OCEANS AND ATMOSPHERE
www.csiro.au

# Update on Yellowfin Tuna Management Procedure Evaluation Oct 2017

Dale Kolody and Paavo Jumppanen

Prepared for the Indian Ocean Tuna Commission Working Party on Methods and Working Party on Tropical Tunas, Seychelles, Oct 2017

**Acknowledgments**

This work was jointly funded by the Common Oceans - Areas Beyond National Jurisdiction Tuna Project (administered by FAO), and Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO - Oceans & Atmosphere).

# Contents

# 1 Summary

This paper summarizes progress on the development of Operating Models (OMs) and evaluation of candidate Management Procedures (MPs) for IOTC yellowfin (YFT) tuna. The phase 1 project finished in Jun 2016, and phase 2 commenced Sep 2017. During the intervening period, various IOTC technical groups provided development requests for the next iteration of the process, including i) refined definitions for yellowfin and bigeye tuna reference set and robustness set Operating Models (OMs), ii) new candidate MP definitions, and iii) MP tuning objectives.

Initial progress on phase 2 has focused on the yellowfin tuna OMs. The revised reference set OM (referred to in aggregate as OM-ref) is composed of an ensemble of 216 stock assessment model configurations, conditioned in relation to the 2016 stock assessment, and representing uncertainty in 6 dimensions in an equally-weighted design:

- 3 X Beverton-Holt stock recruit relationship steepness
- 3 X Natural mortality vectors
- 3 X tag likelihood weighting
- 2 X tag mixing period
- 2 X CPUE standardization method
- 2 X CPUE catchability trend

The revised OM-ref is more optimistic than the phase 1 demonstration case, in part reflecting the improved perception of stock status in the 2016 assessment. However, the central tendency of OM-ref tends to be considerably more optimistic than the 2016 assessment. Most of the model assumptions influence the dynamics in the expected fashion (e.g. lower stock-recruit steepness, lower M and increasing CPUE catchability trend are all generally associated with more pessimistic current stock status). The difference in the quality of fit to CPUE and size composition data does not vary much among the OM-ref models. These data conflict with the tagging data, such that the tag-weighting dimension (tag $\lambda$ = 1.0, 0.1, 0.0) is very influential and a high priority for further consideration. Models with down-weighted tagging data are generally more optimistic in terms of stock status and productivity, with MSY estimated to be greater than double the base case assessment level in 21% of specifications (with an implausible extreme 10 times higher than the base case). The higher productivity scenarios tend to explain the declining CPUE trend as a result of declining trend in recruitment deviations. There are recognized compatibility problems between the tags and the model structure (notably low tag mixing rates), such that full weighting of the tags is questionable, but high tag weighting represents a pragmatic means for obtaining a suite of models that is subjectively consistent with expectations of stationary production dynamics and the perception that the stock has been fully exploited in recent years.

The following OM robustness scenarios were explored:

- Two attempts were made to formulate OM robustness scenarios that admit a potential tendency for longline fisheries to shift toward targeting younger individuals over time: i)

estimating selectivity in 10 year blocks, and ii) estimating changes in selectivity as a monotonic function of time. Neither option resulted in a management situation that was substantially different from the OM-ref stationary selectivity assumption, and hence may not meet the expectations for robustness trials.

- Up-weighting the tagging data (tag λ = 1.5), results in similar, but slightly more pessimistic OM than the 2016 assessment tag weighting assumption (λ = 1.0). It is not clear that the λ = 1.5 robustness scenario adds a fundamentally different challenge for the MP than the λ = 1.0 option. However, it does emphasize the importance of the tag-weighting assumptions in the current model framework, and the need to ensure that MP performance against pessimistic scenarios is explicitly considered (whether in reference or robustness scenarios).

The TCMP identified two initial MP tuning criteria for YFT:

  - Pr(mean(B(2019:2039))/BMSY = 1.0) = 0.5
  - Pr(mean(B(2024))/BMSY = 1.0) = 0.5

Brief testing of candidate MPs suggest that the generally high productivity of OM-ref might result in counter-intuitive performance.  These results are presented for feedback and/or endorsement by the WPTT and WPM, noting that the Commission MSE workplan expects MSE results to be presented to the TCMP for consideration in 2018.

# 2 Introduction

## 2.1 Background

The Indian Ocean Tuna Commission has committed to a path of using Management Strategy Evaluation (MSE) to meet its obligations for adopting the precautionary approach. IOTC Resolution 12/01 *"On the implementation of the precautionary approach"* identifies the need for fishery reference points and harvest strategies that will help to maintain the stock status at a level that is consistent with the reference points. Resolution 13/10 "*On interim target and limit reference points and a decision framework*" identified interim reference points and elaborated on the need to formulate management measures relative to the reference points, using MSE to evaluate harvest strategies in recognition of the various sources of uncertainty in the system. Resolution 15/10 supersedes 13/10 with a renewed mandate for the Scientific Committee to evaluate the performance of harvest control rules with respect to the species-specific interim target and limit reference points, no later than 10 years following the adoption of the reference points, for consideration of the Commission and their eventual adoption. A species-specific workplan was re-affirmed at the 2017 Commission Meeting, outlining the steps required to adopt simulation-tested Management Procedures for the highest priority species (included in Attachment 1). Recognizing the iterative nature of the MSE process, the workplan identifies 2019 as the earliest probable date for MP adoption.

## 2.2 Phase 2 yellowfin and bigeye project

MSE for bigeye and yellowfin tunas has been pursued in parallel, with the first phase of the scientific and technical work described in Kolody and Jumppanen (2016). A second phase project has been established to support progress from Sep2017 to Dec2018. This second phase project is responsible for reporting progress to the IOTC subsidiary bodies (including the TCMP, WPM and WPTT), and implementing feedback to support the technical and scientific needs of the IOTC community. This working paper represents the first reporting of the phase 2 project. Given the recent commencement of this project, progress has been limited to:

i) A "mechanical" update to the YFT reference case OM in line with the feedback from the 2016 IOTC technical working parties, and presentation of common diagnostics for evaluating plausibility. The reference case is intended to encompass the main assessment uncertainties, and provides the main descriptor of expected MP performance, subject to tuning.

ii) Exploration of potential robustness case OMs in line with the feedback from the 2016 IOTC technical working parties. Robustness cases generally include less likely, but potentially troublesome dynamics, and may be used to identify MPs that are more robust to difficult situations.

iii) Presentation of some candidate MP results that meet the initial tuning objectives identified in TCMP (2017). Tuning is the procedure used to attain precise management performance with respect to a single high priority management objective. When multiple MPs are tuned to the same criterion, it is easier to choose among MPs on the basis of secondary or tertiary objectives. Tuning objectives should aim to identify a target location in the main trade-off between conservation risk and economic opportunity, as this is usually the strongest driver of MP performance.

Similar work for bigeye is expected to begin soon after the 2017 WPTT and WPM.

## 2.3 Relationship between the stock assessment and Operating Model

As detailed in Kolody and Jumppanen (2016), the intention has been to maintain a close relationship between the stock assessment modelling and the conditioning of OMs. The two processes are analogous in several respects, i.e. similar population dynamics models are applied to the same data, subject to the same concerns about model formulation and assumption violations, etc. Accordingly, the yellowfin assessment of Langley (2016) provides the core of the OM conditioning process. Key features of the assessment and OM include:

- Implementation with Stock Synthesis 3.24z software
- 4 regions (Figure 1)
- Quarterly dynamics, including recruitment and movement
- 25 fisheries
- Parameter estimation objective function includes
    - Total catch
    - Standardized longline CPUE (one series per region)
    - Size composition data
    - Tags (excluded in some OM scenarios)
    - Recruitment penalties on deviations from stock recruit relationship and mean spatial distribution
- Estimated parameters:
    - Fishery selectivity (shared among some fleets)
    - Longline catchability
    - Virgin recruitment
    - Recruitment deviations from the Beverton-Holt stock-recruit relationship, recruitment spatial partitioning among tropical regions (1,4) and deviations from the mean spatial distribution.
    - Juvenile and adult movement rates

OM conditioning has an increased emphasis on uncertainty quantification and stochastic projections required to develop robust feedback-based MPs through the MSE process. The reference set OM is an ensemble of assessment models that includes several alternative plausible assumptions. The approach to uncertainty quantification adopted here is similar to that used in the CCSBT, in which the emphasis is on model structural uncertainty (including parameters about which the data are expected to be uninformative), and stochastic recruitment uncertainty in the projections. The Maximum Posterior Density Estimates (best point estimates) for the individual models are collated, with the expectation that this source of uncertainty will generally be greater than the parameter estimation uncertainty conditional on any individual model. Once an adequate OM has been defined, it should not need to be updated with the frequency expected for the traditional stock assessment process, unless new evidence emerges to indicate that the uncertainty encompassed by the OM no longer captures reality.

Robustness OMs are generally considered less likely than the reference set, but they are defined to represent plausible, troublesome situations, that may help identify pathological MP behaviour in particular circumstances, and assist in choosing among MPs that are otherwise equivalent.

For the purposes of this paper, we refer to a number of individual models, and OM ensembles as defined in Table 1 and Table 2.

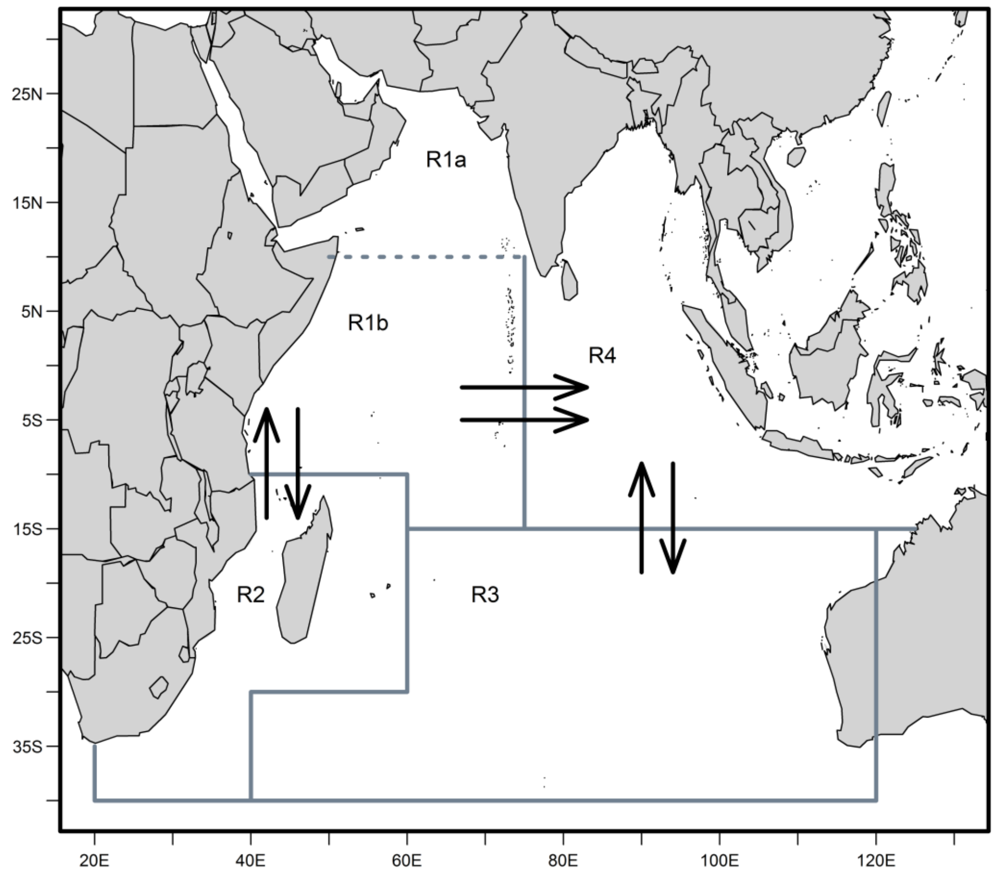



**Figure 1. Spatial structure for yellowfin tuna assessment and all OMs discussed in this report (figure from Langley 2015).**

**Table 1. Model definitions.**

| Model Name | Definition (assumption abbreviations are defined in Table 2) |
|---|---|
| SA-base | The base case assessment from Langley (2016).<br>h80, M10, t10, q0, iH, x3, SS |
| OM-SA-analogue | The single OM specification that most closely resembles SA-base (identical except for no environmental movement link).<br>h80, M10, t10, q0, iH, x3, SS |
| OM-ref | Reference case OM consisting of an ensemble of 216 models, each differing from OM-SA-analogue in 1-6 assumptions. Undefined options as in OM-SA-analogue.<br>h70, h80, h90<br>M10, M08, M06<br>t00, t01, t10<br>q0, q1<br>iH, iC<br>x3, x8 |
| OM-rob-selTrend | A robustness OM consisting of 36 models, designed to look at the implications of temporal variability in selectivity, potentially resulting in a shifted preference toward younger ages. Undefined options as in OM-SA-analogue.<br>M10, M08, M06<br>t01, t10<br>x3, x8<br>SS, NS, ST |
| OM-rob-tagWt | A robustness OM consisting of 36 models, designed to look at the implications of tag-weighing λ options, notably the recommendation of λ = 1.5. Undefined options as in OM-SA-analogue.<br>M10, M08, M06<br>t00, t0001, t001, t01, t10, t15<br>x3, x8 |

**Table 2. Model specification abbreviations. Bold indicates the assessment base case assumption.**

| Abbreviation | Definition |
|---|---|
| | Stock-recruit steepness (*h*) |
| h70 | 0.7 |
| **h80** | **0.8** |
| h90 | 0.9 |
| | Natural mortality multiplier relative to SA-base |
| **M10** | **1.0** |
| M08 | 0.8 |
| M06 | 0.6 |
| | Tag data weighting (tag composition and negative binomial components) |
| t00 | $\lambda = 0$ |
| t0001 | $\lambda = 0.001$ |
| t001 | $\lambda = 0.01$ |
| t01 | $\lambda = 0.1$ |
| **t10** | **$\lambda = 1.0$** |
| t15 | $\lambda = 1.5$ |
| | Assumed CPUE catchability trend (compounded) |
| **q0** | **0% per annum** |
| q1 | 1% per annum |
| | Tropical CPUE standardization method |
| **iH** | **Hooks Between Floats** |
| iC | Cluster analysis |
| | Tag mixing period |
| **x3** | **3 quarters** |
| x8 | 8 quarters |
| | Longline selectivity |
| **SS** | **Stationary** |
| NS | Temporal variability estimated in 10 year blocks |
| ST | Logistic selectivity trend estimated over time |

## 2.4 Management Procedures and MP Tuning

While the emphasis of this report is the revision of yellowfin OMs, results from a small number of candidate MPs are reported as defined in Table 3. See Kolody and Jumppanen (2016) for the full specification of these MPs. This project is aiming for the *sensu stricto* definition of Management Procedures, in which the MP consists of:

i) pre-defined data collection

ii) pre-defined analytical methods (including assessment model specification and data processing)

iii) simulation-tested Harvest Control Rule to specify the management action

Toward this end, the projection component of the OM simulates data that are consistent with the OM conditioning assumptions, and these data are interpreted by the MP to produce the Total Allowable Catch (TAC), subject to "realistic" data and analytical errors.

These MPs were tuned according to the criteria defined in Table 4. For expedience, all tuning for this report was conducted with a minimal set of 216 realizations from OM-ref. As a brief test of the precision associated with tuning to only 216-realizations, 3 tuned MPs were applied to a full set of 2160 realizations. The tuned Pr(green Kobe) = 0.75 was slightly lower than the realized probabilities arising from the full evaluation Pr = 0.76-0.79. This level of tuning precision is considered adequate for the purposes of this report, but the full set of 2160 will be used for the TCMP.

**Table 3. Qualitative definitions of the MPs used in this report.**

| Label | Definition |
|---|---|
| **PT4010** | A catch-based "40:10-type" HCR coupled with a surplus production model. |
| **PT4010F** | An F-based "40:10-type" HCR coupled with a surplus production model. |
| **IC** | A CPUE-based HCR that "aims" for a desirable CPUE target by increasing or decreasing the TAC, depending whether CPUE is above or below the target, and whether it is trending up or down. |
| **CCt** | Constant catch |

**Table 4. MP Tuning objectives applied to yellowfin in this report.**

| Label | Source | Definition |
|---|---|---|
| **T1** | TCMP YFT objective 1 | Pr(mean(SB(2019:2039))/SB(MSY) = 1.0) = 0.5 |
| **T2** | TCMP YFT objective 2 | Pr(mean(SB(2024))/SB(MSY) = 1.0) = 0.5 |
| **T3** | TCMP BET objective 2 | Pr(Green Kobe 2019:2039) = 0.75 |
| **T4** | This report | Pr(mean(SB(2019:2039))/SB(MSY) = 1.5) = 0.5 |

## 2.5 Harvest Control Rules based on catch and fishing mortality

In WPM (2016), it was noted (but perhaps not documented) that one of the MPs tested for yellowfin and bigeye was unusual in that it applied the "40:10-type" Harvest Control Rule to calculate catch, while it is more commonly used to calculate fishing mortality. As shown in Figure 2, the difference in TAC recommendation may be quite different between the two versions. However, the "PT4010-type" MP class was always intended to allow exploration of alternative parameter values. When coupled with a surplus production model, the catch-based MP can easily represent something very similar to the F-based 40:10 MP (i.e. the broken lines in Figure 2), by setting control parameter P2 to a value of 100 and increasing P3.

While catch-based and F-based options are both available for further exploration, we note that the specific form of the MP should be a secondary consideration for the MP selection process. Selection should be focused on the management performance evaluation, which might be counter-intuitive and not consistent with how the HCR "ought" to perform. Comparison of the catch-based and F-based PT4010 MPs (T4 tuning, using parameter P3, evaluated against OM-ref defined below), results in negligible performance differences (Figure 3).

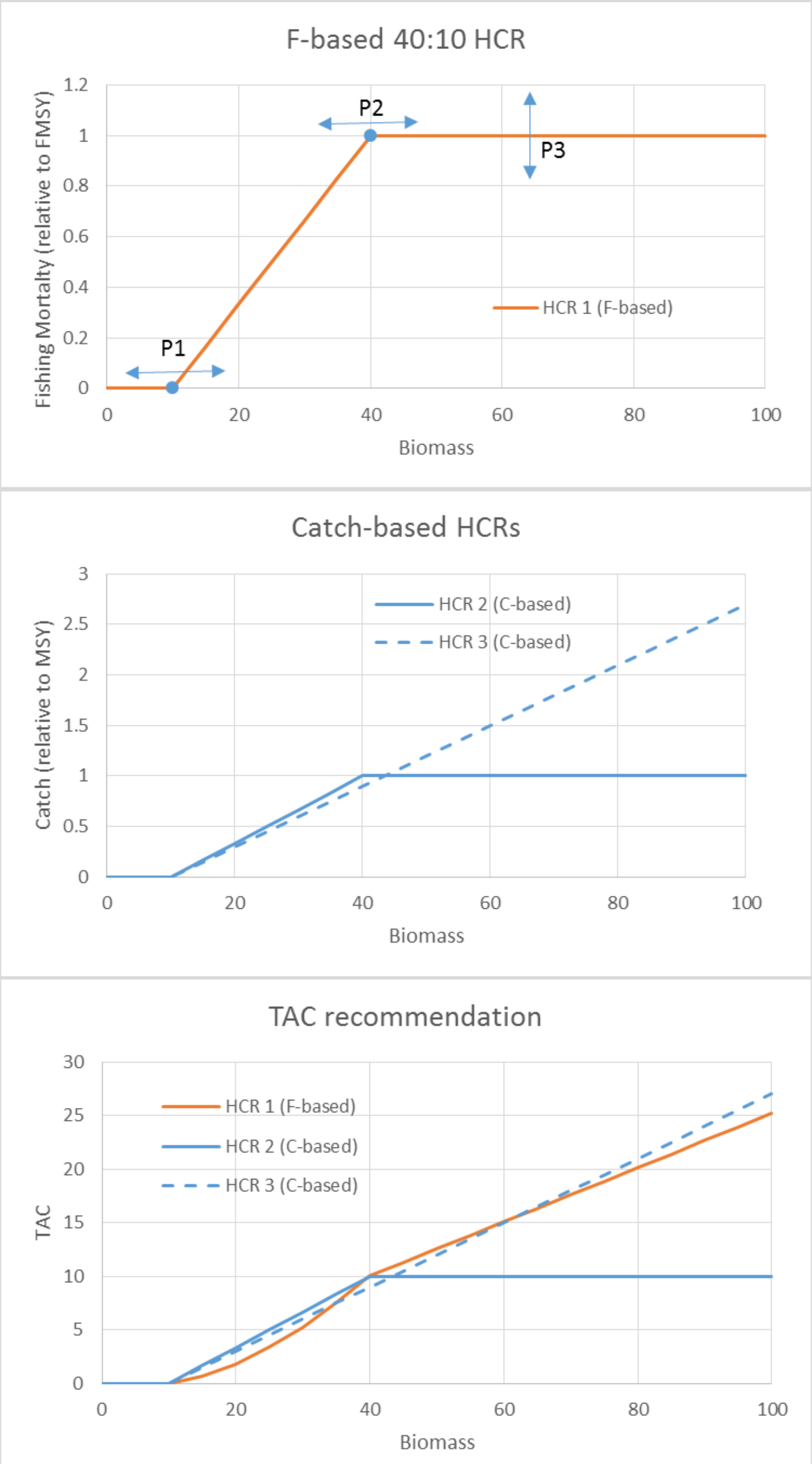



**Figure 2. Comparison of 40:10-type Harvest Control Rules and TAC recommendations using Catch-based and F-based functions (coupled with a surplus production model). The functional form of the TAC recommendation may be very different for the classic functional form (e.g. HCR 1 vs: HCR 2), however, changes to the control parameters may allow very similar outcomes (e.g. HCR 1 vs HCR 3).**

