CCSBT-ESC/1107/12

Updated technical specifications and performance analyses for MP1

Richard Hillary
Ann Preece
CSIRO CMAR & Wealth from Oceans National Research Flagship

Contents

Abstract

Given the most recent CPUE and aerial survey data the model and estimation scheme that form the basis for MP1 are assessed. The underlying biomass random effect model of MP1 explains both the CPUE and aerial survey data well. To assess the consistency of the recruitment estimates in the model we also integrate the SAPUE index into the estimation scheme (though not into the actual MP) and find strong consistency between the aerial survey and the SAPUE data when they overlap, as well as confirmatory information on the low recruitments and high exploitation rates seen in the early 2000s in other data. A minor change to the harvest control rule in MP1 is suggested so as to include as much of the aerial survey data as possible and a comparison with the old structure of MP1 on the updated OM is undertaken.

1 Introduction

This paper is primarily an accompaniment to the MP1 performance paper [1] detailing the estimation performance of the BREM (**B**iomass **R**andom **E**ffects **M**odel) part of that candidate management procedure. The model is updated with the latest MP data and a detailed analysis of the fitting performance and trends from the model is undertaken. The SAPUE index [2] is a useful indicator of trends in juvenile abundance on the surface fishing grounds. We integrate this index into the BREM model (with the CPUE and aerial survey data) to address consistency of trend (with the aerial survey) and what information it might hold on the juvenile part of the stock in the years when the aerial survey was not undertaken. A minor change to the harvest control rule (HCR) underlying MP1 is proposed that uses the random effect structure of the BREM model to ensure that the most up to date CPUE and aerial survey is used in the MP.

2 Estimation performance of BREM framework

To evaluate the performance of the BREM framework on the most recent data sets we first recap the specifics of the population and probability models and then, using Bayesian MCMC and posterior predictive tools, see how well the model fits to and explains the variation in the CPUE and aerial survey data.

2.1 The BREM population and probability model

The core population model is itself very simple: recruitment (R_y) and adult (B_y) biomass are related as follows:

$$B_{y+1} = R_y + g_y B_y, \tag{1}$$

where g_y is the adult biomass net growth effect (encompassing natural mortality, surplus production and exploitation effects). For the recruitment process the following model is assumed:

$$R_y = \exp\left(\mu_R + \epsilon_y^R\right), \tag{2}$$

with $\epsilon_y^R \sim N\left(-\sigma_R^2/2, \sigma_R^2\right)$. For the g_y a conceptually similar model is assumed and

$$g_y = \exp\left(\mu_g + \epsilon_y^g\right), \tag{3}$$

with $\epsilon_y^g \sim N\left(-\sigma_g^2/2, \sigma_g^2\right)$. For the aerial survey data I_y^{AS} a lognormal relationship to the recruiting biomass is assumed but with a one-year delay: $I_y^{AS} \sim LN\left(q^R R_{y+1}, \sigma_{AS}^2\right)$. The reason for this delay is because we assume that the aerial survey covers ages 2 to 4 and that the CPUE covers ages 4 to 12/18. To make sure that we are more likely to detect the movement of a signal in the aerial survey appearing in the CPUE data this delay is assumed as R_y represents the recruitment biomass contribution to the adult biomass (assumed covered by the CPUE). The situation is simpler for the CPUE likelihood and these data are assumed log-normally distributed about the adult biomass: $I_y^B \sim LN\left(q^B B_y, \sigma_B^2\right)$.

The model as it stands is non-identifiable which was explored at length in [3]. Without at least some information as to the ratio of the recruit and adult catchability parameters q^R/q^B then it will be impossible to identify how much recruitment affects biomass trends and how much the net growth of the biomass affects the biomass trends. To solve this problem we look to the output from the OM runs. From the grid runs we can extract the log catchability parameters for both the aerial survey and the CPUE data. Given the grid samples over parameters that will clearly alter this ratio (natural mortality, steepness, age range covered by the CPUE) we bootstrapped the mean difference in the log-catchabilities to obtain a best estimate of this ratio. The bootstrapped mean ratio was very precise (around a 4% CV) with an expected value of $q^{AS}/q^{CPUE} = 13626.28$. However, we need to account for the fact that the CPUE in the OM is in *numbers* but here we are trying to relate biomass to biomass. To take account of this in our catchability ratio consider the following ratio:

$$\psi_q = \frac{\sum\limits_{i=a_l}^{a_u} \pi_i^s w_a}{\sum\limits_{i=a_l}^{a_u} \pi_i^s} \tag{4}$$

where π_i^s is the survival probability from age 0 to age i and a_l and a_u are the minimum and maximum ages observed in the CPUE, respectively. This ratio is readily calculable from the grid files outputted from the OM. For each sampled grid cell this ratio was calculated and then a bootstrapped mean and CV were calculated, to robustify the estimates given the banding by M grid option. As with the q ratio estimates the numbers were very precise: a mean and CV of 0.0616 and 0.026, respectively. Assuming $q^B = 1$ this lead to a value of $q^R = q^{AS}/q^{CPUE} \times \psi_q = 838.21$. In terms of the recruitment biomass variance term a value of $\sigma_R = 0.376$ is employed, as this corresponds to the amount of variation one would expect to see in the aerial survey index (covering ages 2, 3 and 4 for a selectivity of 0.5, 1 and 1, respectively) *only* due to variation in recruitment at age 0 with an assumed SD of 0.6 as per the OM. This was calculated by running a stochastic per-recruit unexploited population for 100 years (with the mean M-vector from the OM) and calculating the SD in the population covered by the aerial survey. The reason for choosing a value of $\sigma_g = 0.246$ was based on a CV of 0.25 which is the mid-point of the process error applied when simulating the CPUE data. In terms of the observation error assumed in the estimation scheme CVs of 0.15 and 0.2 were assumed for the aerial survey and the CPUE data, respectively, given the recent estimates from the aerial survey and the minimum value assumed in the OM conditioning.

