WEALTH FROM OCEANS FLAGSHIP
www.csiro.au

Preliminary reconditioning of the SBT OM with updated data in 2014

Ann Preece, Rich Hillary, Campbell Davies

CCSBT-OMMP/1406/04

CCSBT Fifth Operating Model and Management Procedure Technical Meeting, Seattle, USA

June 2014

Acknowledgments

This work has been funded by the AFMA and CSIRO Wealth from Oceans Flagship.

Contents

Abstract

This paper provides information on preliminary re-conditioning of the SBT operating models (OMs), for consideration at OMMP5 and further refinement prior to a full stock assessment at the extended scientific committee meeting (Sept, 2014). This will be the first stock assessment since 2011, and since the SBT management procedure was implemented and the close-kin project completed. The SBT OMs have been updated with new data provided as part of the 2014 CCSBT data exchange. At technical and scientific committee meetings since the last stock assessment in 2011, the operating models specification, range of values in the reference grid, priors and model sample weights, and new data sources have been refined and their implementation into the SBT OMs agreed. A major change has been the integration of the close-kin data into the operating models. Examination of the diagnostics indicates the updated models fit the data well. The recruitment time-series show high recruitments in the recent years. Consistent with results from OMMP4, the preliminary reconditioning of the SBT OM indicates that stock status has improved since the last assessment in 2011. The projections code has also been updated, and results of preliminary catch and biomass projections using the reference set and the CCSBT MP indicate that the rebuilding objectives for the MP will be met.

1 Introduction

In 2011, the SBT operating models were reconditioned to provide stock status advice and to use in management strategy evaluation of the SBT management procedure that was adopted in the same year. Since 2011, there have been a series of technical OMMP meetings and ESC meetings at which the operating models specification, range of values in the reference grid, priors and model sample weights, and new data sources have been discussed and decisions made on their implementation into the SBT OMs. Given the decisions made since 2011, the OMs have been updated with new data, and this paper describes the preliminary re-conditioning of the OMs. The results presented here are preliminary, and the re-conditioning will be examined at the OMMP meeting and refined and updated to provide scenario based stock status advice at the ESC.

Summaries of OM fits to the new data are presented, along with the posterior predictive analysis of CPUE and Aerial survey data, to see how well they explain the data. The key parameter estimates and grid sampling for the updated OM are summarised. Estimates of trends in historical recruitment and SSB are provided, and projections using the SBT management procedure using updated data are presented.

The most recent aerial survey point is high relative to the rest of the series, and the impact of this high point on the OMs is examined in more detail. Since the close-kin data are to be formally included in the OM for stock assessment purposes this year, we have also examined the impact of these data on the updated OM.

2 Data and code changes

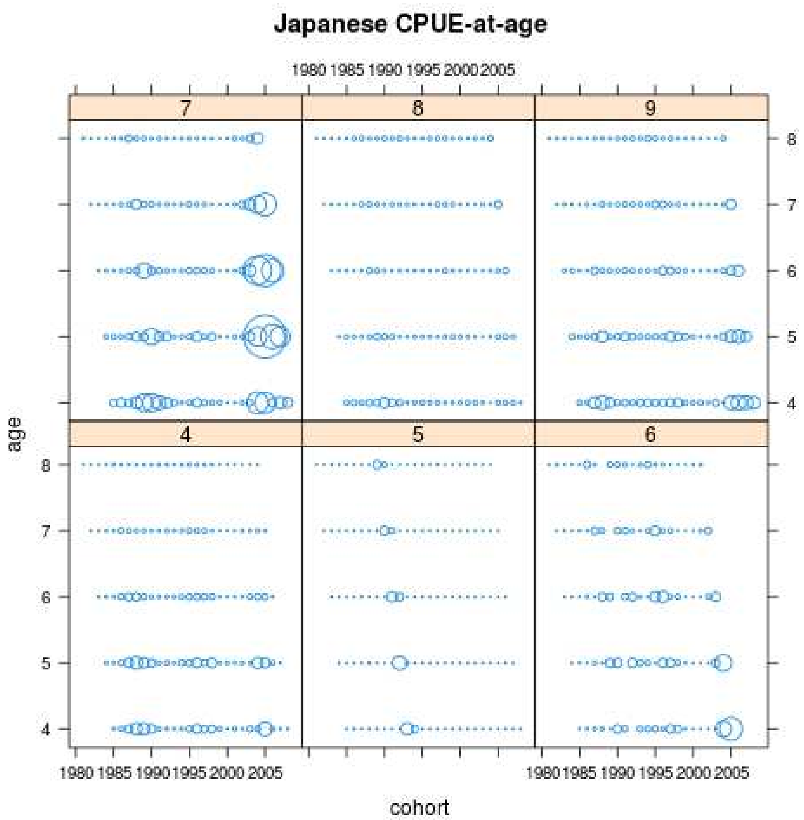
The OMs are fitted to updated data provided during the CCSBT 2014 data exchange. All data update pieces were provided by the 13th June and the outputs shown here are based on the data provided at that time. Data updated in 2014 include: catches by fishery, CPUE, aerial survey, trolling data (not included in these OMs), length and age frequency data. The years in which selectivity estimates for each fishery can change have been altered, primarily to have a 3 year block in the last years for the LL1 fishery.

The close-kin data and method for inclusion in the OM were thoroughly examined in 2012 and 2013, and there are no changes to these data or the method for incorporating them. However, associated with the inclusion of the close-kin data, there has been a change in the method for calculating SSB which was implemented in 2013. This means that SSB as defined in the 2011 reconditioning of the OM is different to the current method. To enable a comparison with 2011 results, the Biomass of fish aged 10 and older is calculated (B10+), which is equivalent to the 2011 method to calculate SSB.

Data and file updates and changes to the projections code have also occurred this year (documented below), and the MP data files have been updated with the new data.

2.1 Exploration of nominal CPUE at age data

In 2011, the nominal and standardised CPUE series data used in the OM were extensively examined to explain the recent increasing trend in CPUE, and to disentangle effects of increased abundance from changes in fishing practices. As noted in Hillary et al (CCSBT-ESC/1107/11) the increased trend in CPUE effects recruitment and steepness in the OMs and can have strong impacts on the predicted future SSB trends and therefore the management procedure.



Fiure 1 Japanese CPUE at age, plotted by cohort, age and region.

Comparing the bubble plots in Figure 1 to the 2011 discussion on these trends we focus on area-specific changes or consistency relative to that analysis:

- Area 4: Similar trends, with an apparent mix of potential recruitment and year-effects (catchability) in the 2008 and 2009 CPUE increases. Also, potentially a signal for a stronger 2005 year-class.
- Area 5: no apparent change.
- Area 6: Data ends in 2006 so better to analyse the New Zealand CPUE which shows a clear year-effect in the 2008 and 2009 data still, as well as a weaker looking year effect for the 2011-2012 CPUE.
- Area 7: Similar trends and a mixture of potential cohort and year effects in recent years with, arguably, better evidence of a stronger 2005 year-class moving through the most recent data.
- Area 8: no obvious shift observed from 2011.
- Area 9: Continued evidence of the catchability year-effect but also stronger evidence of a larger 2005 year-class than in previous analyses.

Based on the most recent data, it is difficult to ascertain whether the catchability increase observed in the raw data and estimated well in the OM for 2008 onwards has continued, though the same signals remain in the data as in 2011 that indicate some kind of noticeable change in catchability did occur around 2008 and 2009. There is now fairly clear evidence in the data of a stronger 2005 year-class moving through the data, but only really in areas 4, 7 and 9. Based on this basic analysis of the raw CPUE-at-age, we suggest that the catchability change - estimated in 2011 to be around 35% - should be re-estimated to see if it still warrants inclusion as a robustness test.

3 Reconditioning the OM

The OMs have been run using the grid defined at ESC 2013, and is defined as base2013 (Table 1). The sqrt file has been updated to include extra years of data and to have a 3 year block for LL1 and LL2 selectivities in the most recent years.

Table 1 The "Base2013" grid used in the diagnostics and results presented in this preliminary reconditioning of the SBT operating models

	LEVELS	VALUES	PRIOR	WEIGHTING
Steepness	5	0.55, 0.64, 0.73, 0.82, 0.9	Uniform	Prior
M0	4	0.36, 0.4, 0.45, 0.5	Uniform	likelihood
M10	4	0.05, 0.075, 0.1, 0.125	Uniform	likelihood
Omega	1	1.0	NA	NA
CPUE	2	2, 3	Uniform	Prior
Q Age Range	2	4-18, 8-12	0.67, 0.33	Prior
Sample Size	1	sqrt	NA	NA

3.1 Fits to the data

3.1.1 SIZE COMPOSITION

Figure 2 -5 show the fits to each of the four longline fisheries' length composition data. Fits seem to be reasonable in recent years, with some lack of fit in historical years for some fisheries which have been noted in the past.

For the LL1 fishery, there is some lack of fit to the earliest years of data, but the rest of the time series appears to fit well (Figure 2). There are some years (2006, 2007) and the most recent year (2013) where spikes in the length frequency are not well fitted.

For LL2, the early years of the series do not fit well, after which fits are reasonable (Figure 3). The fit to the 2010 data are good, but there is some lack of fit in the following years.

For the LL3 fishery the length data fit well in the early years, and the early 2000's data fit well, but not the latter years (Figure 4), which have been down weighted to zero. The years not fitted are: 1972-2004 and 2008-2013.

The LL4 fishery length data fits are reasonable with some spikes indicating lack of fit in some years (Figure 5).

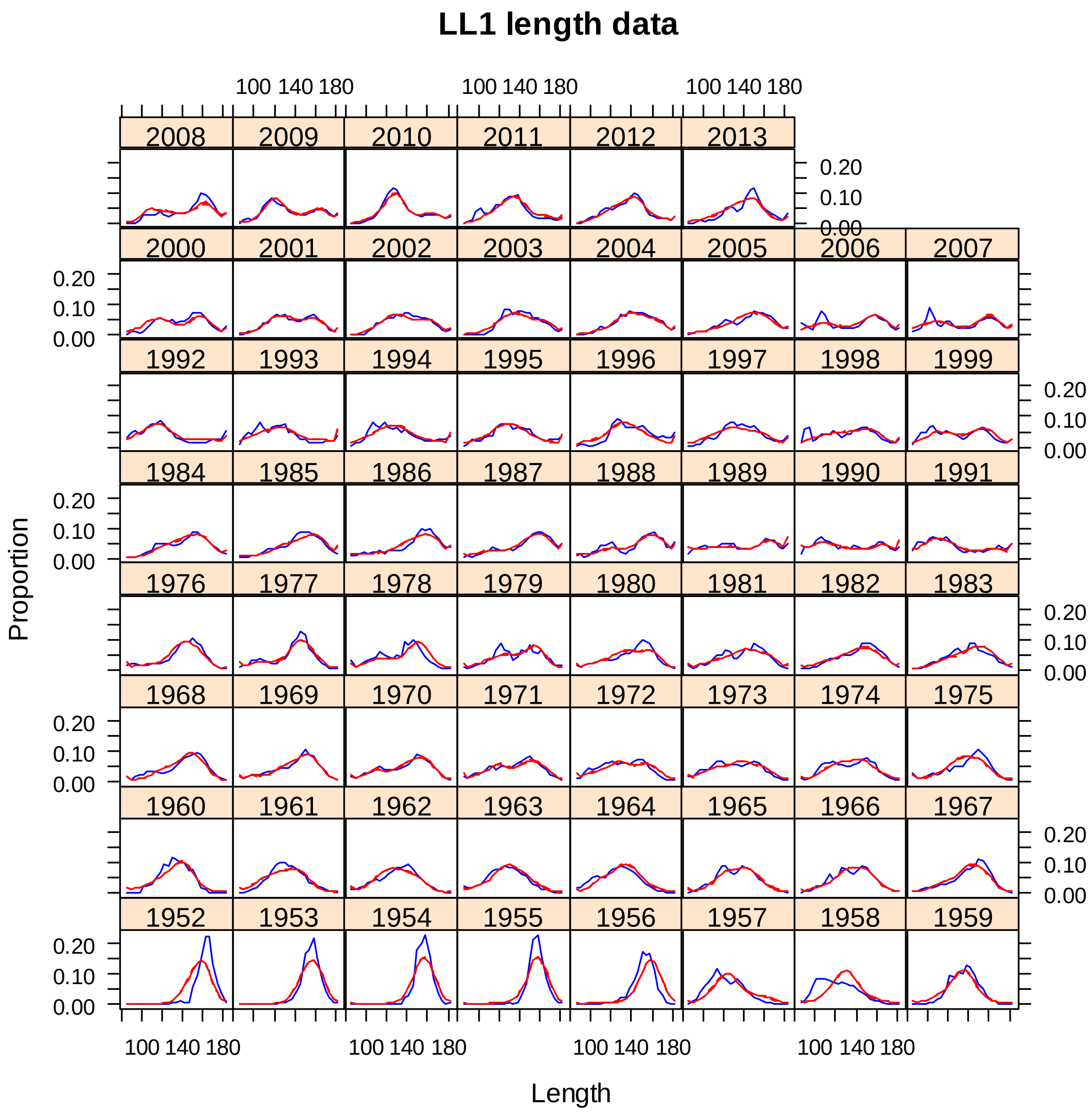


Figure 2 Predicted (red) versus observed (blue) LL1 length composition.