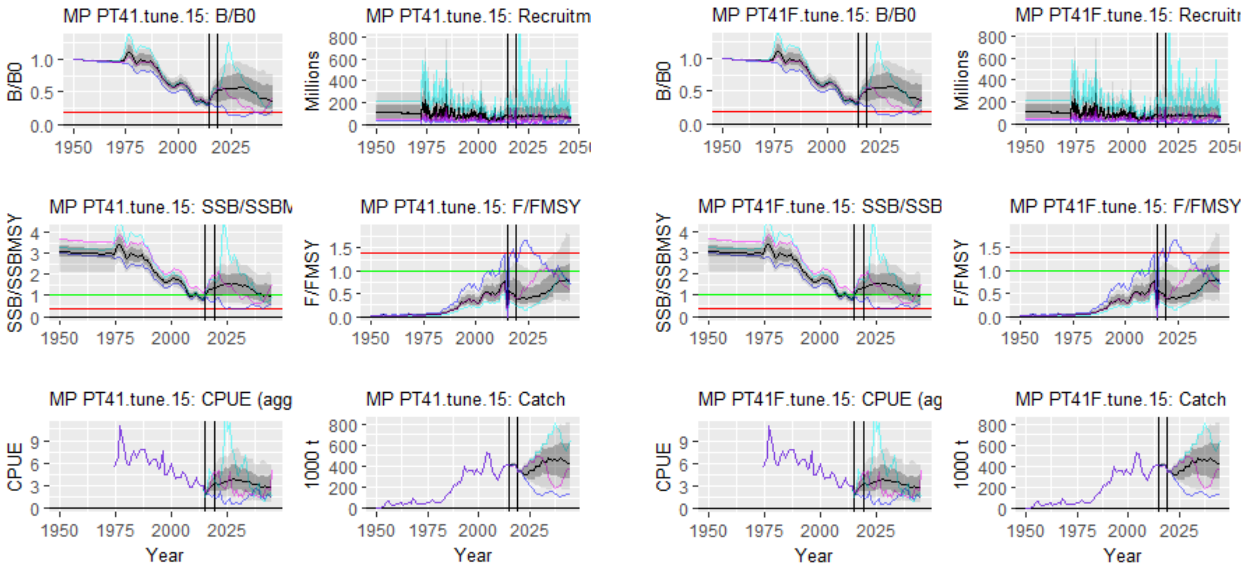



**Figure 3. Comparison of Catch-based (left panels) and F-based (right panels) PT4010 MPs, evaluated against OM-ref, with tuning objective T4, achieved by modifying HCR parameter P3 from Figure 2.**

# 3 Oct 2017 Yellowfin Reference Case Operating Model (OM-ref)

The (Oct 2017) OM-ref incorporates the feedback from the 2016 WPTT and WPM, and represents the first OM definition that was designed with feedback from the IOTC working parties. The key differences from the 2016 demonstration case OM included:

i. An update in relation to the 2016 yellowfin assessment, which most notably included substantially-revised longline CPUE series (plus an additional year of data),

ii. Additional dimensions describing uncertainty in the OM ensemble grid

The OM-ref ensemble is derived from the assessment model that provided the core of the 2016 management advice, detailed in Langley (2015, 2016), and referred to here as SA-base. OM-SA-analogue is the single model specification from OM-ref that is most similar to SA-base, implemented using the same data (with the addition of alternative CPUE scenarios), software (Stock Synthesis 3.24z), and mostly the same assumptions. The other 215 models in OM-ref deviate from OM-SA-analogue as shown in Table 1. As in the original yellowfin OM exploration (Kolody and Jumppanen, 2016), the key difference between OM-ref and SA-base is the movement parameterization discussed in the following section.

## 3.1 Is the environmentally-linked migration in the 2016 assessment required for the yellowfin OM?

SA-ref estimates movement parameters that are linked to quarterly environmental indices. Langley (2015) justified this decision primarily because it improved the fit to the CPUE series. There is clearly seasonal variability in CPUE, which the environmental indices can presumably help describe. Seasonality can be expected to influence the confounded processes of migration and fishery catchability, and within the constraints of the Stock Synthesis temporal configuration (i.e. calendar quarters defined as model years), environmentally-linked movement provides a means to describe these processes.

Environmental variability was not implemented in OM-ref because, i) it adds another layer of complexity in terms of model structure, ii) assuming that SS is designed to accept future environmental co-variates, it could be a non-trivial exercise to forecast environmental variability in the context of MSE projections, iii) it is not clear whether environmental indices are contributing to explaining interannnual variability, or simply seasonality, which is not necessarily important for stock assessment purposes.

Figure 4 shows that including environmentally-linked movement (SA-base) introduces a small seasonal signal to some CPUE predictions, but it is minor relative to the uncertainty assumed in the CPUE series. The fit without environmentally-linked movement (OM-SA-analogue) is very similar in terms of capturing the long term CPUE trends (Figure 5).

Figure 6 shows the similarity between the spawning biomass and fishing mortality time series from SA-base and OM-SA-analogue. Depletion trends in the two series are almost identical, while the recent fishing mortality rate is slightly higher in SA-base. On the basis of these comparisons, we conclude that environmentally-linked migration is probably not a high priority for MP evaluation, and consider removal to be justified for OM-ref.

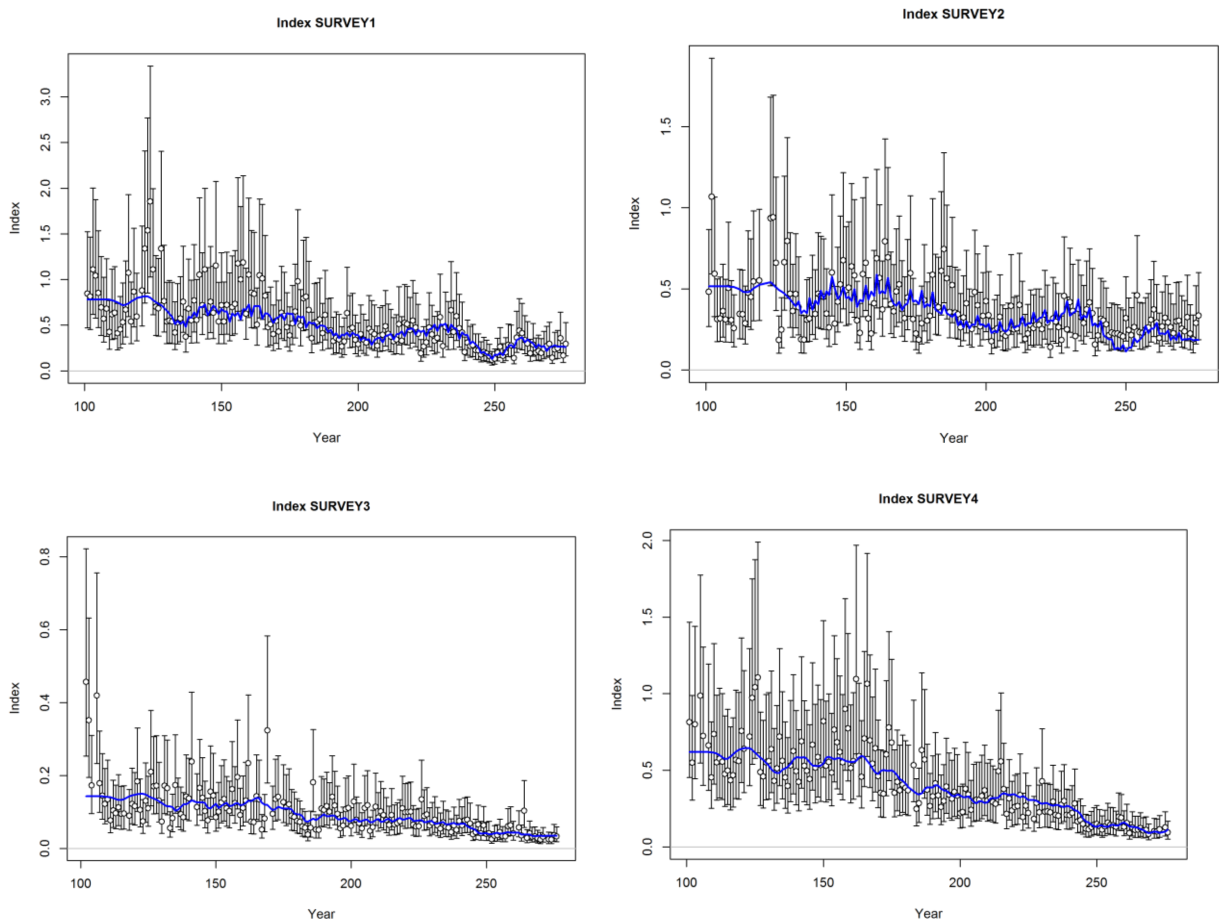



**Figure 4. Comparison of the quality of fit between predicted (lines) and observed (points) CPUE for the 2016 YFT assessment (SA-base) Top panels are the tropical regions, bottom panels are the temperate regions.**

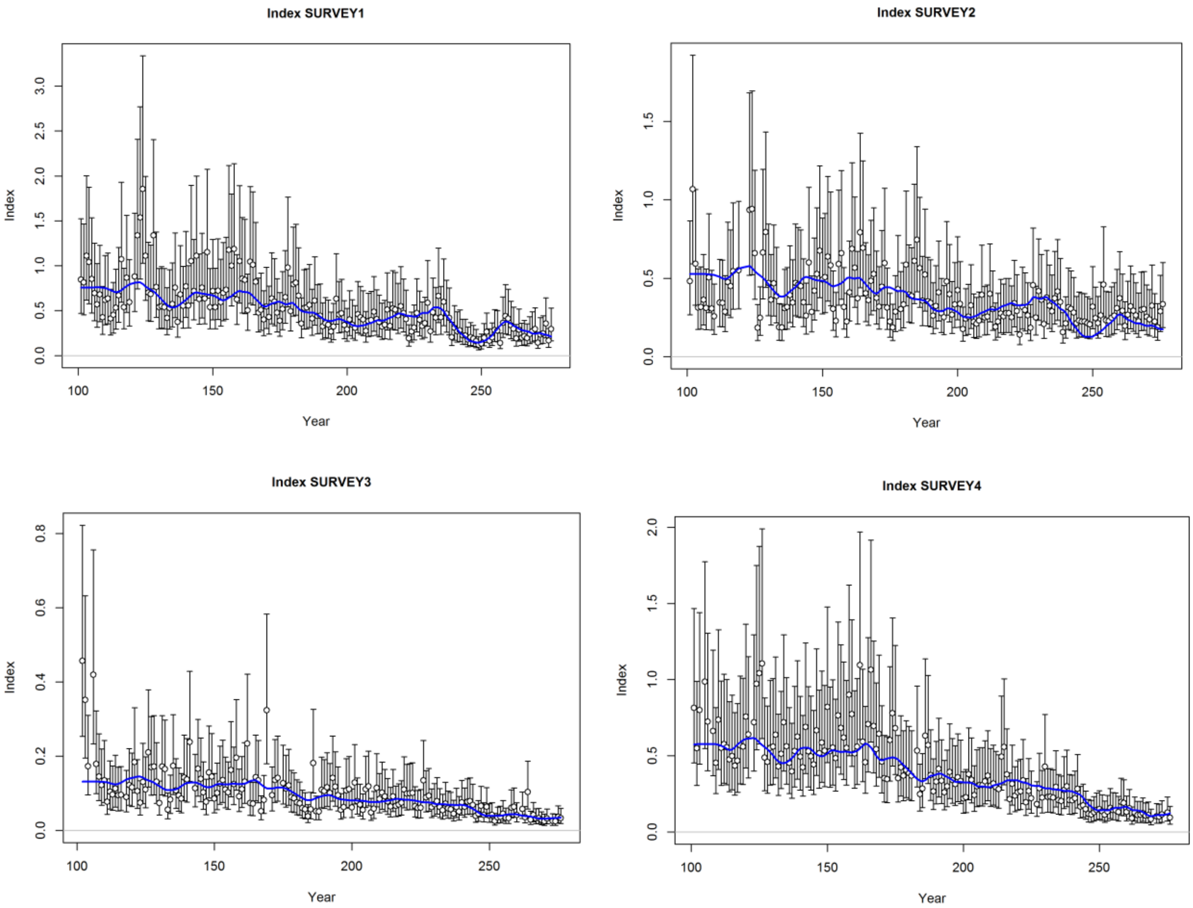



**Figure 5. Comparison of the quality of fit between predicted (lines) and observed (points) CPUE for the OM-SA-analogue (no environmental-linked movement) 2016 YFT assessment (SA-base). Top panels are the tropical regions, bottom panels are the temperate regions.**

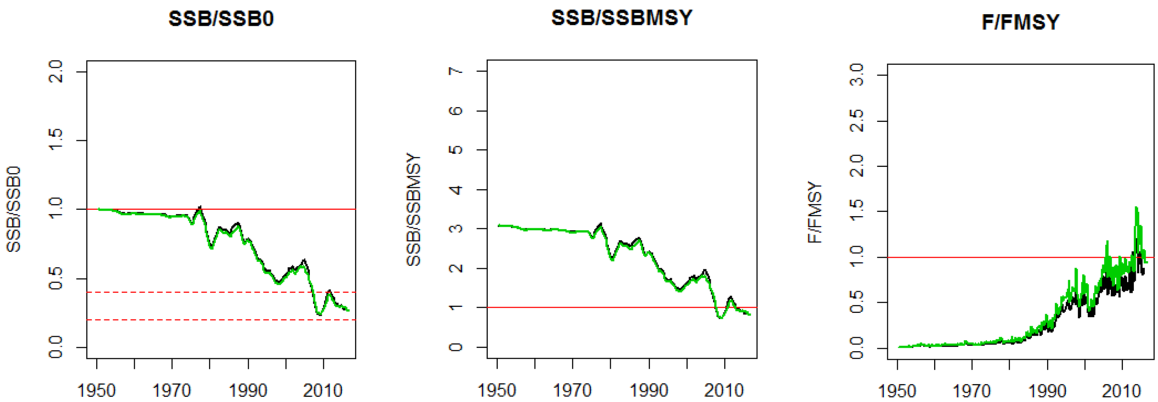



**Figure 6. Comparison of spawning biomass and fishing mortality trends between the 2016 assessment model (SA-ref, green) and the equivalent OM specification without seasonal migration (OM-SA-analogue, black).**

## 3.2 CPUE series

Considerable collaborative work has been undertaken in recent years to improve the understanding of the DWF longline CPUE series, and to provide better relative abundance indices for the stock assessments (Hoyle et al. 2016). The 2016 yellowfin assessment used the latest available studies, but adopted a single (set of area-specific) series as the best available. The WPM encouraged (Attachment 1) the inclusion of an alternative CPUE series in OM-ref, to encompass some of the uncertainty arising from the standardization process. In consultation with IOTC's longline CPUE analysis coordinator (Simon Hoyle, NIWA, New Zealand, pers. comm.), two CPUE series for the tropical regions were selected, with the primary difference being the approach used to account for targeting, either i) Hooks Between Floats (as in SA-base) or ii) cluster analyses on species composition (*std_xTW, Joint_regY_R2_dellog_boat_allyrs, Joint_regY_R5_dellog_boat_allyrs).* The temperate series were not changed from the SA-base assumption, because the species targeting effects were judged to be more important in the temperate zone, such that they really need to be accounted for, and the clustering approach has been judged the best option for achieving this (in the western temperate zone, the eastern temperate zone used HBF for both fleets). The value of the cluster analyses was less clear in the tropical waters, such that including both approaches should better represent the uncertainty.

OM-ref also incorporates catchability trends of 0 or 1% per year (compounded annually, and projected into the future CPUE) on top of the CPUE standardization assumptions, to admit the potential for fishing efficiency improvements related to factors that are not documented in logbooks. The 4 series are shown for the tropical regions in Figure 7. The CPUE decline in the eastern tropical region is not as strong in the eastern region with the clustering analysis.

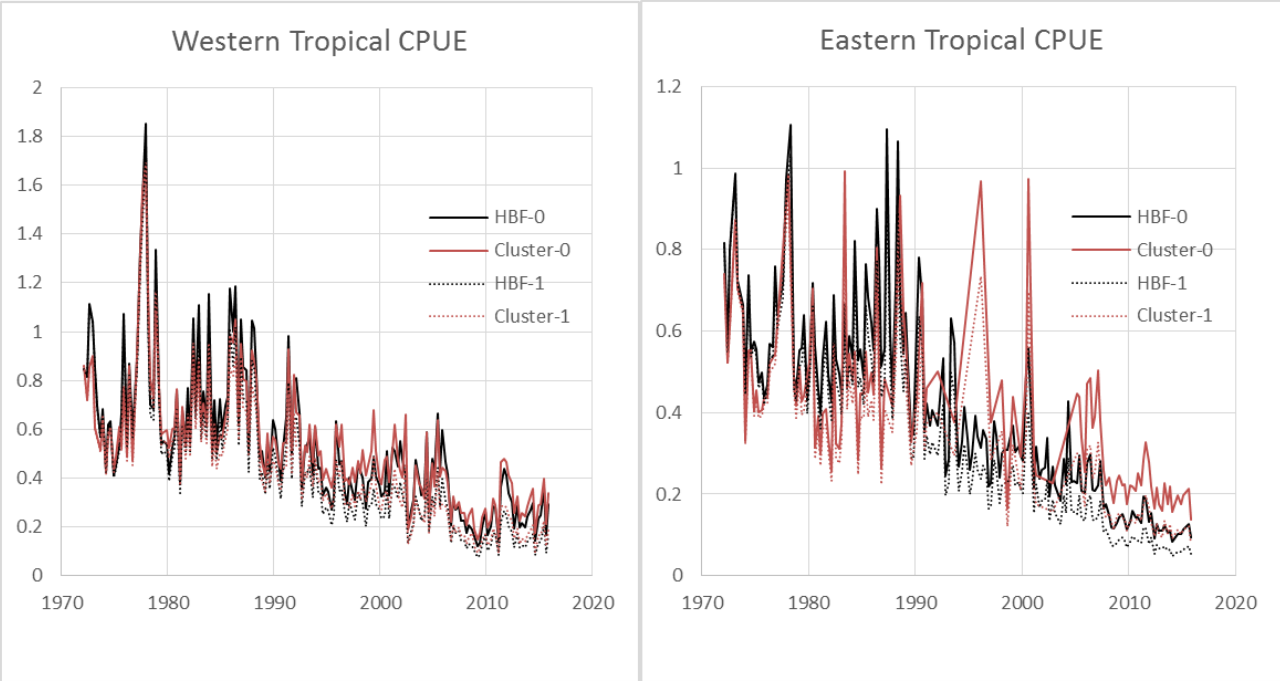



**Figure 7. CPUE series used in the OM grid to represent uncertainty in relative abundance (tropical regions). Legend labels refer to the factors used to account for species targeting shift (HBF or Cluster analyses; iH, iC in the OM grid) combined with the assumed catchability trend (0 or 1% increase compounded annually; q0, q1 in the OM grid).**

## 3.3 Yellowfin OM-ref model fitting summary diagnostics

OM-ref is an ensemble of 216 models, a balanced combination of assumptions which differ in 6 dimensions (Table 1). Each model represents a different assessment model configuration, with a full complement of parameter estimates, model outputs, and potential numerical problems. It is not possible to meaningfully examine all of the model diagnostics in the level of detail that is typically attempted in an assessment. However, it is critical to at least ensure that the individual models in the ensemble are not suffering from pathological convergence failures or describing implausible dynamics. As in Kolody and Jumppanen (2016), we present several summary diagnostics, in a way that allows general trends and outlier behaviour to be rapidly identified. The quality of fit diagnostics reported here are selected to be independent of variance and weighting assumptions (i.e. arbitrarily changing a variance term, assumed sample size or likelihood weighting term can have a large effect on the objective function, even if the model predictions are identical).

The OM-ref models tended to converge reliably, with marginal convergence (i.e. relatively high gradients in the objective function with respect to one or more parameters) evident in only 3 cases (Figure 8, Figure 9). Since these 3 models did not demonstrate any obvious outlier behaviour in terms of quality of fit to the data or stock status, there does not seem to be a pressing need to remove them from the ensemble.

All of the OM-ref models appeared to fit the CPUE series in all regions reasonably well, as indexed by the Root Mean Square Error (RMSE) between predictions and observations (Figure 9, Figure 10). Perfectly consistent agreement would result in RMSE = 0.3 (assumed $\sigma_{CPUE}$ = 0.3), while the observed range was 0.24-0.41. Note that direct comparisons of the RMSE between models that fit to different CPUE series is not very meaningful. The difference in RMSE between the best and worst models was always <0.07. We did not evaluate the auto-correlation in the errors, which may indicate a greater systematic lack of fit in some models than others. Similarly, there is scope to argue about the importance of seasonality in these variance assumptions. However, it appears that concerns about the CPUE fit are less important than the recruitment trend issues discussed below.

Figure 11 illustrates that there is a large degree of variability in the quality of fit to the catch-at-length (CL) data among fisheries (as summarized by the post-fit Effective Sample Sizes - indices that describe the sample size that could be expected to yield the observed quality of fit). However, the degree of variability among models is much smaller. Given the very low assumed CL samples sizes (N=5 for all fisheries), all of these models appear to be consistent with the intent of the assessment, i.e. choose an artificially small sample size that is unlikely to result in an influential conflict with the CPUE (or tag) data.

Figure 12 illustrates the fit to the tagging data (reported as the tag likelihood before the application of tag weighting λ). As would be expected, a higher value of λ results in a better fit to the tagging data (all other things being equal). These results are also partitioned by the tag mixing period (models with different mixing periods are not comparable because they fit to different data). When tag fits are compared across models within a single tag mixing period, it is evident that the tagging data strongly favour the lowest M assumption. This situation was noted in the assessment as well. Since the M06 option is thought to be too low for yellowfin, it may reflect a problem with tag mixing assumptions (SA-base adopts the M10 option).

