Conditioning IOTC Albacore OMs using the ABC approach

R. M. Hillary[1] & I. Mosqueira[2]

[1]CSIRO Environment, Australia.
[2]Wageningen Marine Research, Netherlands.

1 Background

For the current suite of IOTC MSE work, the general approach to conditioning the required set of Operating Models (OMs) has been to use the species-specific stock assessment model structure as the basis for the OMs. A grid of model runs, formulated using a set of alternative assumptions and inputs, is constructed around the base case assessment model. In [1] an alternative and complementary approach was outlined where, instead of the assessment being the basis for conditioning, a suite of possible prior states of historical dynamics and current status are defined. The available, but mostly the more recent, data are included within an estimation scheme built on emerging Approximate Bayesian Computation (ABC) and Synthetic Likelihood (SL) concepts [2, 3]. The aim is to generate a distribution of current abundance, mortality and status that is consistent with both the available data and the suite of possible prior states of nature defined beforehand, some of them informed by the stock assessment results. This set of recent dynamics can then be used to initialise the OMs used to project the stock into the future, and test the candidate MPs.

A stock assessment, in this context, can be viewed as our attempt to do both of these things at once. Ideally, this is arguably a sensible option; however, it is not always successful. The ongoing struggles with the Yellowfin tuna stock assessment, and the conditioning of OMs based upon it, outline this problem: what if you cannot adequately reconcile the data, assessment model structures, and the resultant estimates of current status and future projected dynamics? In [1] we proposed an alternate approach arguing that using a stable, agreed and robust stock assessment was a natural first option, but that the ABC approach was a potentially viable - and scientifically pragmatic - alternative approach if the assessment route was unsuccessful.

In this paper we condition the Indian Ocean Albacore tuna OM that mirrors (biologically and structurally) the most recent stock assessment, utilises length composition and longline CPUE data, and is able to explore a wide range of stock status prior hypotheses, many of them built on information from the results of the stock assessment. The aim of this work was to cover the same range of factors/hypotheses covered in the previous suite of OMs (REFS):

1. Use of - and generation of data from - alternative CPUE series

2. Influence of size data

3. Impact of assumed catchability increases in LL fleets

4. Uncertainty in steepness and natural mortality

5. Uncertainty in recruitment variability

2 Methods

ABC [2] and SL [3] methods can be used to define an *approximate* distribution for the parameters θ we are interested in; subsequently, we obtain an approximate distribution for all the variables that depend on those parameters. Where they differ from more classical frequentist or Bayesian methods is how the data, D, are included. Classical methods posit a likelihood

for the data, given the parameters: $\ell\left(D\,|\,\boldsymbol{\theta}\right)$; for a Bayesian analysis we then define a prior distribution, $\pi(\boldsymbol{\theta})$ to then obtain the posterior distribution of the parameters given the data:

$$\pi\left(\boldsymbol{\theta}\,|\,D\right) = \frac{\ell\left(D\,|\,\boldsymbol{\theta}\right)\pi(\boldsymbol{\theta})}{\pi(D)}$$

The ABC approach relaxes the requirement for a specific likelihood (i.e. data generating probability model) to the idea of a *discrepancy* statistic that measures the difference between the observed data and the model-derived process variables, X, that relate to it. The simplest example would be some distance metric $\rho(D, X)$ whereby we require that this distance between the observed data and our prediction is less than some value $\delta > 0$ (i.e. assumes uniform error on a radius δ). Values whereby $\rho(D, X) \geq \delta$ receive zero probability mass and, in a sampling scheme, would never be accepted. This simple approach will not necessarily work for certain types of data, especially the types we often have in fisheries contexts, but there are natural generalisations of this simple discrepancy idea that can be applied here.

Another very useful thing we can easily embed within a sampling scheme is a set of informative priors on various elements of the process variables, X, which effectively imply a prior on the key parameters, $\pi(\boldsymbol{\theta})$. This means we can also define informative priors for the various stock status variables (MSY ratios, SSB depletion *etc.*) and include the most relevant recent data, without having to fully model the historical dynamics and data and define appropriate likelihoods for these data. If we are departing from the stock assessment approach, this implies that we have not really succeeded in being able to robustly estimate the distribution of these key status variables. This approach takes a step back and instead tries to define scenarios (specifically distributional ones) for these variables at particular periods in time that are consistent with previous assessment experiences and the data. In terms of likely data sources, the two most obvious examples in tuna population models are the CPUE indices of abundance, and the catch at length composition. For some examples we may also have mark-recapture data (e.g. tropical tuna) or other data sources. These can be naturally added to the algorithm, as long as the population and fishery model can generate observations for them and a distance metric can be computed.

The population models and associated biological relationships will be very similar to the stock assessment, given that they need to have the same structure as the existing OMs (which we do not propose need changing at this stage). The main technical challenge is constructing the discrepancy statistics for the data, and the sampling scheme to generate samples from our approximate distribution of the parameter and population dynamic variables. We have moved the technical details of how these are done to Appendix A to avoid an unnecessary amount of technical exposition in the main body of the paper: what we are trying to demonstrate is that we can get the variables we need for OM conditioning, using plausible status scenarios, key data sets and emerging powerful statistical sampling techniques.

