Stock assessment of Indian Ocean yellowfin tuna using a Bayesian implementation of a two-age structured model.

Iago Mosqueira & Richard Hillary
AZTI Tecnalia
Imperial College London

IOTC WPTT, 7th meeting, Phuket, July 2005

1 Description of the model

1.1 Data sources

Catch data compiled by the IOTC Secretariat was used in this analysis. Catch-at-Size (CAS) by gear, fleet and year, and for 2cm length intervals, was generated by the Secretariat from the available size sampling database. These were then divided in two age groups, juveniles and adults, separated at 85cm fork length. This is the length assumed for first maturity by the Working Party, and separates clearly the catches obtained by purse seiners fishing on FADs from all other catches.

The indices of abundance used here have been presented in detail elsewhere (). Three sets of indices were available, for the Japanese longline fleet, Taiwanese longline fleet, and for the European purse seiners. For each of them, a number of possible standardised indices were available. The index selected for this model run was the Japanese longline index for the whole Indian Ocean (19689-2003). This index is assumed to reflect the abundance of the adult part of the stock only.

1.2 Population dynamics

The population model employed here has two age classes: juvenile and adult. In practice, the juvenile class incorporates effectively ages zero and one, with the second age-class incorporating everything else. The distinction, in terms of catches, between these age-classes was made largely on knowledge of where the length-frequency differences between FAD and free-school fishing appear and the assumed length at first maturity - everything to the left of this reference length is assumed to be juvenile fish; everything to the right is adult fish.

In the model, we assume the catch data are correct - for one, this simplifies the model, and improves the identifiability of the model parameters. While, almost certainly, catch-at-length data are not completely known, there are consequences when attempting to incorporate such uncertainty into the estimation procedure - particularly so, given the limited nature of the tuning data at hand.

The population is assumed to begin in unexploited equilibrium, and a Beverton-Holt stock-recruit relationship is assumed, using the virgin biomass-steepness parameterisation. Given the virgin spawning stock biomass, S_0, and the SSB per unit recruit, ρ:

$$\rho = w_1 m_1 + w_2 m_2 \frac{e^{-M_1}}{1 - e^{-M_2}} \tag{1}$$

where w_a, m_a and M_a are the mean weight, maturity ogive, and natural mortality for each of the two age-classes, the unexploited equilibrium recruitment, R_0, is then easily defined as:

$$R_0 = S_0 \rho^{-1}. \tag{2}$$

The initial numbers-at-age, $N_{1,a}$, are defined as follows:

$$N_{1,1} = R_0, \tag{3}$$

$$N_{1,2} = R_0 \frac{e^{-M_1}}{1-e^{-M_2}}. \tag{4}$$

For the rest of the years, the dynamics of the two age classes are expressed as follows. Given the spawning stock biomass, S_y, defined to be

$$S_y = \sum_{a=1}^{A^+} w_a m_s N_{y,a}, \tag{5}$$

where A^+, the plus group, is equal to two, the recruitment in the following year,

$N_{y+1,1}$, is defined by the Beverton-Holt stock-recruit relationship:

$$N_{y+1,1} = \frac{\alpha S_y}{\beta + S_y}, \tag{6}$$

where $C_{y,a}$ is the catch-at-age, and

$$\alpha = \frac{4hR_0}{5h-1} \tag{7}$$
$$\beta = \frac{S_0(1-h)}{5h-1} \tag{8}$$

For the adult ages, we have the following dynamics:

$$N_{y,2} = N_{y-1,1}e^{-M_1}(1-h_{y-1,1}) + N_{y-1,2}e^{-M_2}(1-h_{y-1,2}) \tag{9}$$

where $h_{y,a} = C_{y,a}/N_{y,a}$ is the harvest rate.

1.3 Modelling methods, parameters and assumptions

The perhaps expected stochastic multipler sometimes seen in equations such as Eq. (6) is absent for two reasons: the first is that we have no prior estimate of such a term, from survey information for example; secondly, the variance of this type of term is difficult to estimate without matching year-class strength information in the CPUE and catch data. We have no age-structure in our CPUE - it is only for adults - and trying to estimate such a term with these types of data invariably leads to both poor performance in the Bayesian estimator, and very small values of the variance - the stock-recruit uncertainty can be "explained" by the model parameters, as they stand.

1.3.1 Weight, maturity & natural mortality

Table 1 gives the particular values used in this assessment for the weight, maturity and natural mortality at age. Mean weight for each age was estimated as an average of the weights of 5cm classes, obtained from the length-weight relationship used by IOTC, weighted by the proportion in abundance of each class according to natural mortality decay. The same procedure was used to estimate mean natural mortality for the juvenile age group.

