



# Data weighting, grid configurations and further exploration of the integration of the close-kin data into the SBT operating model

Rich Hillary, Ann Preece and Campbell Davies  
CCSBT-OMMP/1307/5

Prepared for the CCSBT OMMP4: Fourth Operating Model and Management Procedure Technical Workshop, 23-26 July 2013, Portland Maine, USA

Wealth from Oceans Flagship

CSIRO Marine and Atmospheric Research

### Copyright and disclaimer

© 2013 CSIRO To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

### Important disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

### Acknowledgements

This work was funded by CSIRO Wealth from Oceans Flagship and AFMA.

# Contents

1	Introduction .....	2
2	Data weighting .....	3
	2.1 Catch composition .....	3
	2.2 Tag data.....	4
	2.3 Abundance indices.....	5
	2.4 Close-kin data.....	6
3	Grid configuration and re-sampling scheme.....	8
	3.1 Steepness.....	8
	3.2 Natural mortality .....	11
4	Close-kin model development ideas.....	15
5	References.....	16

## Abstract

In this paper we explore several important issues relating to the CCSBT Operating Model in the context of preparations for the 2014 assessment. Data weighting options are explored for all the key data sets: catch composition, mark-recapture, abundance indices and the close-kin data. In the case of the catch composition, mark-recapture, and close-kin data an approach that looks to balance the variance in the standardised residuals is suggested; for the abundance data a predictive-based approach is outlined.

It is timely for the OMMP technical meeting to review the factors included in the grid and the grid resampling. A simple SBT-like simulation model was used to explore how well we expect to estimate steepness in the SBT situation and the potential impact of fixed versus random effects for recruitment deviations. The sampling of M10 and M0 levels in the grid are also discussed.

Potential extensions to the preliminary approach to including the close-kin data (Hillary et al 2012) were also explored. For process error this was dealt with in the data weighting section. Approaches that included length-based elements of the growth and reproductive processes (and related uncertainty) are outlined.

# 1 Introduction

In this paper we explore several important issues relating to the CCSBT Operating Model in the context of preparations for the 2014 assessment. Relating to data weighting, options are explored for all the key data sets: catch composition, mark-recapture, abundance indices and the close-kin data. In the case of the catch composition, mark-recapture, and close-kin data an approach that looks to balance the variance in the standardised residuals is suggested; for the abundance data a predictive-based approach is outlined.

With respect to grid resampling protocols, a simple SBT-like simulation model was used to explore how well we expect to estimate steepness in the SBT situation and the potential impact of fixed versus random effects for recruitment deviations. The sampling of M10 and M0 levels in the grid are also discussed.

Potential extensions to the preliminary approach to including the close-kin data (Hillary et al 2012) were also explored. For process error this was dealt with in the data weighting section. Approaches that included length-based elements of the growth and reproductive processes (and related uncertainty) are outlined.

## 2 Data weighting

All results presented in this section are done illustratively using the OM conditioning presented to the CCSBT ESC in 2012 (Hillary et al. 2012) denoted as *baseCKmk5sqrt*.

### 2.1 Catch composition

Catch composition, by age or length, is modelled using a multinomial distribution in the SBT OM. Apart from the observed length/age composition the key input parameter is the effective sample size,  $N_{eff}$ , which basically defines the effective precision of the catch composition and, ultimately, how much weight it receives in the likelihood. Perhaps the most common approach to calibrating the effective sample sizes in integrated models is as follows:

1. Define (somehow) an initial value of the effective sample size
2. Fit the model to the data given these initial values
3. Analyse the standardised residuals to look for over or under fitting
4. Adjust the initial effective sample sizes given step 3 and re-run the model

For a given effective sample size,  $N_{eff}$ , and observed,  $p$ , and predicted,  $\hat{p}$ , catch composition then the standardised residual is given by the following:

$$\varepsilon = \sqrt{N_{eff}} \frac{(p - \hat{p})}{\sqrt{\hat{p}(1 - \hat{p})}} \quad (1)$$

So, for a given year, if the effective sample size is correct then we would expect that the variance (across age or length),  $\varphi$ , of the standardised residuals would be approximately equal to 1. Values greater than 1 are indications of over-dispersion/additional process error, with values less than 1 suggesting that the effective sample size is not high enough. The updated value of the effective sample size would then be  $N_{eff}/\varphi$ .

So the dispersion factor,  $\varphi_{y,f}$ , is calculable for each fishery and year, for which data is included, and for each grid sample. There are two key issues that need addressing:

1. For which fisheries to use iterative updating as a mechanism for statistical weighting of the catch composition.
2. If we do decide to use iterative re-weighting, would this be on a year-specific basis or a more a general re-weighting factor averaged across all years and grid samples.

Just for information, Table 1 details the year and grid sample averaged dispersion statistics for each of the fisheries.

**Table 1: Dispersion statistics (averaged over years and grid samples) for each of the fisheries defined in the OM.**

Fishery	LL1	LL2	LL3	LL4	Surface	Indonesian
$\varphi$	0.23	0.26	1.11	0.08	0.33	0.3

## 2.2 Tag data

The tag data are modelled using a multinomial-Dirichlet mixed model, which includes an explicit dispersion factor as part of the likelihood which is defined by the user in the *sqrt* file. This dispersion factor was first calculated at the 2009 Seattle OMMP meeting and has remained fixed at this same value of  $\varphi = 2.35$ . This value was averaged over the whole of the mark-recapture data set, but the tag data are actually quite highly disaggregated in the OM and are treated by tagger, cohort and age of release.

The OM output includes the standardised residuals for all of the mark-recapture data so one can look in more detail at how the levels of dispersion vary across the key disaggregation factors: tagger (1-6), cohort (1989-1994) and age of release (1-3). At the lowest aggregation level, i.e. across all factors and grid samples, we have a value of  $\varphi = 0.99$  suggesting that the tagging dispersion factor is still pretty much correct. Table 2 provides the dispersion factor for each tagger averaged across the other factors and across all grid samples.