3   Albacore case study

The basis for the models and data explored in the Albacore example is contained in the most recent stock assessment [4]. In terms of model structure, the model is setup as follows:

- Annually structured but with four seasons and recruitment in a single pre-specified season.

- No explicit spatial structure but with the same areas-as-fleets approach as the stock assessment.

- Sexually-structured population dynamics driven by growth and selectivity-at-age (selectivity-at-length is not sexually structured by fishery).

- Time-frame for conditioning is 2000 – 2020, so as to model all surviving cohorts.

- Model considers 4 long-line fleets, 1 "other" fleet and 1 purse seine fleet. The 16 seasonal long-line fleets are condensed into 4 fleets, each with 4 seasons of catch, effort and size data (e.g. LL1 merges fisheries 1 to 4 in the assessment).

- A Beverton & Holt steepness-unfished recruitment stock-recruit relationship is used with lognormally distributed deviations constrained by a pre-specified $\sigma_r = 0.3$, as per the assessment [4].

The stock assessment model fits to the disaggregated length-frequency data and the seasonal CPUE data from the longline fleets. The initial approach taken in the ABC albacore modelling was not to fit to disaggregated length frequency data, which is very variable and noisy from season to season and year to year. Instead, we aggregate the size data across years and seasons and fit to a mean size frequency data set per fishery. The reason behind this choice is to obtain a representative selectivity relationship for each fishery, and give a clear indication of the "average" size structure in the fisheries over the time period. Additionally, it purposefully reduces the influence of the size data on the overall population abundance scale, which is a known issue already considered in the previous OM grid.

3.1   Stock status priors

These status priors are a key feature of the OM conditioning methodology we propose. To constrain both absolute and relative abundance, as well as fishing pressure we explore a range of possible types (and parameterisations) of status prior. The following four general types are explored (in different forms):

1. Relative SSB: the most general status variable (SSB_y/SSB_0) - for any range of years we allow a prior mean and SD

2. SSB MSY ratio: the ratio of SSB to that at MSY (SSB_y/SSB_{msy}) - for any range of years we allow a prior mean and SD

3. Harvest rate MSY ratio: the ratio of fleet-averaged harvest rates to MSY levels - for any range of years we allow a prior mean and SD

4. Overfishing probability: if the harvest rate exceeds the level at MSY a half-normal penalty can be applied to penalise against trajectories with very high harvest rates (and likely will crash early in the projections)

The status priors are based on information reported in the most recent stock assessment [4]. Different combinations of status priors are employed for different OM conditionings to explore a balance between letting the data inform the model (w.r.t. abundance indices), and forcing the OMs to adhere closer to the assessment outcomes. For relative SSB we define a prior for the first year (2000) centered on 0.5 with a 95% prior CI of 0.3–0.7. For the relative SSB MSY ratios we define a prior for the last two years (2019 and 2020) centered on 2.25 and 2 with an SD of 0.35 for both. For the scenarios where we apply a prior on the harvest rate MSY ratios for the first and last years (2000 and 2020), we assign a prior mean of 0.6 with an SD of 0.2 (so a prior range of around 0.2–1). These are used for scenarios where it is clear there is no information in the abundance indices in relation to population scale. Applying these vaguely informative priors penalises against scenarios where there is effectively no effect of fishing on abundance, the population is very large and the abundance trends are recruitment driven. When the overfishing probability penalty is applied - to avoid overly pessimistic individual runs - it is applied for all years and where the probability of the overfishing ratio exceeding 2 is less than or equal to 0.05.

3.2   Covariant prior for steepness and natural mortality

A common, and fair, criticism of many stock assessments is that they often have fixed values of both steepness and natural mortality. One of the main outcomes of this is that we effectively fix the value of all MSY variables - especially the stock status ratios for both SSB and fishing mortality [5]. The factorial grid approach tries to address this issue by exploring a combination of discrete values for both parameters. A major drawback of this approach is that it is not *a priori* able to avoid combinations of these two parameters that result in pathologically over/under resilient stock-recruit relationship dynamics. More specifically, for the IOTC a common issue seems to be combinations that are overly pessimistic, resulting in population trajectories that very quickly end up crashing [6].

