Oceans & Atmosphere
www.csiro.au

Updated candidate MP performance summary on reconditioned grid

Rich Hillary, Ann Preece, Campbell Davies
18th June 2019

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7000, Tasmania, Australia.

Contents

1 Background

For the 2019 MSE work, the operating models [1] have been reconditioned with updated and new data (CPUE, CKMR, catch biomass and composition,and new gene tagging data). As agreed in 2017, the UAM1 scenario is included in the reference set for testing of candidate MPs. In this paper we detail the key results of the updated OM reconditioning, including data fits and revised status, as well as the performance of the previous suite of candidate MPs [2].

2 Reconditioned OM

The folllowing data updates and new sources are included in the 2019 reconditioning:

- Catch biomass, composition and Japanese longline CPUE up to and including 2018
- CKMR POP and HSP data up to and including sampling year 2017, which would observe the adult population upto and including 2014
- The two gene tagging data points observing age 2 abundance in 2016 and 2017

2.1 New and revised OM settings

Given the inclusion of the gene tagging data in the OM there are a number of new variables required for the probability model for the matches [3]. The new variables are all included in the `sqrt.dat` file:

- `qgt` (q^{gt}): default is set to 1 (and assumed that $q^{\mathrm{gt}} \leq 1$)
- `gtOD` (φ^{gt}): default is set to 1 (and $\varphi^{\mathrm{gt}} \geq 1$)
- `gtsw`: 0/1 switch flag to turn GT data off/on (default set to 1)

The only other alteration to the settings in the `sqrt.dat` file is to set the q^{hsp} parameter to 1 and *not* estimate it (i.e. set the estimation phase to -1). This was a change agreed by the OMMP group last year, based on analyses from the last OM reconditioning [1] that showed there was no mismatch between the POP and HSP data, in relation to their information on absolute adult abundance (which is what this parameter was designed to capture).

2.2 Summary of reconditioned OM

The grid configuration agreed to in 2017 for MP testing is detailed in Table 3.1 and, in line with previous reconditionings, we sample 2,000 models from the current suite of 432 using the resampling scheme outlined in Table 3.1. We summarise the `base18UAM1` grid of operating models, given this is our current reference case for the MSE work. For the best fitting grid element, the fits to the abundance data (CPUE, aerial survey and gene-tagging) see Figure 3.1. The fits to the conventional tagging data are detailed in Figure 3.2; the aggregated fits to the CKMR POP and HSP data (as per [1]) are detailed in Figure 3.3.