**Table 2: Dispersion statistics (averaged over release cohorts, ages and grid samples) for each tagger.**

Tagger	1	2	3	4	5	6
$\varphi$	1.17	0.99	1.73	1.11	0.64	0.36

From Table 2 taggers 1, 2 and 4 all had dispersion factors close to 1, with tagger 3 having a value closer to 2 and tagger 5 and 6 having values below 1 (with tagger 6 being below 0.5 but above 0.25). Table 3 details the dispersion factors for each of the 6 cohorts of release (1989-1994), averaged across tagger, release age and grid sample.

**Table 3: Dispersion statistics (averaged over release ages, tagger and grid samples) for each release cohort.**

Cohort	1989	1990	1991	1992	1993	1994
$\varphi$	0.73	0.6	0.51	1.61	1.26	1.62

In Table 3, while there are no obviously large deviations from 1 there is a clear pattern that the first 3 cohorts seem to be fitted “better” than the last 3. Reasons for this are unclear, it could be issues linked to conflicts with other data in the last 3 cohorts or subtle patterns linked to the tag grouping and age-at-release over which we have averaged in the first place.

Table 4 details the dispersion factors for the three release ages, averaged over tagger, release cohort and grid sample. The three values are all close to 1 and do not suggest any clear or obvious dispersion issues linked to the age of release.

**Table 4: Dispersion statistics (averaged over release cohort, tagger and grid samples) for each release age.**

Release age	1	2	3
$\varphi$	1.24	0.81	0.83

As a final analysis, it was decided to see if there are any effects when considering simple trends in each of the key factors together. The following linear model was used on the standardised residuals (averaged across all grid samples):

$$\text{res} \sim \text{recapture age} + f(\text{release age}) + f(\text{cohort}) + f(\text{tagger})$$

where,  $f()$  denotes factor and so recapture age is treated as a continuous variable. Of all the effects only the tagger 6 and the 1990 release cohort were significant at the 0.05 level (but not lower than 0.01). The recapture age and tagger 5 effects were significant at the 0.1 level, but all the others were above the 0.1

level. All together the explanatory power of the (admittedly simple) model is very low (adjusted  $R$ -squared of 0.04) so one might conclude that there are no obvious and simple trends in the tagging residuals across all the relevant disaggregation factors.

## 2.3 Abundance indices

In Hillary et al. (2011) a predictive approach was used to analyse the performance of the model in terms of fitting to the relative abundance indices. Assuming that the grid samples are drawn from some kind of underlying representative distribution (like the posterior in the Bayesian paradigm) the posterior predictive approach (Meng, 1994) was employed.

The process works by first simulating data from the assumed likelihood, for each deterministic prediction of the relevant abundance index in the grid sample. A suitable discrepancy statistic ("closeness" of the simulated data to the deterministic prediction) is calculated for each grid sample and accompanying simulated data set. In this work the median absolute deviation (non-parametric) is used. Then, the same discrepancy statistic is calculated for each grid sample and now using the observed data.

The probability with which the predicted discrepancy statistic is greater than the observed one is referred to as the Bayesian  $p$ -value, though it is probably better to refer to it as a predictive  $p$ -value as we're not technically dealing with Bayesian posteriors here. Generally,  $p$ -values of around 0.5 are preferred with values outside the extremes of 0.05 and 0.95 generally indicating some kind of issue with the probability model. One also usually plots the bivariate cloud of discrepancy statistics as they can also be informative.

**Figure 1: fits to abundance indices (top; CPUE left, aerial survey right) and predictive performance (bottom; CPUE left, aerial survey right).**

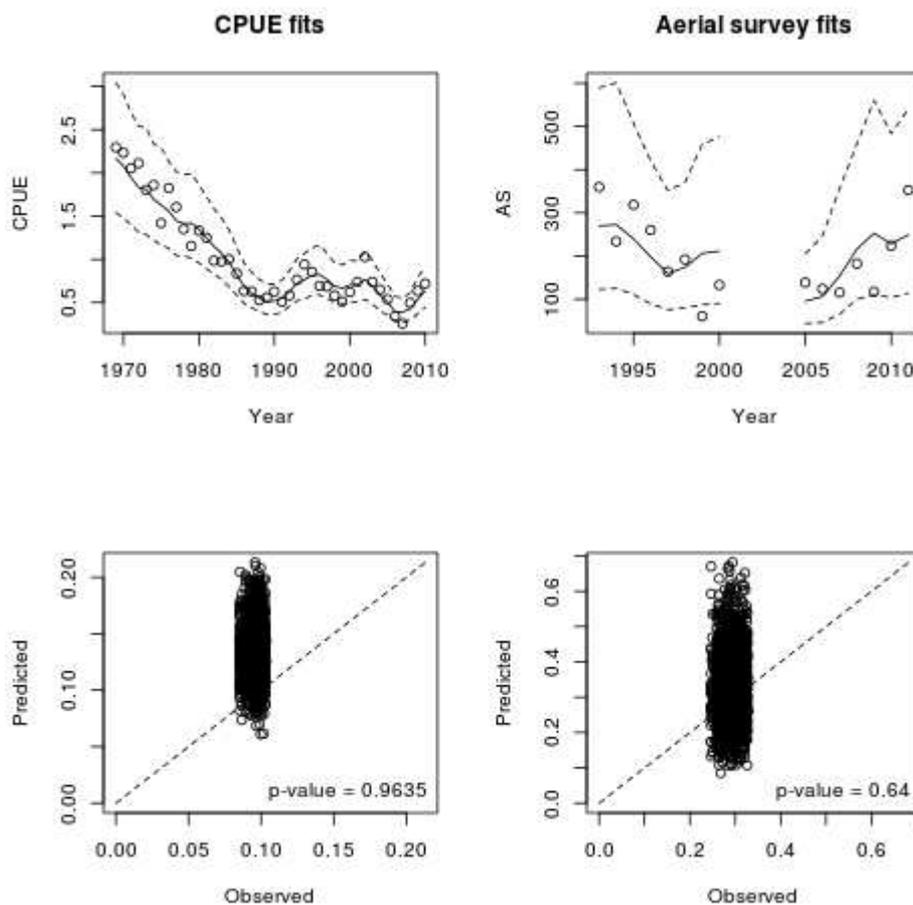


Figure 1 shows the fits to the abundance indices (in terms of observed data and the predictive 90% CI) as well as the predictive  $p$ -values and discrepancy statistics. For the CPUE data, they are generally well fitted

with all but one of the points sitting inside the predictive interval. The aerial survey data are fitted less well, though the general trend is captured and – as with the CPUE – all but one of the points are contained inside the predictive interval. In terms of predictive  $p$ -values the CPUE has a value of 0.96, with the aerial survey a value of 0.63. What this suggests is, from the perspective of the probability model, we are predicting more variability in the CPUE than that observed, but for the aerial survey only slightly so and no obvious problems are apparent. The CV for the CPUE residuals is around 0.14 as opposed to the 0.2 assumed in the OM, which is why we see the high  $p$ -value that we do.

