OCEANS AND ATMOSPHERE

www.csiro.au

A candidate Management Procedure based on a Joint Process and Observation Error Random Effects Production Model

Prepared for the IOTC Working Party on Methods (Oct 2020)

Dale Kolody (Dale.Kolody@csiro.au)

Paavo Jumppanen

CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart TAS 7000, Australia

Citation

Kolody, D, Jumppanen, P. 2020. A candidate Management Procedure based on a Joint Process and Observation Error Random Effects Production Model. Report prepared for the Indian Ocean Tuna Commission Working Party on Methods 2020.

This document will be assigned a number in the IOTC archive at a future date.

Contents

Acknowledgments

This work was jointly funded by the Australian Department of Foreign Affairs and Trade and Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO - Oceans & Atmosphere), administered by IOTC/FAO.

1 Abstract

In this paper, we explore a Pella-Tomlinson Random Effects surplus production model (PTRE) that admits joint process and observation error, as a potential estimation model for use within IOTC Management Procedures. The model is spatially-aggregated and operates on an annual time-step, fitting to catch and CPUE observations. Model parameters include carrying capacity (k), intrinsic population growth rate (r), initial biomass, PT production curve "shape" (p - determines B(MSY)/k), CPUE variance and productivity process error variance. To improve minimization performance, an analytical solution is used for catchability (based on the principle of concentrated likelihood) and MSY replaces r as the leading parameter (to reduce correlation with k, and facilitate intuitive priors). The model is fit with Template Model Builder (TMB) software, which allows the productivity process error to be defined as a latent random variable (which may allow the process error variance to be estimated in principle). The result is conceptually similar to the MCMC numerical integration that is often used in simple assessment models (e.g. JABBA). However, the TMB integration is efficiently implemented with the Laplace approximation and can usually be evaluated within a second (e.g. as opposed to the minutes reported for JABBA), enabling evaluation within formal MP simulation testing. While the model performs very well in general, it was necessary to add several constraints in the context required for MP simulation testing (and these constraints would necessarily be part of the formal MP definition, as they may influence the MP performance). Based on the preliminary MP testing (on the latest Indian Ocean bigeye reference set Operating Model), the approach appears promising, and may offer management performance that is more stable, and less risky, than the other candidate MPs tested to date.

In parallel with the MP development, CSIRO developed an R package to support line by line step-through debugging of TMB (and other C/C++) DLLs (which are notoriously difficult to debug from R).

2 Introduction

This paper is presented as part of the ongoing process to develop simulation-tested Management Procedures for bigeye and yellowfin tunas, as documented in the software and report archive https://github.com/pjumppanen/niMSE-IO-BET-YFT/.

Simple production models with a deterministic production function (as explored to date in IOTC yellowfin and bigeye MPs) often have trouble fitting to long time series of data if there is a systematic lack of fit to the underlying production curve, important transient lag effects due to age-structure, and/or substantial production variability (particularly with serial correlation). This potentially causes two types of problems: i) the model may be badly biased, resulting in poor population inferences (and hence bad management advice), and ii) the model may be numerically unstable with a polymodal likelihood surface, and hence difficult to apply in the automated context required for Management Strategy Evaluation. The observation error models that we have applied to date have required multiple numerical tricks to enable automated testing, including such things as a gridded search over initial parameters values, and differential weighting of the CPUE time series (i.e. high weight on the earliest and most recent CPUE observations, to ensure that the current depletion and recent trend are adequately represented, at the expense of poorly fitting transient dynamics in the middle of the time series). A more stable approach would be preferable.

Joint process and observation error production models may represent a better alternative for MPs. e.g. JABBA (Winker et al 2018) represents such a model that is familiar to some of the IOTC scientific community and has been mentioned as a potential option. However, as far as we understand, JABBA is designed to be applied with a stochastic numerical integration (Markov Chain Monte Carlo methods), which is computationally prohibitive (reportedly 2-3 minutes per model fitting). This is very slow in the context of the testing required for Management Strategy Evaluation, where many thousands of MP evaluations are required. As an alternative, we developed a Pella-Tomlinson production model using Template Model Builder (TMB) software (Kristensen et al 2016., https://kaskr.github.io/adcomp/index.html). TMB uses the Laplace approximation to integrate the likelihood over the productivity process error random effects and uses automatic differentiation to analytically calculate objective function gradients to machine precision. The model provides results that are similar in character to JABBA in terms of the state-space parameterization and integration over random effects (also noting that we found it essential to include parameter priors for numerical stability).

The following describes the TMB-based Pella-Tomlinson Random Effects (PTRE) model, and compares initial MP evaluations to previous approaches.

2.1 BuildSys – a new R package for debugging TMB (and other) C/C++ DLLs

TMB is notoriously difficult to debug, in part because of the constraints imposed by the R environment. In parallel with the MP development, CSIRO (author PJ) developed an R package build system, based on GNU, that supports line by line step-through debugging of C/C++ DLLs, in association with the (free) Microsoft Visual Code development environment. This package is freely available to interested parties from https://github.com/pjumppanen/BuildSys/releases/tag/1.0.4, and is implemented for both windows and linux (though ongoing support will presumably depend on the level of interest).

3 The Joint Process-Observation Error Pella-Tomlinson Random Effects (PTRE) Model

The Pella-Tomlinson production model tested here borrows some of the implementation ideas described in Andre Punt's unpublished lecture notes on Template Model Builder (TMB), and/or James Thorson's https://github.com/James-Thorson/state_space_production_model. The model parameters and variables are described in Table 1. The basic surplus production function is given by:

(1) $$B_t = B_{t-1} + \left(\left(\frac{p+1}{p}\right) rB_{t-1}\left(1 - \left(\frac{B_{t-1}}{k}\right)^p\right) - C_{t-1}\right) \exp\,(\epsilon_t).$$

This is identical to the production function used in previous MPs for IOTC bigeye and yellowfin, except for the addition of the annual production process error (exp (ϵ_t), where ϵ_t is assumed Normal(0, σ_ϵ)). This process error is defined as a latent variable, such that TMB integrates the likelihood (approximately) over all possible values (as opposed to the Errors-in-Variables or penalized likelihood approach commonly used in IOTC Stock Synthesis and most other complicated stock assessments, which seeks to identify the best value for individual recruitment events as individual parameters). As far as we understand, using random effects is more theoretically attractive, particularly when it comes to the estimation of process error variance.