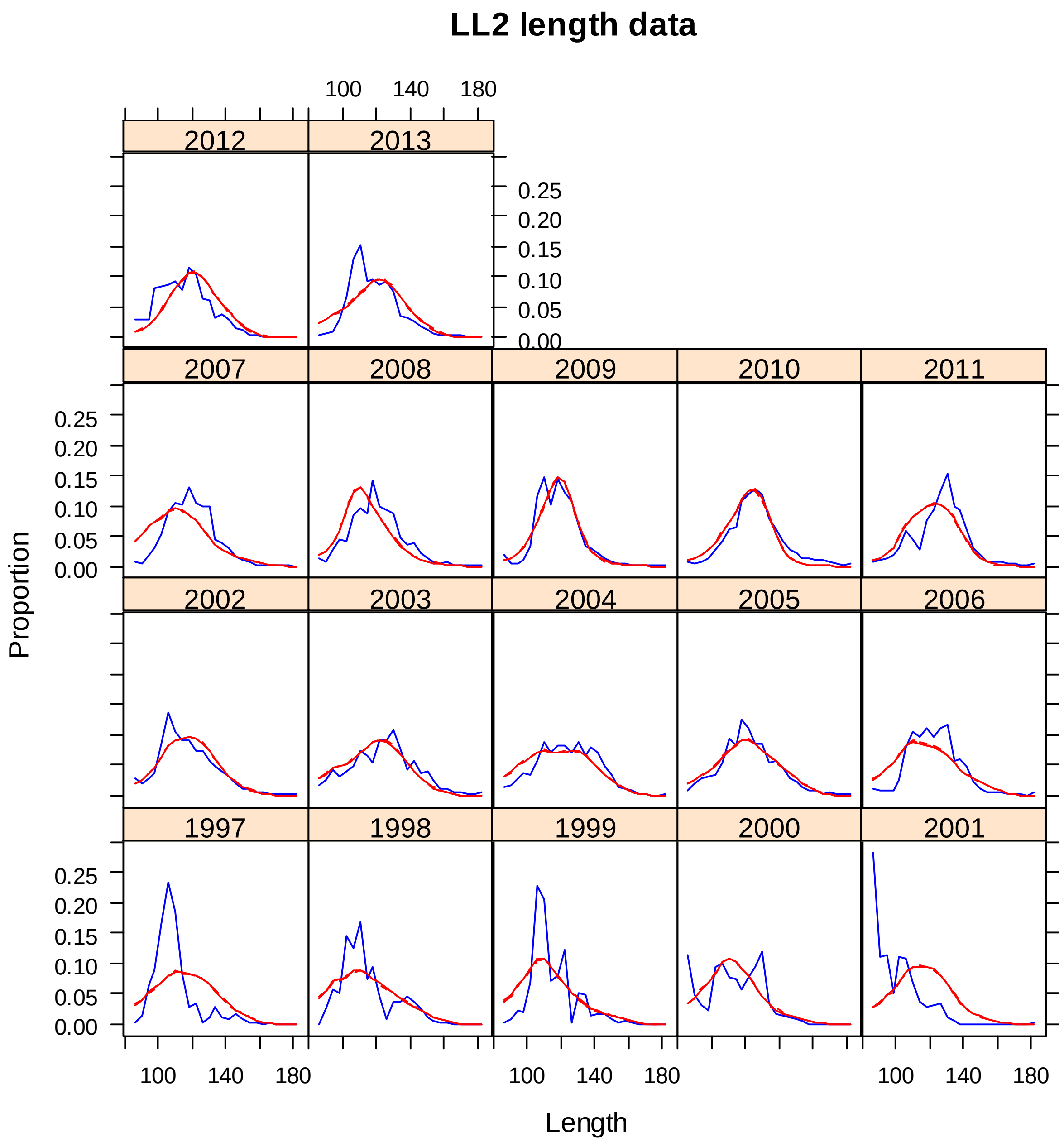


Figure 3 Predicted (red) versus observed (blue) LL2 length composition.

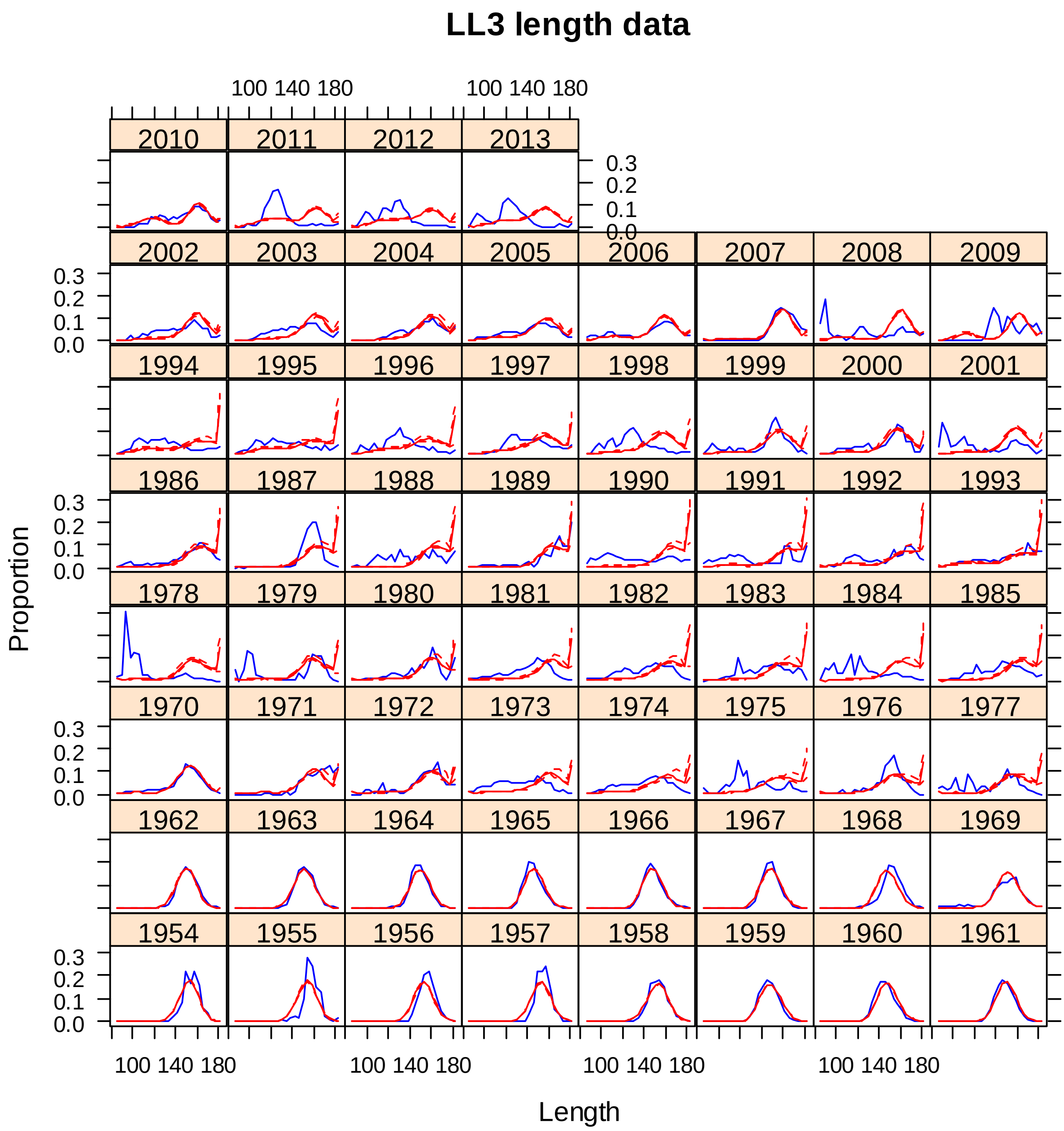


Figure 4 Predicted (red) versus observed (blue) LL3 length composition. Note 1972-2004 and 2008-2013 are not fitted.

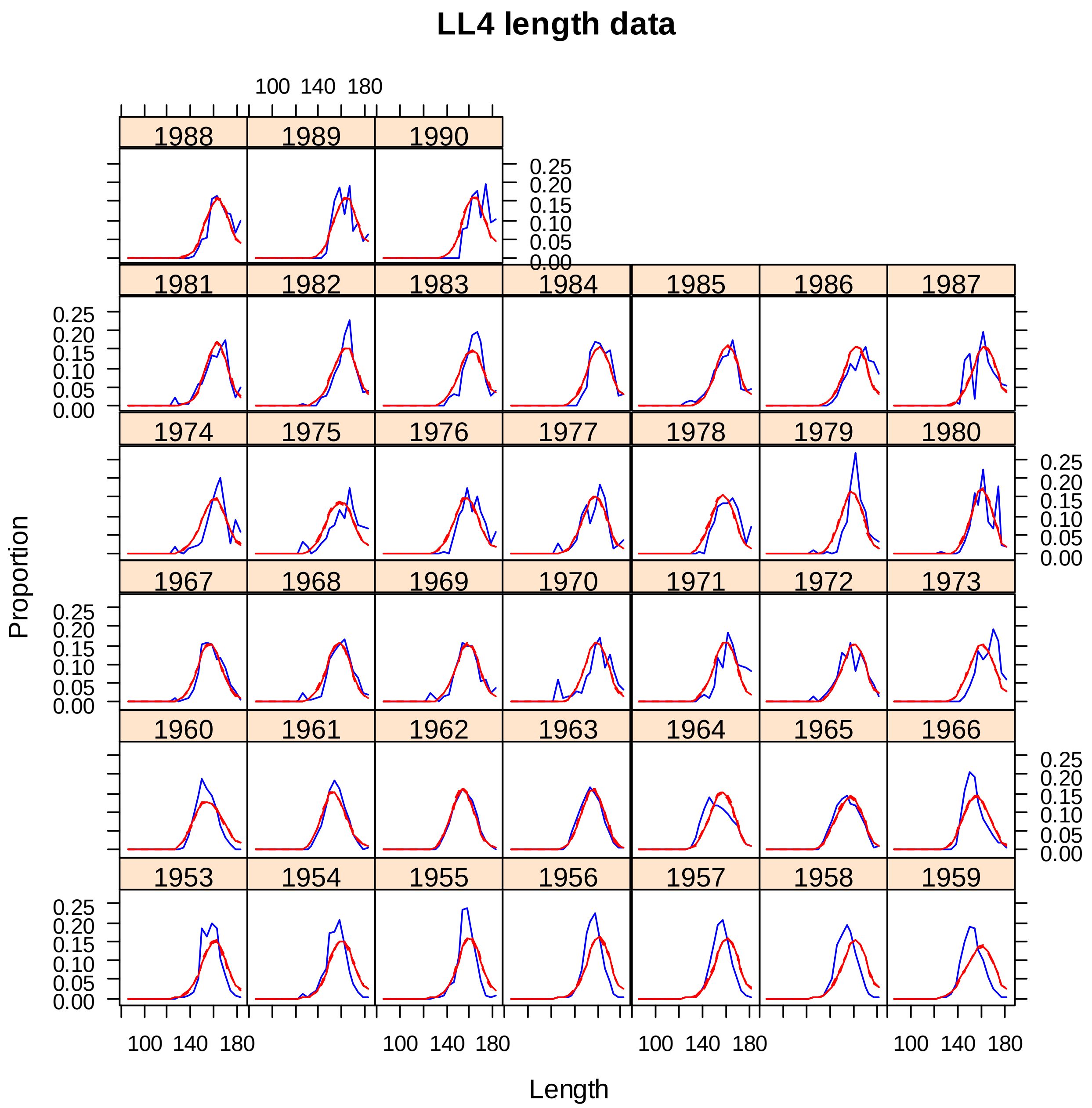


Figure 5 Predicted (red) versus observed (blue) LL4 length composition.