Figure 13 illustrates that the relative OM-ref stock status inferences are qualitatively predictable from many of the input options (i.e. higher steepness, higher M, no catchability trend, Cluster CPUE analysis are all associated with more optimistic status). Less predictably, lower tag weighting (and longer tag mixing period) are associated with more optimistic stock status. We have no *a priori* expectation that the tags would have a pessimistic influence, so this suggests a conflict between the CPUE series (plus the size composition data presumably to a lesser extent) and the tags.

Figure 13 also shows that the median of the current biomass reference points are similar to SA-base. However there is considerable uncertainty represented, and several models are much more optimistic than SA-base. The contrast is particularly evident with the MSY estimates: 83% of the OM-ref models exceed SA-base; 21% of OM-ref results are more than double SA-base; 5 models estimate MSY >4X SA-base. Figure 14 shows the distribution of biomass and fishing mortality time series for OM-ref, relative to SA-base.

The models with very high MSY are affected by a common issue in stock assessment - one can often explain a declining abundance trend through fishery depletion or declining recruitment, and the two mechanisms may be combined along a continuum (e.g. as also observed in the IOTC albacore MSE OM). If the recruitment trend is the result of a systematic trend in deviations from the stock-recruit relationship, this is indicative of a structural inconsistency in the model. Conversely, given that recruitment trends can arise as a modelling artefact, it is a cautionary note that a stock-recruit relationship might be used to explain the recruitment trend (i.e. such that there are no systematic deviations from the SR relationship), but might still be wrong (though internally consistent). The more optimistic models generally show a greater declining trend in recruitment deviations than the optimistic models (Figure 15, Figure 16). The highest productivity models are responsible for the very low fishing mortality scenarios (Figure 14), in which the fishery would have a trivial influence on stock dynamics.

The extremely high MSY scenarios do not show a lack of fit to the CPUE or size composition data that would identify them as obviously incompatible with these data (Figure 17, Figure 18).

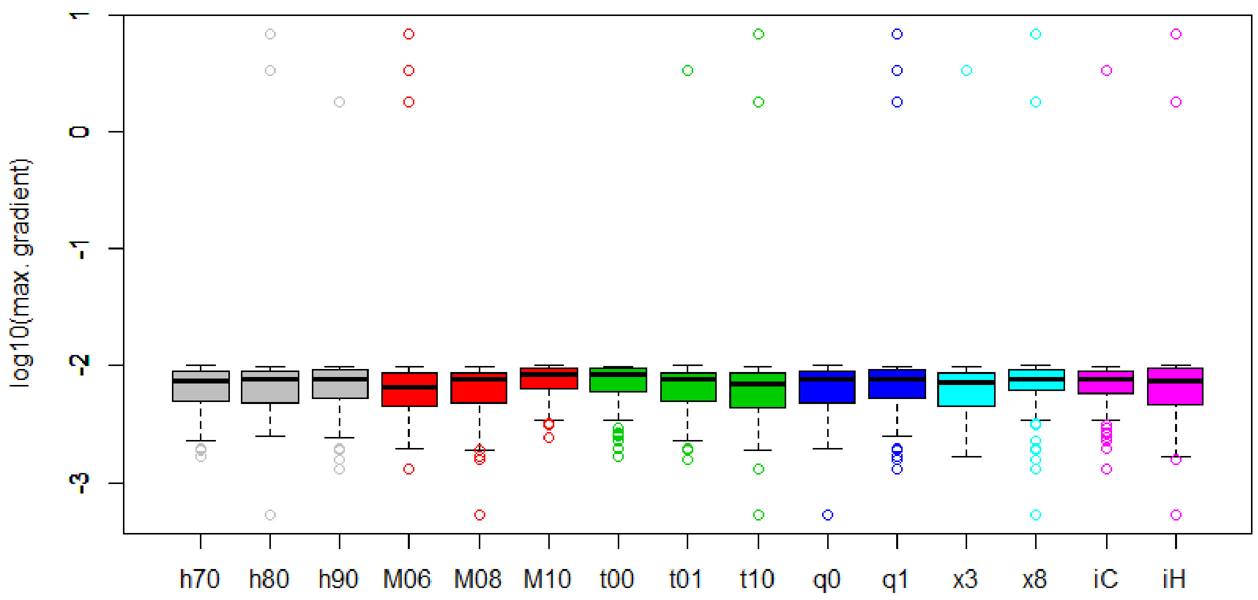



**Figure 8. Yellowfin reference set OM (OM-ref), model convergence summary, indicating marginal performance for 3 of 216 models. All models are represented within each uncertainty dimension (indicated by colours), partitioned by assumptions levels (X-axis labels), marginalized over the other assumption level. e.g. Together, the 3 grey boxes summarize all 216 models, with each boxplot representing the 66 model subset corresponding to the indicated steepness (h) assumption; the red boxes summarize the same 216 models, marginalized over the 3 natural mortality assumptions.**

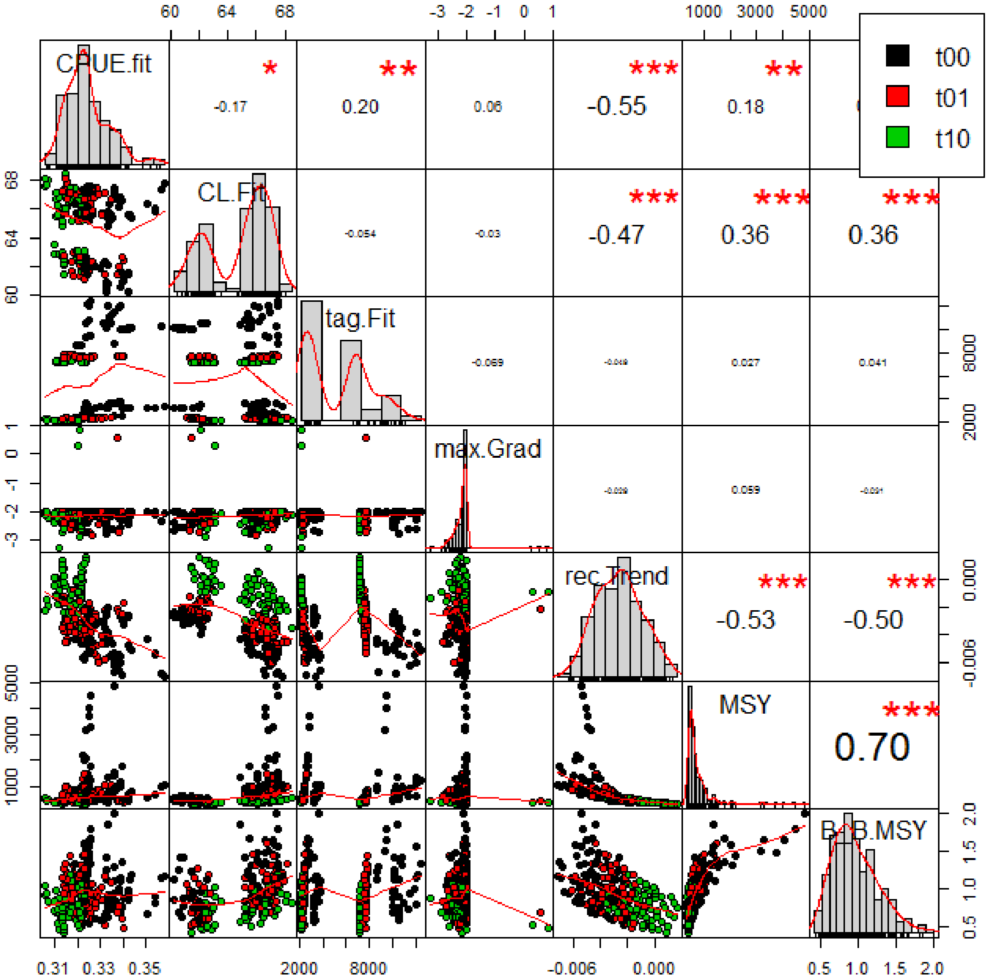



**Figure 9. Correlations among OM-ref summary diagnostics and productivity indices. Scatterplot colours indicate the tag weighting assumption. In this figure, the CPUE fit describes the mean RMSE across all regions, and the size composition fit describes the mean post-fit Effective Sample Size across all fisheries.**

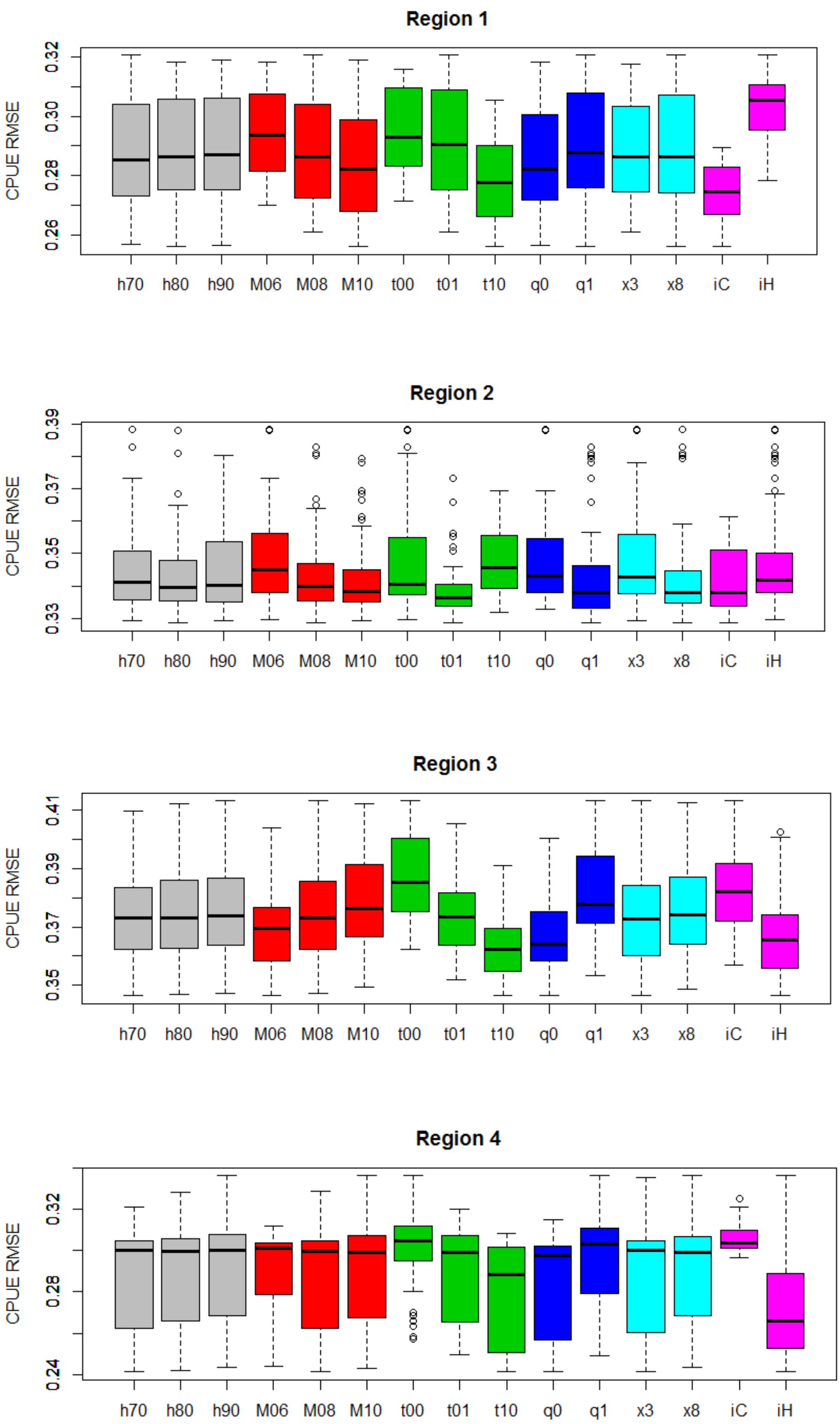



**Figure 10. Yellowfin reference set OM (OM-ref), quality of fit (RMSE) between predicted and observed CPUE (each panel represents a region). Within a panel, all models are represented within each uncertainty dimension (indicated by colours), partitioned by assumptions levels (X-axis labels), marginalized over the other assumption level. e.g. Together, the 3 grey boxes in each panel summarize all 216 models, with each boxplot representing the 72 model subset with the indicated M options; the red boxes summarize the same 216 models, marginalized over the 3 tag weighting options.**

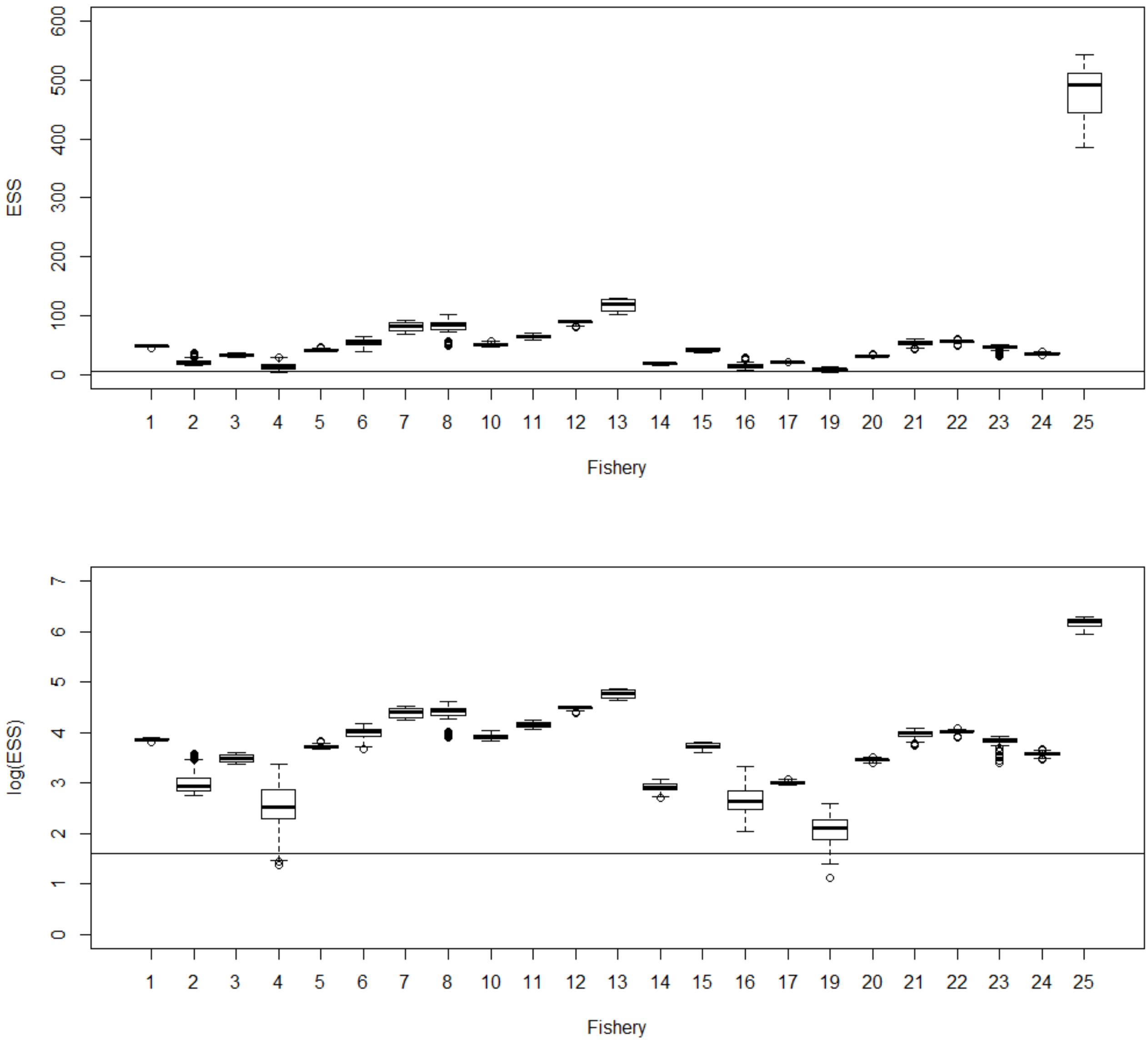



**Figure 11. OM-ref summary of the size composition data quality of fit between model predictions and observations by fishery. Each boxplot represents the distribution of 216 mean (over time) post-fit Effective Sample Sizes (lower panels are log-scale). Reference lines represent the assumed sample sizes input to the model (5 for all fisheries).**

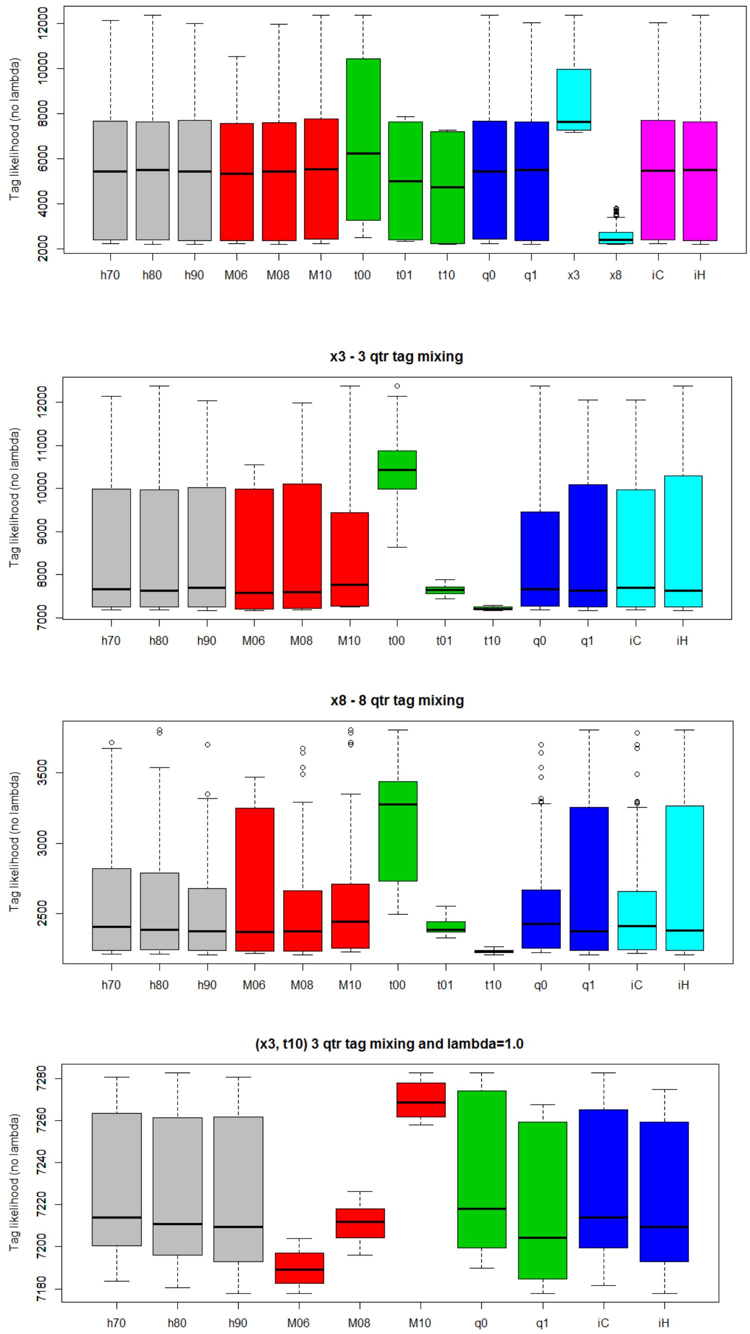



**Figure 12. OM-ref tag recapture likelihoods (i.e. excluding the lambda weighting factor that is applied in the objective function). The top panel represents all 216 models partitioned by all assumptions levels. The second panel represents the 108 models with the short tag mixing period (x3) only, the third panel represents the long tag mixing period (x8) only. The fourth panel represents the 36 models with the short mixing period (x3) and full tag weighting (t10).**

