Commission for the Conservation of Southern Bluefin Tuna

みなみまぐろ保存委員会

Report of the Special Management Procedure Technical Meeting

15-18 February 2005
Seattle, U.S.A.

Report of the Special Management Procedure Technical Meeting

15-18 February 2005

Seattle, U.S.A.

Agenda Item 1. Opening of Meeting

1. The meeting was opened by the Chair, Dr Ana Parma.
2. The Chair presented the Terms of Reference for the meeting. In particular, the Chair noted that there were three outcomes required from this meeting:
   - an agreed operating model;
   - an agreed Reference Set (RS) and Robustness trials (RT) to replace the Core set and Sensitivity trials from the SAG 5 / SC 9; and
   - an agreed report that clearly describes the decisions made and the rationale for these decisions.
3. Participants were introduced and the list of meeting participants is included at **Attachment 1**.

Agenda Item 2. Adoption of agenda

4. The draft agenda circulated prior to the meeting was accepted and is **Attachment 2**.

Agenda Item 3. Appointment of rapporteurs

5. Rapporteurs from members and the panel were appointed to produce the text of the report. Dr Shelton Harley from NZ volunteered to serve as rapporteur.

Agenda Item 4. Admission of documents and finalisation of document list

6. The list of documents for the meeting was considered and classified according to the agenda items for which they were relevant. This list of documents is at **Attachment 3**.

Agenda Item 5. Choice of reference set and robustness trials for final MP testing

5.1 Review of sensitivity analyses conducted by members

7. Dr Jim Ianelli provided a review of the decisions to come from SAG 5 / SC 9. This was closely based on the document "*OMassumptions_being updated Feb05.doc*" that had been provided to participants prior to the meeting. In particular he noted:
   - The core set was based on 270 model runs that took about 6 hours to complete.
   - The weighting of the model runs for the 2000 projections was based on a mixture of prior and prior x likelihood weightings.
   - At SAG 5 it was agreed to update the tagging data so that these were consistent with the definitions for the fisheries seasons and catch estimates used in the MP development work.

8. CCSBT-MPTM/0502/04 provided in its introduction some useful points for consideration of whether a factor should be included in the reference set. These were discussed first with the remainder of the paper discussed later in the meeting. These points can be summarised as:

- Tend to include model hypotheses that represent a relatively high likelihood.
- Tend to include model hypotheses even with a relatively low likelihood if there is a general impression that the model should not be informative with respect to this particular dimension, for example, the different levels of steepness.
- Tend to drop model hypotheses if they add to the size of the GRID, without substantially altering the range of uncertainty in future dynamics (constant current catch projections) (e.g., the probability distribution over plausible scenarios could be adequately represented by a reduced number of such scenarios provided that these were accorded appropriate weights).
- Try to maintain a tractable number of GRID dimensions and levels, such that a sample size of 2000 projections will give a reasonable coverage of models.
- Represent a reasonable probability that CPUE indices may increase or decrease somewhat over the next few years - i.e. explicitly admit that it is not possible to predict with a high degree of certainty the directional change in CPUE in the near future (e.g. avoid a repetition of the problem encountered in the 2004 MP meeting that could undermine credibility in the MP process).

9. There was discussion of these general principles and broad agreement that these should be used to guide the workshop decisions with respect to the definition of the reference set and robustness trials. It was noted that for a factor to be included in the reference set, relative plausibility should remain appreciable over the range considered, and outcomes should be sensitive to different levels of the factor across that range. Conversely, for inclusion in the robustness trials, the hypothesis would be less plausible, but be shown still to have an important impact on performance. If an MP failed to provide adequate feedback control to some of the robustness trials, this could provide guidance on circumstances for which meta-rules may have to be developed. Furthermore, the results could be used to identify research that could be conducted and/or data collected to ensure that hypotheses embedded in sensitive robustness trials were indeed of low plausibility.

10. CCSBT-MPTM/0502/06 described the tuning of the D&M MP to the Core set OM and then applying it to constituent scenarios of the Core set and the Sensitivity trials. This version of the MP did not take account of length distribution information. The presence or absence of appreciable changes in anticipated performance as the levels of each factor were changed in turn was suggested as a guide for their inclusion or exclusion in the reference set and robustness trials. Key findings included:

- Setting the Indonesian selectivity to fixed above age 18 made a considerable difference suggesting it should be included in the reference set.
- The MP had difficulty with the sensitivity trials where a number of consecutive low recruitments was assumed. It was unable to reduce catches enough to allow rebuilding to the tuning level due to a current constraint on maximum reduction in TAC.
- The performance of the MP across the different levels of M0 was very similar suggesting that this factor could be excluded from the reference set.
- There was an appreciable difference with the results for the various CPUE series used in conditioning. In particular the w0.5 series led to appreciably more pessimistic projections. It was noted that this series is qualitatively quite different from the other CPUE series.

11. CCSBT-MPTM/0502/07 provided stock projections under constant catch projections (for the current catch level) for the core set and several sensitivity tests, and also examined some other model structures, particularly ones changing assumptions concerning effective sample size for the likelihood function for the length composition data. Key findings included:

- In general, the projections for the core set under the current constant catch were quite pessimistic.
- There was a close relation between the likelihood of different omega values and the CPUE series. For example, the ST-window series was associated with high omega values while the w0.5 series was associated with low values.
- Inclusion of tagging data led to higher M10 values and more optimistic projection results.
- Recent recruitment estimates showed less variability than expected, with similar values obtained for the various factors and levels in the core set.
- Changing catchability ages (i.e., the ages over which selectivity is normalized) for CPUE to 8-12 from 4-30 led to substantial reduction of biomass ranges and more pessimistic projections.
- Constant selectivity above age 18 for the Indonesian fishery resulted in poorer fits to the data.
- Alternative assumptions for the effective sample sizes for length composition data had substantial impacts on results.

12. In discussing these results, the following key issues were raised:

- Inter-relationships among factors should be examined.
- If there is concern that the magnitudes of the 2000-01 cohorts are being reflected as too well determined, it may be possible to add uncertainty in the projection phase in the same way as applied for the cohorts from 2002 onwards.
- The tag reporting rates are treated as known and without error. The most feasible way to address this, given the current model structure, would be to include alternative reporting rates in the "GRID".
- Dome-shaped selectivity for the Indonesian fishery would not be necessarily inconsistent with current biological knowledge.

13. CCSBT-MPTM/0502/04 described results of constant catch projections over the core set, sensitivity trials, and a range of alternative potential reference sets. It considered factors that could be added and those that could be removed, or the number of levels reduced. In particular it focussed on assumed sample sizes for age and length composition data, CPUE series, the inclusion of tagging data, and alternative catchability age-ranges for the CPUE calculations. Key findings included:

- Using the original sample sizes divided by two (as opposed to square-rooted) led to more optimistic projections (in median terms) and increased variation.
- Results were very sensitive to alternative catchability age-ranges for the CPUE calculations.
- The performance of median CPUE versus the five individual CPUE series together was not that great compared to other factors, particularly the assumption about the ages over which selectivity is normalised, suggesting that it may be possible to reduce the number of CPUE levels in the reference set.
- The inclusion of updated tag data together with a change of the reporting rate option, no longer tended to favour the highest M10 value, and an interaction between the weights assigned to size and age data and the inclusion or exclusion of tag data was noted. The inclusion of tag data generally leads to more optimistic projections in median terms.

14. In its conclusions, CCSBT-MPTM/0502/04 posed the following questions with regard to the choice of the reference set:

- Which sample weights (for the age and length composition data) to use and whether sample weight should be an axis in the grid.
- Whether uncertainty in the catchability age range should be included as an axis in the grid.
- Whether the CPUE axis in the original grid should be reduced to the median CPUE.
- Whether tagging data should be included.

15. CCSBT-MPTM/0502/05 presented revised reporting rates for the tagging data based on the most recent catch estimates and definitions of the fishing seasons. Eight options for reporting rates were reported for each of two assumptions regarding how reporting might change for unobserved sets on observed trips. Key points raised included:

- The eight options do not span the range of uncertainty in the actual reporting rates but provide a measure of the sensitivity of using direct information relative to an assumption for the reporting rate for major unobserved fisheries for which non-zero reporting rate is estimated to be the same as reporting rate for observed fisheries.
- Option 8 is the most plausible option as only in this case are the reporting rates based on direct information for the major non-observed fisheries for which a non-zero reporting rate is estimated.
- There is an unresolvable potential bias in the current estimates as a result of uncertainty about whether 100% tags were returned while an observer was on board independent of whether he was actually observing when a tagged fish was caught. The scaled and unscaled estimates provided in the paper provide a measure of the sensitivity to this uncertainty and the differences between the two are substantive. The current data do not allow the extent of actual bias to be determined, which emphasizes the importance that detailed and accurate data from observers are recorded and made available when observer data are used to estimate the reporting rate.

16. There was considerable discussion on how to treat the updated reporting rate estimates. In the past it was assumed that the presence of an observer on a vessel would result in high, if not 100%, reporting rate from all tagged fish that were recaptured which is the assumption in the scaled estimates. Therefore, the scaled estimates were considered the most appropriate to use in conditioning the operating model. With respect to the selection of a reporting rate option and the conclusion of the paper that option 8 was the most plausible, it was noted that consideration of alternative reporting rate was undertaken as part of the request of the SAG (Attachment 5) to investigate why the tagging data was so informative with respect to M10. It was noted that option 5 had been used in past MP conditioning without any substantive discussion or consideration of the possible sensitivity within conditioning, but simply as the central index in the range of options. It was further noted that model runs that included the option 8 reporting rates did not exhibit many of the problems associated with the option 5 reporting rates (which were likely partly responsible for the decision to exclude the tagging data at SAG 5). Based on these discussions, and the views of those participants familiar with the fisheries and the tagging data, it was concluded that option 8 would be accepted to use in conditioning the operating model.

Notwithstanding the decision to use option 8 reporting rates when including the tagging data, it was noted that a detailed examination of the assumptions for the reporting rate was outside the scope of this meeting and that the decisions made here for the purpose of MP evaluation should not be considered to have implications for the use of reporting rates in future assessments. This issue would need to be considered by the SAG.

5.2 Further analyses

17. Based on the presentations and discussions of the members' papers, several alternative grid runs were undertaken to provide further guidance on the choice of the reference set and robustness trials. Factors considered included: alternative weightings for the age and length composition data, alternative catchability age-range assumptions, and alternative treatment of the tagging data (in and out and alternative reporting rates).

