable of describing the most important uncertainties in a computationally tractable structure. Toward this end, we note that it is not necessary that all models in the ensemble are represented equally (i.e. if there is a justification for differentially weighting models they can be sampled more or less frequently in the stochastic projections). We also note that the dimensionality of the problem can rapidly become unwieldy, so uncertainty dimensions should be expanded sparingly.

The explorations to date suggest a number of features for consideration, discussed under the headings below.

Environmental movement drivers

Langley et al. (2015) defined a reference case which related migration to environmental variables. While there is no doubt that the environment influences migration, the mechanism is poorly understood. We are not proposing to use this approach for OM conditioning at this time, because i) it is a new development that has not been seriously explored, ii) methods for forecasting environmental variables in the projections would require some thought, and iii) the overall stock status estimates with and without the environmental variables do not appear to be very different. It would be interesting to create dummy time series of environmental variables that reflect mean seasonality, to see if the environmental variables are providing any information about interannual variability, or whether they are operating primarily as a seasonal signal (which might be more conveniently handled with a seasonal SS implementation).

Tag dynamics

The Indian Ocean Regional Tuna Tagging Programme was very successful at releasing and recapturing a large number of YFT tags. However, there are a number of concerns about the information content in the tags in an assessment context because i) the tag release design is very unbalanced, ii) tag reporting rates outside of the western purse seine fleets are poor and unknown, iii) tag mixing is much lower than would be expected to meet tag estimator assumptions for estimating M, F, N and reporting rates (e.g. Langley and Million 2012, Kolody and Hoyle 2013, Carruthers et al. 2014).

Accordingly, the OM grid was designed to admit that the tags might be misleading and warrant downweighting (by a factor of 1.0, 0.1, 0.0) in the likelihood. The value of the weighting factor appeared to make made very little difference to the aggregate stock assessment dynamics examined (e.g. Figure 8).

Population structure

The Langley et al. (2015) reference case consisted of 4 areas (Figure 1) with movement estimated among regions. In principle, it is desirable to break the model up into relatively homogeneous units (i.e. to reduce the impact of assessment errors that arise from pooling heterogeneous model features like selectivity), and to explain observations on the basis of appropriate mechanisms (e.g. does CPUE have a seasonal cycle because fish move into and out of a region?). Thus we consider the 4 area assumption for the main OM grid to be reasonable. However, spatial complication comes with a high computational overhead in models, and it is notoriously difficult to estimate movement from the available data. Accordingly, we considered that it would also be useful to consider: i) the implications of movement rates that are fixed at very low and high levels (e.g. model ensembles 2-4), and ii) what are the implications of reducing the spatial complexity to 2 areas (ensemble 4 - West and East, i.e. R1a + R1b + R2 and R3 + R4 from Figure 1). Given the observations from tagging studies, spawning habitat, and seasonal CPUE cycles, we expect that a lot of seasonal North-South migrations occur in the Indian Ocean such that there is considerable mixing in the latitudinal direction.

Trends in the pooled (area-weighted) West and East standardized Japanese longline CPUE indices are very similar (Figure 9). This could indicate a large degree of West-East mixing, or a consequence of the ideal free distribution, i.e. mobile fleets seeking out the highest catch rates can potentially maintain similar population densities across vast areas, even if the populations do not mix. While the model fit to the CPUE series is better when movement is estimated, this may be largely because the model has many more free parameters. When migration rates (and recruitment distribution) are highly constrained in the high and low movement scenarios, the model still has a reasonable capacity to fit the general CPUE trends. The worst fit tends to be the recent decline in the eastern region (though this might reflect temporal changes in targeting).

The stock status estimates of ensemble 4 (2 area models) are similar to the 4 area options, but are somewhat more optimistic in terms of stock productivity. This is not evident from Figure 4, but the upper range of the MSY estimates (364-664K t) are much higher than any of the assessment sensitivity tests reported (309-530K, base case = 402K t). While the uncertainty of the 2 area grid does encompass the reference case dynamics, there has not been a systematic evaluation of the credibility of this suite of models. There appears to be a consistent relationship in that the more optimistic models tend to explain more of the declining CPUE trend as a consequence of a recruitment trend. While this is conceivable, and the magnitude of the effect occurs along a continuum, long term recruitment trends are not very appealing within a stationary recruitment paradigm. While the 2 area approach is attractive from a computational perspective, more work would be required to ensure compatability with the assessment.

Recent assessments have assumed that there is effectively a single aggregate spawning population across the Indian Ocean for stock-recruit purposes. However, there is evidence in the Indian Ocean (e.g. Kolody et al. 2013 and references therein) and Pacific Oceans for yellowfin spawning populations that are genetically distinct and identifiable. The evidence in the Indian Ocean is not definitive, but the uncertainty has been important enough to stimulate a substantial research investment to resolve the issue. At this time, there are no data with which to quantify the Indian

Ocean stock structure. Furthermore, despite the RTTP-IO, there are limited data with which to estimate movement across large areas of the Indian Ocean (i.e. because of the unbalanced tag release design, and poor/unknown tag return rates for most fleets). However, there is the capacity to represent speculative alternatives within the OM, and evaluate how sensitive MP performance is to these alternatives.