The actual parameters to be estimated are μ_R, μ_g, ϵ_y^R and ϵ_y^g. To avoid identification issues with the recruitment in the first year and net growth year effects in the last year, respectively, they were penalised to have mean zero across years (with $-100 \times |\mathbb{E}[\epsilon_y^\bullet]|$ extracted from the log-likelihood). Although maximum posterior density estimates were used in the MP testing, full MCMC (Metropolis-within-Gibbs) routines were developed to explore the parametric and process variable uncertainty in the underlying models in this phase - the chief reason being that we can obtain more detailed information about the variation in the derived trends such as stock growth, recruitment and biomass which are not retrievable from the ADMB runs. While using the term random effect, to be clear this model is more of a Bayesian hierarchical model: a Dirac/point hyperprior is defined for the variance hyperparameters $\sigma_\bullet^2$, which then form the priors for the $\epsilon_y^\bullet$ parameters. This contrasts with the strict view of a random effects model which utilises expectation/maximisation to estimate all the key parameters: expectation where the joint penalised likelihood of the $\mu_\bullet$ and $\epsilon_y^\bullet$ is integrated over the $\epsilon_y^\bullet$ and maximisation where this marginal likelihood is then maximised for the $\mu_\bullet$.

2.2 Performance on historical data

Summarising the marginal posteriors for the parameters μ_R and μ_g; these parameters have mean (and SD) of -1.55 (0.053) and -0.43 (0.048), respectively, with fairly strong negative correlation between these two parameters (-0.53) as one would expect if recruitment makes a significant contribution to the exploitable biomass. The estimated trends in recruit biomass, adult biomass and biomass growth can be seen in Figure 1. For the relative recruitment biomass estimates we observe a sharp decline around 1998 (as seen in 1997 in the aerial survey) hitting the lowest level in 2000. From 2001 to 2004 the estimates are driven by the prior and penalty terms given the absence of data in the aerial survey with the levels of recruitment in 2005 to 2008 staying around the low level but with an upturn in 2009. The aerial survey decreased in 2009 (hence the 2010 reduction in recruitment biomass) but this was followed by an increase to a level above the historical average in 2011 (given the 2010 higher aerial survey) and to levels comparable with the historical maximum in 2012 (given the very high 2011 aerial survey). In the years where there are data to estimate the recruitment trend the CVs ranged from 0.131 to 0.145. For the relative adult biomass estimates we must first point out that we assume that $B_{1994} = I_{1994}^B/q^B$ and that it is known without error (there are no data to estimate it and we assume a relative abundance model anyway). As one would expect the adult biomass trend follows the CPUE series trend (including the gradual decline from 2002-2007 and the sudden upturn in 2008 continuing into 2009 and 2010). The CVs in the estimates (excluding 1994) range from 0.122 to 0.21 with a sustained increase in uncertainty in the middle of the range given the uncertain recruitment dynamics. The biomass growth estimates oscillate below the mean until rising well above it from 1999-2001, after which they show a marked decline as they alone can explain the biomass decline seen in 2002-2007 as recruitment has already dropped to the lower level by 1998. Clearly the sudden increase in 2008 in the biomass cannot be explained by recruitment and so the biomass growth parameter in 2007 increases to a value well above the mean in this year. The continued increase in biomass (given the CPUE) from 2008 to 2009 and 2010 seems to be attributable to biomass growth also - the estimate of biomass growth in 2008 is still above the mean and the recruitment estimate from 2008 is the same as 2007 and not above the average. The increases in biomass from 2010 to 2011 and 2011 to 2012 seems to be mostly driven by the higher than average recruiting biomass in 2011 and very large 2012 recruitment biomass. The estimates of biomass growth in 2010, 2011 and 2012 are driven by both the prior and the penalties and should not be viewed with close scrutiny.

In terms of fits to the data Figure 2 shows a summary of the estimators performance in this regard. For the aerial survey data they are generally fitted quite well but the extremes in these data (the apparently higher variance earlier on) are not fitted so well, presumably given the assumed value of σ_R. For the CPUE data they are also fitted quite well but the model cannot fit the more extreme changes observed in the CPUE - in 2007 the low level sits outside the 95% credible interval - but the median fitted CPUE is much smoother than the observed data. This again is due to the natural constraints placed upon both the recruitment and biomass growth effects via σ_R and σ_g, respectively.

