CCSBT-OMMP/1806/05 (Rev.1)
(OMMP Agenda Item 3)

Oceans & Atmosphere
www.csiro.au

Initial MP structure and performance

Rich Hillary, Ann Preece, Campbell Davies

12 June 2018

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7004, Tasmania, Australia.

Contents

Abstract

We document initial MP structures and their performance on the currently defined revised tuning objectives. At this stage, performance is only shown for the MPs tuned to the objectives for the reference set of OMs. We employ the use of: long-line CPUE, the gene tagging data, as well as the close-kin mark-recapture data (both empirically and in a model-based setting). The four candidate tuning objectives (median TRO of 25%, 30%, 35% and 40% of the unfished state by 2035) were obtained for all but one CMP and showed clear differences in likely performance across objectives and across CMPs. MPs which employed target-limit style rules performed better than those which used only trend information. MPs that used CKMR data (empirically or model-based) resulted in generally higher average TACs and lower variability and incosistency for the 30% and 35% tuning objectives. For the 25% tuning objective all MPs act to rapidly increase the TAC and cause TRO decreases post-2035; for the 40% objective all MPs act to rapidly decrcease the TAC causing a a clear disparity between the TAC and the effective replacement yield at the target objective by 2035.

1 Background

The changes required to the SBT OM projection code (`sbtproj.tpl`) were outlined in [1], which built upon ideas for data generation and MP index construction first explored in [2]. In this paper we first detail the types of data and forms of MP we explored, then their performance when tuned to the four tuning objectives using the reference set of OMs.

2 Data and forms of the CMPs

We use three data sources, thought not all three at the same time, in the various CMPs:

1. Japanese long-line CPUE
2. The gene tagging (GT) data (matches and number of comparisons)
3. The close-kin mark-recapture (CKMR) data (matches and number of comparisons)

2.1 MPs using the CPUE & GT data

The usual trend term has been employed previously [3]:

$$TAC_{y+1} = TAC_y\,(1 + k\lambda_y)$$

where λ is the log-linear trend in the index under consideration could be modified as follows:

$$\begin{aligned}
\Delta_y^{\text{cpue}} &= 1 + k(1+\nu)\lambda_y \quad \text{if} \quad \lambda_y \leq -\tilde{\lambda}, \\
\Delta_y^{\text{cpue}} &= 1 \quad \text{if} \quad \lambda_y \in (-\tilde{\lambda}, \tilde{\lambda}), \\
\Delta_y^{\text{cpue}} &= 1 + k(1-\nu)\lambda_y \quad \text{if} \quad \lambda_y \geq \tilde{\lambda},
\end{aligned}$$

where:

- k is the usual ”gain” parameter
- $\nu \in (-1, 1)$ is an asymmetry parameter that can act more/less strongly on positive/negative trends depending on its relative sign
- $\tilde{\lambda}$ is a ”threshold” trend level: given the time-frame employed to estimate the trend, τ years say, and the CV in the index it would be the lower bound at which the estimates of trend will basically

be driven by the noise. The crucial point being we can construct some simple conditions to set this value given τ and the index CV and auto-correlation

As a simple motivating example for how to estimate $\tilde{\lambda}$ consider $\tau = 7$ years and a CV of 0.2 and autocorrelation of around 0.25 (basically, the settings/values in the Bali Procedure and current OM). If our criterion of performance is the probability that we can estimate the correct **sign** in the trend (i.e. positive or negative) then to reduce the proportion of false negatives to 0.05 we would choose $\tilde{\lambda} = 0.07$ i.e. don't bother with anything less than a trend of $\pm 6\%$ as it has a good chance of being spurious and noise driven. Obviously, a critical false positive rate of 0.05 is fairly stringent - the main point being we can define a critical value of this false positive rate and, given the settings of the OM and the MP, calculate a subsequent value of $\tilde{\lambda}$. Say we wanted to get the right trend 3 times out of 4: the minimum "retained" trend estimate would then be $\tilde{\lambda} = 0.03$.

We already included a simple GT-driven log-linear trend style MP in the `testGTinMP` example uploaded it `github`, so we explore some other options here. The main advantage that the GT data have over the aerial survey is that they are absolute, or should be if everything we think we know about spatial stock structure and recruitment dynamics is about right. The lowest estimated cohorts were (age 0 fish from 2000 to 2002), and by age 2 fish the mean abundance was around 700,000 animals. These recruits very likely caused a rapid decrease in the CPUE in the mid 2000s as they entered the long-line fishery, and left a clear hole in the length frequency data as well as signals in both the tagging and SAPUE data. They represent reasonably well-estimated levels of age 2 fish below which we know bad things can happen, suggesting themselves as a potentially useful limit level - now we have in principle estimates of absolute age 2 fish from the GT program.