CPUE assumptions

The assessment model relies on standardized Japanese longline CPUE as relative abundance indices. It remains unclear the extent to which standardization can account for operational changes in searching, targeting efficiency, skipper and crew skill, etc. Given that most Management Procedures are likely to be critically dependent on CPUE as a relative abundance index, it is important to make sure that the simulated series is not unrealistically informative. As a minimum, projected CPUE observations will have a CV and auto-correlation that is consistent with the deviations between predictions and observations from the individual OMs.

Future Stock Synthesis considerations

The OM conditioning framework was based on the assessment, and there should always be a strong link between the two to ensure general consistency between the two. Some obvious features that might be modified in both frameworks include:

- Annual structure with seasons. The current models defines an SS3 year as a real quarter. This is inconvenient in the sense that timescales are confusing and rate processes (M, F, growth etc.) must be defined in quarterly units rather than annual. Downsides with using SS3 in a year/season configuration include: i) tags are assigned integer ages (i.e. the seasonal model does not appear to be able to admit different quarterly ages of tags within a year), and ii) there is a single annual biomass calculation (which should not really matter unless there is a population with very fast dynamics). However, on the positive side, the year/season configuration allows some seasonal parameters to be estimated (notably migration), and tags might not make much difference to the YFT assessment.
- Temporal variability in selectivity. Some fisheries are subdivided in time to represent the potential change in selectivity. This was done for historical reasons in the MFCL assessments, while SS3 has a convenient structure for introducing temporal variability within a single fishery definition.

Bigeye tuna demonstration case

It is expected that BET conditioning will be pursued when the Secretariat is able to provide the data by the same strata as YFT (presumably late 2015 or early 2016). It is not essential to retain exactly the same structure for YFT and BET, if the Management Procedures are to be pursued and implemented independently. However, we are not aware of any obvious disadvantage in seeking a harmonized structure, as it would simplify the whole process of formulating conditioned OMs, communicating results, and ultimately coming up with Management Procedures that consider the tropical tuna technical interactions.

YFT preliminary MSE results

The software for the basic single species, single population, conditioned operating models defined above (for the non-environmental movement scenarios) is currently functional, and able to evaluate candidate MPs that are managed through aggregate quotas (projections assume proportional effort changes). However, preliminary results are not included in this document, because there are some recognized inconsistencies between the SS3 assessment and the projection model that have not yet been evaluated. We do not want preliminary results to distract from the process of reviewing the current stock assessment. Presentation style of results will conform to the standard established by the WPM and the precedents set by the SKJ and ALB cases.

Candidate Management Procedures

Demonstration MPs will initially consist of relatively simple data-based rules, likely those of which have already been implemented within the ABT framework (but adapted to deal with more sophisticated catch and effort options). If time permits, we would be interested in investigating the relative value of collecting different data for MPs, e.g. genetic-tag-based estimators.

Development Timeline and Priorities

- Oct-Nov2015 –obtain feedback from the IOTC WPM, WPTT
- Dec2015-Mar2016
    - develop YFT OM in relation to WPTT feedback
    - develop demonstration case for multi-stock YFT
    - complete software requirements for single species MSE
    - initial conditioning of BET OM
- ~Mar2016 – Software demonstration and review by the IOTC WPM informal sub-group on MSE
- Apr-Jun2016
    - Address workshop feedback priorities
    - Develop multi-species MP capacity if appropriate
    - Finalize software and documentation
- Jun 2016 – Documentation and software released to the IOTC via Git repository

Request for Feedback

The presentation of this document to the IOTC working parties represents the main opportunity for members of the IOTC scientific community to provide constructive feedback on the development path before the primary phase of the software is completed and released to the IOTC community mid-2016. It is expected that further development on the OM conditioning and MP evaluation will be required after that date, but there is no mechanism in place to support that work, so any opportunities to improve the software should be documented, prioritized and pursued within the existing development timeframe to the extent possible. There is expected to be another opportunity for review of the YFT/BET framework at the informal Methods WG meeting around March 2016.

References

Carruthers, T., Kell, L., Davies,C. 2014. Evaluating management strategies for Atlantic bluefin tuna. Report 1 – contract report to support to BFT assessment (GBYP 02/2014) of the Atlantic-wide research programme on bluefin tuna (ICCAT-GBYP – Phase 4).

Carruthers, T., Fonteneau, A., Hallier, J.P. 2014. Estimating tag reporting rates for tropical tuna fleets ofthe Indian Ocean. Fish. Res. 155 (2014) 20–32

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

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.

Kolody, D., Hillary, R., Preece, A., 2014. Computational Framework to Support Indian Ocean Bigeye and Yellowfin Management Strategy Evaluation: A review of software requirements and options. IOTC-2014-WPM05-07.

Kolody, D., S. Hoyle. 2013. Evaluation of tag mixing assumptions for skipjack, yellowfin and bigeye tuna stock assessments in the western Pacific and Indian Oceans. WCPFC Working Paper: WCPFC-SC9-2013/SA-IP-11

Langley, A., 2015. Stock assessment of yellowfin tuna in the Indian Ocean using stock synthesis. IOTC-2015-WPTT17-XX.

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.

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

Table 1. Model ensemble 1 –conditioned YFT OMs derived from the preliminary Langley (2015) assessment base case.

Feature	Values (ref case in **bold**)	Comment
h - Beverton-Holt steepness	0.7, **0.8**, 0.9	
M - age-specific vector scalar	0.6, 0.8, **1.0**	When estimated, assessment prefers lower values of M
λ_{tags} – tag weight in likelihood	0.0, 0.1, **1.0**	Lower weighting recognizes that tag assumptions are not entirely valid

Table 2. Model ensemble 2 – conditioned YFT OMs derived from the preliminary Langley (2015) assessment base case, including additional migration rate assumptions.