Knife-edge maturity is assumed, which effectively ties the spawning stock biomass to the adult population only.

Table 1: *Weight, maturity and natural mortality values used.*

Variable	Weight (kg)	Maturity	Natural mortality (yr^{-1})
Age 1	10	0	1.2
Age 2	26	1	0.6

1.3.2 Likelihood

Given the single CPUE series we have for the adult fish, we use a normal-log likelihood function for the relative to absolute relationship likelihood. Given the catchability parameter, q, the CPUE, I_y, and the numbers of adults, $N_{y,2}$, the following likelihood function is employed:

$$\mathcal{L}(I_y \,|\, q, S_0, \sigma_q^2, N_{y,a}, C_{y,a}) \propto \prod_{y=1968}^{2003} \frac{1}{\sqrt{2\pi\sigma_q^2}} \exp\left(-\frac{(\ln(I_y) - \ln(qN_{y,2}))^2}{2\sigma_q^2}\right), \tag{10}$$

where σ_q^2 is the variance in the absolute to relative abundance relationship.

1.4 Prior distributions

The Bayesian paradigm requires the definition of prior distributions for each of the parameters. For the catchability, q, we actually estimate $\ln(q)$ in the McMC routine, and we apply a normal prior to this parameter:

$$\pi(\ln(q)) \sim N(0, 100). \tag{11}$$

This prior can be considered non-informative, as the Jeffrey's non-informative prior for such a parameter as q is uniform on a log-scale (Jeffreys, 1961) - the wide normal distribution used here mimics this uniform-log prior and, as we shall see later, gives us a distinct McMC advantage.

For the variance parameter, σ_q^2, an appropriate non-informative distribution is the inverse-gamma distribution of the form:

$$\pi(\sigma_q^2) \sim IG(1.5, 1), \tag{12}$$

which has upper and lower 95% confidence bounds of 9.44 and 0.21, respectively.

For the virgin biomass parameter, assigning a prior distribution is more complicated. There are no obvious non-informative or quasi non-informative priors for this parameter, given its complex presence in the model and the likelihood function. Sometimes a uniform prior for $\ln(S_0)$ is employed but this is neither non-informative nor overtly sensible in some cases. Our approach tries to use previous information on the variability of this parameter, as well as common-sense given the data at hand. Firstly: the virgin recruitment, R_0, must be *a priori* greater than the maximum recorded catch of juveniles of age 0. Thus, given the relationship between virgin recruitment and biomass in Eq. 2, we have an immediate minimum bound on S_0. Running a Schafer surplus production model on the yellowfin data, the suggested CV from the maximum likelihood fit was 0.15. To ensure we are not being too restrictive in this prior we assume a prior CV of 0.3 for S_0. This gives us a suitable standard deviation for $\ln(S_0)$, and so now we have the following log-normal prior for S_0 (in thousands of individuals):

$$\pi(S_0) \sim LN(\ln(15000), 0.3), \tag{13}$$

which has upper and lower 95% confidence limits of 2,463 and 36,992 which satisfies the requirement that R_0 should be at least above the maximum juvenile catch.

Although this is not a "true" non-informative prior, it is in our mind much more applicable than other commonly used options such as a log-uniform prior. Even if we restrict the log-uniform prior with respect to the maximum juvenile catch, we are still giving this, the most unlikely value of the virgin biomass, the highest prior weighting. This is because $\pi(S_0) \sim S_0^{-1}$ for all S_0 above this cut-off value. We shall see later that this prior does not exert undue influence on the posterior values of S_0.

1.4.1 McMC algorithm

The McMC algorithm used in this assessment is written in C++, and implements the Metropolis-within-Gibbs sampling algorithm under the following scheme:

- The log-catchability, $\ln(q)$, is resampled from its normal conditional posterior (because the normal prior is *conjugate* to the likelihood) in a Gibbs move.
- The variance parameter, σ_q^2, is resampled from its inverse-gamma conditional posterior (because the inverse-gamma prior is *conjugate* to the

likelihood) in a Gibbs move.

- The virgin biomass, S_0, is resampled using the random walk Metropolis-Hastings algorithm.

The whole estimation procedure runs in around two minutes on a computer containing a 2.4 GHz processor, and convergence of the resulting Markov chains on the posterior was verified using standard software in R - but visual methods can also be used with experience. This speed of convergence is aided by our choice or priors for $\ln(q)$ and σ_q^2 - because these priors are conditionally conjugate tp the normal-log likelihood, we simply draw each parameter directly from this known conditional posterior, without having to revert to Metropolis-Hastings sampling for all the parameters, which slows down the efficiency of the sampler due to auto-correlation and slow mixing of the chains.

