CCSBT-ESC/1107/11

Australian Government
Australian Bureau of Agricultural and Resource Economics and Sciences

National Research FLAGSHIPS
Wealth from Oceans
CSIRO

Reconditioning of the southern bluefin tuna operating model: exploratory data analysis, fitting performance, and current stock status

R. M. Hillary

A. Preece

C. Davies

B. Barnes

G. Begg

Prepared for the CCSBT Extended Scientific Committee for the 16th Meeting of the Scientific Committee 19-28 July 2011

Bali, Indonesia

Table of Contents

Abstract

The Commission for the Conservation of Southern Bluefin Tuna (CCSBT) Operating Model (OM) forms the basis of the management procedure (MP) work and in this paper we update the OM with the most recent data (i.e. 2010 catch, CPUE, length and age data; 2011 scientific aerial survey data). Trends in recent nominal catch per unit effort (CPUE) data-at-age suggest that the most recent rises in standardised CPUE may not be entirely attributable to increases in SBT abundance. There has been little change in the depletion level in the spawning stock biomass (SSB), with a median (and 90% CI) of 0.05 (0.03-0.07), but recent recruitments (2005-2011) are all estimated to be above the stock-recruit curve and well above the low recruitments estimated in the late 1990s/early 2000s. The higher levels of steepness in the grid (and by correlation M_0 – natural mortality) are being sampled more frequently, relative to the last conditioning of the OM. As with 2009 analyses, length and age frequency data are the dominant influences on the preferred level of steepness. However, there remain contradictory signals for steepness among the fitted data sets and, as such, the information is still equivocal and should be considered with caution. The updated estimates of recent recruitments, coupled with the tendency for higher estimates of steepness, result in a generally more optimistic assessment of the productivity of the stock and levels of recent (last 7 years) recruitments than the 2009 conditioning of the OM.

Introduction

The Commission for the Conservation of Southern Bluefin Tuna (CCSBT) Operating Model (henceforth, OM) is updated to include the most recent catch composition, catch per unit effort (CPUE) and scientific aerial survey data and uses the updated estimates of SBT growth. Given the recent (2-3 years) strong increases in Japanese longline CPUE data – especially in specific areas – we examine the area-specific nominal CPUE-at-age to explore potential drivers of these increases. The OM will interpret these trends (in the standardised CPUE) as increases in the exploitable biomass of the Japanese longline fishery, with similar implications for estimated trends in both historical recruitment and future spawning stock biomass (SSB). Hence, it is considered important to distinguish, to the extent possible, between alternative hypotheses for the observed increases.

The usual summaries are given for the fits of the updated OM (catch composition, tagging data, and the CPUE and scientific aerial survey abundance indices) and a posterior predictive analysis is undertaken for the CPUE and aerial survey data, to see how well they explain the data, in addition to the model fit. Parameter estimates and grid (**basehupsqrt**) samples for the updated OM are summarised and estimated trends in historical recruitment and SSB are provided. Constant catch projections for both the current (9,449 t) and zero TAC levels are detailed.

Reconditioning of the OM

This section is organised as follows: (i) area-specific nominal CPUE-at-age for the Japanese longline fleet is analysed to explore plausible explanations for the most recent CPUE observations, (ii) the fitting and general estimation performance of the OM is detailed; and (iii) parameter estimates, grid samples and historic trends in recruitment and SSB are summarised. Table 1 provides the current specification of the grid for this reconditioning of the OM.

Exploratory analysis of area-disaggregated nominal CPUE-at-age

From 2008 to 2010 we have observed large increases in the nominal and standardised CPUE data used in the OM. In terms of trends in the nominal data, these increases have been area-specific in their relative magnitude, being much larger in areas 6 and 7 than in areas 8 and 9 (Chambers and Boero, 2011). The OM will interpret these trends as increases in the abundance of the population vulnerable to the longline (LL1) fleet, which has implications for estimates of the abundance of corresponding recruitments and future SSB trends.

There are a number of potential explanations for the increasing trend in CPUE – increases in recent recruitment, lower levels of total mortality (fishing and/or natural), or increased catchability. However, as catchability in the OM is effectively fixed over the time-period in question, if the OM is to fit to the CPUE trend (which it does, as we shall see later) this increase will be attributed to increased recruitment, via higher recruitment residuals and/or increased steepness. From previous experience, such changes to recent recruitment will have a strong impact on the predicted future SSB trends and, therefore the performance of the candidate management procedures (MPs) in the simulations. Hence, it is important to understand the extent to which these recent CPUE observations result from increases in abundance or changes in fishing practices, so that the plausible alternative explanations can be included in MP testing.

The nominal area-specific CPUE-at-age is a useful starting point for exploring potential explanations for the recent increases in standardised CPUE. Given the data $I(r,y,a)$ (where r is the area, y the year and a the age) we use the cohort-transformed data, $I(r,y-a,a)$, in this exploratory analyses as it makes it easier to identify the cohort specific patterns related to year class strength.

Figure 1 shows a bubble plot of the cohort-transformed data for areas 4-9, ages 4-8 and for the cohorts from 1980-2006 (with 2006 being the last observed cohort in the CPUE over this age range and corresponding to CPUE up to and including 2010). With plots of this type, weak/strong year classes appear as (relatively) smaller/larger bubbles as one looks up the vertical (age) axis for a given cohort. For example the weaker year-classes of the late 1990s/early 2000s appear as much smaller CPUE levels across all areas. Lower levels of mortality on a cohort as it ages (assuming similar selectivity for that time period) appear as a more gradual decrease in the magnitude of the bubbles in the vertical direction.

What is more evident, are the strong diagonal trends in high levels of CPUE in multiple cohorts at the same time (e.g. area 7). These are year effects that, in this spatial context, most likely have two explanations: (i) annual changes in the spatial distribution of the population, i.e. SBT move from one area to another, and (ii) an increase in catchability in that specific area that is not related to an increase in abundance (from recruitment, lower mortality or immigration) but possibly a change in fishing practices. We cannot be sure which of these alternatives is the cause. However, there are some signs that are more suggestive of one mechanism above the other.

If it is the result of a change in the spatial distribution of the stock then, given we assume a closed population, an increase in one area should accompany a decrease in other area(s). There are some signs of this in some areas (i.e. areas 4 and 5 for the early 1990s cohorts, areas 8 and 9 for the late 1980s cohorts) but not for the very large increases in the most recent years in areas 4, 6, 7 and 9, that form the basis for the increase in overall CPUE from 2008-2010. The increases happen at the same time and across the areas that contain the overwhelming share of SBT abundance at these ages. Given this, we consider it unlikely that the significant increases in CPUE in these areas is the result of SBT moving from one area to another as all the major areas are increasing. This suggests that an increase in catchability in these areas and for these years (2008-2010) is the most likely explanation for the substantial increase in CPUE, rather than an increase in actual abundance.

Subtle changes in selectivity and fishing pressure across areas mean it is unlikely to be informative to attempt to estimate a catchability effect from these trends in nominal CPUE. The potential for changes in catchability have been handled before as robustness trials (e.g. *upq*/*downq*) and were seen to have a significant effect on MP performance – especially an increase in catchability designed to explore the implications of the introduction of the new management measures for the LL1 fleet. The primary candidates for an explanatory mechanism are changes in fishing pattern/efficiency, an increased density of the population in a given area (spatial contraction), or some other unknown environmental factor changing the spatial distribution and resulting in higher CPUE. These trends are apparent in the nominal data and are of sufficient size to be the major determinant of the increase in CPUE seen in the standardised CPUE series. As effects of this nature are not included as factors in the current CPUE standardisation, they have not been 'standardised out' in the series used in the OM conditioning.

We know from previous experience with the *upq* robustness trial that increases in catchability that occur in the period of OM conditioning lead to positively biased abundance estimates. As we shall see below, the increase in CPUE contributes to increased estimates of recent recruitment and steepness, all of which result in a more optimistic outlook in the OM projections and candidate MP performance. As such, it will be important that: a) the influence of these recent CPUE observations on the assessment of the status of the stock and estimated productivity is given due weight; and, b) it is considered carefully and addressed as part of the robustness testing of MPs.

Fits to the data

The OM is fitted to the fishery-specific catch composition, Japanese longline CPUE, scientific aerial survey, and 1990s mark-recapture data.

Size Composition

Figures 2 to 5 show the fits to the size composition data from the longline fisheries (*LL1*, *LL2*, *LL3* and *LL4*, respectively). For *LL1* the lack of fit to the earliest data persists; the OM cannot fit the spikes of smaller fish seen in 2006 and 2007, presumably as they do not exhibit characteristics of being a strong cohort; the most recent data 2008-2010 are fitted well. For the *LL2* fishery there are notable problems with the early data but the more recent data are fitted fairly well. For the *LL3* fishery the OM seems to underestimate the proportion of smaller fish in the catch in 2005 and 2006 but elsewhere the data are fitted fairly well. For the *LL4* fishery in the mid-1950s the OM repeatedly underestimates the proportion of smaller fish in the size composition, with the rest of the data fitted reasonably well, apart from certain isolated spikes.

Age structure of surface fishery and spawning ground catches

The spawning ground fishery, although uneven in certain years, is fitted fairly well by the OM but the proportion of younger fish in the catch in 2008 is significantly over-estimated (Figure 6). The age composition in the surface fishery is fitted well, with only minor mismatches in certain years, and the most recent data are fitted very well (Figure 7).

LL1 CPUE and Scientific Aerial Survey

Figure 8 shows the fits (across the uncertainty in the grid) to the CPUE and scientific aerial survey data. The CPUE data are fitted well in general, although the stronger changes (both up and down) in the series are not. As before, the scientific aerial survey data are not fitted very well in the early years (1993-2000); while the more recent (2005-2011) data are not fitted much better the general increasing trend is replicated, but the strong increase from 2010 to 2011 could not be replicated by the OM.