Feature	Values (ref case in **bold**)	Comment
h - Beverton-Holt steepness	0.7, **0.8**, 0.9	
M - age-specific vector scalar	0.6, 0.8, **1.0**	When estimated, assessment prefers lower values of M
λ_{tags} – tag weight in likelihood	0.0, 0.1, **1.0**	Lower weighting recognizes that tag assumptions are not entirely valid
4 Area, Single population spatial connectivity	**R2P1Es - movement estimated** R2P1Hi - very high mixing R2P1Lo very low mixing	Admits that movement estimates are indirect and possibly not well informed (environmental variables removed from Hi and Lo options)

Table 3. Model ensemble 3 – as model ensemble 1, except derived from Langley (2015) sensitivity trial without environmental movement drivers.

Feature	Values (ref case in **bold**)	Comment
h - Beverton-Holt steepness	0.7, **0.8**, 0.9	
M - age-specific vector scalar	0.6, 0.8, **1.0**	When estimated, assessment prefers lower values of M
λ_{tags} – tag weight in likelihood	0.0, 0.1, **1.0**	Tag assumptions are not entirely valid
Single population spatial connectivity between West and East areas	R2P1Es-movement estimated R2P1Hi-very high mixing R2P1Lo-very low mixing	ref case environmental drivers not used

Table 4. Model ensemble 4 – 2 area grid, excludes tags and environmental movement drivers.

Feature	Values (ref case in **bold**)	Comment
h - Beverton-Holt steepness	0.7, **0.8**, 0.9	
M - age-specific vector scalar	0.8, **1.0**	When estimated, assessment prefers lower values of M
λ_{tags} – tag weight in likelihood	0.0	Tags were not reconfigured for revised spatial assumptions
Single population spatial connectivity between West and East areas	R2P1Hi - very high mixing R2P1Lo- very low mixing	environmental drivers not used

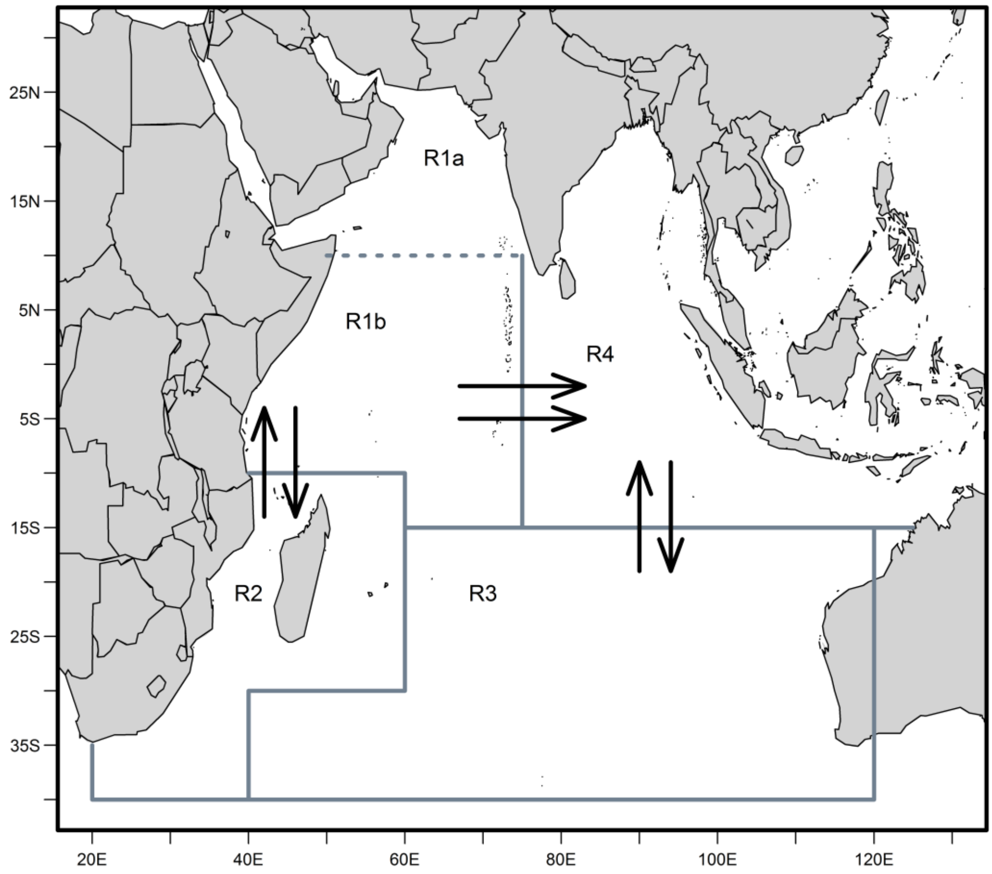

Figure 1. Four area YFT assessment spatial structure from Langley (2015). Two area model discussed in text pools R1a + R1b + R2 in the West and R3 + R4 in the East.