The model operates on an annual time step, where the historical CPUE observations consist of the aggregate of the regionally-scaled CPUE used in the reference case assessment (assuming constant catchability among regions and summed over seasons). The standard relationship between CPUE and biomass is assumed:

(2) $$I_t = qB_t \exp(\delta_t),$$

where δ_t is Normal(0, σ_I). Catchability is calculated using an analytical solution derived from the principle of concentrated likelihood (only years with valid CPUE observations in all 4 seasons and all 4 regions are included):

(3) $$q = \exp\,(\tfrac{1}{T}\textstyle\sum_t \log\,(I_t\,/B_t)).$$

The representation of uncertainty is largely analogous to the result of an MCMC integration (as used in JABBA), but can be achieved very quickly (~1 second for the applications here). However, in practice, there are limited data with which to estimate all of the model parameters, and it was necessary to add parameter constraints to achieve reliable convergence in the context of MP simulation testing.

As presented, the objective function consisted of the two lognormal likelihood components corresponding to the process and observation error above, plus the penalties associated with the four estimated parameters in Table 1. With the exception of k, all log parameters have a normal prior. The k parameter has an ad hoc flat-bottomed U-shaped penalty, which is invoked to be largely uninformative until soft bounds are approached (e.g. Figure 1). This penalty was invoked to prevent some very unrealistic (and rare) results in which the model tended to estimate exceedingly large k and implausibly low r. We do not claim any theoretical justification or historical precedence for this penalty, only that it seems intuitively sensible, and TMB did not seem to have any numerical problem with it (unlike some other approaches tested). The k bounds were approached only rarely, and the MSY and depletion estimates seemed to remain plausible as the bound was approached, so we would not expect this bound to cause anomalous MP behaviour.

The results of two PTRE model fits are shown in Figure 2, for Operating Models from the bigeye reference set that have relatively low and high MSY. These figures illustrate three typical results: i) the PTRE model fits the CPUE very well, ii) a modest level of autocorrelated productivity variability is estimated with a fairly consistent historical pattern, and iii) following a few years of new observations, the PTRE model gains some capacity to distinguish between relatively low and high productivity OMs.

The MP testing to date consisted of fitting the PTRE model, followed by the application of the basic Harvest Control Rule as used in previous model-based MPs (Figure 3). The PTRE model was generally very stable, but instabilities tend to arise under the diverse testing conditions within MSE. The following additional procedures were employed to increase numerical stability:

1. During intial testing, the PTRE model mostly failed when the population was highly depleted, when we might expect the transient age-structured dynamics of the Operating Models to strongly diverge from simplified aggregate biomass dynamics. We added an over-ride to the MP in which the PTRE model is not fit, if the most recent observed CPUE is less than 10% of the mean of the first 3 CPUE observations. If the system reaches this state, it seems clear that management has already failed, and the model is not required to confirm that drastic action is required. Instead, the TAC is simply reduced by the maximum allowable rate until CPUE rebuilds to the 10% threshold.

2. A number of parameter constraints and priors were introduced as in Table 1. It is not realistic to expect the PTRE model to estimate all of the parameters based on catch and CPUE alone. As an example, prior experience has shown that it is usually not practical to estimate p (the production curve shape parameter). We did explore relaxing this parameter with a weak prior (CV log(p) ~25%), and found that the model usually deviated very little from the prior mode. However, it did deviate substantially on rare occasions, causing numerical failures (<1 in 500). Fortunately, we have external information for some of these parameters, and the MP tuning process ensures that the MP performance achieves the expected performance on average, even if there is a bias introduced from our structural assumptions and priors. We would not expect the (tuned) MP performance to be very sensitive to these assumptions, but it is worth further testing.

3. Occasionally, the constrained model still did not converge (max. gradient <0.001) from the initial starting point. In these cases, each PTRE model application included automatic refitting from different initial starting points. Jitter testing suggested that the models that did converge always reached the same mimimum (to within a 0.001 likelihood unit test criterion).

Table 1. Model Parameters and Variables (estimated parameters as tested in bold)

Parameter, variable, data or subscript	Definition	Comment (priors and penalties)
MSY	Maximum Sustainable Yield	Defined as a leading parameter (in place of r) to reduce correlation with k and facilitate intuitive priors prior mode 0.8 C_{max}, σ = 0.25
k	Population carrying capacity	Flat U-shaped prior defined in Figure 1 Range (5 C_{max} – 40 C_{max})
p	"shape" parameter	Determines the shape of the production function, prior mode -0.16, σ = 0.05 (p = -0.16 corresponds to B_{MSY}/k = 0.34)
σ_{ϵ}	Productivity process error	prior mode 0.15, σ = 0.2
σ_I	CPUE observation error	fixed = 0.15
r	Intrinsic population growth rate	Derived from other parameters: $r = \frac{MSY}{k}(p+1)^{1/p}$
q	Catchability	Analytical solution described in text
ϵ	Productivity deviate(t)	Latent variable
B	Biomass(t)	Initial biomass assumed unfished (=k)
C	Catch(t) in mass	Assumed known without error (C_{max}= mean of 5 highest years of catch)
I	CPUE(t)	Annual abundance index (aggregated over seasons and regions)
t	Time index (years)	