If we are to use the absolute nature of the GT data then the general principles would be something like:

- Below the limit level the HCR should act strongly to reduce the TAC
- Above the limit level and up to some pre-specified upper level the GT part of the HCR maintains the TAC where it is
- If recent mean recruitment has been suitably elevated (i.e. above a pre-specified level) then the HCR should act to increase the TAC

To calculate the recent mean age 2 abundance from the GT data consider a weighted moving average approach:

$$\bar{N}_{y,2} = \sum_{i=y-1-\tau}^{y-2} \omega_i \widehat{N}_{i,2}$$

where ω_i is a weighting proportional to the number of matches used to produce the GT estimate $\widehat{N}_{i,2}$ (basically inverse variance weighting). The 2 year delay between having the estimate and what year it actually refers to is factored into the calculation. The multiplier for the GT part of the HCR would then be:

$$\begin{aligned}
\Delta_y^{\text{gt}} &= \left(\frac{\bar{N}_{y,2}}{N_{\text{low}}}\right)^{\alpha} \quad \text{if} \quad \bar{N}_{y,2} \leq N_{\text{low}}, \\
\Delta_y^{\text{gt}} &= 1 \qquad\qquad\quad \text{if} \quad \bar{N}_{y,2} \in (N_{\text{low}}, N_{\text{high}}), \\
\Delta_y^{\text{gt}} &= \left(\frac{\bar{N}_{y,2}}{N_{\text{high}}}\right)^{\beta} \quad \text{if} \quad \bar{N}_{y,2} \geq N_{\text{high}}
\end{aligned}$$

with N_{low} the limit level and N_{high} the upper level at where TAC inrceases are permitted. The exponents α and β are to allow for differential responses depending on the situation: we might expect $\alpha > 1$ as we would want to act strongly on poor recruitment levels; alternatively we might have $\beta < 1$ so that TAC inreases based on increased recruitment are more modest, given increased recruitment does not guarantee the TRO will increase (especically if we increase the Fs they experience as they mature).

2.2 MPs using the CKMR data

From the CKMR data we can derive an empirical index (be it POP, HSP or combined), I_y^{ck} [2], or an actual estimate of the TRO and mean adult Z from the model-based framework. The empirical CKMR indices can clearly capture trends in the true TRO: as it goes up/down, then the number of kin matches (conditional on the sampling regime) should generally go down/up. What complicates this simple interpretation of the trend in the overall number of matches (and what it is telling us) are the following:

1. **Changes in adult "recruitment"**: increases or decreases in the number of sub-adults who survive to become reproductively active due to variations in actual age-0 recruitment and changes in juvenile and sub-adult mortality
2. **Changes in adult mortality** changes in either the TAC or trends in the adult recruitment will clearly alter the harvest rates and, even for an assumed time-invariant M, by extension total mortality/survival

The short-to-medium term effect of these changes can be seen in the empirical indices, driven by the large 2013 recruitment beginning to move its way into and through the adult population. Strong changes in age structure and mortality *have* to have a strong effect on the POP and HSP probabilities . If we could somehow use a model to tease these effects out of the data, there is the *potential* to estimate trends in TRO that correlate far better with the true TRO than the empirical indices can. We could also estimate additional variables of use (like adult recruitment and total mortality).

This is similar to the approach taken when developing the population model behind the Bali Procedure. The specifics would be quite different given the radical difference in the input data (CPUE and aerial survey *vs.* POP and HSP data), but the idea is the same: reduce the variability and bias in the indices by accounting for the underlying population dynamic drivers of the fluctuations causing those biases. We explored a simple age-structured population model:

$$
\begin{aligned}
N_{y,a_{\min}} &= \bar{R}\exp\left(\epsilon_y - \sigma_R^2/2\right), \\
\epsilon_y &\sim N(0,\sigma_R^2), \\
N_{y+1,a+1} &= N_{y,a}\exp\left(-Z_{y,a}\right) \qquad a \in (a_{\min}, a_{\max}), \\
N_{y+1,a_{\max}} &= N_{y,a_{\max}-1}\exp\left(-Z_{y,a_{\max}-1}\right) + N_{y,a_{\max}}\exp\left(-Z_{y,a_{\max}}\right), \\
Z_{y,a} &= Z_y \qquad a \le 25, \\
Z_{y,a} &= Z_y + \frac{a-25}{a_{\max}-25}\left(Z_{a_{\max}} - Z_y\right) \qquad a \in [26, a_{\max}], \\
Z_y &= \frac{Z_{\max}e^{\chi_y} + Z_{\min}}{1+e^{\chi_y}}, \\
\chi_{y+1} &= \chi_y + \zeta_y, \\
\zeta_y &\sim N(0,\sigma_\chi^2), \\
TRO_y &= \sum_a N_{y,a}\varphi_a
\end{aligned}
$$

The main ideas behind the population model are:

- It is only for adults, with a minimum age ($a_{\min}$) defining where fish being to be considered possible reproductively active adults

Parameter	Value
$a_{\min}$	6
$a_{\max}$	30
σ_R	0.2
σ_χ	0.1
$Z_{\min}$	0.05
$Z_{\max}$	0.4
$Z_{a_{\max}}$	0.5
$\mu_{\chi_{\rm init}}$	-1.38
$\sigma_{\chi_{\rm init}}$	0.2
$q_{\rm hsp}$	0.9

Table 2.1: Settings for CKMR MP population model

- Recruitment to the adult stock is assumed to vary randomly around a mean value
- Total mortality, $Z_{y,a}$, has a well-defined age and time structure: between $a_{\min}$ and 25 it is fixed by age, but has a random walk time component
- Between 25 and $a_{\max}$ total mortality rises linearly to a pre-specified maximum $Z_{a_{\max}}$
- The time-varying Z_y variable is constrained to be between $Z_{\min}$ and $Z_{\max}$
- Equilibrium conditions are assumed for the first year age-structure
- A plus group is used for ages above the final true age class, $a_{\max} - 1$
- The total reproductive output (TRO) is defined using a time-invariant ogive, φ_a