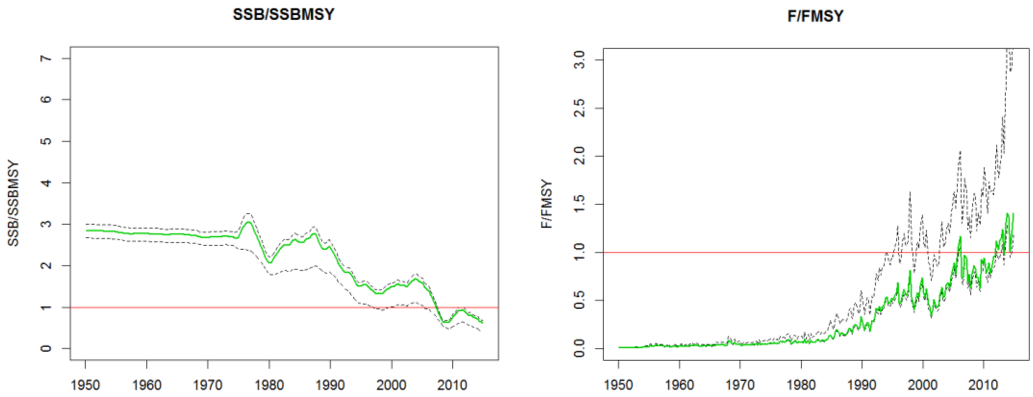

Figure 2. Stock status of the YFT Operating Model ensemble 1 (MPD estimates from 27 models) from Table 1 (4 areas, derived from the Langley (2015) base case). Black lines indicate the median and range of the results; green line indicates the (preliminary) reference case from Langley (2015).

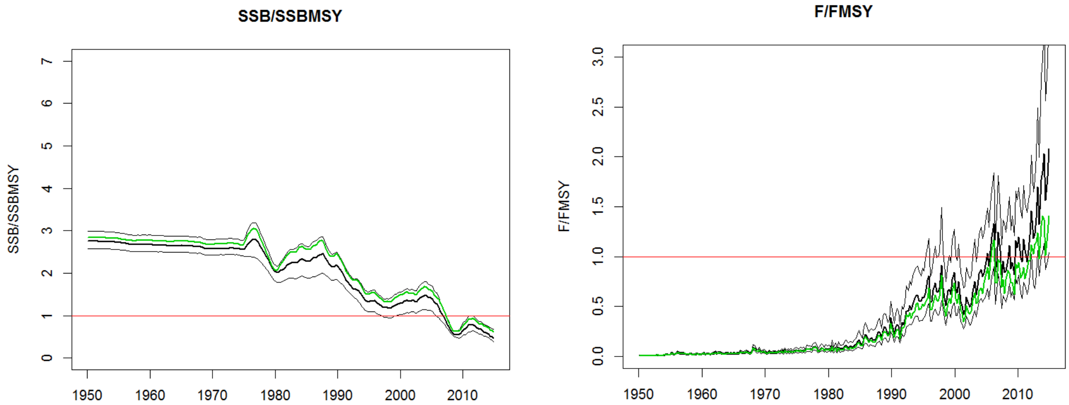

Figure 3. Stock status of the YFT Operating Model ensemble 2 (MPD estimates from 81 models) from Table 2 (4 areas, derived from the Langley (2015) base case). Black lines indicate the median and range of the results; green line indicates the (preliminary) reference case from Langley (2015).

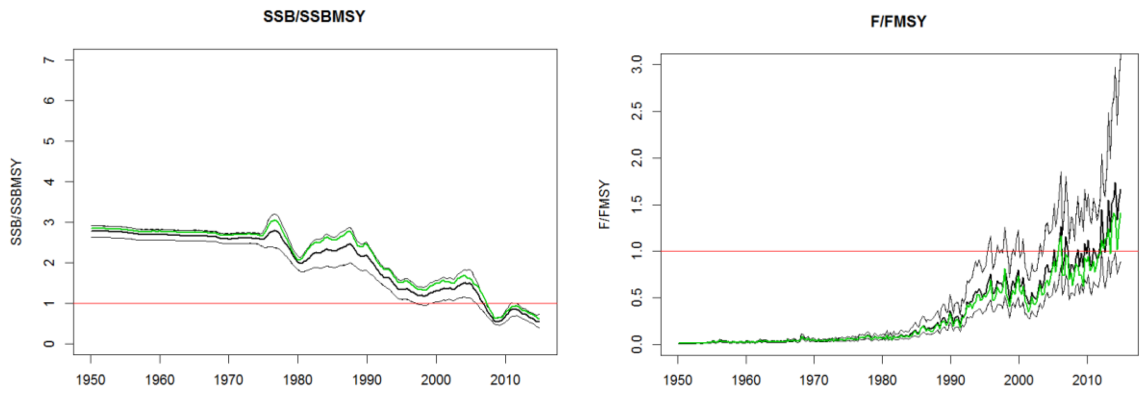

Figure 4. Stock status of the YFT Operating Model ensemble 3 (MPD estimates from 27 models) from Table 3 (4 areas, derived from Langley (2015) sensitivity trial with no environmental movement). Black lines indicate the median and range of the results; green line indicates the (preliminary) reference case from Langley (2015).