From a Bayesian and MCMC perspective, one final analysis is to check the predictive power of the posterior model - how well does the model not just fit but "explain" the data - to be satisfied that the model is at least able to adequately deal with the historically observed data. An established way to do this is to perform a *posterior predictive* analysis [4]: data are simulated from the likelihood (given the posterior sample) and positive discrepancy statistics, Δ (in this case the absolute median deviation), denoting in some way the "closeness" of the simulated and real data to the model prediction, are calculated. Bayesian p-values [5] can then be calculated as the probability that the simulated data are more "extreme" than the real data: $p\left(\Delta^{\text{model}} > \Delta^{\text{data}}\right)$. Bayesian p-values around 0.5 suggest good performance, in that the model is explaining (not just fitting) the data well - values above and below 0.5 can be indicative of the presence of unaccounted for process error and over-fitting, respectively. For the aerial survey data the p-value was 0.43 and for the CPUE data 0.51 so the model is explaining both data sets fairly well. It is often useful to plot the data and model-predicted discrepancy statistics and Figure 2 shows a summary of these - both form a fairly circular cloud around the $y = x$ line as we would like, with no obvious strange visual patterns. On the whole this suggests both the aerial survey and CPUE parts of the probability model are performing well on the historical data.

3 Consistency and information value of SAPUE index

The SAPUE abundance index [2] is a standardised measure of the abundance observed in the commercial operations of the surface fishery on the fishing grounds. While lacking a scientific design like the aerial survey, and with much less coverage of the juvenile population, it is a useful and annually reported index. One other useful feature of the SAPUE index is that it covers some of the years missing - 2001 to 2004 - in the aerial survey.

3.1 Integrating the SAPUE index into the BREM scheme

The SAPUE index, I_y^S, is integrated into the BREM estimation scheme in a very similar manner to the aerial survey and we assume that $I_y^S \sim LN\left(q^S R_{y+1}, \sigma_S^2\right)$. This was considered a reasonable assumption given what the aerial survey observes and what the surface fishery catches in terms of age classes. In this model formulation $\ln q^S$ is assigned a normal prior mean μ_{q^s} and variance $\sigma_{q^S}^2$ which results in the following normal conditional posterior:

$$p\left(\ln q^s\right) \sim N\left(\left(\frac{\mu_{q^s}}{\sigma_{q^S}^2} + \frac{\sum I_y^S / R_{y+1}}{\sigma_S^2}\right) \times \left(\frac{1}{\sigma_{q^S}^2} + \frac{\aleph^S}{\sigma_S^2}\right)^{-1}, \left(\frac{1}{\sigma_{q^S}^2} + \frac{\aleph^S}{\sigma_S^2}\right)^{-1}\right), \tag{5}$$

where $\aleph^S$ is the number of data points used from the SAPUE index. Figure 3 shows the recruitment biomass, sub-adult biomass and net growth parameter summaries with the SAPUE index now included. While there are minor changes in the biomass and biomass growth trends, as well as increases in the precision of both, the major differences with the addition of the SAPUE index are in the recruitment estimates. From Figure 4 the recruitment biomass levels from 2003 to 2005 are all notably lower than those seen in Figure 1 - this is because the aerial survey did not cover those years and previous estimates were driven by the random effect structure and penalties. These low estimates of the sub-adult recruiting biomass in 2003-2005 tally very well with the weak year-classes seen in the length frequency data and estimated in the OM for the years 2000-2003 [6]. Given the year-class estimates from the OM are for age 0 and the SAPUE is assumed to cover ages 2 to 4 this is what one would expect: those weak-year classes moving through the SAPUE data in 2003 and 2004. The decline in the sub-adult biomass in the mid-2000s is now estimated to be more influenced by those weaker year-classes as the growth estimates are not as low in those years as they are in the original model without the SAPUE data (see Figures 1 & 3).

From Figure 4 is is clear that the model can fit the SAPUE data very well, without any clear decrease in the quality of the fits to the CPUE and in particular the aerial survey data. Also, given p-values of 0.42, 0.57 and 0.51 for the aerial survey, SAPUE, and CPUE data, respectively, the model is also still explaining all the data very well. While one can observe by visual inspection something of the agreement between the aerial survey and SAPUE indices in the common years, we can clearly see that from a modelling/assessment-type viewpoint they do seem to match up well. This is encouraging as it gives us some belief that what is happening on the fishing grounds (in terms of relative abundance trends) is not likely to be very different from what is happening in the wider GAB (observed by the aerial survey).

3.2 Information gain from inclusion of the SAPUE index

The estimated low recruitments of the late 1990s/early 2000s have always had a strong influence on the behavior of the OM and previous management advice. While not without interpretation issues, the SRP tagging data of the early 2000s showed high levels of exploitation on these weaker cohorts as they moved through the surface fishery. Given the aerial survey is missing for the majority of the key years in which these weaker cohorts would have been observed we cannot verify that (relative) exploitation rates were elevated in these years. However, given the SAPUE index covers most of these missing years, and with the observed consistency of the index with the aerial survey, we can use our revised estimates of the recruiting and sub-adult biomass (see Figure 3) to explore the issue.

We obtain estimates of the (relative) exploitation rates for the surface and long-line fisheries as follows:

- Surface fishery: given catch biomass, C_y^{surf}, and the (relative) recruiting biomass, R_y, relative exploitation rate is given by $\xi_y^{\text{surf}} = C_y^{\text{surf}}/R_{y+1}$ as recruiting biomass is related to the exploitable surface fishery biomass in the previous year.
- Long-line fisheries: given catch biomass, C_y^{ll}, and the (relative) sub-adult biomass, B_y, relative exploitation rate is given by $\xi_y^{\text{ll}} = C_y^{\text{ll}}/B_y$.