## 2.4 Close-kin data

Whatever the form chosen for the definition of the parental probability calculation in the close-kin work, the base assumption for each parent-juvenile (PJ) comparison is that it comes from a Bernoulli distribution, or a binomial distribution if we aggregate the PJ comparisons into groups (Hillary et al. 2012). While there was very little evidence of correlation between comparisons, via the detection of half-siblings (Bravington et al. 2013), it would be consistent with the treatment of the other data sets to explore the possibility of additional process error via an over-dispersion approach as per the tagging data.

If  $\pi$  denotes the parental probability (Hillary et al., 2012),  $N$  the number of comparisons and  $M$  the number of POP detections then the base likelihood as follows:

$$p(M|\pi) \propto \pi^M (1 - \pi)^{N-M} \quad (2)$$

The over-dispersion part is introduced by then defining  $\pi$  to follow a beta distribution:  $\pi \sim B(\alpha, \beta)$ . If we have the model predicted parental probability,  $\hat{\pi}$ , then

$$\hat{\pi} = \frac{\alpha}{\alpha + \beta} \quad (3)$$

and we define the dispersion parameter,  $\omega$ , as follows:

$$\omega = (\alpha + \beta)^{-1} \quad (4)$$

so that the mean number of POPs will be  $N\hat{\pi}$  and the variance will be

$$N\hat{\pi}(1 - \hat{\pi})(N\omega + 1)(1 + \omega)^{-1} \quad (5)$$

We can now define our over-dispersion factor,  $\varphi$ , as follows:

$$\varphi = (N\omega + 1)(1 + \omega)^{-1} \quad (6)$$

which basically means the expected number of POPs is the same as for the straight binomial model, but the variance is inflated by a given factor (which is what we want). Straight from (6) by defining the over-dispersion factor  $\varphi$  beforehand (as with the tagging data) immediately tells us that

$$\omega = \frac{\varphi - 1}{N - \varphi} \quad (7)$$

making it simple to then use (3) and (4) to solve for  $\alpha$  and  $\beta$  in terms of  $\varphi$  and  $\hat{\pi}$  where

$$\alpha = \frac{(N - \varphi)\hat{\pi}}{(1 - \hat{\pi})(\hat{\pi} + (1 - \hat{\pi})(\varphi - 1))} \quad (8)$$

and

$$\beta = \frac{N - \varphi}{\hat{\pi} + (1 - \hat{\pi})(\varphi - 1)} \quad (9)$$

The final step is to then obtain the marginal likelihood of the POPs by integrating over the (now random variable)  $\pi$  in (2). This thankfully yields a close-form solution for the marginal likelihood of the POPs given the beta-binomial probability model:

$$p(M) \propto \frac{\Gamma(M + \alpha)\Gamma(N - M + \beta)\Gamma(\alpha + \beta)}{\Gamma(N + \alpha + \beta)\Gamma(\alpha)\Gamma(\beta)} \quad (10)$$

where  $\alpha$  and  $\beta$  are as defined in (8) and (9), respectively, and  $\Gamma$  is the gamma function. So, while this represents a well established method of permitting additional variance in the close-kin data via a user-defined over-dispersion factor, how do we go about choosing the right factor?

In certain instances one can estimate the over-dispersion factor directly as you would a process error variance parameter. However, the model itself can scale poorly at values of  $\varphi$  close to 1 – at a value of 1 the actual likelihood in (10) will not be defined. Conceptually, the beta-binomial collapses into the straight binomial at this point but the beta distribution attains an infinite density around  $\hat{\pi}$  like a Dirac delta function. It would seem simpler all round to do what we do with the tagging data and have a user defined value of  $\varphi$  and some simpler manner in which to calculate it.

For both the catch composition and the tagging data we have been using an approach by which we alter  $\varphi$  to balance the variance in the standardised residuals and for the beta-binomial close-kin probability model these are easy to calculate:

$$\varepsilon = \frac{M - N\hat{\pi}}{\sqrt{\varphi N\hat{\pi}(1-\hat{\pi})}} \quad (11)$$

In Hillary et al. (2012) the close-kin data were grouped by the relevant cohort (juvenile birth year), adult capture year and age. It might be sensible to aggregate the standardised residuals a little before doing any such variance calculations just as there will be quite a few zero comparison cases and very low probability comparisons. In Hillary et al. (2012) the normal approximation to the binomial was used to generate manageable fitting summaries for the close-kin data, but there are computational methods to exactly compute the distribution of a sum of binomial variables. After choosing the aggregation level this could be used to generate the set of standardised residuals in (11) used to then estimate appropriate values of  $\varphi$ .

## 3 Grid configuration and re-sampling scheme

The impacts of the inclusion of the close-kin data in the OM were presented in Hillary et al, 2012. These included a shift in the preference for steepness levels, and a discussion of the weighting scheme used to sample models and estimates from the grid.

The suggested changes in the grid in 2012, after inclusion of the close-kin data, were in the levels of the steepness factor because of convergence problems for the combination of the highest M0 (0.45) level and the lowest steepness level (0.55). In the baseCKmk5 set, the lowest steepness level (0.55) was excluded, and the high M0 value was retained because of support in the tagging data for these high M0 values. These and further adjustments to levels of various factors in the grid may need to be considered when the OMs are run using the updated 2013 data input file, during the OMMP4 technical meeting.

