OCEANS AND ATMOSPHERE
www.csiro.au

Update on IOTC Bigeye Tuna Management Procedure Evaluation March 2018

Dale Kolody and Paavo Jumppanen

Prepared for the Indian Ocean Tuna Commission Working Party on Methods informal MSE Meeting, Lisbon, Mar 2018

Copyright

Important disclaimer

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 bigeye (BET) tuna. The BET MSE closely follows the approach developed for YFT (described in a companion paper), but due to time constraints and the prevailing perceptions of better stock status, no progress was made on BET for phase 2 prior to the 2017 IOTC WPTT/WPM meetings. Substantial restructuring of the BET Operating Models was recommended in the 2016 IOTC WPTT/WPM review, to recognize the substantial revision to the BET assessment, including new CPUE analyses, and spatial disaggregation to facilitate a more appropriate inclusion of the tagging data. The corresponding YFT document provides additional background information that is not repeated here.

Progress on phase 2 BET MSE began with a "mechanical" update of the reference case Operating Model (OM) to address the 2016 IOTC WPTT/WPM requests. OMrefB18.1 is composed of an ensemble of 108 stock assessment models, conditioned in relation to the 2016 stock assessment, and representing uncertainty in 5 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 CPUE standardization method
- 2 X CPUE catchability trend

The BET SS assessment model was configured with independent CPUE series in the temperate region for each quarter (in an attempt to compensate for a limitation in the representation of seasonal migration). In the interest of simplifying the OM implementation, we compared an alternative grid (OMrefB18.0) in which the temperate seasonal CPUE series were replaced by a single aggregate (consisting of the 4 independently normalized quarterly series), and found that the stock status differences relative to OMref-B18.1 were negligible across the range of grid assumptions, such that the simplified CPUE assumption was adopted for all simulations presented here.

While OMrefB18.0 is broadly consistent with the assessment, it is more optimistic in terms of productivity and SB/SB(MSY) estimates. As with the YFT OM, the quality of fit to the data diagnostics are not as informative as we might hope for evaluating model plausibility (though the concerns about non-stationary recruitment from YFT were not evident for BET). In the interest of consistency, we proposed an alternative OM based on a new grid which expands OMrefB18.0 in two dimensions (CPUE CV assumptions and CL size composition effective sample sizes), for a total grid size of 432 models. This grid is sampled with replacement, to generate OMrefB18.3, that is simultaneously consistent with the central tendency of the assessment with respect to B(2015)/B(MSY) and MSY (assumed correlation of 0). The sampling CV (0.13) is consistent with the uncertainty presented for the (6 model) BET assessment ensemble, and (coincidentally) very

similar to the YFT sampling approach (in which the CV was arbitrarily chosen to be 3X higher than the YFT reference case assessment).

Projection assumptions for OMs included:

- Initial states (with added error) and most parameters defined by the SS specifications
- temporal variability in selectivity for all fleets
- CPUE CV = 0.2
- annual recruitment CV = 0.6, autocorrelation = 0.5
- first TAC implemented in 2019; bridging catches 2016:2018 = 93Kt (2015 level)
- catch implementation error CV = 0.1

Results of 3 candidate MPs (plus constant catch) are presented for two example BET tuning objectives similar to those identified by the TCMP 2017 (including 3 year TAC setting and 15% TAC change constraints):

- B1: Pr(mean(B(2019:2038)/BMSY) > 1.0) = 0.5
- B2: Pr(Kobe green zone 2019:2038) = 0.75

All of the MPs exhibited undesirable behaviour with the B1 tuning objective. Given that the current stock status has a high probability of SB > SB(MSY), the MPs must rapidly increase fishing effort in the short-term (which may not be in the interest of industry), so that there is a high probability of SB < SB(MSY) later in the time series, to reach SB = SB(MSY) on average from 2019-2038. Achieving this tuning usually results in a downward biomass trend at the end of the 20y evaluation period. In contrast, all of the MPs tend to support very stable population trajectories with the B2 tuning objective. The B1 tuning results are presumably affected by the numerical problems of the high F scenarios that were observed for YFT, while the effect was not obvious for B2 tuning (additional diagnostics will be added for evaluating this problem in the future).

The MSE was repeated for both potential reference case OMs (OM-refB18.0, and OM-refB18.3). We do not present the contrasting results from the two OMs with the intent of suggesting that MP results should be used to choose among OMs (that would be somewhat backward). However, it is a worthwhile demonstration that MP selection will depend on the OM, and we recommend OMref18.3 on the basis of consistency with the approach proposed for yellowfin.

There was not time to explore any BET robustness OM scenarios, however, priorities are proposed for discussion.

We seek guidance and/or endorsement from the MWG informal MSE working group (and BET/YFT MSE project steering committee), with respect to the proposed workplan, notably:

- BET reference case OM (for TCMP 2018 and WPTT/WPM 2018 presentations):
    - Replace 4 seasonal temperate CPUE series with a single aggregate.

  - OM to consist of sampled grid (B18.3) with central tendency of B/BMSY and MSY matching the assessment and a CV to be determined (minimum of 0.17 suggested as the approximate sum of inverse Hessian CV + 6 model MPD grid CV).

- BET Tuning Objectives:

  - Full suite of 5 from the TCMP report

- BET robustness case OMs:

  - Not to be presented to TCMP 2018.

  - Use YFT as a guide to prioritize scenarios for WPTT/WPM 2018.

2 Introduction

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. Recognizing the iterative nature of the MSE process, the workplan identifies 2019 as the earliest probable date for MP adoption.

3 Summary of Bigeye MSE Progress toward the requests from the IOTC Working Parties

MSE for bigeye (BET) and yellowfin (YFT) tunas has been pursued in parallel, with the first phase of the scientific and technical work described in Kolody and Jumppanen (2016). The second phase project commenced in Sep2017, and reported to the WPTT/WPM in Oct2017 (Kolody and Jumppanen 2017), however no new progress on BET had been made at that time, so there was no new guidance for BET development. The recent BET progress reported here is guided by the working group feedback from 2016, and the desire to maintain consistency between the YFT and BET approaches, including:

i) A "mechanical" update to the BET reference case OM in line with the feedback from the 2016 IOTC technical working parties, and presentation of common diagnostics for evaluating plausibility. This OM (OMrefB18.0) is presented in section 5 (including the new CPUE series discussed in section 4). We also defined another OM (OMrefB18.3), which uses the grid-sampling approach proposed for YFT, to come up with an OM that is more consistent with the central tendency of key inferences from the assessment (though with an arbitrarily manipulated variance).

ii) BET robustness set scenarios were proposed along the lines of those for YFT in 2016:

- tag weighting $\lambda = 1.5$
- Testing for the effect of changing selectivity over time (e.g. to check for a shift towards younger selected ages over time).

These have not yet been investigated for BET (but were found to be of negligible importance for YFT, Kolody and Jumppanen 2017). Since there was no BET progress in 2017, robustness scenarios for BET were not discussed further (though one might infer some of the same priorities as for YFT).

iii) Presentation of candidate MP results that meet the initial tuning objectives identified in TCMP (2017) are presented in section 6. The WPM (2016) requested that the "Irate" and "Brule" MPs designed for albacore be applied to YFT and BET to simplify communication if performance is similar. This latter request has not yet been fully met.

4 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 fit to the same data, subject to the same concerns about model formulation and assumption violations, etc. The scientific process has been evolving rapidly in recent years. While the objectives of the two processes are different, it would be difficult to justify the two initiatives evolving in completely different directions. Accordingly, the bigeye 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, using a configuration with calendar seasons defined as model years.
- 15 fisheries (Table 1)
- Beverton-Holt recruitment dynamics
- Parameter estimation objective function includes
    - Standardized longline CPUE (Region 1A and 1B share one series, R2 has one series, and R3 estimates seasonal catchability by splitting the fishery and CPUE by season)
    - 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 (stationary, various functional forms, parameters shared among some fleets)
    - Longline catchability - Regional scaling factors are used to scale relative density to relative abundance among regions, such that 1A, 1B, 2 share catchability. Catchabilities are estimated independently for the 4 seasonal fisheries in region 3.
    - Virgin recruitment

    - Recruitment deviations from the Beverton-Holt stock-recruit relationship, recruitment spatial partitioning among tropical regions (1 and 2) and deviations from the mean spatial distribution. (check for BET)
    - Juvenile and adult movement rates

- Unlike the most recent YFT assessment, the BET assessment and management advice was based on an equally-weighted grid of 6 models in two dimensions:
    - three levels of stock-recruit steepness (h = 0.7, 0.8, 0.9)
    - two tag weighting assumptions (λ = 1.0, 0.1)

The various models, model ensembles and individual assumptions are summarized in Table 2 and Table 3. Projection assumptions were identical to previous iterations of the BET MSE, including:

- Initial states (with added error) and most parameters defined by the SS specifications
- temporal variability in selectivity for all fleets
- CPUE CV = 0.2
- annual recruitment CV = 0.6, autocorrelation = 0.5
- first TAC implemented in 2019; bridging catches 2016:2018 = 93Kt (2015 level)
- catch implementation error CV = 0.1

Table 1. Fishery definitions in the BET 2016 assessment (note that the order does not correspond to the fishery numbering in the assessment files).

Fishery	Time period		
	2012-14	2014	2015
FL2	10,147	13,383	16,153
LL1N	10,416	7,552	5,764
LL1S	16,486	15,868	16,316
LL2	6,394	7,568	4,704
LL3	3,824	4,833	4,886
PSFS1N	1,073	888	2,013
PSFS1S	3,097	3,957	6,753
PSFS2	0	0	200
PSLS1N	5,418	6,958	7,233
PSLS1S	6,359	7,915	8,627
PSLS2	293	159	0
BB1	5,083	6,188	5,717
LINE2	6,395	9,001	8,132
OT1	1,931	2,468	2,492
OT2	3,553	4,313	4,050
Total	80,469	91,052	93,040

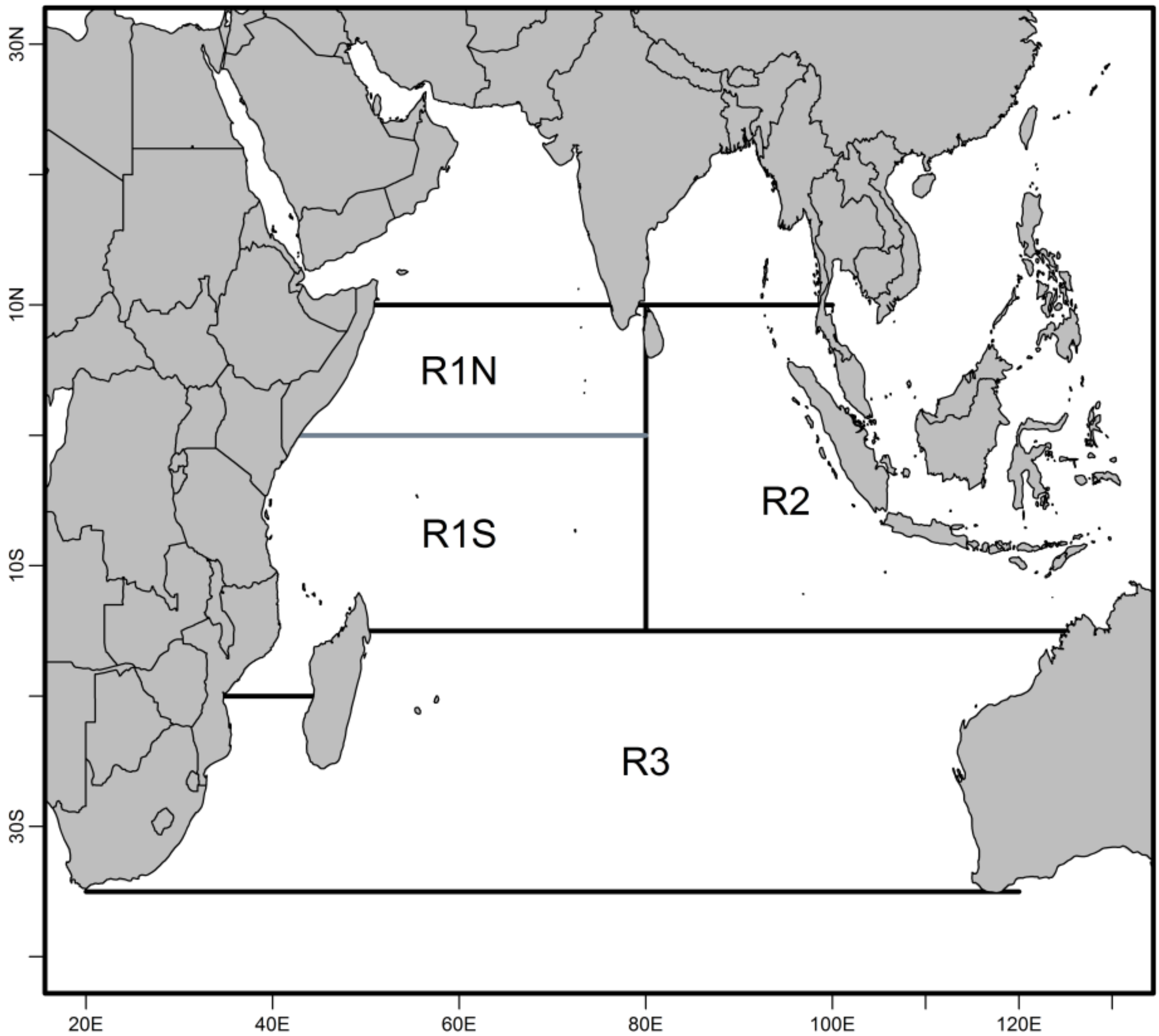


Figure 1. Spatial structure for bigeye tuna assessment and all OMs discussed in this report (figure from Langley 2016).

Table 2. Model definitions. The OMs are listed in the order discussed in the text, reflecting the sequence of development.

Model Name	Definition (assumption abbreviations are defined in Table 3)
SA-refB	Langley (2016) - balanced grid of models used for assessment advice. h70, h80, h90 M10 t10, t01 q0 iH IR3x4
OM-refB	Identical to SA-ref, except that the 4 CPUE series for region 3 (option cpR3x4) have been aggregated into a seasonally averaged mean (option IR3x1).
OM-refB18.0	Identical to OM-refB18.1 except the 4 seasonal temperate CPUE series are combined into a single equally-weighted aggregate.
OM-refB18.1	Reference case OM as proposed by the WPM and WPTT in 2016. Consists of an ensemble of 108 models, derived from the assessment, with uncertainty in 5 dimensions h70, h80, h90 M10, M08, M06 t0001, t01, t10 q0, q1 iH, iC

OM-refB18.2	A grid consisting of an ensemble of 432 models, OM-refB18.0 with additional uncertainty in the weighting assumptions for the CPUE and size composition data. h70, h80, h90 M10, M08, M06 t0001, t01, t10 q0, q1 iH, i10H, iC, i10C ess10, CLRW
OM-refB18.3	An OM derived from OM-refB18.2 with random sampling to achieve the central tendency characteristics of SB/SBMSY and MSY from the assessment, with a 13% CV on each.

Table 3. Model specification abbreviations. Bold indicates the BET assessment assumption(s). Some abbreviations may relate to additional explorations that were not completed, not reported, or pertain to YFT.