Generally speaking, higher/lower values of steepness/natural mortality will tend to result in similar MSY ratios; high/low values of both are often overly optimistic/pessimistic. We attempt to utilise this inter-relationship to construct a covariant joint prior for these two parameters from which we will sample in the ABC-MCMC algorithm outlined later on. The algorithm proceeds as follows:

1. Define marginal distributions for both h and M

2. Calculate SSB depletion, $\tilde{\Delta}$, at MSY for expected values of h and M

3. Simulate h and M from their marginal distributions

4. Define a tolerance interval ε

5. Only accept combinations of h and M within ε of $\tilde{\Delta}$

6. Calculate correlation coefficient in retained bivariate samples

This makes it possible to then define a bivariate normal distribution for h and M using the marginal variances and the estimated correlation coefficient from the algorithm described

above. This then gives us a distribution of values from both h and M with the required marginal variances and a covariant structure that ensures we obtain internally consistent (with respect to MSY ratios) samples that are neither too optimistic or pessimistic relative to the expected MSY ratios.

For this example we take mean values of $h = 0.8$ for steepness and annual $M = 0.3$ as per the most recent stock assessment [4]. We calculate an SD for steepness that results in a prior 95% CI of 0.7–0.9; for M we define a 5% CV so that the prior 95% CI is 0.27–0.33. Using the correlation estimation algorithm described above with a tolerance of $\varepsilon = 5\%$ we obtained a correlation coefficient of -0.58.

3.3 OM scenarios

As stated previously, one of the main intentions of this work is to at least be able to cover all the main axes of uncertainty covered in previous Indian Ocean Albacore OM conditioning [6]:

1. Steepness and natural mortality: covariant joint prior now integrated into the ABC-MCMC algorithm

2. Recruitment variance: alternative OM conditioning that now actively estimates the recruitment variance, σ_r^2 , with an informative prior with mean 0.3 and prior 95% CI of 0.2–0.5

3. Weighting of size data: the issue was their influence on the estimates of population scale. By aggregating the size data across years and seasons, and establishing a conservative weighting strategy in the ABC discrepancy function, we have effectively removed this as an issue

4. Longline catchability trends: we explore a 1% *per annum* increasing trend in the longline CPUE series in alternative OM conditioning runs

5. CPUE series: we explore using the seasonal CPUE indices from longline fleets 1 and 3 - not jointly but separately to avoid data conflicts that will not be resolved without more complex-spatially structured models

4 Results

The key OM stock status variables we summarise are relative SSB, $\Delta_y = SSB_y/SSB_0$, SSB relative to MSY, $\tilde{\Delta}_y = SSB_y = SSB_y/SSB_{\mathrm{msy}}$, and harvest rate relative to MSY, $\mathcal{H}_y = H_y/H_{\mathrm{msy}}$. The seven different OM conditioning scenarios thatr have been developed are as follows:

- **R1**: Uses CPUE from LL fleet 1 with SSB priors but no harvest rate priors

- **R1a**: Uses CPUE from LL fleet 3 with SSB and harvest rate MSY priors

- **R1b**: Uses CPUE from LL fleet 1 with SSB and overfishing probability priors

- **R2**: Same as R1 but σ_r^2 is estimated

- **R2a**: Same as R1a but σ_r^2 is estimated

- **R3**: Same as R1 but assumes 1% increasing catchability trend

- **R3a**: Same as R1a but assumes 1% increasing catchability trend

Before outlining the results it will be helpful to focus on the important qualitative differences between the OM scenarios. Table 4.1 summarises the status of each of the seven OM conditioning scenarios described above. Scenario R1 is arguably where the data - specifically the CPUE abundance indices - have the most influence of both population scale and abundance, biomass and mortality trends. Scenarios R1a, R2a, and R3a are all penalised to be "close" to the high-level stock status summaries coming from the stock assessment [4] - in particular the harvest rate MSY ratio priors in the first and final years are applied in these cases. The main reason for this lies in the lack of information in the CPUE from fleet 3 in relation to population scale, and an apparent production function emerging from the models. For these runs we intermittently obtained very high abundance samples (i.e. a very skewed right-hand tail in the R_0 posterior distribution) where the trends in abundance were just as easily explained by recruitment variability versus fishing pressure. As a result, the relative SSB estimates are still concordant with the status priors but fishing mortality is effectively zero for these samples. By applying not just the biomass-based status priors but additionally the harvest rate MSY ratio priors (scenarios R*a), we avoided these pathological high-abundance runs because the harvest rates could not reduce to such low levels a priori. This is not, perhaps, surprising given the apparent lack of correlation between abundance trends and overall catch biomass - this is quite different for the fleet 1 CPUE and the results attest to this fundamental difference between the two fleet CPUE indices. For the R1b run the overfishing probability penalty was applied due to the appearance of high levels of overfishing in certain runs in the last 5 years of the model time-frame. As already stated, if we do not penalise against these runs they are likely to result in rapid declines in abundance in the first years of the projections at current catch levels. As can clearly be seen from Table 4.1 applying this prior does not really change the biomass status results, but it does avoid the high levels of overfishing in the final conditioning years. What this overfishing penalty does is to increase the correlation between the various parameters and avoid those combinations that result in very high overfishing rates. The 1% p.a. increasing trend in catchability scenarios (R3 & R3a) results in slightly lower median biomass ratios by the end of the modelling time-frame, stronger overall declines in abundance over the time-frame, and most clearly the lowest/highest quantiles of the status ratios - depending on the ratio and what quantile is pessimistic relative to the median - are lower/higher for these scenarios.