In addition to reviewing the levels of factors in the grid, it is timely for the technical meeting and ESC to review the method and weightings used to select a simulation set of models by sampling the operating model results (the weighted sampling of the grid combinations). As noted in Hillary et al., 2012, the sampling scheme uses a combination of prior and posteriors to weight various factors. When posterior weightings are used (for steepness, M10 and M0), these weightings are currently using the value of the objective function which is a combination of both the negative log-likelihood and the penalty factors in the objective function. The description of these posteriors has in the past been described as the likelihood weighting, but should more correctly be described as the objective function weighting. Hillary et al 2012, pointed out that the penalties used in the OMs are having an influence primarily on the sampling of the steepness levels, and to a lesser extent on M10. M0 is also weighted using the objective function, but the penalty component of the objective function is having very little contributing effect on the sampling of the M0 levels. These three factors are explored below.

### 3.1 Steepness

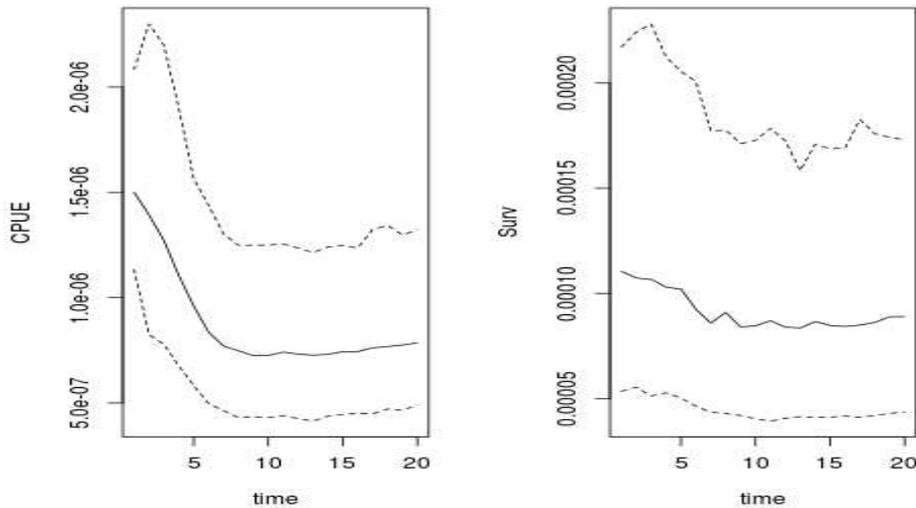
Clearly information on steepness is equivocal – see Hillary et al. (2011) versus Hillary et al. (2012) – and changed notably with the introduction of the close-kin data. The main informant on steepness now is the recruitment deviation penalty, consistently preferring lower levels of steepness in recent years. This information, while not necessarily spurious, is secondary, as it is coming via a subjective prior not directly from the data and should be reviewed for this reason. Two key questions in this regard are:

1. Is this information likely to be robust?
2. Is it influenced by the penalised likelihood versus proper random effect treatment of the recruitment deviations in the SBT OM?

Neither of these questions can be addressed directly using the OM given its complexity, so a simple SBT-like test model was used to explore the questions. The test model had two age-classes (mimicking juvenile and adult parts of population), with a recruitment survey and an adult CPUE index. The precision of the CPUE index was given a CV of 0.15; the recruitment survey a CV of 0.2; and the recruitment deviations a CV of 0.5. True steepness was defined to be  $h = 0.75$  with a simple selectivity ogive of 0.5 for the juveniles and 1 for the adults and an indicative natural mortality of 0.2. The model was run for 1,000 iterations and for 21 years with an effort profile that mimics the strong adult abundance decline in SBT followed by an arrest in

the decline and very minimal increase at the end of the simulation period – see Fig. 2 for the CPUE and recruitment survey summaries and Fig.3 for the spawning abundance summary.

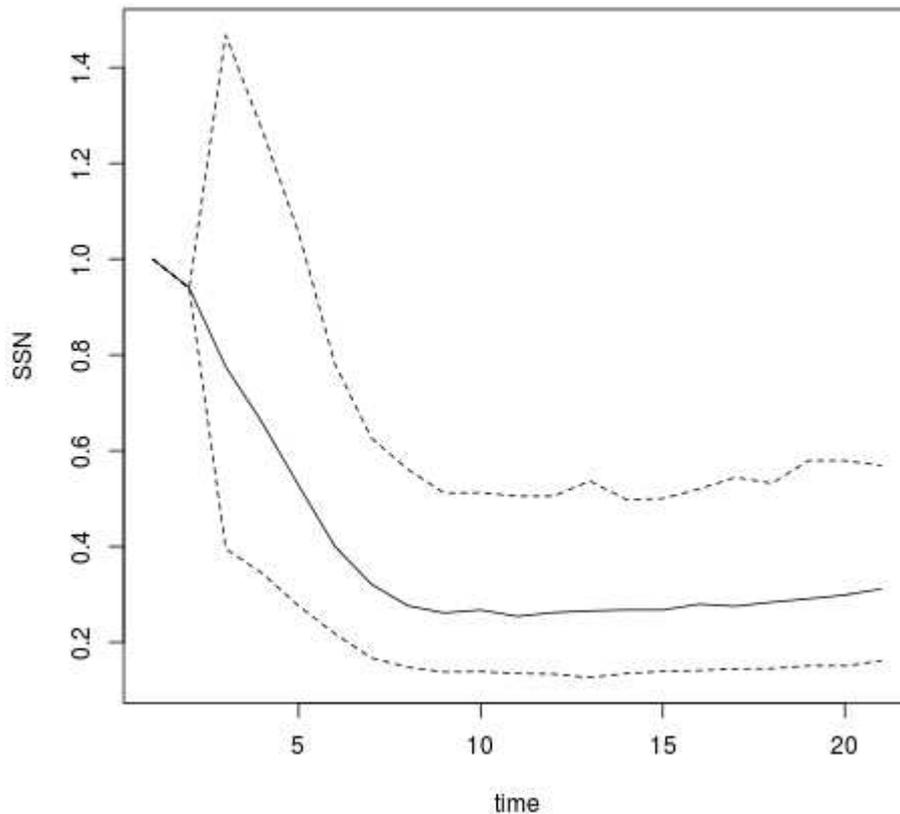
**Figure 2: Abundance indices (CPUE, left; recruitment survey, right) for simple SBT-like two-age model**