2 Results

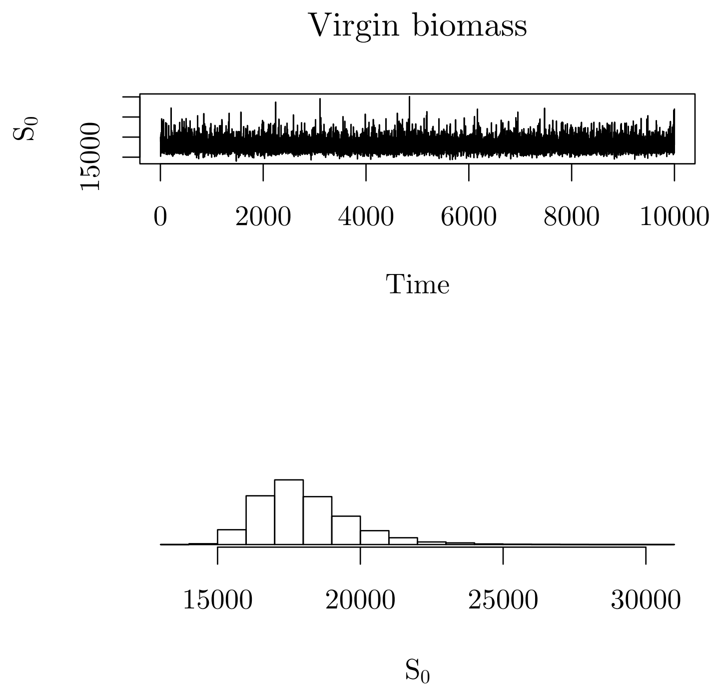

Figure 1: McMC chain and posterior pdf of the estimated virgin biomass (S_0).