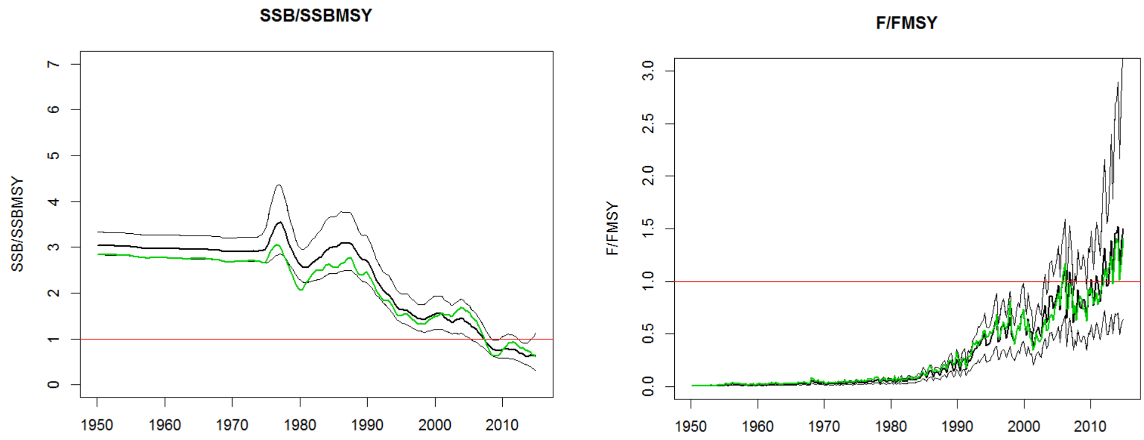

Figure 5. Key features of the YFT Operating Model ensemble 4 (MPD estimates from 27 models) from Table 4 (2 areas, no tags or environmental movement). Black lines indicate the median and range of the results; green line indicates the (preliminary) reference case from Langley (2015).

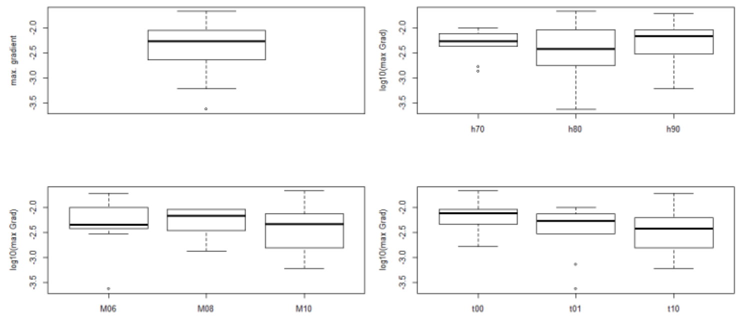

Figure 6. Convergence indicator (maximum gradient in the objective function) diagnostics for model ensemble 3.

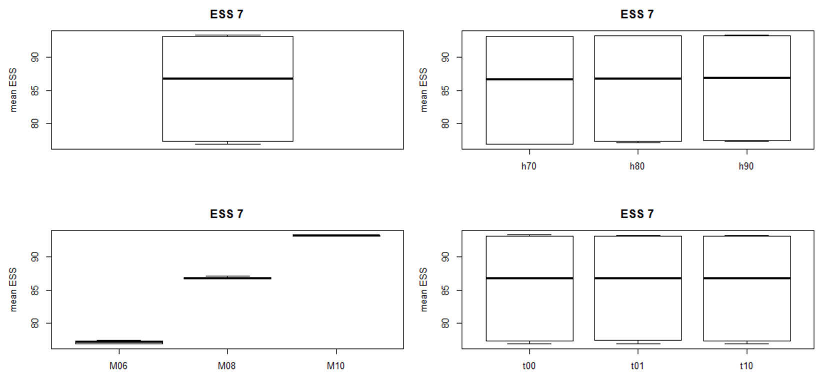

Figure 7. Quality of fit indicator for model ensemble 3, comparing predicted and observed CL fits for fishery 7 (in this case indicating that higher M has a better fit than lower M, while steepness and tag weight have little influence.

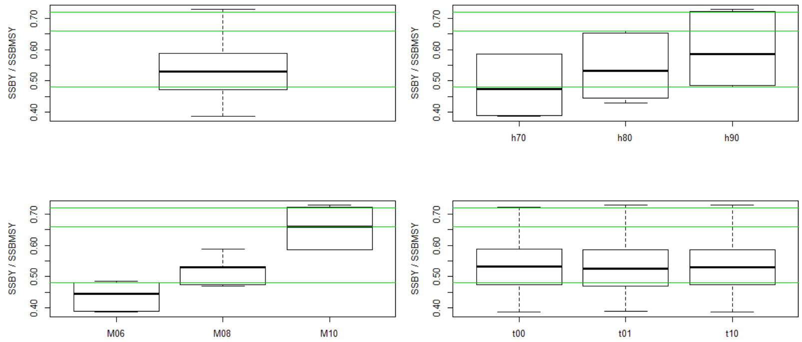

Figure 8. Summary of SSB(2014)/SSBMSY stock status estimate for model ensemble 3, indicating that M and h have a strong influence (while the tags do not). Green lines indicate the median and range described from the preliminary Langley (2015) assessment.

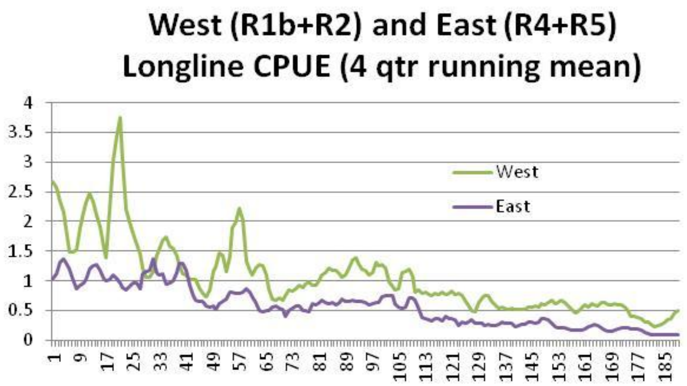

Figure 9. Comparison of (area-weighted) standardized Japanese LL CPUE from the Western and Eastern Indian Ocean.