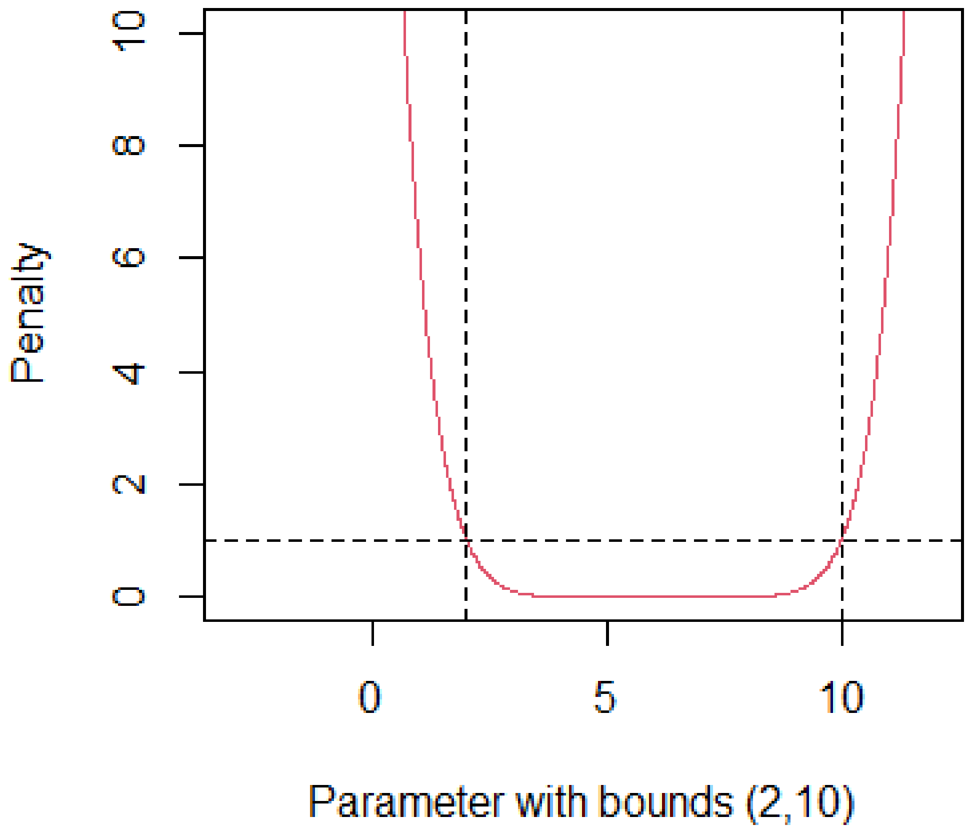
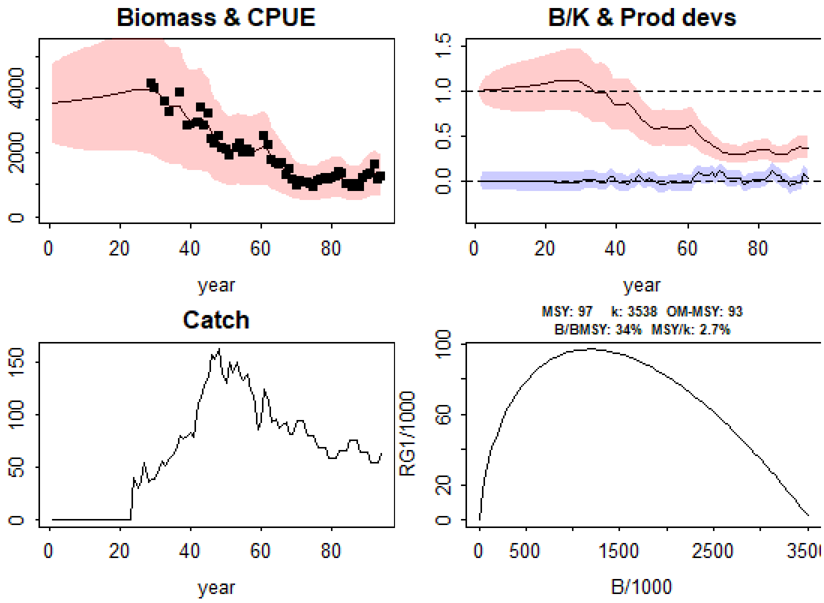
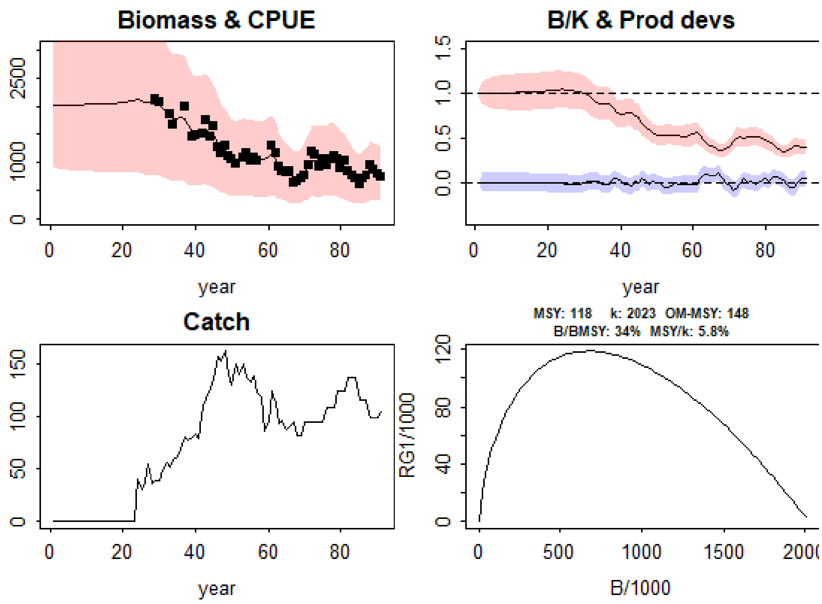


Figure 1. The ad hoc U-shaped penalty applied to constrain the k parameter (arbitrary parameter units for illustration). The y-axis is log-likelihood units. The function is simply a transformation of x^8, such that the penalty is one likelihood unit on each of the "bounds".



A) Operating Model MSY = 93Kt.



B) Operating Model MSY = 148Kt.

Figure 2. Examples of PTRE model fitting for contrasting bigeye OMs, following 20 years of projections (t=70 corresponds to 2020). Shaded regions indicate +/- 1 SE. Bottom right panels in each quadrant show the estimated production function, and compare the PTRE estimate of MSY with the Stock Synthesis value from the underlying OM.

4 Candidate MP Evaluation Results

The PTRE model was tested in conjunction with the default Harvest Control Rule (HCR) used in previous model-based MPs (Figure 3), together forming the candidate MP referred to as TMB.B2. TMB.B2 is compared with three other candidate MPs (as previously applied in Kolody et al 2020): Constant Catch (C.B2), a CPUE-based MP that aims for a CPUE target (D.B2)), and the original model-based MP which uses an observation error Pella-Tomlinson model (M.B2). M.B2 and TMB.B2 use the same Harvest Control Rule. MP evaluation conditions included:

- Evaluated against the most recent reference case bigeye OM – OmRefB20.1 (500 realizations based on a random sample of a fractional factorial grid of 72 Stock Synthesis models)
- All MPs were tuned to bigeye tuning objective 2 identified by the 2019 TCMP (60% probability of being in the Kobe green zone over the period 2030-2034)
- TAC is set every three years (and held constant), starting in 2021 (allocations set by the recent historical catch distribution)
- TAC changes were capped at +/- 15% per setting

The 4 MPs are compared in Figure 4 - Figure 10, using the standard TCMP MP evaluation outputs. These tests suggest that the initial application of the PTRE MP provides the most attractive MP behaviour, with less tendency to overshoot the biomass target in the long-term (Figure 4), and a considerably lower probability of exceeding biomass and fishing mortality limits (Figure 4, Figure 5). Evaluation speed is typically faster than the simpler observation error models that we have used to date (because the latter were implemented directly in R, used the finite difference gradient approximation rather than analytical derivatives, and required a systematic grid search of starting parameters).

These initial results are very promising, however, we would be reluctant to make definitive conclusions at this time. A concerted effort to develop the best version of any of these MP classes is still pending further clarification about objectives from the Commission/TCMP (e.g. a unique tuning objective has not yet been agreed, and it has been recognized that post-2035 MP behaviour is of secondary importance).