5.3 Choice of final reference set of operating models and robustness trials

5.3.1 Grid axes, weights and integration of results

18. Based on the early results, and the additional analyses that were undertaken, the following proposals were discussed in relation to the possible composition of the reference set:

- <u>Reducing the dimensions of M0 in the reference set</u>: While results from a specific MP were not sensitive to alternative levels of M0, the projections based on constant catch were quite sensitive to such levels. **Decision: leave M0 unchanged (three levels)**.
- <u>Estimating omega</u>: It was noted that omega was strongly associated with CPUE series (some favouring the low value of omega and others the higher value of omega, for all other factors the same), and could therefore potentially be estimated instead of being part of the grid. While this might reduce the dimensions of the reference set, it was recognized that this would require changes to the code and could lead to model instability. **Decision: leave omega unchanged (two levels)**.
- <u>Reducing the number of levels for M10 in the reference set</u>: It was noted that none of the core set and sensitivity trial runs examined (including runs which exclude tag data) provided much weight to the low level of M10 (0.07). (Figures 1, 2 and 9 show examples demonstrating this). **Decision: remove the low level for M10 (0.07) and proceed with two levels**.
- <u>Include alternative catchability age-range assumptions in the reference set</u>: There was considerable discussion relating to the ages predominately captured in the LL1 fishery, changes in selectivity, and the possible effects of historical discarding. It was agreed that, *a priori*, the current range (4-30) was not amongst the most suitable options. Following extensive discussions about the 8-12 option, it was decided that this should be included as one option in the reference set. It was noted that CCSBT-MPTM/0502/04 provided results for 6-18, so an analysis based on 4-18 was undertaken for comparison. Little difference was found between 6-18 and 4-18 and it was decided that 4-18 was the more appropriate option to include in the reference set, but with a double weight to reflect that the 4-18 and 6-18 options were considered equally likely to the 8-12. (Figs. 10 and 11). **Decision: to add catchability age-range as an axis in the reference set with two levels, 4-18 and 8-12, with prior weights of 0.67 and 0.33 respectively**.
- <u>Include constant Indonesian selectivity above age 18 as an option in the reference set:</u> While this was shown to have an important effect on results in some of the analyses, the fits under this option were found to be poor, particularly to the plus group in the Indonesian fishery age composition data. Furthermore it was not clear how the alternative hypotheses already included in the reference set, but difficult to reconcile with such constant selectivity, could continue to be accommodated. In addition, there are consequent marked changes in the likelihoods of the different M10 values and the possibility of solving some of these problems by allowing for changes in natural mortality at older ages was not possible to accommodate computationally at this stage. **Decision: keep this selectivity option in the robustness trials**.
- <u>Include tagging data in the operating model for the reference set:</u> With the updated data and change in reporting rates (to option 8 from option 5), the undesirable properties (large preference for high values of M10) seen when the tagging data were previously included were less evident

(Fig. 1 compared to Fig. 3). **Decision: include tagging data in the reference set and add a "no tagging data" run to the robustness trials**.

- Include alternative weightings for the length and age composition data as levels in the reference set: There was considerable discussion over the inclusion of this as an axis in the reference set, and in particular how many levels to include. Three options were considered: (1) SQRT (the square root of the original sample sizes multiplied by 5 – this was what had been chosen by the panel for use in the operating model) (Fig. 10), (2) SQRT2 (double the SQRT values); and (3) ORIG0.5 (the original samples sizes divided by two) (Fig. 11). The key differences between SQRT and ORIG0.5 were how the early data are weighted relative to the more recent data. SQRT2 provided an option where the overall sample sizes were generally larger. ***The average samples sizes for each fishery under each option are provided in Table 1***. Several model outputs were compared for the options. The standard deviations of the normalised residuals for the fit to the size composition data increased going from SQRT to SQRT2. The value for LL1 fishery increased, but did not get that close to a value of 1 (from 0.46 to 0.62). For some of the fisheries (LL2 and LL4 fisheries for which there are early data) the standard deviations were somewhat greater than one in the SQRT option and these moved substantially further away from one under SQRT2. This was not considered to be a desirable property because it tended to give the early data too much weight. It was difficult to distinguish between SQRT and ORIG0.5 as the differences were hidden because they are related to changes in relative weightings over time. Close examination of the projections under SQRT and ORIG0.5 indicated that these options encompassed a wide range of immediate past and immediate future stock trends (Fig. 7). **Decision: include weightings of sample sizes as an axis in the reference set with two levels, SQRT and ORIG0.5 with equal weight**.

Table 1. Average effective sample sizes specified by fishery and option.

Option	LL1	LL2	LL3	LL4	Indonesia	Surface
SQRT	51	31	18	28	69	25
SQRT 2	102	62	37	56	139	49
ORIG 0.5	74	21	17	24	114	22

- Add the "four years low recruitment" option as an axis to the reference set: After a short discussion it was decided that the autocorrelation included in the projections probably provided a good enough reflection of some indications that immediate future recruitments will also be low**.** **Decision: keep "four years low recruitment" as an option for the robustness trials**.
- Explicitly consider uncertainty in the estimates of the 2000 and 2001 year classes in the reference set: It was noted that the 2000 and 2001 year classes were included in the projections with no added uncertainty even though there were few observations of them in the data and the estimates were associated with high coefficients of variation (cv). A range of options for adding extra noise was discussed. Based on inferences from the Hessian matrix estimates of uncertainty and consideration of the levels of error generally encountered in other sources of data, it was considered that adding error independently to each estimate (based on a cv of 0.4 and a lognormal distribution corrected for bias) was an appropriate approach to address the underestimation of uncertainty. **Decision: additional error was added to the estimates of these cohorts in the projections**.

19. A summary of the reference set as a result of these decisions is presented in Table 2. Full documentation of the reference set specification is provided in **Attachment 5**.

Table 2. Specification of grid axes for reference set specifications:

	Levels	Cumul N	Values			Prior	Simulation Weights
Steepness	3	3	0.385	0.55	0.73	0.2, 0.6, 0.2	Prior
M0	3	9	0.3	0.4	0.5	Uniform	Posterior
M10	2	18		0.1	0.14	Uniform	Posterior
Omega	2	36		0.75	1	0.4, 0.6	Posterior
CPUE	5	180				Uniform	Prior
q Age-range	2	360		4-18	8-12	0.67, 0.33	Prior
Sample Size	2	720		Sqrt	Original/2	Uniform	Prior

5.3.2 Robustness trials

20. The starting point for discussions of the components of the robustness trials was the table of sensitivity tests in the SAG report (Attachment 5). Based on this table, decisions on the structure of the reference set (see above), and other discussions, the following changes were made to the sensitivity tests from SAG 5 to make up the robustness trials:

- <u>Excluding tagging data</u>: As the tagging data are now included in the reference set it was considered useful to include "no tagging data" as a robustness test. It was noted that previously excluding these data had provided greater weight to the lowest level of M10 (which was now removed from the reference set). **Decision: include "no tagging data" as a robustness test and include all three levels of M10 (i.e. including 0.07), but only for the SQRT sample size weighting option**.
- <u>Series of years of low recruitment</u>: It was noted that this was for additional years after the weak 2000 and 2001 cohorts. Also, it was noted that this option did not require running a grid as it applied only to the projection period. **Decision: include two levels for this robustness test - (a) 4 extra years (2002-05), and (b) 2 extra years (2002-03), and apply across the full reference set**.
- <u>Changes in future catchability</u>: It was noted that both increases and decreases had been considered in the sensitivity test, but from a practical standpoint only increasing catchability impacted resource recovery negatively. **Decision: remove decreasing catchability from the robustness trials**.
- <u>Increasing the strength of the 2000 and 2001 cohorts</u>: It was noted there are several reasons why these recruitment estimates may be biased low (estimates recruitment lower than the true). It was considered important to test the robustness of the various MP's to this scenario, i.e., will the MP's recognize that things aren't as bad and not make unnecessary early TAC cuts. It was considered that improved catch performance by MP's for this scenario compared to the reference set would improve Manager confidence in them. **Decision: Include a run where the 2000-01 recruitments are increased by a factor of three**.
- <u>Penalty for high exploitation on age 3</u>: When considering the uncertainty in the strength of the 2000 and 2001 cohorts, it was noted that a useful "ground-truthing" exercise would be to examine the estimated exploitation rate for age 3s (the predominant age class taken in the surface fishery) for the period 2004-06. For 2004, it was found that for the reference set, the exploitation rates were very high. There was concern that these high exploitation rates may not be feasible or consistent with recent experience in the surface fishery. It was considered important to develop a robustness test that addresses these concerns. The specific details of the implementation is provided in **Attachment 6**. Briefly, a penalty was added to the age 3 surface fishery exploitation rate s in 2004 to ensure that this does not exceed exploitation rates estimated for the mid 1980s (a period where it is recognized that exploitation rates for juvenile fish appeared to be near the

maximum feasible). It was noted that this was a new robustness test and there was no opportunity to evaluate the results of this run at the meeting. **Decision: add a robustness test which penalizes high exploitation rates for age 3s in 2004**.

21. Summary of these decisions and a description of what is undertaken in each of the robustness trials is presented in **Attachment 6**.

Agenda Item 6. Performance statistics

22. The Chair presented the paper "*Performance_stats_Feb05.doc*" which had been provided to participants prior to the meeting. The document was revised based on the discussions and an updated version is provided as **Attachment 7**. In summary, the group noted:

- That it would be useful to provide a generalised function to calculate the risk statistic (i.e., user defined thresholds). It was also noted that the index as originally defined accounted only linearly for the distance below the threshold line and it was considered important to modify the index so that biomasses further below the threshold line are given proportionally greater weight. To achieve this a power term was added to the equation to allow some flexibility. It was recommended that discussions of choices for thresholds levels and values for the power term could be discussed at the MP4 workshop. This function was to be added to the graphics package so that users could investigate options.
- It was noted that some performance statistics (e.g., Equation 14) were based on a first TAC change in 2006. Rather than attempt to generalize the equations, a sentence was added to the text to indicate that Equation 14 constituted a particular example and could be adapted.
- The starting year for equation 12 for AAV was changed from 2003 to 2005 so that only future years were considered.

It was again recognized how important it was to easily view and compare results from different analyses with the same plots etc. It is important to maintain/update the graphics package.

Agenda Item 7. Tuning levels

23. There was a need to check the achievability of the 1.3 tuning level for the new reference set. To provide guidance on this two sets of projections were undertaken; (1) zero catch projections from 2008; and (2) making TAC reductions of 5000t every three years starting in 2008 (hence the TAC would reduced to zero by 2014). Results of the second set of projections gave a median B2022/B2004 corresponding to a tuning level of 1.98. This suggests that a tuning level of 1.3 was possible given the current options for maximum TAC reductions and schedules for TAC changes. It was agreed to use tuning levels of 0.9, 1.1, and 1.3.

24. Robustness trials will need to be run only with the MPs that have been tuned to the reference set at the 1.1 tuning level for TAC schedule change "option *b*" (see below).