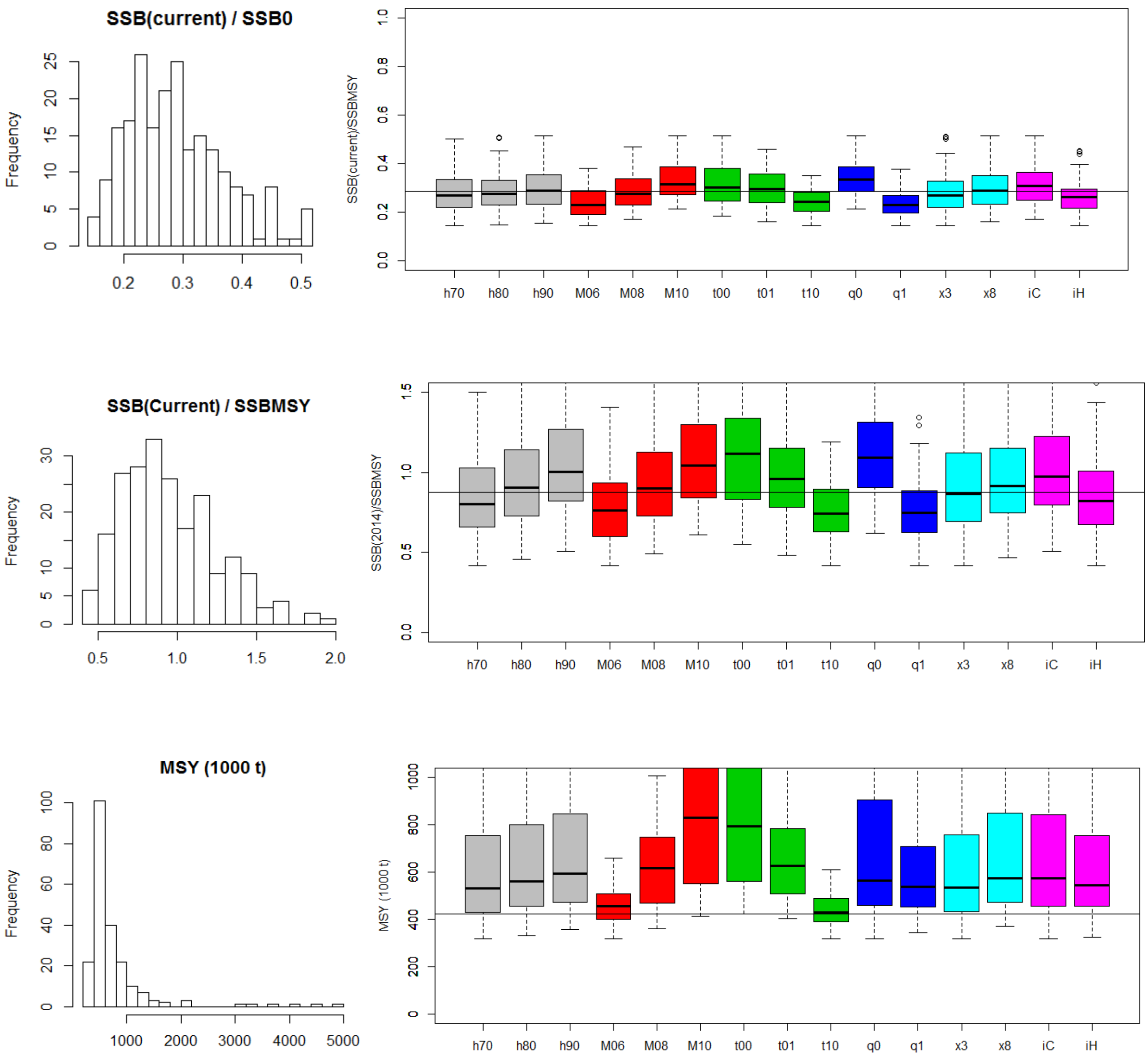



**Figure 13. OM-ref stock status summary plots indicating how the uncertainty in the model ensemble is partitioned according to assumptions. Horizontal line is the value from SA-base.**

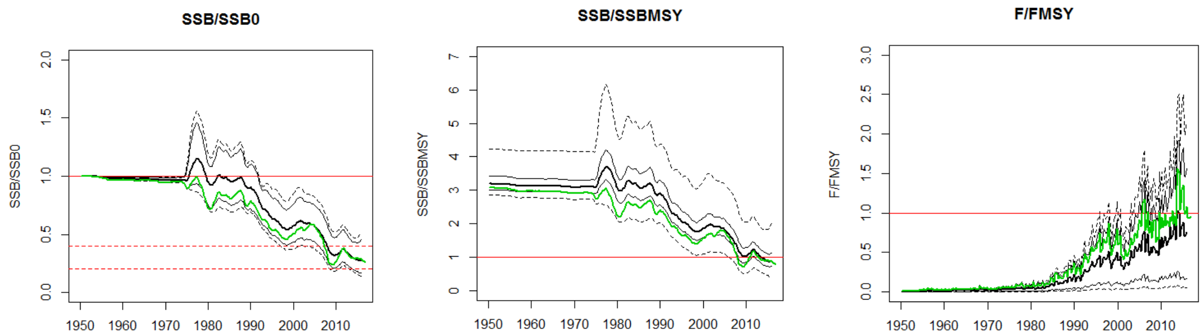



**Figure 14. Time series plots comparing the OM-ref ensemble (black - 0, 25, 50, 75, 100th percentiles) with SA-base (green line)**

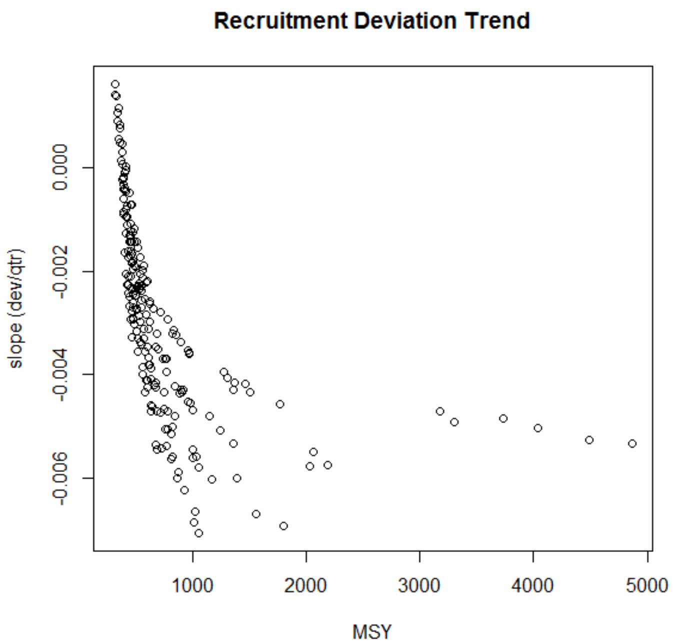
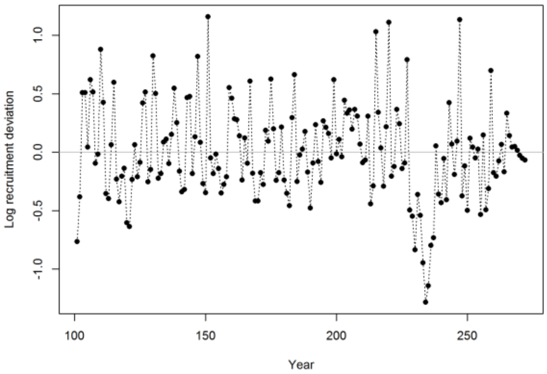
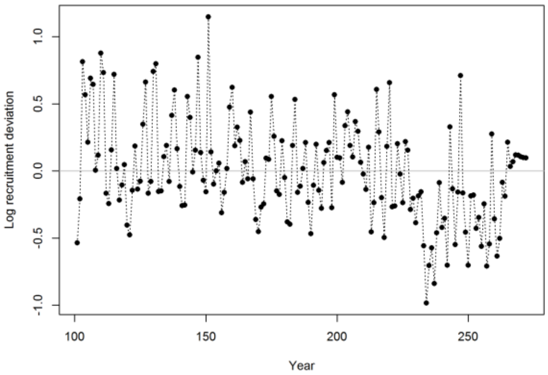



**Figure 15. Relationship between the recruitment deviation time series trend and MSY for OM-ref models.**

A) SA-base-analogue, MSY = 422 Kt




B) OM-ref model (h80, M08, t01, q0,x8, iC), MSY = 813 Kt




C) OM-ref model (h80, M08, t01, q0,x8, iC), MSY = 4800 Kt

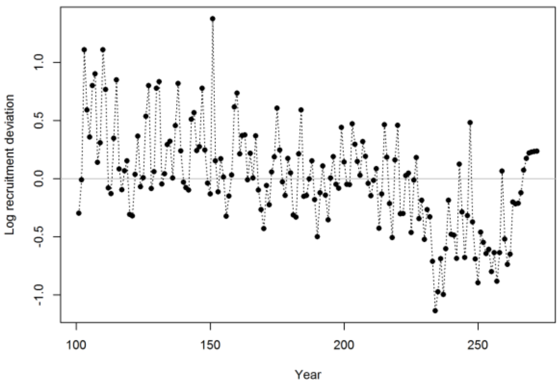



**Figure 16. Recruitment deviation time series for the SA-base-analogue, an OM-ref model with MSY about double the SA-base-analogue and the OM-ref model with the highest MSY.**

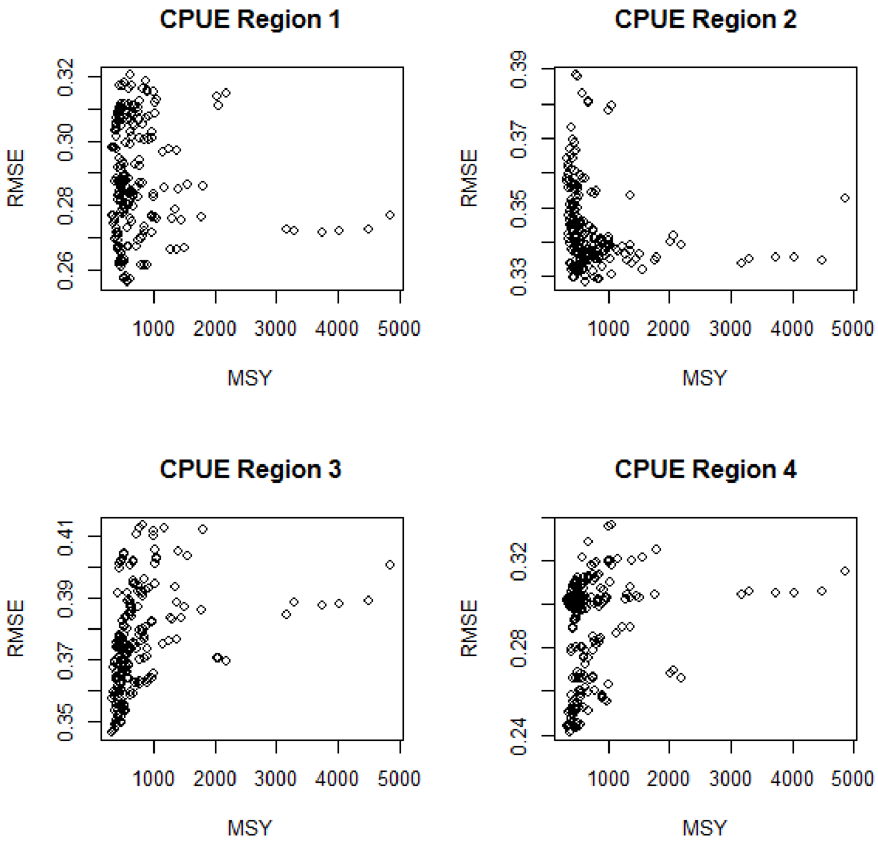



**Figure 17. Relationship between MSY and the quality of fit to the CPUE indices for OM-ref models.**

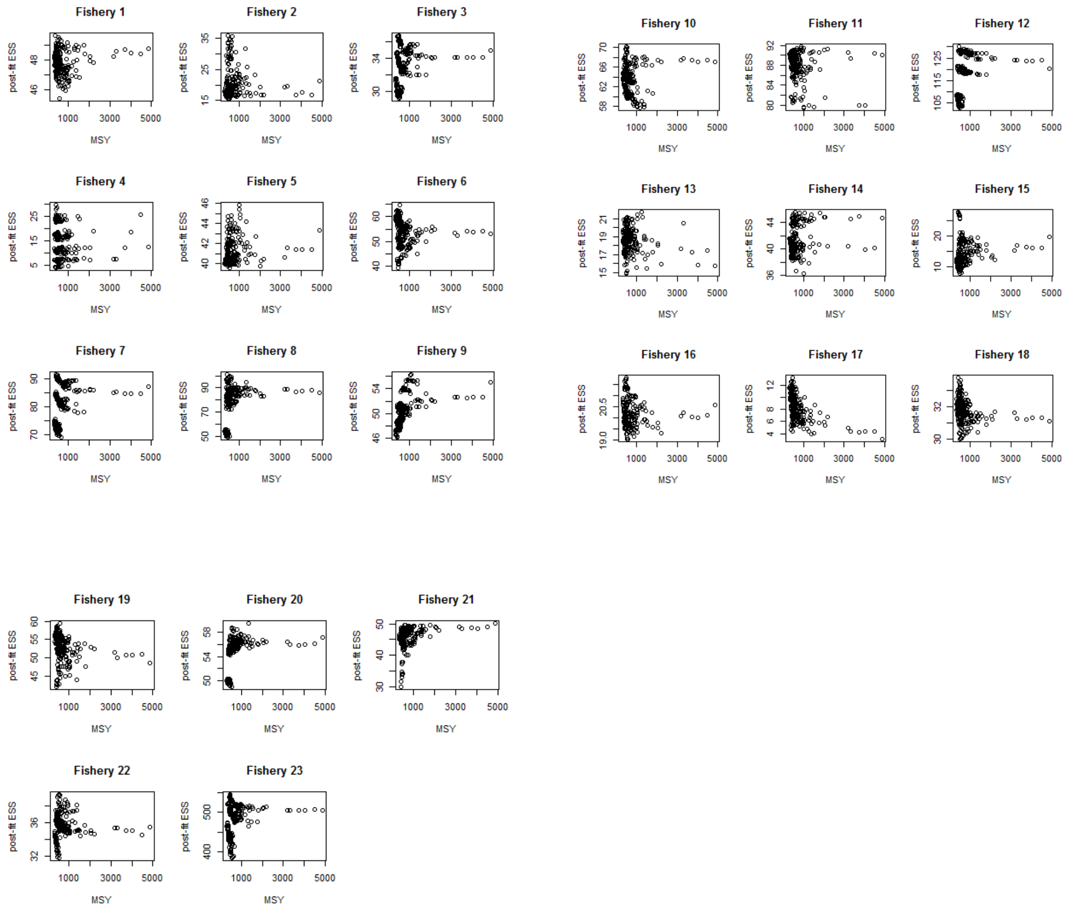



**Figure 18. Relationship between MSY and the quality of fit to the size composition data (post-fit effective sample sizes) for OM-ref models.**

# 4 MP performance when evaluated with the yellowfin reference case OM.

The focus of this WP is the exploration of robustness set and reference set OMs, however, given that MP tuning levels have been defined by the TCMP, we had an initial look at the interaction between OM-ref and the tuning levels. Time series results of 3 MPs (one being constant catch for comparison), are shown in Figure 19 - Figure 22, tuned to the 4 tuning levels from Table 4.

Tuning levels T1 and T2, (initial proposals from the TCMP, inferred from IOTC resolutions) potentially have some unattractive features. The productivity of the OM-ref ensemble is high enough that the MPs need to increase catches to bring biomass down to the targets. This creates a long term pattern in which biomass is generally above the target in the early years, then below in later years.  There is a general declining biomass trend at the end of the projection period.

In contrast, tuning levels T3 (identified by the TCMP for bigeye) and T4 (arbitrarily defined by the authors) are more conservative with respect to biomass, and suggest that the MPs can deliver more stable catches over the projection period, leaving the biomass at healthy levels at the end of the projection period.

Example MP performance for the T1 tuning is shown for two MPs (Figure 23) partitioned for the highest ($\lambda = 1.0$) and lowest ($\lambda = 0.0$) tag weighting from OM-ref. This indicates that the IT MP struggles with the more pessimistic ($\lambda = 1.0$) option, while the PT4010 MP mostly manages to avoid a catastrophic stock collapse. Both MPs may fail to achieve stable dynamics in the ($\lambda = 0.0$) scenario as well, due to rapid catch increases. Figure 24 shows more stable performance for both MPs with the more conservative T3 tuning.

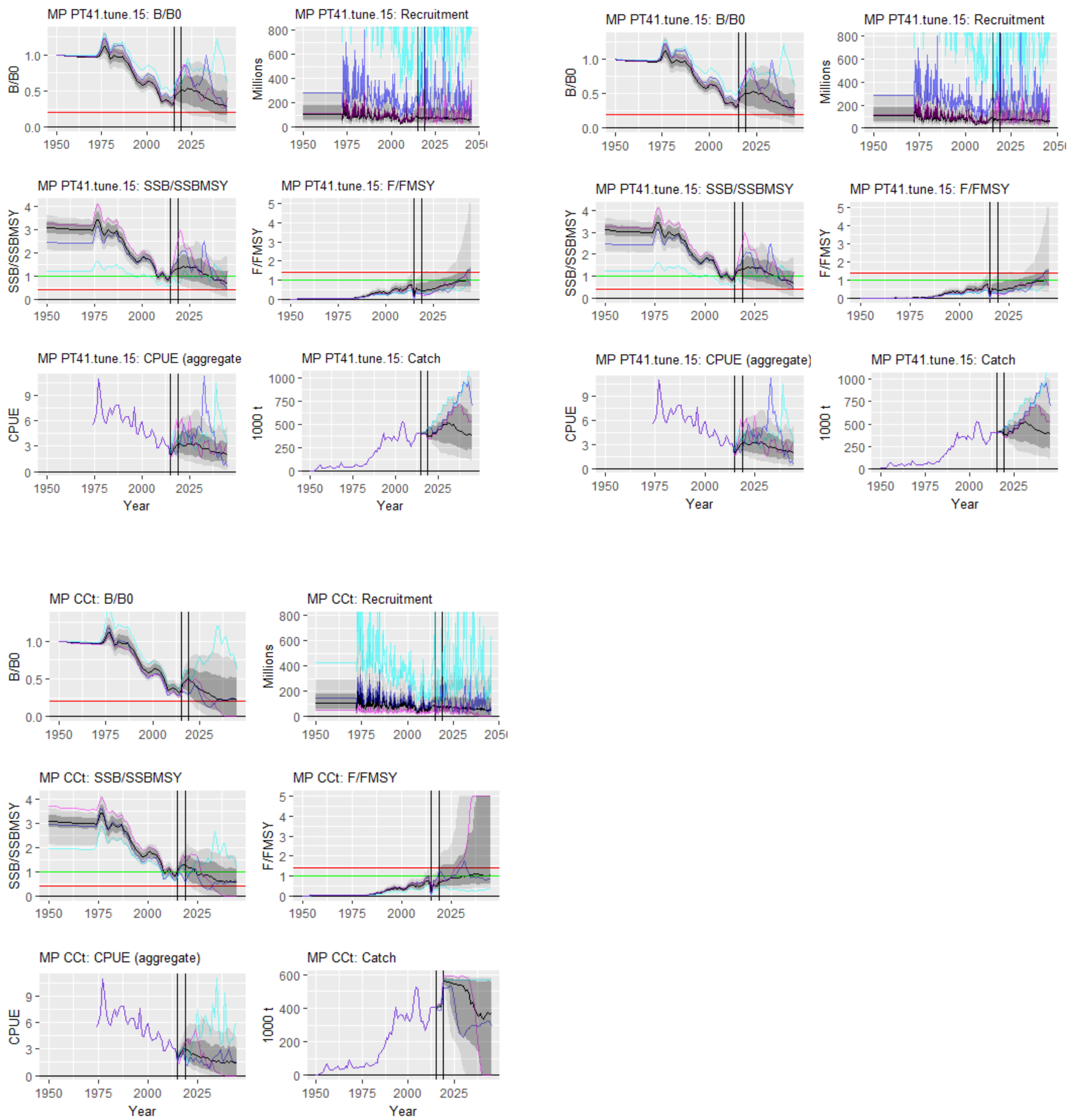



**Figure 19. MSE summary plots for three MPs (IT = CPUE Target, CC = Constant Catch), for tuning objective T1 (OM-ref, 216 realizations, Pr(mean(B(2019:2039)/BMSY)=1.0) = 0.5).**

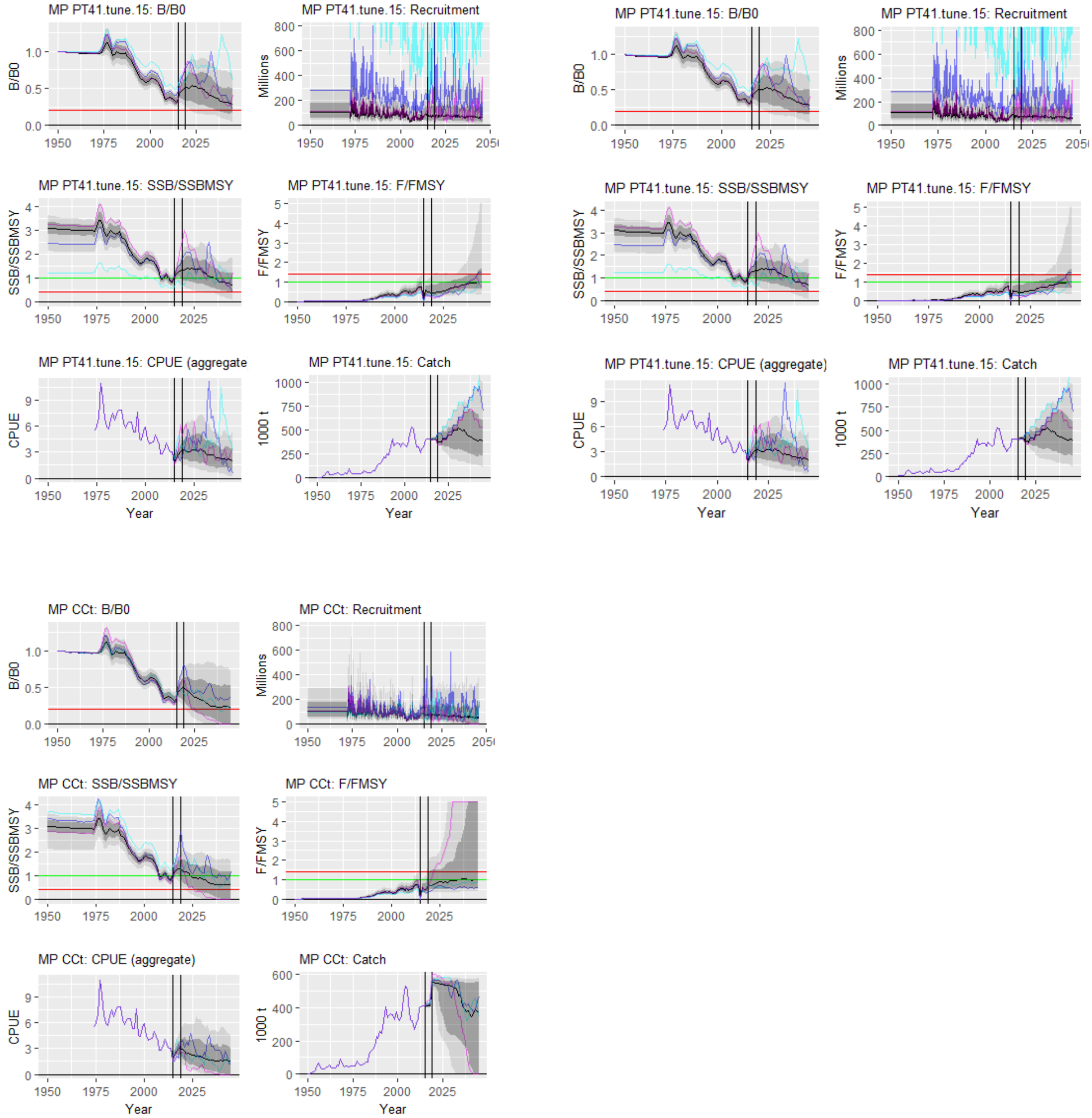



**Figure 20. MSE summary plots for three MPs (IT = CPUE Target, CC = Constant Catch), for tuning objective T2 (OM-ref, 216 realizations, Pr(mean(B(2024)/BMSY)=1.0) = 0.5).**