3.1.2 AGE COMPOSITION AUSTRALIAN SURFACE FISHERY AND INDONESIAN FISHERY

The fit to the Australian Age frequency data is good throughout the series, and fits very well in the most recent 15 years (Figure 6).

The fits to the Indonesian age data are reasonable for most years (Figure 7), with some spikes in the age frequency that aren't fitted as well in the most recent 3 years. The age frequency in the most recent year in the OM (data from season 12/13) is quite different from previous years with a large number of small fish observed. The OM has fitted to this data reasonably well, but underestimates the numbers of small fish. It is not yet known whether these small fish observed in the catch monitoring program are from the spawning ground or possibly from further south, as has occurred in some earlier years (2005-2007). Preliminary data for 2013/14 show even larger numbers of small fish but these data are not included in the updated OMs. The issue is being further investigated (Craig Proctor, pers comm.), but no additional information is currently available and this should be further discussed at the ESC.

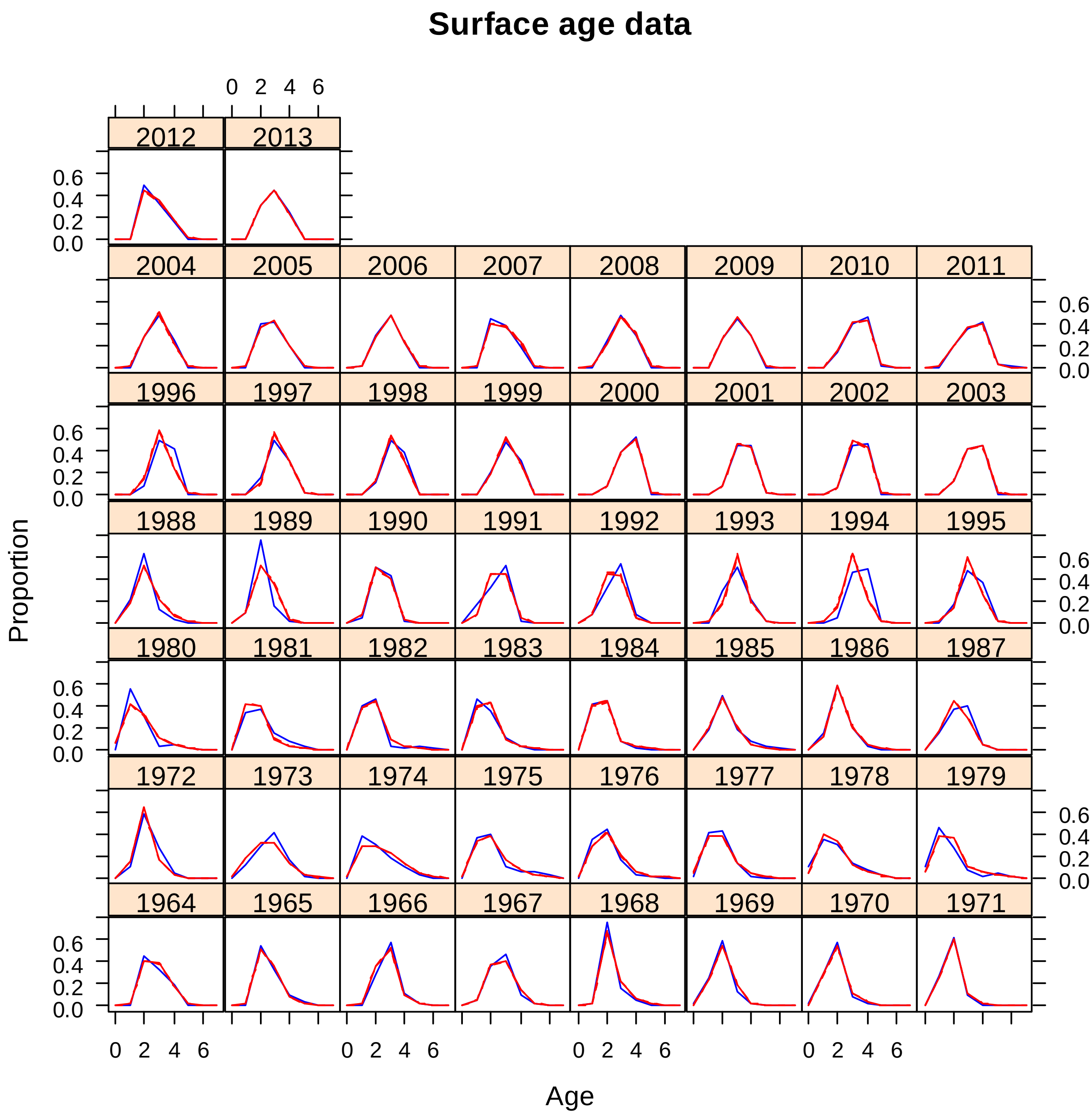


Figure 6 Predicted (red) versus observed (blue) surface fishery age composition.

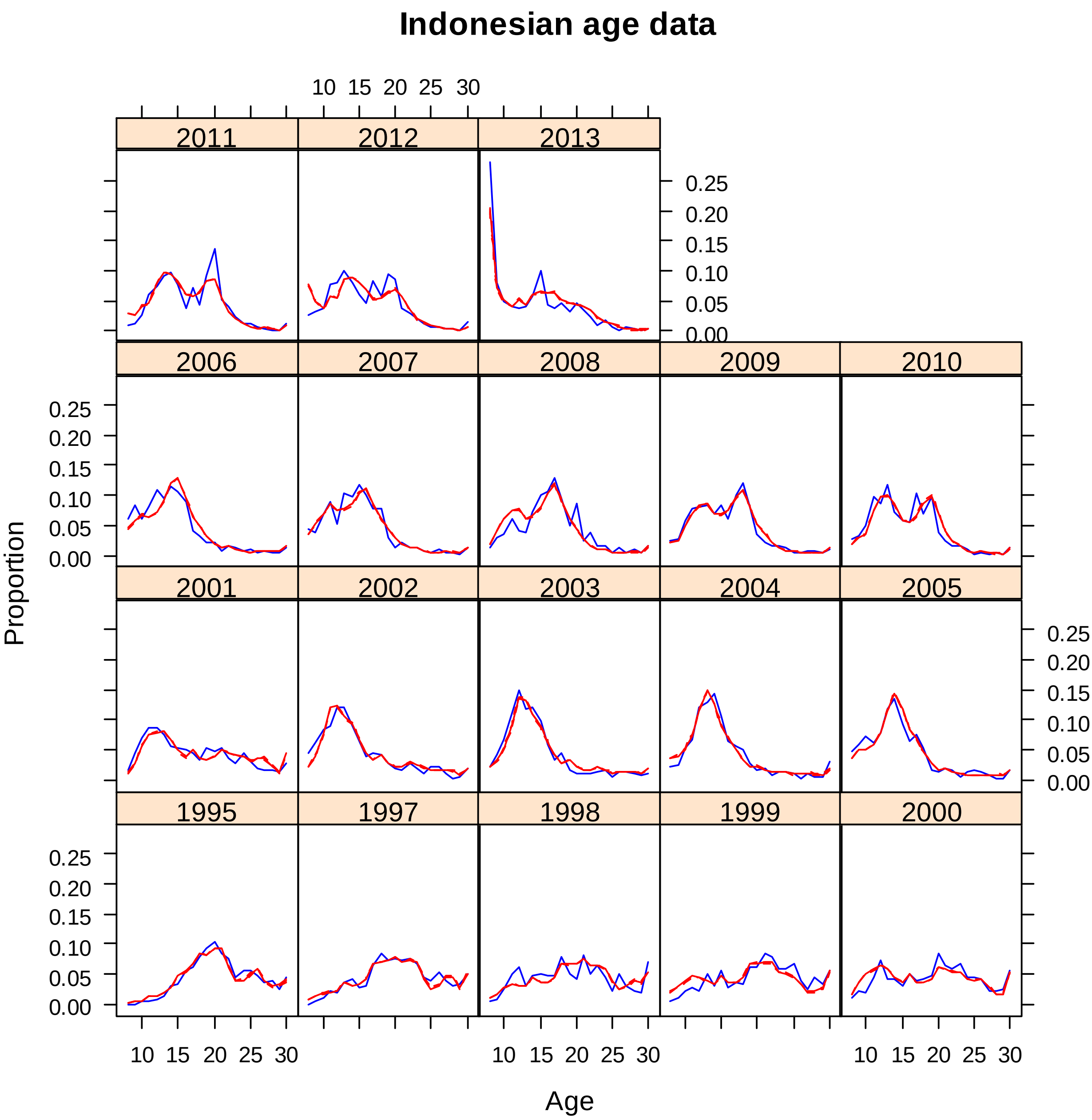


Figure 7 Predicted (red) versus observed (blue) Indonesian fishery age composition.

3.1.3 LL1 CPUE AND SCIENTIFIC AERIAL SURVEY DATA

The fits to the LL1 CPUE data and aerial survey data relative to the expected values from the reference set of OMs in the grid (Figure 8), indicates that the CPUE data fit reasonably well, but the aerial survey data fit less well. The fluctuations in the recent years in the aerial survey are not well fitted in the OMs, with the most recent highs and lows outside of the 90% range, but the increasing trend is replicated.

The posterior predictive analysis presented in previous papers (Hillary et al 2011; CCSBT-ESC/1107/11) has been repeated here (Figure 8). The posterior predictive analysis (Gelman et al, 1994) allows us to examine how well the OM explain the observed data, by examining residuals from simulated data and the observed data, and calculating the probability (p-values) that the predicted median deviation is more extreme than the observed median deviation (see Hillary et al, 2011, for a more detailed explanation). Values around 0.5 are the ideal, with p-values greater or less than 0.5 indicating that the OM predictions are more/less variable than the actual data, and values outside the 90% interval suggest there is something wrong with the model and likelihood (Hillary et al, 2011). The p-value of (0.88) indicates that the OM predicts CPUE data that are consistently more variable than the observations. The predicted Aerial Survey data are significantly less variable than the observed aerial survey data.