A posterior predictive analysis (Gelman *et al*., 1995) was undertaken in addition to examination of OM fit to the input data series. Though the samples from the grid are not strictly samples from a Bayesian posterior distribution, we effectively treat them in this way, and a posterior predictive analysis allows us to see how well the OM is not just fitting but explaining the data, in terms of the variance structure of the data and that which we assume in the likelihood. One simulates data, given the posterior sample and the likelihood, and generates two sets of residuals: one for the simulated data and one for the real data. A suitable summary statistic (in this case the absolute median deviation) is calculated and one then computes the probability that the predicted residuals are more "extreme" (i.e. greater than) the observed residuals. This probability is known as a Bayesian *p*-value (Meng, 1994) and values of around 0.5 are the ideal; values greater/less than 0.5 suggest that the OM predictions are more/less variable than the actual data; values outside of the 90% credible interval (< 0.05 or > 0.95) suggest something fundamentally wrong with the assumed model and likelihood. For the CPUE data the *p*-value was 0.91 – while the OM fits the data fairly well the assumed CV of 0.2 is higher than the residual CV of around 0.13, so the predicted CPUE data are more variable than the actual CPUE data. For the scientific aerial survey data the *p*-value was 0.65 suggesting that the OM is doing a fairly good job explaining the data, and that the assumed likelihood and process error estimates are acceptable.

1990s Tagging

The tagging data are disaggregated by cohort, age-at-release and tagger group in the OM likelihood (Eveson 2009). Figure 9 shows (for the most likely grid sample) the observed and predicted total recaptures-at-age, for each year of release. Tags released in 1996 and 1997 display by far the highest number of recaptures and are fitted well. The other years (1992-1995), although with less recaptures, are fitted fairly well with no recurring over/under estimation problems across years.

For a more detailed view of how the OM is fitting to the data at the level specified in the likelihood Figure 10 shows a plot of the fits to the tagging data aggregated at the level of cohort, age of release and tagger group. In Figure 10 the red level indicates recaptures of 100 fish or greater, orange between 20 and 100, yellow 5-20 and pale yellow less than 5. As one would expect, there is more variation in predicted versus observed recapture numbers at the tagger group/cohort-at-release/age-at-release level than at the fully aggregated level seen in Figure 9. There are no apparent systematic trends across tagger groups (especially the major ones). This suggests that the tag shedding rates are reasonably well characterised, as well as the data being consistent across these groups in relation to abundance and mortality rates. All the major recapture events are fitted fairly well. In general one expects the larger recapture events to be fitted better and that lack of fit to the minor recapture events cannot be viewed as closely, given the much lower precision in the information content of these events. Overall, the tagging data appear to be an internally consistent data set that the model is able to fit fairly well.

Parameter estimates and stock status

A summary of the sampling of the various parameter levels of the current grid (**basehupsqrt**) is provided in Figure 11. In comparison to the previous OM reference grid (**base5hsqrt**) higher levels of both steepness (h) and M_0 are now being sampled. There is a fundamental, yet complicated, inter-dependence between steepness, initial abundance (B_0), and natural mortality that cannot be seen from the level plots. Figure 12 shows a scatter plot of the unexploited SSB (B_0), steepness (h), juvenile (M_4) and oldest (M_{30}) natural mortality – the key point is that apart from steepness (which is part of the grid) all the other parameters are estimated. As is often the case, there is negative correlation between steepness and unexploited biomass; steepness and M_4 are positively correlated; unexploited biomass and M_{30} are strongly positively correlated. Figures 11 and 12 demonstrate that when information on these key parameters changes (such as steepness this year) it has an effect on all these other key parameters, to differing degrees.

Figure 13 shows the likelihood profiles for steepness, for each of the fitted data sets and for the objective function as a whole, conditional on the three levels of M_{10} (0.07, 0.1, 0.14). The information on steepness, in terms of level and strength, for each of the data sets varies greatly, and often with clear sensitivity on the level of M_{10}. As in previous analyses (Anon, 2009; Eveson and Davies 2009), the information is also equivocal across data sets for which clear information on steepness is observed.

For the length data: the *LL1* catch composition prefers lower levels of steepness across all levels of M_{10}; the *LL2* catch composition appears to have no preferential levels; the *LL3* catch composition prefers higher levels of steepness, especially for lower levels of M_{10}; the *LL4* catch composition prefers higher levels of steepness for higher levels of M_{10}, but lower levels for lower values of M_{10}. For the age data: the spawning ground fishery, in general, prefers higher levels of steepness but more so for lower values of M_{10}; the surface fishery shows weak preference for higher steepness at medium to high levels of M_{10} and no real preference for low M_{10}. The CPUE data shows marginal preference for low to mid range levels of steepness for high M_{10} but a general preference for mid to high level steepness levels for low to mid levels of M_{10} (with a clear three tier structure probably driven by the age range and omega elements of the grid). The tagging data are fairly non-informative on steepness. The aerial survey data appears to prefer mid to higher levels of steepness, and across all M_{10} scenarios.

We note that the preferred level of steepness and sampled range (0.42-0.85; 0.55-0.9) has increased in the current grid (**basehupsqrt**) relative to the previous grid (**base5hsqrt**) with a mean of 0.77 relative to 0.66, but it has also become more precise with a CV of 0.1 relative to 0.14.

The question is what has changed, across all the data sets that appear to inform on steepness, from the previous to the current OM? Looking at the likelihood profiles for the previous grid there is little change in the profiles for the size composition data (*LL1* to *LL4*). The spawning ground age data seem to show a slightly higher preference for higher levels of steepness, though not by much. The surface fishery age data has gone from a weak preference for low steepness to a weak preference for higher steepness. The CPUE data show a change in preference from lower to mid levels of steepness (0.55-0.65) to a general preference for mid to higher levels of steepness (0.75-0.9). The tagging data seem less informative even than the previous grid (even with no new data as this is data from the 1990s tagging program). The aerial survey data show a stronger preference for higher levels of steepness this time, relative to the previous grid (Anon., 2009).

The size and age composition data of the various fisheries (apart from *LL2*) are still the dominant data sets within the objective function that contribute information on steepness and, as before, their preferences are all different and with some dependence on the M_{10} level. Both the surface fishery and the aerial survey data have moved to a preference for higher steepness levels (given the higher apparent level of younger fish in the Great Australian Bight) but their influence on the likelihood is not so strong. For example, the CPUE has shown a noticeable shift towards a preference for higher levels of steepness (and for the most often sampled levels of M_{10}) and has more influence in the likelihood than the aerial survey. This suggests that these small shifts in the surface fishery and aerial survey steepness preferences, in combination with the stronger influence of the CPUE and its shift towards higher steepness levels, is causing the shift in the estimated levels of steepness sampled in the updated grid.

Figure 14 shows a summary (median and 90% CI) for the historical SSB and recruitment estimates. Current estimates of median (and 90%ile) SSB depletion are 0.05 (0.03-0.07).

Current SSB depletion is still at around 5% of unexploited SSB (similar to the previous Om conditioning estimate) with a stable flat trend over the last few years. Recent recruitments (2005-2011) are all significantly higher than the low recruitments estimated during the 1999-2002 period; in particular the 2006 recruitment, estimated to be low in the previous OM conditioning, has been revised upwards substantially.

Figure 15 summarises the stock-recruit relationship (median stock-recruit relationship and median and 90% CIs for actual SSB and recruitment estimates). The most recent recruitments (at the lowest SSB levels) are all above the median stock-recruit curve.

Constant catch projections

In this section we explore constant catch projections for both the current (9,449 t) and zero TAC for comparative purposes. Figure 16 summarises the most recent and future SSB dynamics for these two constant catch scenarios, at the median, 30th and 5th percentiles. The reason for including the 30th percentile is that it provides an estimate of when the interim rebuilding level and MP tuning target (20% B_0) will be achieved. For the current catch level, the tuning target ($p(SSB > 0.2\ B_0) = 0.7$) is reached in 2026, and for the zero TAC case it is reached in 2021. The faster than previously projected recovery of the future SSB is clearly evident, driven by both the higher estimates of recruitment from 2005-2011 (causing the sudden rise at around 2015) and the increase in the frequency with which the higher levels of steepness in the grid are being sampled. This predicted increased rate of recovery has implications for the current tuning criteria and operational constraints for MP testing and selection.

Summary

The SBT OM has been updated to include the most recent catch, catch composition, CPUE and scientific aerial survey data. Overall, the OM fits reasonaly well to the data, with no trends in residuals that have not been observed in previous conditioning exercises. Information on steepness is still dominated by the length composition data, and the contradictory preferences among the fitted data sets observed in the previous conditioning exercise remain. Overall, higher levels of steepness (and by correlation M_0) are being sampled in the most recent grid (**basehupsqrt**). This is driven by a preference for higher steepness levels in the CPUE, surface catch composition, and scientific aerial survey data.

Current SSB is estimated to be at around 5% of unexploited biomass, consistent with the 2009 OM conditioning. However, recent recruitments (2005-2011) are estimated to be higher than previous conditioning and above the expected stock-recruit curve. These estimates are driven by both the recent increases in CPUE and the scientific aerial survey data.

In the case of the CPUE, there is some concern that these increases may be an artifact of changes in fleet dynamics. There are indications in the nominal CPUE-at-age data of a period of increased catchability, which may be driving the increase in the standardised CPUE trend and associated estimates of higher levels of recruitment and preference for higher levels of steepness.

Preliminary constant catch scenarios (both current and zero TAC) indicate a stock that is expected to recover faster than was predicted on the basis of the previous OM conditioning, given the higher recruitments and steepness estimates.Overall, while the estimated state of the spawning stock is similar to the previous conditioning, the updated OM is more optimistic with respect to the level of recruitment over the most recent 5 years and the rate of recovery of the stock under constant catch levels than the OM used in the last round of MP work in 2010. The predicted increases in the likely productivity and, hence, rate of recovery of the stock mean that the current tuning criteria and operational constraints for MP testing and selection may require revision.

Acknowledgements

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

References

Anonymous. (2009) Report of the Operating Model and Management Procedure Technical Meeting. Seattle, USA.

Chambers, M., and Boero, V. (2011) Examination of CPUE indices for southern bluefin tuna. CCSBT-ESC/1107/10, Indonesia, Bali

Eveson, P., and Davies, C.R. (2009) Re-conditioning of the CCSBT Operating Model: impacts of revised natural mortality schedule and interaction with steepness. Working Paper CCSBT-ESC/0909/40to the 14th Meeting of the Extended Scientific Committee, Busan, Korea, 5-11 September 2009,

Gelman, A., Carlin, B. J., Stern, H. S., and Rubin R. B. (1994). Bayesian Data Analysis, Chapman and Hall, London.

Meng, X. (1994). Posterior predictive *p*-values. Ann. Stat. **22**: 1142-1160.

Tables

Table 1: The specifications used for the updated (**basehupsqrt**) OM grid.