Two model configurations were explored: (i) with recruitment deviations treated as fixed effects via penalised likelihood, and (ii) where recruitment deviations are treated as random effects and estimated accordingly. One difficult factor to account for with an SBT-like model is we now have effectively two absolute abundance data sets in the tagging and the close-kin in the OM. This would have been very difficult to simulate in a transparent and efficient manner in the simple test model so the following approach was taken:

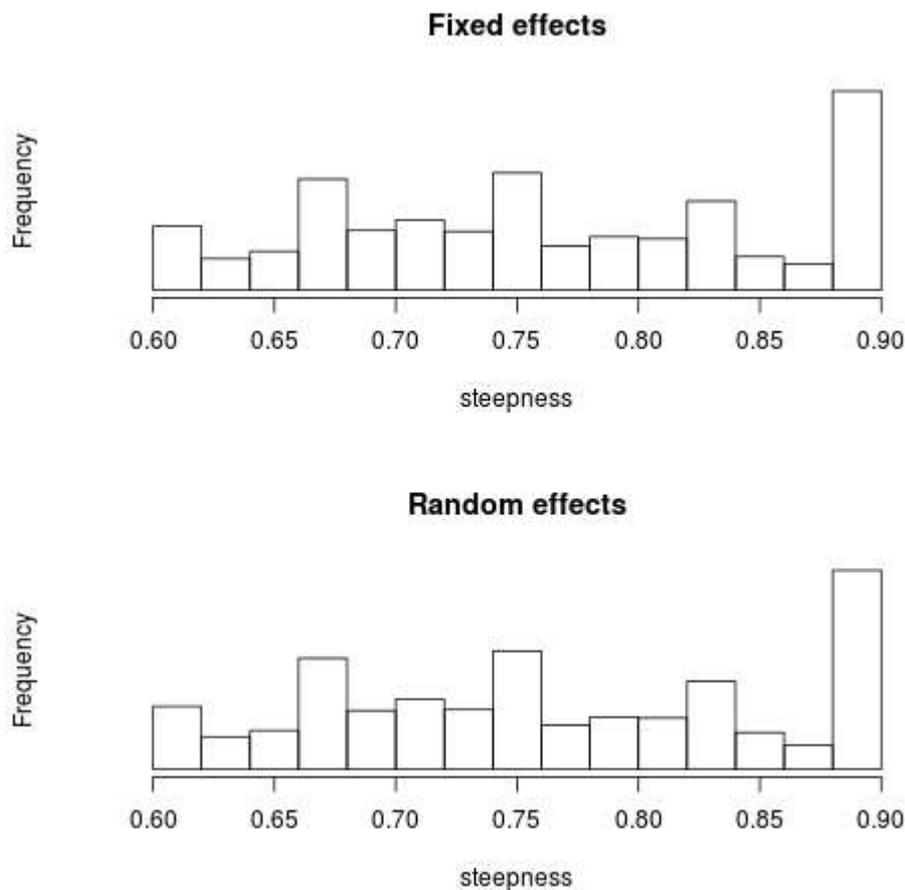
The key absolute abundance parameter  $B_0$  was fixed at 1 so the model estimates only catchability of the  $F$ -effort relationship (catches are estimated in the model structure), steepness, and the recruitment deviations. The arbitrary catchabilities for the CPUE and survey indices are calculated as nuisance variables. This very rough implementation essentially assumes that absolute abundance is “taken care of” and we are asking the model to estimate the key density-dependence parameter (steepness) along with recruitment deviations and “relative” catches. This is, perhaps, a less challenging situation than we have with SBT OM, as we are effectively removing the extra impediment of correlation between  $h$  and  $B_0$  from the estimation process; but is considered informative none the less.

**Figure 3: Relative abundance of the adult spawning population for the simple SBT-like two-age model.**



For each of the 1,000 simulated population and observation realisations the steepness estimates (for both fixed and random effect model configurations) were recorded and plotted in Fig. 4. Clearly, even with the strong decline and slight upturn in adult abundance, the true steepness is not recovered in the back-estimates well at all. In these results the most probable estimate is actually on the upper boundary of 0.9 (but it often hits the lower boundary too). There is also a smaller, but less well defined, density centred round the true value of 0.75, with the within-boundary estimates muddled and multi-modal. What is perhaps surprising is that the results are the same for both fixed and random effect recruitment models, and for both configurations the key influence of the recruitment penalty is most often to push the estimates of steepness to the boundaries (which happens 20% of the time here).

The conclusion for this simplified test model is that, with such an adult abundance history, the data do not give us any clear information on steepness. This has been the view of the SC for a long time. The second more interesting part is that, be it for fixed or random effect recruitment models, we can expect that the recruitment prior will very often result in strong but spurious information on steepness and this will dominate the data in terms of information, resulting in the estimates hitting the upper and lower boundaries a significant proportion of the time. This model is a *very* rough imitation of some features of the SBT system, but it is interesting that we also see the strong role the recruitment prior plays in terms of information on steepness.

**Figure 4: Back-estimates of steepness for the fixed (top) and random (bottom) effect recruitment models.**

The current penalty profile for steepness looks quite like the situation we see in the SBT OM now the close-kin data are included, with a preference for steepness at the lower end of the range and no information coming from the data. We cannot definitively say what situation we are in based on this simple model exploration but it does suggest we need to re-evaluate whether we want steepness to be resampled via the objective function, or whether a return to some kind of disperse or weakly informative prior is preferable.

### 3.2 Natural mortality

Natural mortality appears to be mainly informed by tagging data in the OM, and to a lesser extent by the close-kin and age/length frequency data. Natural mortality is also highly correlated with steepness levels, and therefore sampling of the natural mortality factors is likely to be affected by decisions on how to sample the steepness factor (see Eveson and Davies, 2009).