As noted in Hillary et al. (2013) these results merely confirm the assumptions and input settings we define in the OM for these abundance indices. For CPUE, we define the minimum CV of these data to be 0.2, whereas the empirical estimate of the total error CV is actually lower (ca. 0.14); as a result, when simulating these data from the OM with a CV of 0.2 they are more variable than the observations. For the aerial survey, we restrict the process error CV to be 0.18, whereas if estimated it would be somewhat higher (ca. 0.34); as a result, when we simulate data from the OM it is less variable than the observations. The posterior predictive analyses confirm that what we assume in the OM structure is indeed what we obtain post-fitting but, because the reasons for these assumptions have been discussed and agreed, they do not suggest something systemically wrong in the fits to these indices. What is perhaps more important, and clear from the graphs, is that the OM prediction intervals generally encompass the data and that the trends are fairly well estimated.

To check whether the 2014 aerial survey point is outside the range tested in 2011, we first need to appropriately rescale the current aerial survey point given the unavoidable changes in overall magnitude in the survey from year to year. This catchability ratio approach to the problem has been documented in detail in paper CCSBT-OMMP/1307/4 and is used here. For the base grid used in the MP tuning in 2011 (MP3_2035_3000_inc_base) the probability that the (rescaled) 2014 aerial survey point is greater than the simulated 2014 aerial survey used in 2011 is 0.991. So, in isolation, this indicates it would meet the exceptional circumstances criterion in the meta-rules process. However, following a similar argument to why the 2011 low aerial survey point was considered not to meet the exceptional circumstances criterion, the following robustness trials run in the MP testing (Laslett, troll, highaerialCV) all had values of the 2014 aerial survey greater than the current observation more than 5% of the time. So, it could be argued that the MP testing in 2011 did cover the range of values of the aerial survey currently being observed.

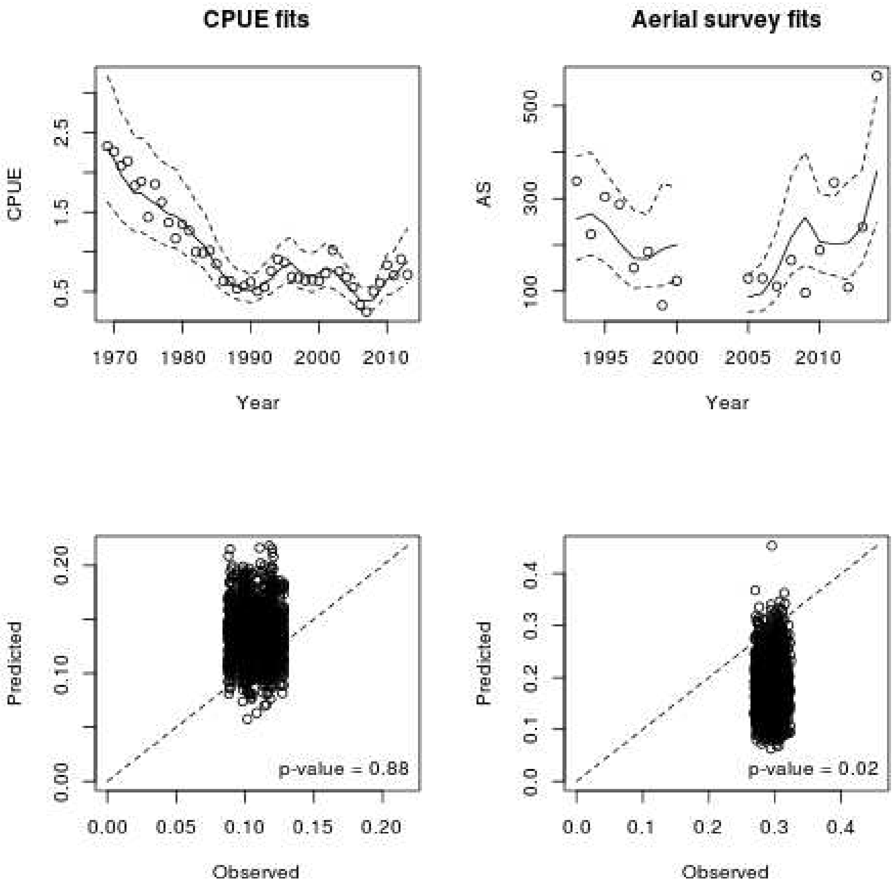


Figure 8 Median (full) and 95% (dotted) predictive interval and observed values (circles) for the CPUE (top left) and aerial survey (top right) data. Posterior predictive p-values and the observed and predicted discrepancy values for the CPUE (bottom right) and the aerial survey (bottom left).

3.1.4 1990'S TAGGING DATA

The fits to the tagging data are similar to previous re-conditioning, with the observed and predicted total recaptures-at-age for each year of release fitting well (based on the most likely model from the grid – h2m3M2O2C3a2) (Figure 9). The largest numbers of releases were in 1996 and 1997 and these data fit well. In the earlier years the fits are also reasonably good.

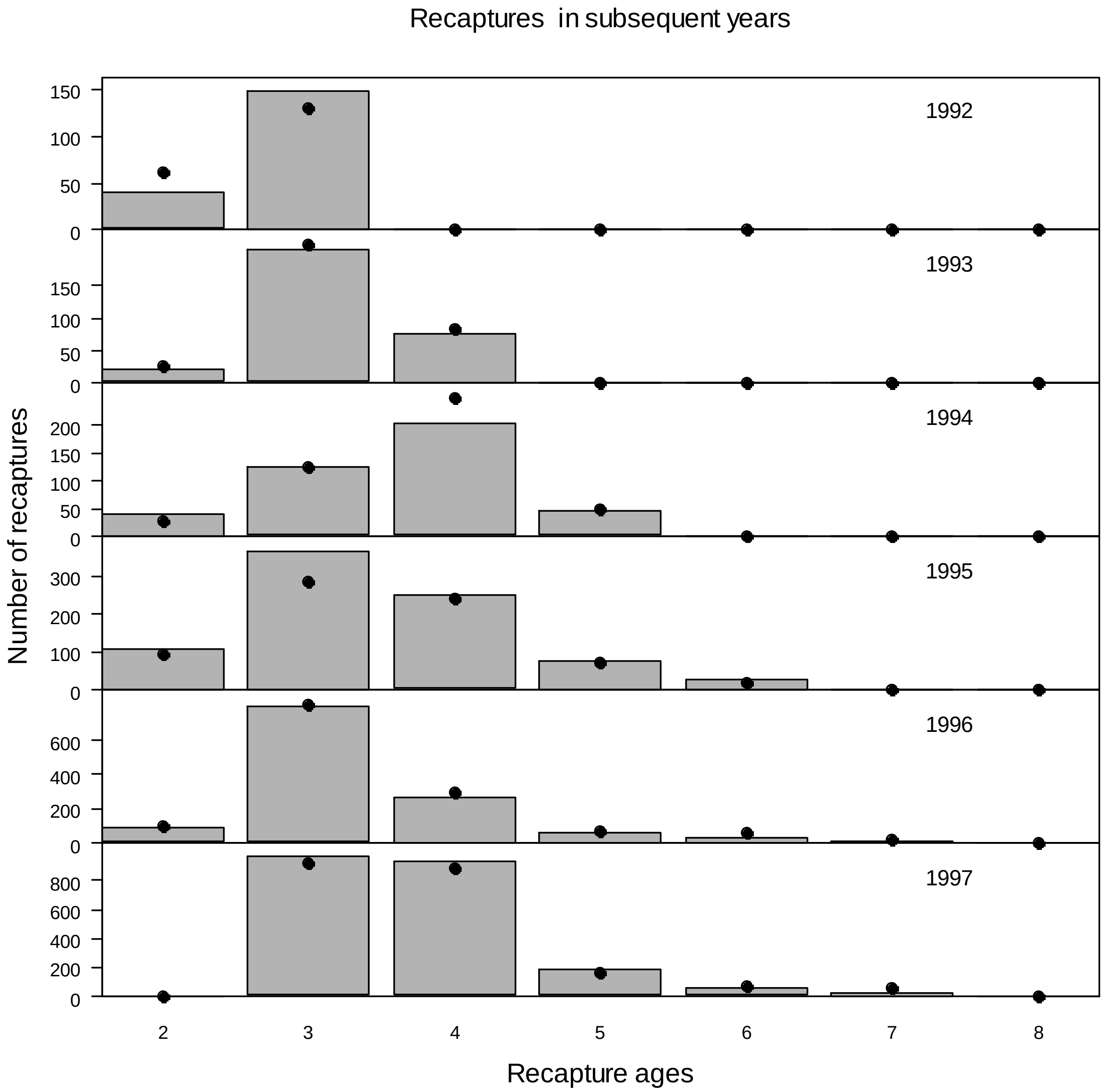


Figure 9 Predicted (bars) versus observed (circles) total recaptures for each year of release for the 1990s tagging data and for the grid configuration with the best overall fit.

3.1.5 FITS TO CLOSE-KIN DATA

The method for integrating the close-kin data into the OM and the fits to the data were explored in detail in 2012 and 2013 (Hillary et al., 2012; Hillary et al., 2013), and the ESC agreed that these data should be included in the 2014 stock assessment models. A new maturity schedule based on available biological information (Davis et al, 2001 - CCSBT-ESC/0108/16) associated with inclusion of the close-kin data was adopted at the 2013 ESC (ESC report: Anon, 2013).

The fits to the close-kin data in Figure 10 are reasonable. There are some data points that the model does not fit well, but the trends are captured and we should always be mindful of the very low sample sizes in terms of the number of POPs at this aggregation level. The ability of the models to explain the close-kin data are shown in the posterior predictive analysis plots in Figure 11. In a similar way to the analysis of the aerial survey and CPUE abundance estimates described above and in Hillary et al (2013), the posterior predictive analysis provides information on the models ability to explain the close-kin data; the p-value of 0.36 indicates that the model is adequately fitting the data, and the distribution of points in the figure indicates that the predicted values have similar absolute median deviation to the observed values. This also suggests that there is very little if any clear over-dispersion in these data, conditional on the OM structure. If this *p*-value was very low we would suspect that the OM was over-fitting these data and, hence, highly likely that the data were over-dispersed with respect to the OM. Since this feature is not apparent, and the data are fairly well explained by the variability and structure of the OM, we see no reason to move to a more complicated likelihood for the close-kin data this year.

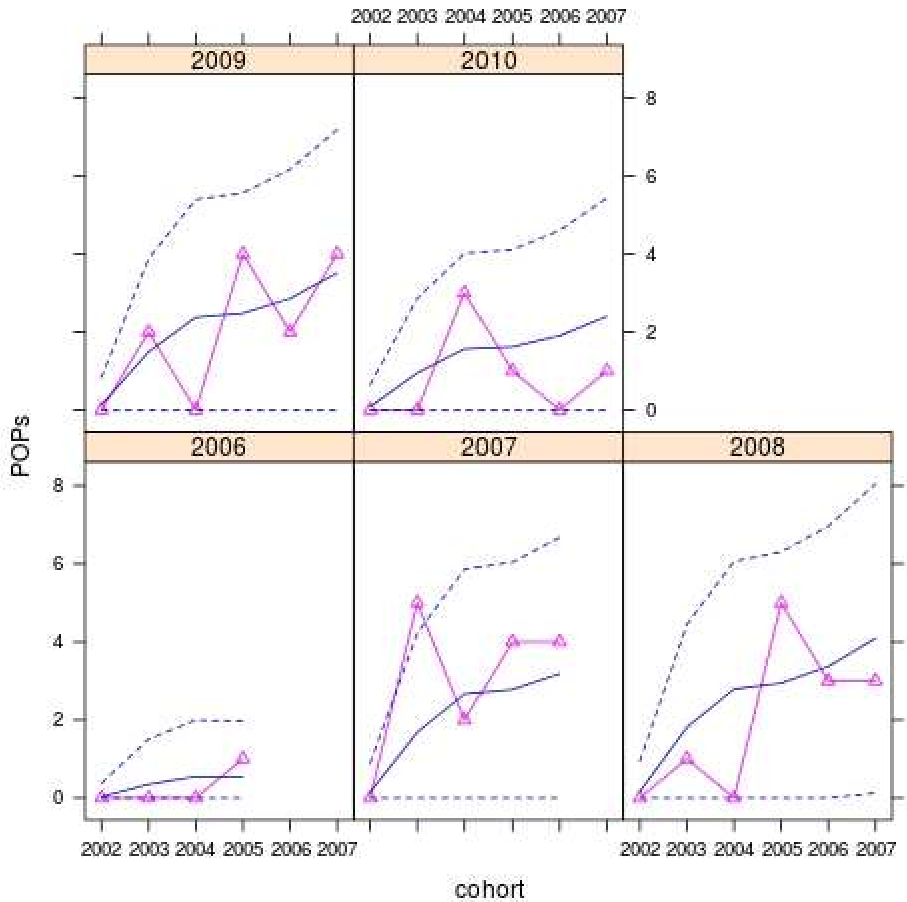


Figure 10 Predicted (blue, median and 95% CI) versus observed (magenta triangles) number of POPs for the close-kin data aggregated to the cohort level (i.e. across both adult capture year and age).