The estimate parameters of this model are:

1. The mean adult recruitment, $\bar{R}$
2. The adult recruitment deviations, ϵ_y
3. The initial value, $\chi_{\rm init}$, that "starts" the random walk for Z_y (with an associated normal prior mean and SD)
4. The random walk deviations ζ_y

This is basically the same number of parameters estimated in the Bali Procedure population model [3], so we are not talking about a large number of model parameters, and many of them are going to be constrained deviation parameters. The likelihood model for the POP and HSP data are basically the same as those used in the SBT OM, but where M_a and the harvest rates are replaced by $Z_{y,a}$ to estimate cumulative survival in the HSP likelihood.

Using the same fixed TAC (17,647t) projection we used to explore the utility of the empirical CKMR indices we undertook a full simualtion evaluation of the simplified CKMR estimation model with the settings as laid out in Table 2.1.

Figure 2.1 outlined the estimation performance of the CKMR MP population model, given the settings in Table 2.1 and assumed a future fixed TAC of 17,647t. Clearly, while under-estimating the much later TRO and mean adult Z, it captures the true trends very well. The median correlation between the estimated and true TRO and mean adult Z was 0.97 and 0.86, respectively.

The values for the fixed and prior parameters in the CKMR model are clearly informed by the current OM. The value of the mimimum age, 6, is chosen given the youngest age of a detected parent in the POPs. The maximum age is the maximum age assumed in the OM, and the total mortality at the maximum age is basically the estimated level (around 0.5). Average adult Z at the start of the CKMR data (around the early 2000s) is somewhere around the 0.15 level and the prior mean and SD of $\chi_{\rm init}$ are chosen to reflect this, albeit fairly weakly given the prior variance. The minimum value of Z for the ages less than 25 is fixed at 0.05 which is the lowest level of M_{10} assumed in the OM and the maximum Z is assumed to be lower

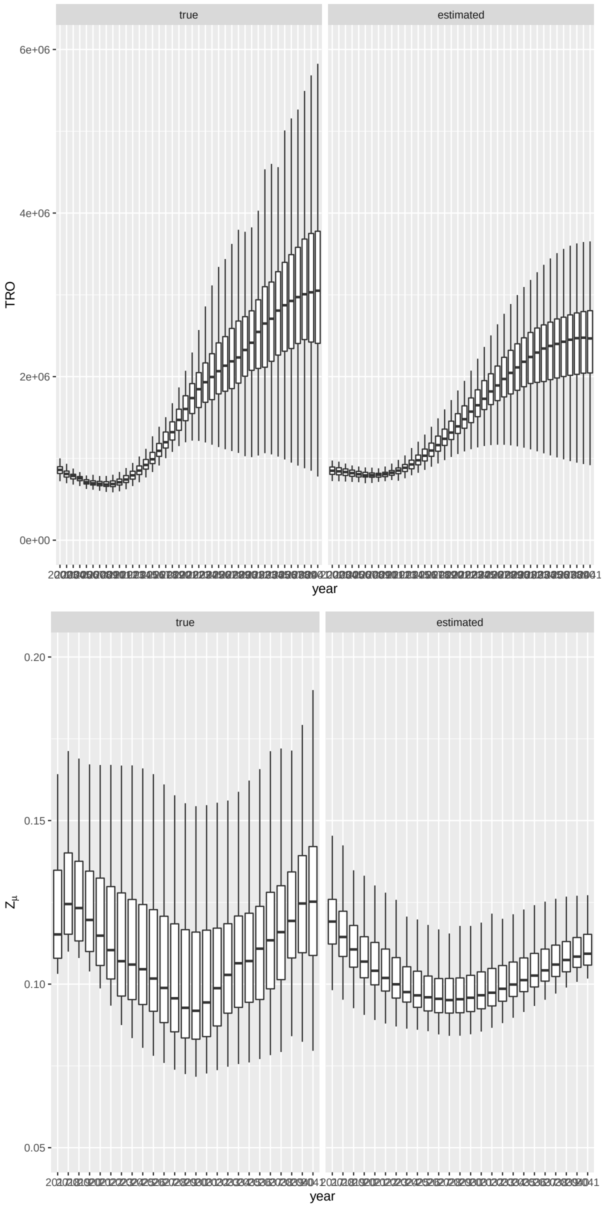

Figure 2.1: *True (left) and estimated (right) TRO (top) and mean adult Z (bottom) for the CKMR MP population model outlined in Table 2.1*

than the Z at the maximum age. The way that Z increases to the maximum level is also driven by the OM: permitted to increase linearly from age 25 to the maximum age. The value of q_{hsp} is also informed heavily by the OM estimates at around 0.9. So the model rests heavily on things we think we now know about some of these key variables given the inclusion of the historical CKMR data (which is valid in an MP context), but does not assume to know things it simply could not about future parameters and data (which is arguably the most important part of using models in an MP context).