The fits to the CPUE are similar to previous years, and the notable increase in CPUE in 2018 is fitted well (driven by the already large estimate of recruitment in 2013 driven by the 2016 aerial survey). The fits to the aerial survey haven't changed since the previous assessment [1] and the fit to conventional tagging data are also similar to previous years. The fits to the gene tagging data are good and the observed number of matches are well within the approximate 95%
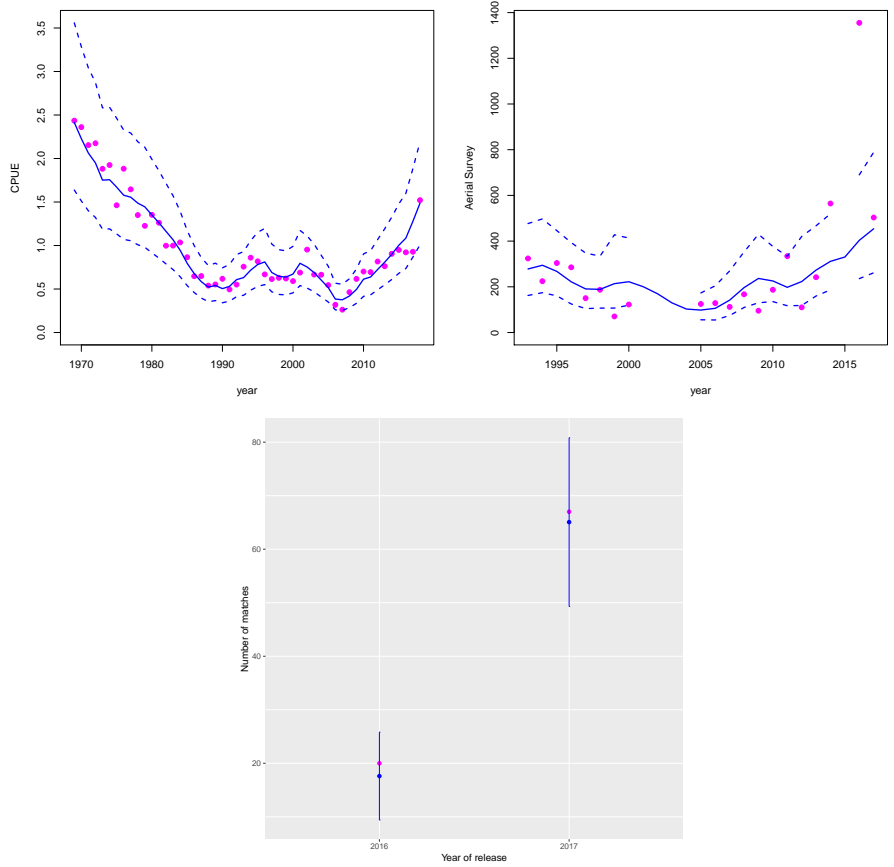
Figure 2.1: *Observed (magenta) and predicted median and 95% CI (blue) for the Japanese longline CPUE (top left) and aerial survey (top right) indices, and the gene tagging matches (bottom).*
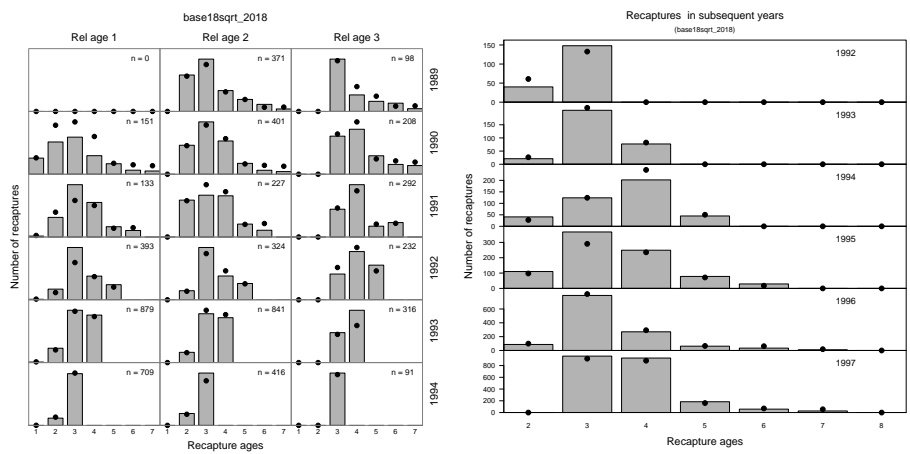

Figure 2.2: *Disaggregated (left) and pooled (right) 1990s tagging data fitting summaries.*

Parameter	Values	Prior	Resampling	CumulN
Steepness	$\{0.6, 0.7, 0.8\}$	Uniform	Prior	3
M_0	$\{0.35, 0.4, 0.45, 0.5\}$	Uniform	Objf	12
M_{10}	$\{0.0.5, 0.085, 0.12\}$	Uniform	Objf	36
ω	$\{1\}$	Uniform	Prior	36
CPUE ind.	$\{2, 3\}$	Uniform	Prior	72
CPUE ages	$\{4, 18\}$ & $\{8, 12\}$	$\{0.67, 0.33\}$	Prior	144
ψ	$\{1.5, 1.75, 2\}$	$\{0.25, 0.5, 0.25\}$	Prior	432