Appendix 1. Do we need to use the Baranov Catch Equations?

The Baranov catch equations provide a description of fishery population dynamics assuming that fishing and natural mortality rates are continuous (within the timestep of a difference model):

$N_{t+1,a+1} = N_{t,a}\exp(-(F_{t,a} + M_{t,a}))$,

where: N (numbers), F (fishing mortality) and M (natural mortality) are subscripted by t (time) and a (age) (subscripts are mostly implicit below), and catch (C) is given by:

$C = F/(F+M)N(1-\exp(-(F+M)))$.

There is no analytical solution that allows one to solve the Baranov equations for a specified level of catch, as required for projecting a quota extraction (i.e. if one knows C and M, an iterative method must be used to solve for N_{t+1} and F, to an arbitrary level of precision). Iterative solutions are computationally time consuming and not practical to implement for high volume MSE simulations when using a high level interpretive language such as R, though efficient low level subroutines can greatly reduce the speed burden. In the interest of speed and simplicity for code distribution, we have initially used the simpler approach of adopting an approximation of the Baranov equations, with the expectation that all of the MSE code can be maintained in R, and a more efficient low level sub-routine can be called from R at a later date. Here we consider the implications of different approaches that avoid the iterative solution. A common approach is to assume that catch extractions occur at a single instant, at some point in the year, while M is continuous before and after the fishery:

$N(preTAC,a) = N(y,a)\exp(-T(M))$

$N(postTAC,a) = N(preTAC,a)-C(TAC)$

$N(y+1,a+1) = N(postTAC,a)\exp(-(1-T)(M))$,

where T = timing of the instantaneous TAC fishery and catches from the quota-managed fisheries are represented by C(TAC). Pope's approximation assumes that the fishery occurs midway through the year (T=0.5). The Carruthers et al. (2014) source code developed for Atlantic bluefin, assumed that all catch occurred at the beginning of the timestep (T=0), before any natural mortality. Since it remains unclear how the IOTC plans to implement management actions, the BET/YFT framework needs to be able to represent a mix of quota- and effort-regulated fisheries. We are not sure if anyone has gone down this route before, but an obvious option for this extension can be achieved with the approximation:

$N(preTAC,a) = N(y,a)\exp(-T(M + F(TAE)))$

$N(postTAC,a) = N(preTAC,a)-C(TAC)$

$N(y+1,a+1) = N(postTAC,a)\exp(-(1-T)(M + F(TAE)))$

$C(TAE) = TF(TAE)/(TF(TAE)+TM)*N(y,a)(1-\exp(-T(F(TAE+M)))) +$
$TF(TAE)/(TF(TAE)+TM)*N(postTAC,a)(1-\exp(-(1-T)(F(TAE+M))))$

where fishing mortality and catch from the effort-managed fisheries are represented by F(TAE) and C(TAE). The same initial numbers, catch quota and effort will result in different N(t+1), F(TAC) and C(TAE) than the Baranov equations. Pope's approach is usually described as an approximation to the Baranov equations, however, this approach may actually be a closer representation to reality for many fisheries, depending on the specific timing of the fishery (i.e. many fisheries do not have uniform effort distribution within a year (or season). Furthermore, when coupled with a spatial model that assumes instantaneous movement (the most common assessment approach), the Baranov approach will suffer from low resolution time-step inconsistencies anyway. Rather than being concerned about whether or not Pope's approximation is 'correct' in some sense, the bigger concern is that i) it may not be consistent with the conditioning model (i.e most stock assessments use the Baranov equations), and ii) the equations above represent a mix of effort- and quota-managed fisheries that will not yield a consistent relationship between fishing mortality and catch. For the results presented below, we assumed catch at the beginning of the timestep, rather than mid-timestep, as this results in a simpler error structure (not shown - Pope's approximation is most similar to Baranov with M~0.2, which perhaps not coincidentally reflects the historical needs of the North Atlantic).

The question is how big is the difference between Baranov and the approximation for the current application? If the difference is small, the inconsistency is not important. Fig. A1 illustrates contours of the deviations between the two (defined as 100% X (Approx-Baranov)/Baranov) for:

1) $N_{t+1,}$
2) F calculated for a TAC-based fishery ,
3) Catch calculated for a TAE-based fishery,

presented for a broad range of TAC (expressed as harvest rate (C/N), F (for effort managed fisheries) and M. The deviations are clearly large for much of this space. Contours of N_{t+1} are also included in Fig. A1 as an indicator that the ridiculously high deviations are associated with very low survival rates that would probably not be relevant in the context of any sensible harvest strategy. These inconsistencies may cause non-trivial concerns in an MSE evaluation. Options for reducing the inconsistency include:

1) Increase the temporal resolution. Note that BET and YFT simulations will probably be run quarterly (to capture the seasonality described in assessments), while F and M values are usually reported as annualized rates (i.e. the deviations corresponding to an annual M=0.8 would correspond to M=0.2 in Fig. A1). On the basis of these contours, it is not obvious that increasing the temporal resolution by a factor of 2-4 would be sufficient to alleviate the problem.
2) Apply a deviation correction to get the approximate Baranov (~B) solutions, e.g. $N_{\sim B,\, t+1} = N_{Approx,\, t+1}$ + b(H,F,M), where b() is a function of the TAC harvest rate, F(TAE) and M (alternatives to Pope's approximation have been described in the literature and may use similar approaches).
3) Implement an efficient low level sub-routine for the projections that uses the Baranov equations. FLASHER from the Fisheries Library in R is one possibility that is currently being extended to handle multi-population, multi-season, multi-area projections, with

simultaneous TAC and TAE management control options. Older versions of FLASHER could be implemented on each season-area strata independently.