25. The Working group concluded that the agreed reference set provides a sufficient basis to allow final recommendations on the choice of an MP. However, the basis for advice on a tuning level is dependent, at least in part, on the estimates for the 2000 and 2001 cohorts. The workshop noted that the 2000 and 2001 recruitment estimates and data supporting them represent a situation outside the "norm" in the historic data (e.g., the very low numbers of 3 and 4 year olds in the recent Japanese longline catches). This, combined with the potential

sensitivity of performance at a given tuning level to the estimates for these two cohorts, means that final advice on the consequences of different tuning levels may require taking into account projection results from both reference and some of the robustness set. The need or otherwise for this will depend on new data and the indicator analyses to be considered at SAG 6. How best to combine reference set and robustness test results to provide advice on tuning levels is an important question that will need to be discussed at the Management Procedure workshop in May.

Agenda Item 8. Schedules for TAC changes

26. Chair presented the paper "*OPtions for catch stability_Feb05.doc*." It was noted that there was a mistake in the paper as originally distributed–maximum catch reduction in options a and b should be 5000t not 3000t. It was noted that, given that the decision year for option a) was originally specified as 2004, it excluded the catch and CPUE data for 2004 in computing the TAC for 2006. Although this is sufficient for MP testing, the May meeting needs to clarify which data will be used in the implementation of the chosen MP, should the Commission subsequently select option a). The revised paper is provided as **Attachment 8**.

27. Options a and c (Appendix 8) need to be tested using only the 1.1 tuning level.

Agenda Item 9. Other computing issues

28. No other computing issues were identified at this meeting.

Agenda Item 10. Workplan and timetable

29. The new projection code, reference set, and files needed for MP testing will be distributed by March 15 at the latest. Components will be circulated earlier as they are completed. By 21 March (or earlier if possible), developers will provide feedback on the performance of the operating model from tuning their MP's based on catch stability option b and 1.1 and 1.3 tuning levels.

30. Considering that the meeting in May will not allow sufficient time to evaluate a large number of MP's, the workshop decided that developers should present at the most three variants on each of their MPs with an indication of their preferred choice.

Agenda Item 11. Finalisation and adoption of meeting report

31. The report of the meeting was adopted.

32. The meeting closed at 8:15pm, 18 February 2005.

List of Attachments

Attachments

1 List of participants

2 Agenda

3 List of documents

4 Results of alternative “grid” explorations

5 Configuration of operating model for the reference set

6 Robustness trials

7 List of performance statistics

8 Options for schedule of TAC changes

Attachment 1

List of Participants
Special Management Procedure Technical Meeting
15-18 February 2005
Seattle, U.S.A.

CHAIR
Dr Ana PARMA
Centro Nacional Patagonico
Pueto Madryn, Chubut
Argentina
Phone: +54 2965 451024
Fax: +54 2965 451543
Email: parma@cenpat.edu.ar

ADVISORY PANEL
Dr James IANELLI
REFM Division
7600 Sand Pt Way NE
Seattle, WA 98115
USA
Phone: +1 206 526 6510
Fax: +1 206 526 6723
Email: jim.ianelli@noaa.gov

Prof John POPE
The Old Rectory
Burgh St Peter
Norfolk, NR34 0BT
UK
Phone: +44 1502 677377
Fax: +44 1502 677377
Email: PopeJG@aol.com

Dr Trevor BRANCH
Department of Mathematics and Applied Mathematics
University of Cape Town
Rondebosch 7701
South Africa
Phone: +27 21 6502336
Fax: +27 21 6860477
Email: tbranch@maths.uct.ac.za

AUSTRALIA

Dr Tom POLACHECK
Senior Principal Research Scientist
Pelagic Ecosystems Sub-Program
Division of Marine Research
CSIRO P.O. Box 1538
Hobart, TAS 7001
Phone: +61 3 6232 5312
Fax: +61 3 6232 5012
Email: tom.polacheck@csiro.au

Dr Marinelle BASSON
Senior Research Scientist
Pelagic Ecosystems Sub-Program
Division of Marine Research
CSIRO P.O. Box 1538
Hobart, Tas 7001
Phone: +61 3 6232 5336
Fax: +61 3 6232 5012
Email: marinelle.basson@csiro.au

FISHING ENTITY OF TAIWAN

Dr. Chin-Hwa SUN (Jenny)
Professor and Director
Institute of Applied Economics,
National Taiwan Ocean University
2 Pei-Ning Road, Keelung 20224
Taiwan
Phone: +886 2 2462 2324
Fax: +886 2 2462 7396
Email: jsun@mail.ntou.edu.tw

JAPAN

Dr Sachiko TSUJI
Section Chief
Temperate Tuna Section
National Research Institute of
Far Seas Fisheries
5-7-1 Orido, Shimizu, Shizuoka,
Shizuoka 424-8633
Phone: +81 543 36 6042
Fax: +81 543 35 9642
Email: tsuji@fra.affrc.go.jp

Dr Hiroyuki KUROTA
Temperate Tuna Section
National Research Institute of
Far Seas Fisheries
5-7-1 Orido, Shimizu, Shizuoka,
Shizuoka 424-8633
Phone: +81 543 36 6043
Fax: +81 543 35 9642
Email: kurota@affrc.go.jp

Prof Doug BUTTERWORTH
Department of Mathematics and Applied
Mathematics
University of Cape Town
Rondebosch 7701
South Africa
Phone: +27 21 650 2343
Fax: +27 21 650 2334
Email: dll@maths.uct.ac.za

NEW ZEALAND

Dr Shelton HARLEY
Senior Scientist
Ministry of Fisheries
PO Box 1020, Wellington
Phone: +64 4 494 8267
Fax: +64 4 494 8261
Email: harleys@fish.govt.nz

REPUBLIC OF KOREA

Dr Dae-Yeon MOON
Senior Scientist
Distant-water Fisheries Research Div.
National Fisheries R & D Institute
408-1 Shirang-ri, Kijang-gun
Pusan 619-902
Tel: +82 51 720 2320
Fax: +82 51 720 2337
Email: dymoon@nfrda.re.kr

Attachment 2

Agenda

Special Management Procedure Technical Meeting

15-18 February 2005

Seattle, U.S.A.

1. Opening

   Terms of Reference:

   Choose operating models and robustness trials for final Management Procedure (MP) testing and specify details of trials.

2. Adoption of agenda
3. Appointment of rapporteurs
4. Admission of documents and finalisation of document list
5. Choice of reference set and robustness trials for final MP testing

   5.1 Review of sensitivity analyses conducted by members
   5.2 Further analyses
   5.3 Choice of final reference set of operating models and robustness trials
   - grid axes, weights and integration of results
   - robustness trials

6. Performance statistics
7. Tuning levels
8. Schedules for TAC changes
9. Other computing issues
10. Workplan and timetable
11. Finalisation and adoption of meeting report

Attachment 3

List of Documents

Special Management Procedure Technical Meeting

(CCSBT-MPTM/0502/)

01. Draft Agenda of MPTM.
02. Draft List of Participants of MPTM.
03. Draft List of Documents of MPTM.
04. (Australia) Exploration of the SBT operating model with implications for the selection of the core set and robustness trials.: D. Kolody, M. Basson, A. Preece, J. Hartog and T. Polacheck.
05. (Australia) Updated estimates of tag reporting rates for the 1990's tagging experiments.: P. Eveson and T. Polacheck.
06. (Japan) Application of the D&M management procedure to the core and sensitivity trials to assist identify factors to which MP performance is likely to be the most sensitive.: D.S. Butterworth and M. Mori.
07. (Japan) Further Exploration of the Operating Model for the Management Procedure Evaluation.: H. Kurota.

Attachment 4

Results of alternative "grid" explorations

Effect of model selection on M10

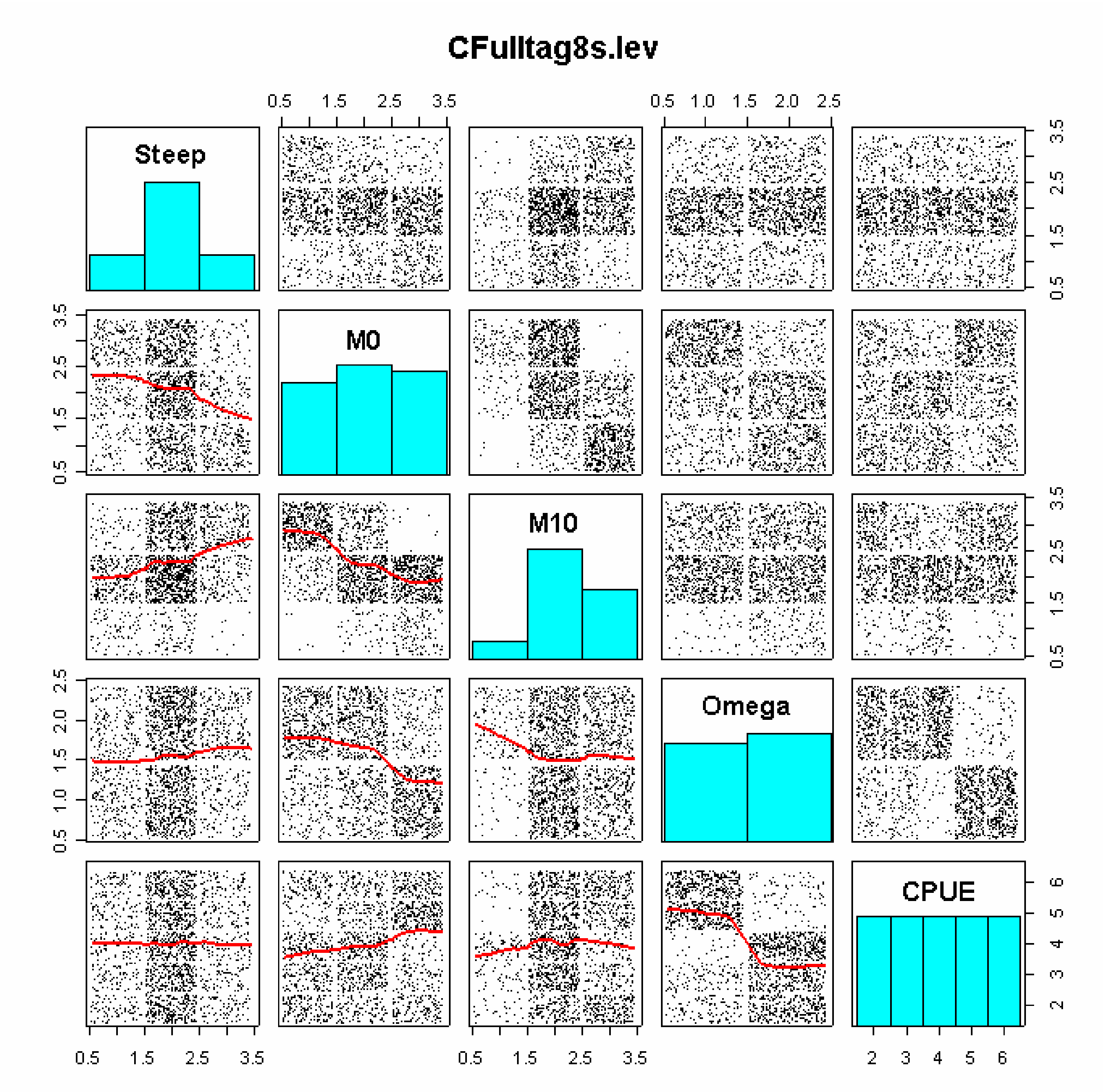


Figure 1. Pairwise plots of 2000 samples drawn from an MPD grid compiled for tag-reporting rate option 8 (updated prior to February 2005)