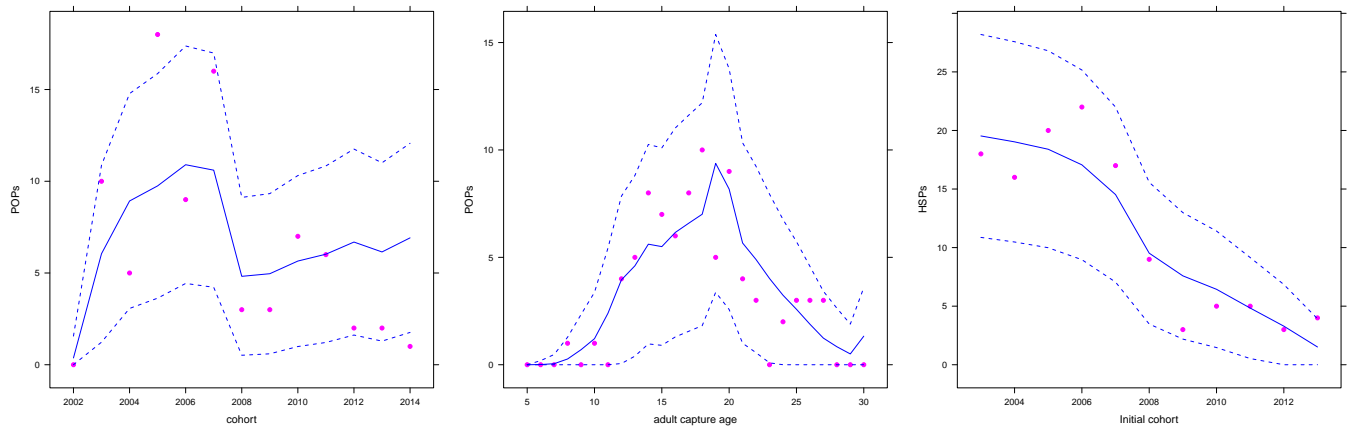Table 2.1: *Summary of the agreed grid configuration for the 2019 reconditioning*

Figure 2.3: *Observed (magenta) and predicted median and 95% CI (blue) for fits to the POP data aggregated to the cohort (left) and adult capture age (middle) levels, and the HSP data aggregated to the initial comparison cohort level (right).*

confidence interval. A more detailed summary of these fits can be found in [3]. The fits to the aggregated POP data are similar to previous years but there is a slight trend in number of POPs for the most recent juvenile birth years (2012–2014) being over-estimated. Apart from the last point which is just outside the bounds, there is no clear significant misfit, and the data for these cohorts will be not be static - in the coming years we will compare new adults to juveniles born in these years and so could detect more matches which will change this trend. The fits to the adult capture age level of the POPs is good as are the HSP fits when aggregated to the initial cohort level. It is also worth noting the sample sizes for the CKMR monitoring are based on previous OMs. Given the updated estimates of status and population dynamics since the original design study and the use of CKMR for stock assessment, monitoring the rebuilding plan and input to candidate MPs, a review of this monitoring program and associated sample sizes should be a priority to ensure appropriate samples sizes in the future.

Variable	TRO depletion	B_{10+} depletion	F/F_{msy}	B/B_{msy}	B_{msy}/B_0
Summary	0.17 (0.15–0.21)	0.14 (0.12–0.17)	0.55 (0.41–0.74)	0.64 (0.47–0.91)	0.27 (0.22–0.32)

Table 2.2: *Population dynamic summaries (median and 90% CI) for the reconditioned OM.*

The main population dynamic summaries can be found in Table 3.2 (for the reference set with UAM1 scenario included). Current TRO depletion has a median (and 90% CI) of 0.17 (0.15–0.21) so higher than the 0.13 estimate of 2017 [1] but in line with the projections done in both

2017 and 2018. Current estimates of F are just above half of F_{msy} with a very low probability of exceeding it. The ratio of the adult biomass at MSY relative to the unfished level is consistent with previous estimates: 0.27 (0.22–0.32).