Abbreviation	Definition
	Stock-recruit function (*h* = steepness)
h70	**Beverton-Holt, *h* = 0.7**
h80	**Beverton-Holt, *h* = 0.8**
h90	**Beverton-Holt, *h* = 0.9**
Rh70	Ricker, *h* = 0.7
Rh80	Ricker, *h* = 0.8
Rh90	Ricker, *h* = 0.9
	Recruitment deviation penalty
sr4	$\sigma_R = 0.4$
sr6	**$\sigma_R = 0.6$**
sr8	$\sigma_R = 0.8$

	Future recruit failure
r55	3 years of poor recruitment (2019-2022); mean dev = -0.55, consistent with YFT assessment
	Natural mortality multiplier relative to SA-base
M10	**1.0**
M08	0.8
M06	0.6
	Tag recapture data weighting (tag composition and negative binomial)
t00	$\lambda = 0$
t0001	$\lambda = 0.001$
t001	$\lambda = 0.01$
t01	**$\lambda = 0.1$**
t10	**$\lambda = 1.0$**
t15	$\lambda = 1.5$
	Assumed longline CPUE catchability trend (compounded)
q0	**0% per annum**
q1	1% per annum
q3	3% per annum
q5	5% per annum
	Tropical CPUE standardization method (error assumption for all series)
iH	**Hooks Between Floats ($\sigma_{CPUE} = 0.2$)**
i10H	Hooks Between Floats ($\sigma_{CPUE} = 0.1$)
iC	Cluster analysis ($\sigma_{CPUE} = 0.2$)
i10C	Cluster analysis ($\sigma_{CPUE} = 0.1$)

	Tag mixing period
x3	3 quarters
x4	**4 quarters**
x8	8 quarters
	Longline selectivity
SS	**Stationary, logistic, shared among areas**
S4	LL selectivity independent among areas
NS	Temporal variability estimated in 10 year blocks
ST	Logistic selectivity trend estimated over time
Sdev	15 years of selectivity deviations estimated (XXX-XXX)
Sspl	Cubic spline function (to admit possibility of dome-shape)
	Size composition input Effective Sample Sizes (ESS)
ESS2	ESS = 2, all fisheries
ESS5	ESS = 5, all fisheries
ESS10	**ESS = 10, all fisheries**
CLRW	ESS = One iteration of re-weighting; the output ESS from a reference case assessment specification (mean over time by fishery, capped at 100)

4.1 Bigeye standardized CPUE series

Considerable collaborative work has been undertaken in recent years to improve the understanding of the DWF longline CPUE series (e.g. Hoyle et al. 2016), and to provide better relative abundance indices for the stock assessments. The 2016 bigeye assessment used the latest available studies, but adopted a single (set of area-specific) series as the best available (and this is the series that all candidate MPs use in the MSE). The WPM requested the inclusion of alternative standardized CPUE series in the reference case operating models for bigeye and yellowfin, 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-ref) or ii) cluster analyses on species composition (*std_xTW, Joint_regB2_R1_dellog_boat_allyrs, Joint_regB2_R5_dellog_boat_allyrs).* The temperate series were not changed from the SA-ref 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. The value of the cluster analyses was less clear in the tropical waters, and represents a reasonable alternative for the purposes of representing uncertainty.

Catchability trends of 0 or 1% per year (compounded annually, and projected into the future CPUE) are OM assumption options added on top of the CPUE standardization method option, 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 2. The CPUE decline using the cluster analysis does not appear to be as steep as the SA-ref (HBF). The additional variability introduced by adding the 1% CPUE catchability trend appears greater than that introduced through the alternative standardization methods.

The treatment of the temperate zone CPUE series was simplified in the OM relative to the assessment. SAB-ref partitioned the temperate zone longline fishery by quarter into 4 fleets. This was reportedly done to admit different characteristics of the fleet by season, and attempts to compensate for the fact that the calendar-seasons-as-SS-model-years configuration does not describe seasonal migration. However, since selectivity was shared among these 4 fleets, and catchability is estimated independently for each of the 4 fleets, it is not immediately obvious that 4 fleets offers any additional insight. Since this adds additional complication to how the OM outputs annual aggregate CPUE for MP testing, we re-aggregated the temperate longline fleet in the OM conditioning (with the CPUE from each season independently re-normalized). Comparison of a full grid of models with disaggregated and aggregated temperate CPUE series shows a negligible difference in terms of MSY and SSB2015/SSBMSY (Figure 3 and Figure 4), so we judged it appropriate to proceed with the simplified CPUE approach.

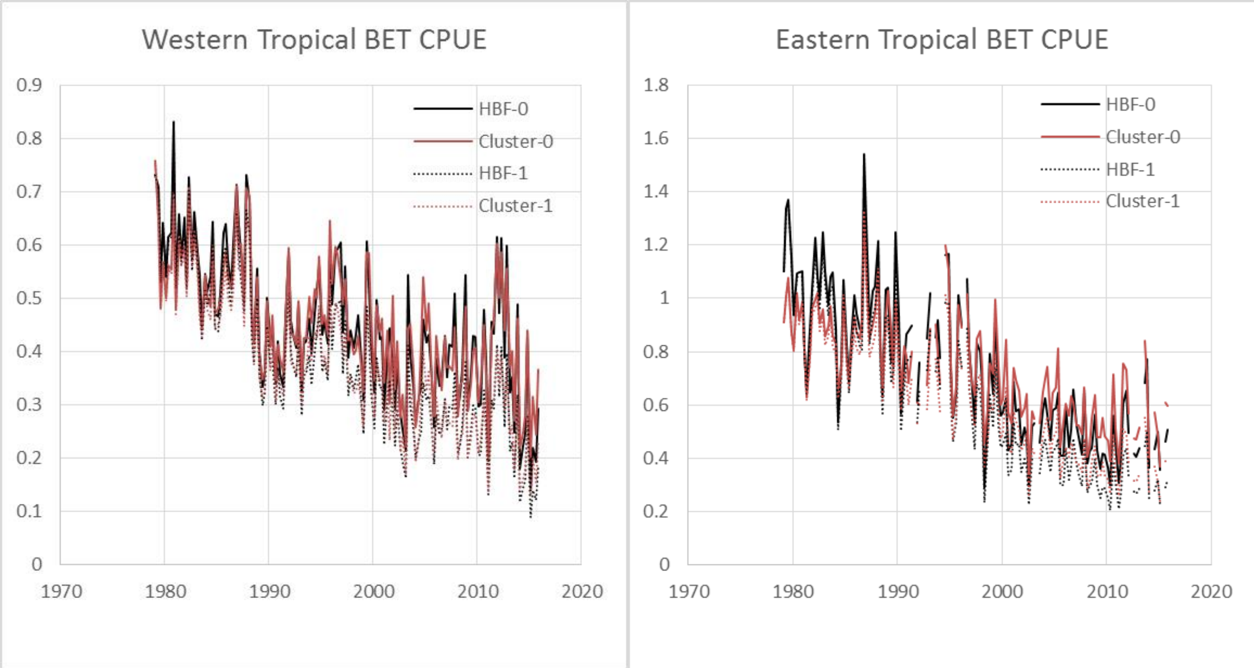
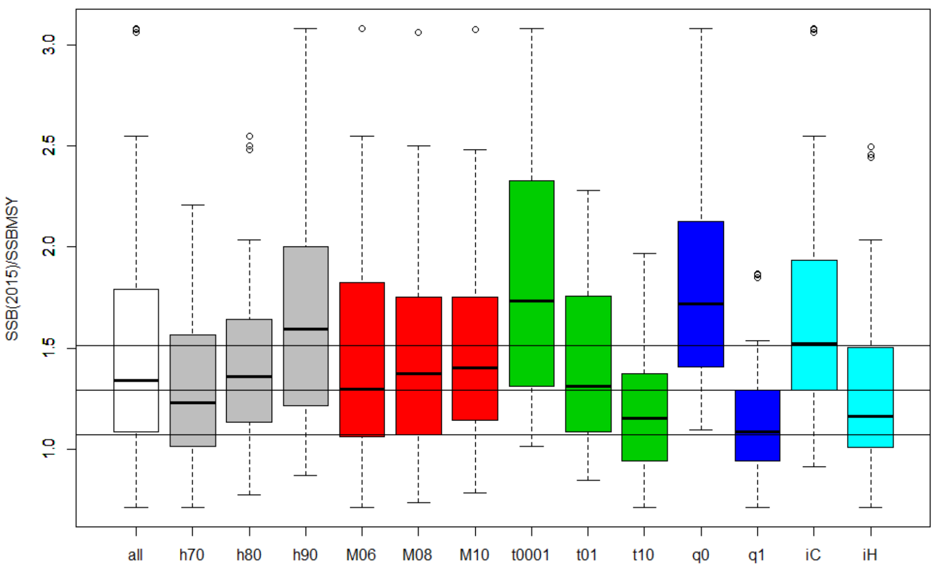


Figure 2. BET standardized longline CPUE series used in the OM grid to represent uncertainty in relative abundance. Legend labels refer to the factors used to account for species targeting shift (HBF or Cluster analyses; iH, iC in the OMs) combined with the assumed catchability trend (0 or 1% increase compounded annually; q0, q1 in the OM grid).



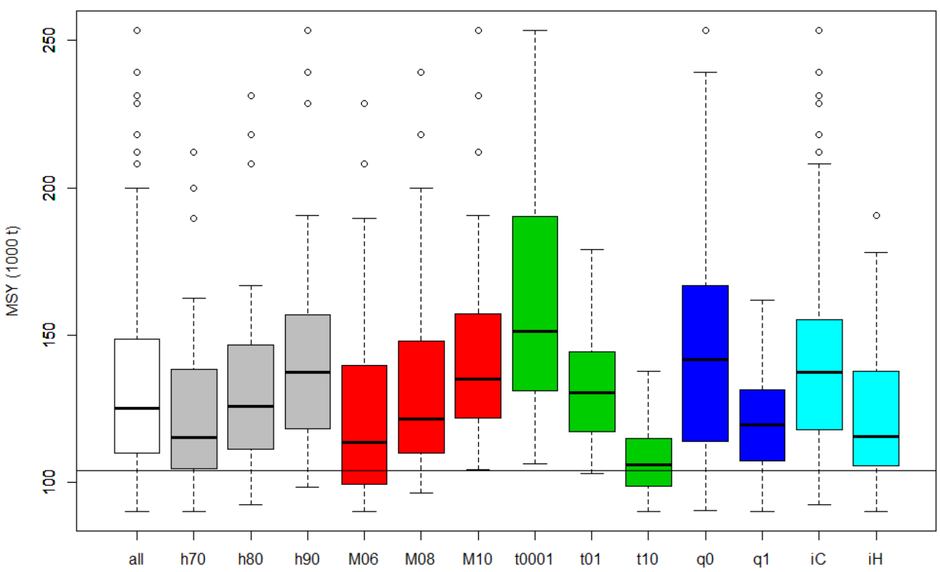


Figure 3. Key stock status indicators from OMref18.1, a grid of 108 BET models using the original CPUE configuration (temperate region partitioned into 4 independent series by season). Reference lines indicate the reported values from the stock assessment. Note that the q1 results here are based on an additive, rather than multiplicative catchability trend, but this is consistent with Figure 4 for comparison purposes.

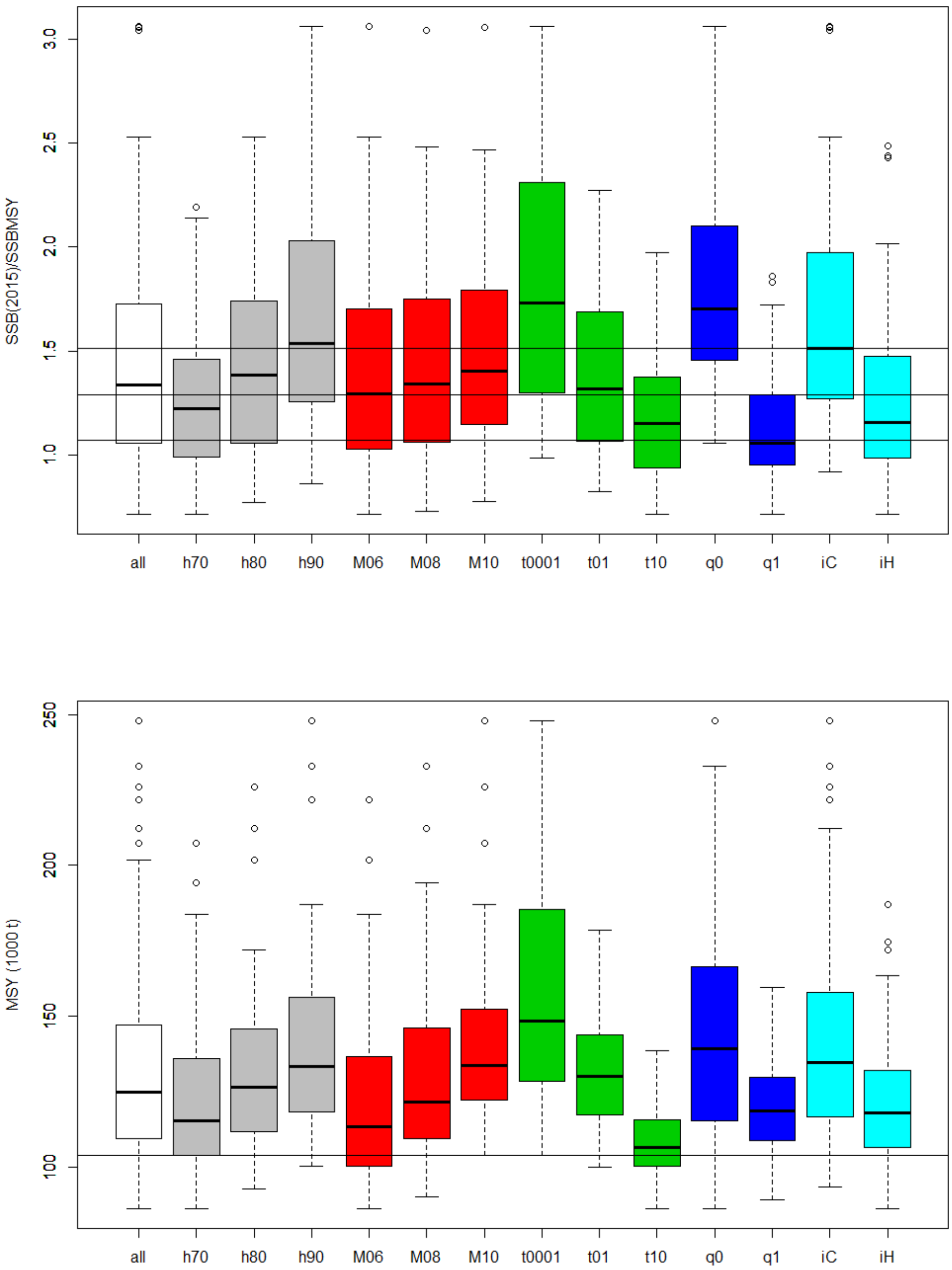


Figure 4. Key stock status indicators from OMref18.0, a grid of 108 BET models with the temperate CPUE aggregated into one series (the mean of the seasonal indices independently renormalized). Reference lines indicate the reported values from the stock assessment. Note that the q1 results here are based on an additive, rather than the intended multiplicative catchability trend, but this is consistent with Figure 3 for comparison purposes, and fixed for all subsequent results.

5 Revised BET reference set OMs

The proposed BET reference case OM was developed using the same general approach and logic used for the YFT.

5.1 OMref18.0 - The original balanced grid proposed in 2016

The reference case OM proposed by the WPTT and WPM in 2016 (OM-refB18.1) was fit and compared with OM-refB18.0, and found to be essentially indistinguishable in terms of stock status inferences (e.g. Figure 3 and Figure 4), which appears to justify the CPUE simplification carried forward in subsequent OM conditioning mentioned previously. Key diagnostic plots for OM-refB18.0 are shown in Figure 5 - Figure 9, from which we note:

- On the basis of the maximum gradient of the objective function with respect to parameters, one model indicated failed convergence, two others were probably marginal (0.1 < gradient < 1)
- The central tendency of OM-refB18.0 appears to be generally more optimistic than the assessment in terms of current stock status (SB/SBMSY) and MSY. As might be expected, lower steepness, lower M, and 1% catchability trend were generally associated with more pessimistic stock status. As with YFT, lower tag weighting and the alternate tropical CPUE series (cluster analysis) were associated with more optimistic outcomes. In contrast, the central tendency of OMrefB18.0 appears to be somewhat more pessimistic in terms of B/B0 depletion.
- The quality of fit to the CPUE series does not offer much information to distinguish among models. The models tend to fit the tropical CPUE cluster analysis series better than the HBF analysis. Higher weighting of the tags results in a slightly worse fit to the CPUE data, as might be expected. Neither q option is obviously better fit.
- The quality of the fit to the CL data does not appear to offer much power to distinguish among models, though the observed range for fishery 11 is substantial, and the fit to some fisheries is universally poor (7, 8, 12) which (presumably) reflects the limited CL data sampling.
- Unlike YFT, there is no substantial tendency for recruitment deviation trends - the OM models that deviate most from the assessment (in terms of high MSY) have the lowest trends.
- The tag downweighting option used in the BET assessment ($\lambda = 0.1$) still results in considerable tag influence in the models (e.g. compared with $\lambda = 0.001$).