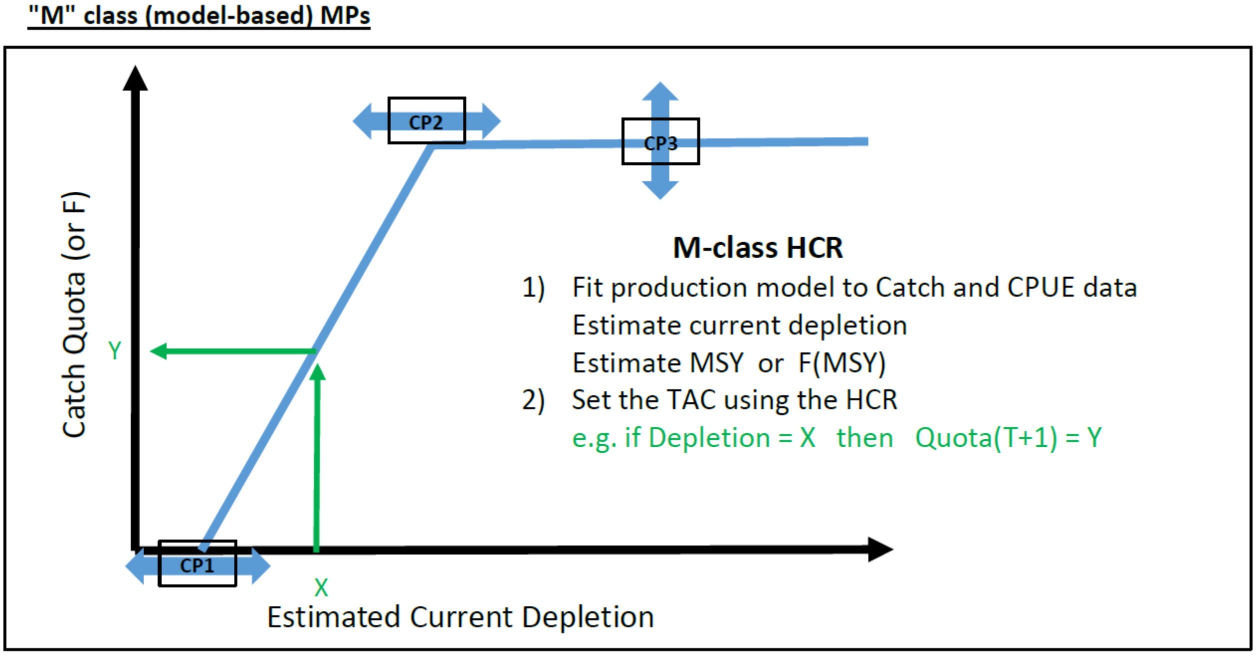


Figure 3. Cartoon representation of the "model-based" MPs (PTRE and the original observation-error PT model). In this case, control parameter CP3 was adjusted to tune the MP to achieve 60% probability of being in the Kobe green zone over the period 2030-2034; CP1 and CP2 were fixed at 0.1 and 0.4 respectively (depletion relative to unfished).

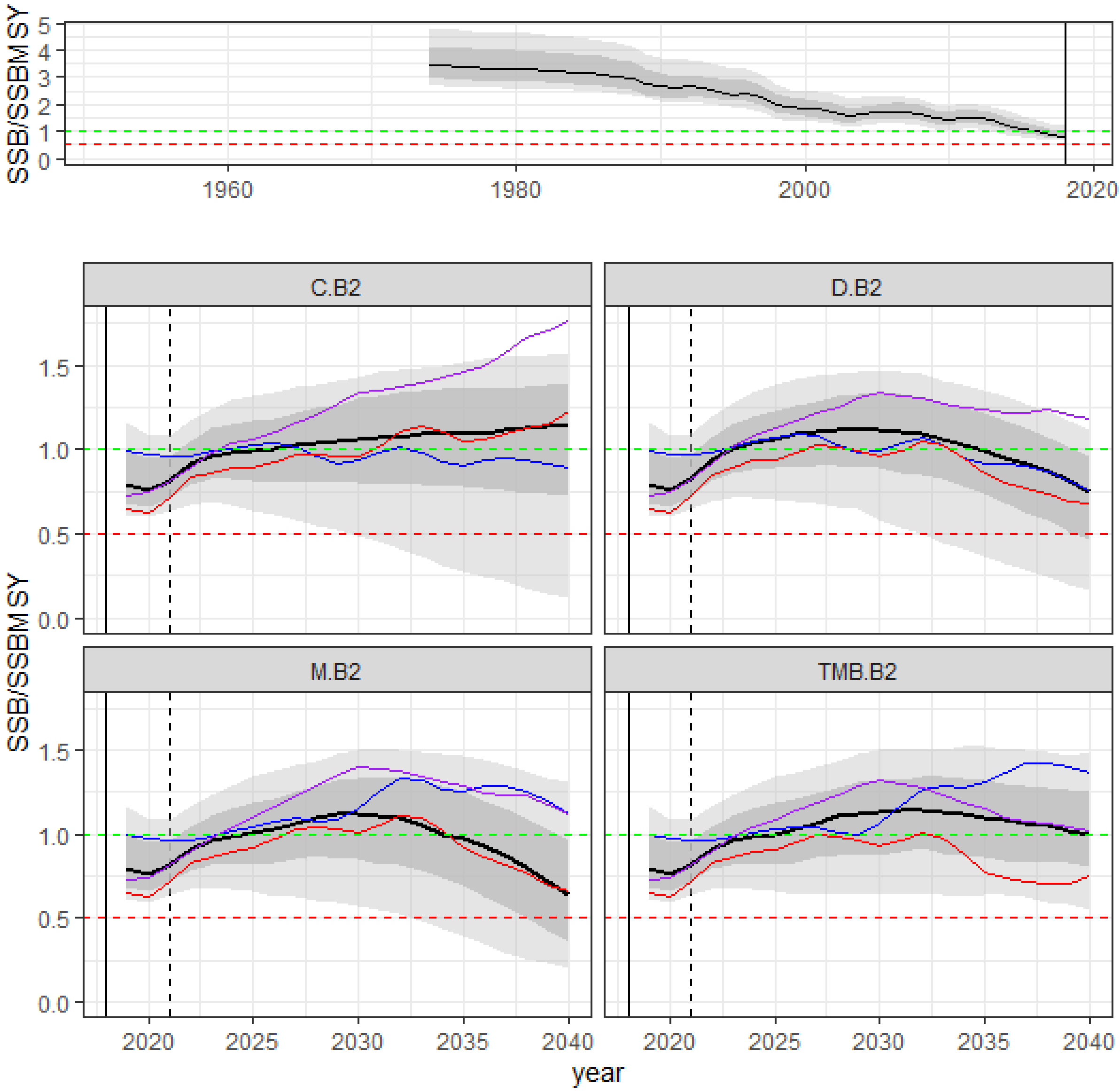


Figure 4. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Time series of relative spawning stock size for the candidate MPs. The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that the variability of any individual trajectory will generally exceed the median.