As a quick exploration of option 2, we used a simple least squares approach to fit a third order polynomial (21 terms) to each set of deviations, i.e.

deviation = b(H, F, M) = b(H, F, M) = $a_1 + a_2H + a_3F + a_4M + a_5H^2 \ldots a_x(H^2F) \ldots a_y(HFM)$

Fitting was limited to observations where Z (combined mortality from all sources) was less than 1.6 (survival>20%).

Fig. A2 illustrates the catch equation deviations with and without the added correction terms. Fig. A2 includes a balanced grid of points (on a linear scale) with H(0.001 – 0.8), F(0.001 – 1.0) and M(0.025 – 0.5). With the correction terms, deviations can be kept below 2% for N (Z<2) and F and C (Z<4).

Since we know each of H, F and M, its trivial to calculate the deviation correction term, and avoids searching for iterative solutions. At this time, we expect that Pope's approximation with a deviation adjustment provides a computationally efficient and numerically sufficient approach for the YFT/BET simulations. However, we recognize that an independently coded option using an efficient low level sub-routine should be pursued in parallel.

Finally we note that any quota-based simulation may be faced with an impossible task if the quota exceeds the vulnerable population, or if the quota can only be obatined with an unrealistic level of effort. For the simulations, we have applied an arbitrary hard limit to the catch (if C>0.8N then C=0.8N), such that the catch quota is not met. This is not a realistic solution, but any sensible MP should probably not be required to operate under these circumstances unless something catastrophic has happened.