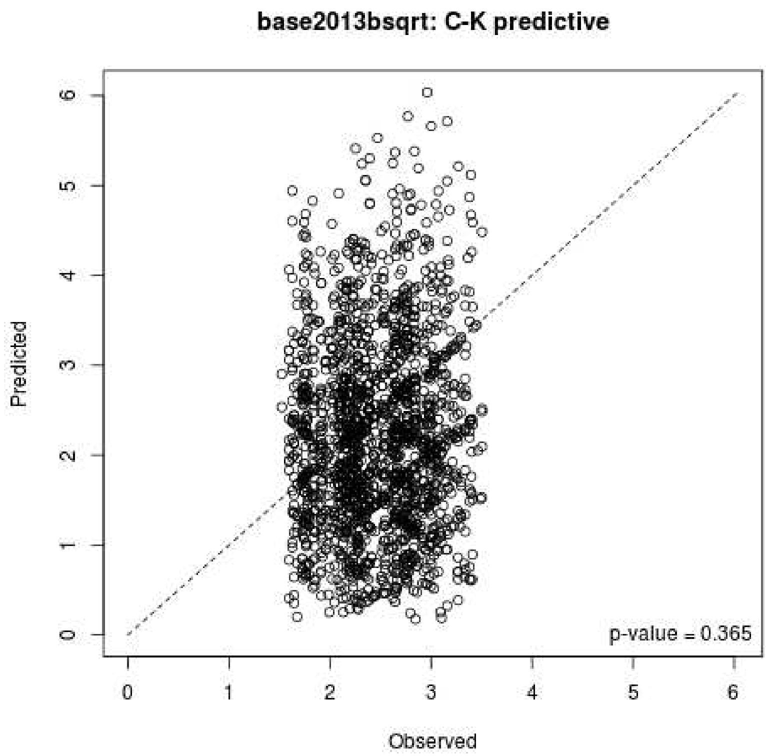


Figure 11 Posterior predictive performance (p-value and discrepancy plot) for the close-kin data aggregated to the cohort level (i.e. across both adult capture year and age).

3.2 Parameter estimates and preliminary stock status estimates

The summary of the grid sampling is shown in Figure 12. The sampling across the grid looks similar to the exploratory work undertaken in 2013, but with higher sampling of the lowest M10 value in the base grid. At the 2013 ESC lower M10 values were explored, but excluded from the current grid. The range of values used for M10 in the grid may need to be further evaluated.

Scatter plots of the likelihood profiles for steepness, M0 and M10, for each of the fitted datasets and the penalty functions are shown in Figures 13 to 18. These figures indicate that there is a mild preference for higher steepness values from the aerial survey and LL1 data, but the recruitment penalty function preference is for lower steepness value. The tag data has a slight preference for M0=0.45, and least preference for the lowest M0 value. The tag data has a strong preference for lower M10 values, and the Australian and Indonesian Age frequency data indicate preferences for higher M10 values. The penalty functions don't seem to be affecting M10 preference. M10, M0 and steepness are all strongly correlated. Steepness is uniformly weighted in the grid sampling as agreed in 2013.

The depletion estimates from this preliminary re-conditioning of the operating model indicate that Stock is at 6-9% of initial biomass, median 7%, based on the definition of SSB used in 2011. Using the new SSB measure which uses the new maturity schedule adopted in 2013 associated with incorporating the close-kin data, the depletion estimate is 8-12%, median 9%.

The Recruitment time series shows very high recruitments in the most recent years.

Current fishing mortality estimates do not indicate overfishing: F/Fmsy is 0.64 (80^{th} percentile range: 0.38 - 0.94).

Current biomass is estimated to be less than current estimates of BMSY: Bcurrent/Bmsy is 0.38 (80^{th} percentile range: 0.27- 0.69)

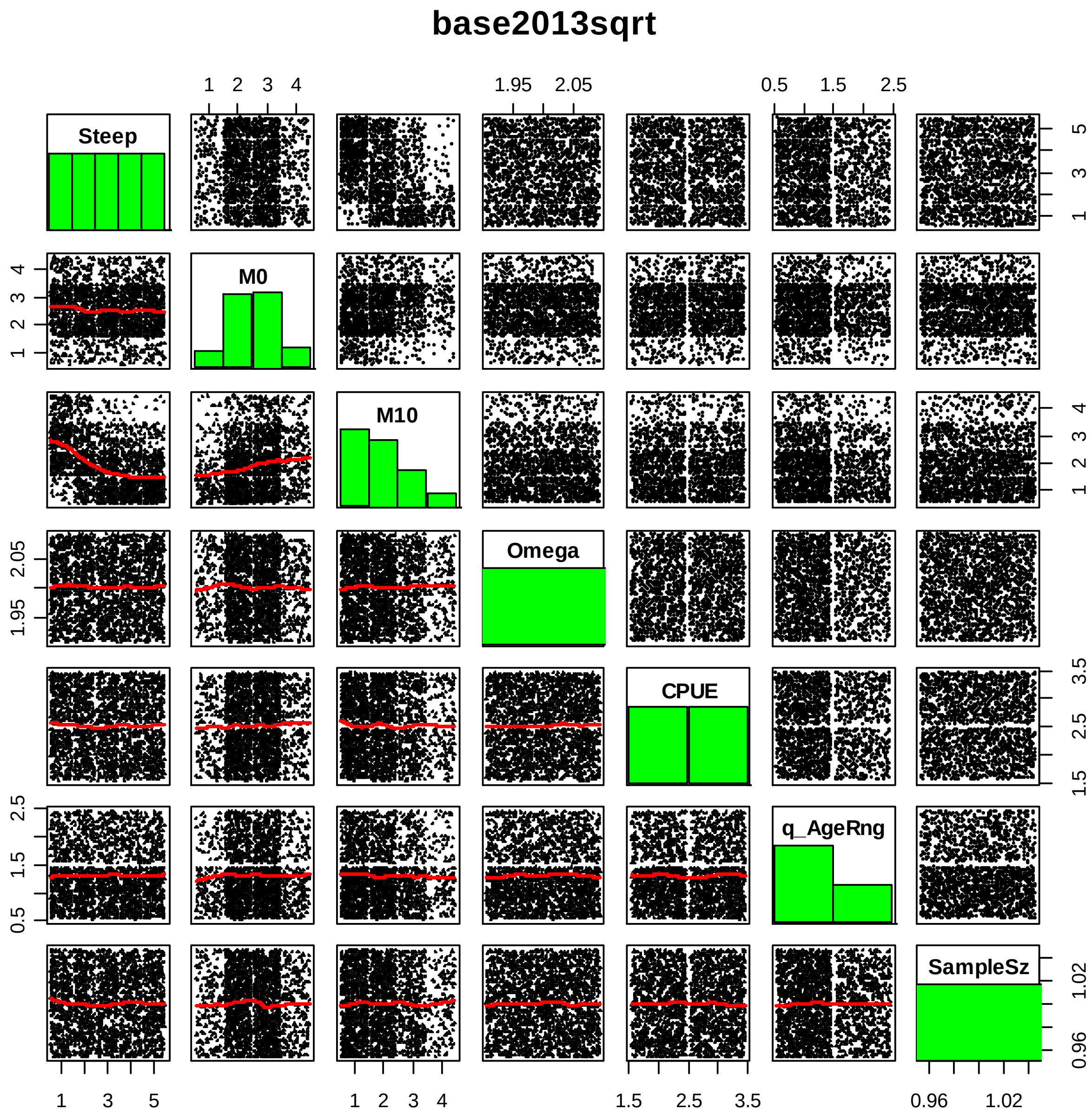


Figure 12 Grid sampling summary for the base grid on the updated data.

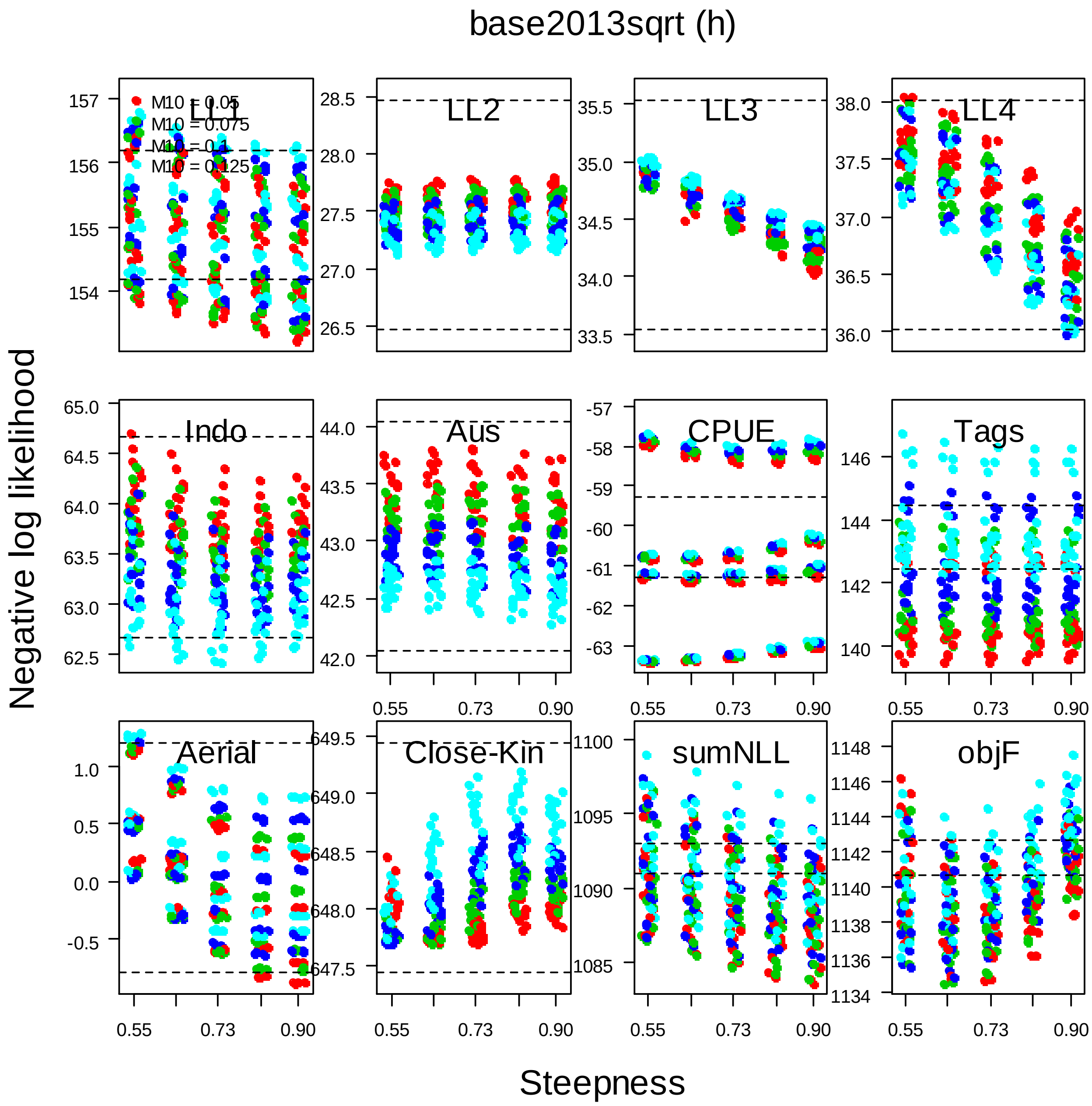


Figure 13 Likelihood profile plots for steepness.