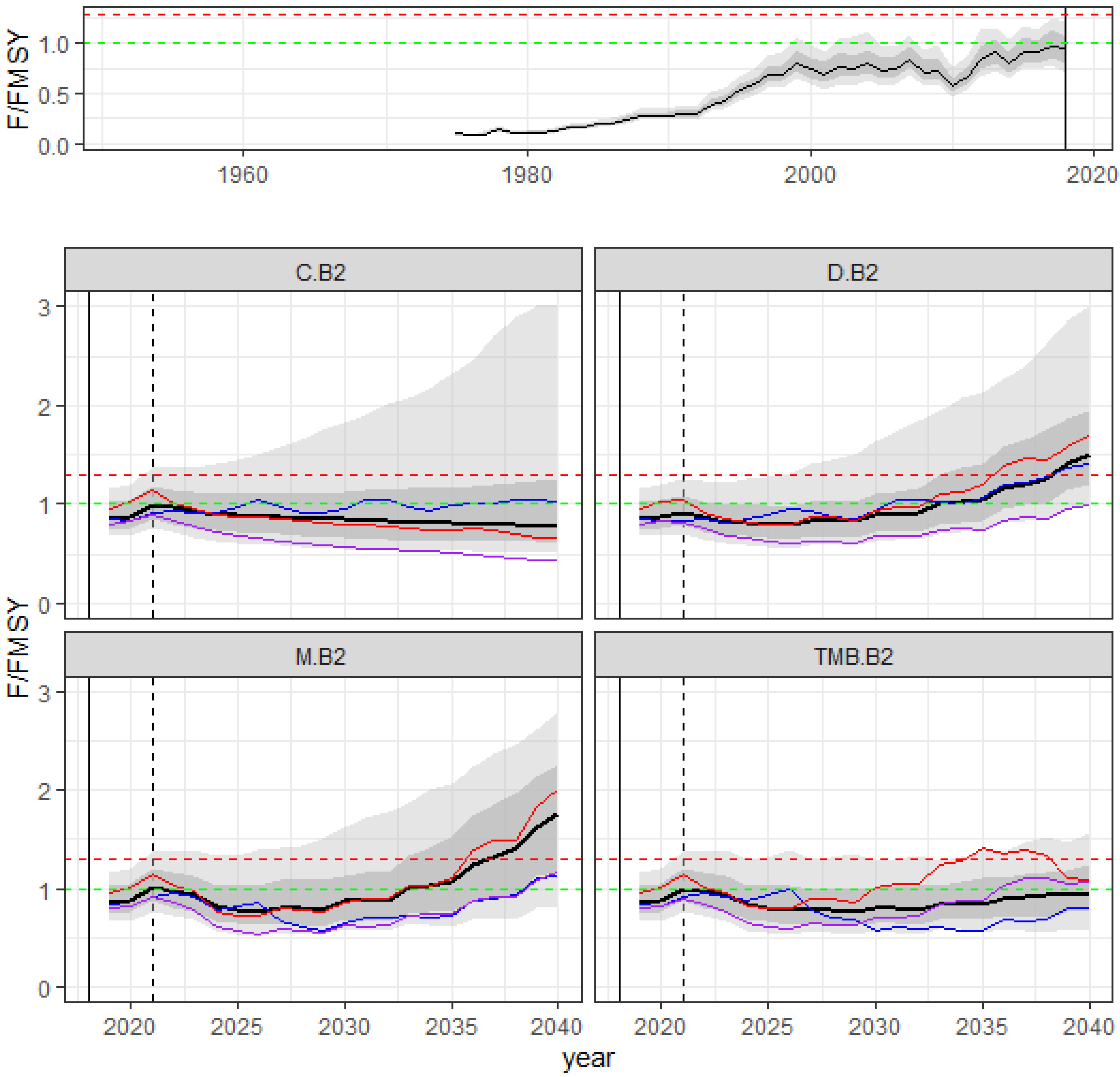


Figure 5. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Time series of fishing intensity (Upper bound truncated at F = 3) for the candidate MPs. The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that the variability of any individual trajectory will generally exceed the median.