Looking at Figure 5, beyond the high estimate of relative harvest rate in 1999 (driven by the low recruitment biomass estimated in 2000), the results seem to agree with the tagging data (both 1990s and 2000s). In the 1990s the exploitation rates were lower and as the weak year classes move into the surface fishery in 2003 and 2004 we see a large rise in the exploitation rates to 60-80% of mean levels (from 1993-2011). These weaker year-classes, which experienced higher levels of fishing pressure compounding their weakness, then move into the long-line exploitable biomass in the mid-2000s reducing the abundance and giving rise to higher than average (50-60%) levels of exploitation rate. Quota cuts in 2006 and then 2009, along with a return to better year-classes in the latter part of the decade, then serve to decrease the relative exploitation rates to and even below (for the surface fishery in 2010 and 2011) the average level of around the last 20 years.

4 Update to MP1 HCR and comparison with previous version

A minor alteration is proposed for the HCR in MP1 that utilises the random-effect structure of the model to include as much data as possible. The core structure of the HCR is the same, but there is a small change to the years when back-averages or reference levels are calculated. As with the original version of MP1 the TAC is weighted average of the previous TAC and the TAC from the BREM HCR:

$$TAC_y = \psi_y TAC_{y-1} + (1-\psi_y) TAC_y^{\text{brem}}, \tag{6}$$

where $\psi_y \in [0,1]$ is a memory weighting term set at $\psi_y \equiv 0.5$ for all years. As before

$$TAC_y^{\text{brem}} = C_y^{\text{targ}} \times \Delta_y^R \times \Delta_y^g, \tag{7}$$

but the biomass relative to the target level, B^*, is calculated in year y not $y-2$ as before.

$$C_y^{\text{targ}} = \begin{cases} \delta \left[\dfrac{B_y}{B^*}\right]^{1-\varepsilon_b} & \text{for} \quad B_y \geq B^* \\ \delta \left[\dfrac{B_y}{B^*}\right]^{1+\varepsilon_b} & \text{for} \quad B_y < B^* \end{cases} \tag{8}$$

and $\varepsilon_b \in [0,1]$ represents the degree to which the response to a biomass level above or below the target level B^* is asymmetric.

The recruitment adjustment Δ_y^R is defined as follows:

$$\Delta_y^R = \begin{cases} \left[\dfrac{\bar{R}}{\mathcal{R}}\right]^{1-\varepsilon_r} & \text{for} \quad \bar{R} \geq \mathcal{R} \\ \left[\dfrac{\bar{R}}{\mathcal{R}}\right]^{1+\varepsilon_r} & \text{for} \quad \bar{R} < \mathcal{R} \end{cases} \tag{9}$$

and $\varepsilon_r \in [0,1]$ is the level of asymmetry in response to the current moving (arithmetic) average - and this has been changed to include up to year y - recruitment levels, $\bar{R}$:

$$\bar{R} = \frac{1}{\tau} \sum_{i=y-\tau+1}^{y} R_i, \tag{10}$$

of length τ relative to the average, $\mathcal{R}$, calculated over the years for which the estimates are based on observed data. The final term is the stock growth term and no asymmetry in action is assumed so

$$\Delta_y^g = \left[\frac{\bar{g}}{\mathcal{G}}\right]^\gamma, \tag{11}$$

where

$$\bar{g} = \frac{1}{\tau}\sum_{i=y-\tau+1}^{y} g_i, \tag{12}$$

and $\mathcal{G}$ is the mean value of g_y over which the estimates are based on real data. The term $\gamma \in [0, 1]$ in Eq. 11 is an importance weighting term.

The reference catch level, δ, is the tuning parameter. Key HCR parameters that are kept fixed are:

- B^*: target relative biomass level (effectively a target CPUE level) of 1.2 as before.
- τ: length of the moving averages for the recruitment and biomass growth parameters set to 5 in all cases.
- The biomass asymmetry parameter ε_b is set equal to 0.5
- The recruitment asymmetry parameter ε_r is set equal to 0.75
- γ: importance weighting of the biomass growth adjustment in the HCR set to 1 in all cases.

To explore the implications of this minor change to MP1 both the new and previous versions were tuned with a rebuild probability of 0.7 by 2035 and 2040 with an assumed 1-year lag, using the updated OM grid [6] `basehupsqrt`. Figure 6 shows the SSB and catch performance summaries for the old and new MP1 HCR formulations. The comparative behaviour of each of the formulations is consistent across the two tuning criteria. The old version increases the catch levels at a slower rate in the early years (it has an innate lag in the detection of positive signals on recruitment); as a result, several years later we see slightly lower SSB rebuilding with the new formulation. The new version begins to reduce the TAC increases in the lower quantiles sooner than the old version and, while attaining slightly lower SSB rebuilding by 2035 or 2040, does achieve higher lower quantiles for the SSB as a result.

5 Summary

The BREM part of MP1 was fitted to the updated CPUE and aerial survey data using a Bayesian framework, to assess the estimation and predictive performance of the model. Using a posterior predictive analysis it was clear that the model explains both data sets well - it fits the series well and and the predicted data exhibit both structure and variance levels very similar to those observed in the real data. The conclusion is that there is nothing to suggest that the BREM approach is inappropriate from a purely statistical viewpoint.