Effect of catchability age-ranges for CPUE

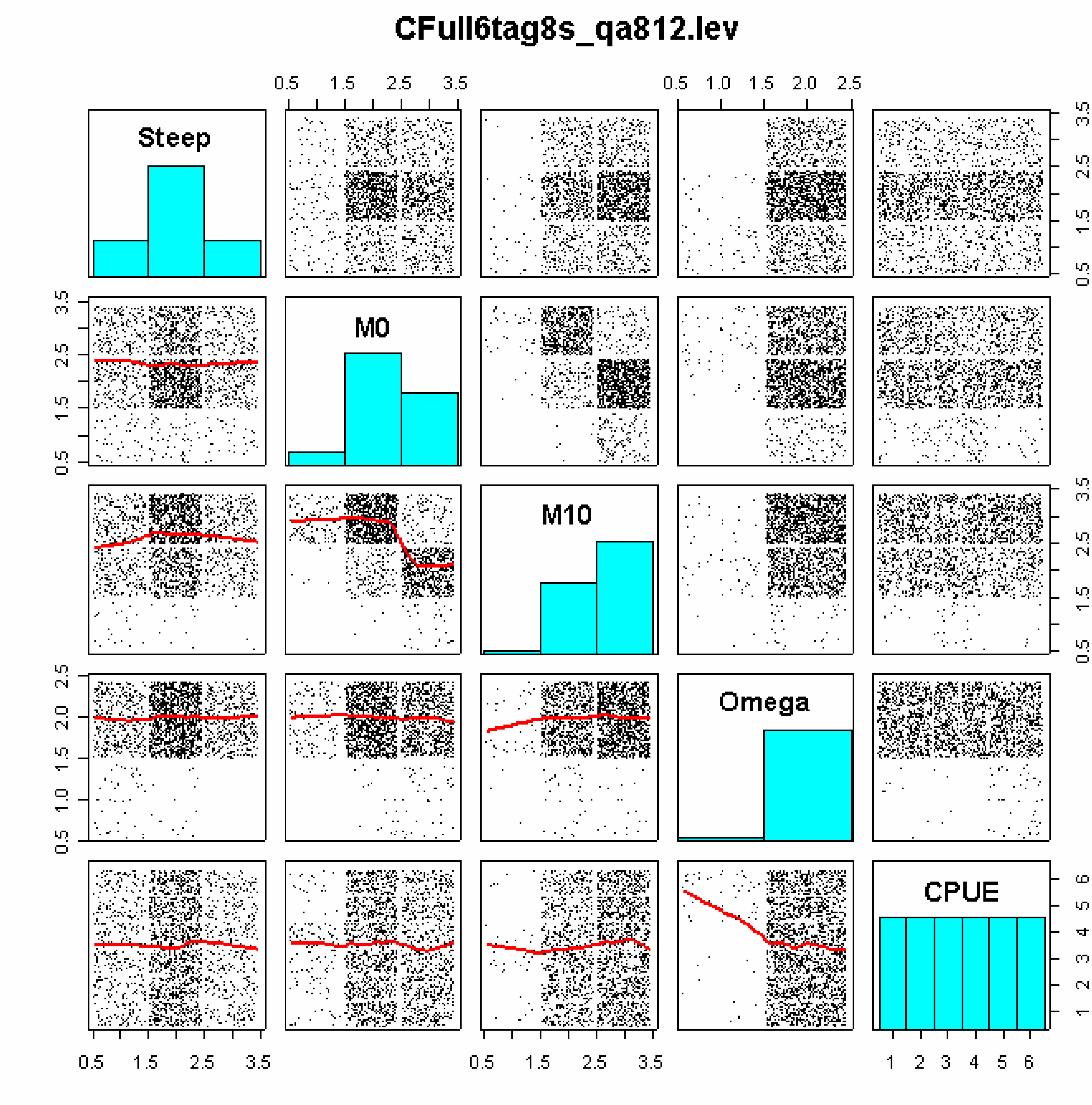


Figure 2. Pairwise plots of 2000 samples drawn from an MPD grid compiled for tag-reporting rate option 8 (updated prior to February 2005) and using the age range 8-12 to normalize CPUE. Note that this run used preliminary reporting rates and also included the median CPUE series (series 1).

Effect of tagging data

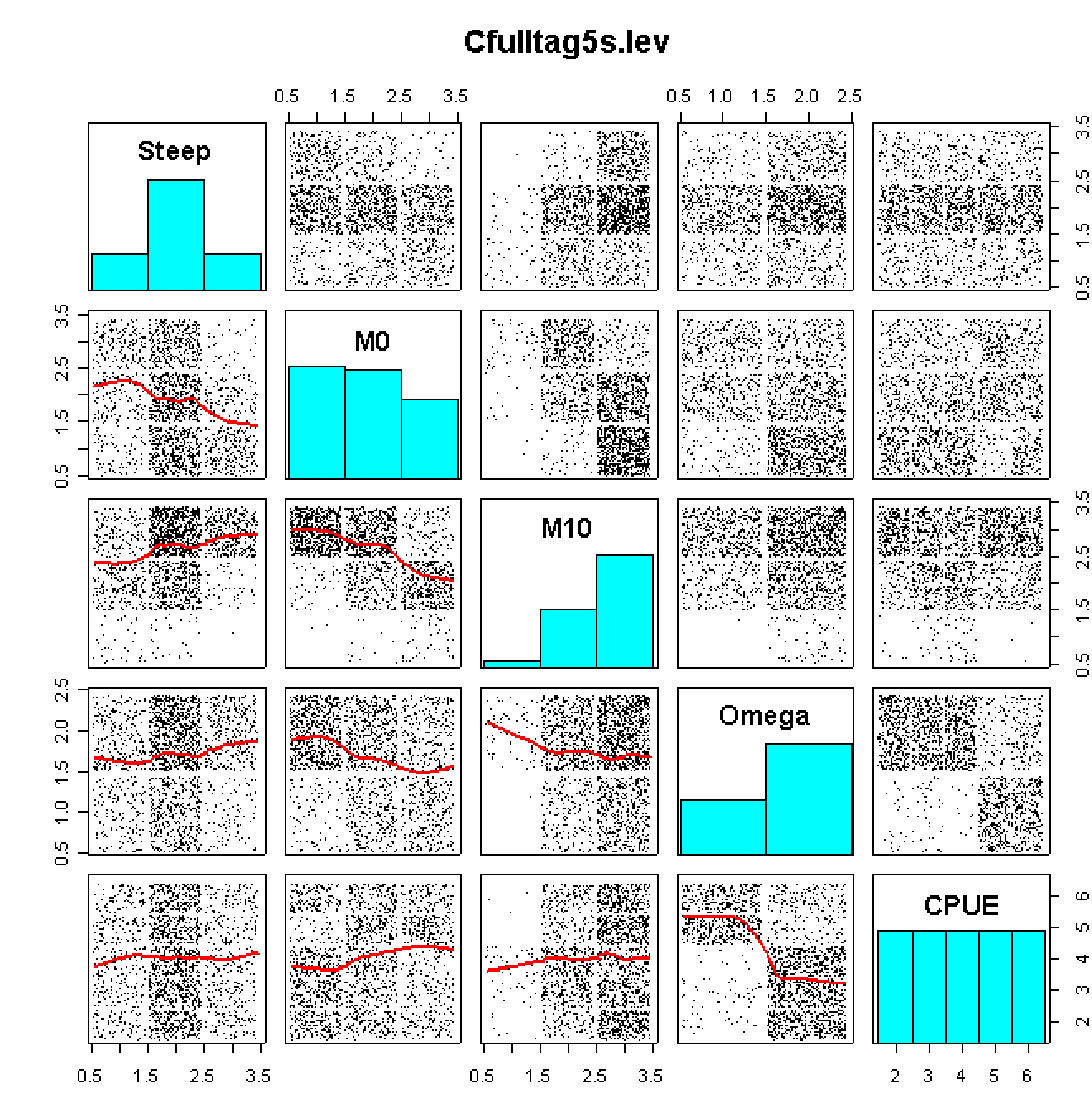


Figure 3. Pairwise plots of 2000 samples drawn from an MPD grid compiled for tag-reporting rate option 5 (updated prior to February 2005)