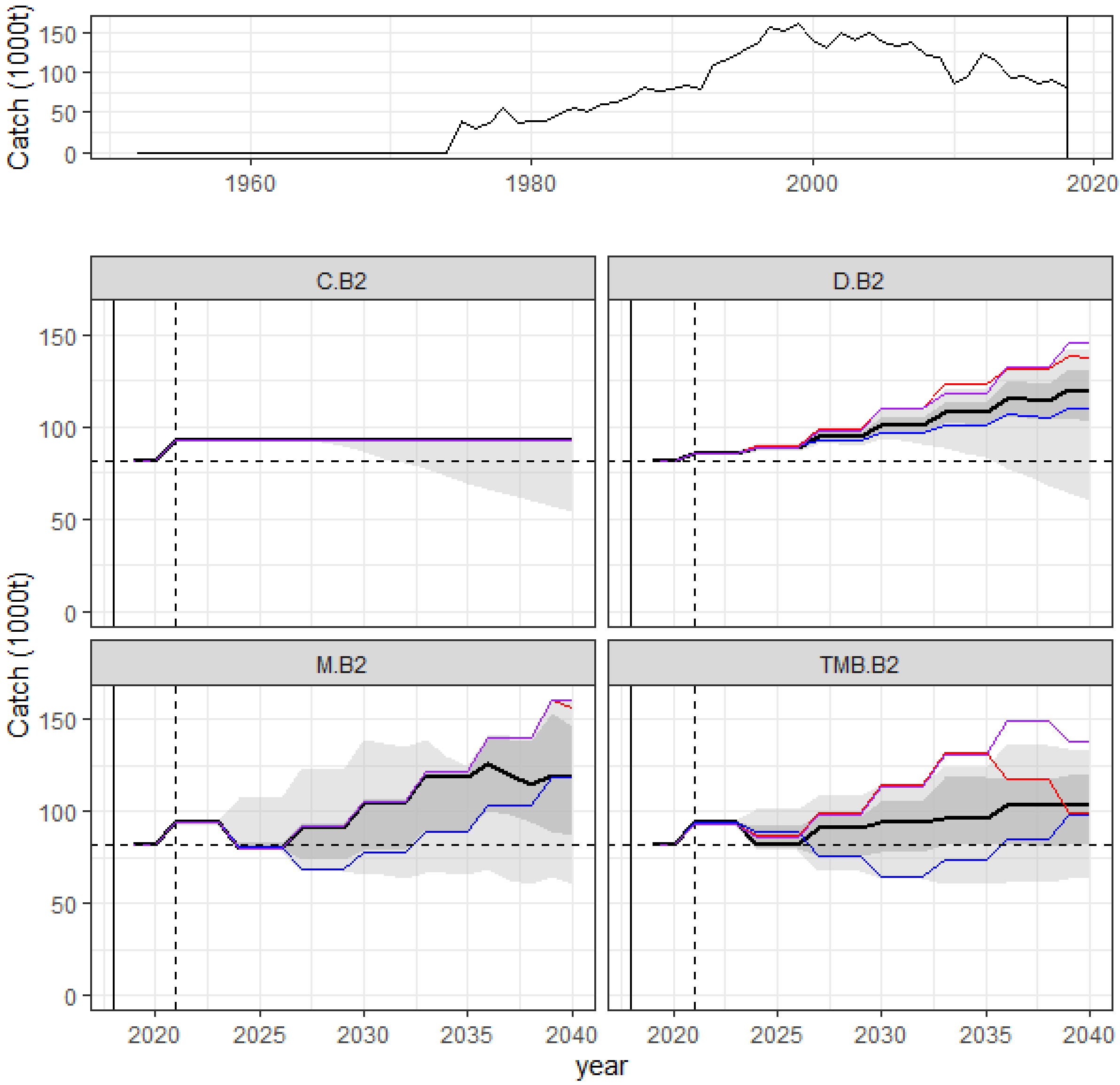


Figure 6. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Time series of catch for the candidate MPs. The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2018) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that the variability of any individual trajectory will generally exceed the median.

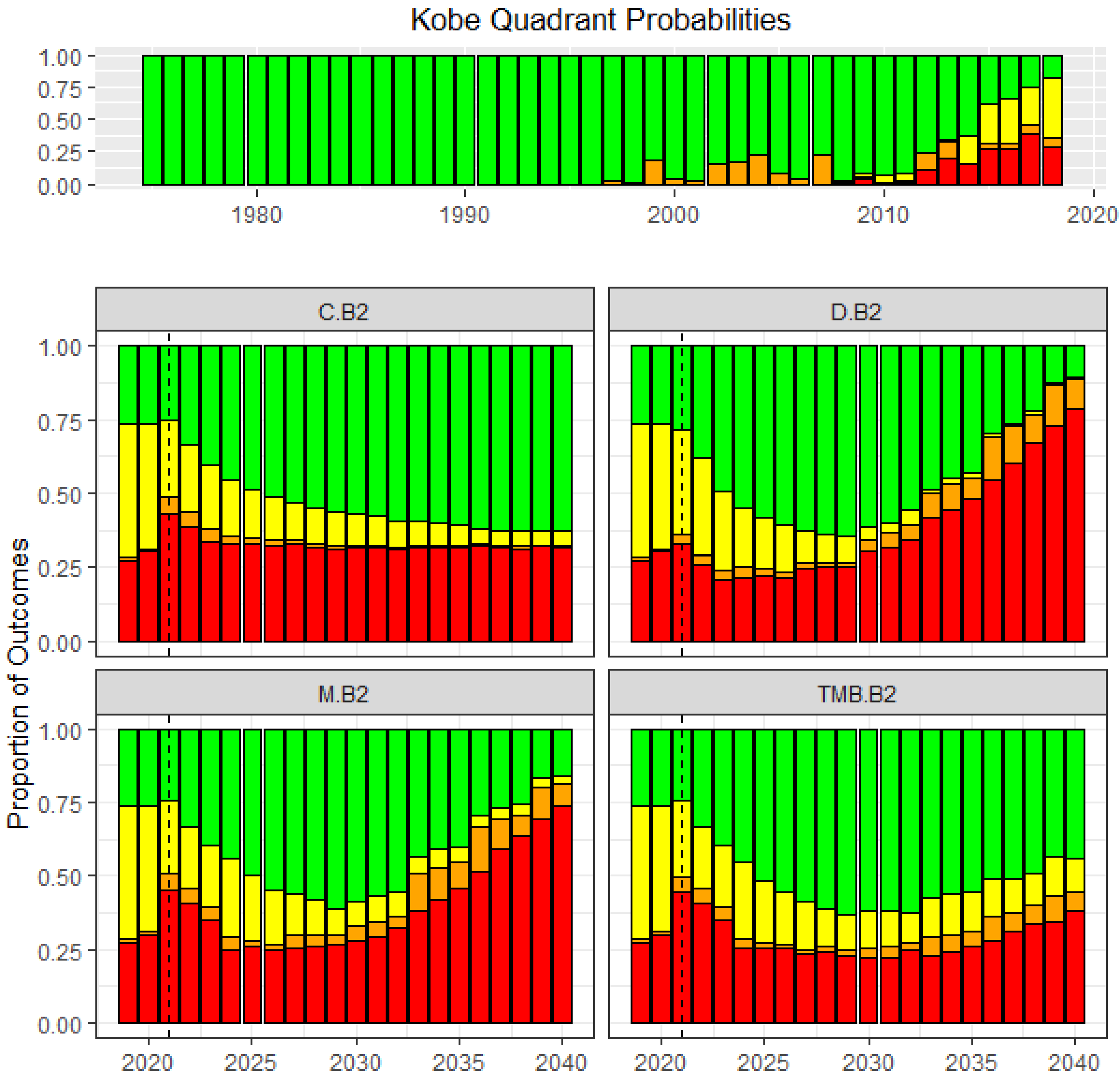


Figure 7. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Proportion of simulations in each of the Kobe quadrants over time for each of the candidate MPs. Historical estimates are included in the top panel. The lower panels are projections, with the first MP application indicated by the broken vertical line (2021).