The projection dynamics for OMref18.0 with a fishing moratorium (starting in 2019) show that the central tendency of the population recovers to slightly above that of the early period of the fishery, which is to be expected since the population dynamics start from a lightly exploited state (Figure 10). The current catch projections clearly predict that current catches are not only sustainable, but that the population would increase to levels considerably higher than current (Figure 10). This appears to be consistent with the assessment projections. The assessment K2MSM reports P(B(2018)>B(MSY)) = 20% and P(B(2025)>B(MSY)) = 25% (Table 4), while Figure 10 suggests 10% < P(B/BMSY) < 25% for both dates.

The default BET OM (proposed by the WPTT/WPM in 2016, without results to examine) appears to be better behaved than the default YFT OM with respect to recruitment deviation trends, but many of the same concerns remain. Notably, i) the uncertainty is high, ii) the results tend to be more optimistic than the stock status perceptions arising from the assessment, iii) the tags are very influential in anchoring the assessment in line with preconceived notions about stock status (e.g. the stock is probably near full exploitation and recent catches are probably near MSY), iv) there are reasons to doubt the tag mixing assumptions (e.g. Kolody and Hoyle 2015), and v) the quality of fit to the CPUE and size composition data are not very helpful for assigning relative plausibility of models, and vi) the choice of assumptions can skew the resulting inferences to a seemingly arbitrary degree. Because of these concerns, we have proposed an alternative reference case OM below , along the lines of that used for YFT.

Table 4. Bigeye K2MSM from the 2016 assessment.

IOTC–2016–WPTT18–R[E]

	80% (74,432t)	100% (93,040t)	120% (111,648t)	140% (130,256t)
$B_{2018} < B_{MSY}$	11	20	30	40
$F_{2018} > F_{MSY}$	2	19	40	61
$B_{2025} < B_{MSY}$	6	25	49	60
$F_{2025} > F_{MSY}$	1	19	42	53
Reference point and projection timeframe	**Alternative catch projections (relative to the catch level from 2015) and probability (%) of violating MSY-based limit reference points ($B_{lim} = 0.5\ B_{MSY}$; $F_{Lim} = 1.3\ F_{MSY}$)**			
	80% (74,432t)	**100%** (93,040t)	**120%** (111,648t)	**140%** (130,256t)
$B_{2018} < B_{LIM}$	0	0	0	0
$F_{2018} > F_{LIM}$	0	4	18	37
$B_{2025} < B_{LIM}$	0	1	12	33
$F_{2025} > F_{LIM}$	0	9	30	48

** Minor differences in the 2015 catch estimates between the Kobe II Strategy Matrix and management quantities in Table 1, are due to updates in the nominal catch published prior to the Working Party on Tropical Tunas.*

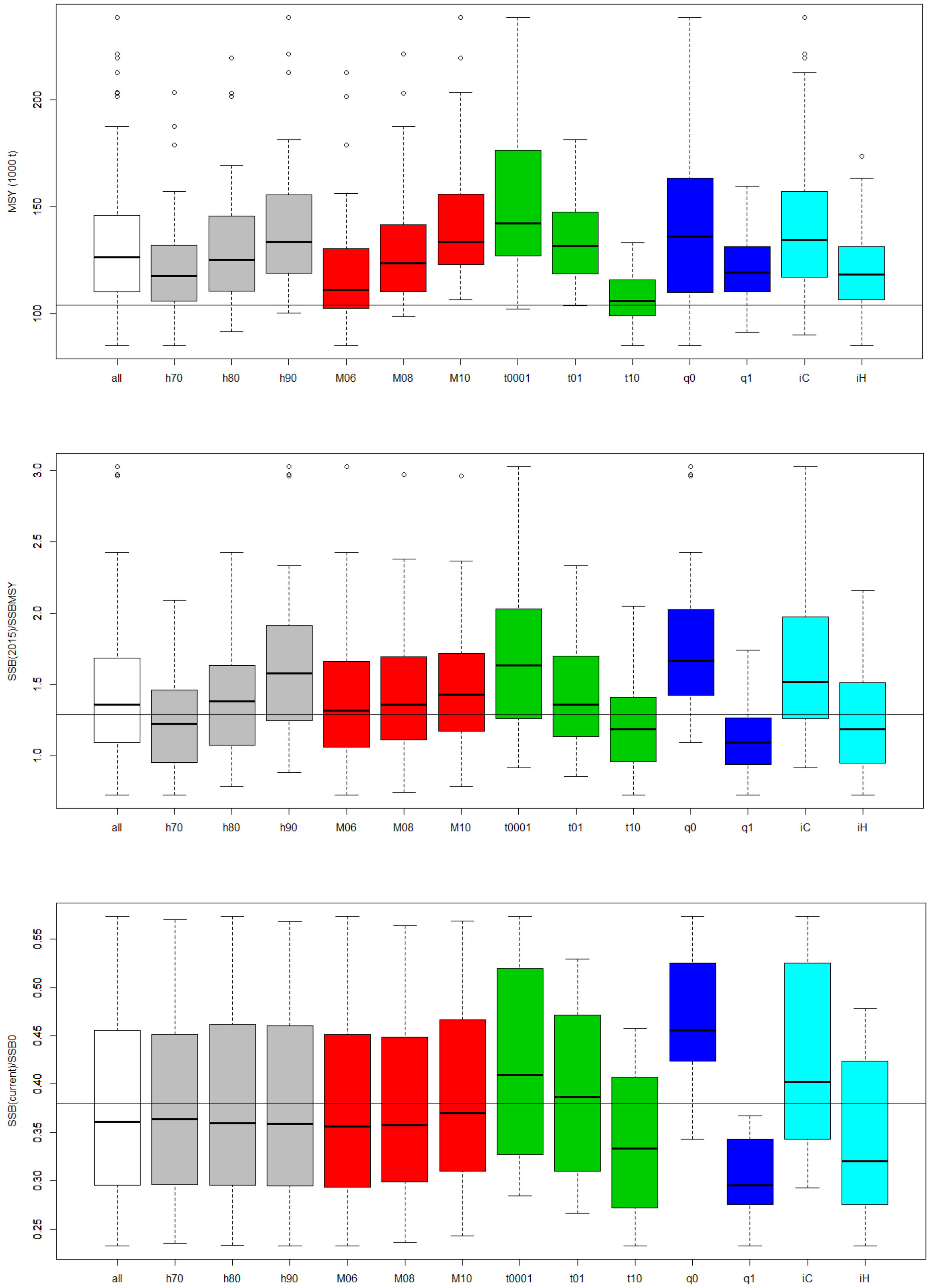


Figure 5. OM-refB18.0 MPD stock status indicators, partitioned by model assumptions. Each colour represents the full suite of 108 models, each boxplot represents all the models with the indicated option.

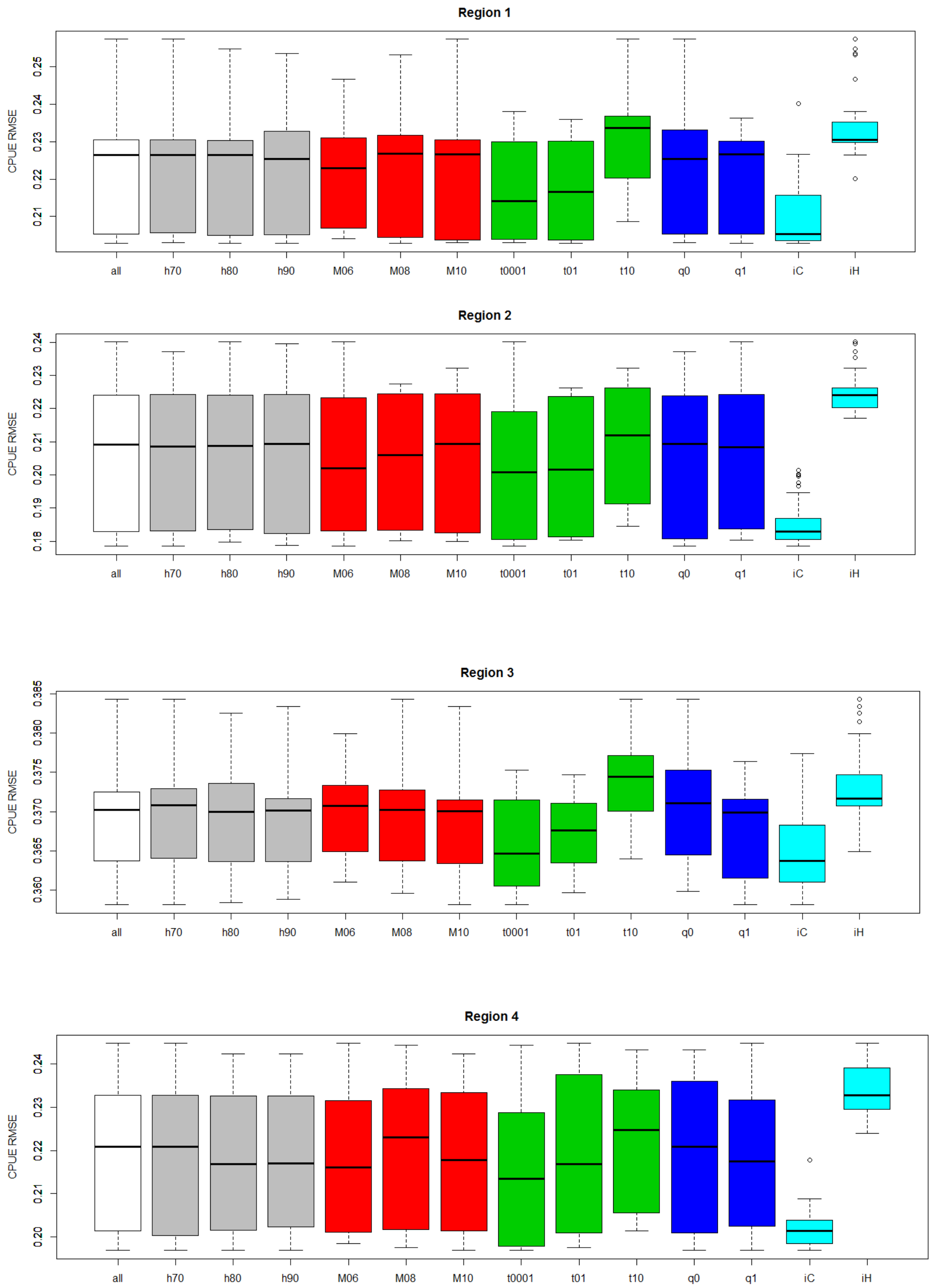


Figure 6. OM-refB18.0 MPD quality of fit to the CPUE series summaries by region, partitioned by model assumption.

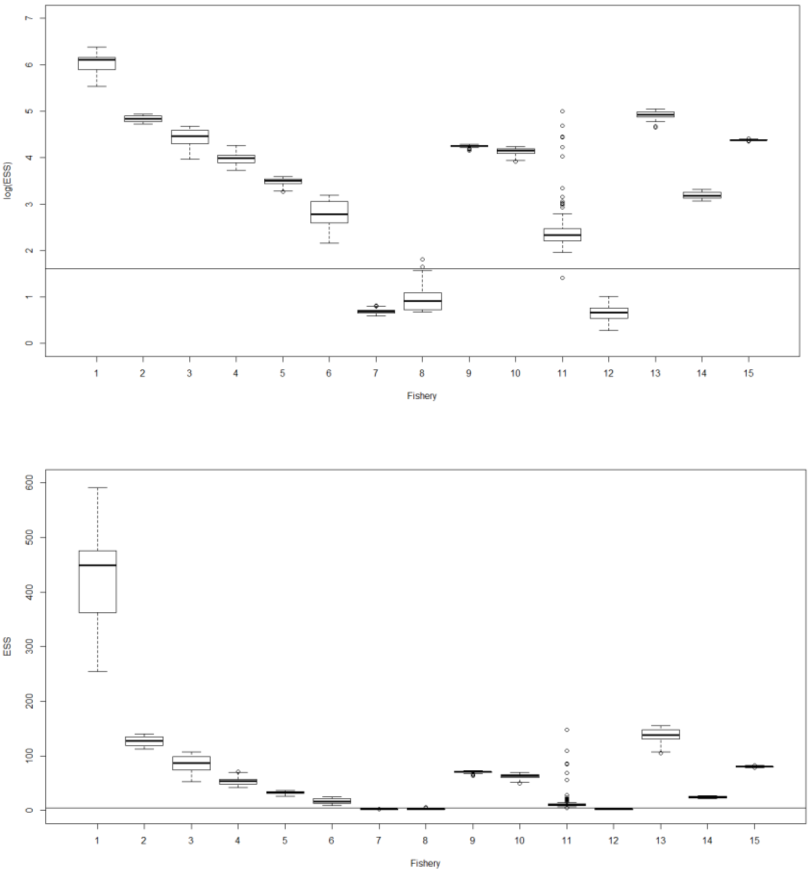


Figure 7. OMrefB18.0 quality of fit to the CPUE series (post-fit effective sample size), partitioned by fishery, but pooled over all model assumptions. The reference line (10) is the assumption in the assessment.

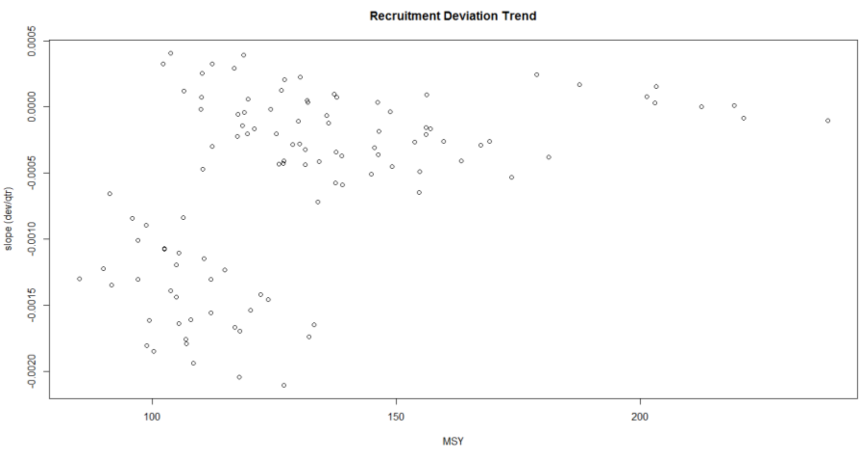


Figure 8. OMrefB18.0 relationship between the trend in recruitment deviations and MSY.

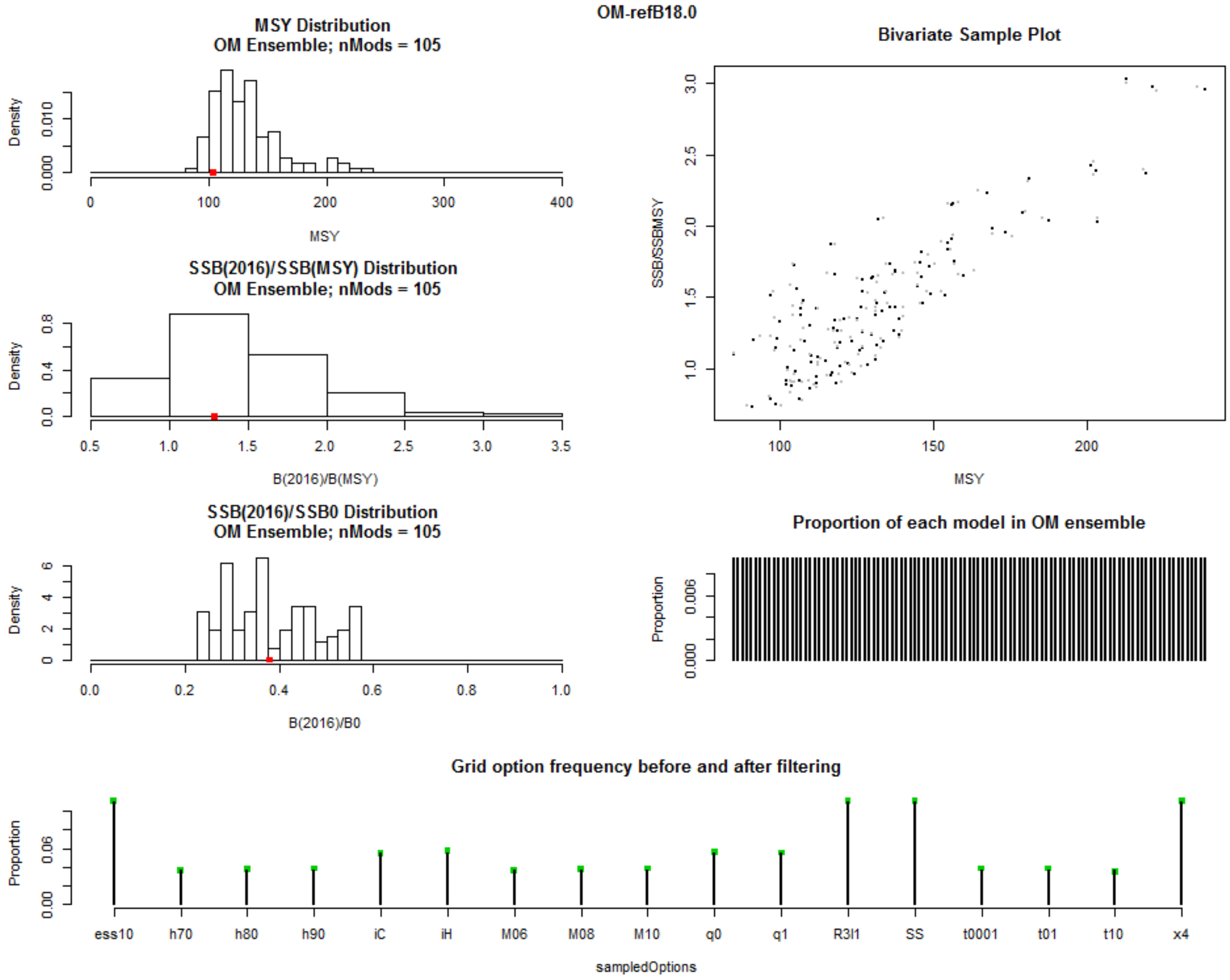


Figure 9. Characteristics of OMrefB18.0. Red circles indicate the point estimates from the 2016 assessment. The top right panel indicates the relationship between MSY and SBY/SBMSY (grey points are jitters to emphasize repeat sampling frequency). The middle right panel indicates the relative frequency of the models sampled (uniform in this case). The bottom panel indicates the relative proportion of the individual assumptions in the ensemble (green points) relative to the original grid (black lines) - essentially identical except for convergence failures.