To assess both the consistency and information content in the SAPUE index it was integrated into the BREM framework as a secondary (relative) abundance index for the juvenile biomass. The series was explained well within the augmented model (again using a posterior predictive analysis), showed very good consistency with the aerial survey in years of overlap, and also provided information on low recruitments in the late 1990s/early 2000s not seen in the aerial survey data because the survey was not active at this time. This adds to the already existing evidence of these weak year-classes seen in various data sets and estimated in the OM [6]. In line with the tagging data from the 2000s it also suggested that the exploitation rates on these weak year-classes as they moved through the surface then the long-line fisheries were also well above historical (1993-2010) average levels by around 50-80%.

A minor change to the HCR employed in MP1, which basically brings forward the times at which biomass reference levels and recruitment back-averages are calculated, was outlined. This change sought to make use of the random effect structure of the model to both use the most recent data possible and make TAC decisions based on the most up to date estimates of the juvenile and sub-adult biomass levels. Both the previous and suggested updated versions of MP1 were tuned to the updated OM reference grid `basehupsrt` assuming an interim rebuild probability of 0.7 by 2035 and 2040 with a one-year lag. The new version reacted sooner to the positive recovery signals seen recently and had slightly higher early TACs with a minimal negative impact on SSB rebuilding over the short (2022/2025) term. In the longer term the new version had lower TACs at the lower quantiles (as it reacts faster to negative signals in the lower SSB quantiles) and was able to maintain the lower SSB quantiles at a higher level than the old version. In summary, the change to the MP1 HCR acts as we expected (acts faster on positive and negative signals) and shows no apparent decrease in overall rebuilding performance.

6 Acknowledgements

This work was funded by the Fisheries Research and Development Corporation (FRDC) and the Wealth from Oceans National Research Flagship.

References

[1] B, Barnes, R. M. Hillary, M. Chambers, P. Tennant, A. Preece, C. Davies and G. Begg. (2011) Results of the performance of MP1. CCSBT-ESC/1107/13, Bali, Indonesia.

[2] J. Farley and M, Basson. (2011) Commercial spotting in the Australian surface fishery, updated to include the 2010/11 fishing season. CCSBT-ESC/1107/16, Bali, Indonesia.

[3] Trenkel, V. (2008) A two-stage biomass random effects model for stock assessment without catches: What can be estimated using only biomass survey indices? Can. J. Fish. Aquat. Sci. **65**(6): 1024–1035

[4] Gelman, A., Carlin, B.J., Stern, H.S. and Rubin, R.B. (1995) Bayesian data analysis. Chapman and Hall, London.

[5] Meng, X. (1994) Posterior predictive p-values Ann. Stat. 22, 1142–1160.

[6] R.M. Hillary, A. Preece, B. Barnes, C. Davies and G. Begg. (2011) Reconditioning of the SBT operating model: exploratory data analysis, fitting performance, and current stock status. CCSBT-ESC/1107/11, Bali, Indonesia.

7 Figures

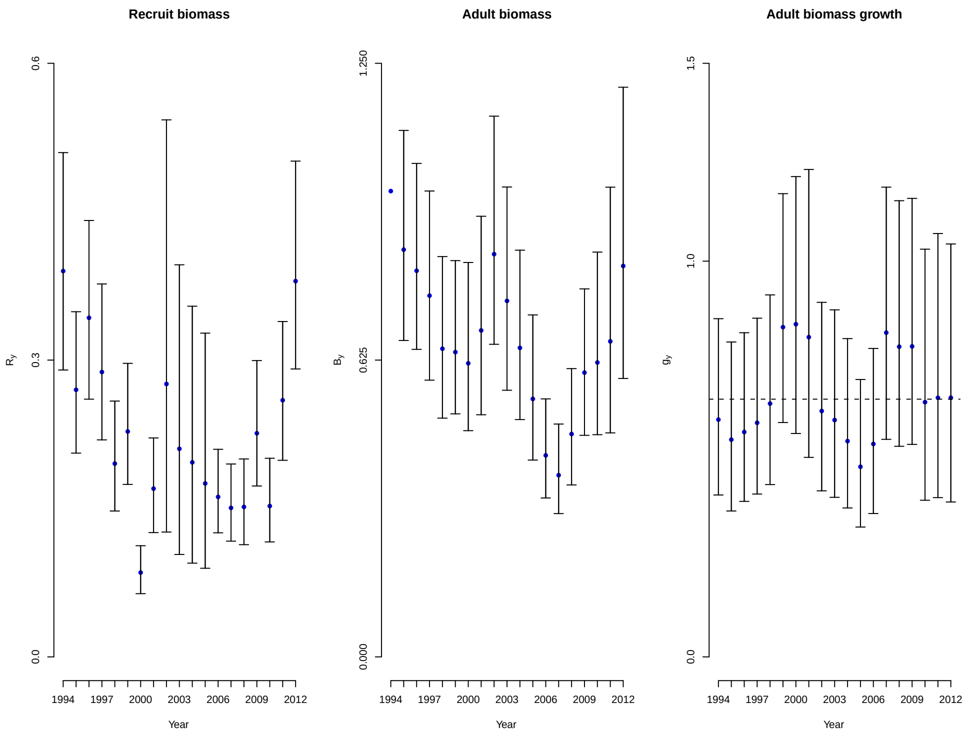

Figure 1: *Summary (median, circle; whiskers, 95% credible interval) of the relative recruitment biomass (left), relative adult biomass (middle) and net biomass growth (right, dotted line denotes the mean) using the aerial survey and the CPUE data.*