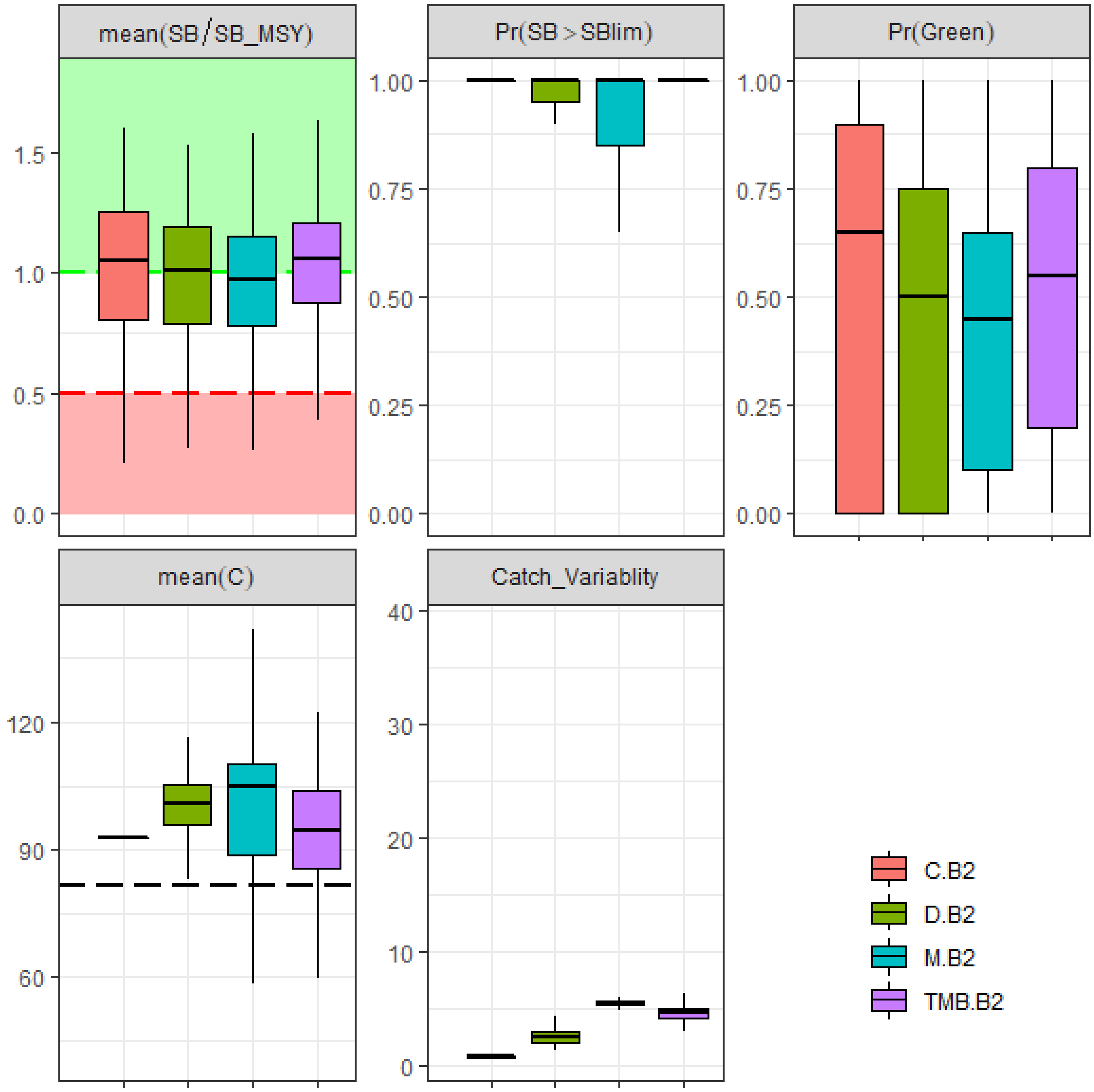


Figure 8. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Boxplots compare candidate MPs with respect to key performance measures averaged over the period 2021 - 2040. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers represent 10th - 90th percentiles. Red and green horizontal lines represent the interim limit and target reference points for the mean SB/SBMSY performance measure. The horizontal dashed black line is 2018 catch.

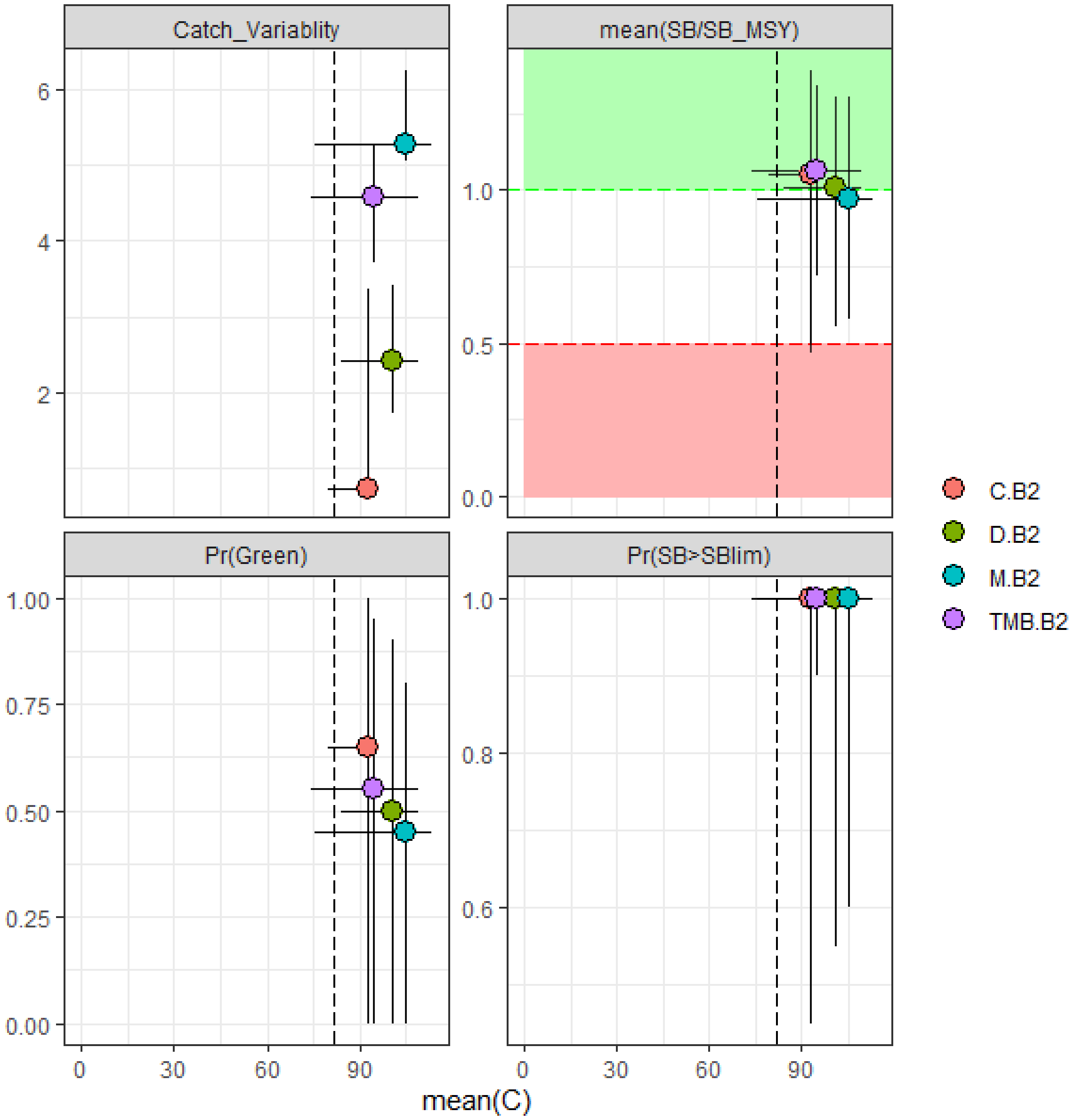


Figure 9. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Trade-off plots comparing candidate MPs with respect to catch on the X-axis, and 4 other key performance measures on the Y-axis, each averaged over the period 2019 - 2038. Circle is the median, lines represent 10th-90th percentiles. Red and green horizontal lines represent the interim limit and target reference points for the mean SB/SBMSY performance measure. The dashed vertical black line is 2016 catch.