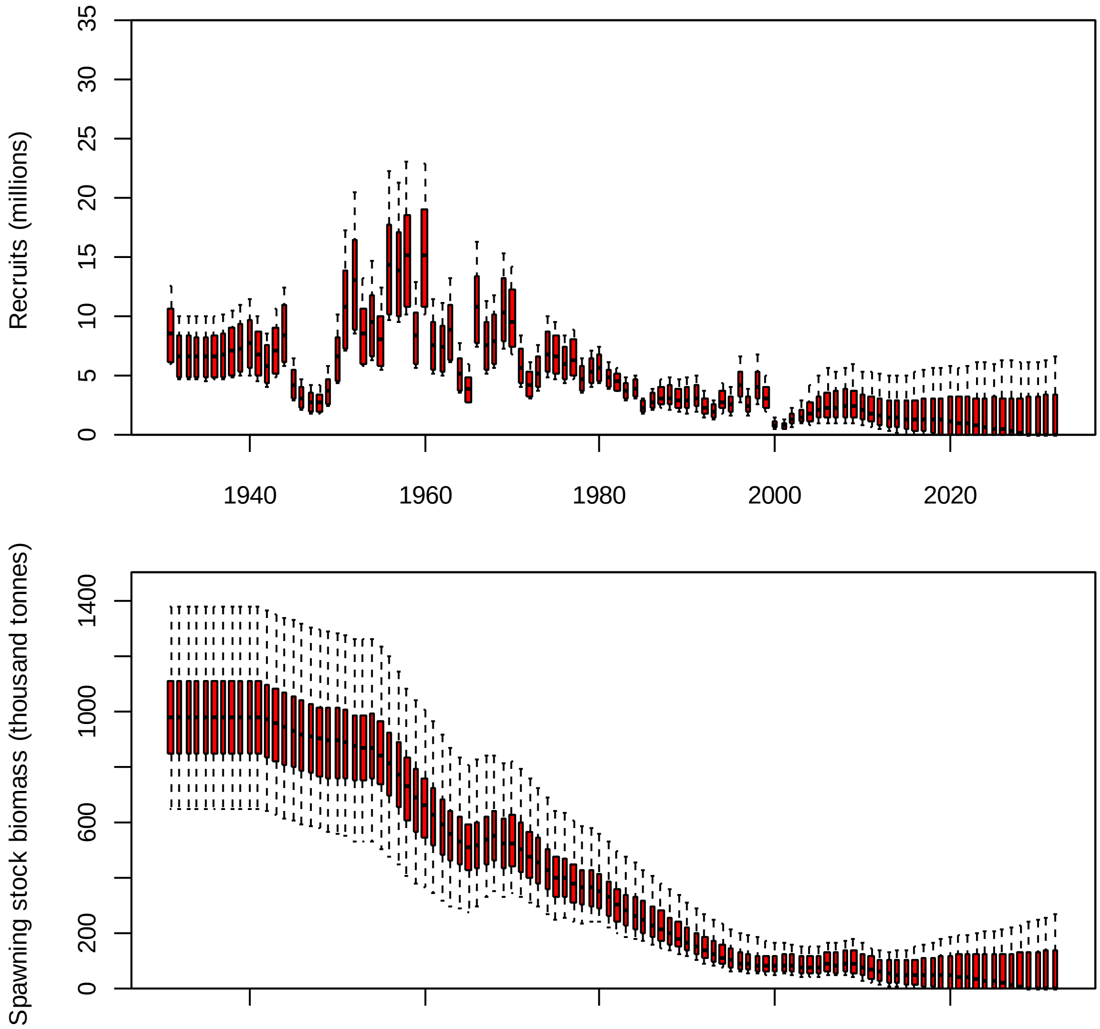


Figure 4. Constant catch projections using the Cfullnotag set, i.e., the full set prior to the meeting. This set did not include the tagging data. Indicates the 5%, 25%, 50%, 75% and 95%-iles.

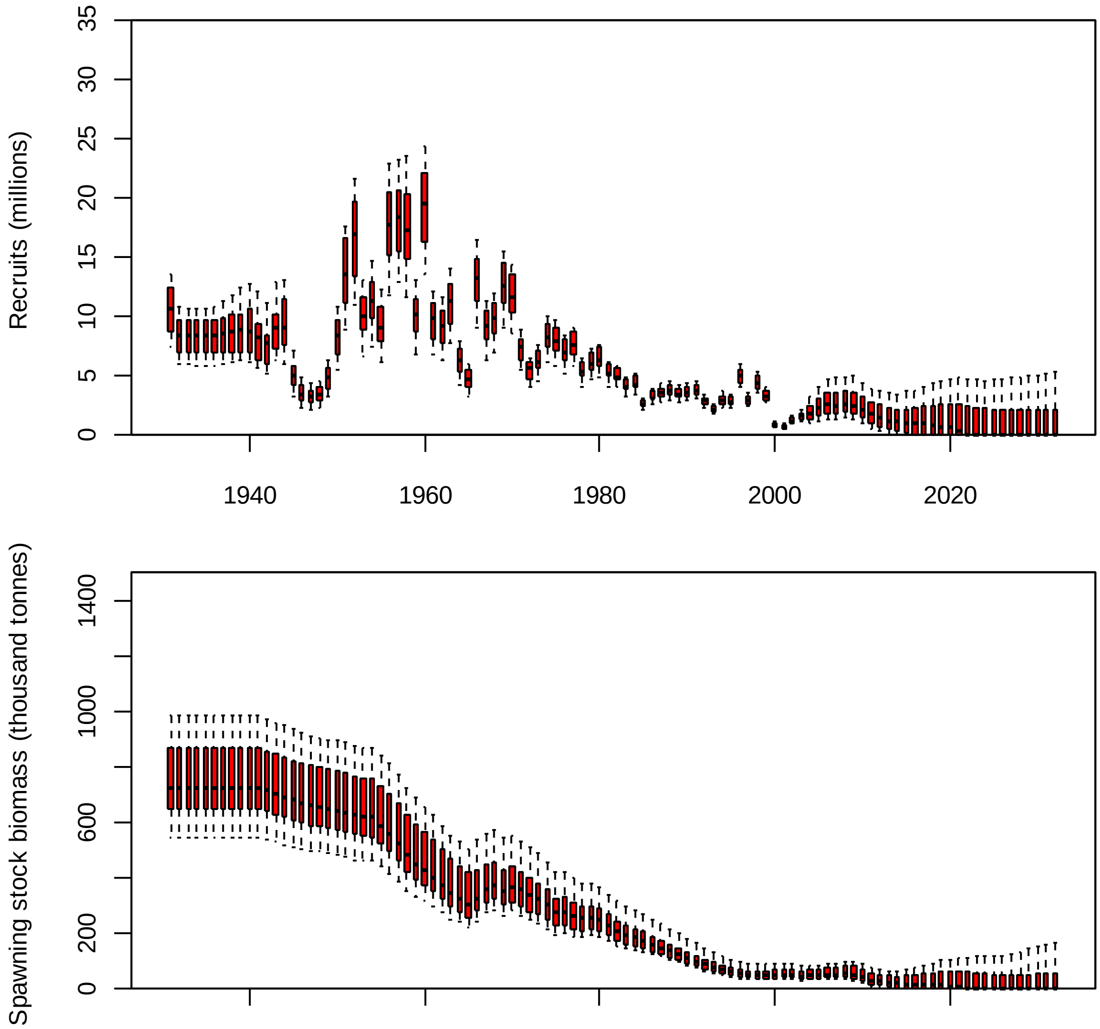


Figure 5. Constant catch projections using the Cfull2 set, i.e., the full set agreed after the meeting. Tagging data were included. The new tag reporting rates (option 8) were used in these analyses, as well as two additional grid axes.

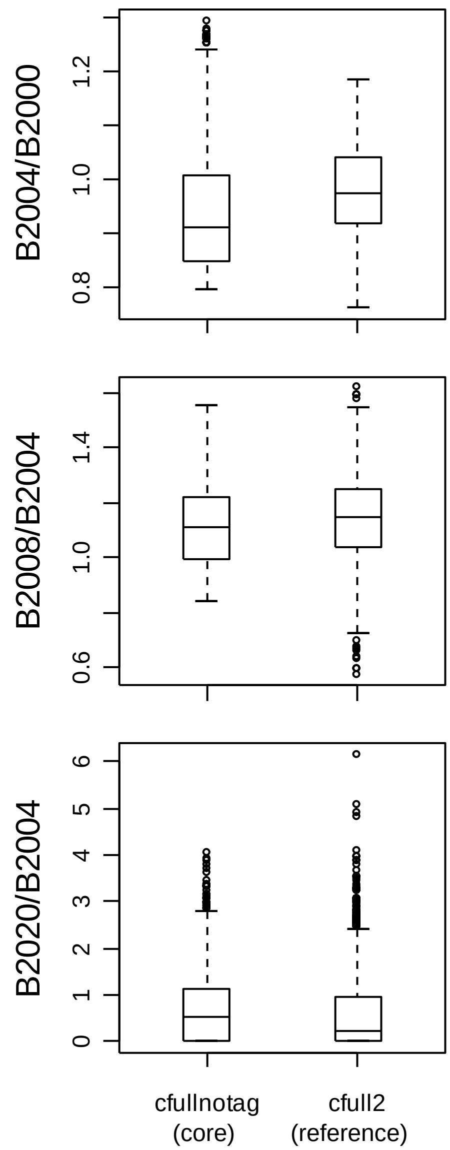


Figure 6. Comparison of trends in spawning biomass between the original reference set (cfullnotag) and the new reference set (cfull2). The box indicates the median and the 25% and 75%-iles. The whiskers are 1.5* the distance between the 25% and 75%-iles, and the circles indicate any values outside the range of the whiskers.

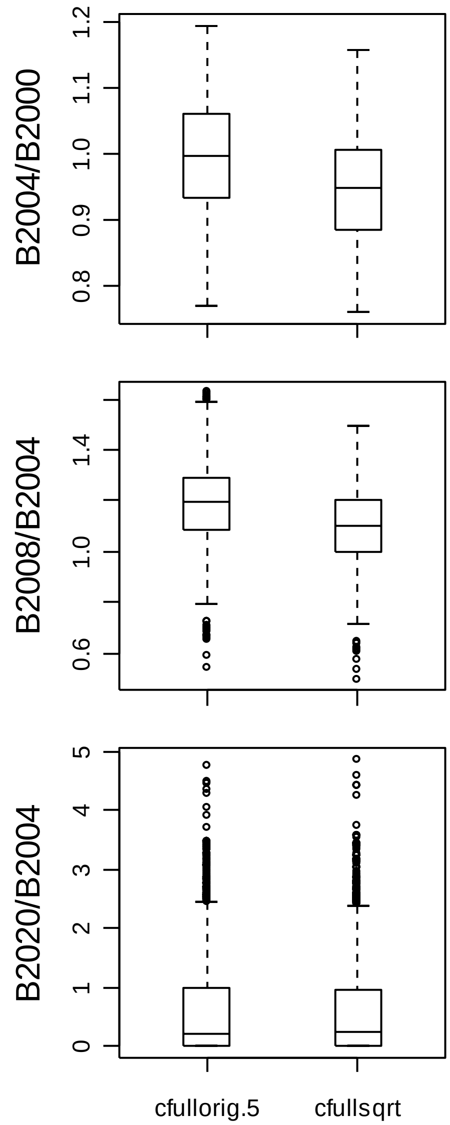


Figure 7. Comparison of the performance of two sets of measures of the weights on length frequency data.