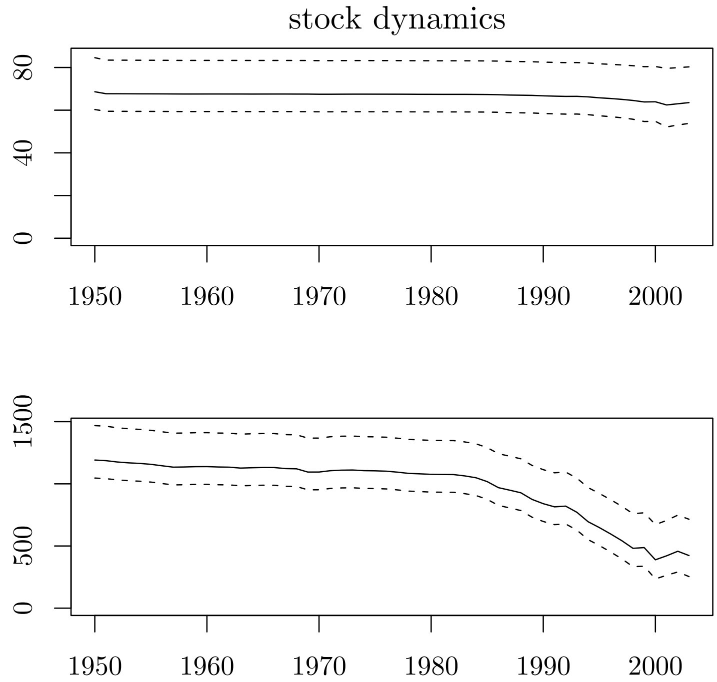

Figure 2: Estimated population trajectories for both age groups

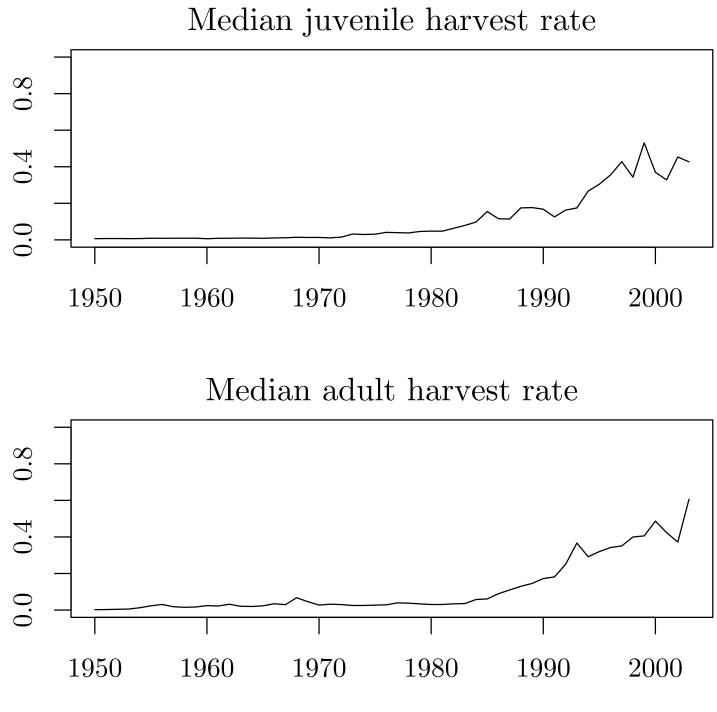

Figure 3: Estimated harvest rates (proportion of the stock being caught) for both age groups

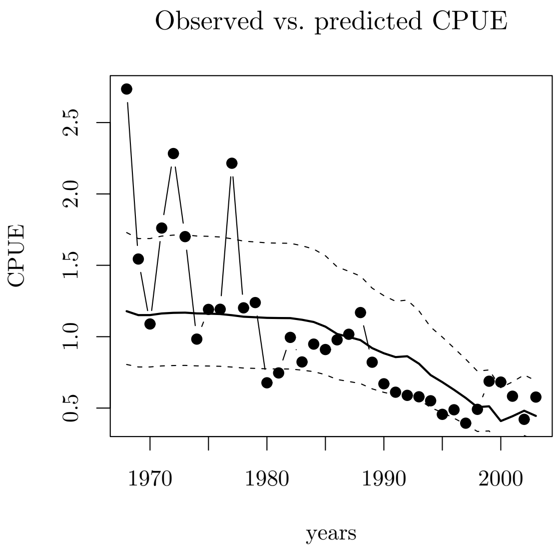

Figure 4: Observed (full line) and predicted index of abundance, with 95% credibility intervals

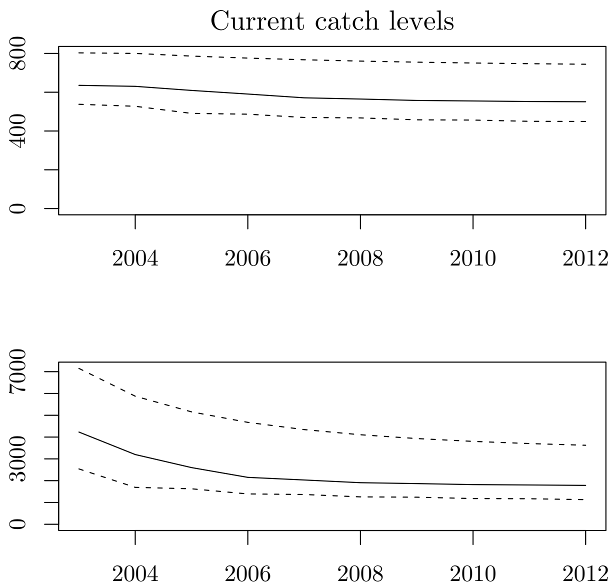

Figure 5: Projections of future biomass for both ages under current catch levels.

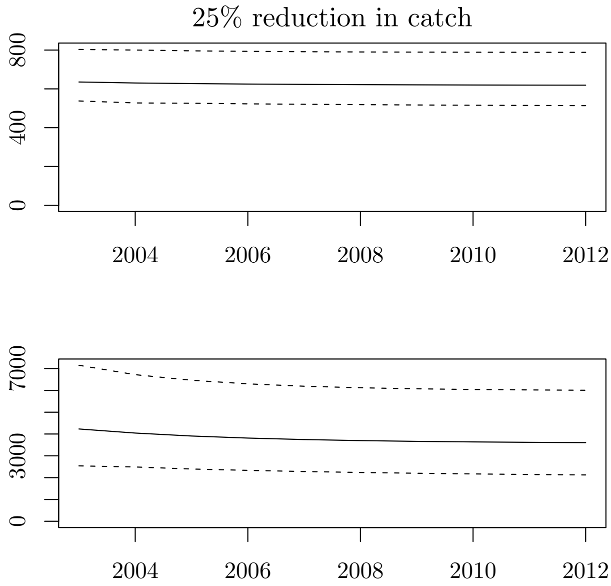

Figure 6: Projections of future biomass for both ages if current catch levels were reduced by 25%