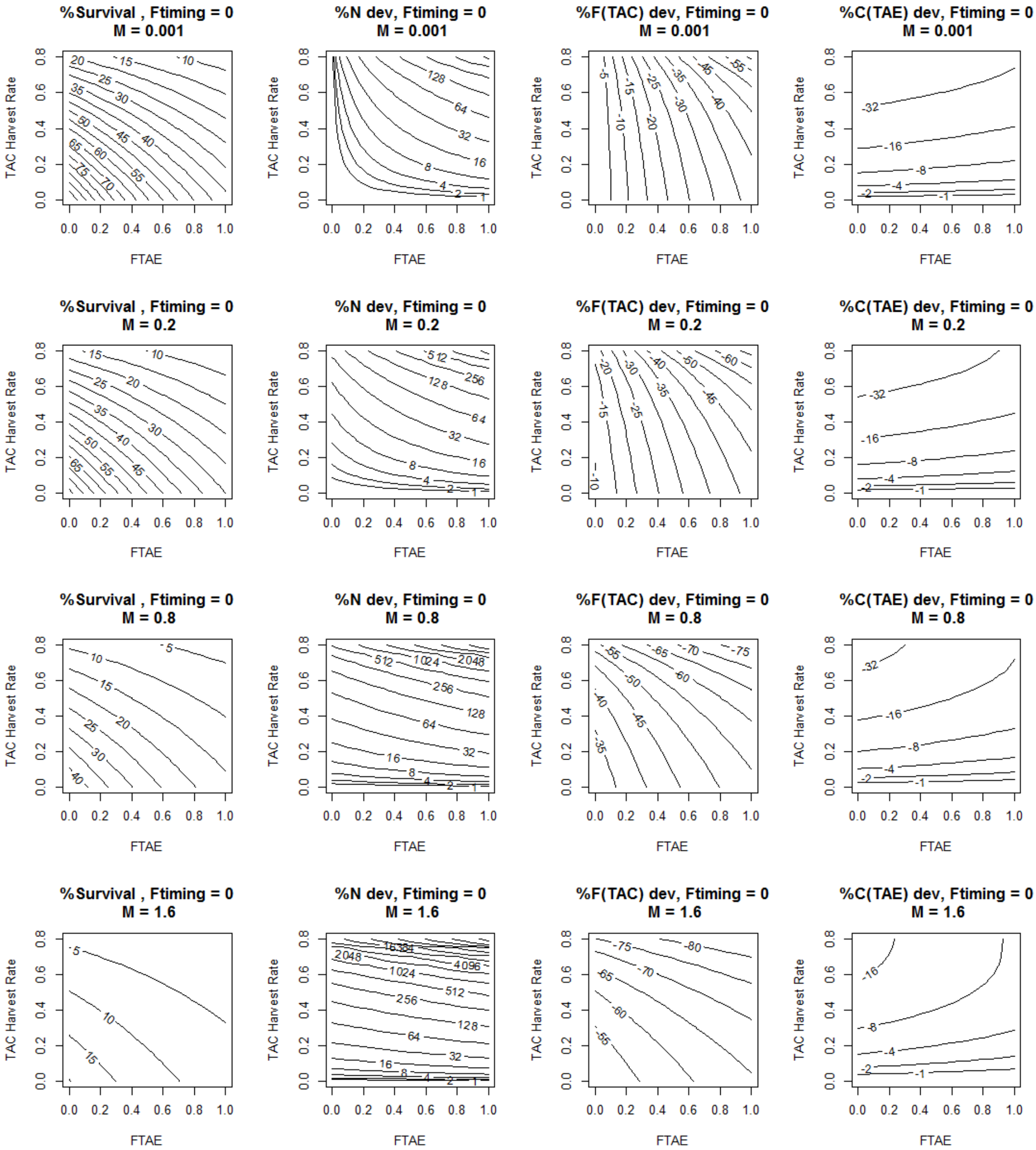

Fig. A1. Deviations in N(t+1), F for the TAC-managed fishery, and Catch of the TAE managed fishery for a grid of harvest rate X FTAE and 4 levels of M. (Note irregular contour levels)

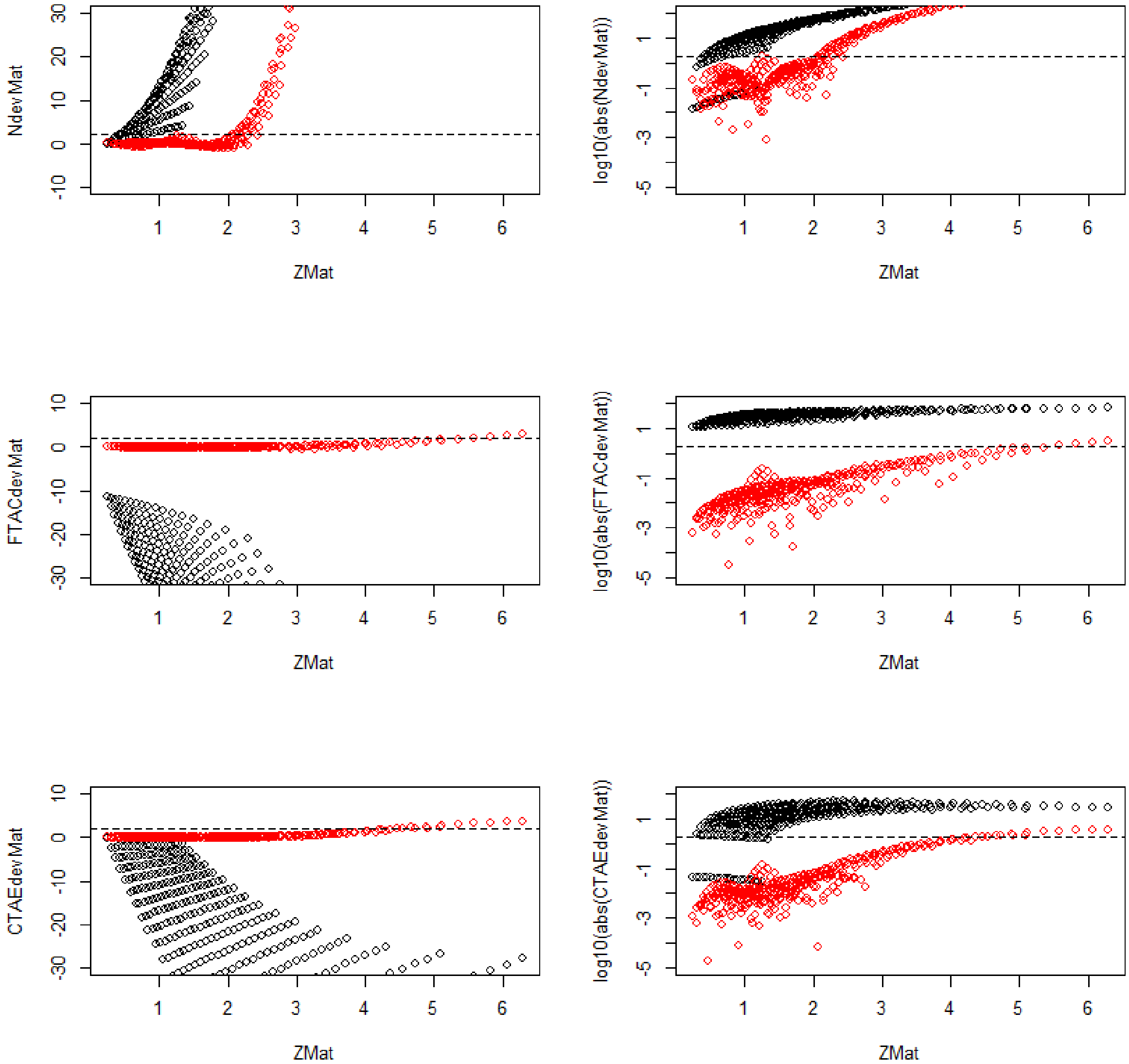

Fig. A2. Comparison of the deviations between Pope's approximation (beginning year fishery timing) and Baranov equations (100X(Pope-Baranov)/Baranov) for N_{t+1} (top), F derived from TACs (middle) and Catch derived from TAEs (bottom) for a range of TAC, TAE and M combinations (indicated by the aggregate Baranov Z on the X-axis). Black points represent the calculated deviation (many points off-scale), red points represent the calculated deviation with a "Baranov correction term" added (estimated by fitting a 3^{rd} degree polynomial surface as f(TAC Harvest Rate, F_{TAE}, M)). Right panels are log-scale. Broken lines represents 2% deviation.