4.1   Fits to the data inputs

The main difference in the scenarios is with respect to which CPUE data is being used: quarterly abundance indices from fleet 1 or fleet 3. In both cases we assume a seasonal structure to the catchability parameters - this results in clear improvements in the fits to the data in both cases, but especially for the fleet 3 scenario. Clearly from Table 4.1, the choice of CPUE series has a clear effect on all the status levels and trends. However, the fits to the

data are not meaningfully affected by either estimating σ_r^2 or the assumption of a 1% annual increasing catchability trend, so we only graphically display the fits to the CPUE data for scenarios R1 and R1a in Figure 4.1.

For fleet 1 the abundance trends are arguably best captured in seasons 1 and 4 with very few observations outside the 95% predictive intervals. For fleet 2 seasons 1, 3 and 4 seem to be the best explained with the model struggling to fully capture the up and down trend in season 2. The difference is not great, but in general the OMs seem to fit better to CPUE from fleet 1 versus that from fleet 3.
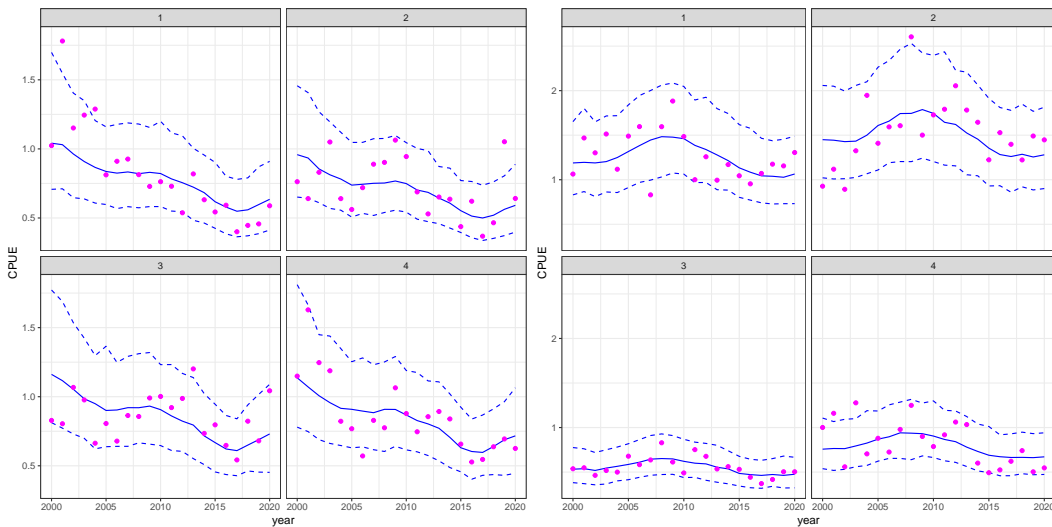

Figure 1: *Fits to quarterly CPUE data (magenta circles) in terms of approximate posterior median and 95% credible interval (blue full and dotted lines) for fleet 1 (left) and 3 (right). In both cases seasonal catchability parameters are estimated.*

As with the CPUE series scenarios, the fits to the size data are not obviously different when estimating σ_r^2 or including catchability trends, so in Figure 4.2 we compare the fits to the size data for scenarios R1 and R1a only. For all the five fisheries with size data (aggregated across seasons and years) we see an acceptable fit to the length observations. Generally speaking, the fits in scenario R1 (fleet 1 CPUE) are moderately better than those for R1a (fleet 3 CPUE). In both scenarios the length data are the same, but Figure 4.3 shows the comparative selectivity-at-length relationships for both scenarios. There are some subtle differences for 2, 3 and 4 but no major ones - this suggests that the differences in the fit to the length data is attributable at some level to the other key population dynamic and fishery variables in the model.

4.2   Estimates of σ_r^2

Runs R2 and R2a both explored the estimation of the recruitment variance parameter via the conjugate prior approach outlined in the Appendix. The informative prior (assumed inverse-Gamma) parameters where chosen to give a prior mean of 0.3 (value assumed in the assessment) and 95% credible interval between 0.2 and 0.5. The aim was not to definitely estimate this parameter but to explore, when given the freedom to deviate from the assumed

value, if there was strong information on it in the data. The prior and posterior distributions for σ_r^2 for both these OM scenarios can be seen in Figure 4.4. For the R2 scenario the posterior median (and 95% credible interval) was 0.41 (0.26–0.63); for the R2a case it was 0.31 (0.23–0.47). For the R2 scenario the posterior estimates of σ_r^2 were higher than assumed in the assessment or the prior distribution employed herein. However, there is considerable overlap in the prior and posterior distributions and the impact on the key high-level status summaries of this increase in recruitment variability was minimal (see Table 4.1). For the R2a scenario there was little to no change in the overall posterior median relative to the prior - the only obvious change was an increase in the precision of σ_r^2 in the posterior. This is a clear signal that the data in this scenario are clearly very compatible with the underlying assumed value of 0.3 in the assessment.