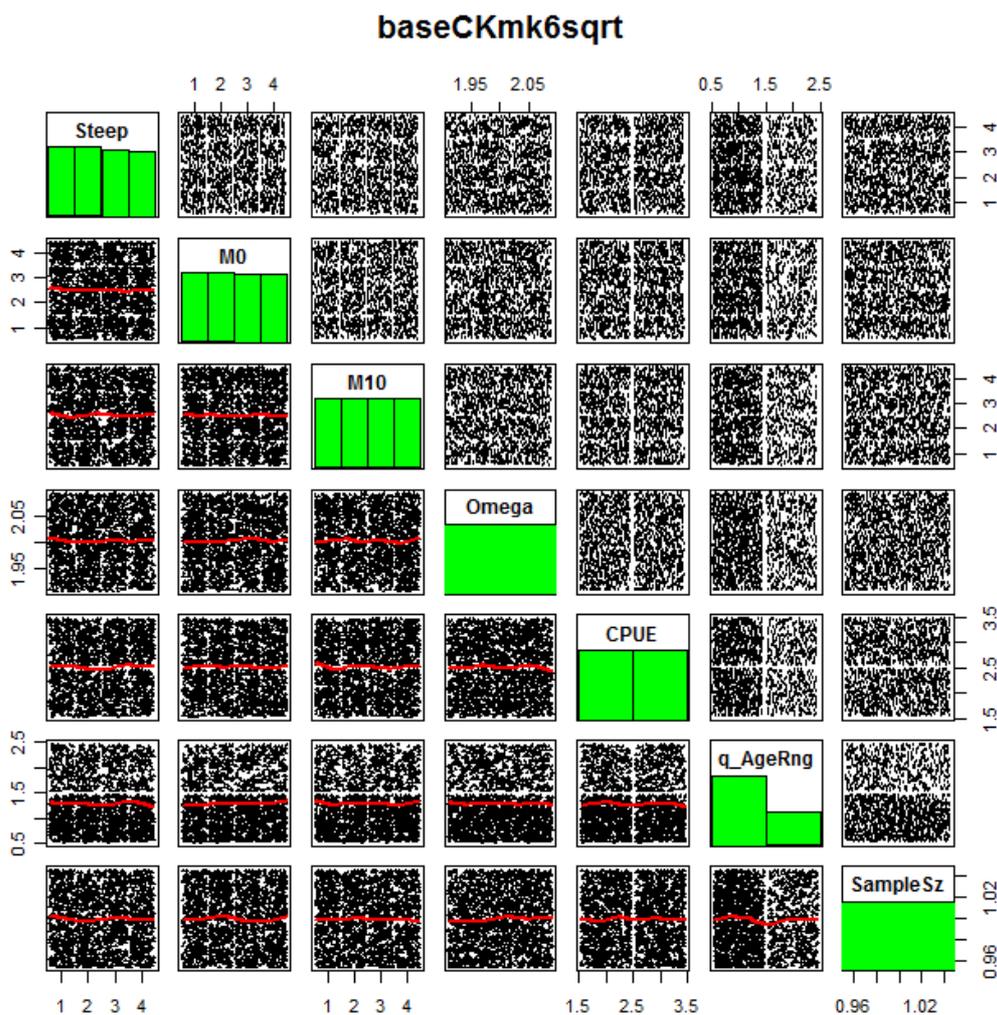
The sampling of the M10 levels in particular was shown to be influenced by the sampling of steepness levels in Hillary et al, 2012. To examine the preferences for the M10 and M0 levels by the likelihood and penalty components of the objective function in the absence of the effects of sampling of the other factors, the grid was run for case baseCKmk6 with uniform prior for sampling steepness, M0 and M10. We are not proposing that uniform priors be used for all the factors in sampling the grid – only for this exploration process. The grid for baseCKmk6 is the same as baseCKmk5 in Hillary et al, 2012, where the lowest steepness value is excluded. The level plot and M0 and M10 likelihood and penalty profiles are shown in figures 5 – 9.

There is a small effect from the penalties on the weighting of the M10 levels, which appears to be from the selectivity change and recruitment penalty, and mixed weak information from the data

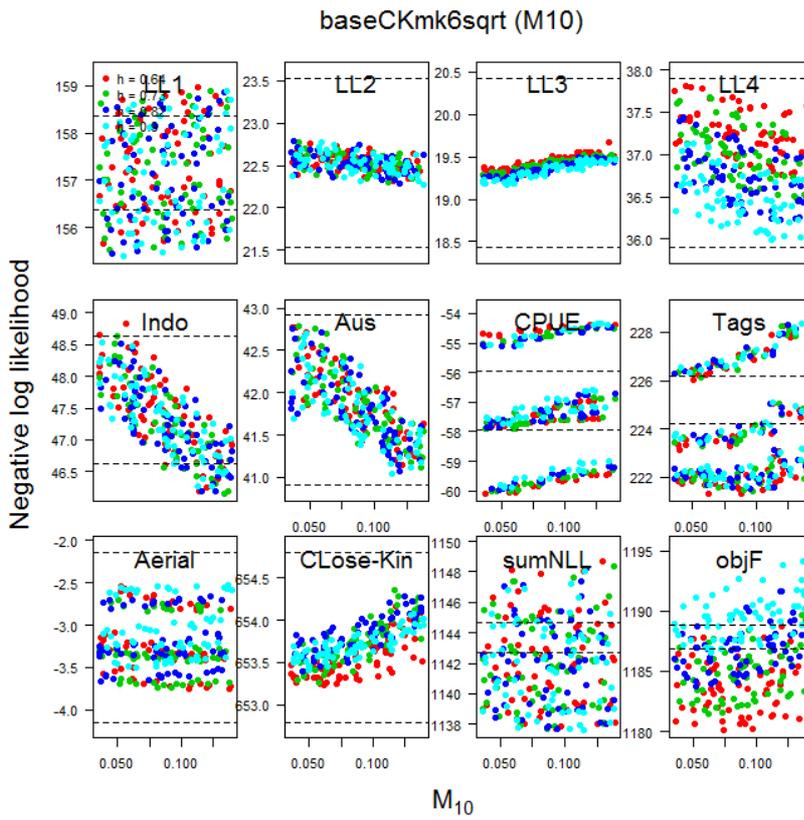
sources (figure 6 and 7). It is suggested that for this case that the technical meeting and ESC should review whether these penalties should be influencing sampling preferences for M10, or whether a uniform prior or the negative log likelihood component only of the objective function should be used for weighting the sampling of the M10 levels.