	Levels	Cumul N	Values	Prior	Weighting
Steepness (*h*)	5	5	0.55, 0.64, 0.73, 0.82, 0.9	uniform	likelihood
M_0	4	20	0.3, 0.35, 0.4, 0.45	uniform	likelihood
M_{10}	3	60	0.07, 0.1, 0.14	uniform	likelihood
Omega (ω)	1	60	1	NA	NA
CPUE	2	120	w.5, w.8	uniform	prior
***q* age-range**	2	240	4-18, 8-12	0.67, 0.33	prior
Sample size	1	240	SQRT	NA	NA

Figures

Figure 1: Nominal Japanese CPUE-at-age by area and cohort (not year). Bubble diameter is a measure of the size of the CPUE. Data shown are from months 4-9 only. Each panel is a specific area, with the *x*-axis denoting the particular cohort, and the *y*-axis is the age of the fish in that specific cohort.

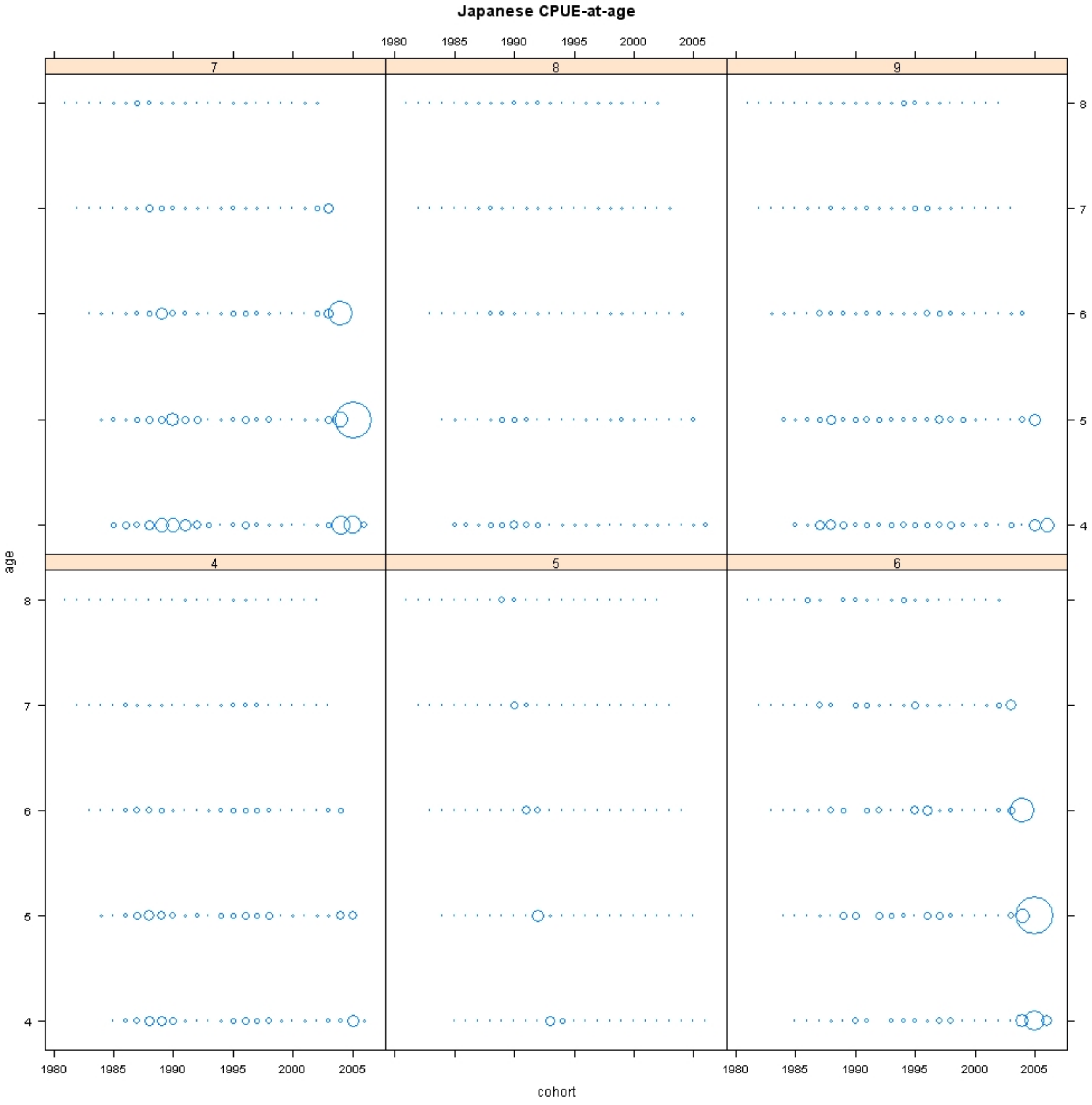


Figure 2: Observed (blue) and predicted (red, full line median dotted line spread of 90% CI) length composition for the *LL1* fishery.

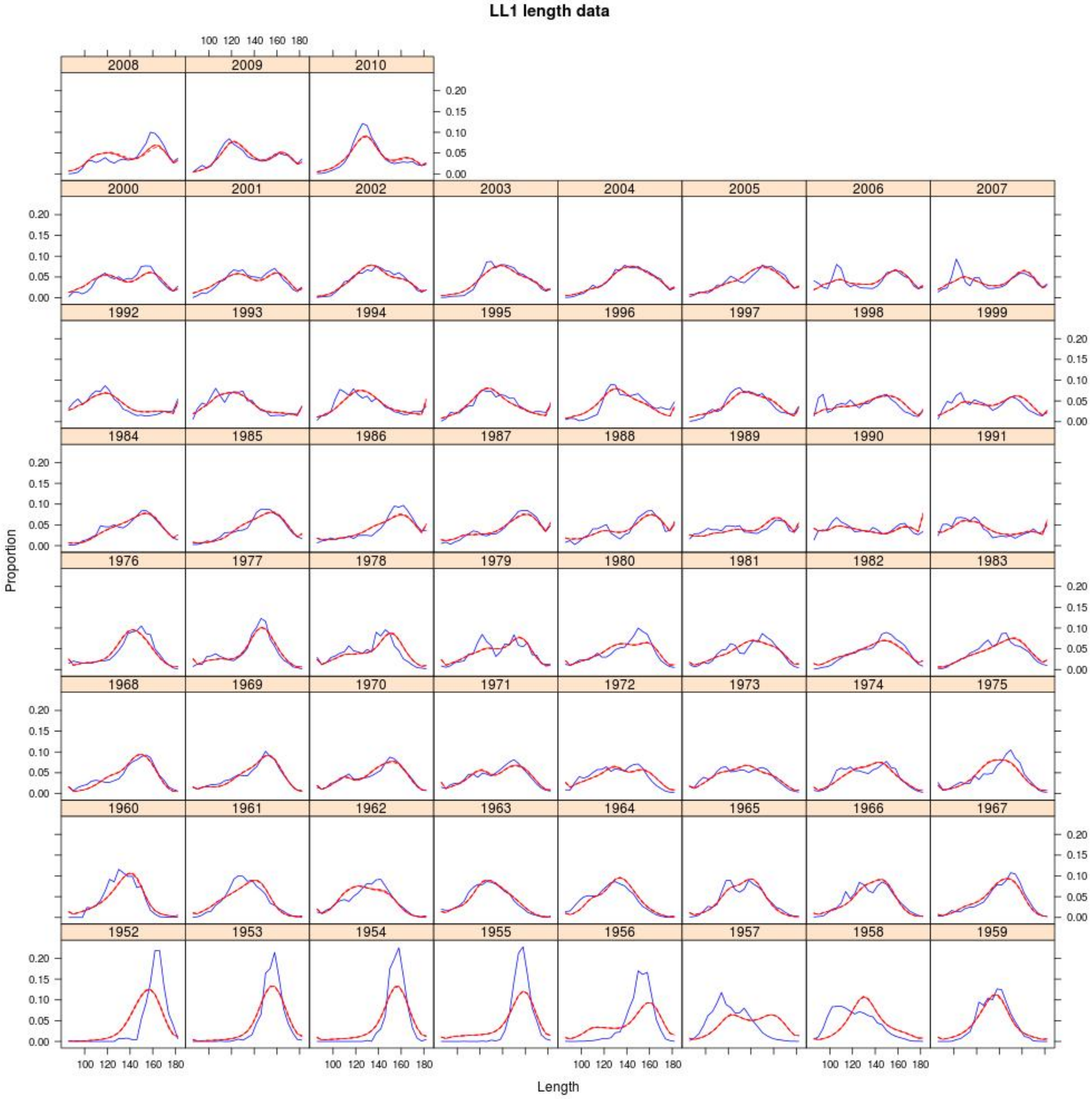


Figure 3: Observed (blue) and predicted (red, full line median dotted line spread of 90% CI) length composition for the *LL2* fishery.