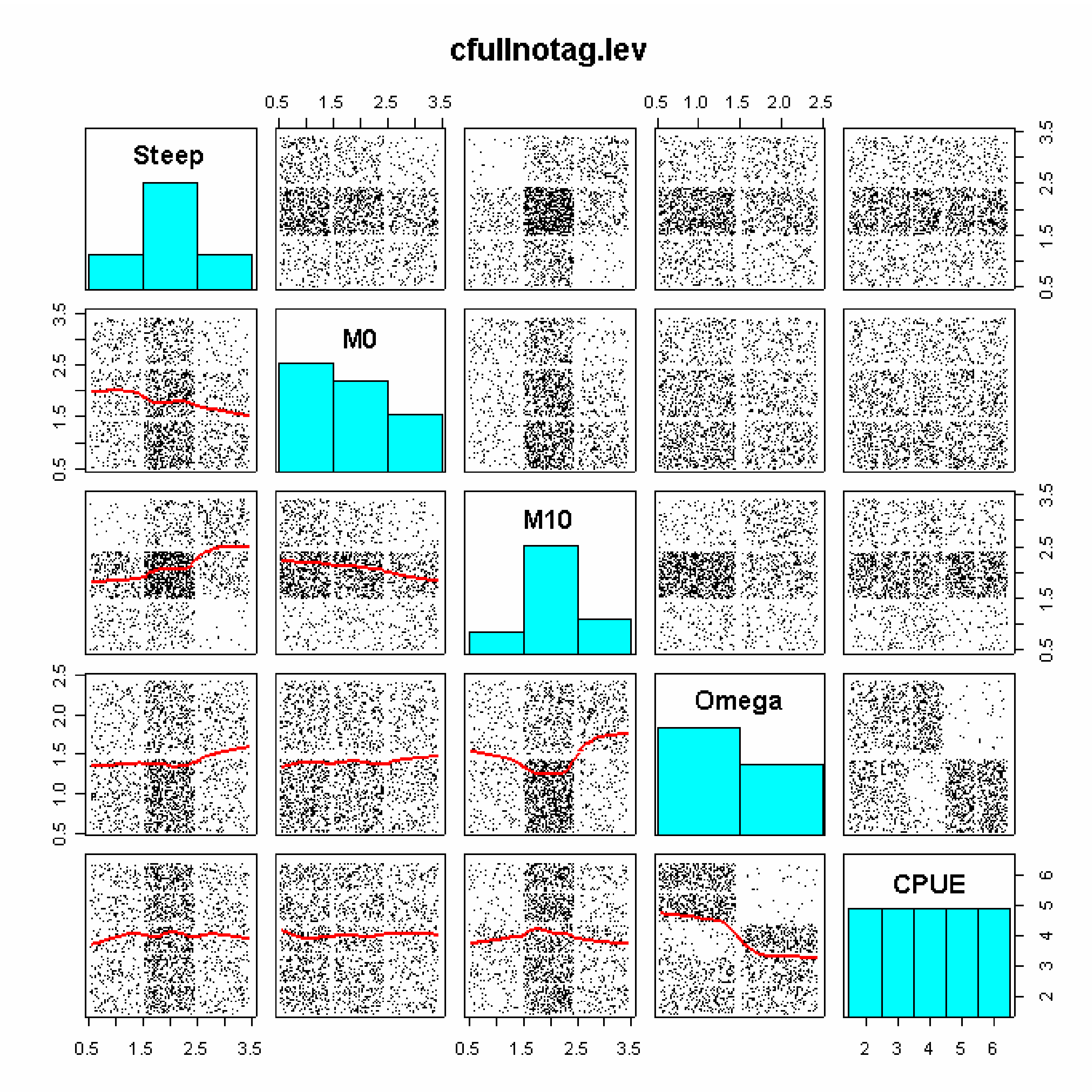


Figure 8. Pairwise plots of 2000 samples drawn from an MPD grid excluding the tagging data.

Sample sizes

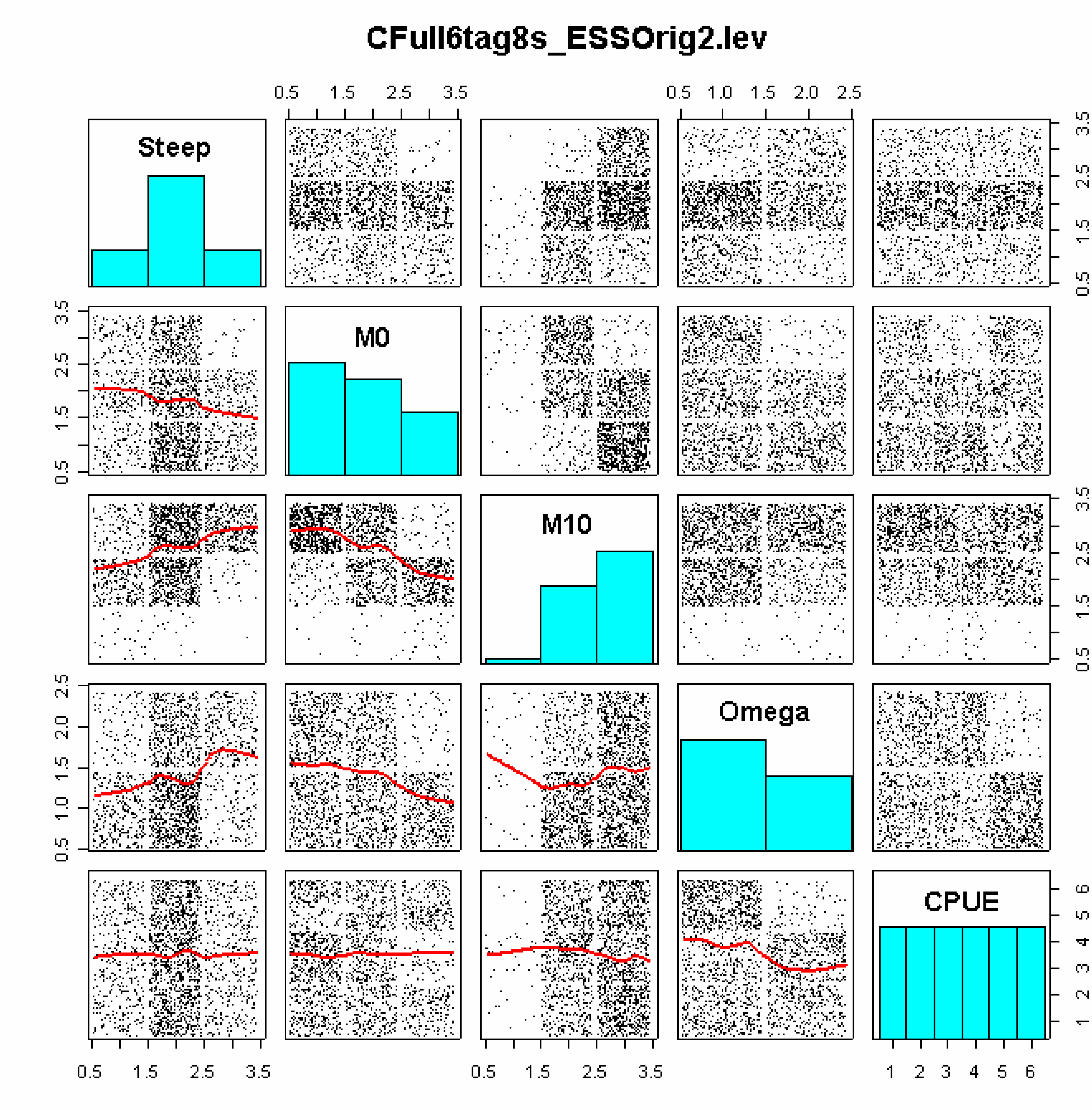


Figure 9. Pairwise plots of 2000 samples drawn from an MPD grid compiled for tag-reporting rate option 8 (updated prior to February 2005) and with the original sample sizes doubled.

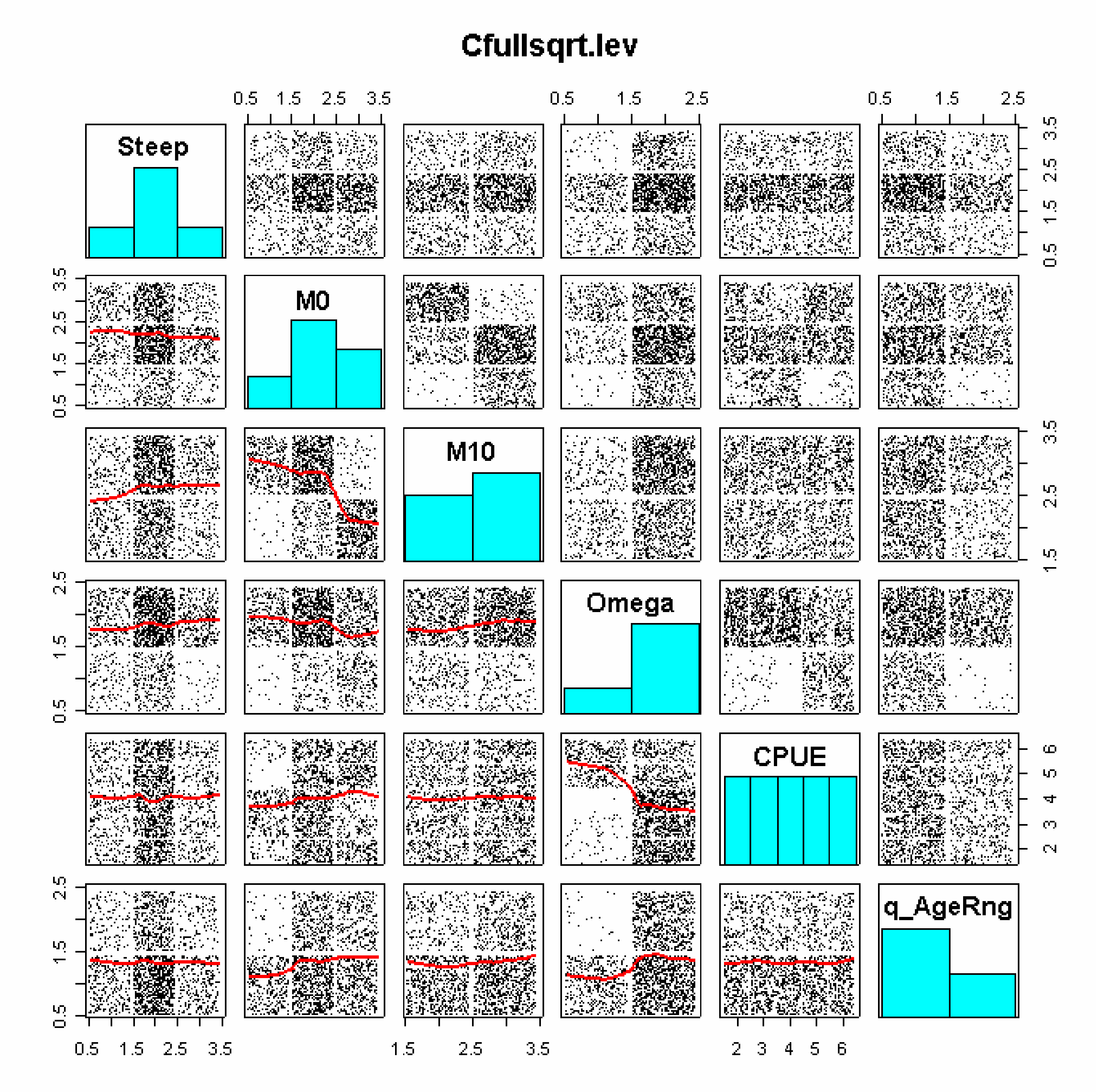


Figure 10. Pairwise plots of 2000 samples drawn from an MPD jgrid compiled for tag-reporting rate option 8 (updated prior to February 2005) and “sqrt” sample sizes proposed by the Panel in July 2004 and including two different age ranges for normalizing CPUE (4-18 and 8-12).

.