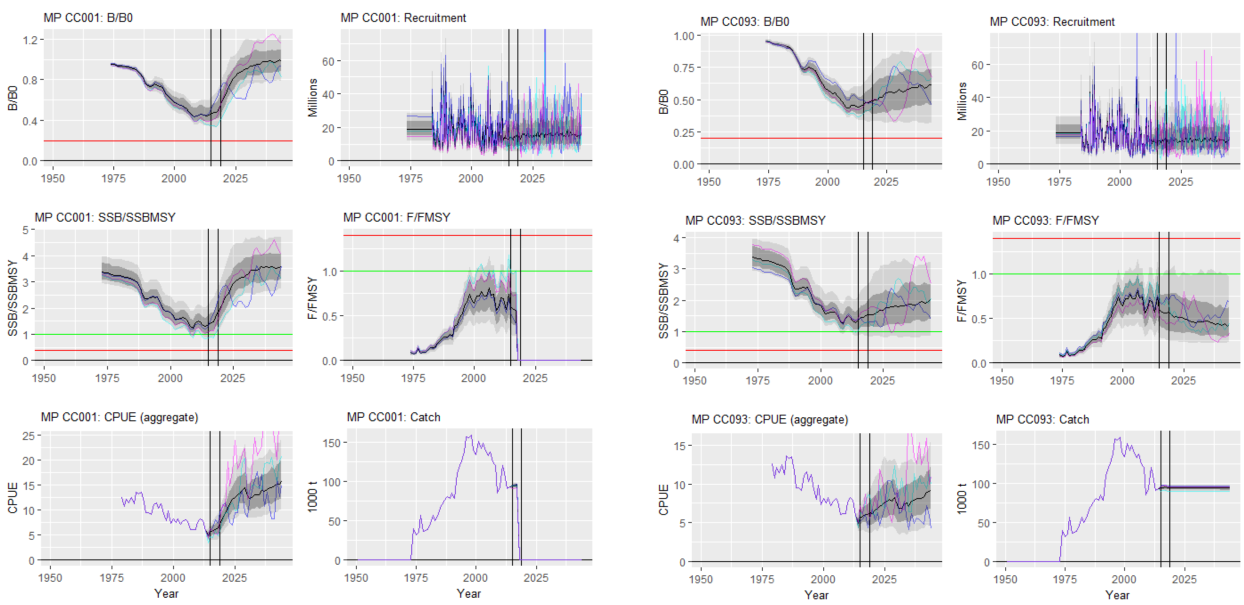


Figure 10. OMref18.0 projections with a fishing moratorium starting 2019 (left panel) and with constant current catches at 93 000t (right panel).

5.2 OMref18.3 - a proposed BET reference OM based on sub-sampling for consistency with the assessment

A new reference case grid OM-refB18.2 was created by expanding OM-refB18.0 with additional options for data weighting - one step of iterative re-weighting for the CL data (CLRW) based on the original assessment specification, and a lower CPUE CV = 0.1 (i10C and i10H), forming an ensemble of 432 models. The new dimensions recognize concerns about the arbitrary nature of the weighting for the various data series, and were primarily intended to fill in more of the uncertainty space in a continuum that can be sub-sampled. The new options were admittedly arbitrarily chosen, but both shift the central tendency of MSY and SB/SBMSY toward the values reported in the assessment (Figure 11).

OM-refB18.3 followed the YFT approach of sampling (with replacement) from refB18.2 to achieve the central tendency of the BET assessment in terms of SSBY/SSBMSY = 1.29 and MSY = 104 Kt (and removing 48 of the 432 models did not meet the convergence criterion, max. gradient < 0.1). Assuming a normal approximation for the aggregated 80% confidence intervals reported in the assessment (from WPTT 2017) results in an MSY CV = 0.127 and SB2015/SBMSY CV = 0.136. These CVs are considerably larger than the values derived from the YFT assessment, because the BET assessment uncertainty encompasses a grid of 6 models. We propose using these CVs for the BET OM, because they are similar in magnitude to that used for the proposed reference case YFT OM OMref17.2 (i.e. ~3X the CVs from the YFT reference case assessment). However, we recognize a certain arbitrariness in this approach, and welcome alternative suggestions. We were not able to ascertain how the BET assessment CVs were calculated. If we assume that it was based on only the MPD estimates and variances are additive, we might suggest that a minimum CV for MSY might be 0.127 + ~0.045 (Hessian CV for YFT) = 0.172.

The BET sampling approach was also intended to have 50% representation for each of the catchability trend options.

A comparison of key characteristics from OM-refB18.0, OM-refB18.2 and OMrefB18.3 are shown in Figure 9 - Figure 13.

The projection dynamics for OMref18.3 with a fishing moratorium (starting in 2019) show that the central tendency of the population recovers to slightly above that of the early period of the fishery, which is to be expected since the population dynamics start from a lightly exploited state (Figure 10). The current catch projections suggest that current catches are probably sustainable, but unlikely to result in increasing biomass (like OMref18.0). OMref18.2 and OMref18.0 both suggest 10% < P(B/BMSY) < 25% for 2018 and 2025, which is compatible with the K2MSM predictions.

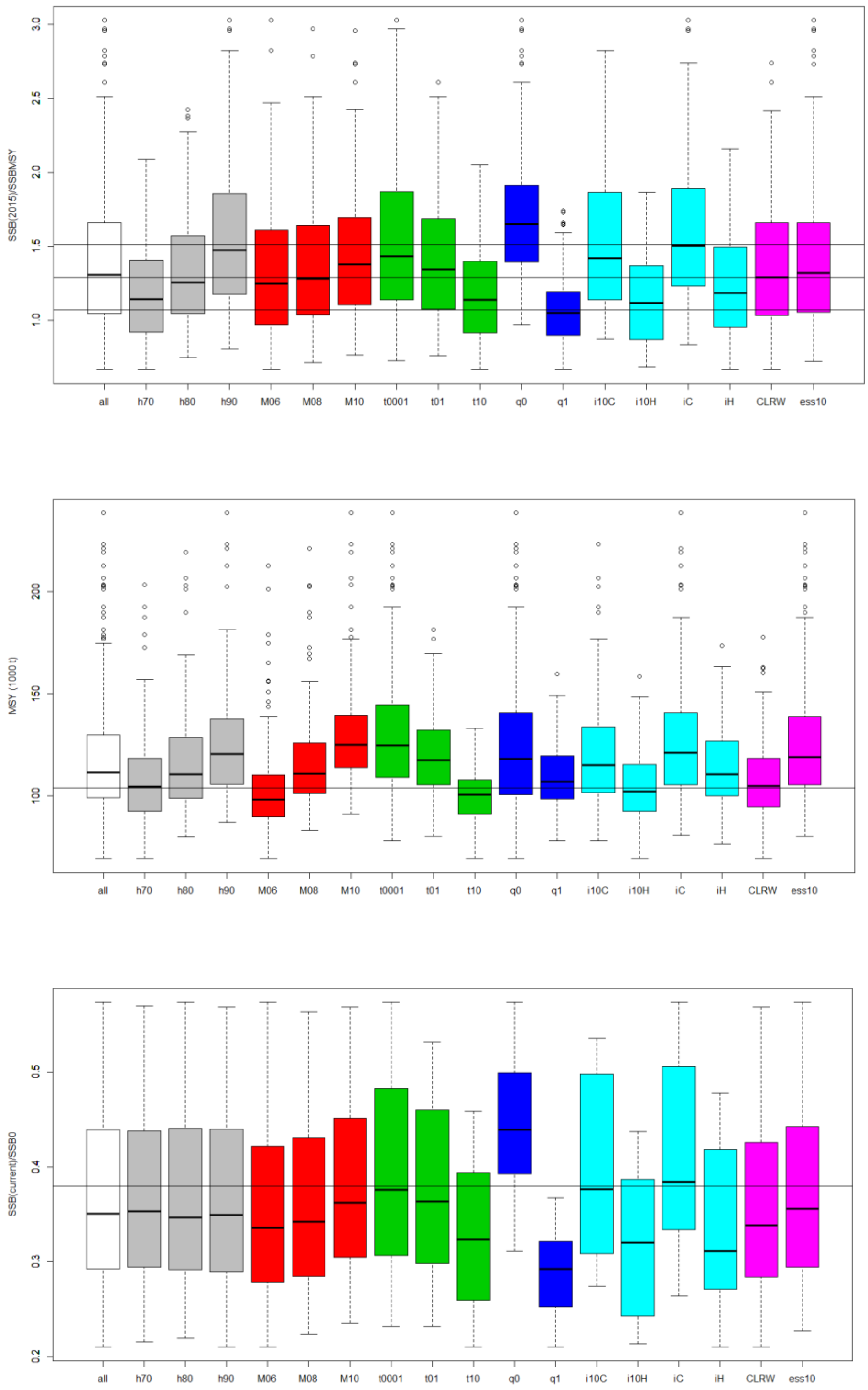


Figure 11. Stock status characteristics of OMrefB18.2, the unweighted grid from which OMrefB18.3 is sampled.

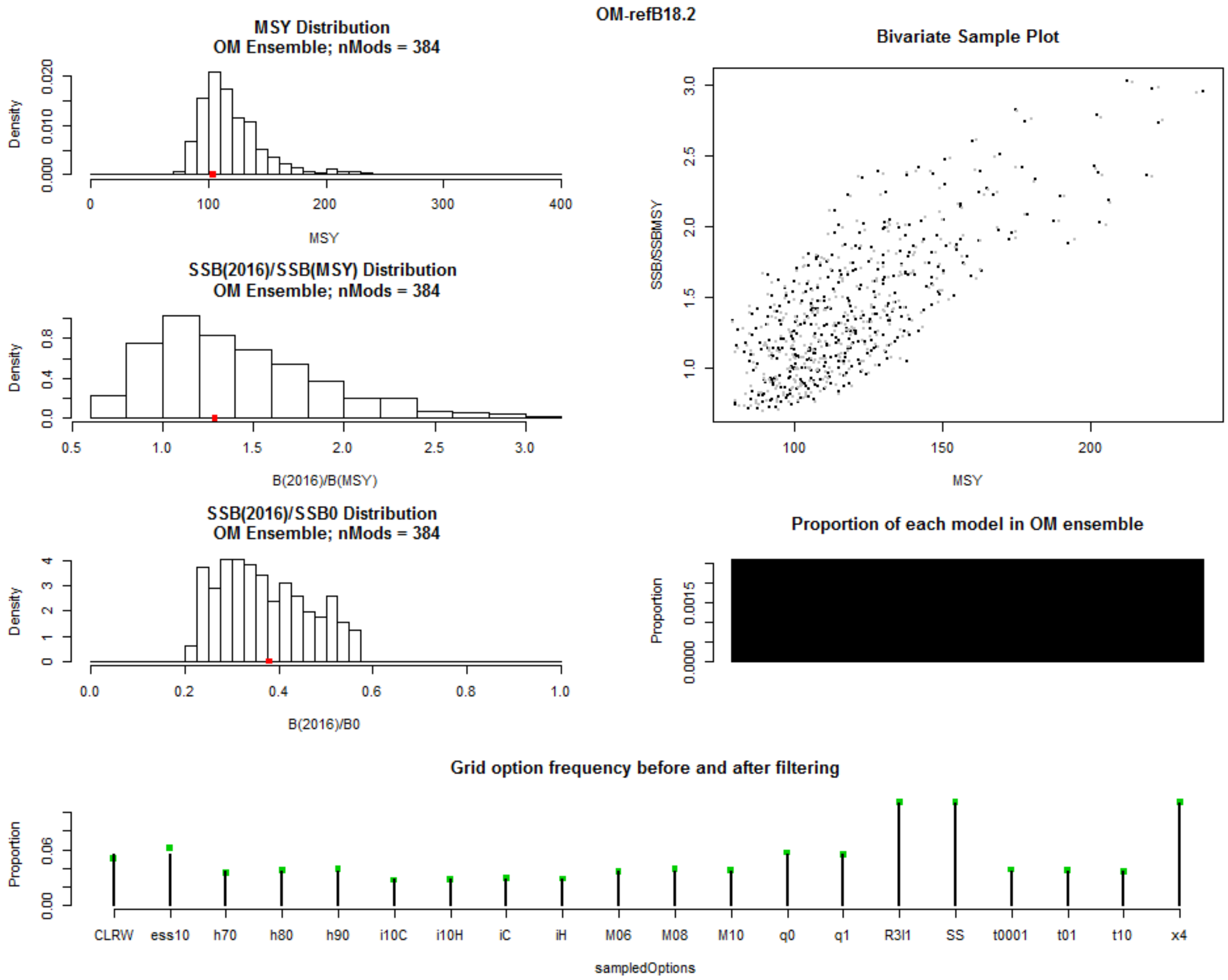


Figure 12. Characteristics of OMrefB18.2. Red circles indicate the point estimates from the 2016 assessment. The top right panel indicates the relationship between MSY and SBY/SBMSY (grey points are jitters to emphasize repeat sampling frequency). The middle right panel indicates the relative frequency of the models sampled (uniform in this case). The bottom panel indicates the relative proportion of the individual assumptions in the ensemble (green points) relative to the original grid (black lines) - essentially identical except for convergence failures.