5   Discussion

This paper outlines a full conditioning of a suite of OMs for Indian Ocean Albacore based on ideas originally raised in [1] and further explored in [9], with respect to methods for conditioning OMs that are not directly built upon the stock assessment model structures and output. The approach uses much, if not all, of the same demographic and life-history parameters the assessment does, and often key subsets of the same data if they are required to be simulated in the OMs. It also includes key prior information on current/recent/historical status variables (MSY ratios, depletion *etc.*) to help estimate the main abundance, biomass and mortality variables required in the OMs. There is already an IOTC precedent for these type of approach - the original Skipjack tuna OM was conditioned in a similar fashion [10] using the methodology outlined in [11].

A number of alternative formulations for the OM have been explored, so that the original uncertainty grid employed for this stock is effectively included by the algorithm. Incorporating the correlation among two important variables in the grid, h and M, allows this OM to avoid generating population trajectories that could not be employed for MSE projections. Adopting a simpler model structure has allowed to better understand the interaction between data sources, again to ensure the OM obtained is fit for purpose.

The methodology outlined here appears to provide a solid foundation for an OM for this stock. Tests of the candidate MPs are being currently carried out.

6   Acknowledgements

OM scenario	Δ_{2000}	Δ_{2020}	$\tilde{\Delta}_{2020}$	$\mathcal{H}_{2020}$
Prior	0.5 (0.3–0.7)	n/a	2 (1.3–2.7)	0.6 (0.2–1)
R1	0.76 (0.61–0.97)	0.41 (0.22-0.56)	1.9 (1.02–2.58)	1.13 (0.5-3.78)
R1a	0.56 (0.4–0.72)	0.48 (0.3–0.69)	2.01 (1.42–2.61)	0.68 (0.33–1.38)
R1b	0.74 (0.61–0.9)	0.42 (0.24–0.54)	1.98 (1.16–2.49)	0.98 (0.46–2.4)
R2	0.71 (0.57–0.84)	0.41 (0.21–0.55)	1.91 (0.99–2.53)	1.22 (0.45–3.57)
R2a	0.56 (0.41–0.72)	0.47 (0.28–0.71)	2.03 (1.3–2.54)	0.65 (0.34–1.4)
R3	0.78 (0.6–0.91)	0.38 (0.15–0.52)	1.77 (0.7–2.44)	1.4 (0.58–5.06)
R3a	0.63 (0.48–0.77)	0.42 (0.25–0.59)	1.94 (1.23–2.5)	0.71 (0.35–1.45)

Table 1: *Overall OM conditioning status summary in terms of the approximate posterior median and 95% credible interval in brackets. Note the harvest rate MSY ratio is only applied in the Rxa scenarios; the overfishing probability is only applied for the Rxb scenarios.*
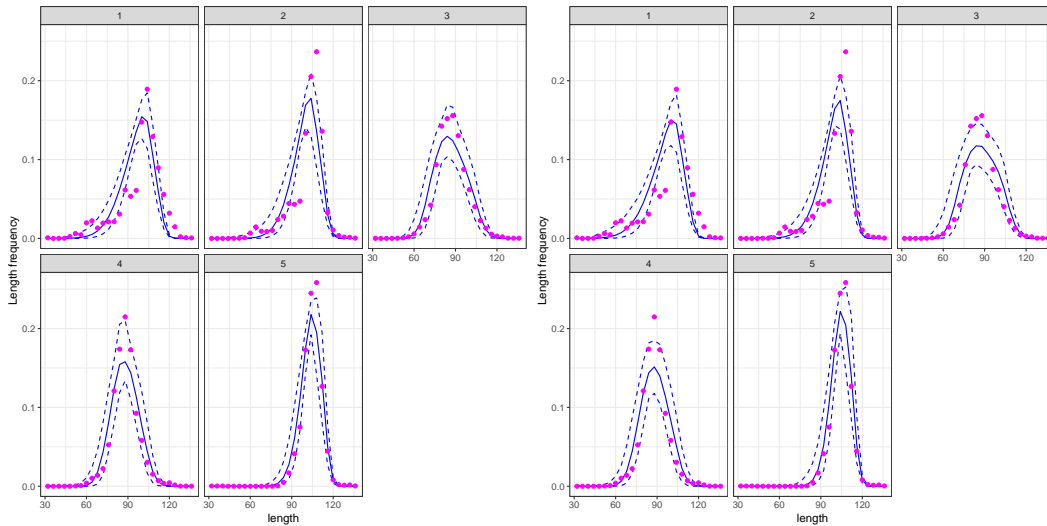