3   Candidate MP structures

We have not, at this stage, modified the structures of any of the previous candidate MPs [2]; instead we have simply chosen to retune them given the revised OMs. The details of the candidate MP structure we explore in this work (previously referred to as **rh12** [2]) can be found in the Appendix but we outline the general qualitative features of the MP structure:

- The MP uses CPUE, gene tagging and CKMR (POP and HSP) data

- For the CPUE part of the HCR a simple trend form is used

- For the CKMR part a simplified adult population model (abundance and total mortality) is fitted to the CKMR data. The log-linear trend in TRO, λ^{ck}, is then used in the HCR (albeit with a minimum initial increasing trend in TRO encouraged)

- For both the CPUE and CKMR trend terms the gain parameter is density-dependent - specifically on the relative level of TRO compared to the TRO estimated given the actual observed data. For a given level of TRO rebuilding (relative to the recent estimates) the gain parameter is stronger prior to reaching the rebuilding then decreases as the TRO reaches the target level. This ensures reactivity when needed (in the rebuilding phase) but stability when it is reached

- For the gene tagging term a limit-type approach is used: (i) for values of the current 5 year average 2-year old abundance below the limit strong (supralinear) decreases in TAC are enacted; (ii) for values above the upper level weaker (sublinear) increases in TAC are permitted; (iii) for values between the two nothing is done to the TAC. A crucial difference for the GT part of the HCR is that there is *no* inertia: once the values appear outside the bounds of inaction the TAC is proportionally changed

4   Robustness tests

A number of key robustness tests were outlined in last years ESC report with high (H), medium (M), and low (L) rankings:

1. **reclow5** (H): 5 years of future mean recruitment @ 50% of expected level

2. **cpuew0** (L): alternative weighting for main CPUE series

3. **as2016** (H): remove 2016 aerial survey index

4. **h55** (M): reduced grid with only $h = 0.55$ steepness value

with the combination robustness **as2016reclow5** also getting a high ranking. Given issues around getting data sets updated and running, we focus only on the high ranked robustness tests.
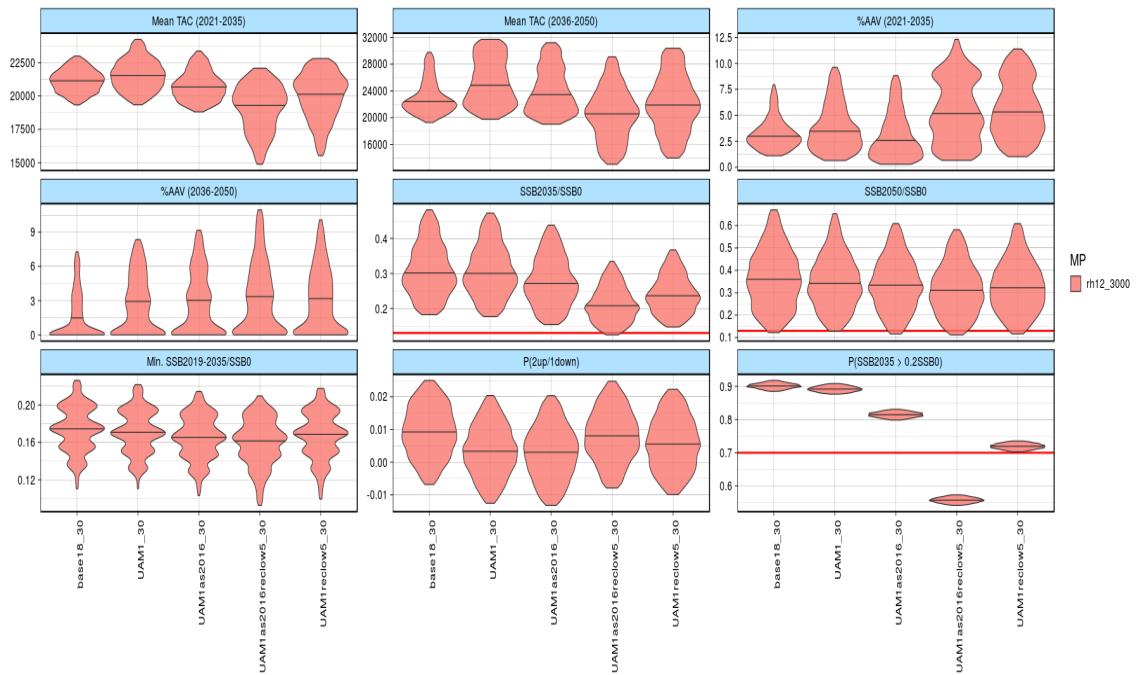
Figure 5.1: *Performance summary for the 30% by 2035 tuning objective.*