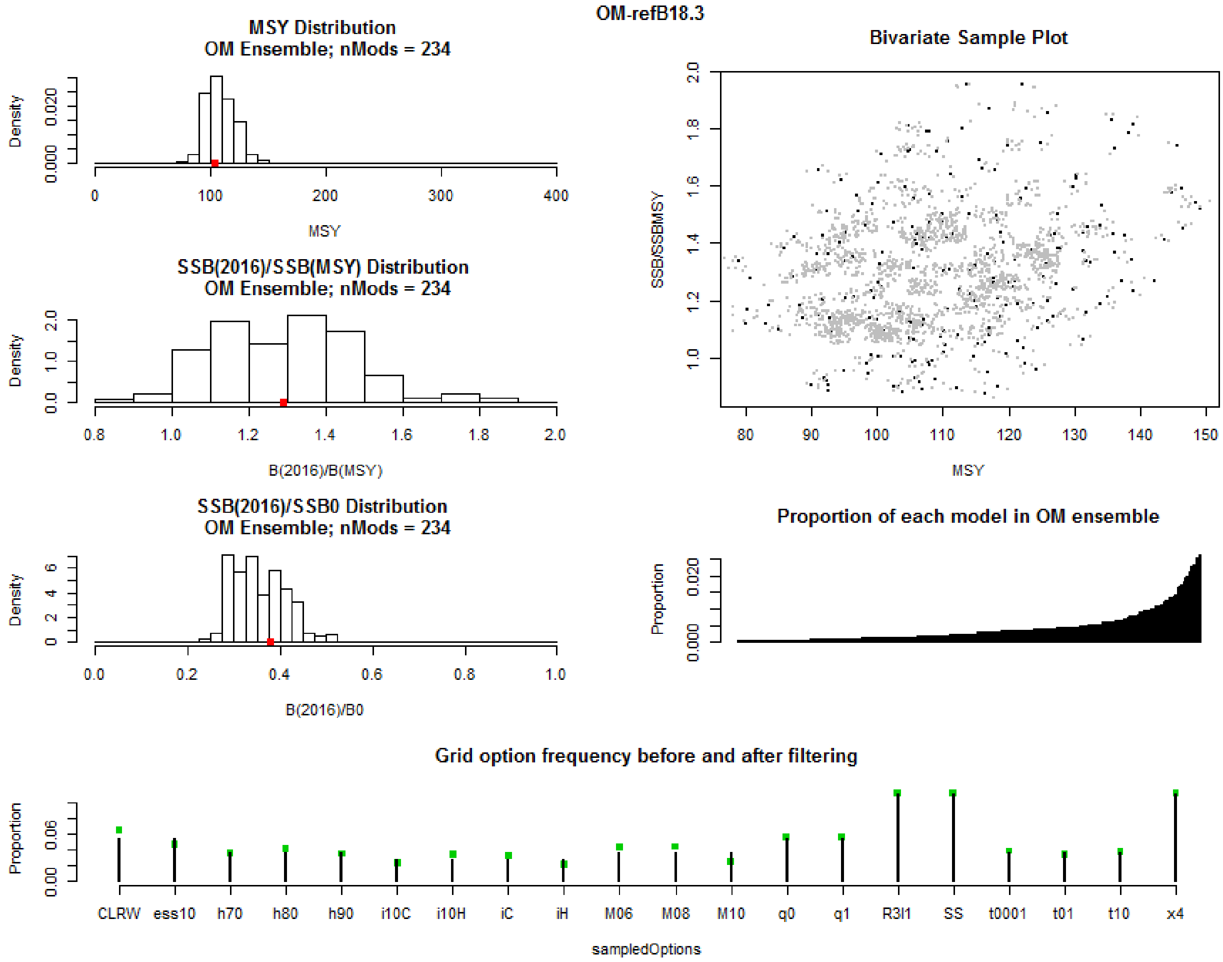


Figure 13. Characteristics of OM-refB18.3, the proposed reference set OM derived using the bivariate sampling. Red circles indicate the point estimates from the 2016 assessment. The top right panel indicates the relationship between MSY and SBY/SBMSY (grey points are jitters to emphasize repeat sampling frequency). The middle right panel indicates the relative frequency of the models sampled. The bottom panel indicates the relative proportion of the individual assumptions in the ensemble (green points) relative to the original grid (black lines).

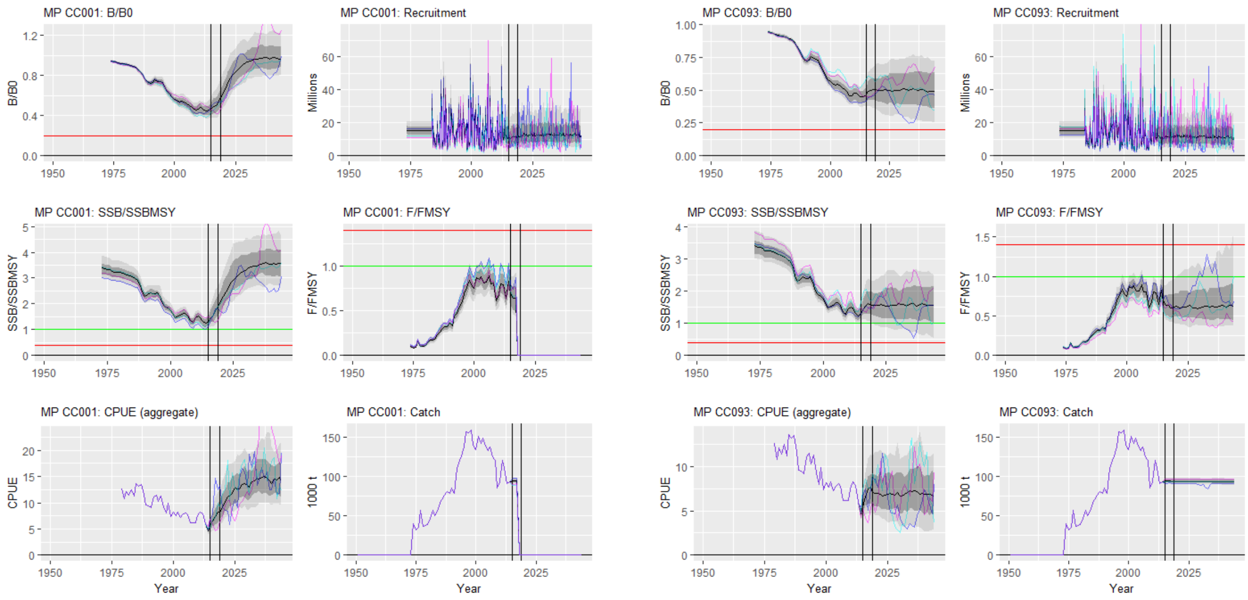


Figure 14. OMref18.3 projections with a fishing moratorium starting 2019 (left panel) and with constant current catches at 93 000t (right panel).

6 MP performance when evaluated with the bigeye reference case OMs.

Results from candidate MPs are reported as defined in Table 5. See Kolody and Jumppanen (2016) for the full specification of these MPs, noting that the PT4010F option was added as described in Kolody and Jumppanen (2017). 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 or data processing)

iii) Harvest Control Rule to specify the management action

All three elements of the MP are simulation-tested together.

These MPs were tuned according to the criteria defined in Table 6. For expedience, all tuning for this report was conducted with 200 (or 216) realizations. Previous testing indicated tuning with only 200 realizations resulted in performance within 5% of the tuning objective when subsequently applied to a full suite of 2000 realizations. This level of tuning precision is considered adequate for the purposes of this report, but the full set of 2000 will be used for the TCMP.

Table 5. Qualitative definitions of the MPs used in this report.

Label	Definition
PT4010	A catch-based "40:10-type" HCR coupled with an observation error surplus production model.
PT4010F	An F-based "40:10-type" HCR coupled with an observation error surplus production model.
IT	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 (i.e. "ballistic" HCR that ignores feedback)

Table 6. The example MP Tuning objectives defined for BET and YFT are similar to those agreed by the TCMP, but not identical

Label	Source	Definition
Y1	YFT objective 1	Pr(mean(SB(2019:2038)/SB(MSY)) > 1.0) = 0.5
B1	BET objective 1	(as Y1)
Y2	YFT objective 2	Pr(SB(2024)/SB(MSY) > 1.0) = 0.5
B2	BET objective 2	Pr(Green Kobe 2019:2038) = 0.75

A single implementation of each of the MPs listed in Table 5 was tuned for tuning objectives B1 and B2 for the default OM specification (OMrefB18.0), and our preferred OM (OMref18.3). No time was available to try and improve performance by manipulating control parameters (other than the tuning parameter). However, it is immediately evident that the choice between the two tuning levels has a much greater impact on performance than the individual MPs.

OMref18.0 candidate MP performance time series plots for tuning level B1 are shown in Figure 15, indicating undesirable behaviour in all cases. Given that the initial stock status has a high probability of being above the target, attaining the target over the 20 year period requires biomass below the target elsewhere in the time period. All the tested MPs drive down the population, such that biomass usually remains on a downward trend at the end of the 20y evaluation period. It is possible that an MP could be constructed which would stabilize behaviour by the end of the 20 year period, but this may require removal of the 15% TAC change constraint (to greatly increase initial catches). Given that there is currently no management restriction preventing the bigeye fishery from catching more fish, it is unlikely that they would be inspired to meet the high TACs that these MPs would recommend. Accordingly, adoption of this tuning level would probably require careful consideration of implementation error. The impact of the high F problem observed for YFT (i.e. the MPs frequently unable to remove the TAC) may also be occurring, but the undesirable behaviour of the B1 tuning will occur regardless.

In contrast, all of the MPs tend to support very stable population trajectories (on average) with the B2 tuning objective (Figure 16). Corresponding Udon-Soba plots are shown in Figure 17.

The MSE was repeated for the authors' preferred OM, OM-refB18.3. The intent is not to suggest that MP results should be used to choose among OMs (that would be somewhat backward). Rather, the intent was to explore just how sensitive the MP selection process would be to the choice of OM. The MP performance for OMref18.3 tuning B1 is qualitatively similar (and equally unappealing) as OMref18.0 (Figure 18). The MP performance for tuning B2 is much more sensible (Figure 19), though notably does not suggest the rebuilding trend observed for OMref18.0. Figure 20 shows the OMref18.3 Udon-Soba plots for B2 tuning only.

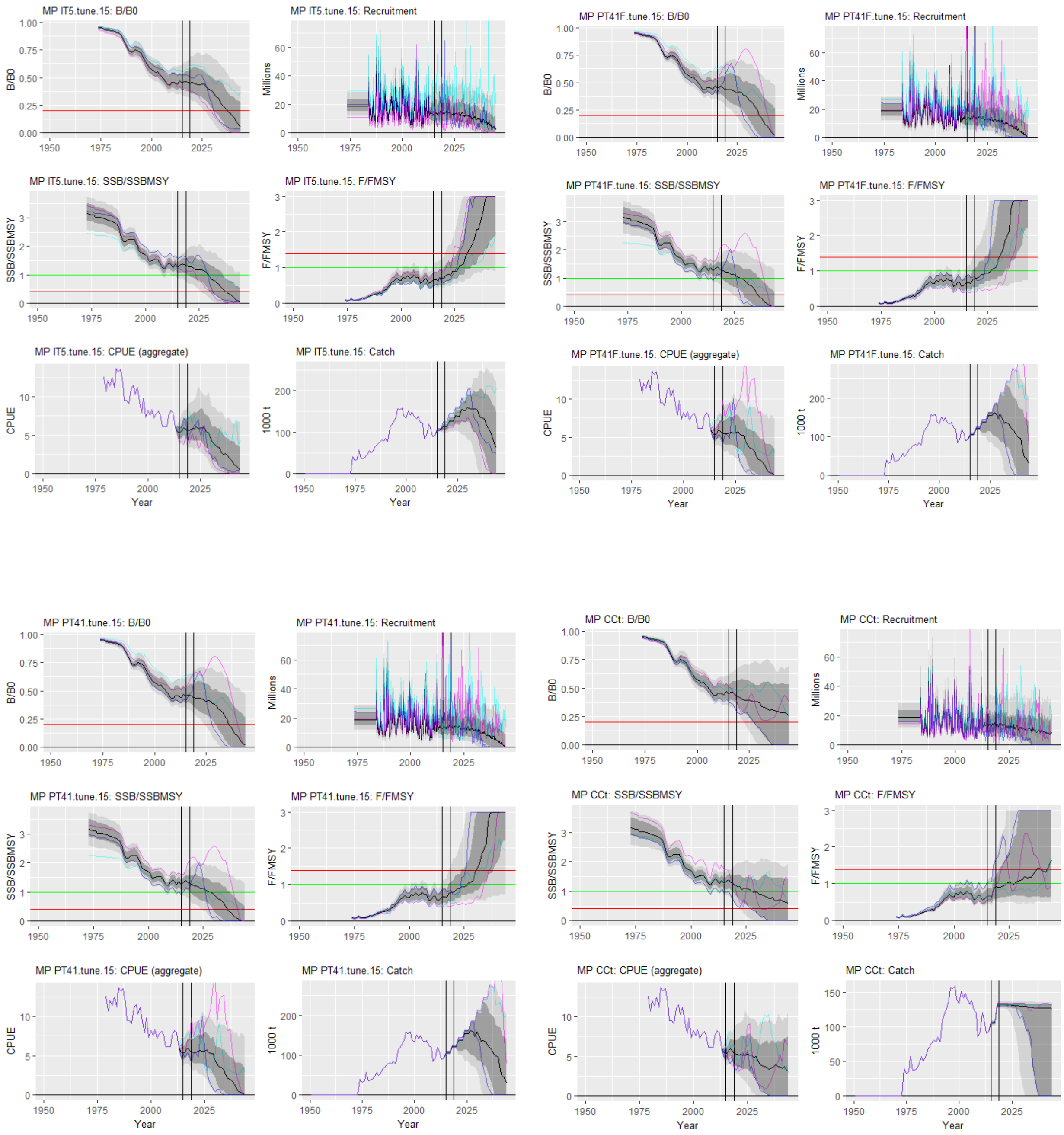


Figure 15. MP summary results for OMref18.0, with tuning level B1.

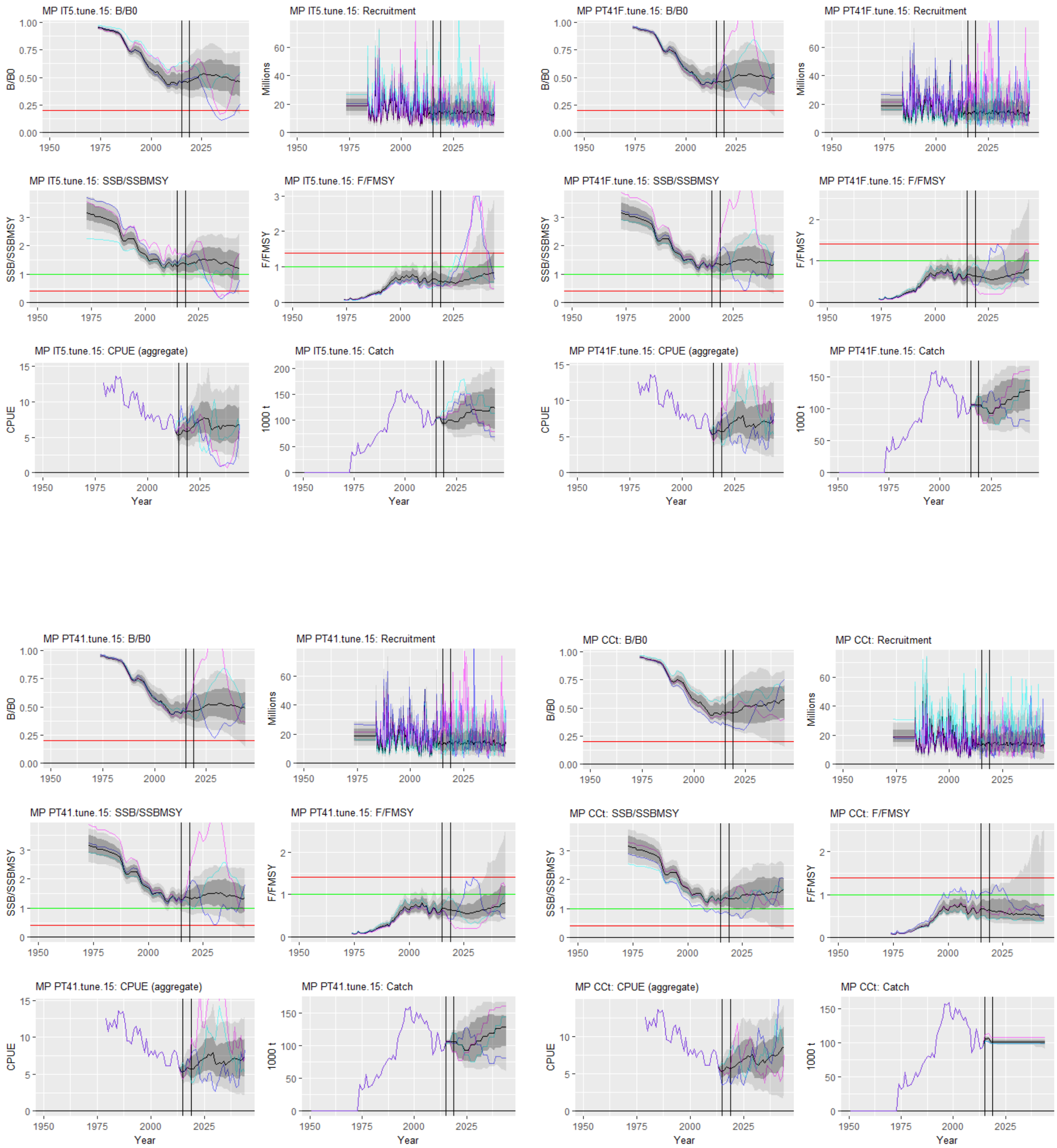


Figure 16. MP summary results for OMref18.0, with tuning level B2.