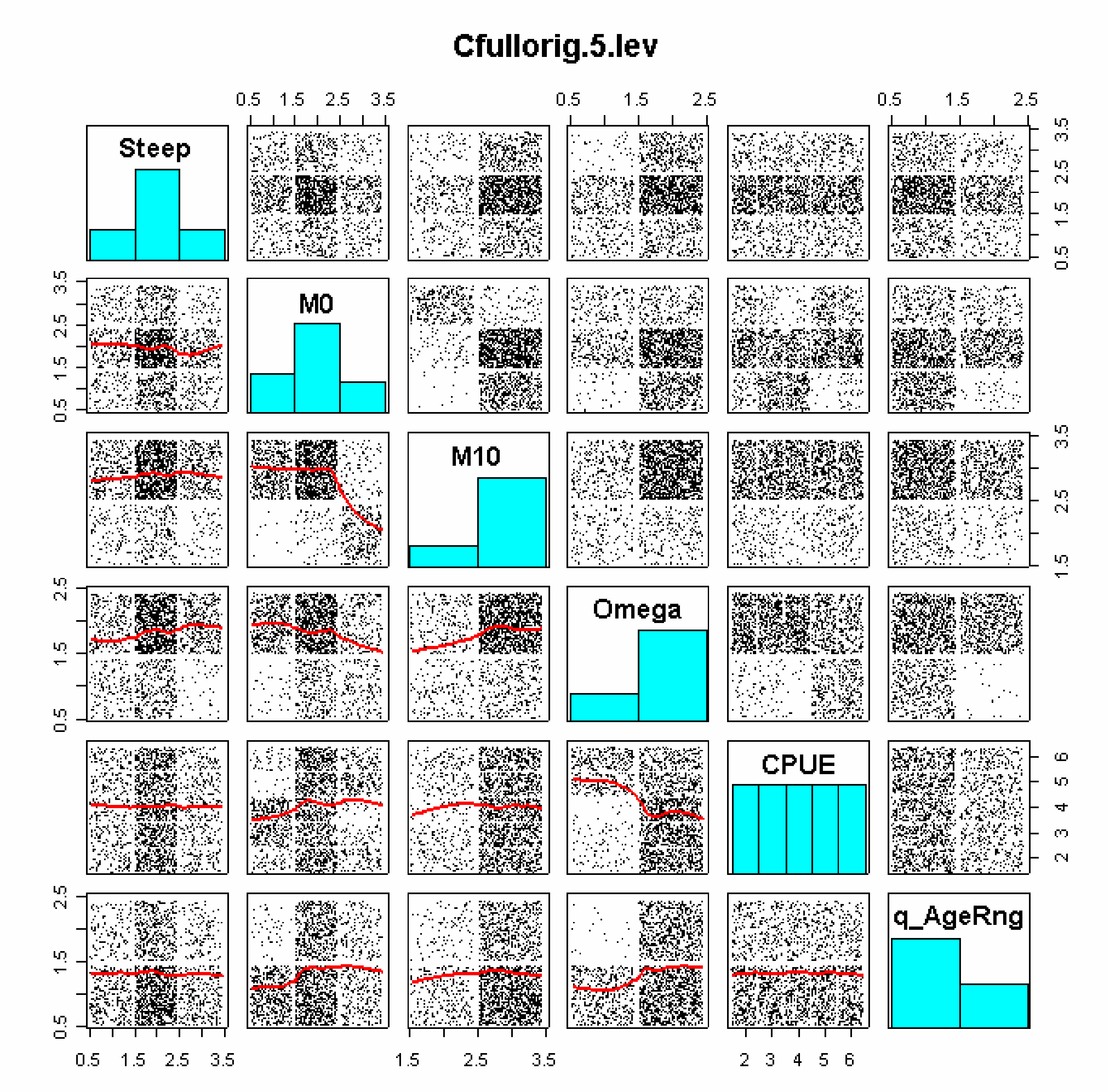


Figure 11. Pairwise plots of 2000 samples drawn from an MPD grid compiled for tag-reporting rate option 8 (updated prior to February 2005), with the original sample size specification multiplied by 0.5 and including two different age ranges for normalizing CPUE (4-18 and 8-12)..

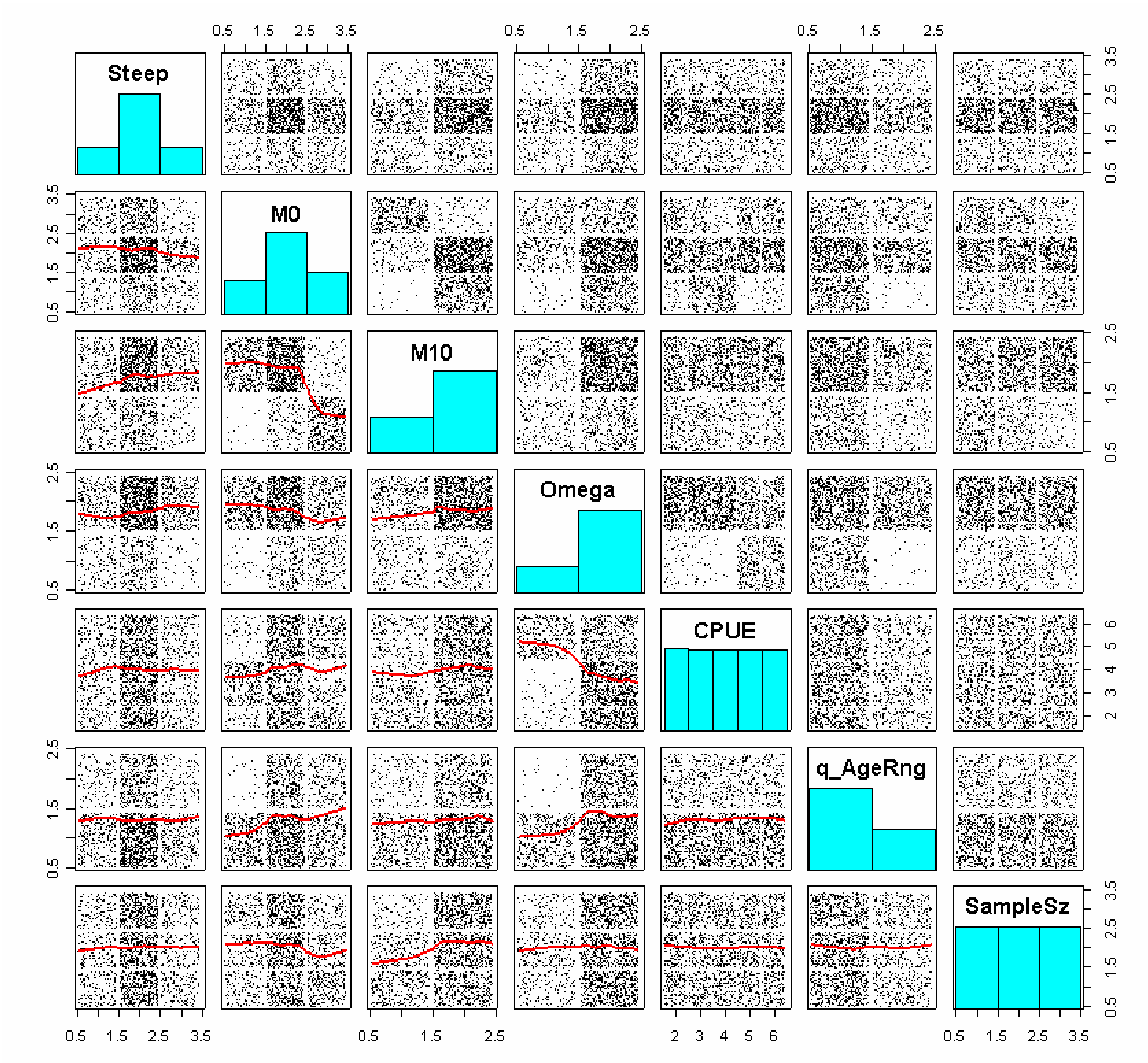


Figure 12. Pairwise plots of 2000 samples drawn from an MPD grid compiled for tag-reporting rate option 8 (updated prior to February 2005) and integrated over three sample-size scenarios and two CPUE catchability age ranges (4-18 and 8-12).

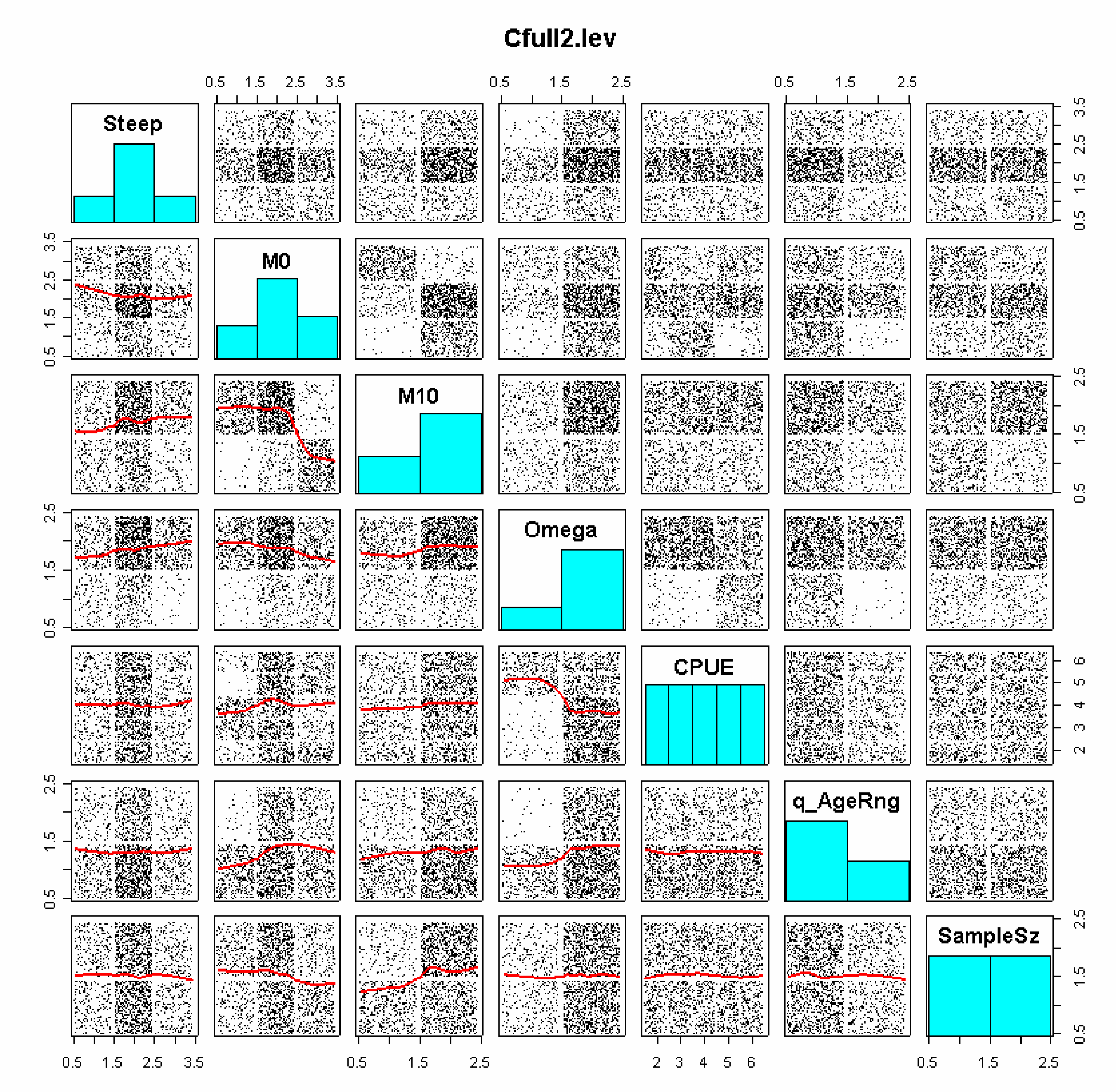


Figure 13. Pairwise plots of 2000 samples drawn from an MPD jgrid compiled for tag-reporting rate option 8 (updated prior to February 2005) and integrated over **two** sample-size scenarios and two CPUE catchability age ranges (4-18 and 8-12). **This was selected as the final reference set to be used for MP evaluations.**