5   Results

At this stage the candiate MP was tuned to the 30% by 2035 and 35% by 2040 tuning objectives and the results are split by each of these tunings given they do result in MPs with different behaviours. The base grid in all cases is the UAM1 option *but* we do include the tuned grid without the additional UAM to see the effect of this inclusion.

5.1   30% by 2035 tuning level

Figure 5.1 shows the SBT shiny app violin plot performance summary for this tuning level. Figure 5.2 shows the associated TAC and TRO worm plots for the base UAM1 grid. For the base tuning average TACs (for 2021–2035 period) range from 19,000 to 23,000t; AAV is low (median of just over 3%) and never seems to exceed 10%; for the period after the tuning the AAV is even lower as build into the MP structure; the probability of 2 TAC increases then a decrease is very low; and the probability of being above 20% of the unfished level in 2035 is around 0.9 (so well above the previous 0.7 tuning objective).

For the **as2016** robustness test, this generally results in slightly lower average TACs over the tuning period, slightly lower AAV (as big 2013 recruitment is reduced in influence in projections), and just misses the actual tuning objective getting to around 27% with probability 0.5. The original tuning objective is still exceeded (just over 0.8).

For the **reclow5** robustness test, this results in lower TACs over the tuning period and specifically an asymmetric distribution in the average TAC to levels down to around 15,000t at the lowest given the limit-type nature of the gene tagging part of the HCR. The median value of depletion by 2035 is around 0.24 and the original tuning objective is still achieved.

For the **as2016reclow5** combination robustness test this results in the most pessimistic projections, as one might expect. Average TAC levels are similar but a little lower than the **reclow5**
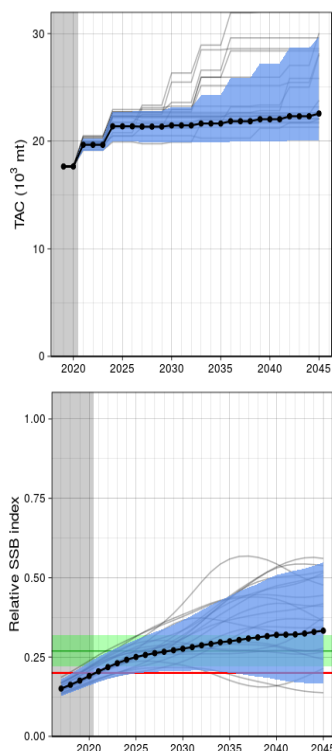
Figure 5.2: *Worms plots for the base UAM1 grid (tuned to 30% by 2035) for TAC (top) and TRO (bottom) and 20 random worms are shown.*

case, wIth median TRO levels of around 0.21 by 2035 - so it misses the original tuning objective but does get the depletion to 20% with a greater than 50% probability by 2035.

5.2 35% by 2040 tuning level

Figure 5.3 shows the SBT shiny app violin plot performance summary for this tuning level. Figure 5.4 shows the associated TAC and TRO worm plots for the base UAM1 grid. For the base tuning average TACs (for 2021–2035 period) range from 18,000 to 21,000t; AAV is low (median of just over 2%) and never seems to exceed 10%; for the period after the tuning the AAV is even lower as build into the MP structure; the probability of 2 TAC increases then a decrease is very low; and the probability of being above 20% of the unfished level in 2035 is around 0.95 (so well above the previous 0.7 tuning objective).