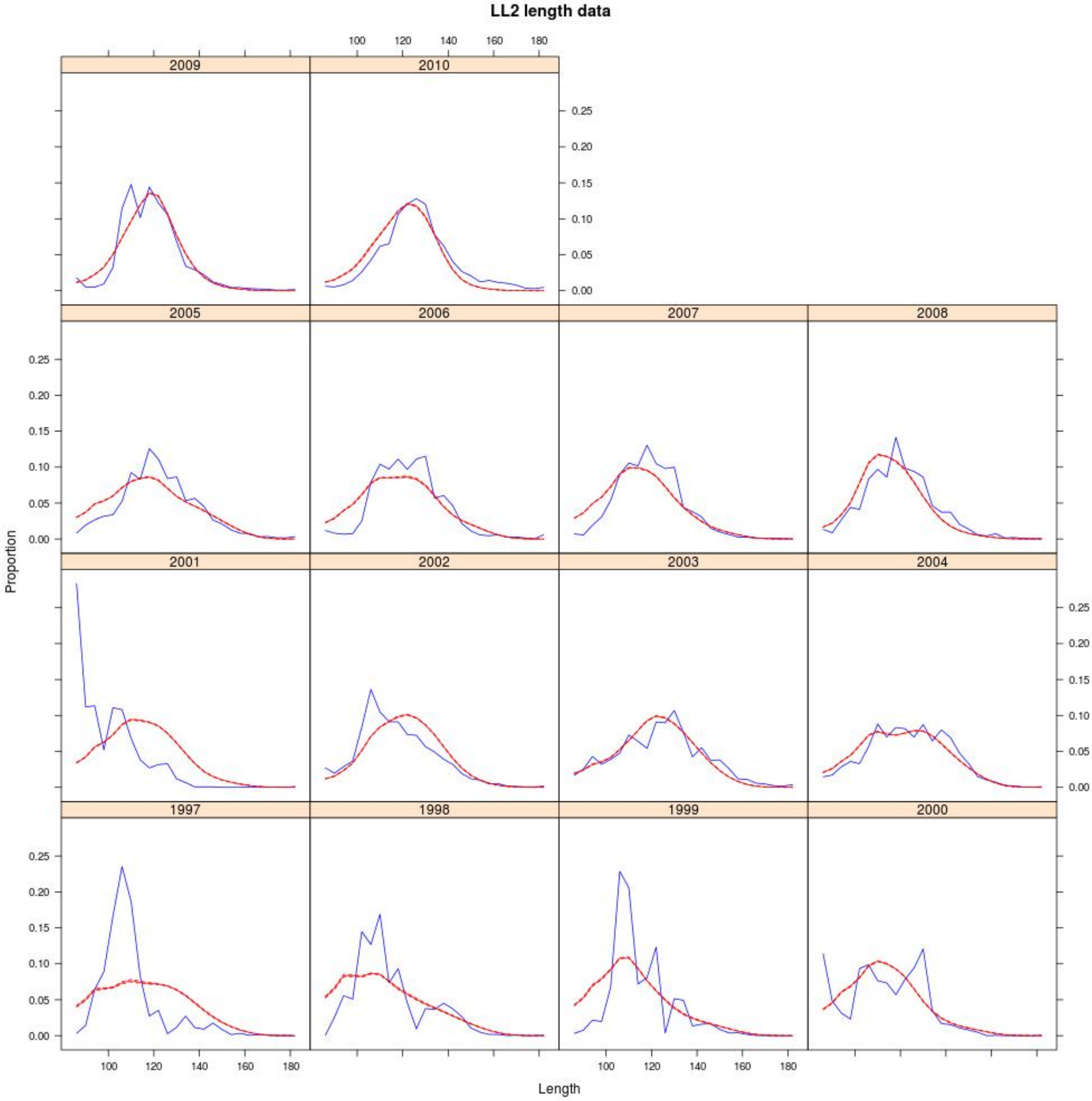


Figure 4: Observed (blue) and predicted (red, full line median dotted line spread of 90% CI) length composition for the *LL3* fishery.

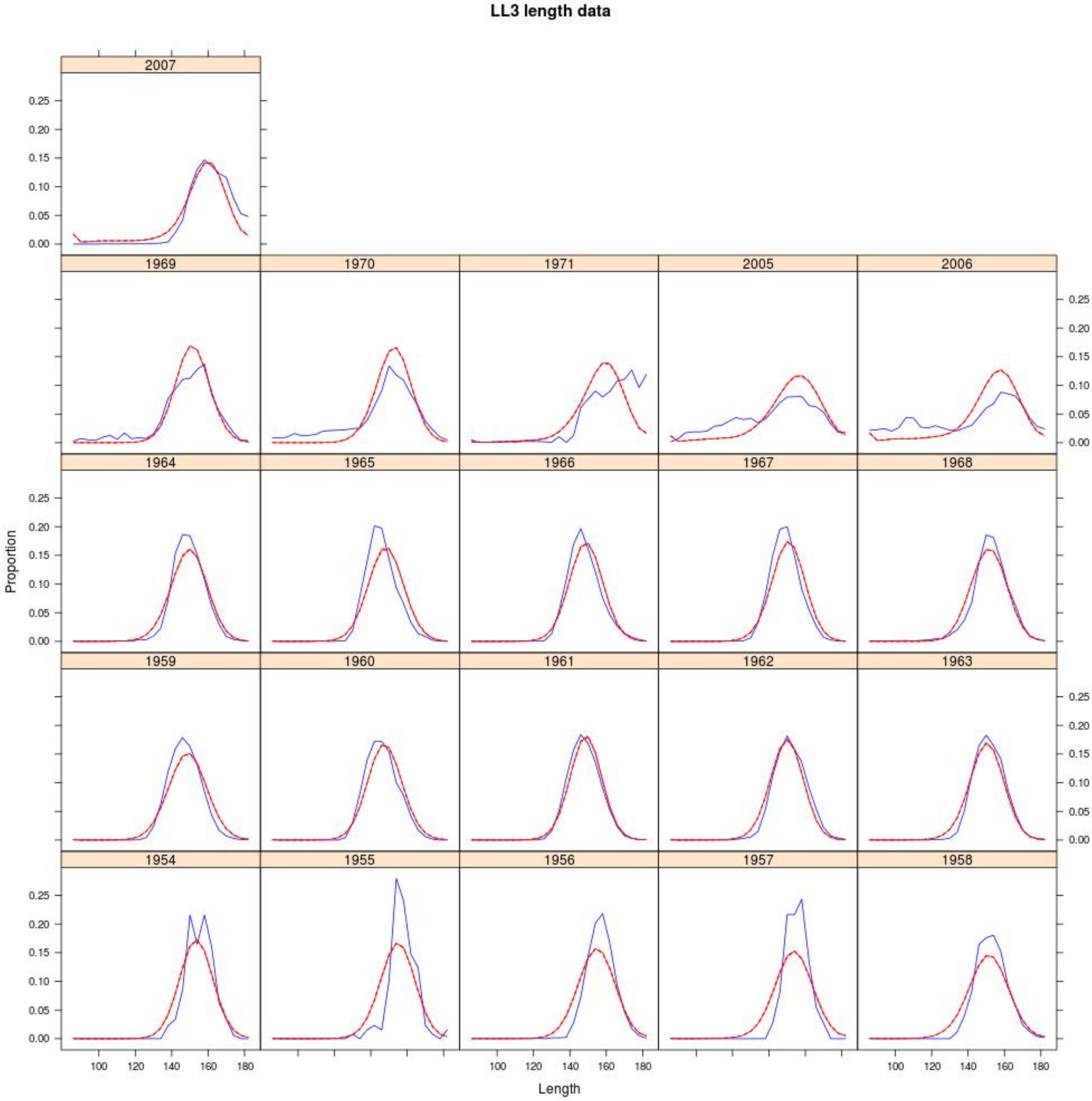


Figure 5: Observed (blue) and predicted (red, full line median dotted line spread of 90% CI) length composition for the *LL4* fishery.

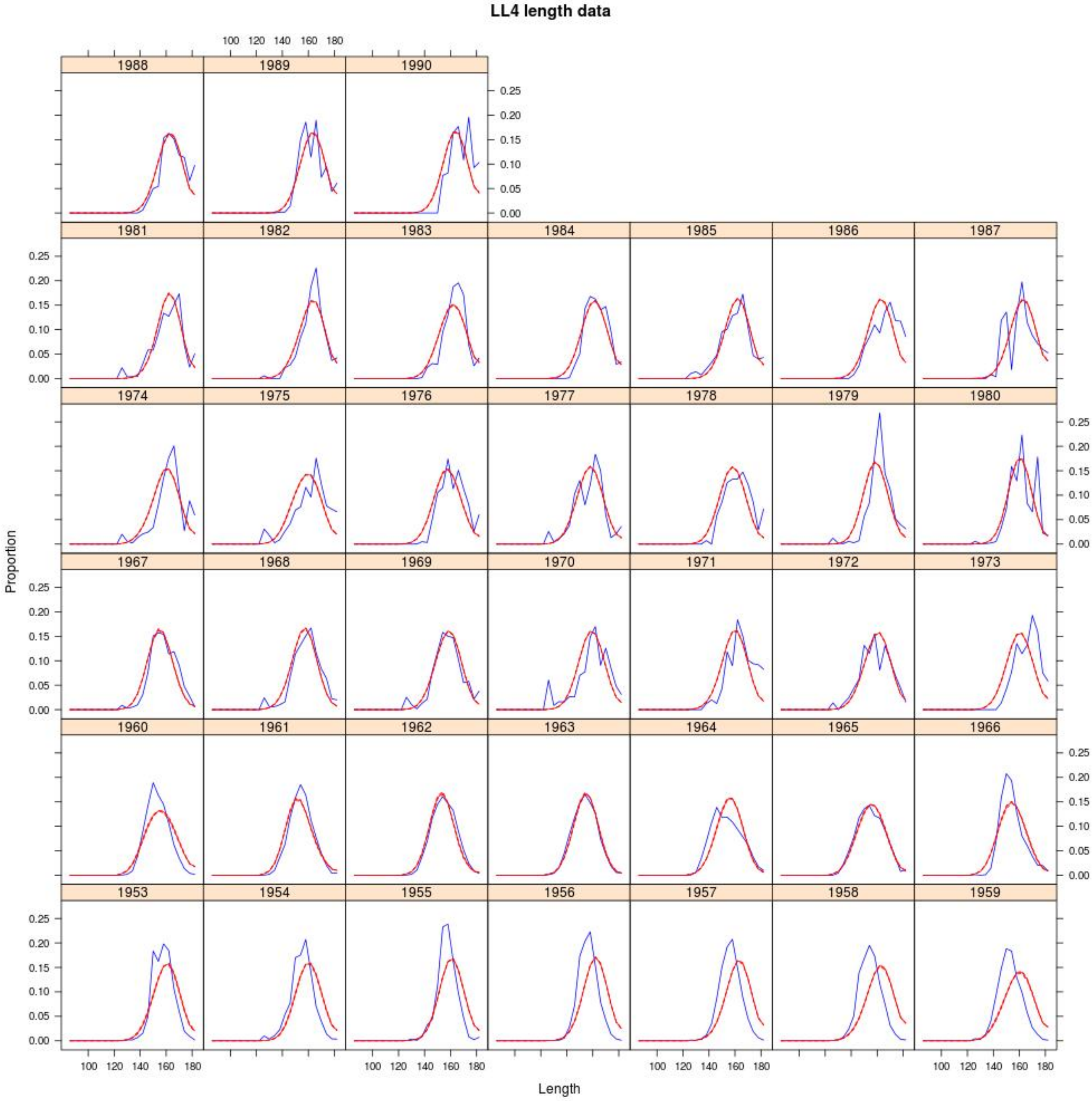


Figure 6: Observed (blue) and predicted (red, full line median dotted line spread of 90% CI) age composition for the Indonesian fishery.