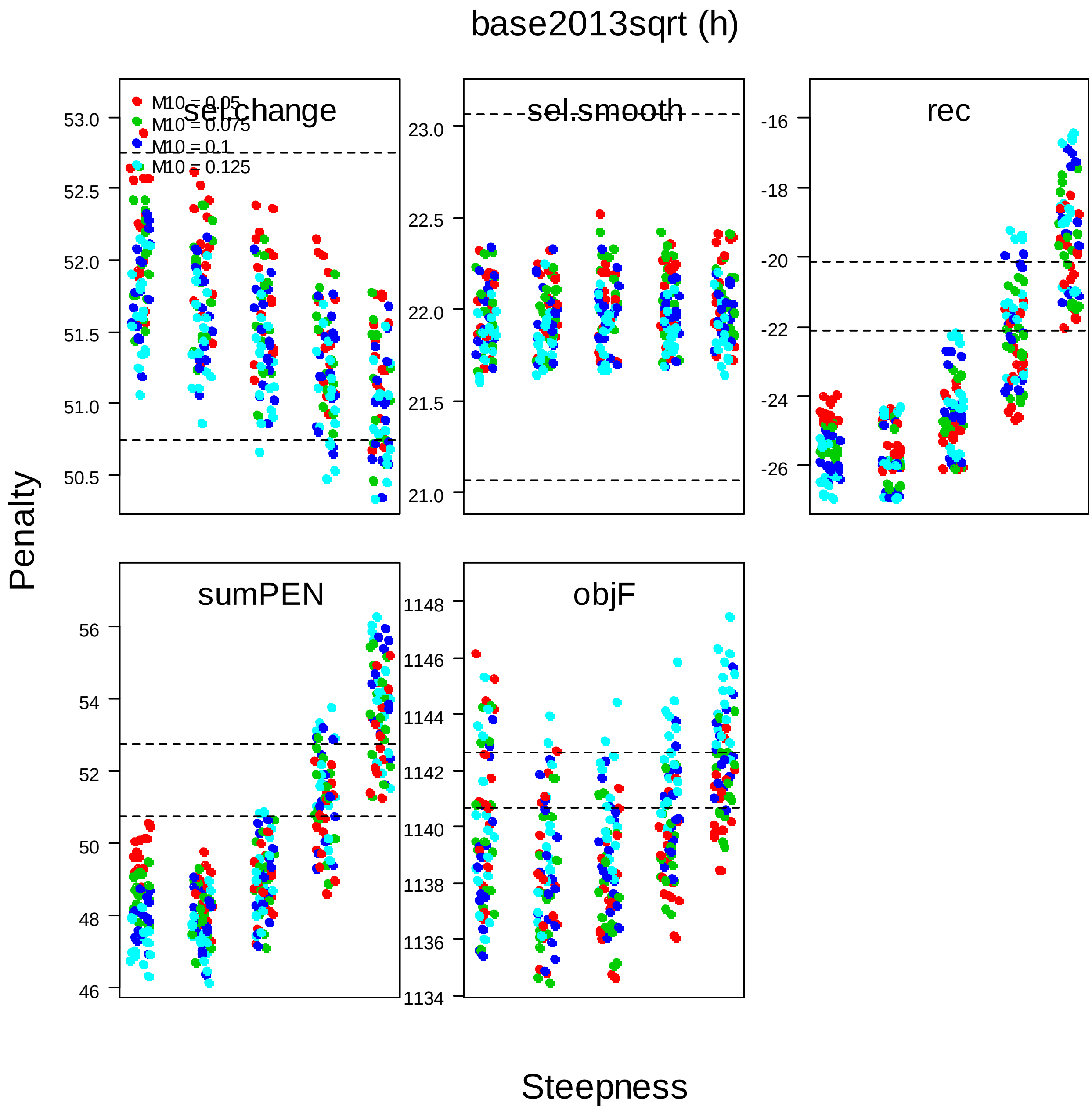


Figure 14 Penalty profile plots for steepness.

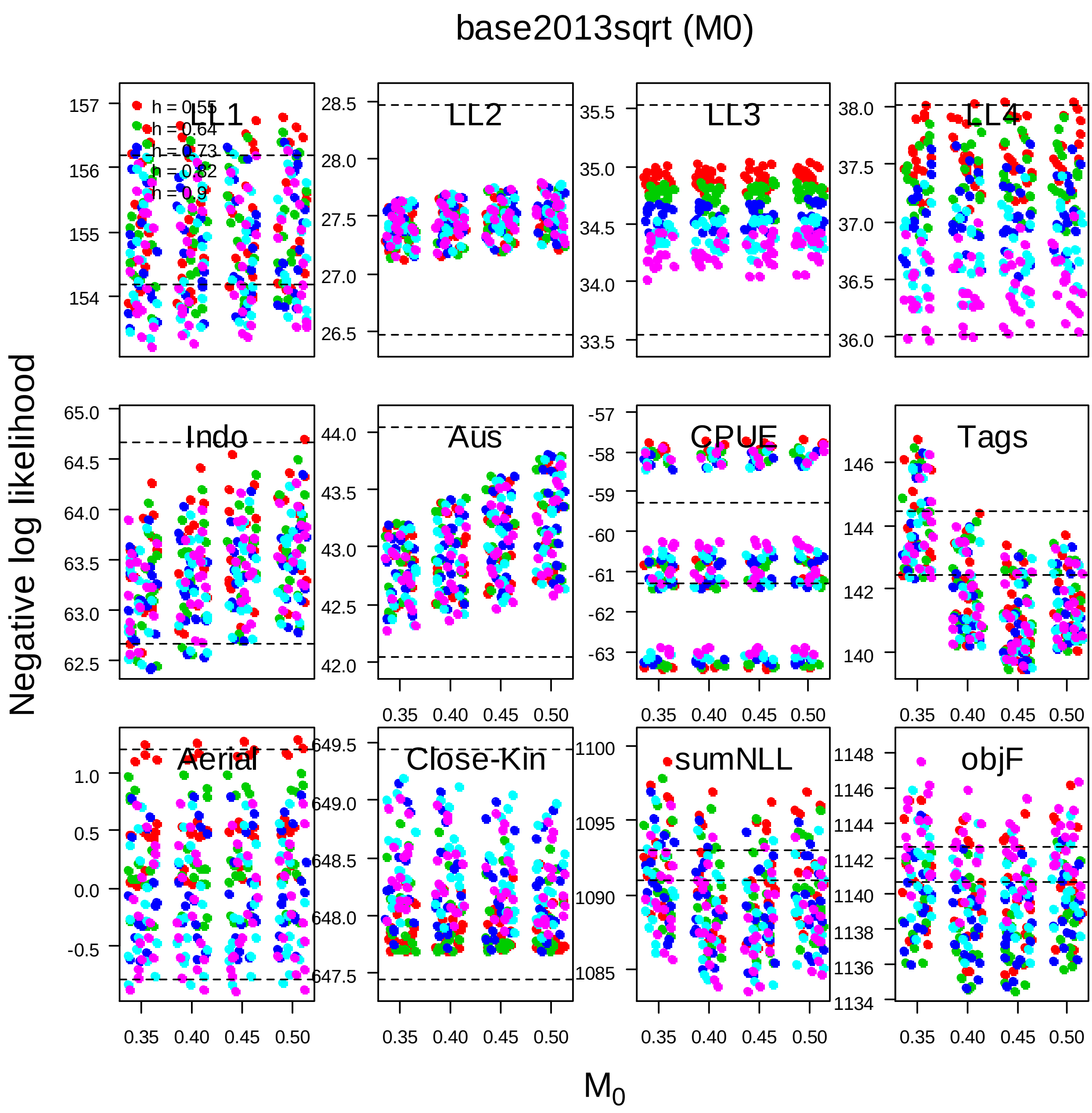


Figure 15 Likelihood profile plots for M_0.

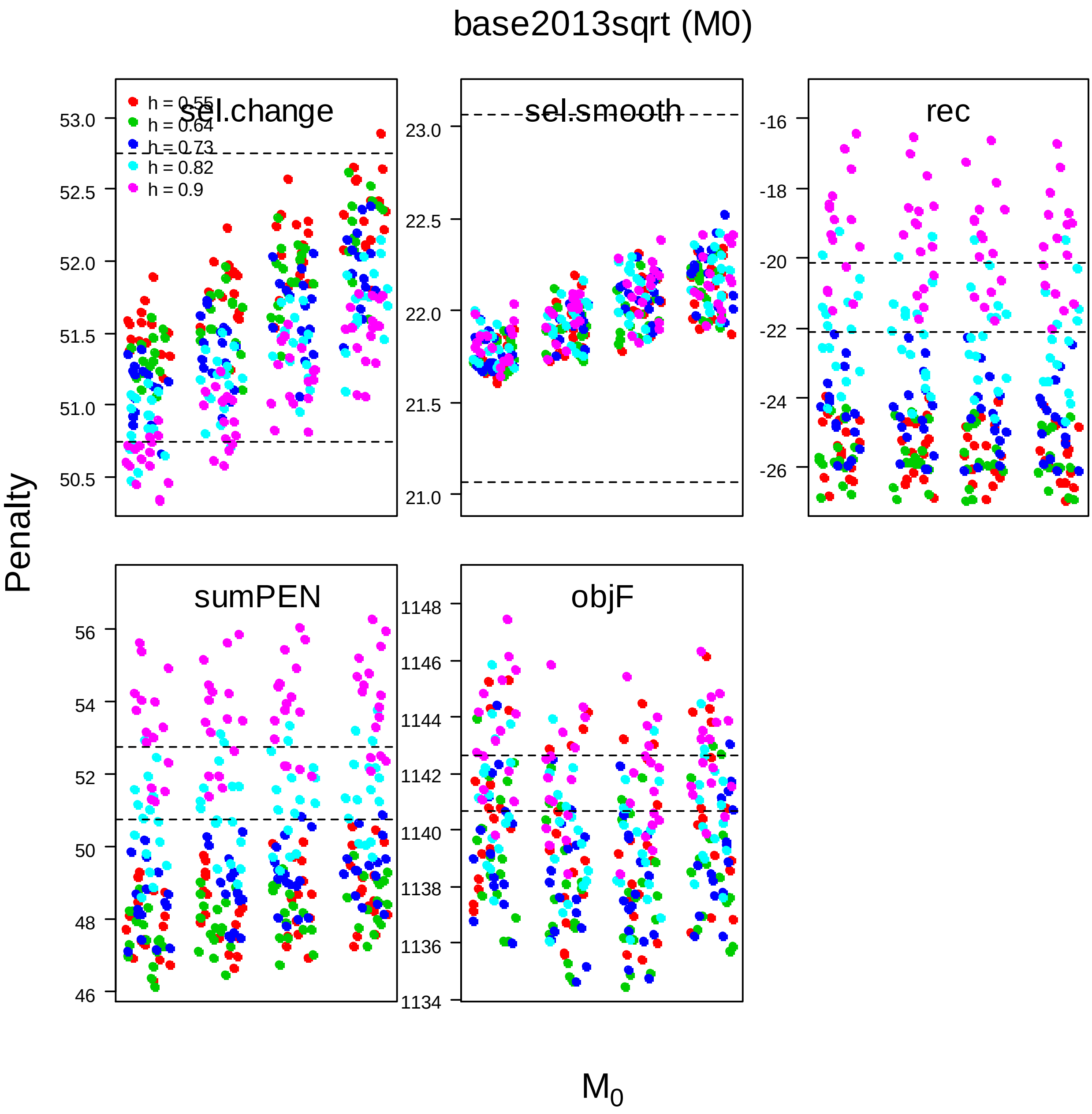


Figure 16 Penalty profile plots for M_0.