The main thing we do know about where we want the future TRO to go, relative to the present, is that we want it to go up. So, given a reference level of the empirical index centered around the present, $\tilde{I}$, the target level of the TRO we wish to rebuild the stock to will be some multiple of this, $\gamma > 1$. This formulation lends itself to the construction of a target-type HCR element for the CKMR data:

$$
\begin{aligned}
\Delta_y^{\mathrm{ck}} &= (1 - w^{\mathrm{ck}}) + w^{\mathrm{ck}} \frac{I_{y-4}^{\mathrm{ck}}}{\gamma \tilde{I}(1-\delta)} && \text{if} \quad I_{y-4}^{\mathrm{ck}} \leq \gamma \tilde{I}(1-\delta), \\
\Delta_y^{\mathrm{ck}} &= 1 && \text{if} \quad I_{y-4}^{\mathrm{ck}} \in (\gamma \tilde{I}(1-\delta), \gamma \tilde{I}(1-\delta)) \\
\Delta_y^{\mathrm{ck}} &= (1 - w^{\mathrm{ck}}) + w^{\mathrm{ck}} \frac{I_{y-4}^{\mathrm{ck}}}{\gamma \tilde{I}(1+\delta)} && \text{if} \quad I_{y-4}^{\mathrm{ck}} \geq \gamma \tilde{I}(1+\delta),
\end{aligned}
$$

Summarising the key control parameters:

1. The term $w^{\mathrm{ck}} \in (0, 1)$ is a weighting term: closer to zero and there is a lot of ”inertia” in this part of the HCR and TACs will move little, regardless of trends; closer to 1 and the HCR moves to being almost a pure target style HCR driven only by the distance of the current CK index relative to the target, $\gamma \tilde{I}$
2. The term $\delta \in (0, 1)$ is a ”buffer” term, so that the HCR is instructed to do nothing if the current index above or below the target by a factor of $(1 + \delta)$ or $(1 - \delta)$, respectively
3. The term $\tilde{I}$ is a suitably defined estimate of the ”recent” index -best calculated for the period for which we have actual CKMR data
4. The term γ is the index ”inflation” value: the amount by which we want to increase the CKMR index from the $\tilde{I}$ level. For example, if we have a reasonably well correlated 1:1 relationship between the CKMR index and TRO, then γ would be the tuning target TRO depletion level divided by the current TRO depletion level. In practice this terms seems well suited to being a tuning parameter, given an array of possible future TRO tuning targets

With the model-based option we will also obtain a potentially usable estimate of mean adult Z, and obviously using the trend in this data may also prove fruitful. Initial runs using the initial CKMR part of the HCR in the model-based context actually showed strange (at first) but after consideration (given how well the model tracks the trends) obvious behaviour. The model accurately estimates when the true TRO is above or below the target level $\gamma \tilde{I}$, and cuts the TAC even if the trend in TRO is going up. This causes it to have to then increase the TAC rapidly once the TRO approaches the target level, and some oscillatory behaviour can emerge.

A modification to the HCR for the model-based CKMR MP was required so that it was able to track the trend in the TRO when *below* the target level (and reduce it if it is not going up fast enough). The minimum expected (log-scale) rate of increase is actually fairly simple to define. Given a time-frame, T, and the inflation factor γ, then $\tilde{\lambda}^{\mathrm{ck}} = \gamma^{T^{-1}} - 1$, and we chose the time-frame as 2007 to 2035.

The modified HCR was defined as follows:

$$
\begin{aligned}
&\Delta_y^{\text{ck}} = 1 + k^{\text{ck}}\left(\lambda_y^{\text{ck}} - \tilde{\lambda}^{\text{ck}}\right) \quad \text{if} \quad I_{y-4}^{\text{ck}} \le \gamma\tilde{I}(1-\delta) \quad \text{and} \quad \lambda_y^{\text{ck}} < \tilde{\lambda}^{\text{ck}},\\
&\Delta_y^{\text{ck}} = 1 \quad \text{if} \quad I_{y-4}^{\text{ck}} \le \gamma\tilde{I}(1-\delta) \quad \text{and} \quad \lambda_y^{\text{ck}} \ge \tilde{\lambda}^{\text{ck}},\\
&\Delta_y^{\text{ck}} = 1 \qquad\qquad\qquad\qquad \text{if} \quad I_{y-4}^{\text{ck}} \in (\gamma\tilde{I}(1-\delta), \gamma\tilde{I}(1-\delta))\\
&\Delta_y^{\text{ck}} = \left((1-w^{\text{ck}}) + w^{\text{ck}}\frac{I_{y-4}^{\text{ck}}}{\gamma\tilde{I}(1+\delta)}\right)^{\beta^{\text{ck}}} \quad \text{if} \quad I_{y-4}^{\text{ck}} \ge \gamma\tilde{I}(1+\delta) \quad \text{and} \quad \lambda_y^{\text{ck}} < 0,\\
&\Delta_y^{\text{ck}} = \left((1+k^{\text{ck}}\lambda_y^{\text{ck}})(1-w^{\text{ck}}) + w^{\text{ck}}\frac{I_{y-4}^{\text{ck}}}{\gamma\tilde{I}(1+\delta)}\right)^{\beta^{\text{ck}}} \quad \text{if} \quad I_{y-4}^{\text{ck}} \ge \gamma\tilde{I}(1+\delta) \quad \text{and} \quad \lambda_y^{\text{ck}} \ge 0.
\end{aligned}
$$