Figure 2: *Fits to aggregated length data (magenta circles) in terms of approximate posterior median and 95% credible interval (blue full and dotted lines) for fleet 1 (left) and 3 (right).*
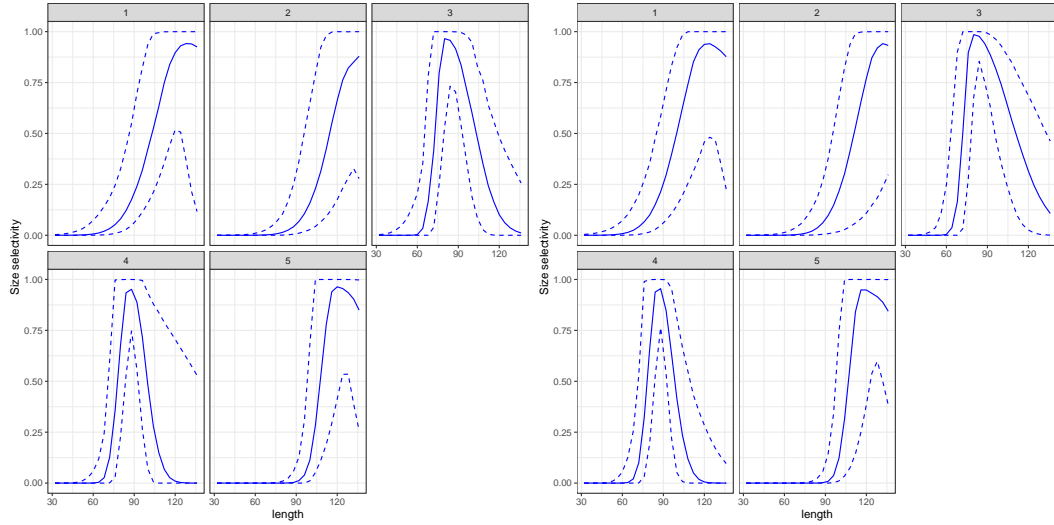
Figure 3: *Selectivity (assumed double normal) relationships-at-length for each fishery with length data: blue line is the median and the dotted blue lines are the 95% credible intervals.*
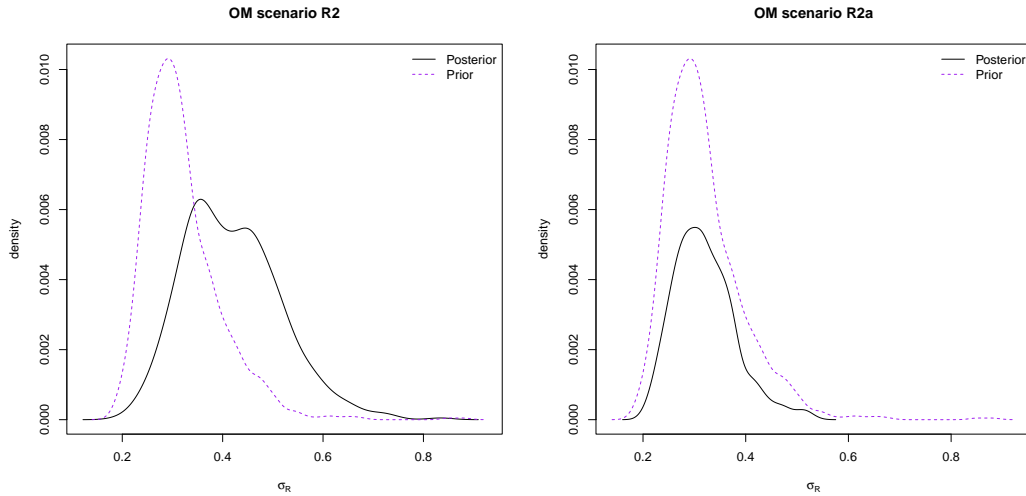

Figure 4: *Prior (dotted purple line) and posterior (full black line) for σ_r^2 for OM scenarios R2 (left) and R2a (right).*

References

[1] Hillary, R., Mosqueira, I., Preece, A., Rosa, D., Williams, A. and Edwards, C. (2021) Strategies for the conditioning of operating models for IOTC stocks, *IOTC–2021–WPM12–16*

[2] Sunnaker, M., Busetto, A., Numminen, E., Corander, J., Foll, M., *et al.* (2013). Approximate Bayesian Computation. *PLoS Comput. Biol.* **9**(1):e1002803. doi:10.1371/journal.pcbi.1002803.

[3] Wood, S. (2010) Statistical inference for noisy nonlinear ecological dynamic systems. *Nature.* **466**: 1102–1107. doi: 10.1038/nature09319.

[4] Rice, J. (2022) Stock assessment of albacore tuna (*Thunnus alalunga*) in the Indian Ocean using Stock Synthesis. *IOTC–2022–WPTmT08–09.*

[5] Mangel, M. *et al.* (2013) A perspective on steepness, reference points, and stock assessment. *Can. J. Fish. Aquat. Sci.* **70**: 930–940.