For the **as2016** robustness test, this generally results in slightly lower average TACs over the tuning period, slightly lower AAV (as big 2013 recruitment is reduced in influence in projections), and just misses the actual tuning objective getting to around 29% with probability 0.5. The original tuning objective is still exceeded (just over 0.86).

For the **reclow5** robustness test, this results in lower TACs over the tuning period and specifically an asymmetric distribution in the average TAC to levels down to around 15,000t at the lowest given the limit-type nature of the gene tagging part of the HCR. The median value of depletion by 2035 is around 0.26 and the original tuning objective is still achieved.

For the **as2016reclow5** combination robustness test this results in the most pessimistic projections, as one might expect. Average TAC levels are similar but a little lower than the **reclow5**
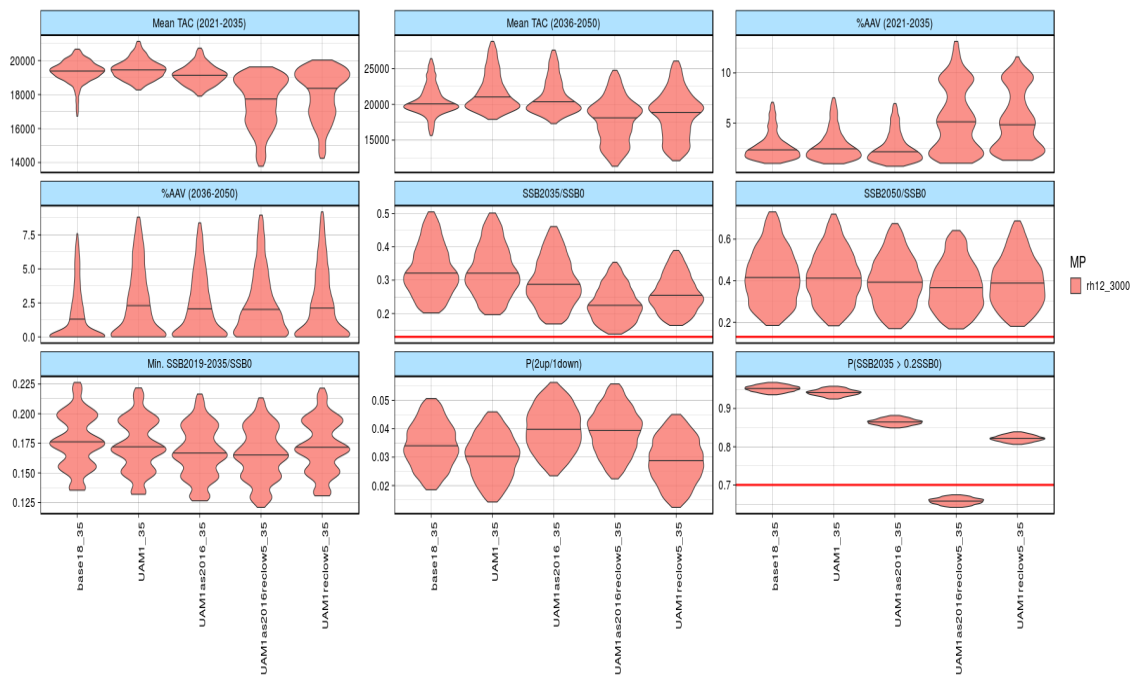
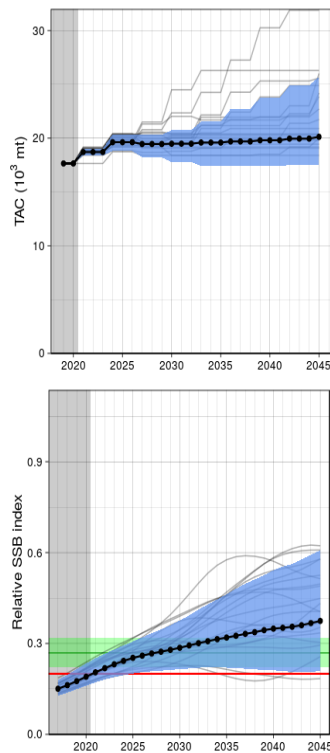Figure 5.3: *Performance summary for the 35% by 2040 tuning objective.*

Figure 5.4: *Worms plots for the base UAM1 grid (tuned to 35% by 2040) for TAC (top) and TRO (bottom) and 20 random worms are shown.*

case, wth median TRO levels of around 0.23 by 2035 - so it misses the original tuning objective but does get the depletion to 20% with a greater than 60% probability by 2035.