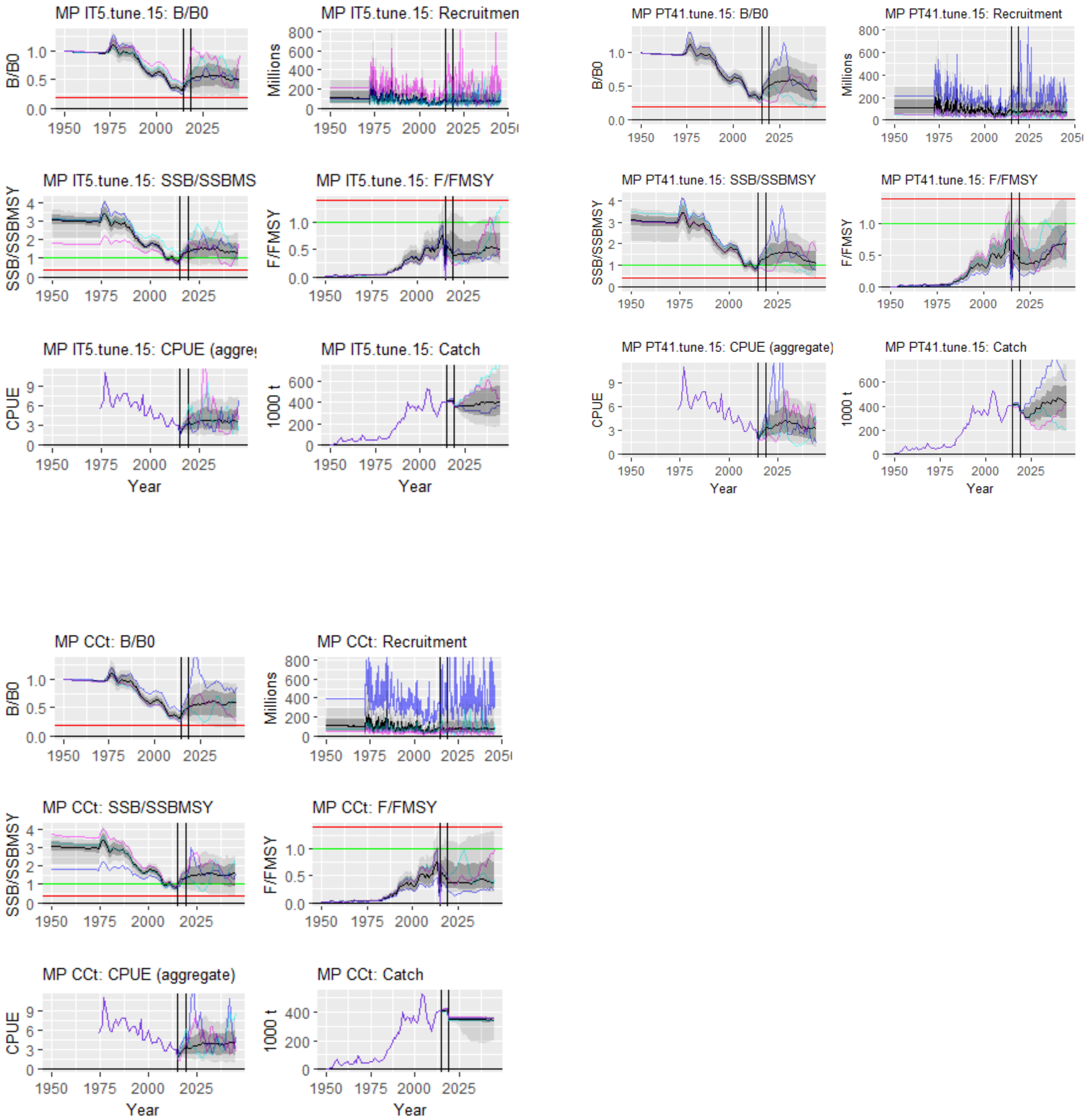



**Figure 21. MSE summary plots for three MPs (IT = CPUE Target, CC = Constant Catch), for tuning objective T3 (OM-ref, 216 realizations, Pr(green Kobe) = 0.75).**

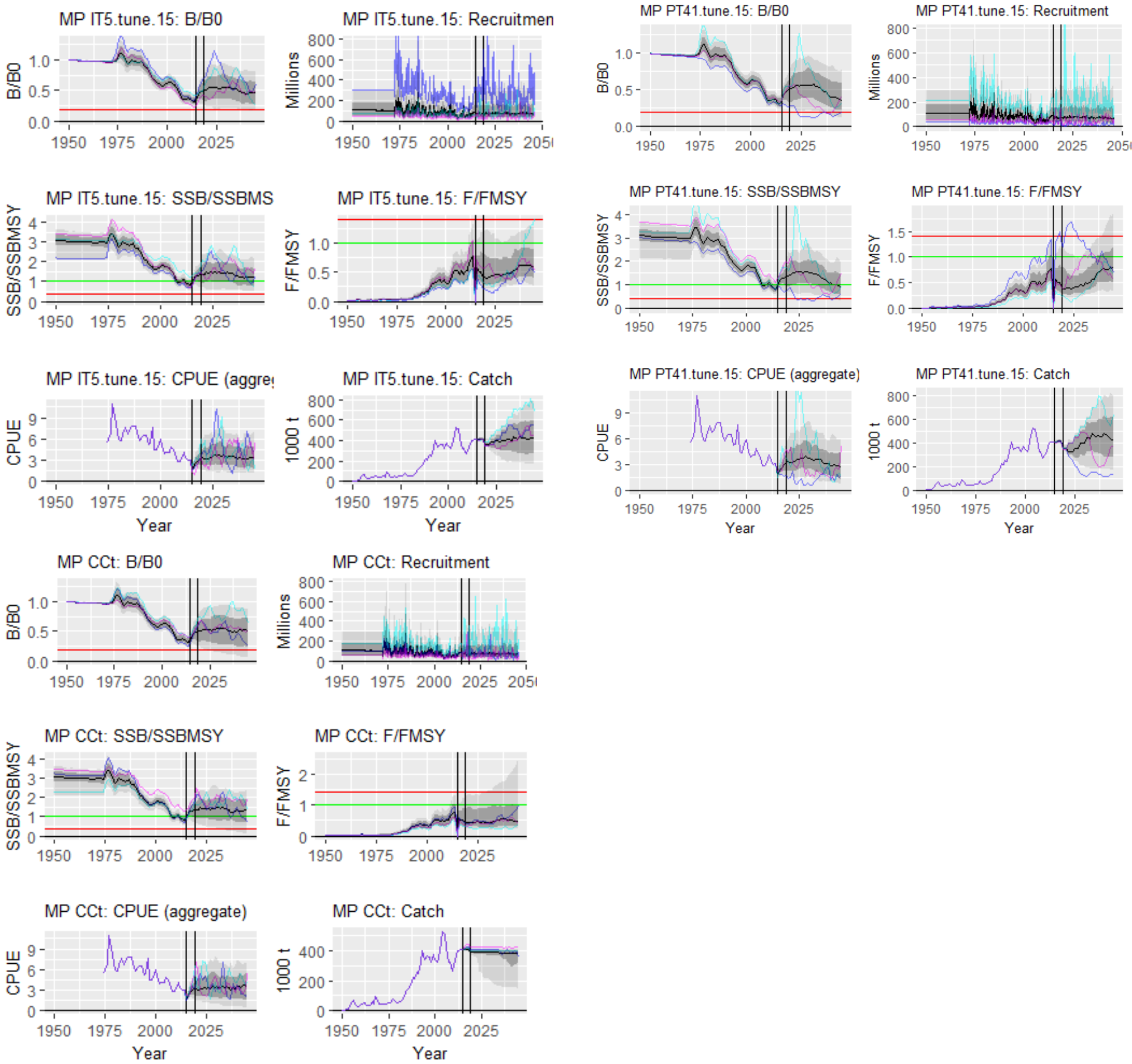



**Figure 22. MSE summary plots for three MPs, for tuning objective T4 (OM-ref, 216 realizations, Pr(mean(B(2019:2039)/BMSY)=1.5) = 0.5).**

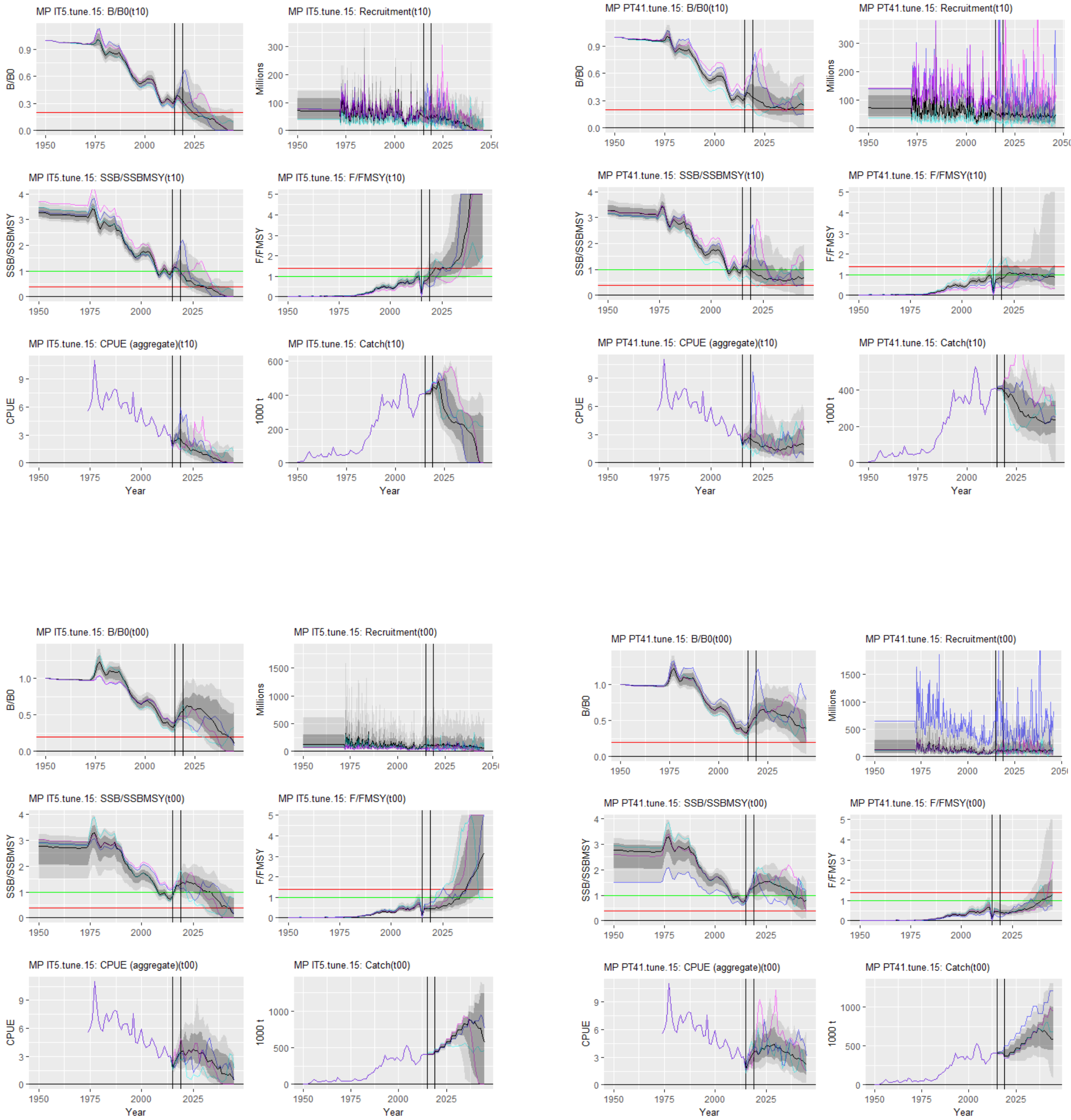



**Figure 23. MSE summary plots for two MPs (IT = CPUE Target, CC = Constant Catch), for tuning objective T1 (OM-ref, 216 realizations, Pr(mean(B(2019:2039)/BMSY)=1.0) = 0.5), partitioned by the highest (top panels) and lowest (bottom) tag weighting assumptions (λ = 1.0 (t10), 0.0 (t00)).**

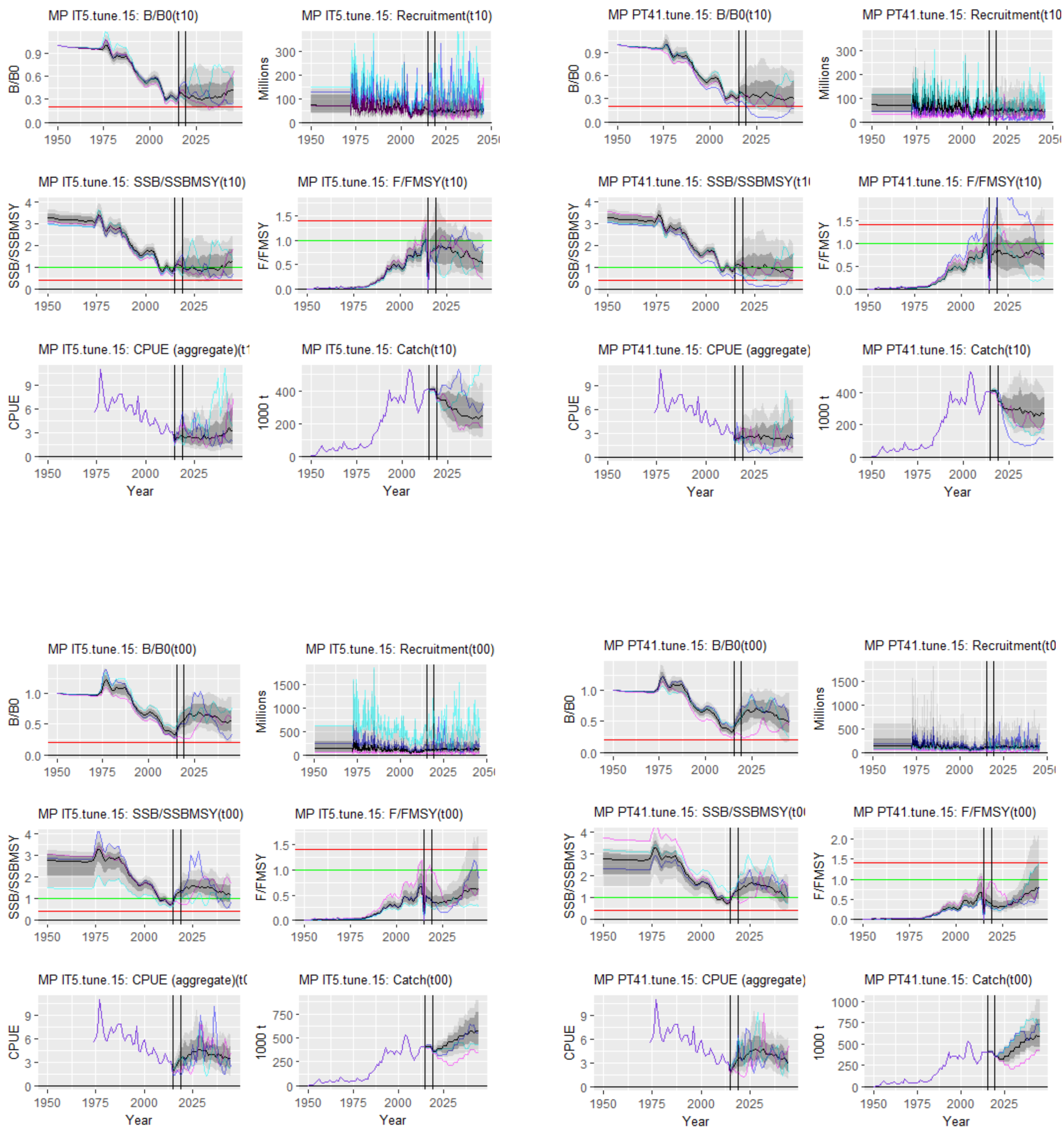



**Figure 24. MSE summary plots for two MPs, for tuning objective T3 (OM-ref, 216 realizations, Pr(green Kobe) = 0.75) (OM-ref, 216 realizations, Pr(mean(B(2019:2039)/BMSY)=1.5) = 0.5), partitioned by the highest (top) and lowest (bottom) tag weighting assumptions.**

# 5 Yellowfin robustness set OMs

WPM (2016) proposed some yellowfin OM robustness scenarios, which were further clarified in the *Informal Methods working group on MSE* (MSE 2017), resulting in the interpretations explored below.

## 5.1 Potential "juvenilization" of longline fishery selectivity

In WPTT (2016) it was noted that there was evidence for gradually shifting of fishery selectivity toward younger sizes/ages for some tuna species in the Atlantic Ocean. This could introduce bias to the assessment through several paths, including misleading catch size composition and relative abundance indices. The potential effects of this scenario was explored by contrasting 3 longline selectivity options in a grid of 36 models defined as OM-robSelTrend (Table 1):

1) SS = Stationary Selectivity (all fleets) is the assumption in OM-ref and SA-base.
2) NS = Non-Stationary longline selectivity (other fleets stationary) - this scenario estimates time series variability in longline selectivity, with independent selectivity parameters estimated every 10 years. This is clearly not a detailed effort to understand and describe the complicated issues involved with the confounding of movement and catchability, and the implications of sharing selectivity and catchability among fleets and regions through the dynamic history of this fishery. It is simply an attempt to see if there is evidence for temporal variability in LL selectivity, and if the estimated changes are likely to have an important effect on MP performance.
3) ST = Selectivity Trend in longline fleets (other fleets stationary) - this scenario estimates selectivity with a continuous change over some or all of the fishery history. Stock Synthesis achieves this by making the selectivity parameters a logistic function of time (the function is defined by two parameters - the date of the inflection point, and the rate of change). Note that the OM-ref longline selectivity happens to be a logistic function of age, so ST describes selectivity as logistic functions of time and age (subject to the standard re-scaling such that max(selectivity)=1.0 in every year).

Figure 25 illustrates the longline selectivity estimates for the NS scenario (other assumptions as in OM-SA-analogue). The model estimates a substantial shift toward targeting younger ages over the first 30 years, and a lesser shift back toward older ages in the latter 30 years.

Figure 26 illustrates the longline selectivity estimates for the ST scenario (other assumptions as in OM-SA-analogue). The model estimates a rapid shift toward targeting younger ages over the first 10 years, within minimal change afterward. Unlike the NS scenario, the ST parameterization is a monotonic function, i.e. it lacks the flexibility to estimate a reversal of the selectivity trend.

Figure 27 - Figure 29 show the OM-robSelTrend quality of fit indices, which suggests that the added flexibility from the non-stationary selectivity does not substantially change the fit to the data.

Figure 30 compares the estimated stock status associated with the different selectivity assumptions, and Figure 31 compares biomass projections from the MSE results, partitioned by selectivity option. In both cases, it appears that the non-stationary longline selectivity scenarios are not much different (NS is slightly more optimistic), than the OM-ref stationary selectivity assumption.

Since the non-stationary longline selectivity dynamics do not appear to add any substantial challenge to the OM, we would not recommend maintaining them in future MSE analyses as a robustness scenario (at least not with the NS and ST parameterizations tested here).