In terms of naming conventions of MPs (descriptive ones at least) for CPUE we have:

1. **C1**: trend-type CPUE HCR
2. **C2**: target-type CPUE HCR

For GT data we have:

1. **GT1**: limit-type GT HCR
2. **GT2**: trend-type GT HCR

For CKMR data we have:

1. **CK1a**: empircal POP-based target HCR
2. **CK1b**: empircal HSP-based target HCR
3. **CK1c**: empircal POP+HSP combined target HCR
4. **CK2**: model-based (inc. POP and HSP) based
5. **CK2z**: as with **CK2** but including mean adult Z trend

So an MP file called `C1GT2CK1a.tpl` would include CPUE (trend), gene-tagging (limit) and a CKMR empirical POP index (target) HCR. For this initial phase of the work we explored seven CMPs:

1. C1GT1 (CPUE trend, GT limit): **rh1**
2. C1GT2 (CPUE trend, GT trend): **rh2**
3. C2GT2 (CPUE target, GT limit): **rh3**
4. C2GT2 (CPUE target, GT trend): **rh4**
5. C1GTCK1a (CPUE trend, GT trend, CKMR POP index): **rh5**
6. C1GTCK1c (CPUE trend, GT trend, CKMR POP+HSP index): **rh7**
7. C1GTCK2 (CPUE trend, GT trend, CKMR TRO estimate): **rh8**

with their short names used in graphical and other summaries in bold.

2.3 Overall formulation for TAC change

For CMPs using both CPUE and gene tagging data the TAC change would be as follows:

$$TAC_{y+1} = TAC_y \times \Delta_y^{\text{cpue}} \times \Delta_y^{\text{gt}},$$

and for CMPs using the CPUE, gene tagging and the CKMR data it would be defined by

$$TAC_{y+1} = TAC_y \times \Delta_y^{\text{cpue}} \times \Delta_y^{\text{gt}} \times \Delta_y^{\text{ckmr}}.$$

It is worth noting the slightly different manner in which "inertia" in the TAC is included in this formulation. It is actually embedded within the changes (the Δ's), not within the main TAC equation. This was done to make it possible to effectively introduce inertia into the CPUE and CKMR parts of the HCR but actively exclude it from the GT part of the HCR (so it was not impeded in reducing the TAC for low recruitments).

3 MP performance on the reference set

Each of the CMPs was tuned to the four tuning objectives, with a tolerance of 1% (i.e. the MP parameters were altered until the probability of meeting the objective was 0.5 ± 0.005). In terms of summary statistics we use the following initial set relating to TRO:

- $\mathbb{P}\left(TRO_{2035} > 0.2TRO_0\right)$: the previous MP tuning objective
- $\mathbb{P}\left(TRO_{2035} > TRO_{2017}\right)$: probability that the TRO in the tuning year is greater than the current level
- $\mathbb{P}\left(TRO_{2040} > TRO_{2035}\right)$: probability that the TRO five years after the tuning year is above that in the tuning year, to catch MPs which increase the TAC too high/fast when attaining the tuning objective and cause a future "undershoot"
- Log-linear trend in TRO from 2021 to 2035: to see if there are trajectories which actually on average go down from when the first TAC change occurs to when the tuning year is reached

and the following set for TACs:

- Mean TAC (across years) from 2021 to 2035
- AAV (from 2021 to 2035)
- Maximum TAC decrease (from 2021 to 2035)
- $\mathbb{P}\left(TAC_{r+3} < TAC_{r+2}\right)$ *if* $TAC_{r+2} > TAC_{r+1}$ *and* $TAC_{r+1} > TAC_r$, for the r^{th} TAC decision (default is currently $r = 1$)
- Mean (averaged over years and from 2021 to 2035) of the lower 10%ile in the TAC

The one perhaps unfamiliar tweak to these is for the TAC up/down statistic. The *expected* probability that this happens is merely the number of times the TAC goes up twice *and then goes down*, divided by the number of times it went up twice. For some tuning objectives and MPs, the TAC going up twice might be a very common occurrence, but where the TAC going down again is not quite as common. For others, the TAC going up twice might be *very* rare, but it coming down again is comparably quite common. For the first case we would have a low and accurate estimate of the up/down statistic; for the second case we would have a very high but very uncertain estimate that doesn't really tell is much at all. If we simply plot the expected probability of going up then down, without some measure of the uncertainty thereof, we are likely to make misguided inferences about MP performance on this statistic. To address this we used a simple Bayesian approach to plotting the likely distribution of this statistic, not just its expectation. For a given MP run, if n_1 is the number of runs where the TAC goes up twice, and n_2 is the number of runs where it subsequently goes down again, assuming an uninformative prior for the probability this happens, $p^{\uparrow\downarrow}$, then the posterior distribution would be a beta distribution: $p^{\uparrow\downarrow} \sim \mathrm{B}(0.5 + n_2, 0.5 + n_1)$. We draw a sample from this distribution when graphically displaying this statistic.

Figure 3.1 shows the performance of the 7 CMPs, for each of the 4 tuning objectives, for the TRO statistics relating to probabilities. Figure 3.2 shows the log-linear TRO growth summary. Figure 3.3 shows the CMP performance for the TAC summary relating to actual catch amount (mean, maximum decrease, and lower 10%ile). Figure 3.4 shows the AAV and up/down catch probability statistics.