[6] Mosqueira, I. (2021) Update of the Indian Ocean Albacore Operating Model. *IOTC-2021-WPM12-09.*

[7] Wilkinson, R. D. (2013) Approximate Bayesian Computation (ABC) gives exact results under the assumption of model error. *Stat. Appl. Genet. Mol. Biol.* **12**(2): 129–141.

[8] Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. (1993) Handbook of Markov Chain Monte Carlo. *Chapman & Hall/CRC.*

[9] Hillary, R.M, and Mosqueira, I. (2023) Exploring the ABC approach for IOTC Albacore OM conditioning. *IOTC–2023–MSETaskForce*

[10] Bentley, N., and Adam. S. (2016) Management strategy evaluation for the Indian Ocean skipjack tuna fishery. Working paper IOTC-2016-WPM07-15 Rev1 presented to the 7 th IOTC Working Party on Methods, Mahe, Seychelles 11-13 November.

[11] Bentley, N. and Langley, A. (2011). Feasible stock trajectories: a flexible and efficient sequential estimator for use in fisheries management procedures. *CJFAS.* **69**(1): 161–177

[12] Kullback, S., and Leibler, R. A. (1951) On information and sufficiency. *Annal. Math. Stats.* **22**: 79–96.

Appendix A

Approximate McMC ABC algorithm

There are a wide variety of possible algorithms that can be used to generate a sample from the approximate posterior distribution [2, 3, 7]. Given the relative complexity of our likely suite of models, we consider that Algorithm D from [7] is the most applicable. It is basically an ABC-configured Metropolis-Hastings accept-reject algorithm as used in classic Bayesian McMC contexts. At time t, we have our joint parameter and process variable state $\Xi_t = \{\boldsymbol{\theta}_t, X_t\}$. We generate a proposal for a new parameter vector $\boldsymbol{\theta}'$ (and $X' = f(\boldsymbol{\theta}')$) from the pre-specified transition kernel $q(\boldsymbol{\theta}_t, \boldsymbol{\theta}')$. We define the following acceptance probability for $\Xi' = \{\boldsymbol{\theta}', X'\}$:

$$\alpha(\Xi_t, \Xi') = \min\left(1, \frac{\pi(D, X')q(\boldsymbol{\theta}', \boldsymbol{\theta}_t)\pi(\boldsymbol{\theta}')}{\pi(D, X_t)q(\boldsymbol{\theta}_t, \boldsymbol{\theta}')\pi(\boldsymbol{\theta}_t)}\right)$$

and generate a random variable $u \sim U[0, 1]$. If $\alpha(\Xi_t, \Xi') > u$ we accept the proposal and $\Xi_{t+1} = \Xi'$; if $\alpha(\Xi_t, \Xi') \leq u$ we reject the proposal and set $\Xi_{t+1} = \Xi_t$. By choosing a symmetric normally distributed transition kernel $q(\boldsymbol{\theta}_t, \boldsymbol{\theta}') = q(\boldsymbol{\theta}', \boldsymbol{\theta}_t)$ this term disappears from the acceptance rate calculations. The prior distribution, $\pi(\boldsymbol{\theta})$, is for the estimated parameters. The $\pi(D, X')$ term is our likelihood analogue or discrepancy function - it includes both the observed data (catch length composition, abundance indices) and our additional information on the key status variables contained in the suite of process variables, X (e.g. MSY and depletion ratios). The Markov chain transition kernels $q()$ are defined to use the random walk approach for sampling the posterior surface. To make the McMC algorithm more efficient we implement a Metropolis-within-Gibbs sampling approach [8], where parameters are grouped together depending on expected correlation. Each block is updated using the Metropolis-Hastings algorithm, conditional on the parameters not included in that block being fixed at their most recent value. After doing this for each block of parameters (the Gibbs sampling part of the algorithm) we have fully updated all the parameters of the model and repeat the same process many times. Random walk variances are adjusted to achieve acceptance rates around 40% - generally considered as optimally efficient [8]. A suitable burn in period is used to get the sampler moving on the surface before we decide to keep the samples, and these are then thinned to remove autocorrelation in the Markov chain. When this is done we have 1,000 samples from the approximate posterior distribution of interest and we use the Geweke statistic [8] to test for non-convergence of the Markov chains.

To augment the MCMC algorithm to accommodate the sampling of both h and M we first define $\boldsymbol{\xi} = \{h, M\}$ and an extended parameter vector $\tilde{\boldsymbol{\theta}} = \{\boldsymbol{\theta}, \boldsymbol{\xi}\}$. We factorise the extended approximate posterior as follows:

$$\pi(\tilde{X}, D) = \pi(X, D \,|\, \boldsymbol{\xi})\pi(\boldsymbol{\xi})$$

For the sampling algorithm we first sample $\boldsymbol{\xi}$ from the joint prior $\pi(\boldsymbol{\xi})$ and, conditional on this sample, we then apply the ABC algorithm described above to sample from $\pi(X, D \,|\, \boldsymbol{\xi})$. This two-stage algorithm then obtains a sample from the augmented approximate posterior distribution of $\tilde{\boldsymbol{\theta}}$. This approach thus imposes the joint prior distribution but also preserves the correlation between the steepness, natural mortality and all the other model parameters

in the overall MCMC sample vector.