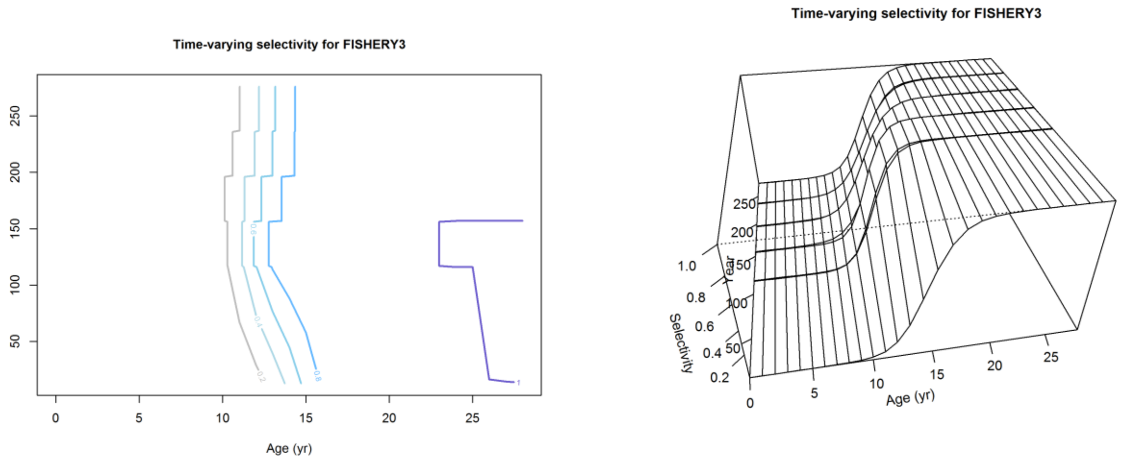



**Figure 25. Longline selectivity estimated for robustness scenario NS - showing a shift toward the preference of younger ages over the first 35 years, and a shift toward older ages over the final 30 years (time axes are in quarters). Other assumptions defined as in OM-SA-analogue.**

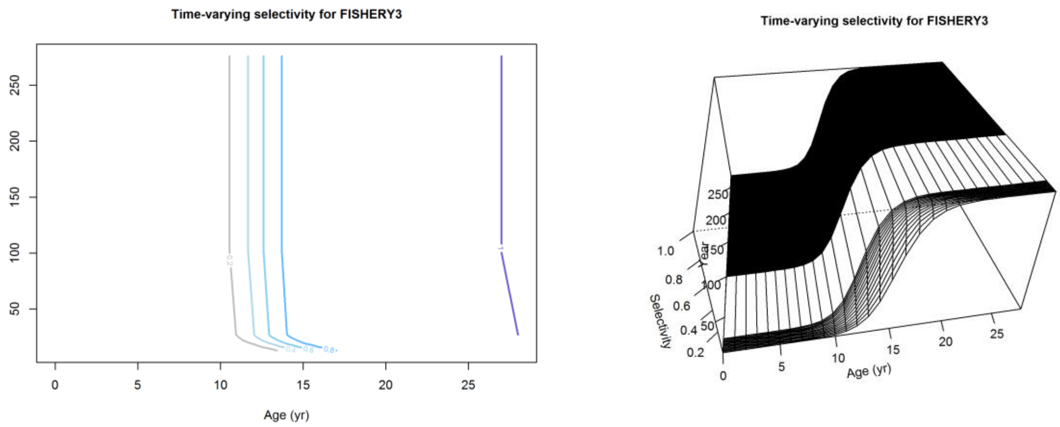



**Figure 26. Longline selectivity estimated for robustness scenario TS - showing a shift toward the preference of younger ages, but primarily over the first 6 years only (time axes are in quarters). Other assumptions defined as in OM-SA-analogue.**

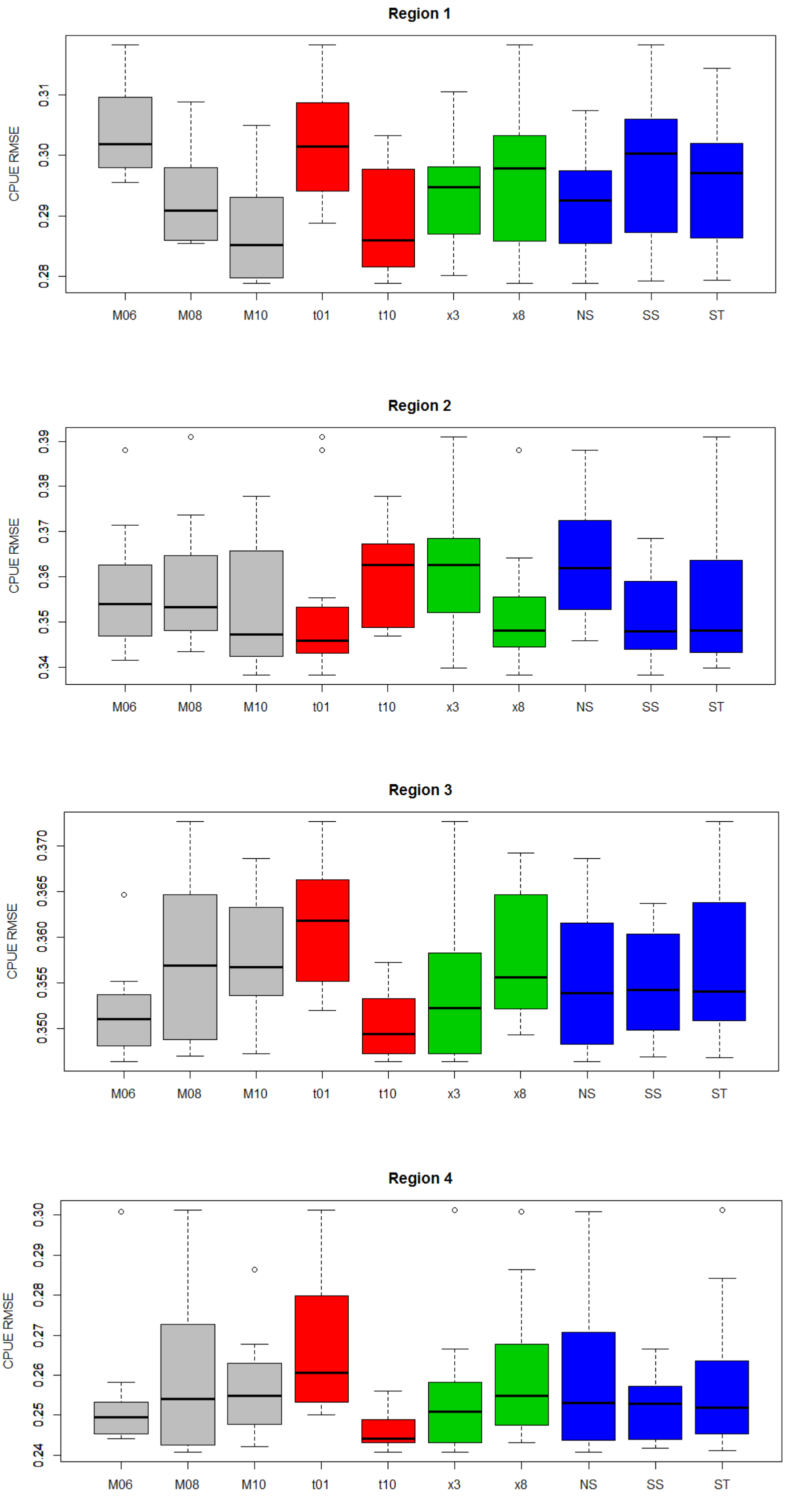
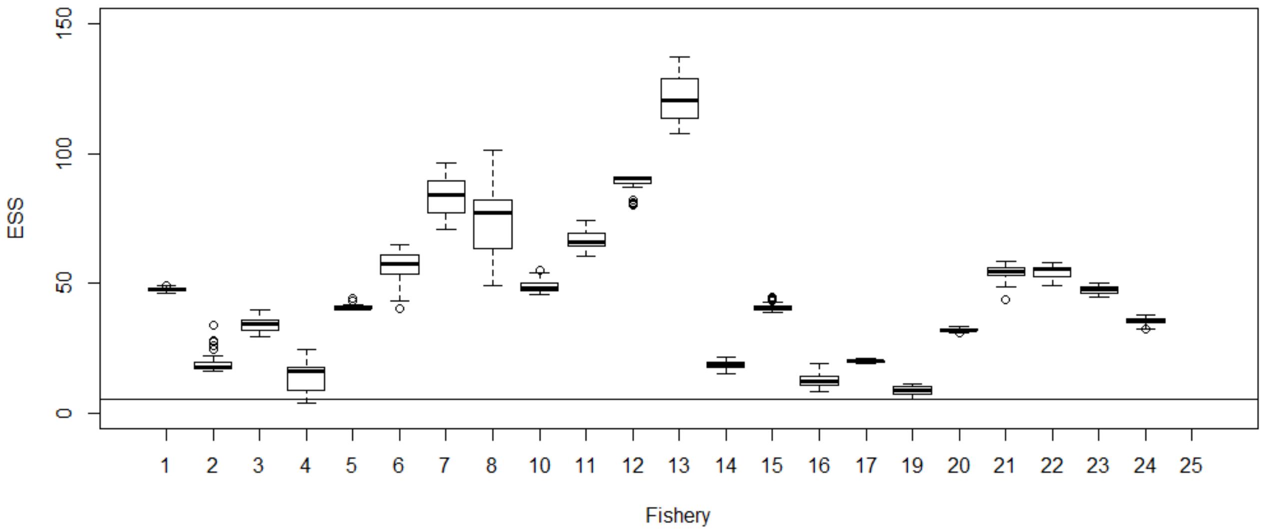



**Figure 27. Yellowfin robustness set OM-robSelTrend, quality of fit (RMSE) between predicted and observed CPUE (each panel represents a region). Within a panel, all models are represented within each uncertainty dimension (indicated by colours), partitioned by assumptions levels (X-axis labels), marginalized over the other assumption level. e.g. Together, the 3 grey boxes in each panel summarize all 36 models, with each boxplot representing the 12 model subset with the indicated M assumption; the red boxes summarize the same 36 models, marginalized over the 6 tag weighting assumption.**



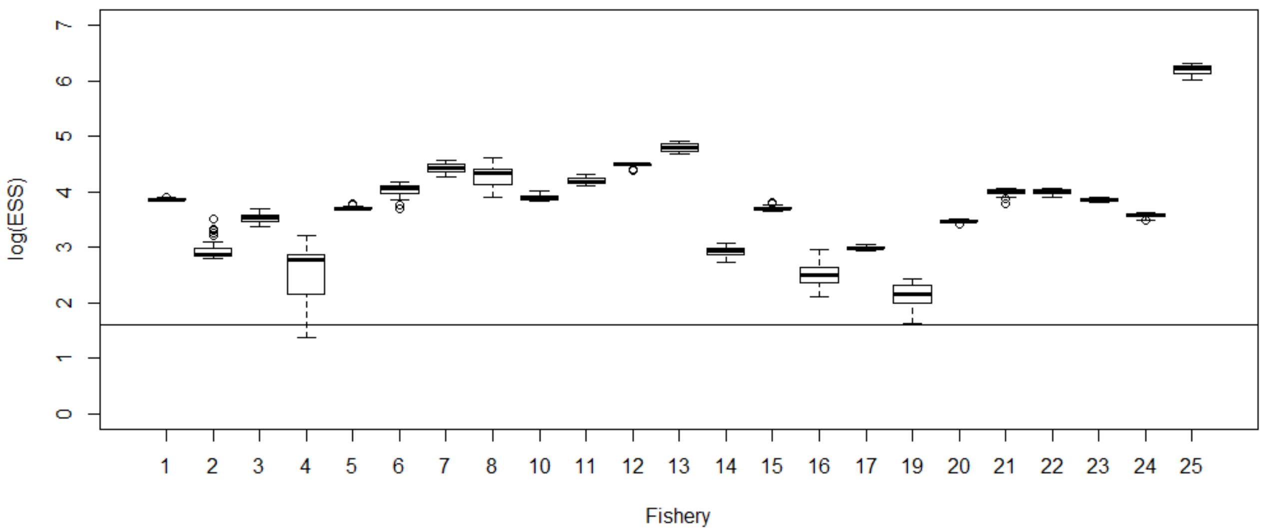
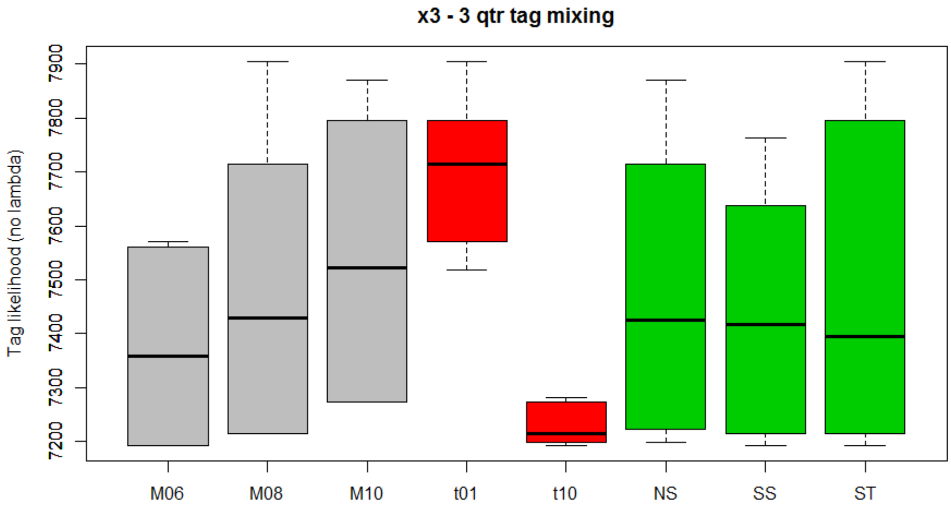




**Figure 28. OM-robSelTrend summary of the size composition data quality of fit between model predictions and observations by fishery. Each boxplot represents the distribution of 216 mean (post-fit) Effective Sample Sizes (arithmetic mean and log(mean) over all years). Reference lines represent the assumed sample sizes in the model (5 for all fisheries).**



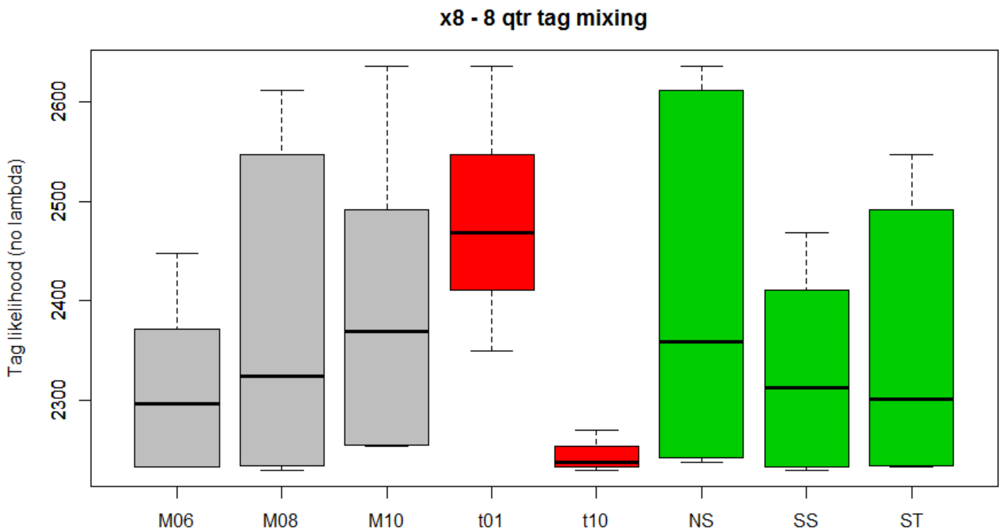
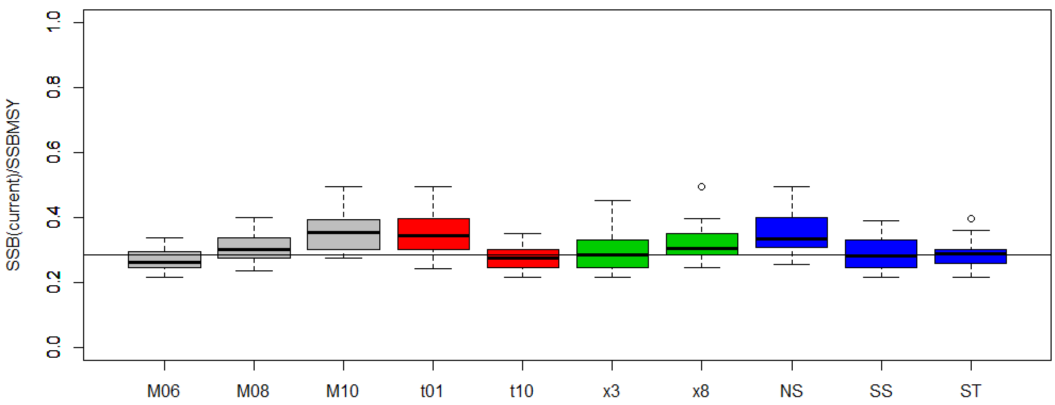
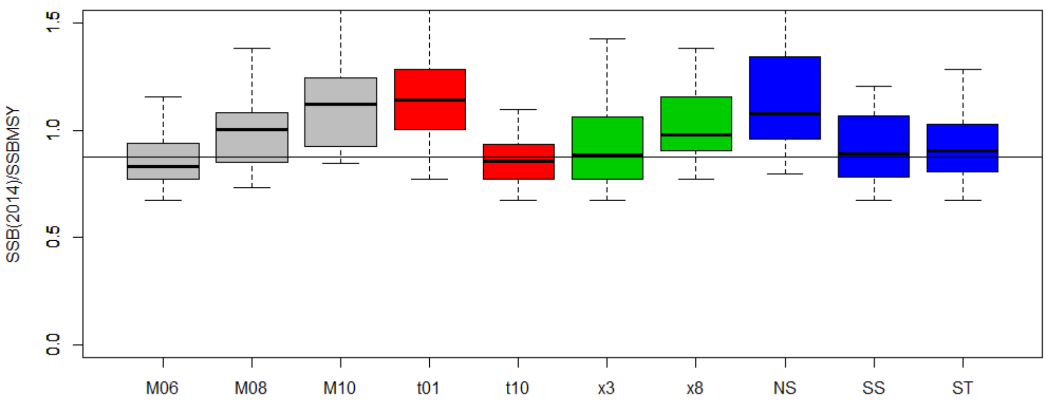




**Figure 29. Robustness ensemble OM-rob-SelTrend tag recapture likelihoods (before the lambda weighting factor is applied), partitioned by assumptions. Top panel corresponds to the short tag mixing period, bottom panel is the long mixing period.**



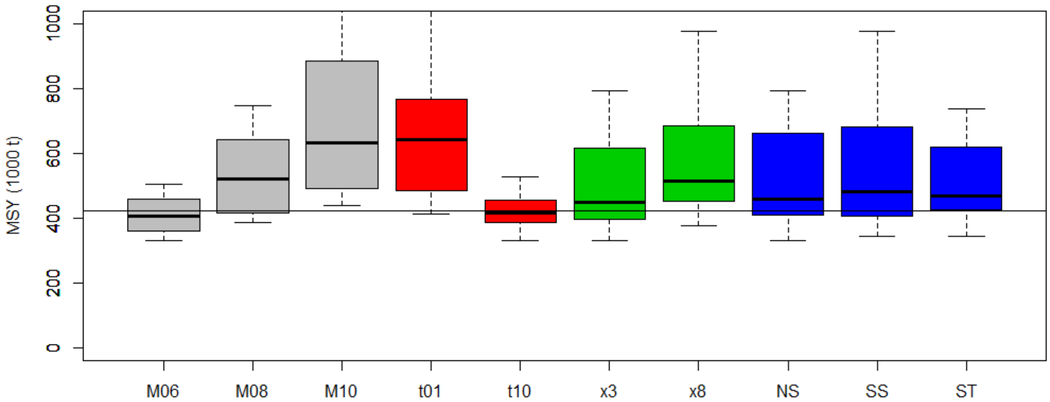




**Figure 30. OM-robSelTrend stock status summary plots indicating how the uncertainty in the model ensemble is partitioned according to assumptions. Horizontal line is the value from SA-base.**

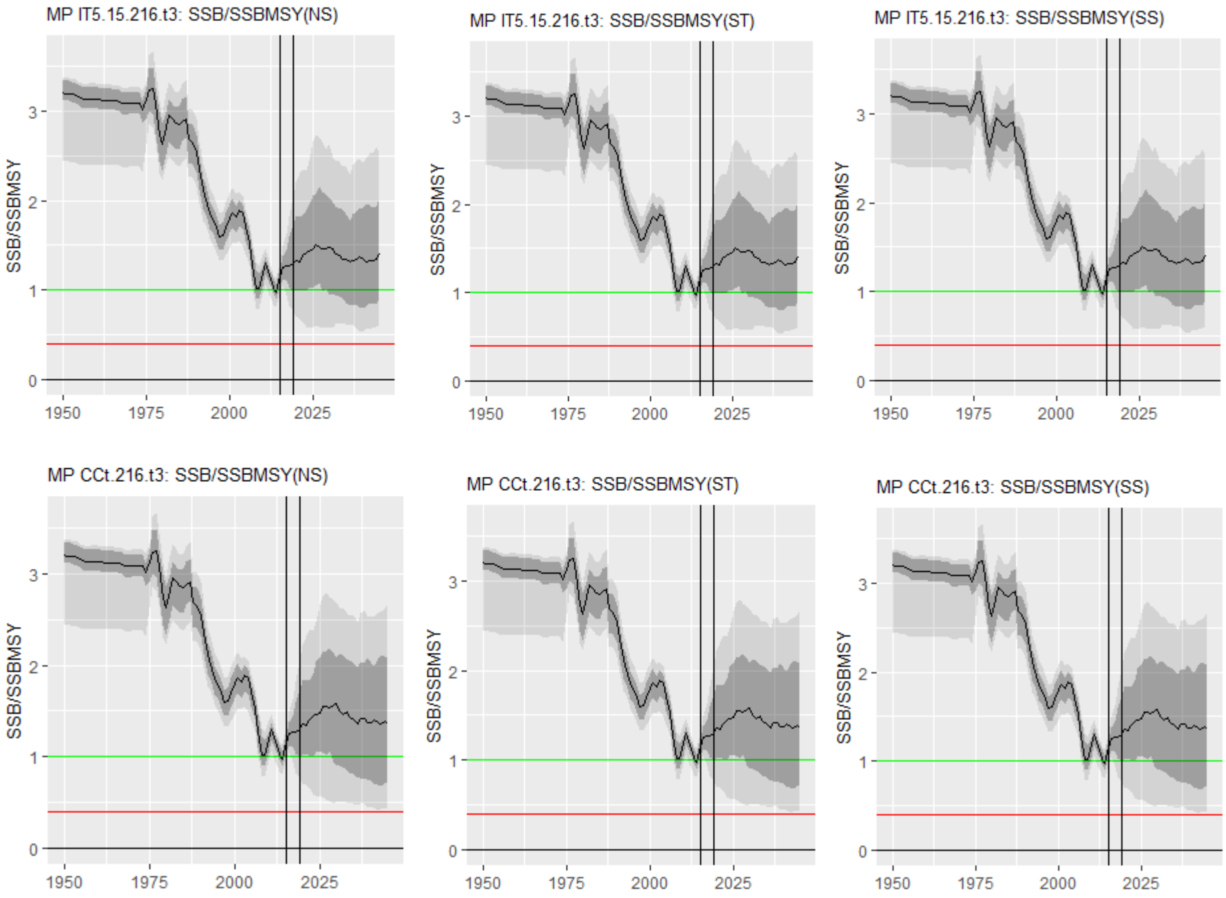



**Figure 31. Comparison of MP evaluations from a subset of OM-rob-selTrend. Columns represents three different longline selectivity assumptions (NS = selectivity estimated in 10 year blocks, ST = selectivity trend estimated as a logistic function of time, SS = stationary - the OM-ref default). Top row is a feedback-based MP (Index Target), bottom row is constant catch, both tuned to OM-ref (216 realizations) with tuning objective T3.**

## 5.2 OM-rob-tagWt: OM robustness scenario exploring increased weighting of the tagging data

This robustness scenario (defined in WPM 2016), explores what might happen if the tag data is disproportionately more informative than all other data in the model (t15 = tag weighting λ = 1.5). This is a curious scenario given the known problems with the tagging data (i.e. in particular, it is recognized that the low mixing rates of tagged and untagged fish relative to the scale of the assessment regions, can be expected to bias tag estimators). In this case, up-weighting the tagging data was suggested as being analogous to down-weighting all the other data. This should be a reasonable means of testing implications on parameter point estimates (but it will artificially increase parameter estimation precision). Since tag weights were found to be an influential uncertainty in OM-ref additional down-weighting levels were added to better illustrate the continuum of tag influence (λ = 1.5, 1.0, 0.1, 0.01, 0.001, 0.0). Tag weighting assumptions were evaluated in a grid of 36 models defined as OM-rob-tagWt in Table 1.