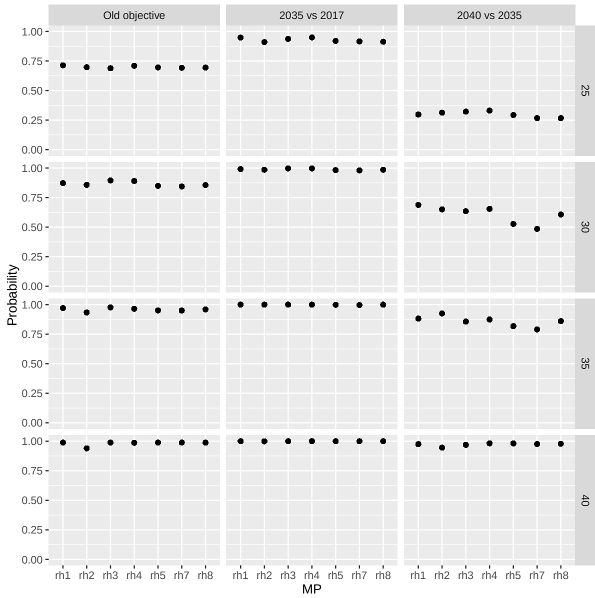

Figure 3.1: *TRO probability summary for each of the CMPs and for each tuning objective.*

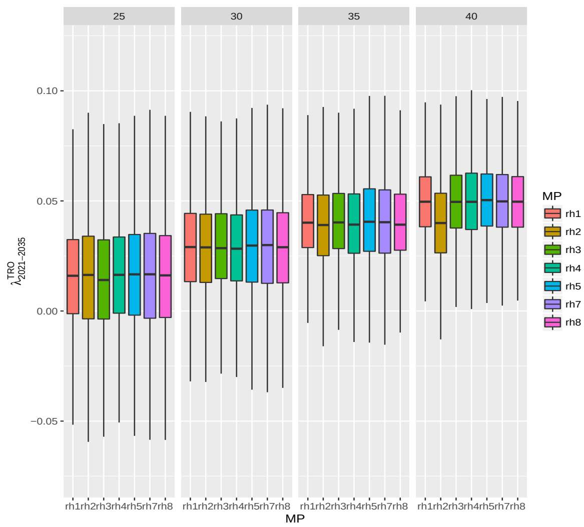

Figure 3.2: *TRO log-linear growth summary for each of the CMPs and for each tuning objective.*

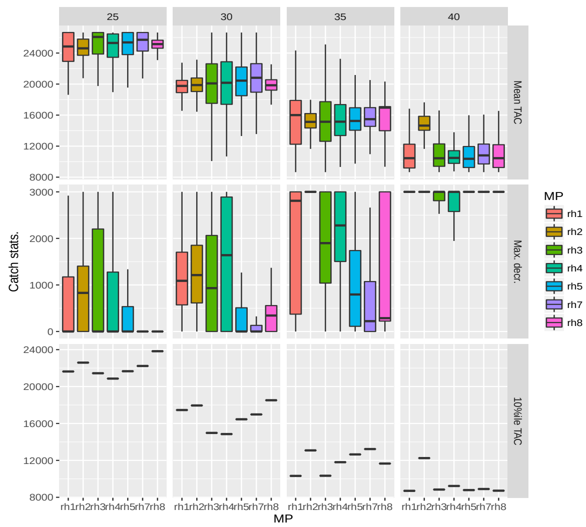

Figure 3.3: *TAC summary for the statistics relating to actual catch amounts: mean TAC (top), maximum decrease (middle), and lower 10%ile (bottom).*

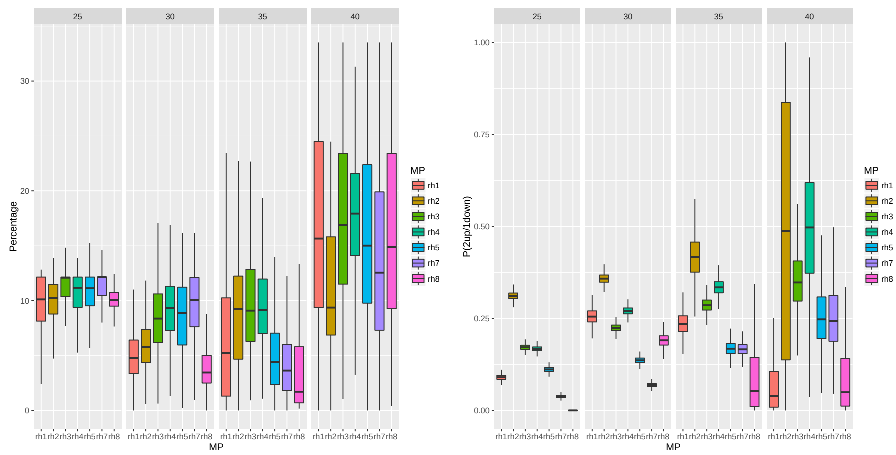

Figure 3.4: *TAC summary for the statistics relating to actual catch amounts: mean TAC (top), maximum decrease (middle), and lower 10%ile (bottom).*