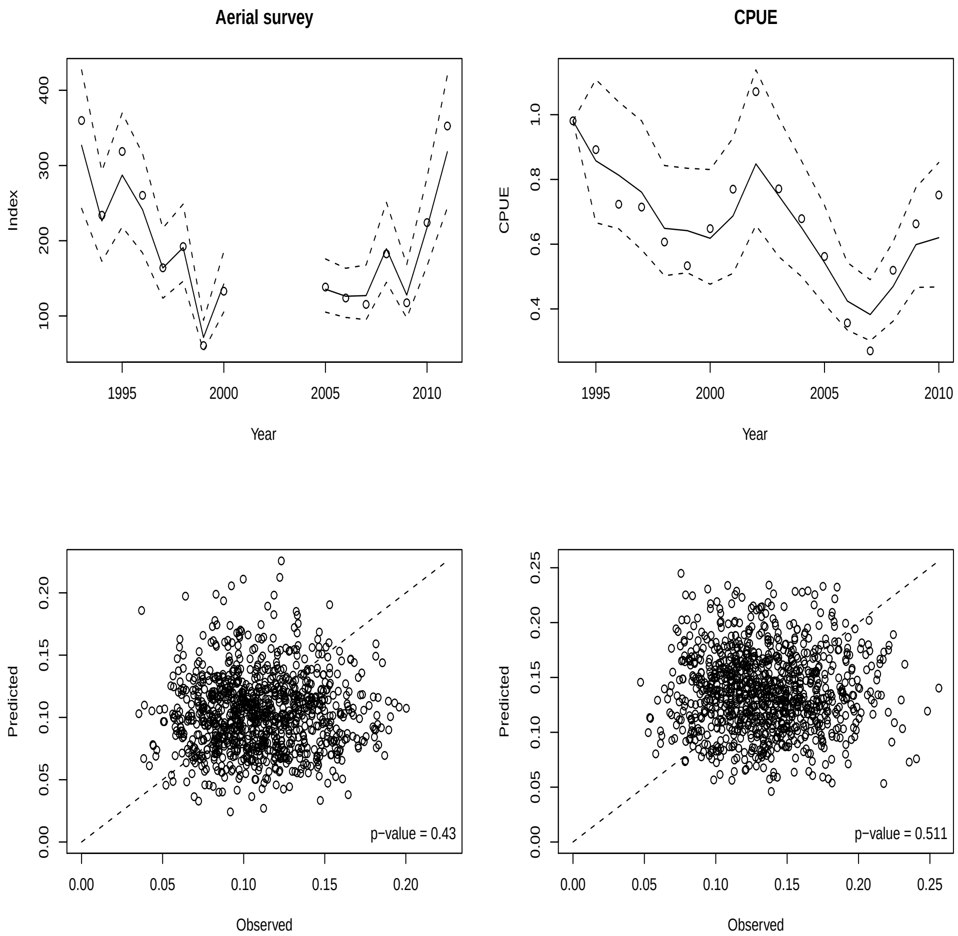

Figure 2: *Fitting and posterior predictive summaries the aerial survey data (left) and the commercial CPUE data (right). The points are the data with the full and dashed lines representing the median and 95% credible intervals, respectively.*

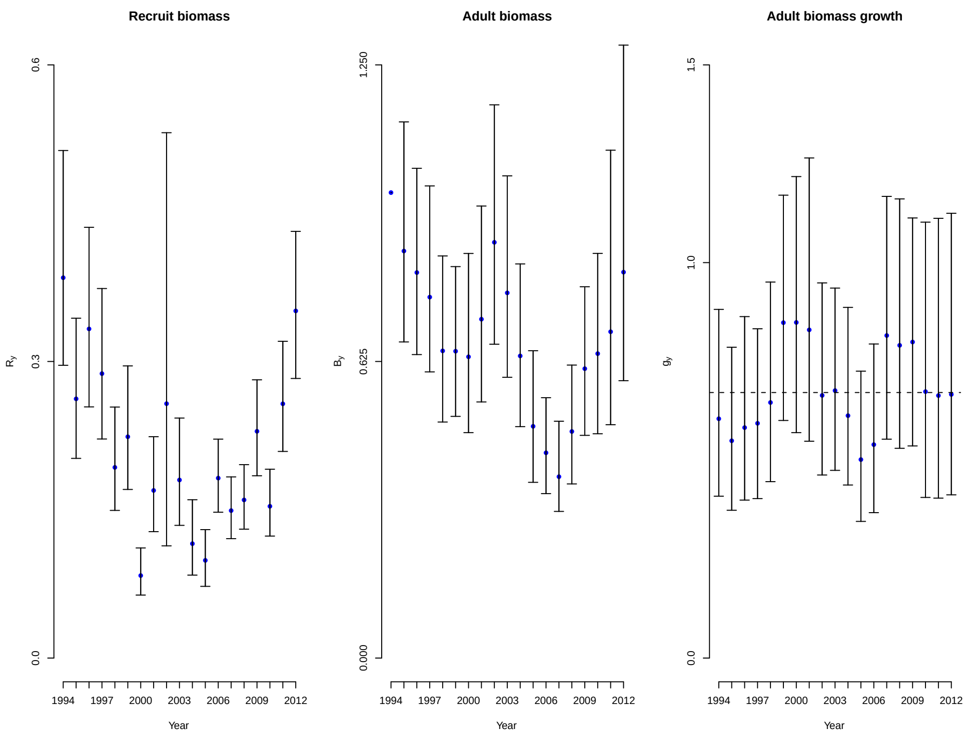

Figure 3: *Summary (median, circle; whiskers, 95% credible interval) of the relative recruitment biomass (left), relative adult biomass (middle) and net biomass growth (right, dotted line denotes the mean) using the aerial survey, SAPUE index and the CPUE data.*