Figure 32 - Figure 35 illustrate the quality of fit diagnostics for OM-rob-tagWt. The tags appear to be more compatible with CPUE from regions 1,3 and 4, and most conflicting with region 2 (Figure 32). The fits to the aggregate tagging data are very poor and visually indistinguishable for tag weighting λ = 0.01, 0.001, 0.0. The fit is much better (though not great) and similar for λ = 1.0, 1.5, while λ = 0.1 is distinctly intermediate between the other groups (Figure 34). The stock status groupings are similar (Figure 36), with higher tag weight being more pessimistic.

Figure 37 illustrates MP evaluations for 4 tag weighting assumptions in OM-rob-tagWt, and two MPs (Index Target and Constant Catch, tuned with OM-ref (216 realizations) for tuning objective T3. The projected biomass (and other performance characteristics not shown) are consistent with the assessment diagnostics (and OM-ref MP outcomes). Tag-weighting λ = 1.5 is similar to, but slightly more pessimistic than λ = 1.0, and both are considerably more pessimistic than the down-weighted scenarios (λ = 0.1, 0.0 in this case).

The exploration again emphasizes that the tag weighting is very influential to the stock status and MSE. It is not clear that the difference between λ = 1.5 and λ = 1.0 is distinct enough to justify a new robustness scenario, but it is clear that within the current assessment structure, scenarios with high tag weighting should form a core part of the MSE evaluation.

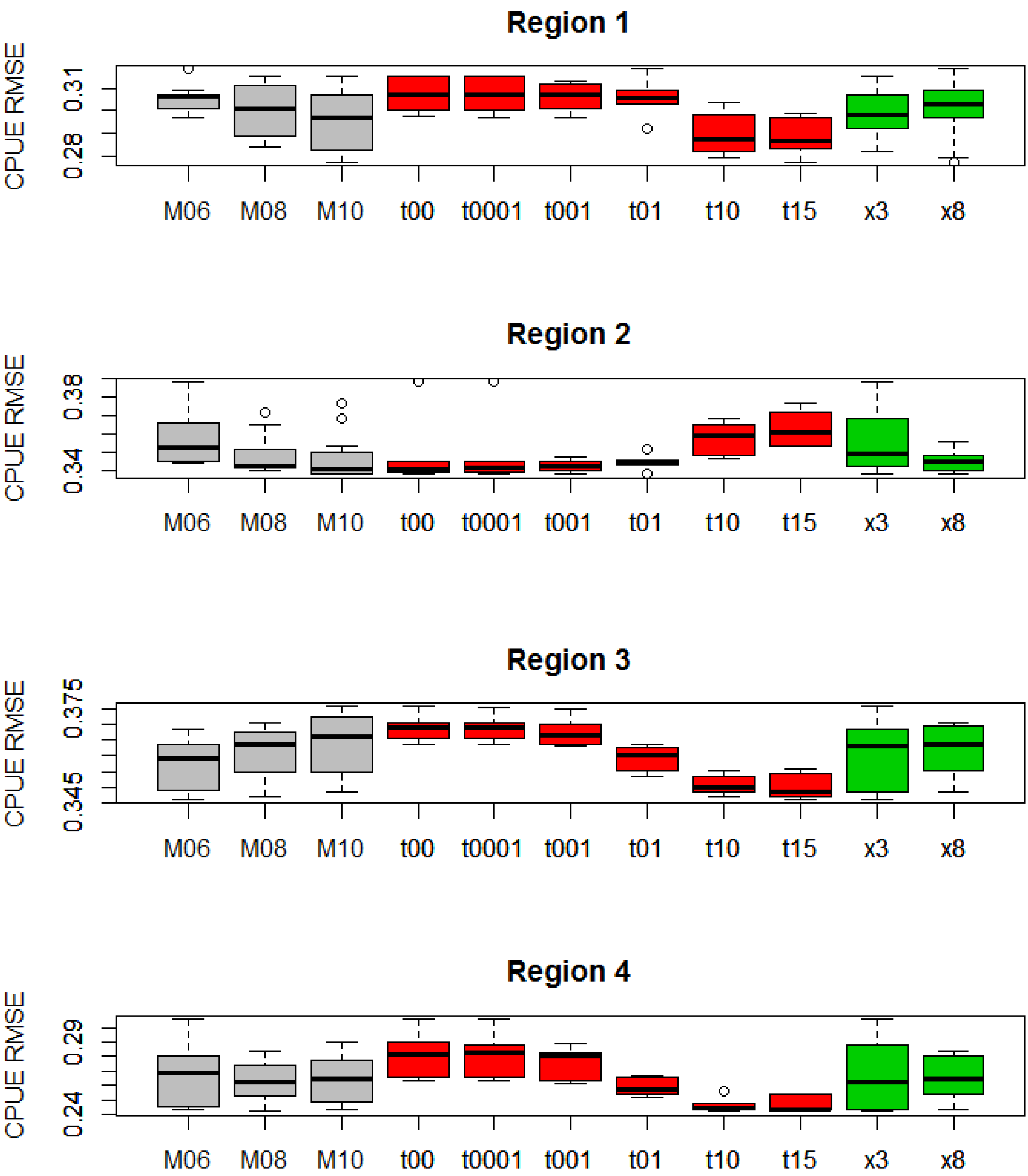



**Figure 32. Yellowfin robustness scenario OM-rob-tagWt, quality of fit (RMSE) between predicted and observed CPUE (each panel represents a region). Within a panel, all models are represented within each uncertainty dimension (indicated by colours), partitioned by assumptions levels (X-axis labels), marginalized over the other assumption level. e.g. Together, the 3 grey boxes in each panel summarize all 36 models, with each boxplot representing the 12 model subset with the indicated M assumption; the red boxes summarize the same 36 models, marginalized over the 6 tag weighting assumption.**

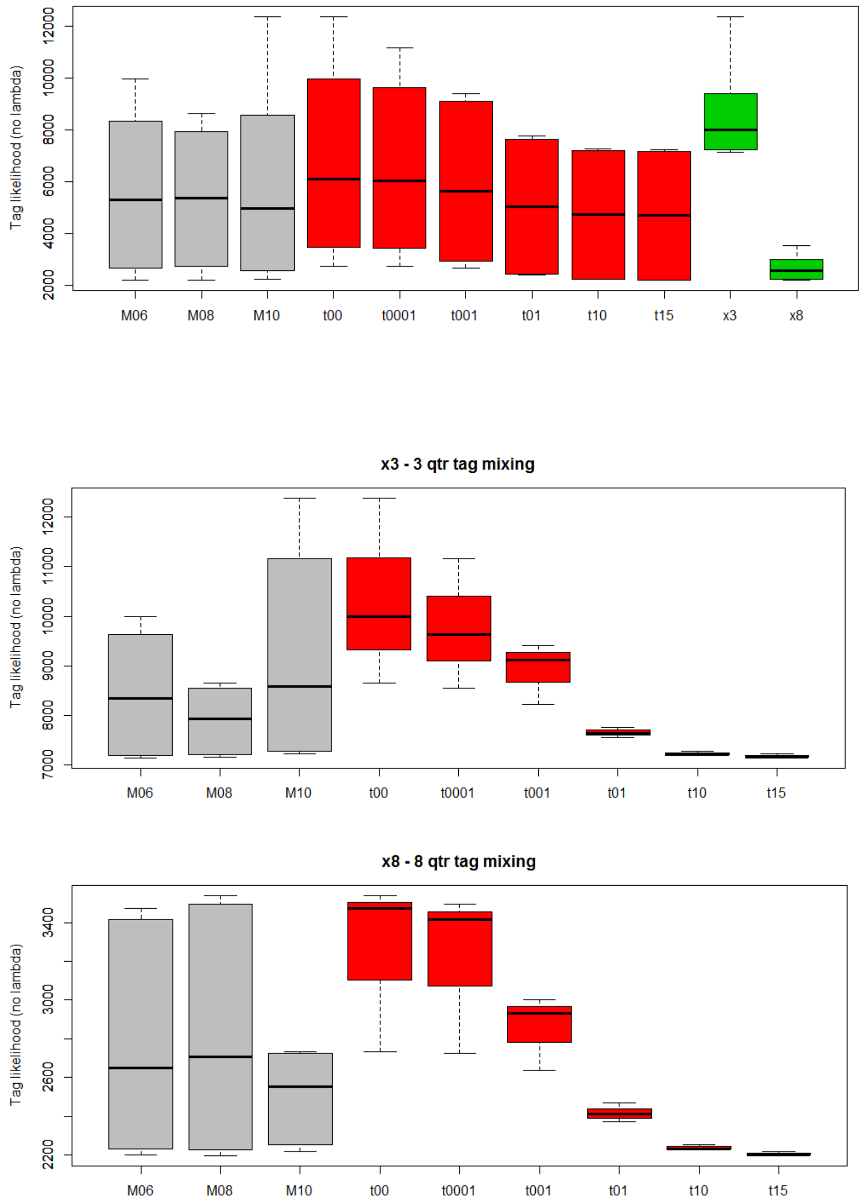



**Figure 33. OM-rob-tagWt tag recapture likelihoods (before the lambda weighting factor is applied). The top panel represents all 36 models partitioned by all assumptions levels. The second panel represents the 18 models with the short tag mixing period (x3), the third panel represents the long tag mixing period (x8).**

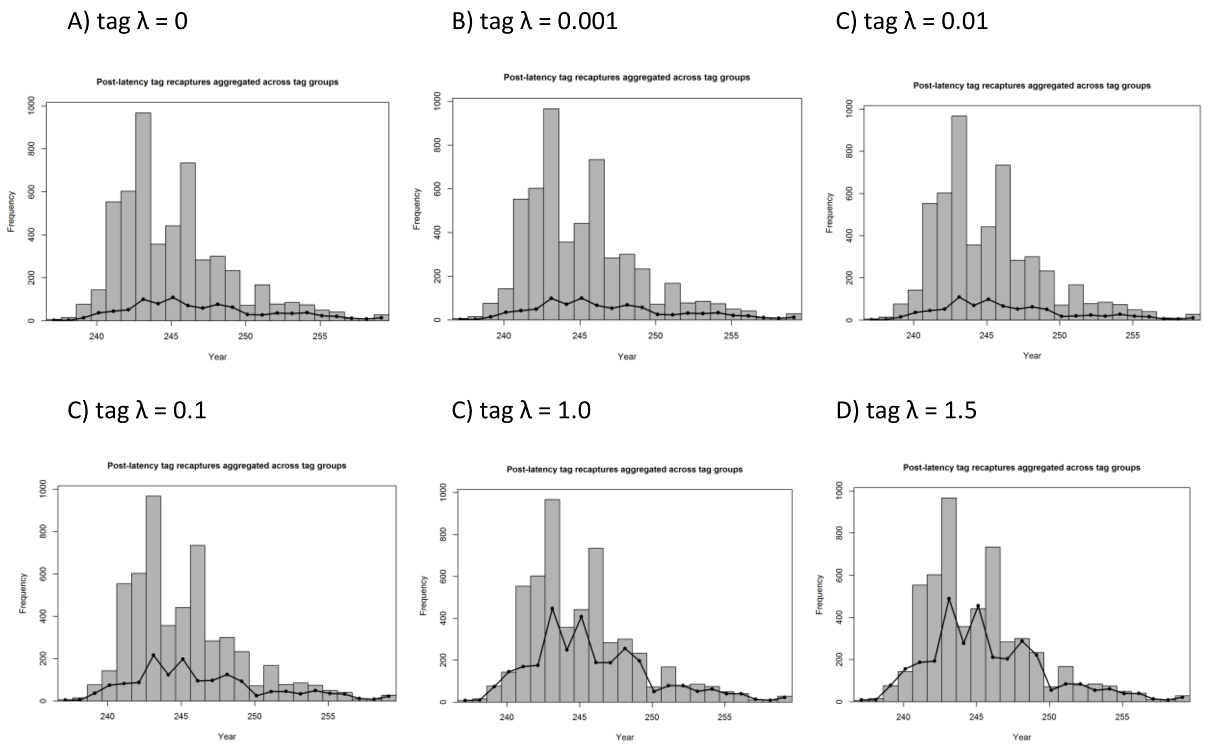
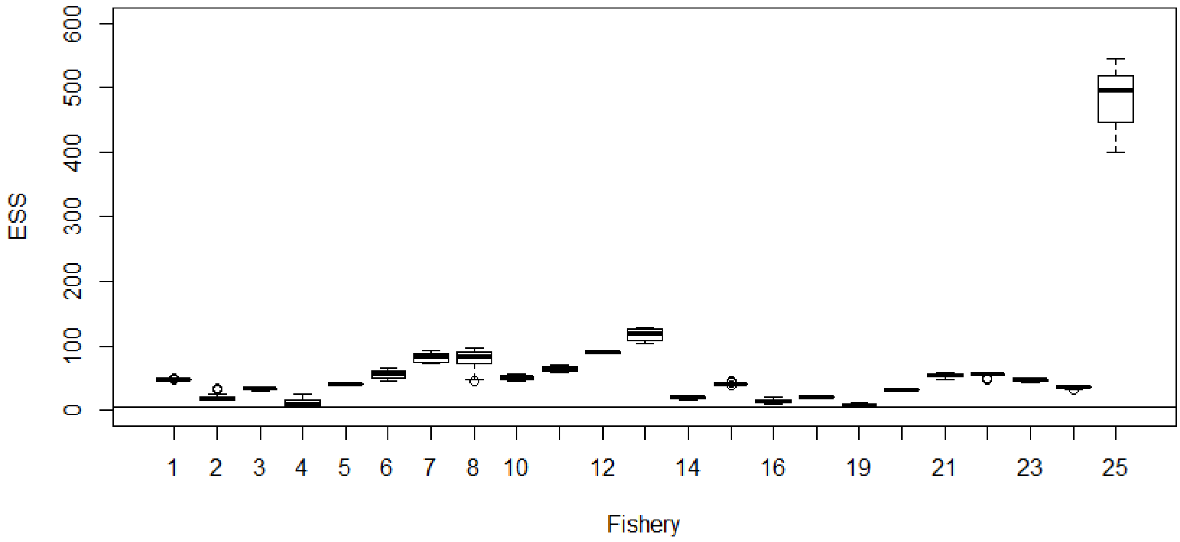



**Figure 34. Time series of aggregate tag recapture predictions and observations. Panels contrast tag weighting lambda options while other assumptions correspond to the OM-ref-SA-analogue.**



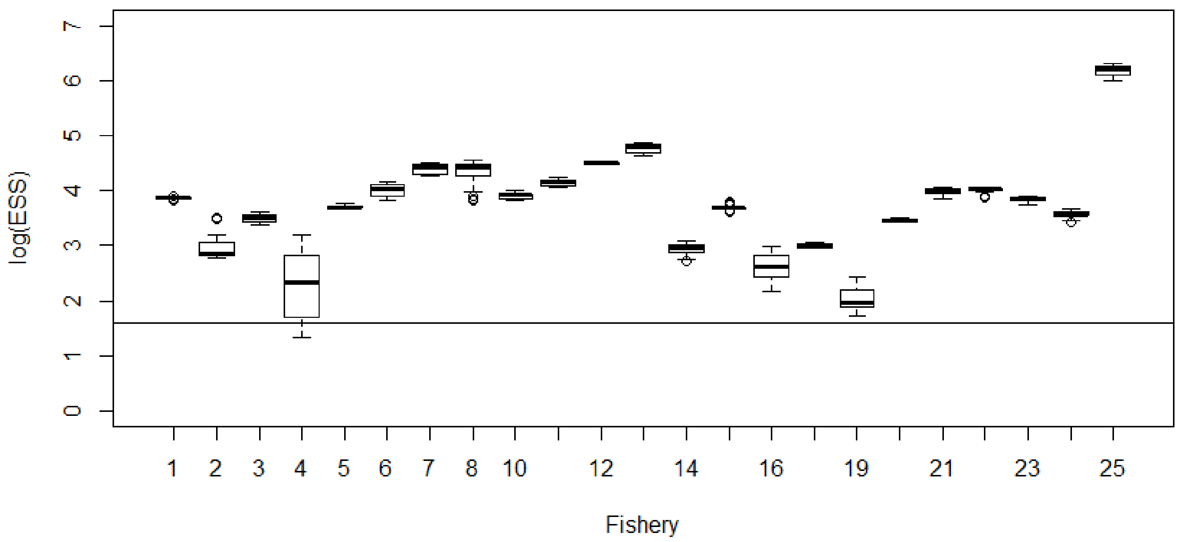




**Figure 35. OM-rob-tagWt summary of the size composition data quality of fit between model predictions and observations by fishery. Each boxplot represents the distribution of 36 mean (post-fit) Effective Sample Sizes (arithmetic mean and log(mean) over all years). Reference lines represent the assumed sample sizes in the model (5 for all fisheries).**

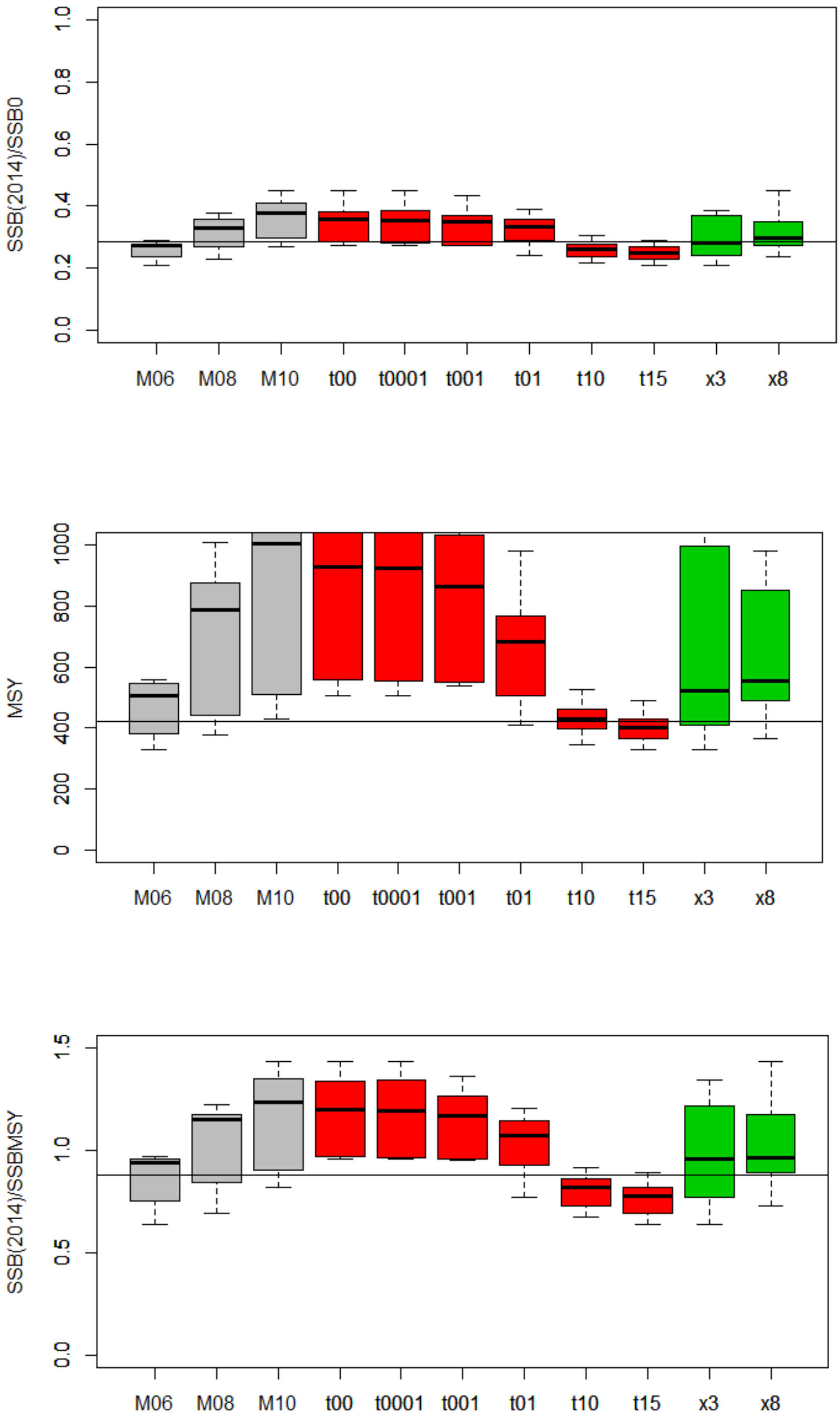



**Figure 36. OM-rob-tagWt stock status summary plots indicating how the uncertainty in the model ensemble is partitioned according to assumptions. Horizontal line is the value from SA-base.**