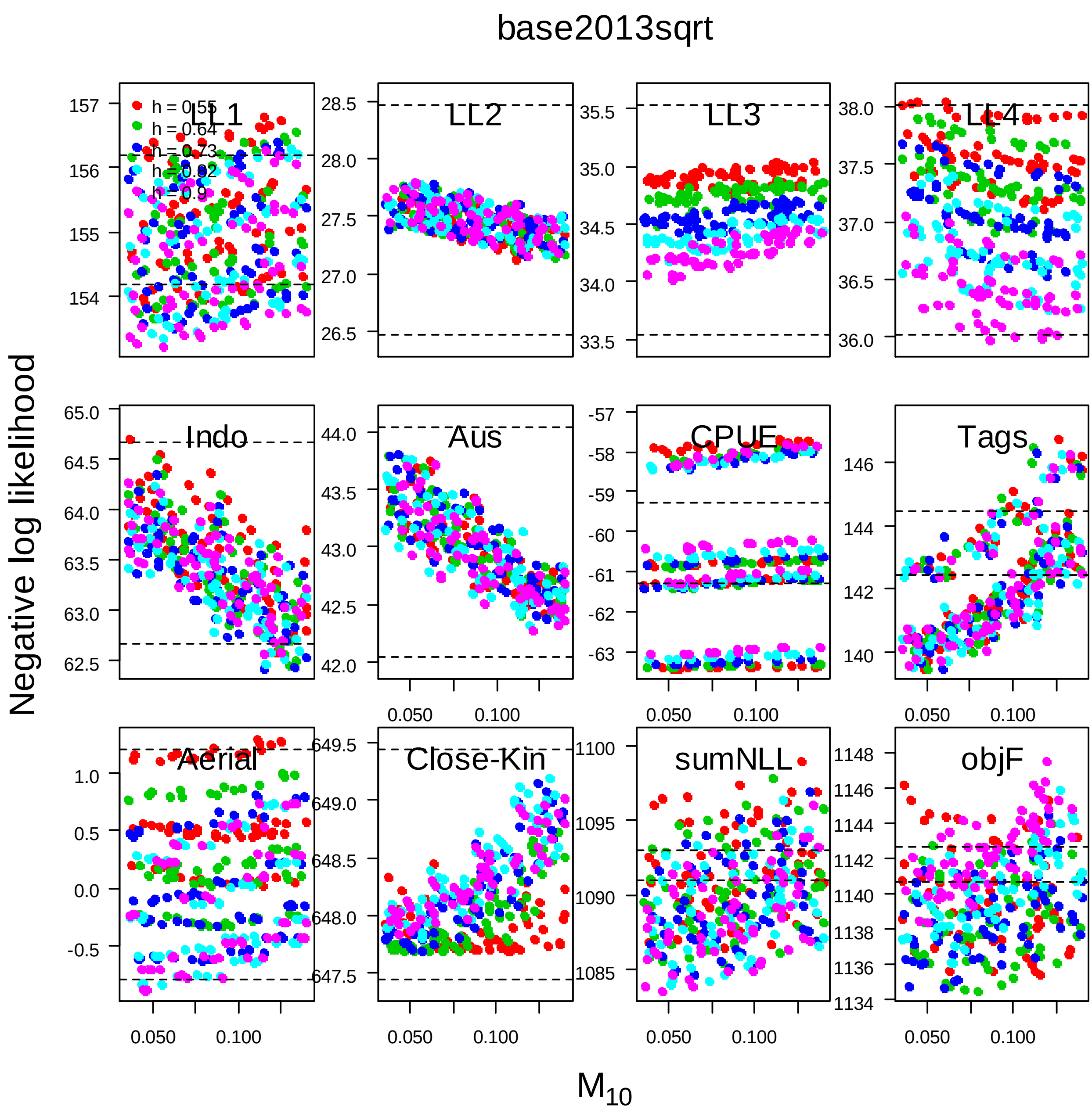


Figure 17 Likelihood profile plots for M_{10}.

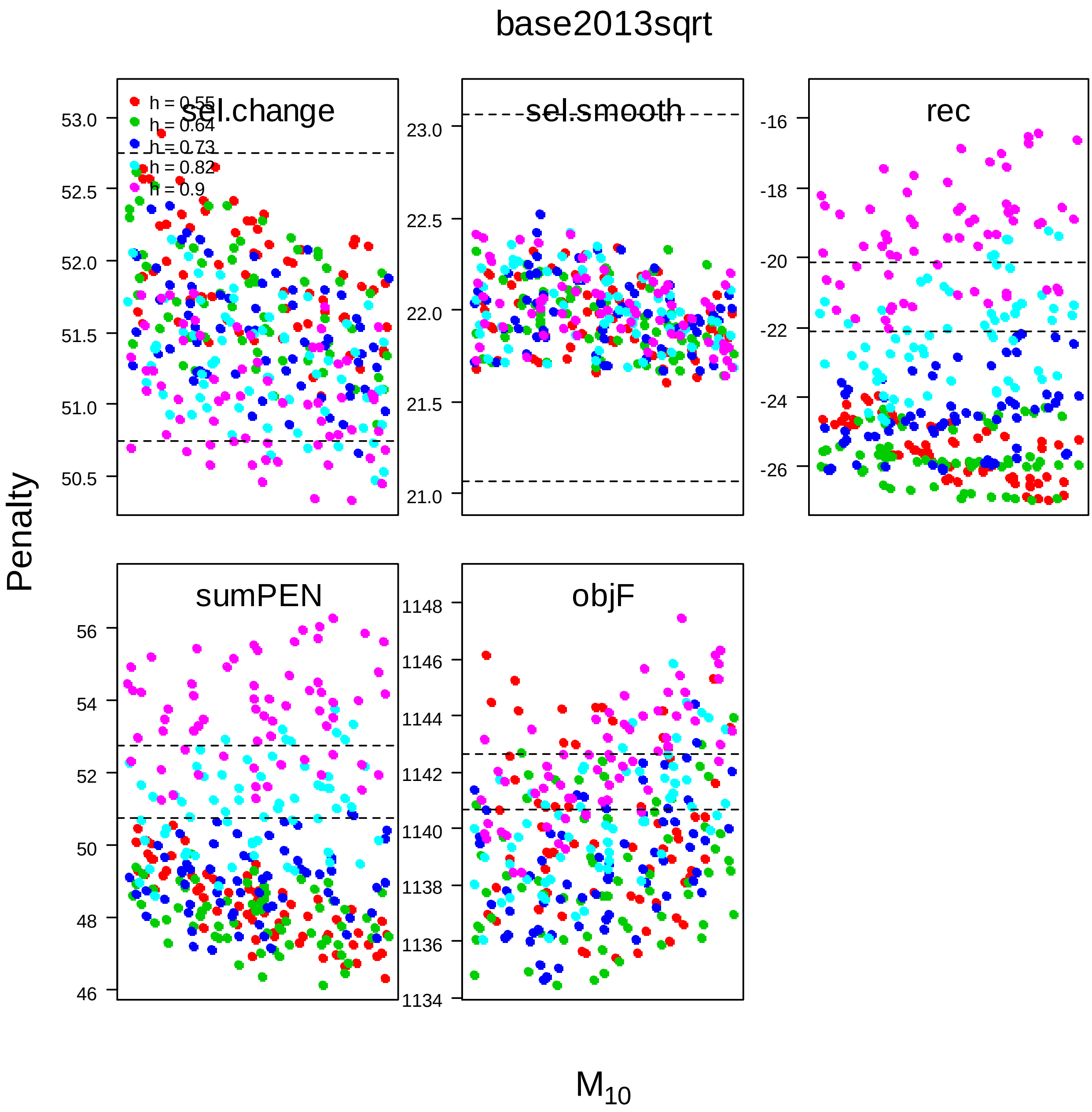


Figure 18 Penalty profile plots for M_{10}.

4 Projections

The projections code has been updated to address: 1) changes to C++ string handling, 2) changes in the SSB calculations and to provide outputs of both SSB and B10+, 3) changes from updated conditioning models, and 4) to start in year 2014. There remain a few hardwired pieces that were also updated or no longer apply. The changed code has been pushed to github, which is the version control system currently in place.

The Management Procedure is used in projections to set three year blocks of TAC. The MP code and historical data file have been updated to account for: 1) additional years of data, 2) the new time series for CPUE (average of the base series and adjusted for overcatch scenarios), 3) the new time series of aerial survey data, and 4) the adjustment to the qratio value required in the MP because of the new aerial survey time series (Preece et al, 2013: MP paper).

The recent high recruitments and positive trends observed in the SSB from the conditioning models have resulted in projections, that use the MP to set the TACs, which show rapid rebuilding of the SBT stock (Figure 19). This is a result of updated data and inclusion of the close-kin data in the models. The probability of rebuilding to 20% of SSB(0) by 2035 is 92%. The management procedure appears to be operating as anticipated.

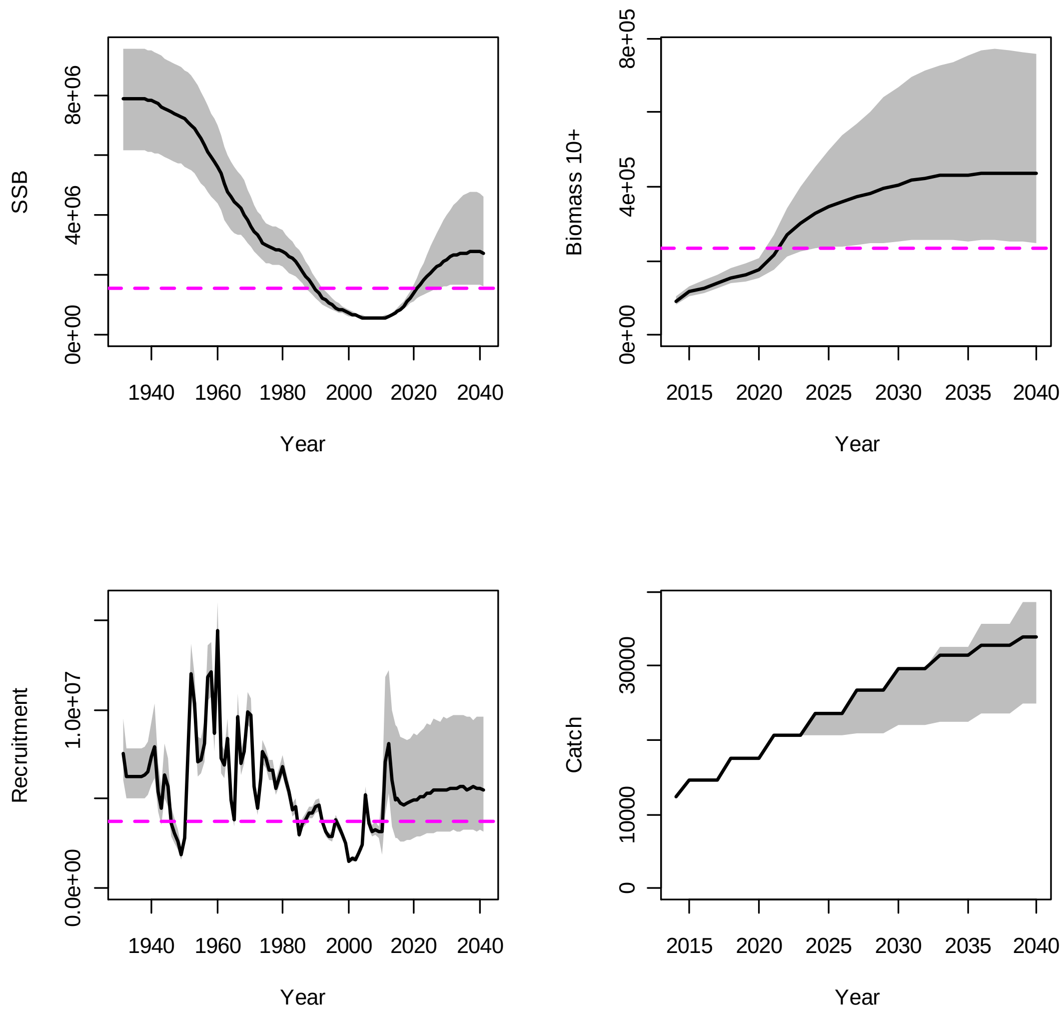


Figure 19 Projections, using the MP to set TACs. SSB (new definition from inclusion of close-kin), Biomass 10+ (definition used in 2011), Recruitment and TAC. In each plot the black line is the median of the simulations, grey area is 80th percentile, and the pink dashed line is 0.2*SSB(0) in the SSB plot, 0.2*median B10+(0) in the Biomass10+ plot, and 0.5*median R(0) in the recruitment plot.