6 Discussion

The CCSBT OM has been reconditioned for data up to and including 2018 as well as the first inclusion of the two gene tagging data points. The base grid agreed at the previous ESC was used and the UAM1 unaccounted mortality scenario was used to create the reference set of OMs used in the MSE work. The data were generally fitted well - including the new gene tagging data - and there were no obvious issues with the resulting OM that would suggest it could not be used in the MSE work this year.

The best performing of our suite of candidate MPs from the previous year, **rh12** [2], was retuned to the reference grid for the 30% by 2035 and 35% by 2040 tunuing objectives. The key robustness tests explored were those ranked as high in 2018: **reclow5**, **as2016**, and the combination **as2016reclow5**. Performance was similar to previous analyses:

- For the reference grid, average catches were in the 18,000-23,000t range (across both tunings) over the tuning period and the previous tuning objective was achieved with high probability (90% or above). Both AAV and the probability of having two TAC increases followed by a decrease were low

- For the **reclow5** and **as2016** the average catches were lower, and the original tuning criteria were undershot, but the previous tuning objective was met with a greater than 70% probability and future declines below current levels were very unlikely

- The combination robustness test **as2016reclow5** was the only test where the MP failed to achieve the old tuning objective, but did result in a greater than 50% probability of being above 20% of B_0 in 2035 in both tuning cases

- In general, the 35% by 2040 tuning results in more conservative projections than the 30% by 2035 tuning

7 Acknowledgements

References

[1] R. M. Hillary *et al.* (2017) Reconditioning of the CCSBT Operating Model in 2017. *CCSBT–ESC/1708/14*.

[2] R. M. Hillary, A. Preece, C. R. Davies (2018) Revised MP structures and performance. *CCSBT–ESC/1709/20*.

[3] R. M. Hillary, A. Preece, C. R. Davies (2019) Changes to SBT OM conditioning code. *CCSBT–OMMP/1906/**.

Appendix

We explored a modified version of the original adult-focused age-structured population model, now with auto-correlated "recruitment" deviations:

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

The estimate parameters of this model are:

1. The mean adult recruitment, $\bar{R}$

2. The adult recruitment deviations, ϵ_y

3. The initial value, χ_{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 similar to the number of parameters estimated in the Bali Procedure population model. There are not 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. The assumed settings for the CKMR MP population model are detailed in Table 8.1.

The general structure of the revised MP is as follows:

$$
TAC_{y+1} = TAC_y \left(\omega^{\text{cpue}} \left(\Delta_y^{\text{cpue}} - 1\right) + \omega^{\text{ck}} \left(\Delta_y^{\text{ck}} - 1\right)\right) \times \Delta_y^{\text{gt}},
$$

where the inertial terms for the CPUE and CKMR parts of the HCR are now additive, not multiplicative as previously explored. This avoids the quadratic term in the multiplicative case where both trends are consistently positive subtley but consistently making the TAC increases larger than for the additive case, despite the trends being the same in both cases.

Before detailing the changed form of the HCR we recap some useful variables:

Parameter	Value
$a_{\min}$	6
$a_{\max}$	30
σ_R	0.25
ρ	0.5
σ_χ	0.1
$Z_{\min}$	0.05
$Z_{\max}$	0.4
$Z_{a_{\max}}$	0.5
$\mu_{\chi_{\text{init}}}$	-1.38
$\sigma_{\chi_{\text{init}}}$	0.15
q_{hsp}	0.9

Table 7.1: Settings for CKMR MP population model

- I_y^{ck}: moving average of the estimated TRO from the MP population model (now pushed forward to the current year using the model to project forward for 4 years to avoid too much intertia in the signal when you need it)