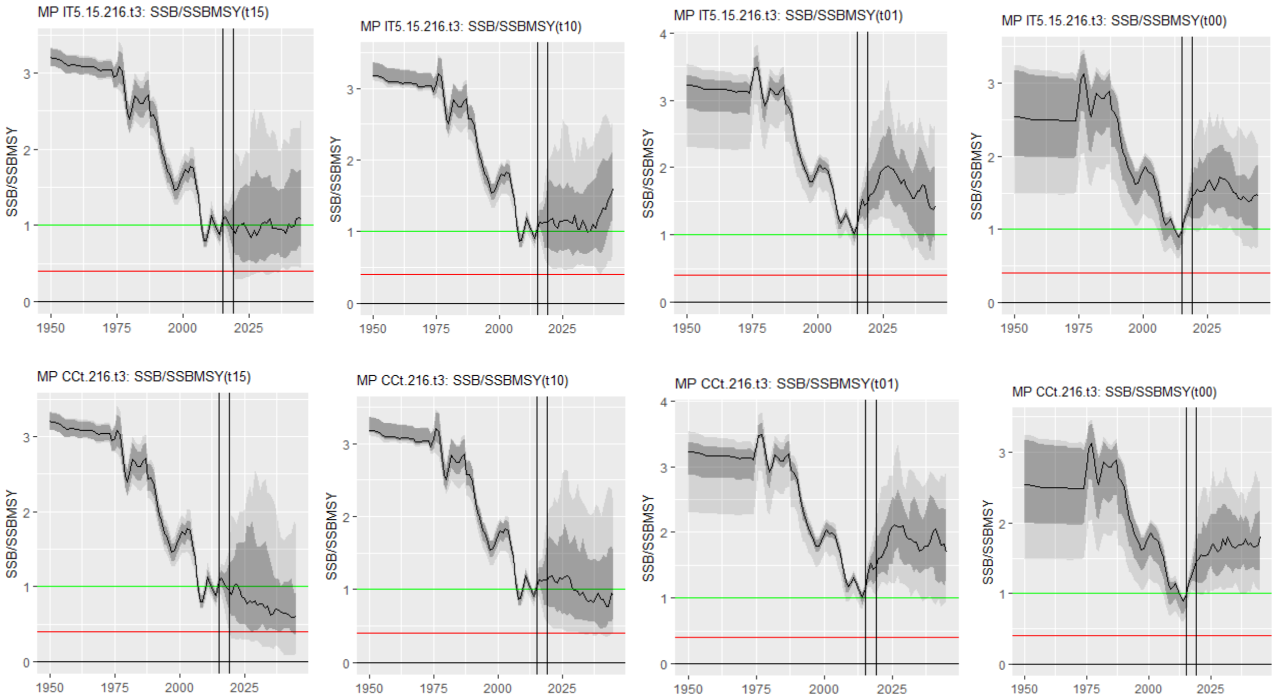



**Figure 37. Comparison of MP evaluations from a subset of OM-rob-tagWt. The first column represents the proposed robustness OM tag weighting (t15, tag λ = 1.5); the remaining columns represent the tag weighting in the reference set (t10, t01, t00, λ = 1.0, 0.1, 0.0 respectively). Top row is a feedback-based MP (Index Target), bottom row is constant catch, both were tuned to OM-ref (216 realizations).**

## 5.3 Extend Standardized CPUE time series back to 1972

WPM (2016) identified an additional robustness scenarios which requested exploring the inclusion of standardized CPUE back to 1972. This was presumably targeted at bigeye, because yellowfin SA-ref and OM-ref already use the CPUE data starting from 1972.

# 6 Discussion

This WP summarizes progress on the revised yellowfin OM, and represents the last chance for the full WPTT and WPM to provide feedback before results are presented to the TCMP for consideration in 2018. The current support project is scheduled for completion in Dec 2018, such that there is some misalignment in the two timelines. We will need to present MSE results to the TCMP in 2018 on the basis of a substantially revised bigeye OM which will not have been reviewed by the WPTT/WPM and SC. The review process will depend on the informal MSE meetings. There will be a final opportunity within the project for review of the bigeye and yellowfin OMs by the WPTT and WPM in 2018. However, the current contract will finish before subsequent revisions can be made, and before the TCMP 2019, which is identified as the first date at which the Commission is likely to contemplate selecting among MPs.

The yellowfin stock assessment and OM represent an ambitious attempt to link together diverse data, most of which have non-trivial problems of interpretation, and structural assumptions which represent numerous compromises and simplifications. The following discussion takes the pragmatic view that formulation problems will have to be resolved within the existing modelling framework, though problems identified should be revisited in future studies (outside of the current MSE timeline), perhaps with new analyses, different modelling frameworks and/or the collection of additional data.

Within the context of the uncertainties that have been explored in the OM to date, it is clear that the weighting of the tagging data is highly influential. There are a number of points to consider about the yellowfin tagging data in the assessment / OM conditioning:

i) Langley (2015, 2016 and references therein) describe strong evidence for incomplete mixing between tagged and untagged fish. This is a fundamental underpinning assumption for tag-based abundance, movement and mortality estimators. Kolody and Hoyle (2015) use simulations to show that poor tag mixing can be expected to cause substantial biases in these estimators, but the magnitude (and even direction) may not be easy to predict.

ii) The tagging data strongly support natural mortality levels that are considered to be low for yellowfin (e.g. Figure 12). Langley (2015, 2016) also observed this result, but adopted the highest M assumption for the base case assessment. This is a recognition that there are some internal contradiction in the perceived value of information in the tagging data.

iii) Down-weighting or removing the tags leads to a number of very optimistic OM scenarios, in which MSY is estimated to be considerably higher than the highest ever catches and the SA-base estimates (absurdly high in a few cases). The high MSY outcomes tend to explain an increasing degree of CPUE decline via a downward trend in recruitment deviations. This is a fairly common occurrence in assessment models that is not consistent with the stationary production dynamics paradigm that most single species assessments adhere to. However, it is not an impossible outcome, and a question of degree. It is not easy to defined a boundary between plausible and implausible.

iv) We would not expect the tagging data to be very informative about long term recruitment trends, since there were only 3 years of releases, a decade ago. The tagging data influence is probably indirect, providing an anchor on absolute abundance that forces the model into a particular parameter space.

It is desirable to make use of the RTTP-IO tagging data, because it was an expensive fisheries-independent research programme, with many successful outcomes. However, if the model is not structured adequately to accommodate the tag dynamics, their use might be counter-productive. Higher weights to the tagging data tend to produce models that are more attractive in the usual "everything-ought-to-be-stationary" assessment paradigm, and most fisheries management tools are geared toward this paradigm. We would argue from the above points that there are good reasons to not weight the tags very highly in and of themselves. If the tags are to be given a high weighting, the argument seems to be more about identifying production dynamics that are consistent with preconceived expectations, rather than considerations about information content and quality of fit to the various data sources. This is an explicit admission that some very influential subjective decisions need to be made, despite the desire for statistical objectivity.

For the MSE to progress, the question is how much further effort should be spent trying to understand the complicated model interactions, and the different ways in which non-stationary recruitment might arise, or alternative ways in which stationary recruitment might be imposed. Other avenues remain largely unexplored. Temporal changes in selectivity, or biases in size composition sampling might artificially introduce a recruitment trend. Spatial variability in selectivity will also affect the influence of the longline CPUE indices. While the OM-SA-analogue model appears to have attractive stationary recruitment dynamics estimates (Figure 16), it is perhaps notable that it also has non-stationary patterns in the spatial deviations of recruitment (Figure 38). It is not clear that this is any more or less likely than stationary recruitment abundance trends, and could also easily be an artefact of CPUE standardization problems or selectivity variability. The effects of unbalanced movement among regions add an additional degree of freedom for the model to fit abundance trends, which has never been considered in detail.

In addition to the issues identified above, we welcome feedback on any other aspect of the OM formulation or candidate MPs. It should be recognized that a number of subjective decisions need to be made is an MSE process, and the earlier that the broader views of the IOTC technical working parties can be accommodated, the greater confidence that all can have in the process going forward. Ideally, MSE in an RFMO context should be undertaken with the active engagement of many parties, including at the technical level, to represent the broad scientific experience within the working parties. Toward that end, we would continue to encourage other member scientists to download the source code, and scrutinize OM assumptions, performance characteristics and MP formulations, and present alternative views where appropriate.

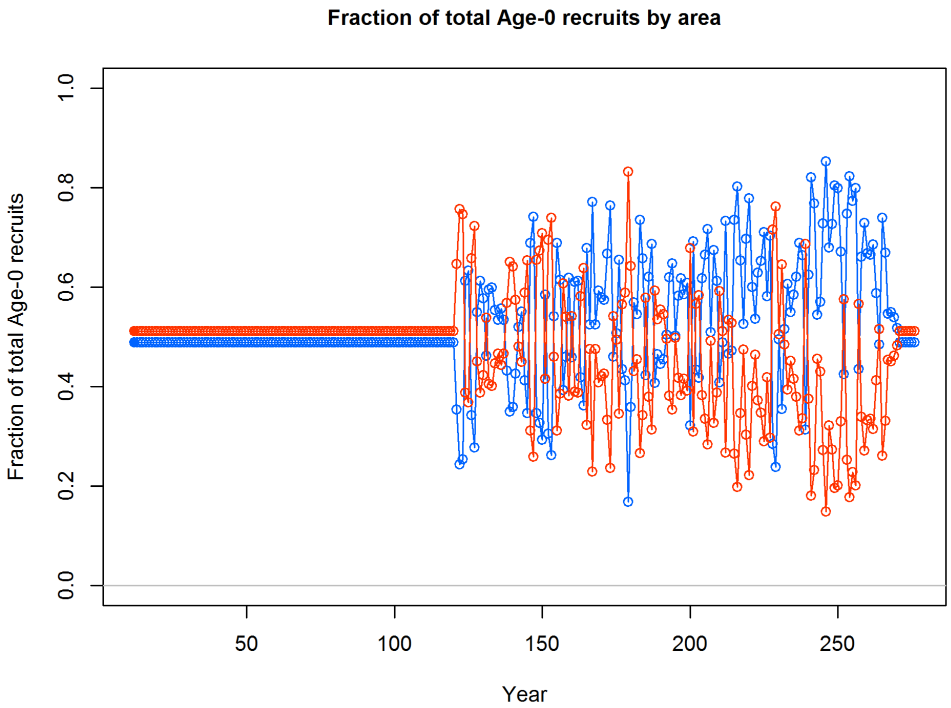



**Figure 38. OM-SA-analogue spatial recruitment deviations.**

# 7 References

Hoyle, SD, Kim, DN, Lee, SI, Matsumoto,T, Satoh, K, Yeh, Y-M. 2016. Collaborative study of tropical tuna CPUE from multiple Indian Ocean longline fleets in 2016. IOTC–2016–WPM07–11

Kolody, D, Jumppanen, P. 2016. IOTC Yellowfin and Bigeye Tuna Management Strategy Evaluation: Phase 1 Technical Support Project Final Report. Indian Ocean Tuna Commission Working Paper IOTC-2016-WPTT18-32.

Kolody, D., Hoyle, S. 2015. Evaluation of tag mixing assumptions in western Pacific Ocean skipjack tuna stock assessment models. Fisheries Research 163 (2015) 127–140.

Langley, A. 2015. Stock assessment of yellowfin tuna in the Indian Ocean using stock synthesis. IOTC-2015-WPTT17-30.

Langley, A. 2016. An update of the 2015 Indian Ocean Yellowfin Tuna stock assessment for 2016. IOTC-2016-WPTT18-27.

IOTC 2017. Report of the 21st Session of the Indian Ocean Tuna Commission Yogyakarta, Indonesia, 22–26 May 2017. IOTC–2017–S21–R[E]

MSE 2017. Report of the 6th Workshop on MSE of IOTC WPM Scientists, Bangkok, Thailand, 1-4 April 2017.

TCMP 2017. Chair report of the 1st IOTC Technical Committee on Management Procedures. Yogyakarta, Indonesia, 20 May 2017. IOTC–2017–TCMP01–R[E]

WPM 2016. Report of the 7th Session of the IOTC Working Party on Methods. Victoria, Seychelles, 11–13 November 2016. IOTC–2016–WPM07–R[E]

WPTT 2016. Report of the 18th Session of the IOTC Working Party on Tropical Tunas. Seychelles, 5-10 November 2016. IOTC–2016–WPTT18–R[E]

# Attachment 1: Compilation of yellowfin and bigeye MSE development requests from the IOTC.

**2016 Working Party on Methods** - the following extracts highlight the MSE changes requested for bigeye and yellowfin tuna MSE in the next iteration, with context provided in square braces. The 2017 WPM requests subsume the 2017 WPTT requests.

[Clarification for bigeye and yellowfin reference case OM ensembles (OM-ref)]

- Weighting factors (lambdas) for both tag likelihoods would need to be modified in the uncertainty grid (i.e. the abundance estimator and the movement estimator).

[Balanced grid for bigeye and yellowfin Reference case OM ensembles (OM-ref)]

Table 2. Summary of the grids agreed by the WPM for YFT and BET OM ensembles.

| YFT OM ensemble | 3 x 3 x 3 x 2 x 2 x 2 options = 216 configurations | | | |
|---|---|---|---|---|
| Grid Dimension | Levels | | | Notes |
| Natural mortality | 0.6 | 0.8 | 1.0 | Multiplier relative to 2015 reference case assessment |
| Steepness | 0.7 | 0.8 | 0.9 | Beverton-Holt relationship |
| Tag lambdas | 0.0 | 0.1 | 1.0 | weighting factor for both components of tag likelihood |
| Tag Mixing period | 3 | 8 | | Quarters |
| Relative abundance (CPUE) bias | 0.0 | 1.0 | | Historical and future catchability trend (percent per annum compounded) |
| CPUE analysis for tropical tuna targeting | CLU | HBF | | CLU = cluster analysis (no HBF)<br>HBF = hooks between floats |
| **BET OM ensemble** | **3 x 3 x 3 x 1 x 2 x 2 options = 108 configurations** | | | |
| Grid Dimension | Levels | | | Notes |
| Natural mortality | 0.6 | 0.8 | 1.0 | Multiplier for M(a=4) relative to 2013 reference case |
| Steepness | 0.7 | 0.8 | 0.9 | Beverton-Holt relationship |
| Tag lambdas | 0.0 | 0.1 | 1.0 | weighting factor for both components of tag likelihood |
| Tag Mixing period | 4 | | | Quarters |
| Relative abundance (CPUE) bias | 0.0 | 1.0 | | Historical and future catchability trend (percent per annum compounded) |
| CPUE analysis for tropical tuna targeting | CLU | HBF | | CLU = cluster analysis (no HBF)<br>HBF = hooks between floats |

[Bigeye and yellowfin robustness OM proposals for exploration]

- Testing for the effect of changing selectivity over time (e.g. to check for a shift towards younger selected ages over time).
- A model structure that would permit accommodating the CPUE information prior to 1979 should be considered.
- Tag λ=1.5. This was proposed on the basis that the tags might be more informative and reliable than the other data, despite the fact that the tags are known to violate mixing assumptions and are expected cause biases.

**2017 Working Party on Methods informal working group on MSE** - the following extracts highlight the MSE changes requested for bigeye and yellowfin tuna MSE in the next iteration.

> The priorities for Phase 2 were discussed, which included (1) update of the OMs and Reference Set, (2) spatial dynamics, (3) MPs and (4) tuning.
>
> With regards to the OMs, they will be updated in relation to new assessments (particularly revised CPUE indices, and new spatial structure for BET). Details on the OM Reference set defined by 2016 WPM for BET and YFT were reviewed and Robustness scenarios were clarified by the group, including the selectivity changes that would be explored by admitting temporal variability in the OM conditioning. A specific proposal for pre-1979 CPUE series is still required.

> At this time, there are a small number of robustness scenarios proposed for albacore, yellowfin and bigeye, which have not been tested. These scenarios were proposed in the spirit of curiousity, with neither strong support nor opposition, so the group did not consider it a priority to communicate these scenarios to the TCMP in this round.

**2017 Technical Committee on Management Procedures** identified default MP specifications, including tuning criteria, to be reported against in 2018.

> 47. The TCMP **NOTED** the default Yellowfin tuna MP assumptions, including 3 year TAC setting, 15% TAC change constraint, and tuning objectives proposed for phase 2:
> a) 50% probability of rebuilding to B(target) by 2024 (interpretation from Resolution 16/01[3])
> b) 50% probability B>B(target) from 2019-2039 (interpretation from Resolution 15/10)
>
> 48. The TCMP **NOTED** the default bigeye tuna MP assumptions, including 3 year TAC setting, 15% TAC change constraint, and tuning objectives proposed for phase 2:
> a) 50% probability B>B(target) from 2019-2039 (interpretation from Resolution 15/10)
> b) 75% probability in Kobe green zone from 2019-2039 (interpretation from Resolution 15/10)

## **2017 Commission meeting** - MSE workplan.

**Schedule of work for the development of management procedures for key species in the IOTC Area.** *A more detailed explanation of the roles of the Working Parties (WPs), Scientific Committee (SC), Technical Committee on Management Procedures (TCMP) and the Commission are provided at Annex 1.*

| Year | Albacore | Skipjack | Yellowfin | Bigeye | Swordfish |
|---|---|---|---|---|---|
| 2017 | **WPs/SC**: Undertake MSE and provide advice on the performance of candidate MPs. Identify issues which might need specific guidance from the Commission, including how to interpret objectives, timelines and acceptable levels of risk. | **WPs/SC**: Apply HCR using results from 2017 stock assessment to calculate the total annual catch limit. Secretariat to advise CPCs of catch limit. | **WPs/SC**: Undertake MSE and provide advice on the performance of candidate MPs. Identify issues which might need specific guidance from the Commission, including how to interpret objectives, timelines and acceptable levels of risk. | **WPs/SC**: Undertake MSE and provide advice on the performance of candidate MPs. Identify issues which might need specific guidance from the Commission, including how to interpret objectives, timelines and acceptable levels of risk. | **WPs/SC**: Develop framework and seek funding for MSE. Advise TCMP and Commission on progress |
| 2018 | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to the Commission on any outstanding issues resulting from the application of the HCR if required | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | |
| | **Commission**: Consider work and advice from subsidiary bodies and provide direction to the WPs/SC on the need to undertake further MSE of candidate or alternative MPs | **Commission**: Provide direction to the WPs/SC on the need to refine the HCR and/or MSE | **Commission**: Consider work and advice from subsidiary bodies and provide direction to the WPs/SC on the need to undertake further MSE of candidate or alternative MPs | **Commission**: Consider work and advice from subsidiary bodies and provide direction to the WPs/SC on the need to undertake further MSE of candidate or alternative MPs | |
| | **WPs/SC**: Consider recommendations from the Commission and undertake MSE to provide advice on the performance of candidate MPs | **WPs/SC**: Consider recommendations from the Commission and further refine the HCR through MSE as directed | **WPs/SC**: Consider recommendations from the Commission and undertake MSE to provide advice on the performance of candidate MPs | **WPs/SC**: Consider recommendations from the Commission and undertake MSE to provide advice on the performance of candidate MPs | **WPs/SC**: Develop initial operating model and undertake MSE to provide initial advice on the performance of candidate MPs. Identify issues which might need specific guidance from the Commission, including how to interpret objectives, timelines and acceptable levels of risk. |
| 2019 | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to Commission on any outstanding issues resulting from the application of the HCR if required | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives |
| 2019 | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs | **Commission**: Consider work and advice from subsidiary bodies and review Resolution 16/02. | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs | **Commission**: Consider work and advice from subsidiary bodies and provide direction to the WPs/SC on the need to undertake further MSE of candidate or alternative MPs |
| | **WPs/SC**: Undertake MSE and provide advice on the performance of candidate MPs | | **WPs/SC**: Undertake MSE and provide advice on the performance of candidate MPs | **WPs/SC**: Undertake MSE and provide advice on the performance of candidate MPs | **WPs/SC**: Consider recommendations from the Commission and undertake MSE to provide advice on the performance of candidate MPs |
| 2020 | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives | **TCMP**: Provide advice to the Commission on elements of candidate MPs that require a decision by the Commission, including the performance of candidate MPs against Commission objectives |
| | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs until an MP is adopted. | | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs until an MP is adopted. | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs until an MP is adopted. | **Commission**: Consider work and advice from subsidiary bodies. Decision and adoption of an MP or provide direction to the WPs/SC on the need for further MSE of candidate or alternative MPs until an MP is adopted. |

CONTACT US

**t** 1300 363 400
+61 3 9545 2176
**e** csiroenquiries@csiro.au
**w** www.csiro.au

AT CSIRO, WE DO THE
EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE.
WE INNOVATE.

FOR FURTHER INFORMATION

**Oceans & Atmosphere**
Dale Kolody
**t** +61 6 6232 5121
**e** Dale.Kolody@csiro.au
**w** www.csiro.au