5 Additional requests from the 2013 ESC

5.1 Recalculation of the over-dispersion factor in the mark-recapture data

At the highest aggregation level (across cohorts, release and recapture ages, tagger groups) the estimated variance in the standardised residuals was 0.77. For “ideally” weighted data (i.e. a correctly specified over-dispersion coefficient) this value would be equal to 1, so the indication is that the tag data are currently being under-weighted in the model. It was agreed at the 2013 ESC that, because the previous weighting was based on similar quantitative analyses, then any change to the empirical over-dispersion coefficient (current set at 2.35) should be reflected in the values used in the OM. So, based on this the new value of the over-dispersion factor in the OM conditioning should be 0.77 * 2.35 = 1.82.

5.2 Potential over-dispersion in the close-kin data with the updated OM

This issue is dealt with in section 3 in the close-kin sub-section. Detailed posterior predictive analyses, and standardised residual bootstrap analyses as detailed in CCSBT-OMMP/1307/5, both show no convincing evidence for over-dispersion in the close-kin data, conditional on the current OM structure.

5.3 Comparison of OM and stand-alone close-kin assessment for SSB and total mortality

Figure 20 shows the comparison of the estimated biomass of animals age 10 and above for the OM with the close-kin data included and for the stand-alone close-kin assessment. Figure 21 shows the total mortality estimates (Z) for ages 8 and above for the OM with the close-kin data included and for the stand-alone close-kin assessment. In terms of biomass of animals age 10 and over (old SSB definition) the stand-alone close-kin estimates are noticeable higher earlier on (2002-2006) although still there is over-lap of the 95% CIs in both cases; by 2008-2010 the estimates are clearly close with around a 25-40% difference (with the stand-alone close-kin estimates the highest). In terms of total mortality, the stand-alone close-kin data are consistently higher than those from the OM (and, by definition, time-invariant) with no clear over-lap of the confidence intervals over time. The Indonesian age data are known to favour higher mortality rates at these ages, and are the dominant source of information on survival in the stand-alone assessment; in the OM (earlier in this paper) this preference is tempered by both the close-kin and the mark-recapture data and so lower M_{10} values (and, subsequently, lower total mortality) are sampled in the OM. So, while the differences are clear, the story itself is consistent: both can explain the close-kin data reasonably well but the standalone assessment does this by higher abundance and lower survival; the OM, constrained by much more data than the stand-alone assessment, explains the data by lower abundance but higher survival.

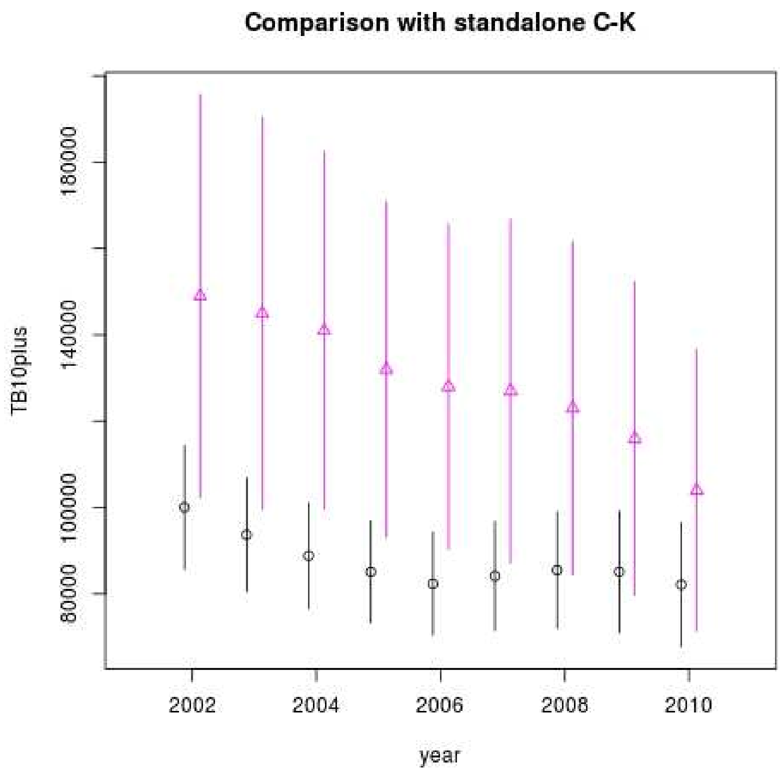


Figure 20 Comparison of biomass of animals age 10 and over for the stand-alone close-kin (magenta, median and 95%iles) and the OM (black, median and 95%iles).

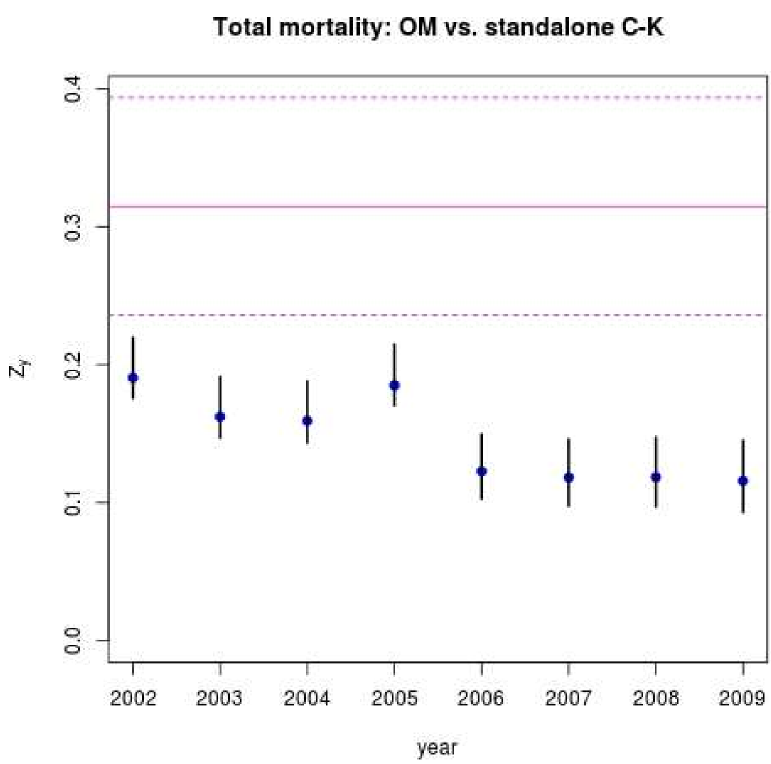


Figure 21 Comparison of total mortality age 8 and over for the stand-alone close-kin (magenta, median and 95%iles) and the OM (black, median and 95%iles).

6 Summary

The SBT OM has been reconditioned with the data supplied through the 2014 data exchange and the integration of the close-kin data into the operating models. A summary of initial examination of results and diagnostics is given below along with the issues they raise for consideration at the OMMP Working Group meeting:

- The updated models fit the size and age data well. Recent changes in the Indonesian age frequency need to be discussed and additional information gathered for the ESC to make more informed assessment of whether the higher than usual abundance of smaller size classes represent mature fish caught on the spawning ground, or immature fish caught to the further to the south, as has been observed previously.
- A bubble plot based analysis of the raw CPUE-at-age data to examine changes in catchability as requested by the ESC, suggests that catchability increase in the upq2008 robustness test should be re-estimated to determine whether it still warrants inclusion in future MP evaluations.
- The fits to the LL1 CPUE data and aerial survey data indicate that the CPUE data fit reasonably well. The aerial survey data fit less well; the 2014 point being the highest in the series and a few earlier points (1999, 2009 & 2012) being low and also outside the estimated Confidence Interval (CI). The 2014 point may qualify for exceptional circumstances, however, the robustness test conducted in the MP testing in 2011 encompassed such a value. Given the point is high, and therefore positive, the appropriate action may be to wait for additional observations of these cohorts from the AS and longline fisheries to provide more precise estimates of their abundance.
- The fits to the tagging data are similar to previous re-conditioning, and fit well.
- The fits to the close-kin data are reasonable. There are some data points that the model does not fit well (e.g. 2004 year class in 2008 and 2009), but the trends are captured. The posterior predictive analysis indicates that the model fits the data well and that here is likely to be little, if any over-dispersion.
- Comparison of the OM estimates of SSB and mortality with the close-kin independent assessment indicate some differences, which can be explained by use of the available data, and constraints in the OM from additional data sources not used in the independent assessment.
- The over-dispersion factor for the tagging data was recalculated and the lower result suggests that consideration should be given to revising this in the update of the OM.
- Parameter estimates and sampling of the grid are similar to the preliminary work done at OMMP4 in 2013.
- Consistent with results from recent years, the stock status is low but appears to be improving. The recruitment time-series show high recruitments in the recent years, which have a significant impact on projected catches. The current level of uncertainty in the strength of these recent year classes should decrease as they become fully selected to the longline fisheries.
- The projections code has also been updated, and results of preliminary catch and biomass projections using the reference set and the CCSBT MP indicate that the rebuilding objectives for the MP will be met.

References

Anon 2013. Report of the extended scientific committee meeting.

Davis, T., Jessica Farley and John Gunn (2001) Size and Age at 50% Maturity in SBT: An integrated view from published information and new data from the spawning ground. CCSBT/0108/16

Gelman et al. (1994) Bayesian data analysis. Chapman and Hall, London, UK.

Hillary et al. (2011) Reconditioning of the southern bluefin tuna operating model: exploratory data analysis, fitting performance, and current stock status. CCSBT-ESC/1107/11

Hillary et al. (2012) Initial exploration of options for inclusion of the close-kin data into the SBT operating model. CCSBT-ESC/1208/21

Hillary et al. (2013) Data weighting, grid configurations and further exploration of the integration of the close-kin data into the SBT operating model. CCSBT-OMMP/1307/5

CONTACT US

t 1300 363 400
+61 3 9545 2176
e enquiries@csiro.au
w 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 / Wealth from Oceans Flagship
Ann Preece, Rich Hillary, Campbell Davies
t +61 3 6232 5222
e ann.preece@csiro.au
w www.csiro.au/wfo