In the case of M0, it is primarily the likelihood weighting that is influencing the selection of the higher M0 levels, and the penalty effects are not having a large impact on this (figures 8 and 9). The selectivity penalties show a weak preference for lower M0 levels. As for the M10 sampling above, the OMMP meeting could consider which components of the objective function should be included in the sampling of M0 levels.

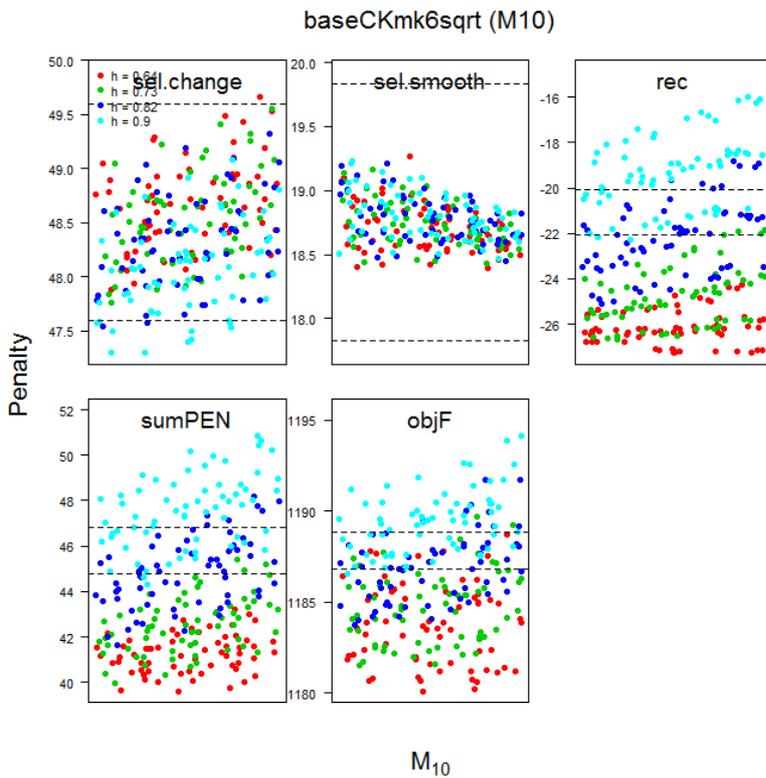
**Figure 5. The level plot for the baseCKmk6 grid where steepness, M0 and M10 are sampled based on uniform prior weights for the purpose of examination of the preferences in likelihood and penalty components of the objective function without the correlations from sampling of the other factors (e.g. without the effects from M and steepness correlation. These figures use the 2011 input data file and haven't been updated with the most recent data, and the grid is the same as that defined for baseCKmk5 in Hillary et al, 2012, where the lowest steepness (0.55) level has been excluded.**



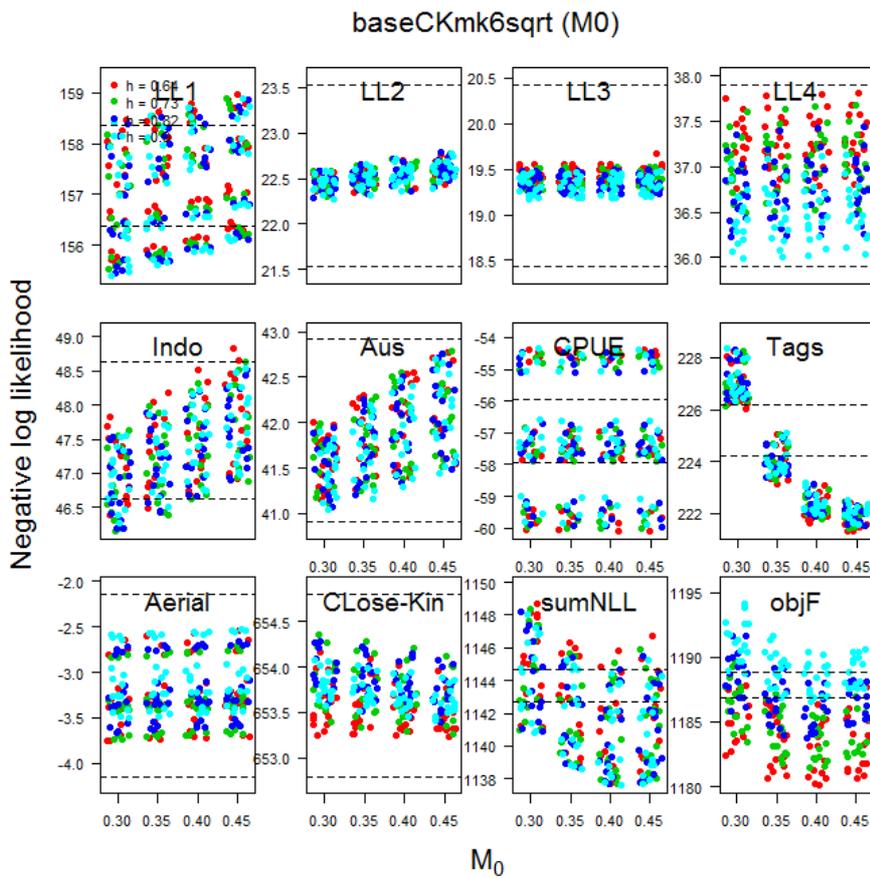
**Figure 6. Likelihood profile for M10 (with colour split for levels of steepness), for grid baseCKmk6 sqrt**