- $\tilde{I}$: average estimated TRO from 2003 to 2012 (reference period w.r.t. relative rebuilding criterion)

- γ: proportional amount of TRO rebuilding we wish to achieve

We are interested in the following ratio: $\delta = I_y^{\text{ck}}/(\gamma\tilde{I})$. To get from the current average level of TRO to the 30% level we would consider $\gamma \approx 2$; for the 35% level $\gamma \approx 2.5$. As the ratio δ approaches 1 (i.e. we *think* we are at or close to the target TRO), we would like to have the potential to morph (continuously and possibly smoothly) the behaviour of the MP. It seems that MPs need to be fairly reactive in the first 10–15 years (3–4 TAC decisions) of the projections to be able to tune to the 30% target by 2035, but afterwards that embedded reactivity might be giving rise to continued TAC increases to levels likely to cause the TRO to come back down again post-target year. For the CPUE trend part of the HCR we explore a density-dependent gain parameter:

$$k^{\text{cpue}}(\eta) = k_1^{\text{cpue}}\left(1 - \left(1 + e^{-2\kappa\eta}\right)^{-1}\right) + k_2^{\text{cpue}}\left(1 + e^{-2\kappa\eta}\right)^{-1}$$

where $\eta = \delta - 1$. This is using the logistic function approximation to the Heaviside step function $H[\eta]$ ($H[\eta < 0] = 0$, $H[\eta \geq 0] = 1$). We set $\kappa = 20$ so the transition between the two gain parameters, given η, happens within $\pm 5\%$ of $\delta = 1$. The CPUE multiplier is then just defined as follows:

$$\Delta_y^{\text{cpue}} = k^{\text{cpue}}(\eta)(1 + \nu)\lambda^{\text{cpue}} \quad \text{if} \quad \lambda^{\text{cpue}} \leq 0,$$
$$\Delta_y^{\text{cpue}} = k^{\text{cpue}}(\eta)(1 - \nu)\lambda^{\text{cpue}} \quad \text{if} \quad \lambda^{\text{cpue}} > 0$$

For the CKMR part of the HCR we try to preserve the main elements of the previous candidate MP (**rh8**): ensure a minimum rate of increase in the TRO *beneath* the target level, and once it is achieved we would like to maintain the TRO at that level. To include this kind of behaviour in the HCR we also include some density-dependence in the log-linear growth rate at which the HCR moves from a TAC increase to a TAC decrease:

$$\Delta_y^{\mathrm{ck}} = 1 + k^{\mathrm{ck}}(\eta) \left(\tilde{\lambda}(\eta) - \lambda^{\mathrm{ck}} \right),$$

$$k^{\mathrm{ck}}(\eta) = k_1^{\mathrm{ck}} \left(1 - \left(1 + e^{-2\kappa\eta} \right)^{-1} \right) + k_2^{\mathrm{ck}} \left(1 + e^{-2\kappa\eta} \right)^{-1},$$

$$\tilde{\lambda}(\eta) = \lambda_{\min} \left(1 - \left(1 + e^{-2\kappa\eta} \right)^{-1} \right)$$

The threshold level at which a trend goes from a TAC decrease to an increase essentially begins at $\lambda_{\min} > 0$ and, as the estimated TRO approaches the target level, this rapidly dercreases to zero (in a similar way to the CPUE trend term). This is to ensure that a minimum level of rebuilding is encouraged for **all** trajectories below the target, and where above the target the *status quo* is preferred.

Along with embedding a kind of switching mechanism in both **rh11** and **rh12**, in terms of behaviour once the target is met, we also introduce a maximum TAC value. This is again to avoid short-term increases to levels of TAC (and, hence, total catch including UAM) that are not sustainable in the long-term, even for the most optimistic grid combinations and future trajectories, and will definitely require large TAC decreases in the future. The value chosen for the maximum TAC was 32,000t. Including UAM (which is approximately and consistently 20% of the TAC) this value would be a total catch of around 36,000t.

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.