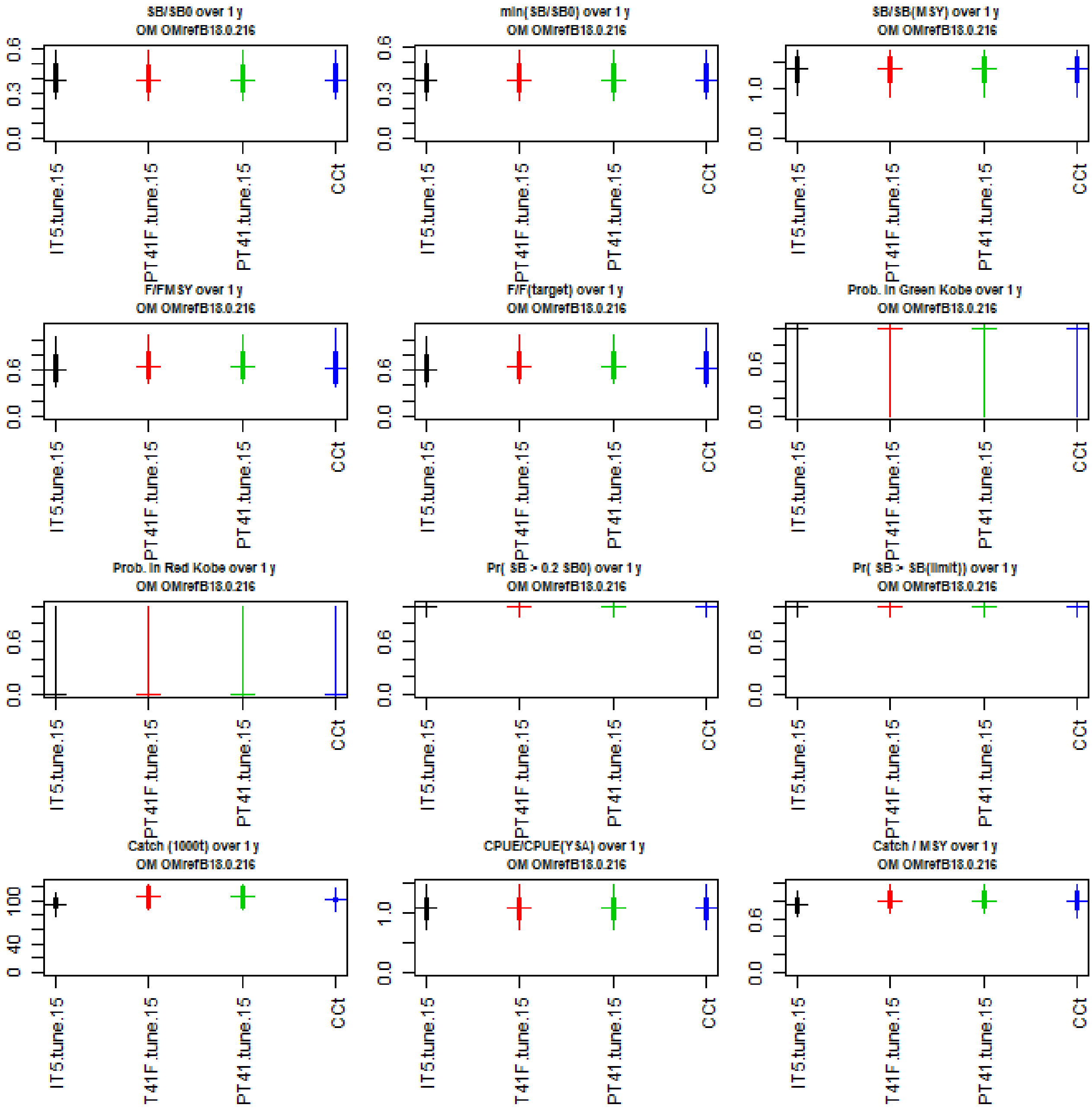
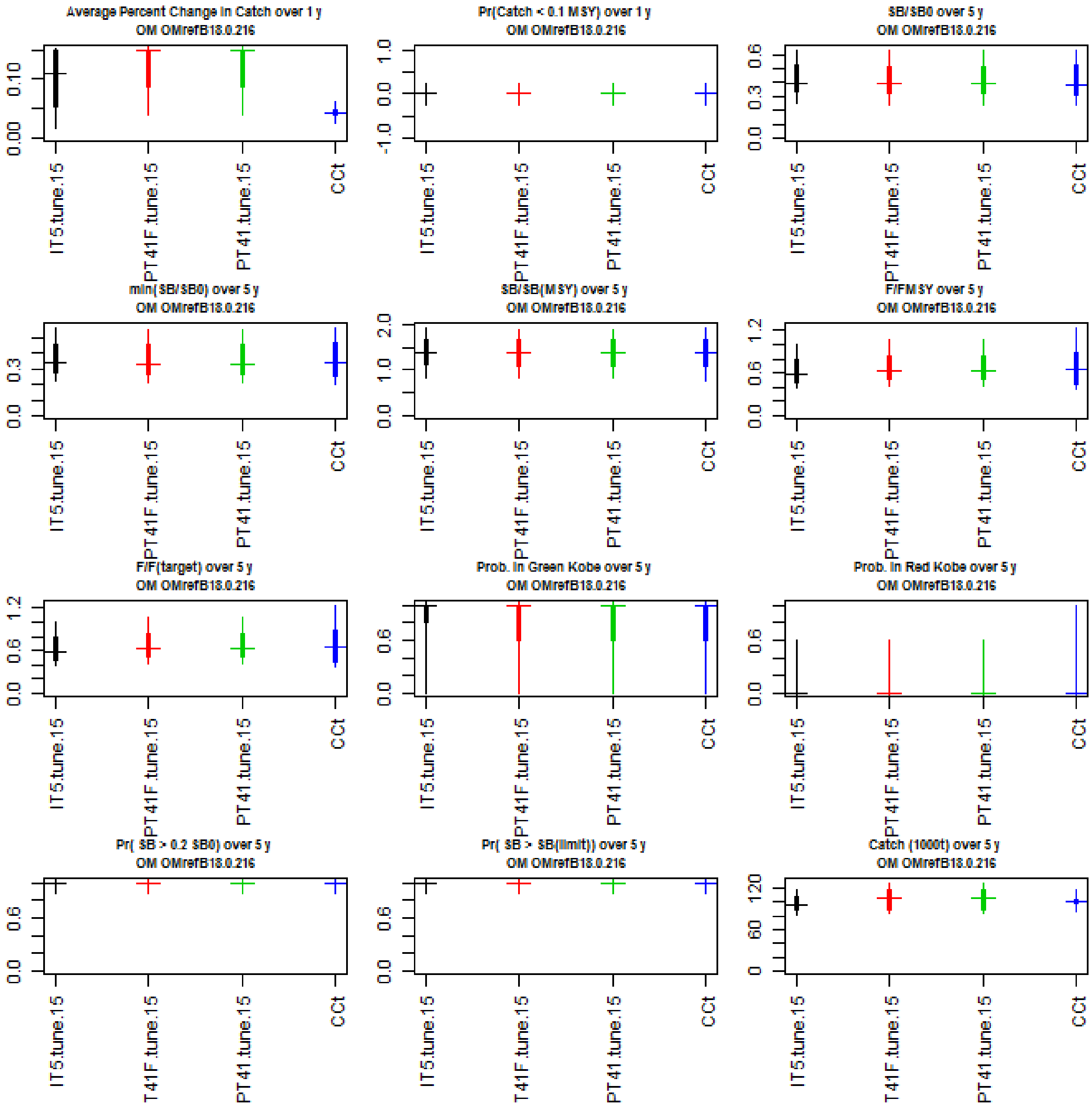
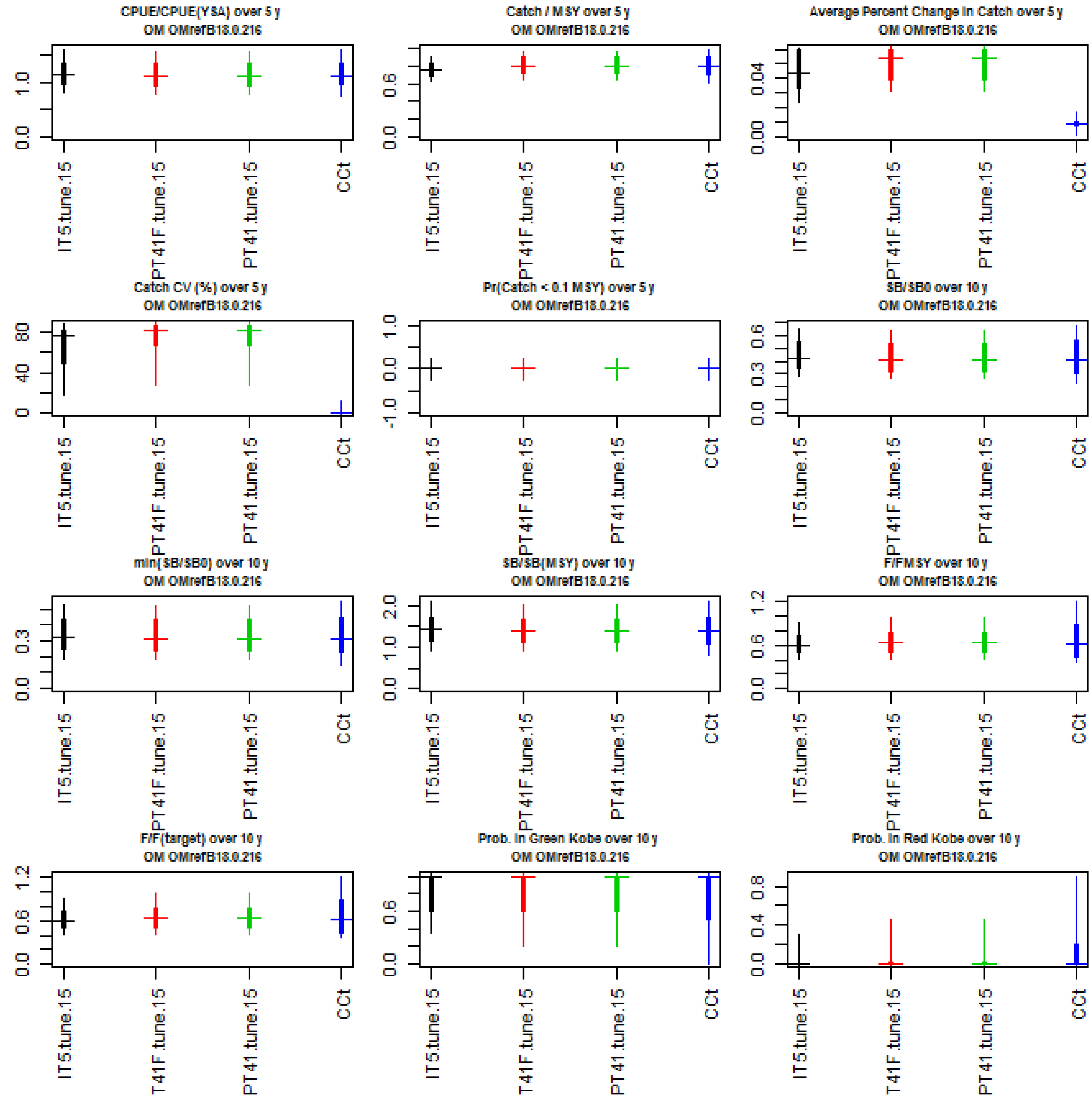


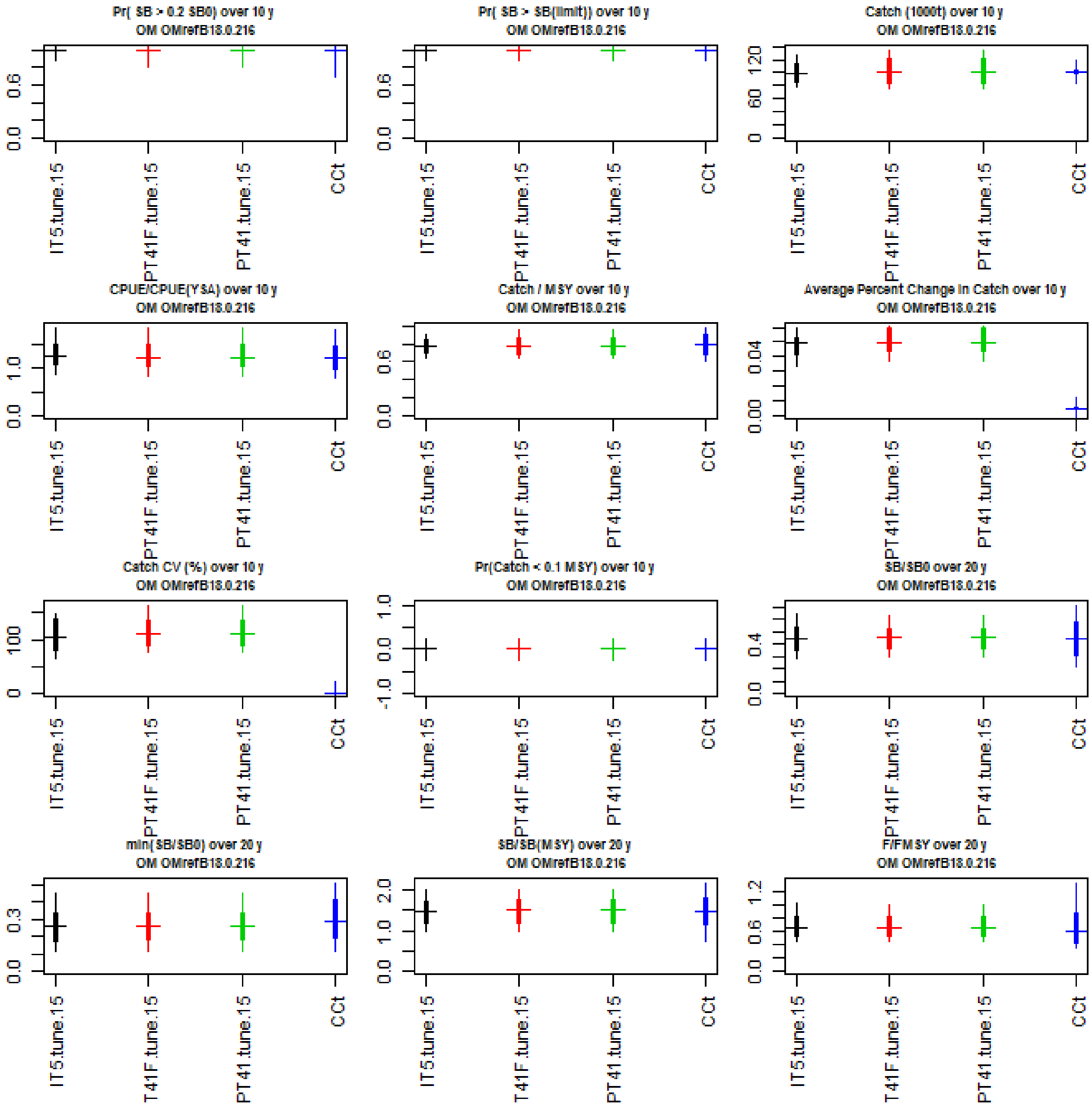
Figure 17. Udon-soba plots comparing the 4 MPs evaluated against OMref18.0 for tuning B2.