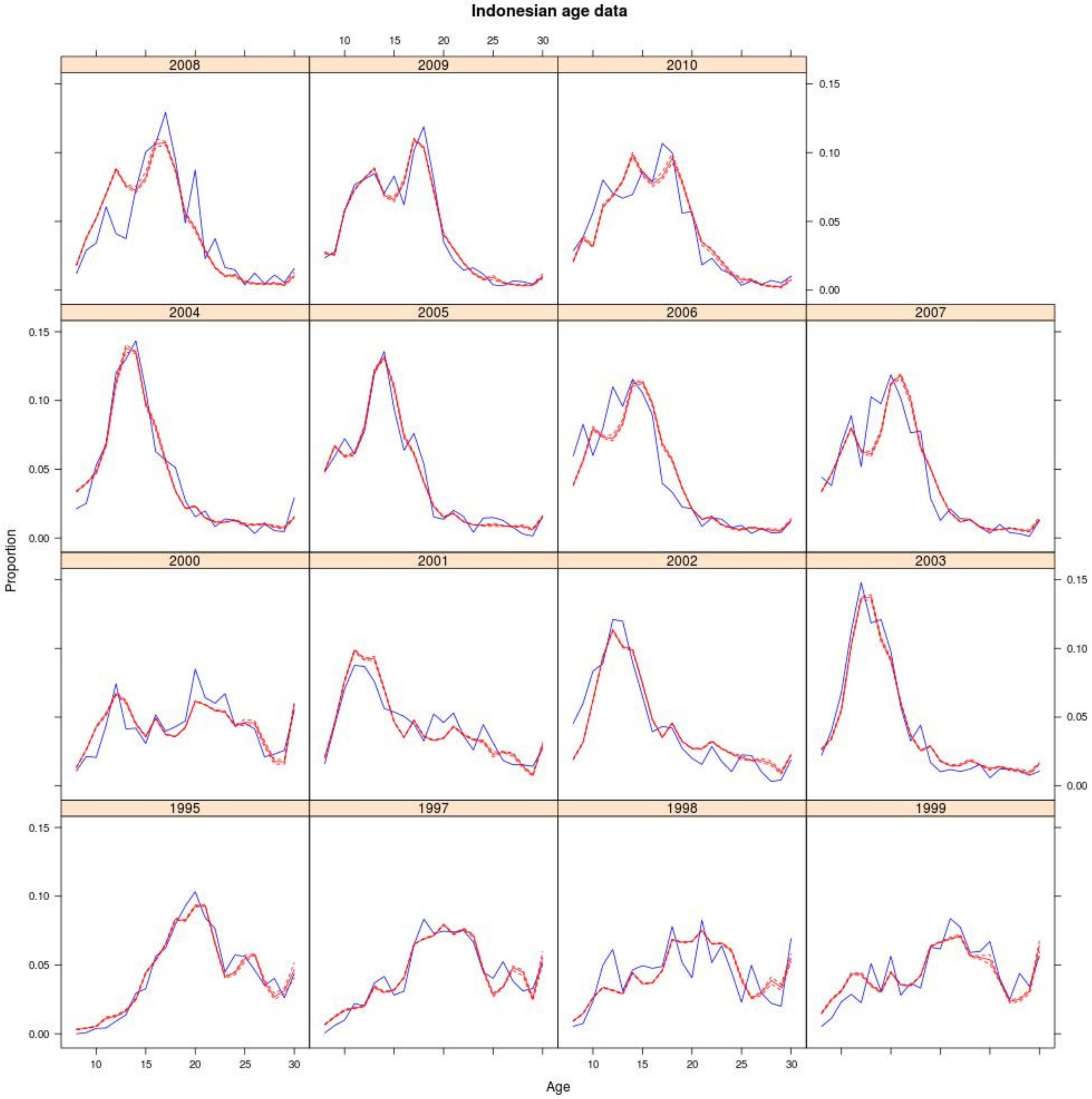


Figure 7: Observed (blue) and predicted (red, full line median dotted line spread of 90% CI) age composition for the surface fishery.

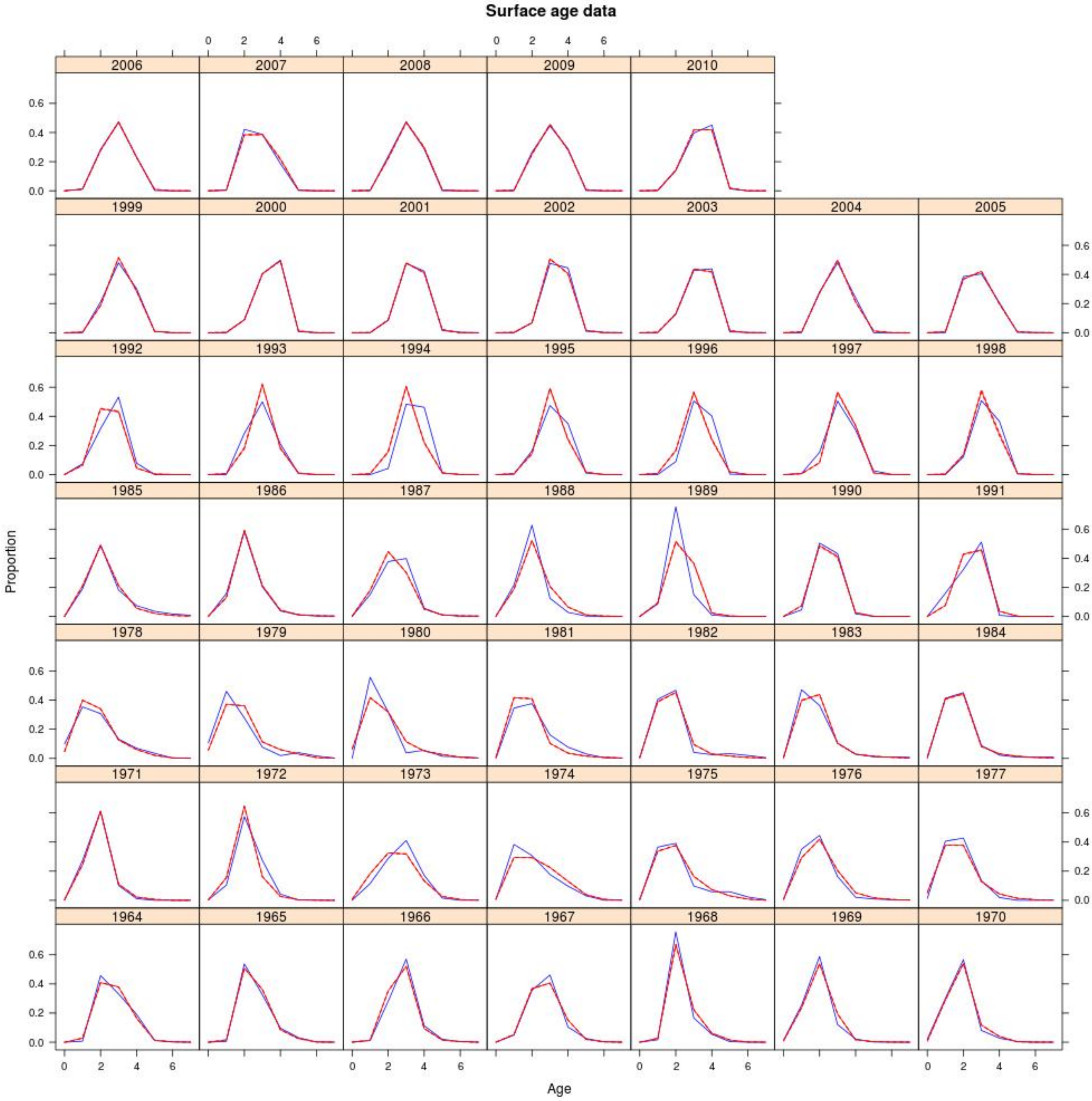


Figure 8: Fits (top) and posterior predictive summaries (bottom) for the CPUE (left) and scientific aerial survey (right) abundance indices. For the fits the circles are the observed data, with the full and dotted lines being the median and upper and lower 5 %iles, respectively.

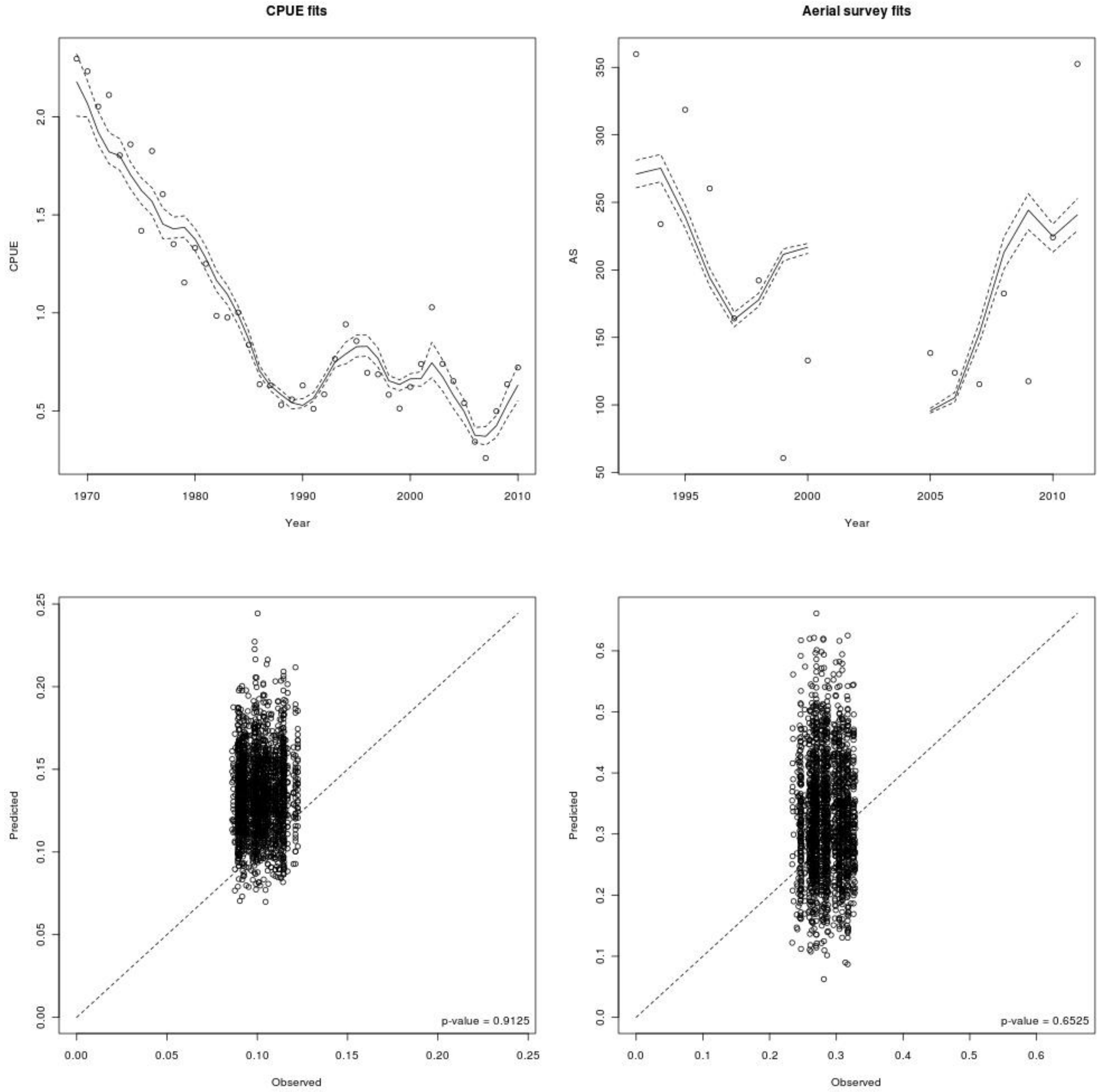


Figure 9: Fits to tag recaptures-at-age pooled into release years. Results are shown for the most likely grid sample.