3.1 General performance across tuning objectives

At the extremes of the tuning objectives (25% and 40%) all MPs basically have to act to either rapidly increase the TAC (25% objective), or rapidly decrease the TAC (40% objective). This is not surprising given that the constant TAC levels for these two cases are 23,850t (25%) and 11,803t (40%) - both of which are basically 2 maximum TAC changes (for 3,000t max. change) away from the last fixed TAC of 17,647t. In the 25% case all the MPs result in catch levels that causes the TRO in 2040 to have a 65-75% chance of being below the 2035 level. For the 40% objective the rapid and sustained reductions in TAC required result in TRO levels in 2040 that are greater than those in the target year (2035) almost with probability 1, and TAC levels that are on average around 30–35% of MSY levels, even with a TRO in 2040 that is well above the MSY level. So, for the 25% objective all MPs increase the catch above the level at which we would expect the tuning objective to be maintained (it goes down in the future); there is also a significant chance of stock trajectories that go down on average over the rebuilding period. For the 40% objective all MPs rapidly decrease TACs, where the TRO continues rising rapidly after the target year, and there is a large distance between the average TAC and MSY, even with a TRO level significantly above the level assumed to produce MSY. Henceforth, we only outline the more specific performance of the MPs for the 30% and 35% tuning objectives, given there is more apparent behavioural contrast for these two cases, relative to the extrema.

3.2 MP performance on the TRO statistics for 30% and 35% tuning targets

All the MPs clearly exceed the original tuning objective for the Bali Procedure, so there are no issues there. For the 30% objective, the TRO in 2040 is greater than that in the tuning year (2035) with a probability of around 50% or above for all MPs - though it is basically at 50% for **rh5** and **rh7** (the CKMR empirical MPs). For the 35% objective these probabilities are all at 75% or more - all MPs have a better than average chance of further increasing the TRO post tuning year. For the 30% tuning objective there is a small chance of trajectories that go down on average between 2021 and 2035 with no trend across MPs; for the 35% objective this probability is very low and basically the same across MPs. On average, the CKMR driven MPs seem to result in a more stable TRO trajectory post tuning year (but also with a small chance of it coming down again by 2045), but other than this there is little to clearly separate the MPs looking at TRO summary statistics alone.

3.3 MP performance on the TAC statistics for 30% and 35% tuning targets

Contrasting the MPs with trend vs. target forms for CPUE first: the trend driven MPs (**rh1**, **rh2**) show lower variability in both mean TAC and AAV relative to the target-driven ones (**rh3**, **rh4**). Conversely they have a higher propensity to increase the TAC twice and then decrease it. For the 30% tuning objective the trend-based MPs show lower maximum decreases, relative to their target-based counterparts, but this behaviour reverses for the 35% objective. The MPs including CKMR (**rh5**, **rh6**, and **rh7**) tend to have higher mean TACs (relative to CPUE/GT only MPs), with a significantly and consistently lower chance of decreasing the TAC after 2 increases. For the empirical CKMR MPs (**rh5**, **rh7**) the AAV is comparable to the CPUE/GT MPs for the 30% objective, but lower for the 35% objective. Across all tuning objectives the CKMR including MPs have clearly lower maximum TAC decrease statistics, relative to CPUE/GT MPs. The model-based CKMR MP (**rh8**) consistently has both the lowest AAV and chance of a TAC decrease after two increases, with better than average levels of the lower 10%ile TAC. Mean TAC levels, while more variable for target CPUE-drive MPs and empirical CKMR MPs, relative to the CPUE trend or model-based CKMR MP, were very similar in terms of medians across tuning objectives.

4 Overall summary of initial MP performance

Figures 3.5 and 3.6 show the TAC and TRO worm plot summaries for MPs **rh1**, **rh3**, **rh7**, and **rh8** for the 30% objective to outline the general differences between MPs with different HCRs. The CPUE trend,

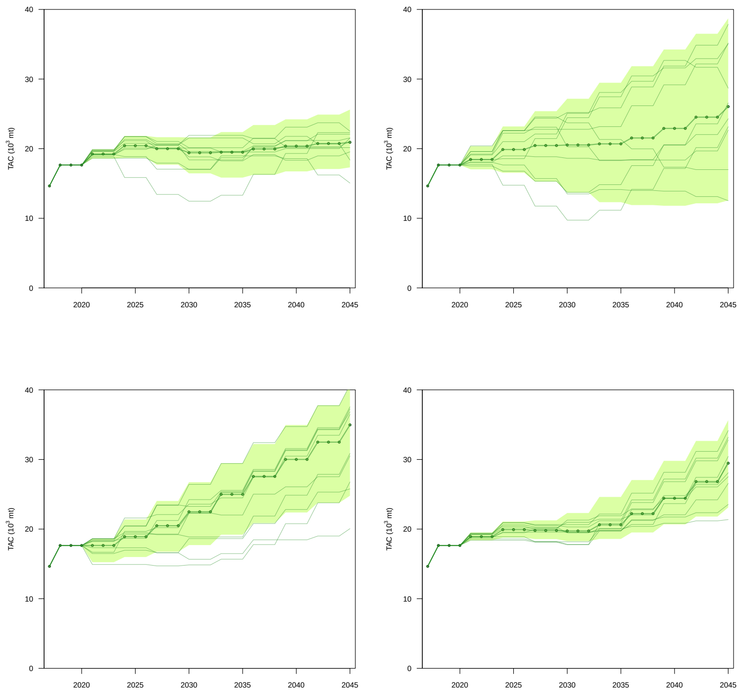

Figure 3.5: *Worm and quantile summary for TAC and MPs **rh1** (top right), **rh3** (top right), **rh7** (bottom left), and **rh8** (bottom right).*