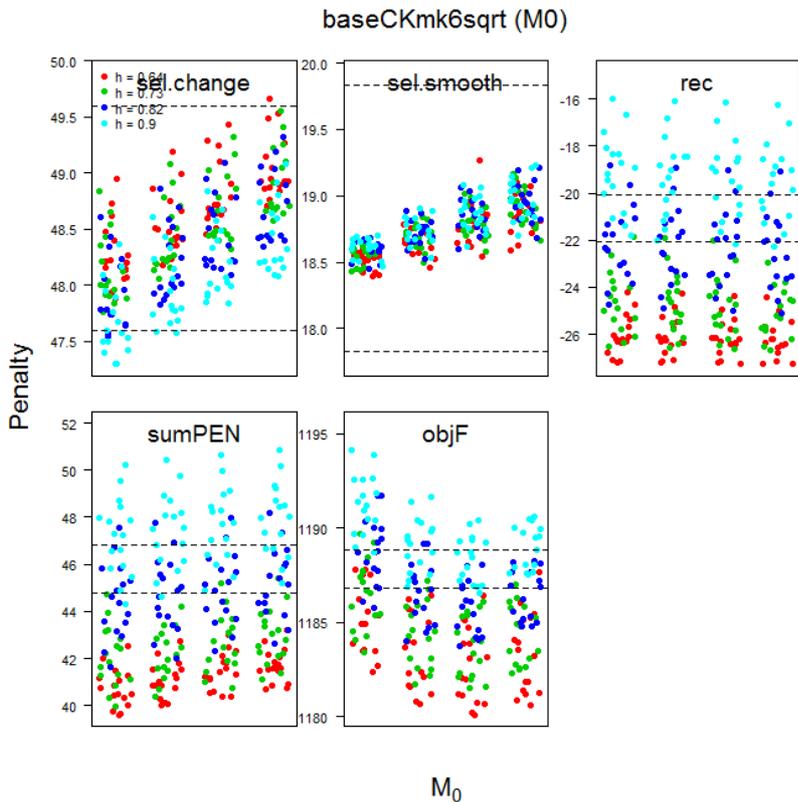
**Figure 7. Penalty profile for M10 (with colour split for levels of steepness), for grid baseCKmk6 sqrt**



**Figure 8. Likelihood profile for M0 (with colour split for levels of steepness), for grid baseCKmk6 sqrt**



**Figure 9. Penalty profile for M0 (with colour split for levels of steepness), for grid baseCKmk6 sqrt**



## 4 Close-kin model development ideas

While no further work has been done beyond the approach outlined in Hillary et al. (2012), and the over-dispersion approach detailed in this paper, there are some things that would be worth exploring. Perhaps the most important factor is a more detailed treatment of the role of length in the process. In Hillary et al. (2012) the data were pre-processed to be age-based – including the effective reproductive contribution ogive. The variation in length-at-age was dealt with in terms of estimating expected ages and so on, but doing this externally will perhaps give more certainty to the data than if this was treated internally.

The way the data were collected mean that juvenile age (and therefore cohort) are most likely known with a much higher degree of accuracy than adult age-at-capture (which we only have otoliths for when the adult is a known parent). Adult length is the key observable in the overwhelming majority of comparisons and so we can infer from this the adult length-at-comparison (i.e. in the juvenile birth year not adult capture year) from the estimated growth distribution.

The way the current parental probability is calculated is age-based:

$$\pi_{ij} = \frac{2\varphi_{a_j - a_i}}{\sum N_{y_i - a_i, a} \varphi_a} \quad (12)$$

where  $a_j$  is the adult capture age,  $a_i$  is the juvenile capture age,  $y_i$  is the adult capture year and  $\varphi_a$  is the relative reproductive ogive-at-age. The currency of the OM in terms of population dynamics is age, but length is a key facet of the mechanics of the close-kin components. In the numerator of (12) there will be an effective distribution of possible lengths the adult could be in the juvenile birth year, given it was length  $l_j$  at capture:  $p_{ij}^{l_j}$ . Similarly for the denominator, each age will have a distribution of possible lengths:  $p_a^l$ . Assuming a finite set of discrete length bins we can reformulate the parental probability in (12) into a more length-based quantity:

$$\pi_{ij} = \frac{2 \sum_l \varphi_l p_{ij}^{l_j}}{\sum_a \sum_l N_{y_i - a_i, a} p_a^l \varphi_l} \quad (13)$$

This minor adjustment to the parental probability just avoids having to compute expectations of age conditional on length outside of the model and thereby include this uncertainty within the model structure. Note this would not change the likelihood structure in the model but would alter the grouping of adult capture age in the data, as now it would be defined in terms of adult capture length. The other key groups, adult capture year and juvenile capture age, would remain the same.

## 5 References

- Bravington, M., Grewe, P., and Davies, C.R. 2013. Fishery independent estimate of spawning biomass of southern Bluefin tuna through identification of close-kin using genetic markers. Information paper to CCSBT OMMP Working Group Meeting, Portland, Maine, USA, 23-26 July 2013.
- Eveson, P. and Davies, C.R. 2009. Re-conditioning of the CCSBT Operating Model: impacts of revised natural mortality schedule and interaction with steepness. *CCSBT-ESC/0909/40*.
- Hillary, R. M., Preece, A., Davies, C.R., Barnes, B., and Begg, G. 2011. Reconditioning of the southern bluefin tuna operating model: exploratory data analysis, fitting performance, and current stock status. *CCSBT-ESC/1107/11*.
- Hillary, R.M., Preece, A., Davies, C.R., Bravington, M.V., Eveson, J.P., and Basson, M. 2012. Initial exploration of options for inclusion of the close-kin data into the SBT operating model. *CCSBT-ESC/1208/21*.



CONTACT US

**t** 1300 363 400  
+61 3 9545 2176  
**e** [enquiries@csiro.au](mailto:enquiries@csiro.au)  
**w** [www.csiro.au](http://www.csiro.au)

YOUR CSIRO

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills for building prosperity, growth, health and sustainability. It serves governments, industries, business and communities across the nation.

FOR FURTHER INFORMATION

**Marine and Atmospheric Research**  
Rich Hillary  
**t** +61 3 6232 5222  
**e** [rich.hillary@csiro.au](mailto:rich.hillary@csiro.au)  
**w** [www.csiro.au/wfo](http://www.csiro.au/wfo)