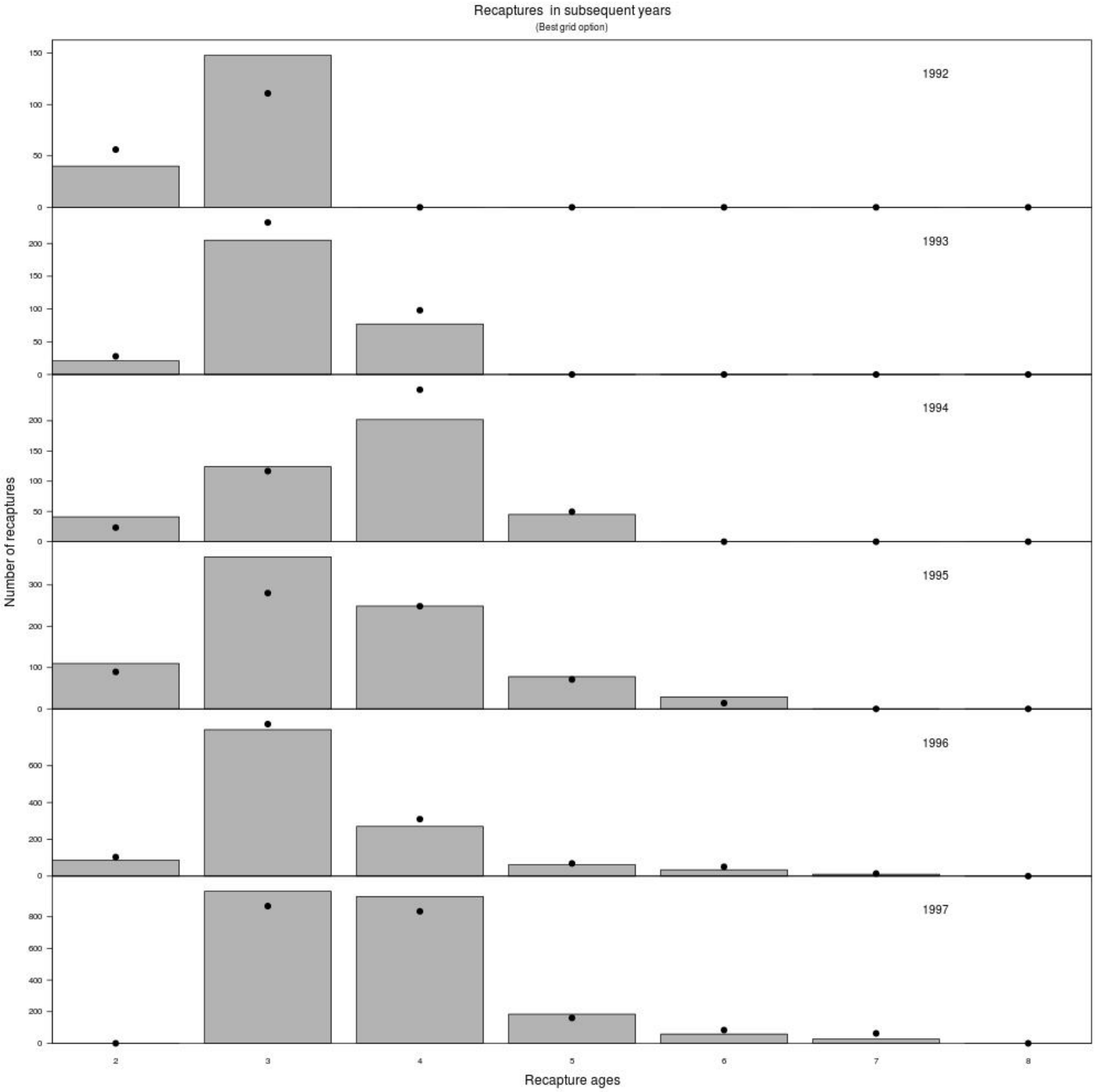


Figure 10: Fits to tag recaptures-at-age at the resolution assumed in the likelihood: disaggregated by tagger group (1-6), cohort tagged (1989-1994), and age-at-release (1-3). Results are shown for the most likely grid sample and the red level indicates recaptures of 100 fish or greater, orange between 20 and 100, yellow 5-20 and pale yellow less than 5.

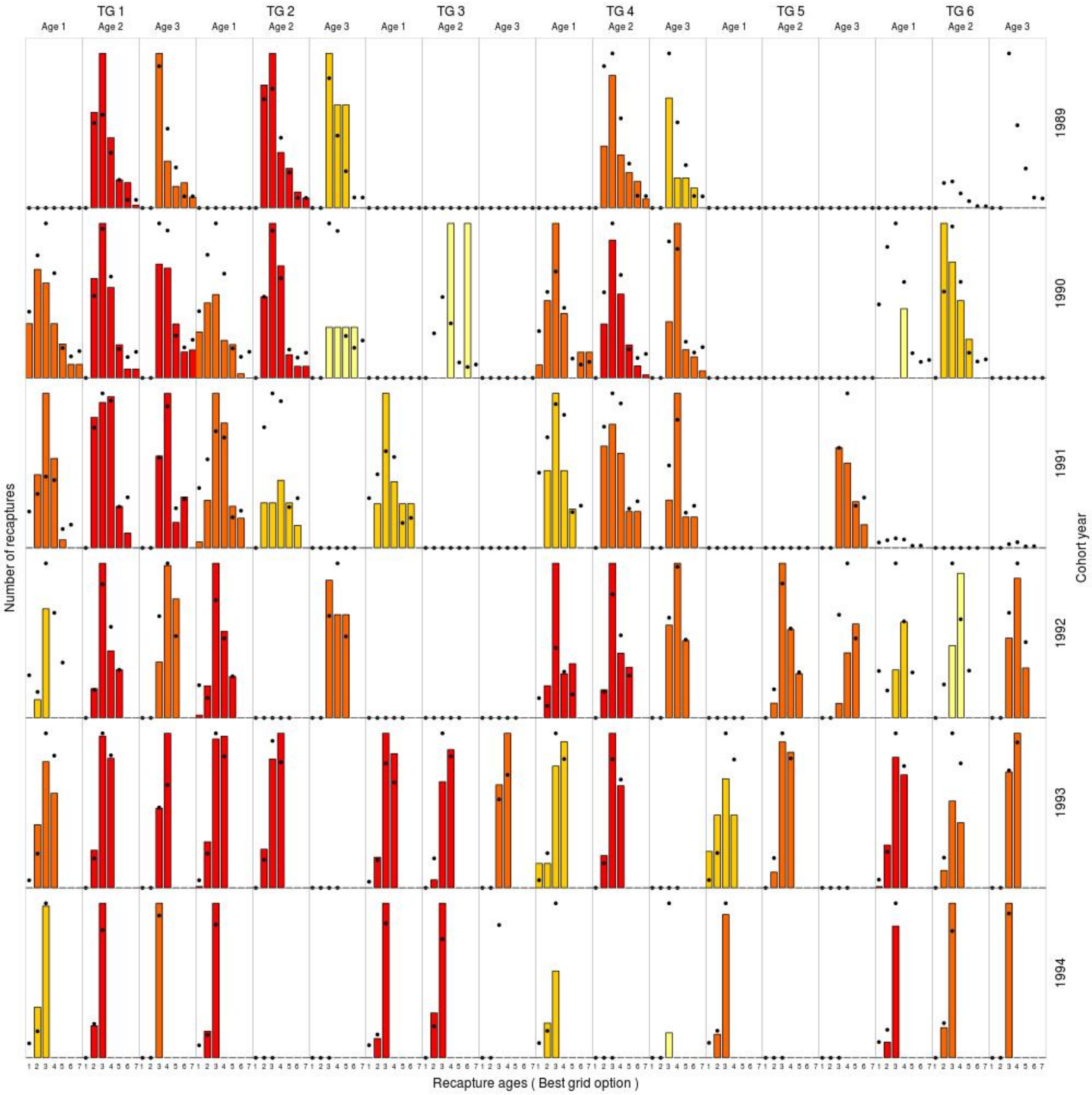


Figure 11: Summary of the levels sampled and inter-dependence of the grid elements for **basehupsqrt**.