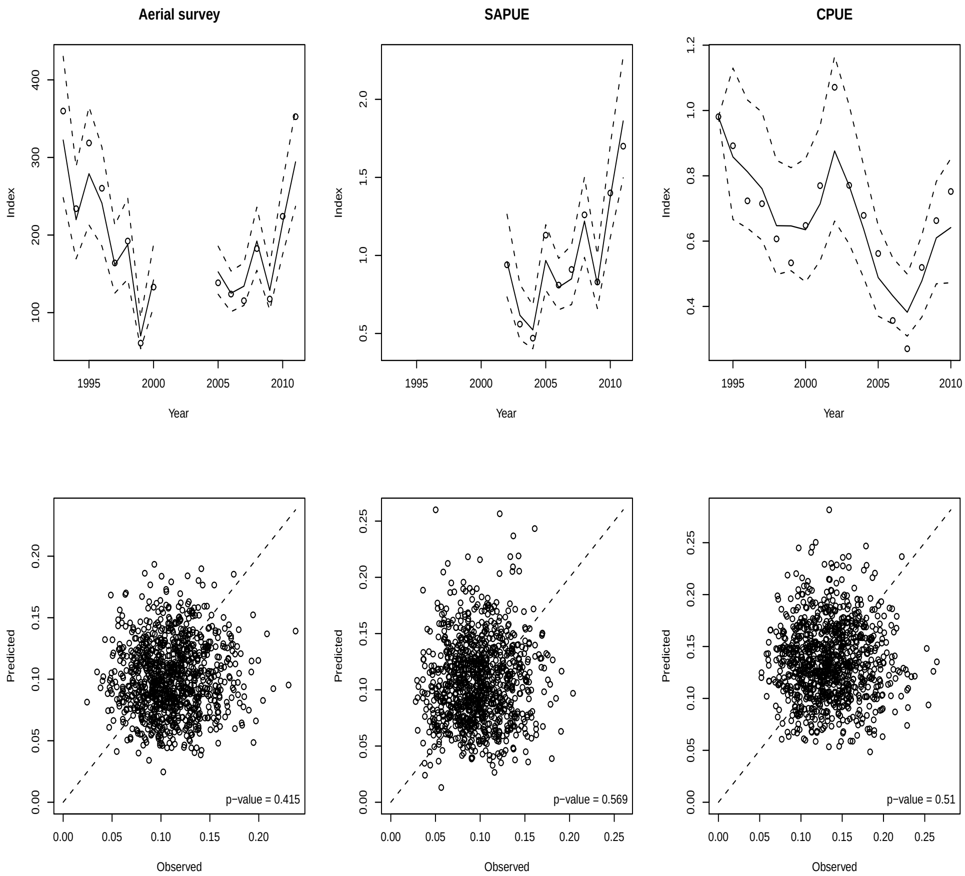

Figure 4: *Fitting and posterior predictive summaries the aerial survey data (left), SAPUE index (middle) and the commercial CPUE data (right). The points are the data with the full and dashed lines representing the median and 95% credible intervals, respectively.*