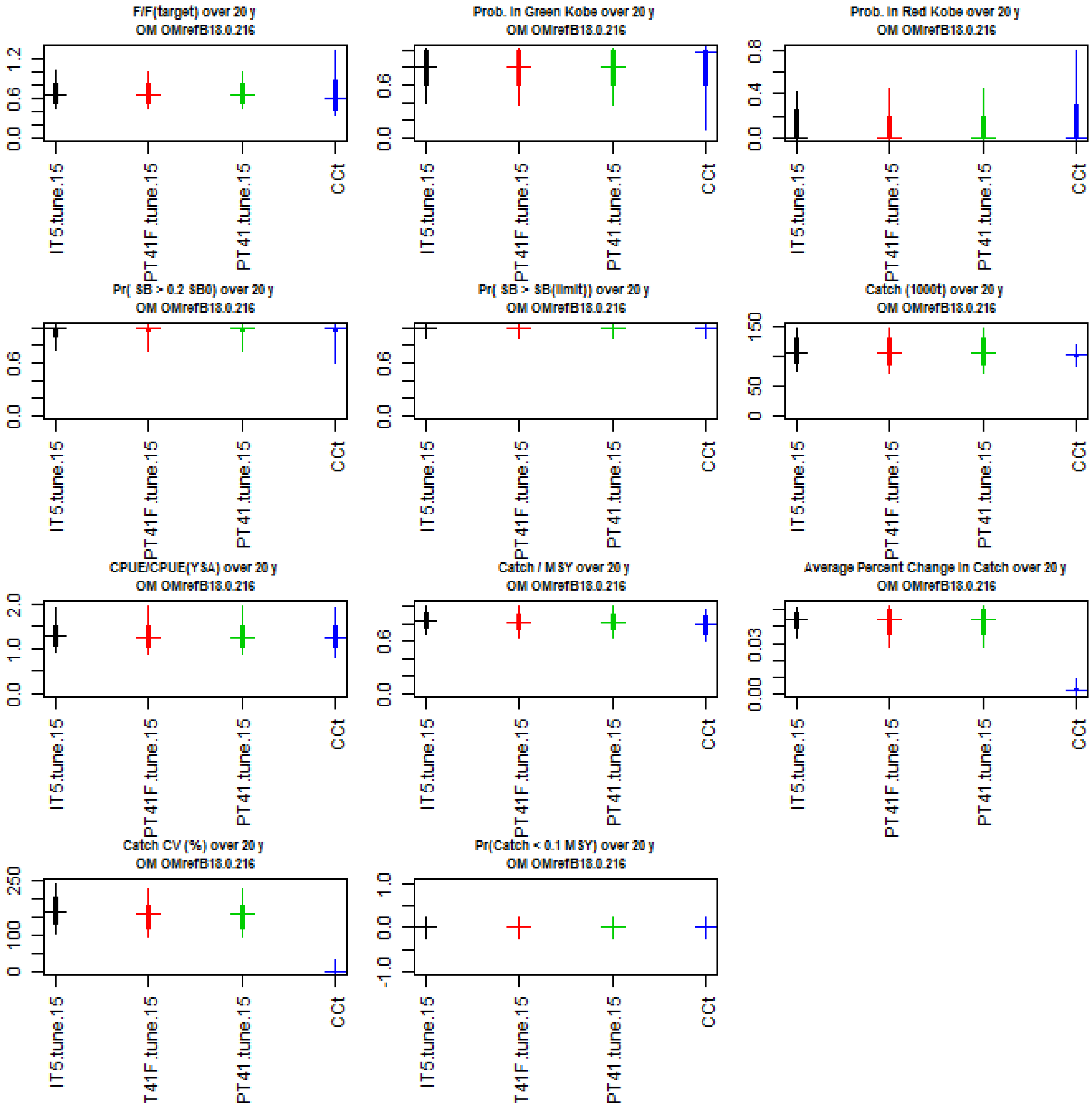
(Figure 17 cont.)



(Figure 17 cont.)



(Figure 17 cont.)



(Figure 17 cont.)

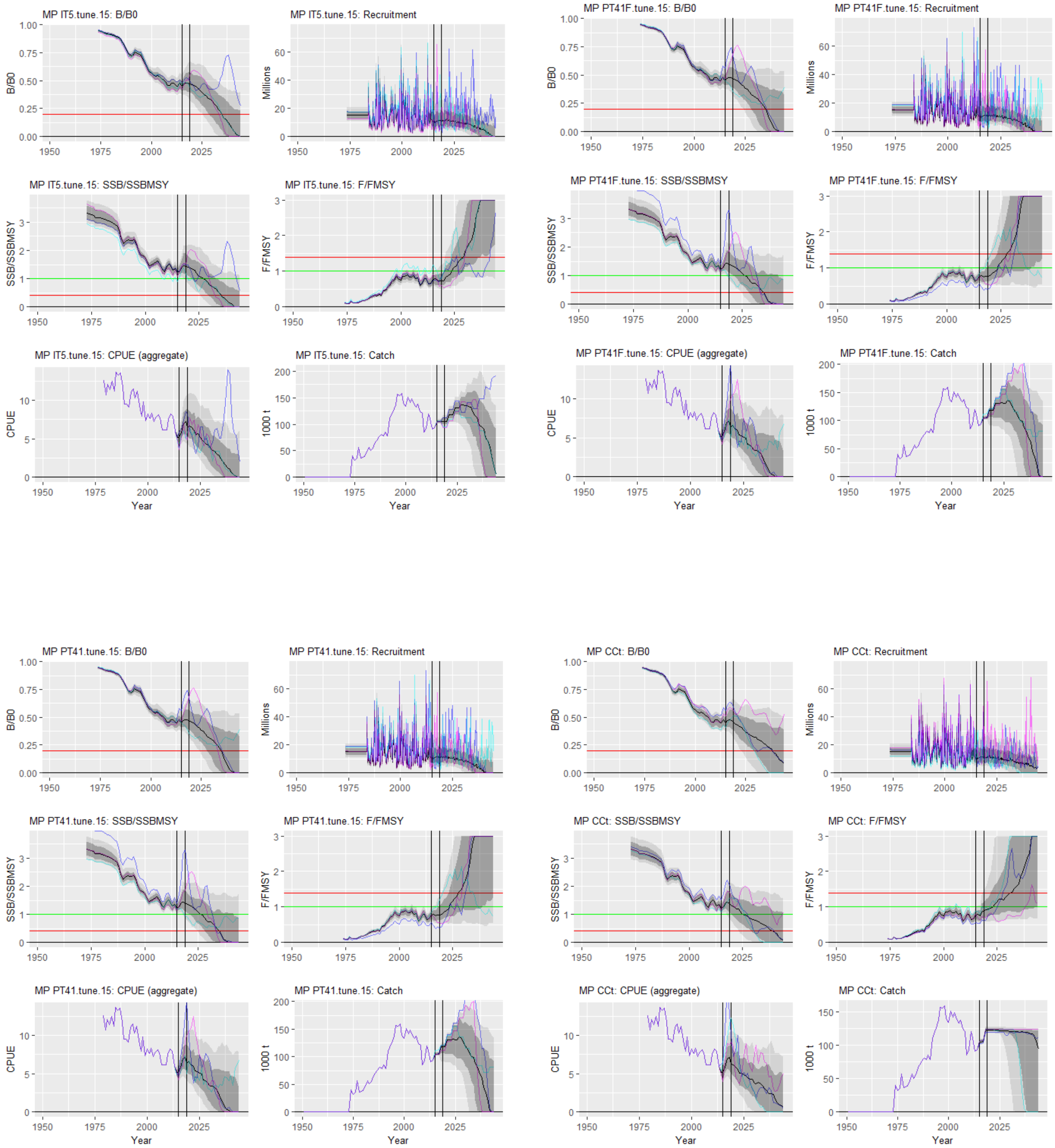


Figure 18. MP summary results for OMref18.3, with tuning level B1.

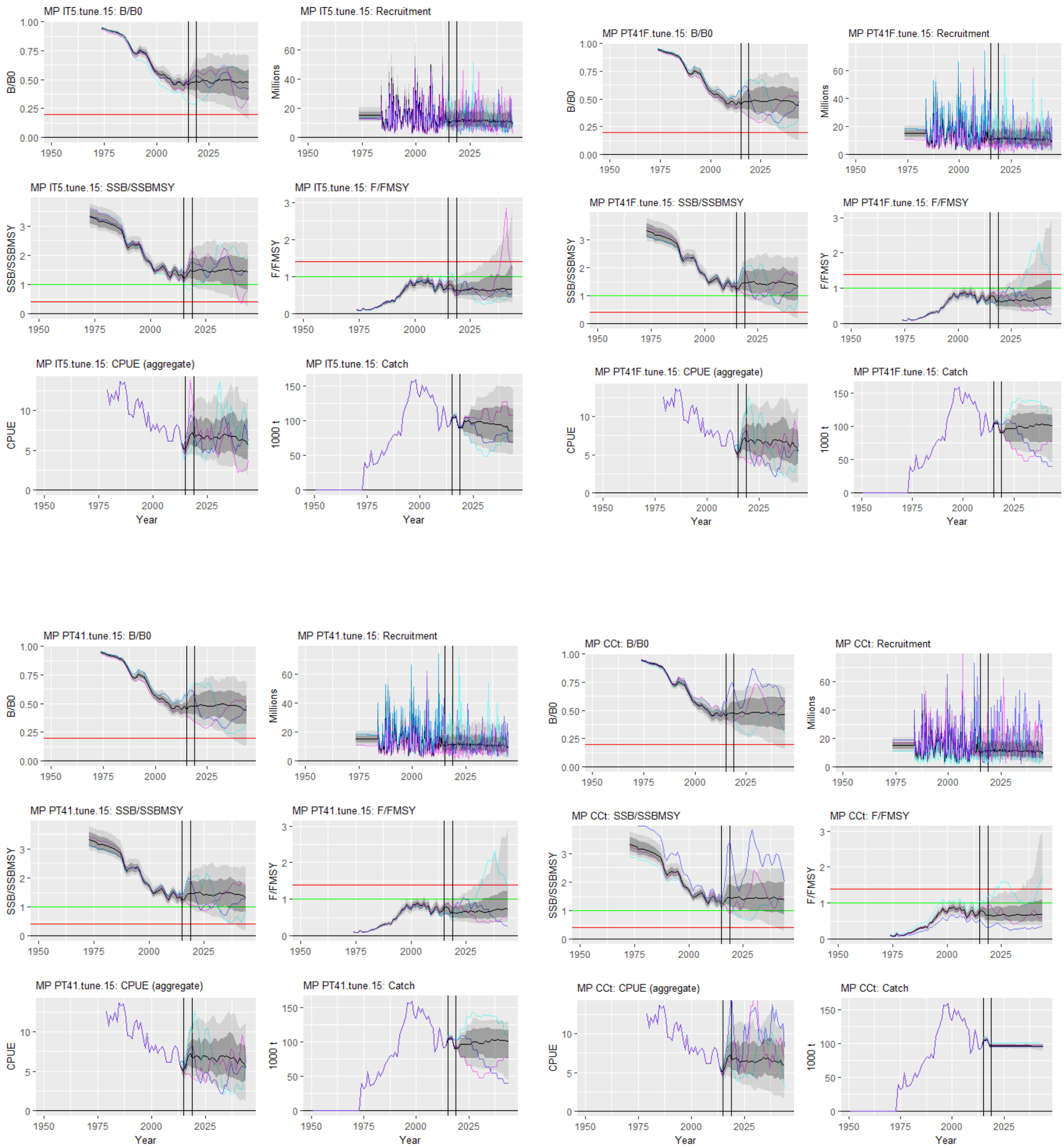


Figure 19. MP summary results for OMref18.3, with tuning level B2.