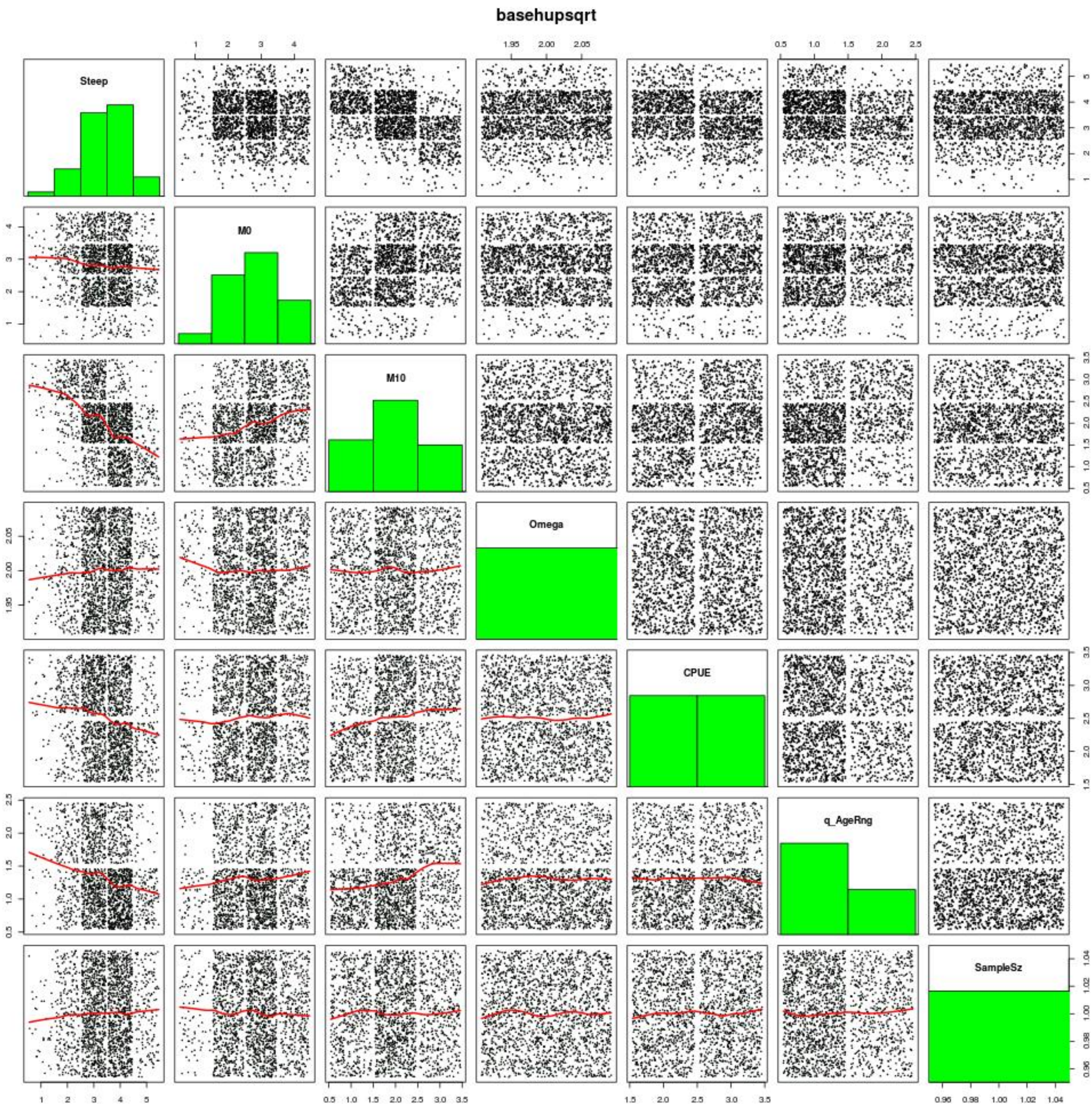


Figure 12: Scatter plot summary of the correlation between steepness (*h*), virgin biomass (B_0) and the two estimated natural mortality parameters (M_4 and M_{30}) for **basehupsqrt**.

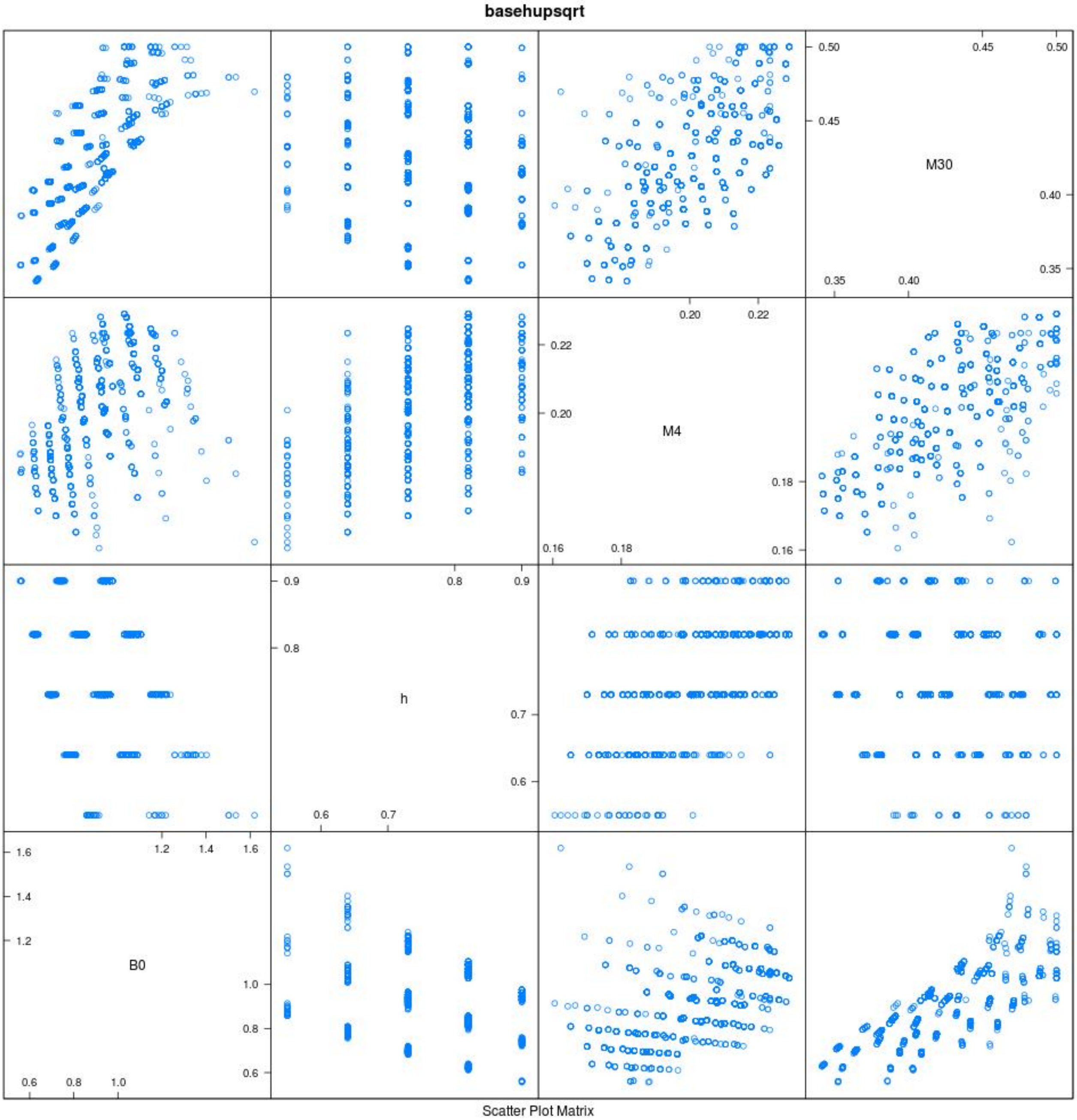


Figure 13: Likelihood profiles for steepness for each of the data sets and each sample from the **basehupsqrt** grid, and conditional on the three grid levels of M_{10}.

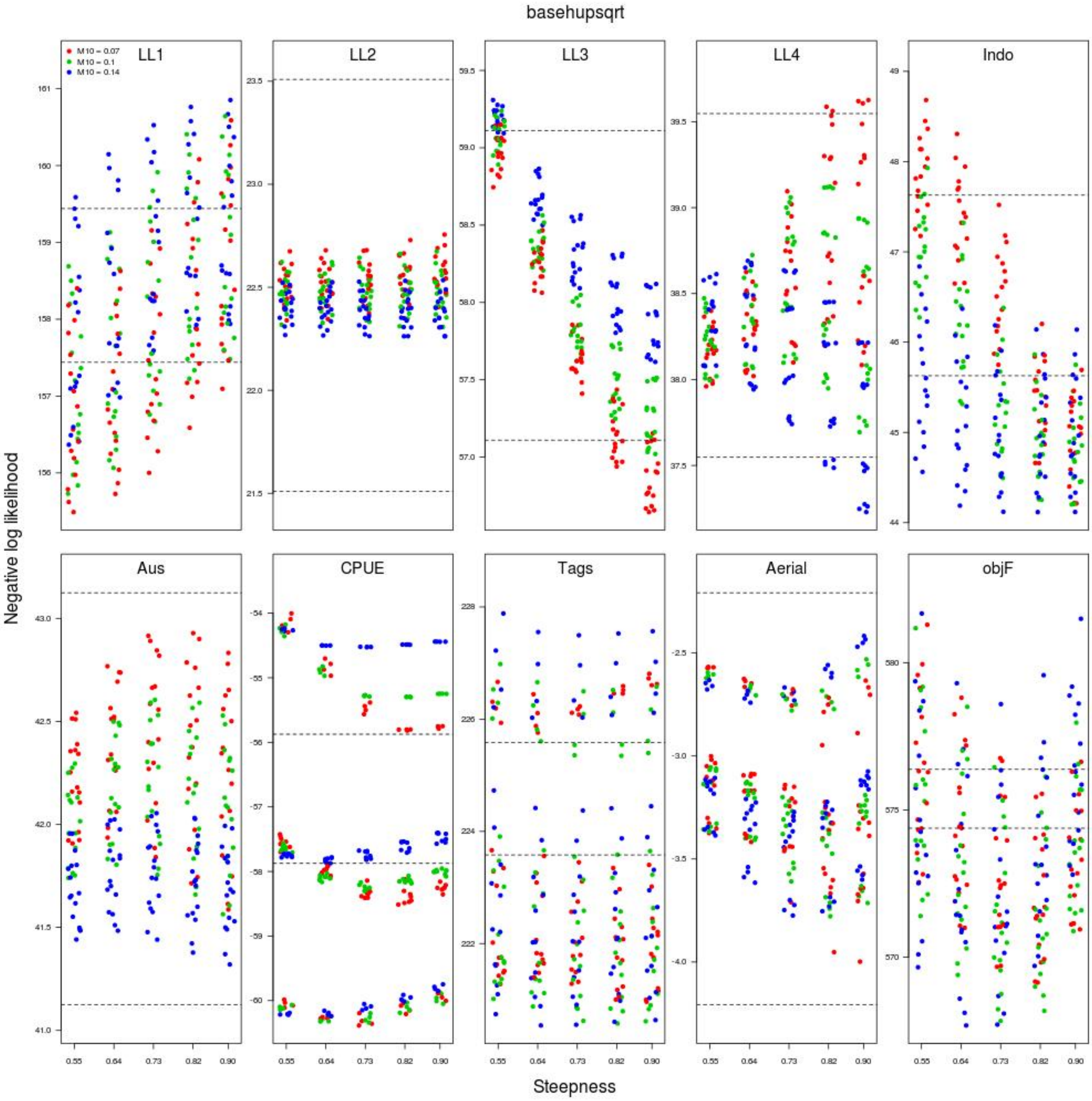
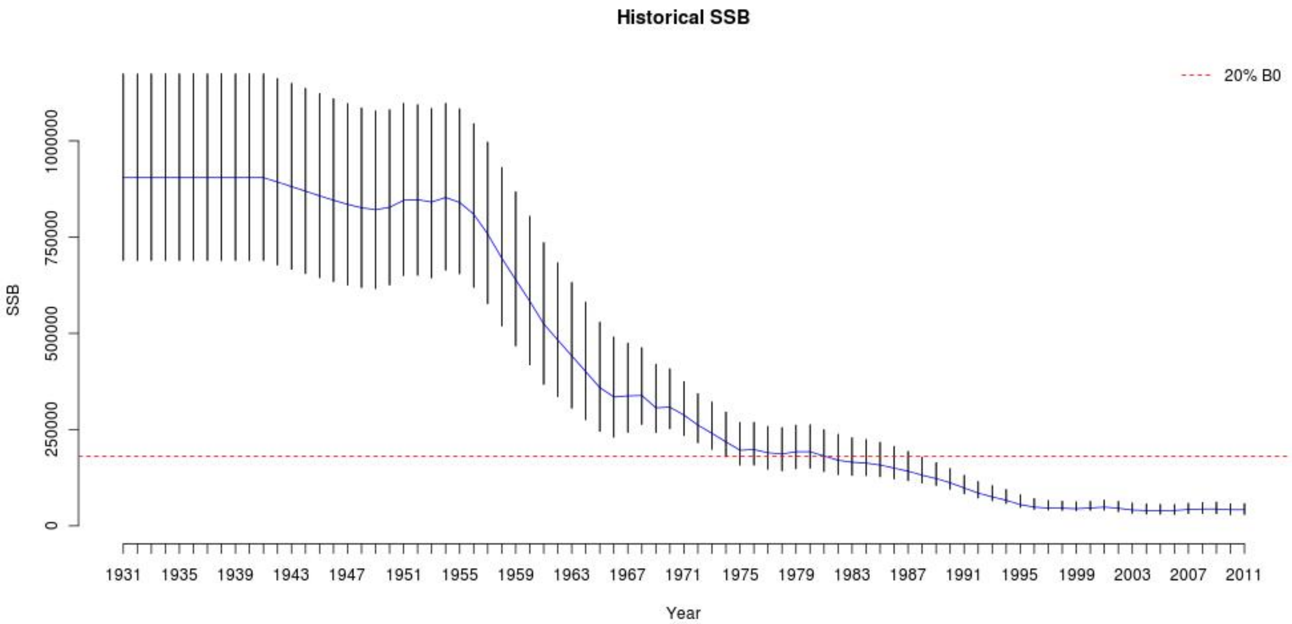