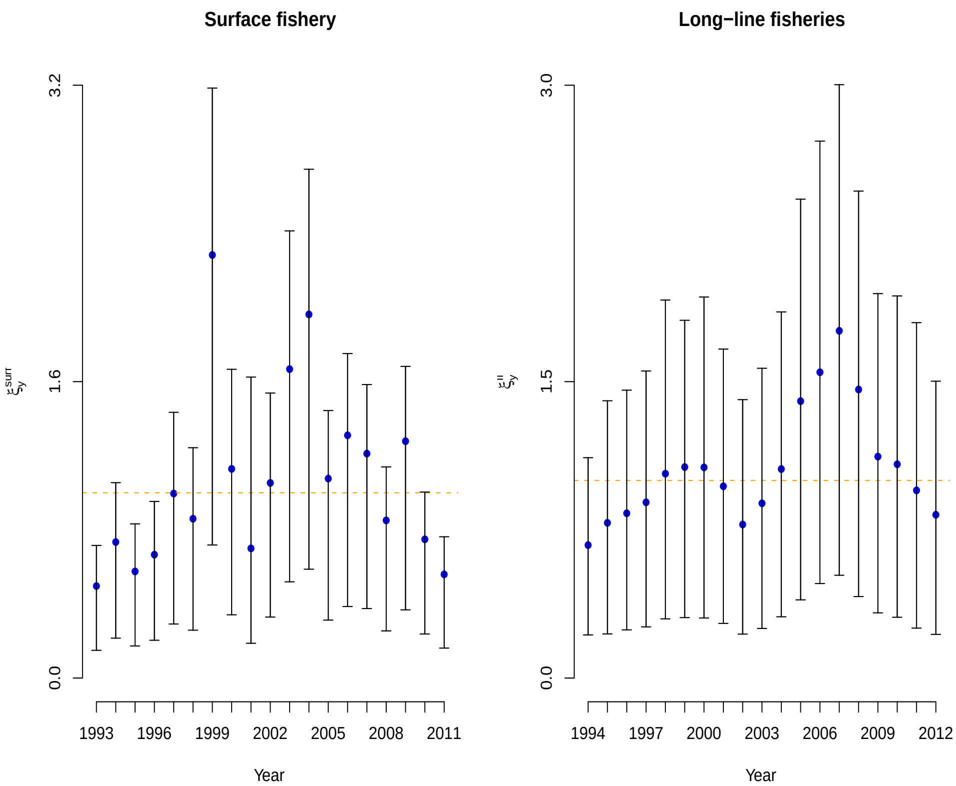

Figure 5: *Mean-standardised relative harvest rates for the surface (left) and long-line (right) fisheries. Shown are the median and 95% credible intervals with dotted line the mean (equal to 1).*

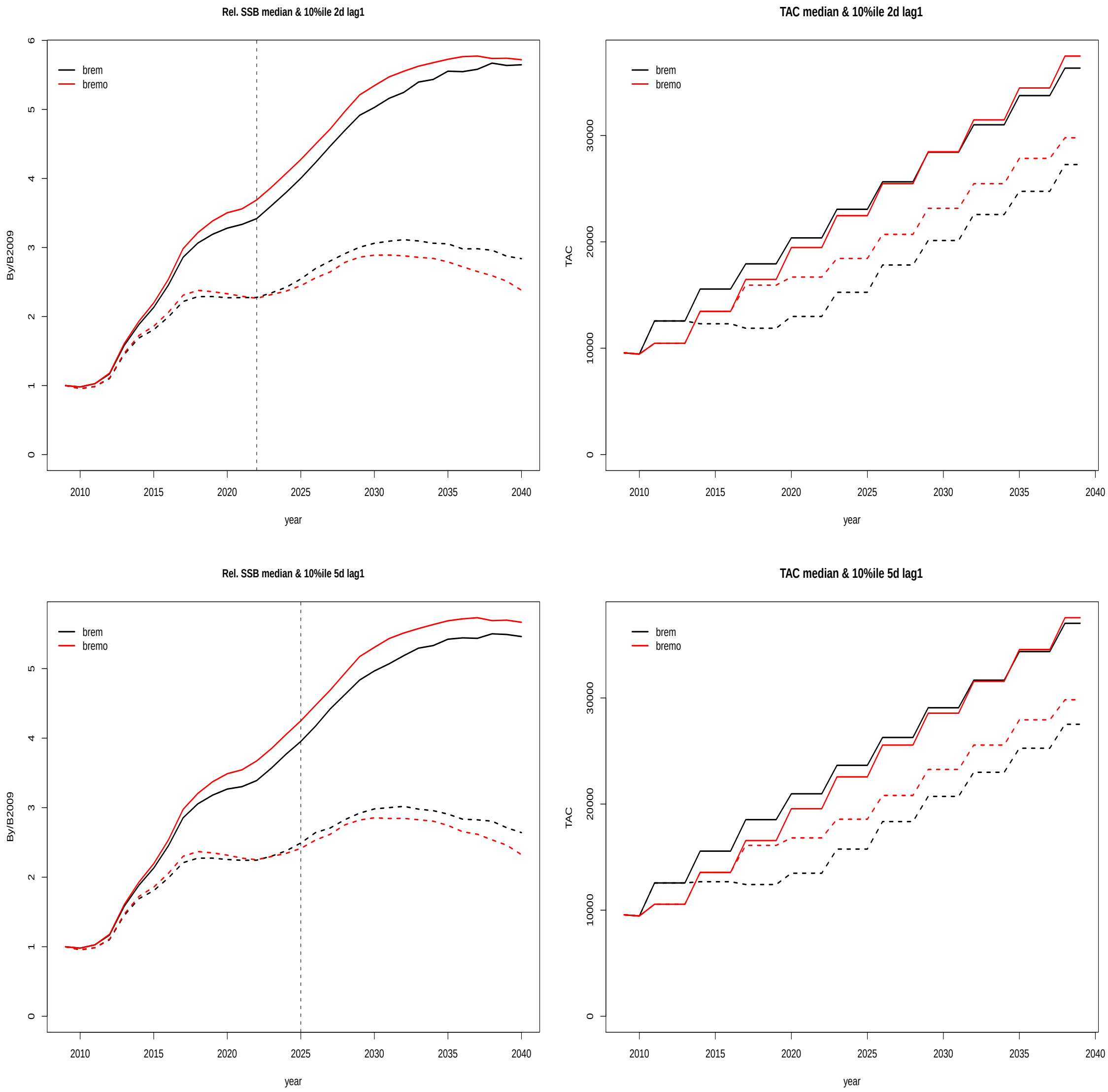

Figure 6: *Old (*bremo*) and new (*brem*) MP1 HCR formulations. The SSB (left) and catch (right) performance summaries for the tuning levels 2 (top) and 5 (bottom) assuming a 1-year lag.*