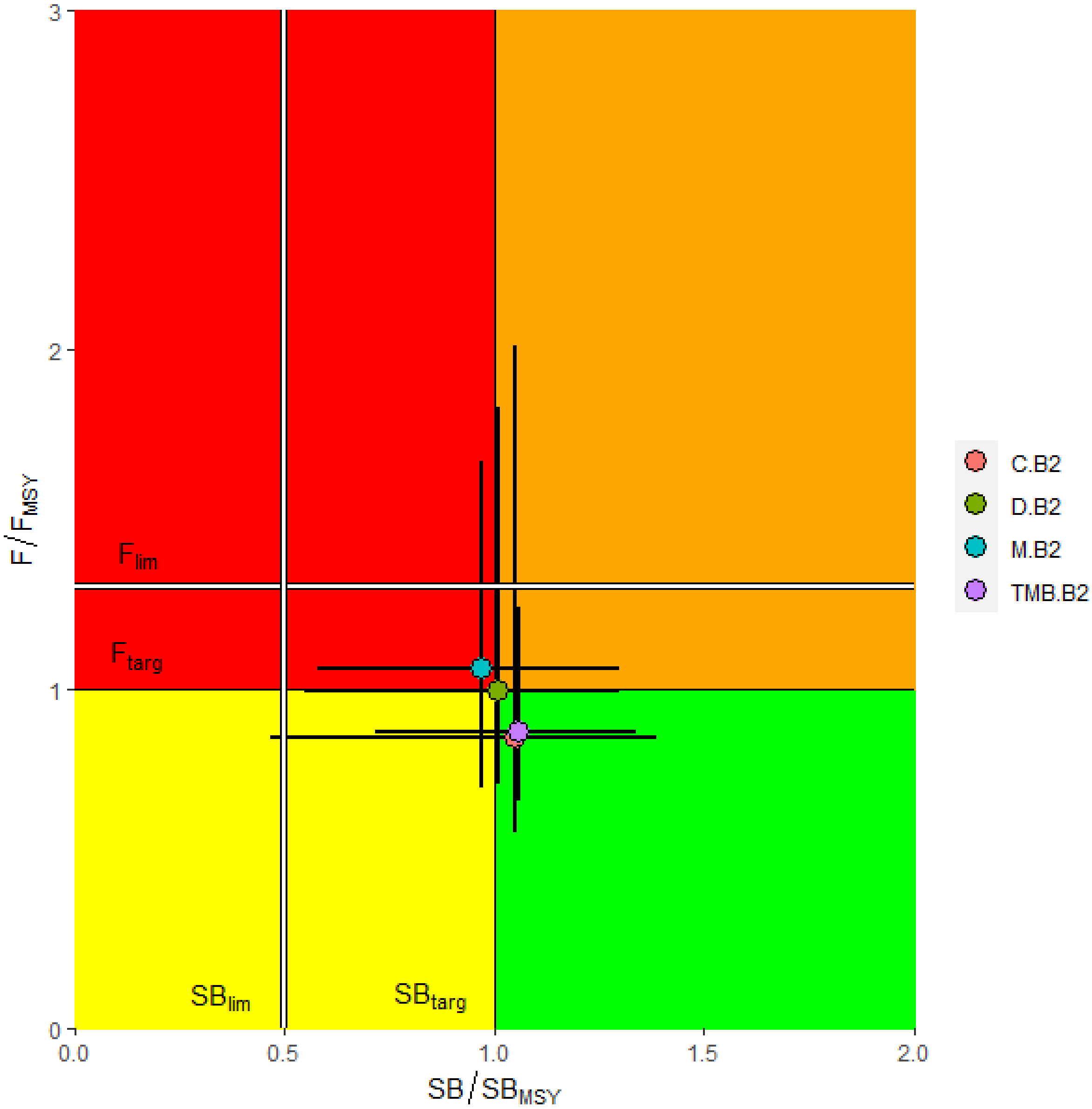


Figure 10. MP evaluation summaries from the Bigeye reference set OM OMrefB20.1. Kobe plot comparing candidate MPs on the basis of the expected 20 year average (2019-2038) performance. Circle is the median, lines represent 10th-90th percentiles.

5 Discussion

The PTRE-based MP is theoretically attractive, it appears to be numerically stable, very fast in the testing to date, and it is capable of distinguishing among OM productivity levels over a medium time horizon. We would hope that the introduction of process error should improve the ability of these models to make useful inferences about complicated age-structured populations in an MP context, but it is premature to conclude that PTRE-based MPs are definitively better. Conversely, there is also no obvious reason not to move to this platform as the primary basis for further developing model-based MPs.

One of the obvious concerns in the model (and true of any of the MPs tested to date) is the need to fix or constrain some parameters in the interest of numerical stability, particularly in automated fitting. The priors and bounds are somewhat arbitrary, and we have not tested how sensitive the MP results are to these priors. However, the MPs have been tuned to achieve the tuning objectives requested by the TCMP. As such, the effects of informative priors represent part of the formal MP specification. During MP implementation, it might be tempting to alter particular priors or constraints, because they are not strictly required to achieve convergence, or there may be a strong theoretical arguments or new information as to why there is a better option. But this sort of ad hoc modification should be discouraged - it goes against the spirit of MP pre-specification and would be expected to alter the MP performance with unknown consequences.

We expect to undertake further development of this approach and present results to the TCMP 2021. As always, we welcome suggestions for further development, and encourage others to consider developing their own competing MPs. The code for implementing this MP will be available on github in the near future.

References

Kolody, D, Day, J, Jumppanen, P. 2020. Indian Ocean bigeye tuna Management Procedure Evaluation update March 2020. Report prepared for the Indian Ocean Tuna Commission Informal Management Strategy Evaluation workshop 2020. IOTC–2020–WPTTXX-XX.

Kristensen, K, Nielsen, A, Berg, CW, Skaug, HJ, and Bell, BM. 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Softw. 70(5):1-21.

Winker, H, Carvalho, F, Kapur, M. 2018. JABBA: Just Another Bayesian Biomass Assessment. Fisheries Research 204: 275-288. https://doi.org/10.1016/j.fishres.2018.03.010.

CONTACT US

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

AT CSIRO WE SHAPE THE FUTURE

We do this by using science to solve real issues. Our research makes a difference to industry, people and the planet.

As Australia's national science agency we've been pushing the edge of what's possible for over 85 years. Today we have more than 5,000 talented people working out of 50-plus centres in Australia and internationally. Our people work closely with industry and communities to leave a lasting legacy. Collectively, our innovation and excellence places us in the top ten applied research agencies in the world.

WE ASK, WE SEEK AND WE SOLVE

FOR FURTHER INFORMATION

CSIRO Oceans and Atmosphere
Dale Kolody
t +61 3 6232 5121
e dale.kolody@csiro.au
w https://www.csiro.au/en/Research/OandA