Figure 14: Historical SSB (top) and recruitment (bottom) dynamics for SBT for the **basehupsqrt** grid. Blue-line is the median and the whiskers extend out to show the 90% CI in both cases. For the SSB plot the interim rebuilding target (20% of B_0) is shown also and the median (and 90% CI) estimate of SSB depletion is 0.05 (0.03-0.07).



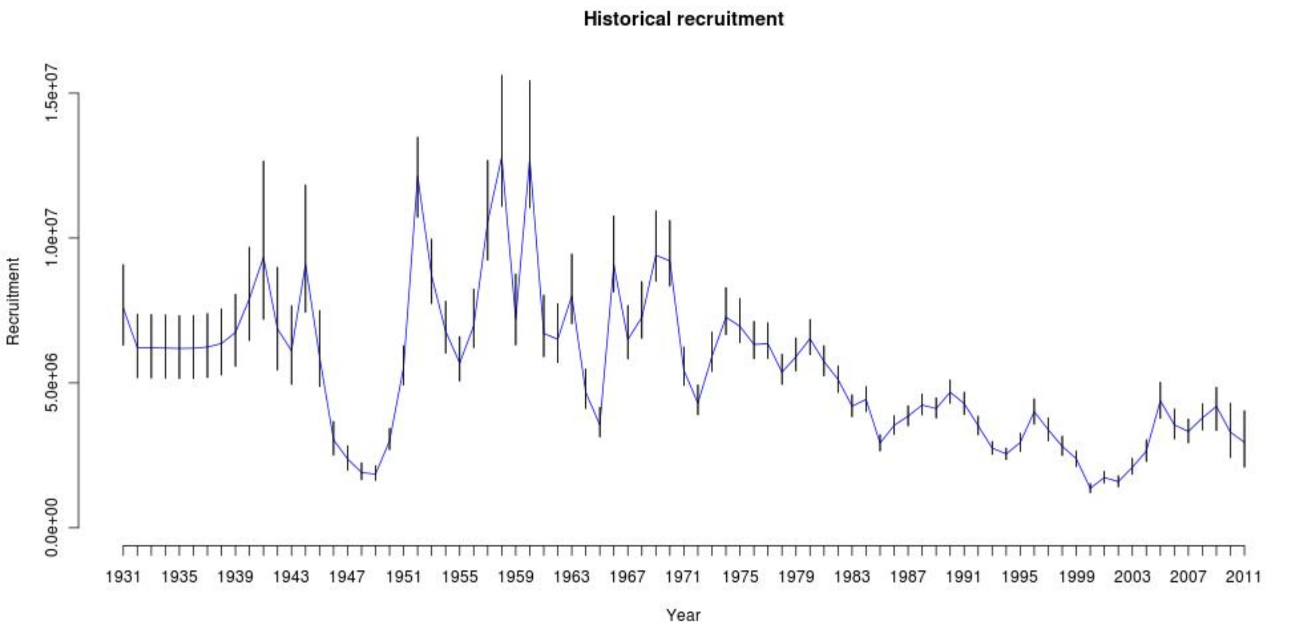


Figure 15: Estimated median stock-recruit relationship (dotted red line) and median (blue circle) and 90% CI (whiskers) actual stock-recruit relationship.

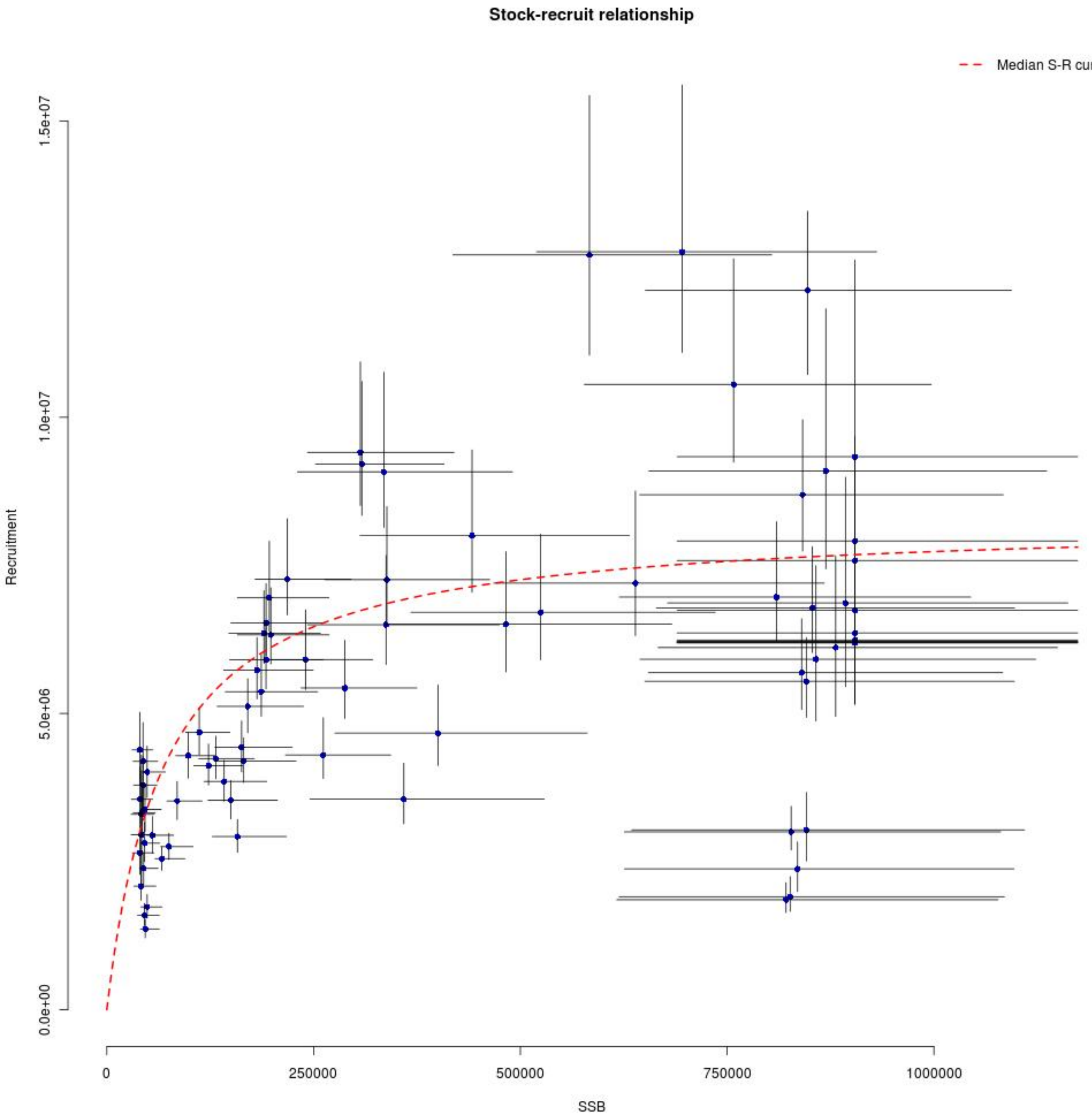


Figure 16: Constant catch scenarios (zero and 9,449 t) and their projected SSB dynamics. The full line is the median, the longer dotted line the 30th %ile, and the shorter dotted line the 5th %ile.