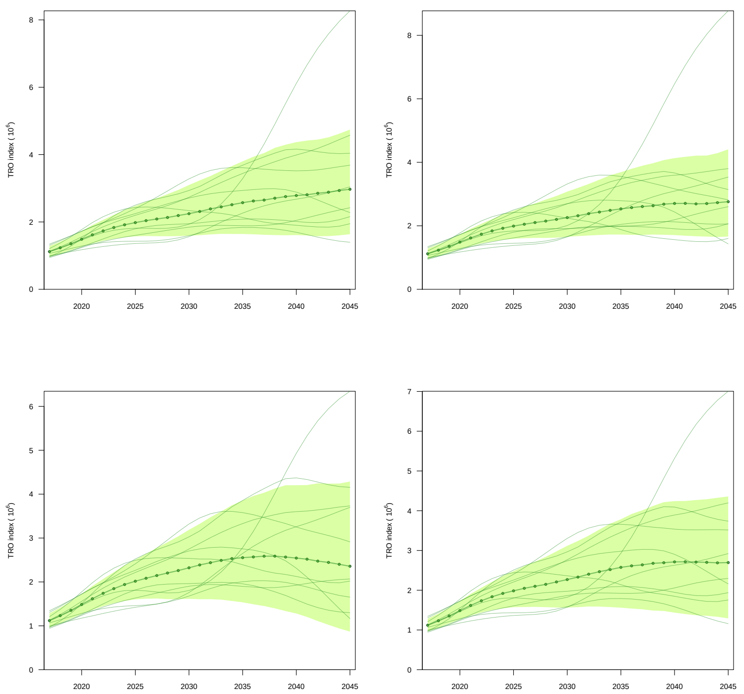

Figure 3.6: *Worm and quantile summary for TRO and MPs **rh1** (top right), **rh3** (top right), **rh7** (bottom left), and **rh8** (bottom right)..*

GT limit MP (**rh1**) shows low TAC variability, but more prominet up/down TAC dynamics (given the 2013 recruitment-driven trend in future CPUE) with TRO still increasing post-tuning objective year. The CPUE target, GT limit MP (**rh3**) shows more TAC variability (both mean and AAV) than **rh1**, but higher TACs later in time and a more stabilising TRO trend post tuning year. The empirical CPUE trend/CKMR target/GT limit MP **rh7** shows higher average catches, little up/down TAC behaviour but with TAC still increasing post-tuning year and the TRO just beginning to come down again by 2045. The model-based CKMR/CPUE/GT MP **rh8** shows a gradual increase in mean TAC early with later slight faster increases. It has the best up/down TAC and AAV performance across all MPs as well as the lowest level of maximum TAC decreases. After the tuning year the TRO begins to stabilise at the target level.

These initial MP explorations have yielded some informative inital results:

- The 25% and 40% tuning objectives (by 2035 with a 3,000t max change in TAC) show the least contrast across MPs as they all have to either rapidly increase (25%) or decrease (40%) the TAC
- The 30% and 35% tuning objectives allow for more behavioural contrast across the CMPs
- Both trend and target features explored for CPUE: the former show low TAC variability but higher up/down TAC behaviour; the latter show more overall TAC varability but comparable up/down behaviour - overall they seem a bit *overreactive* but this can be fixed
- The use of the GT data in the limit form (GT1) versus the trend form (GT2) - even at this stage - looks more promising
- The use of CKMR data is clearly potentially very useful. In the empirical form the tuning to the 30% objective looks a tough aggressive (in terms of TAC increases later in time) but this can be tempered. It is the model-based use of CKMR that seems to show very promising results, with the best AAV stats, the lowest maximum TAC decreases, and the least propensity for decreasing TAC after two increases, as well as mean TAC levels at or above the other CMPs.

For the 30% and 35% objectives, it is clear that when tuning to the reference set (and for the 3-year TAC maximum 3,000t change settings) that a wide array of potential CMPs can be tuned to the objectives, and that they can and do have differing performance characteristics. Choices around how best to use the CPUE (trend, target, less reactive) and the GT data will be best explored via the low recruitment and CPUE-related robustness trials. The CKMR data clearly show promise also in both their empirical and model-based forms.

5 Acknowledgements

This work was funded by CSIRO and the Department of Agriculture and Water Resources.

References

[1] R. M Hillary, A. Preece, and C. R. Davies (2018) Data generation & changes to SBT OM. *CCSBT–OMMP/1809/**

[2] R. M Hillary, A. Preece, and C. R. Davies (2016) Methods for data generation in projections. *CCSBT–OMMP/1609/07*

[3] R M. Hillary *et al.* (2016) A scientific alternative to moratoria for rebuilding depleted international tuna stocks. *Fish and Fisheries*. **17**:469–482.

CONTACT US

t 1300 363 400
+61 3 9545 2176

e csiroenquiries@csiro.au

w www.csiro.au

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 Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE. WE INNOVATE.