Resampling recruitment variance

We additionally explore the estimation of the recruitment variance, σ_r^2, using conjugate priors. We assign a Gamma prior to the precision (the inverse of the variance) with prior parameters α_r and β_r. Given the recruitment deviations $\gamma_y \sim N(0, \sigma_r^2)$, the (conditional) posterior of $1/\sigma_r^2$ is also a Gamma distribution with parameters:

$$\tilde{\alpha} = \alpha_r + n_y/2,$$

$$\tilde{\beta} = \beta_r + \sum_{i=1}^{n_y} \gamma_i^2/2$$

where n_y is the number of years for which recruitment variations are estimated. So to generate a sample of σ_r^2, conditional on the γ_y, we sample a precision variate from this Gamma distribution defined by $\tilde{\alpha}$ and $\tilde{\beta}$ and simply invert it to obtain the next sample of σ_r^2.

Construction of discrepancy statistics

For the CPUE based abundance index, in the IOTC MSE context CPUE is the key candidate MP data input so we are going to have to simulate the CPUE data in the OMs (which requires catchability and variance parameters). For this reason we actually used a modified log-normal distribution for these data. The data were log transformed and standardised to have mean zero (this removes the need for estimating catchability at estimation run time). The catchability can be efficiently estimated post-estimation so this treatment doesn't cause issues later on in the OMs. The *overall* variance in the distribution was estimated by fitting a LOESS smoother to the log-transformed index and using the residual SD in the likelihood - in this regard this part of the discrepancy function is closer to SL than ABC.

For the length composition data we took a nonparametric approach using the concept of the Kullback-Leibler divergence (KLD, [12]): this is a measure of the divergence between in this case a discrete probability distribution P_i, relative to a reference distribution Q_i. It is defined as follows:

$$D_{KL}(P \parallel Q) = \sum_i P_i \ln\left(\frac{P_i}{Q_i}\right) \geq 0$$

with the convention that

$$\lim_{x \to 0^+} x \ln(x) = 0.$$

The KLD serves as a potentially very useful option in the ABC sense for the following reasons:

1. It is nonparametric so the underlying generating distribution of the length data does not need to be assumed

14

2. It reduces to zero when $P_i \equiv Q_i$ and increases the further Q_i diverges from P_i - by defining P_i as our observed data and Q_i our model prediction it makes an obvious candidate as a discrepancy measure of lack of fit

3. The units are interpretable. For natural logarithms the units are called nats - 1 nat is basically a difference in probability of $1/e$. This means we can set tolerance levels for how much divergence we are willing to accept in our predicted data that have a grounding in information theory

In practice we use the *negative* KLD - it reaches a maximum at perfect prediction of the data and decreases as this gets progressively worse, much like a likelihood does. If we define $p_{f,l}$ as our fishery-specific observed (annually and seasonally aggregated) length data and $\widehat{p}_{f,l}$ as our predicted length composition this part of the discrepancy function can be defined as

$$\mathcal{D}_{LF} = -\sum_f \sum_l p_{f,l} \ln \left(\frac{p_{f,l}}{\widehat{p}_{f,l}} \right)$$

We also set an effective maximum KL value of 0.8, so that no proposals of new parameter vectors are accepted if the KL value is above this maximum. The rational for a value of 0.8 is that this would roughly align with the upper confidence level of the KL value if the length data were truly distributed around the observed distribution with a multinomial effective sample size of 20.

For the CPUE data we assume an effectively lognormal distribution for the discrepancy function. For seasonal catchability models we calculate the catchability coefficient as follows:

$$\ln q_s = \mathbb{E}^y \left[\ln \left(I_{y,s} / \widehat{X}_{y,s} \right) \right]$$

and $\widehat{X}_{y,s}$ is the seasonal exploitable biomass. For the scenario with a constant catchability across all seasons it is calculated as follows:

$$\ln q = \mathbb{E}^{y,s} \left[\ln \left(I_{y,s} / \widehat{X}_{y,s} \right) \right]$$

The standard deviation for the CPUE discrepancy is taken by fitting a LOESS smoother to the seasonal log-transformed observed CPUE data and calculating the standard deviation in the residuals. The rationale being we want an overall (i.e. both observation and process error) measure of likely variation in the observed CPUE, not an observation only estimate.

For the status variables (MSY and depletion ratios) we simply defined quadratic kernels for the log-transformed variables averaged over the relevant time frame:

$$K(x, y \,|\, \epsilon) = \frac{\| \, x - y \, \|^2}{\epsilon^2}$$

In this case the tolerance for each kernel ϵ can basically be interpreted as twice the standard deviation of a normal distribution. The summation of the CPUE and length composition discrepancies and the status prior kernels makes up the overall discrepancy function $\pi(D, X)$.