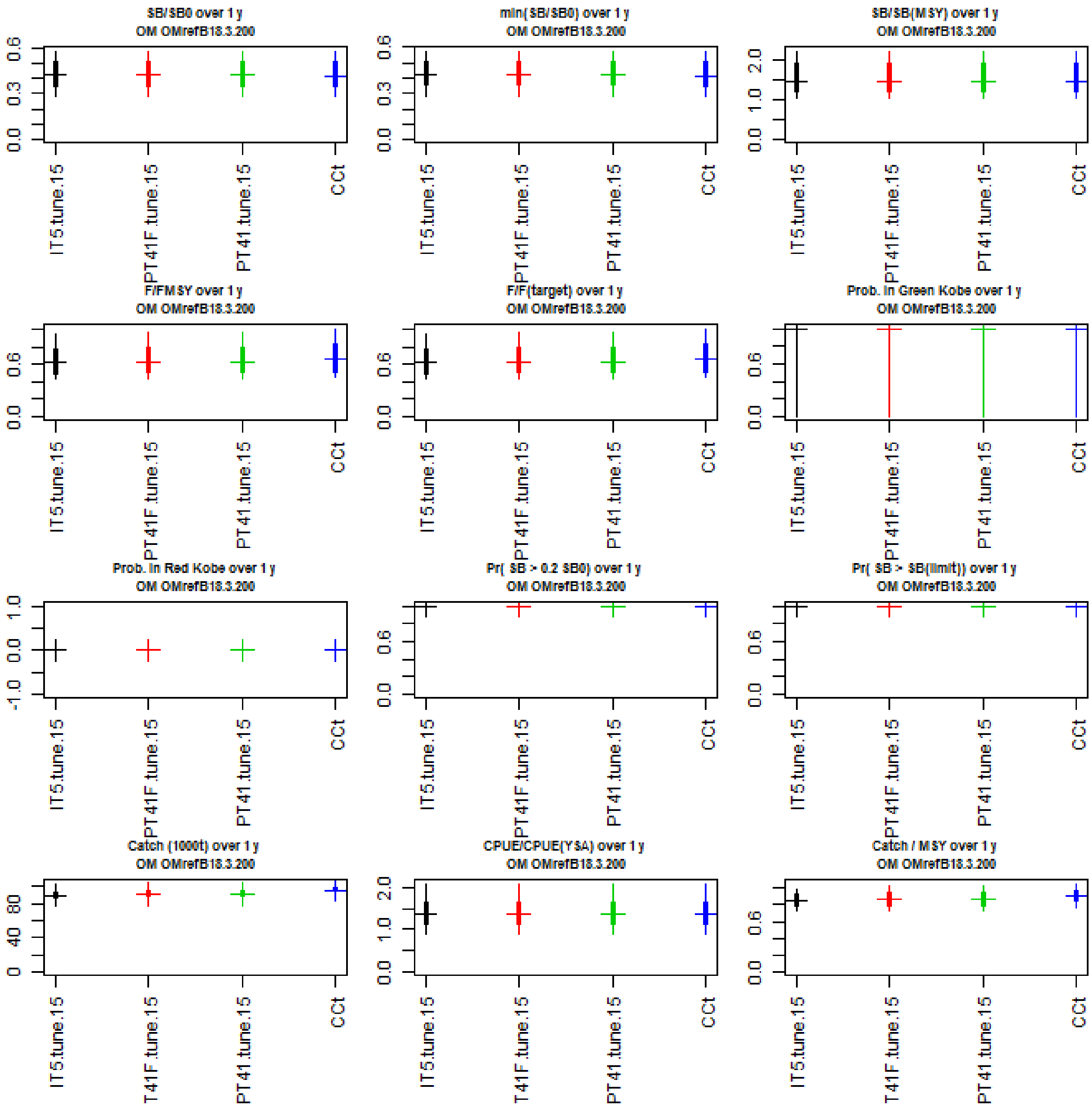
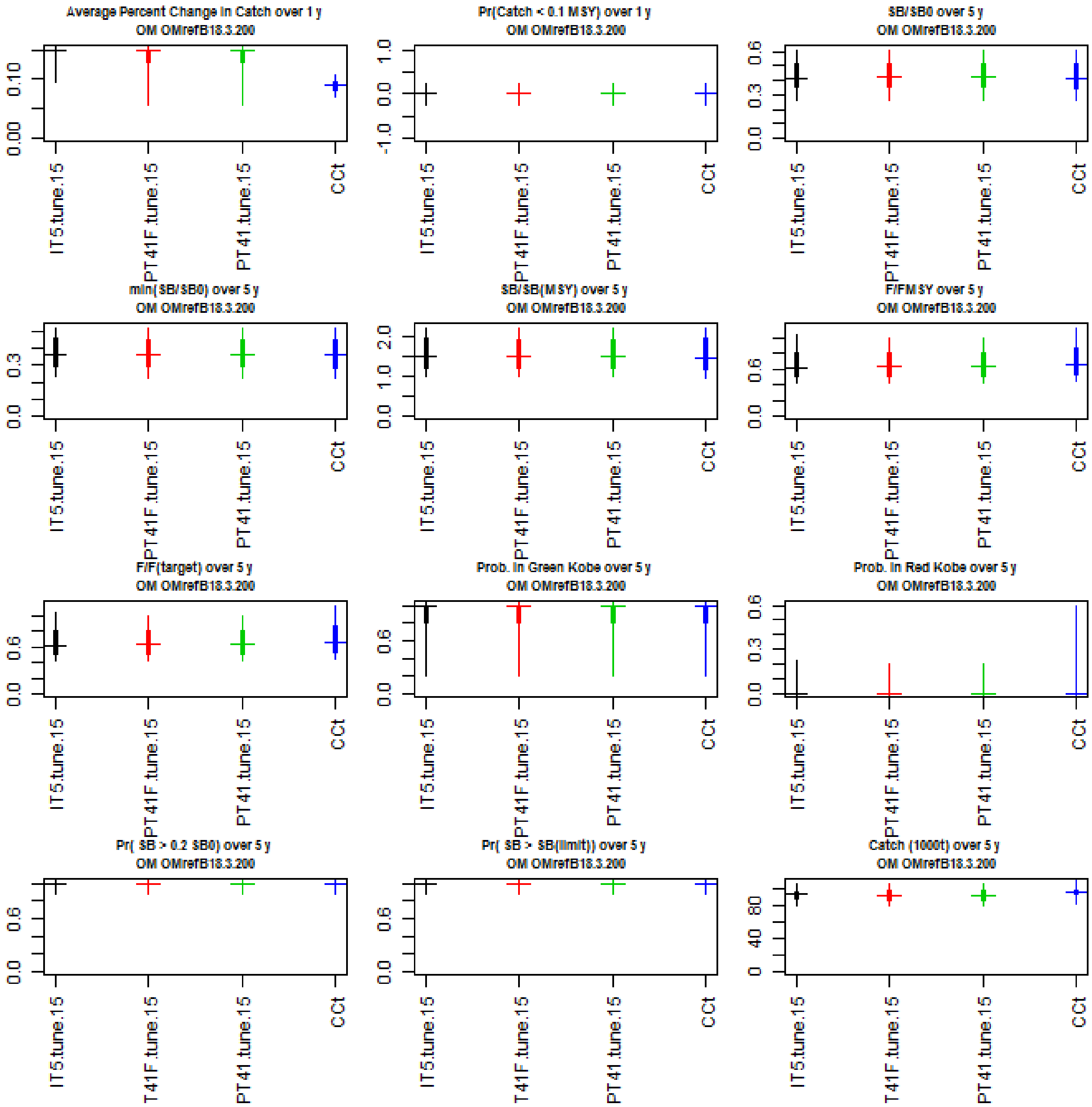
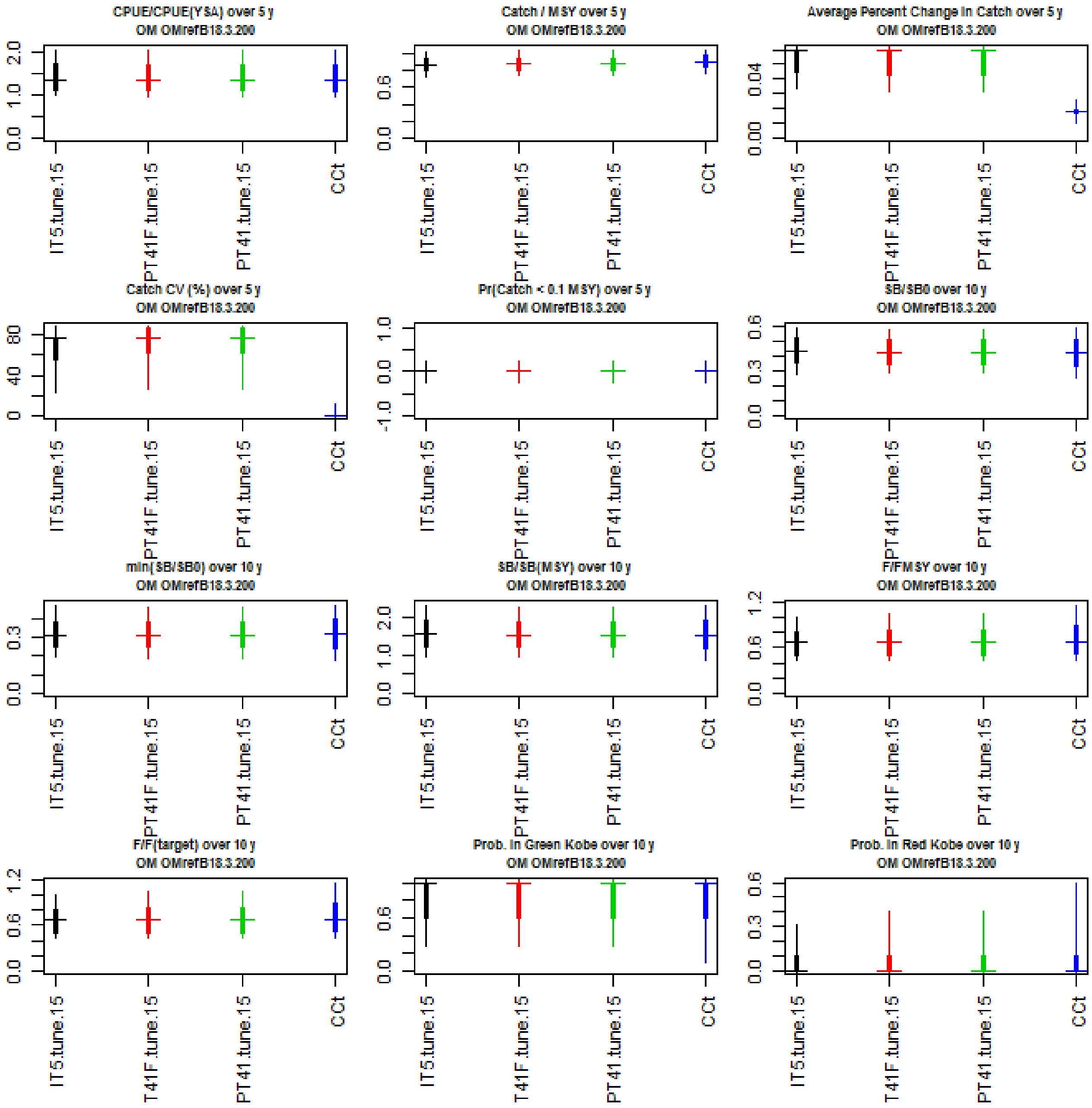


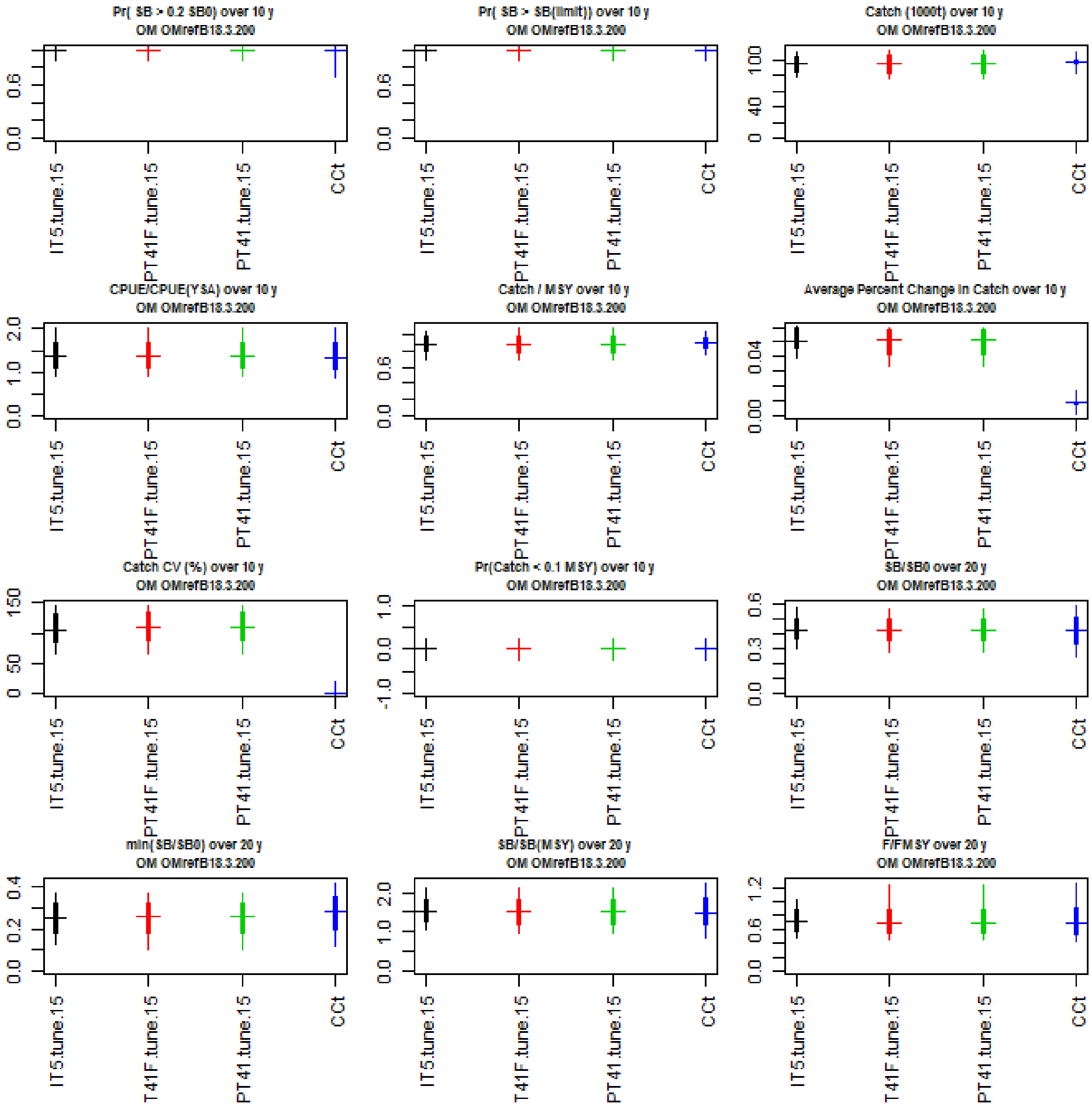
Figure 20. Udon-soba plots comparing the 4 MPs evaluated against OMref18.3 for tuning B2.



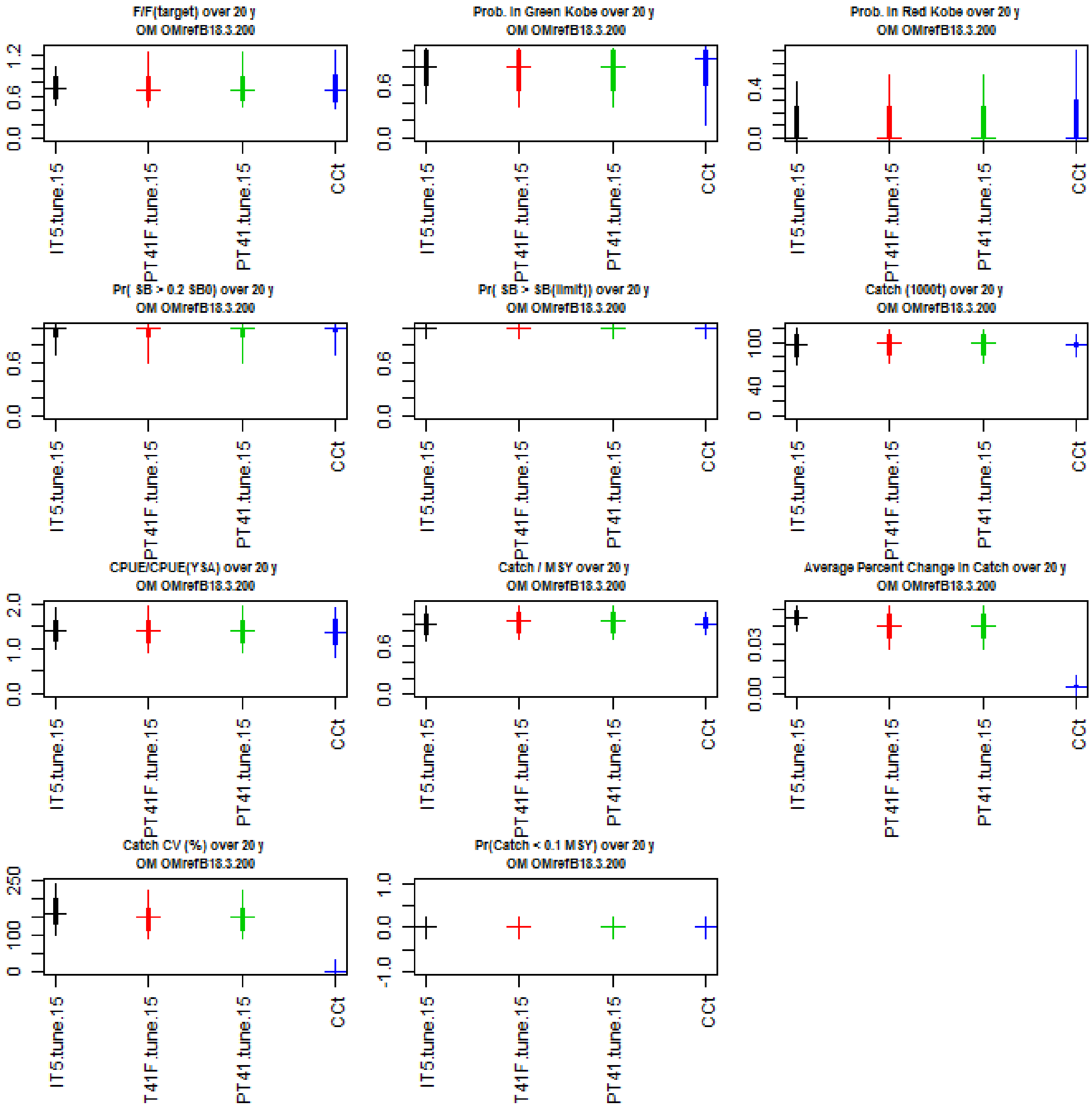
(Figure 17 cont.)



(Figure 17 cont.)



(Figure 17 cont.)



(Figure 17 cont.)

7 BET Robustness scenarios

There was not adequate time to properly consider the proposed BET robustness scenarios in time for the informal MWG MSE meeting. We will be seeking to clarify the priority ranking for the WPTT/WPM 2018 (and the TCMP 2018 if this is judged to be useful at this point).

8 Discussion

This WP summarizes progress on the revised BET OM, and initial MP evaluations, representing the last chance for feedback from the broader IOTC community before results are presented to the TCMP for consideration in 2018. We consider the following priority points for feedback/endorsement for the phase 2 BET MSE to move forward:

1) BET reference case OM

- Is there any objection to replacing 4 seasonal temperate CPUE series with a single aggregate?
    - The stock status inferences were almost identical, but the OM implementation with 4 series would require further development work.
- Should we retain the default BET OM defined by the WPTT/WPM 2016 (for TCMP 2018 and WPTT/WPM 2018 presentations), or move to a weighted re-sampling grid?
    - The potential problem with the default is that it tends to be more optimistic than the assessment, and the uncertainty space moves around considerably with the grid assumptions in ways that may not be predictable or desirable.
    - Adopting the re-sampling approach ensures consistency with the central tendency of B/BMSY and MSY in the assessment. The 13% CV on each dimension is consistent with the small assessment grid, and the YFT proposal, but the specific value is worth further debate. The structural uncertainty would be much greater than the assessment, including inferences from models with the tagging data essentially eliminated.
- In light of the undesirable tuning results for B1 should we be pro-active in proposing alternative MP tuning objectives to present to the TCMP?
- What are the priorities for the BET robustness scenarios for review in WPTT/WPM 2018, and should they be presented to the TCMP?

- We would be reluctant to present robustness scenarios to the TCMP at this stage, because i) they may confuse the learning process, and ii) they will not have been reviewed by the broader IOTC scientific community.
- The list of potential BET robustness scenarios proposed by the WPTT/WPM in 2016 and 2018 (mostly for YFT and listed in the YFT companion document) will be discussed at the informal MSE working group for prioritization.

- While not as urgent as for YFT, we will be adding additional diagnostics to the OM reporting to evaluate the impact of the high F problem (i.e. summarizing the difference between the TAC and the realized catch)

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, Jumppanen, P. 2017. Update on Yellowfin Tuna Management Procedure Evaluation Oct 2017. Indian Ocean Tuna Commission Working Paper IOTC-2017-WPTT19-49.

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. 2016. Stock assessment of bigeye tuna in the Indian Ocean for 2016 — model development and evaluation. IOTC-2016-WPTT18-20.

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]

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