Attachment 5

Configuration of operating model for the reference set

Specification of grid axes:

	Levels	Cumul N		Values		Prior	Simulation Weights
Steepness	3	3	0.385	0.55	0.73	0.2, 0.6, 0.2	Prior
M0	3	9	0.3	0.4	0.5	Uniform	Posterior
M10	2	18		0.1	0.14	Uniform	Posterior
Omega	2	36		0.75	1	0.4, 0.6	Posterior
CPUE	5	180				Uniform	Prior
q Age-range	2	360		4-18	8-12	0.67, 0.33	Prior
Sample Size	2	720		Sqrt	Original/2	Uniform	Prior

Note: when different series are used in conditioning, the MPs tested in projections will still use the median CPUE for the historical period.

Process for assigning weights and integrating over grid cells

The approach used to assign weights to the different cells differs depending on the grid axis, as detailed in the Table above. For some axes (M0, M10 and omega) the weights are based on Likelihood × prior. In other axes, the prior weights override the likelihood. The latter is the case of all factors related to changes in data input (CPUE series and sample sizes), and also the case of the steepness parameter. Given problems in model structure, the likelihood was not considered an adequate basis to assign weights to steepness. In particular, the lack of account for autocorrelation in recruitment in the likelihood and the use of a Beverton-Holt curve were discussed in connection to steepness.

Once weights are assigned, the code samples cells with probabilities based on these weights. In order to assure adequate coverage on the steepness, CPUE and sampling size axes, the number of realizations for each h ×CPUE×SS stratum is fixed and values for other axes are sampled within each stratum.

Sample sizes are computed as separate grids and outputs are combined in the *.grid file but they are not randomized. In case users want to use only a subset of the 2000 realizations (e.g. 500 to start tuning procedures), they will need to randomize the records to avoid biases due to the change in sample size assumptions.

Basic model assumptions

Projection period: 2004-2031 (last biomass computed for beginning of 2032)

Catch split: based on average of 2001-2003

Fisheries in projection:

1: LL1 fishery (second season).
2: LL2 fishery (second season).
3: Indonesian spawning fishery (first season).
4: Australian surface fishery (first season).

Sensitivity of conditioning to a number of model parameters was examined, including changing the weight given to tagging data, sample sizes for age/length composition data, age after which Indonesian selectivity is assumed constant, age range used for CPUE standardization, variability in Indonesian and LL1 selectivity. The reference set operating model is based on the following assumptions:

Name	Weight to tagging	Weight to Age/Len comp data	Age plus for Indonesian selectivity	Indonesian selectivity	CPUE age-range	Time-Variability LL 1 selectivity
Cfull	3.2	Reduced sample size (details below)	30	Variable	4-18	As in old reference set

Sample sizes for age/length composition data used in the old reference set were considered to be too large (e.g., *n*=500 for LL1 in final years) to be used in conjunction with constrained changes in selectivity as assumed in the model. Two different assumptions about sample sizes are included in the grid: (a) the sqrt weights proposed the Panel in the sets developed in July 2004 (CCSBT-ESC/0409/42) and (b) half the sample sizes in the old reference set. The sqrt weights were computed by taking sqrt of *n* times 5 for all fisheries and years. This reduces the contrast in sample sizes over time.

The code has a command line argument that changes the sample sizes used in projections. By default, the projection uses the final year sample sizes.

- samplescaler xx

where for xx>0, all age-comp sample sizes are scaled by the value xx .

Selectivities

Conditioning

- LL1 selectivity changes (CV=0.5) every 4 years, with change in 1997 and 2001 (last block is only three years).
- LL2, LL3 and LL4 are constant.
- Indonesian selectivity is constant up to 1996 when it starts changing every two years with CV=0.5.
- Australian selectivity changes (CV=2) by blocks of 4 years up to 1997 when it starts changing every year.

Projections

Randomness in future selectivities were added so that the age-composition data were not unrealistically informative. Random-walk processes as assumed in conditioning are not appropriate because they may result in the selectivities wandering off into implausible regions.

The following lognormal formulation is used for LL1 fishery (note that first subscript corresponds to fishery *f*=1):

$$s_{1,a,y} = s_{1,a,2003} e^{\varepsilon_{a,y}} \quad \text{for} \quad a_1^{\min s} \geq a \geq a_1^{\max s} \quad \text{where} \quad a_1^{\min s} = 2,\ a_1^{\max s} = 17$$

$$\varepsilon_{2,y} = \eta_{2,y}$$

$$\varepsilon_{a+1,y} = \rho_{\text{sel1}}\, \varepsilon_{a,y} + \sqrt{1-\rho_{\text{sel1}}^{\;2}}\, \eta_{a,y}, \qquad \text{where} \quad \eta_{a,y} \sim N(0,0.2^2) \quad \text{and} \quad \rho_{\text{sel1}} = 0.7$$

and selectivities only change every four years so that $s_{1,a,y+3} = s_{1,a,y+2} = s_{1,a,y+1} = s_{a,y}$, with change in 1997 and 2001 (last block is 3 yrs)

For the Australian surface fishery, lognormal variability combined with targeting on age 3 is assumed as follows:

Define

$$P_{3,y} = \frac{N_{3,y}}{\sum_{a=1}^{5} N_{a,y}} \quad \text{and} \quad \bar{P}_3 = \frac{1}{10}\sum_{y=1994}^{y=2003} P_{3,y}$$

If $P_{3,y} \geq \bar{P}_3$

$$s_{6,a,y} = s_{6,a,2003}e^{\varepsilon_{6,a,y}} \quad \text{for } a = 1,2,3,4,5 \text{ where } \varepsilon_{6,a,y} \sim N(0, 0.1^2)$$

Otherwise, increase selectivity of age 3:

$$s_{6,3,y} = s_{6,3,2003}e^{\varepsilon_{6,a,y}} \left(1 + 0.5\frac{\bar{P}_3 - P_{3,y}}{\bar{P}_3}\right)$$

$$s_{6,a,y} = s_{6,a,2003}e^{\varepsilon_{6,a,y}} \quad \text{for } a = 1,2,4,5$$

Natural Mortality

Natural mortality at ages 0 and 10 are included as grid axes. Mortality is assumed constant for ages 10 and older.

Stock-recruitment issues

Steepness

Conditioning and Projections

Steepness is included as one of the main grid axis, with values 0.385, 0.55, and 0.73.

Recent recruitments

Conditioning

The likelihood assumes no autocorrelation except for the last two years. The empirical autocorrelation in the recruitment residuals estimated over the period 1965-1998 is applied from 2001 onward.

Projection

Lognormal autocorrelated error is added to initial abundances (numbers at age 0 through 2 in 2004) within the projection code.

Let τ_y represent the lognormal recruitment deviate in year *y* and $\hat{\tau}_y$ its MPD estimate. The initial abundances passed from the conditioning code correspond to

1) $\hat{\tau}_{2001}$ estimated from model fit

$$\hat{\tau}_{2002} = \hat{\rho}\,\hat{\tau}_{2001}$$

$$\hat{\tau}_{2003} = \hat{\rho}^2\,\hat{\tau}_{2001}$$

$$\hat{\tau}_{2004} = \hat{\rho}^3\,\hat{\tau}_{2001}$$

where $\hat{\rho}$ is the empirical estimate of autocorrelation based on recruitments for years 1965-1998.

2) Stochastic projections

$$N_{2004,4} = \hat{N}_{2004,4}\exp\{0.4\,z - 0.08\}$$

$$N_{2004,3} = \hat{N}_{2004,3}\exp\{0.4\,z - 0.08\}$$

$$N_{2004,2} = \hat{N}_{2004,2}\exp\{\varepsilon_{2002}\}$$

$$N_{2004,1} = \hat{N}_{2004,1} \exp\{\hat{\rho}\,\varepsilon_{2002} + \varepsilon_{2003}\}$$
$$N_{2004,0} = \hat{N}_{2004,0} \exp\{\hat{\rho}^2\,\varepsilon_{2002} + \hat{\rho}\,\varepsilon_{2003} + \varepsilon_{2004}\}$$

where $z \sim N(0,1)$ and $\varepsilon_y \sim N\left(0,(1-\hat{\rho}^2)\sigma_R^2\right)$, where σ_R=0.6. Note that log-normal error with s.d.=0.4 has been added to account for uncertainty around $\hat{N}_{2000,0}$ and $\hat{N}_{2001,0}$.

These equations imply that:

$$\tau_{2004} = \hat{\tau}_{2004} + \hat{\rho}^2\,\varepsilon_{2002} + \hat{\rho}\,\varepsilon_{2003} + \varepsilon_{2004}$$

which is used to generate

$$\tau_{2005} = \hat{\rho}\tau_{2004} + \varepsilon_{2005}$$

and so on. This formulation amounts to assuming autocorrelated recruitment starting in 2002.

Treatment of σ_R

Conditioning

Initially σ_R was estimated with a lower bound of 0.40. Since the SAG of 2003, the value has been fixed at 0.6. This was originally done to help achieve a close-to-uniform MCMC distribution of *h* values within each bin.

Projections

Use value passed from conditioning.

Trends in carrying capacity

Conditioning

A suggestion was made that one reason many assessments show a low value of steepness may be attributed to changes in carrying capacity. It was suggested that the Aleutian low (i.e., large-scale climate/oceanographic regime shifts) may affect spawning grounds for SBT (but in a way that is not directly obvious). The shift was identified in 1977 and the suggestion was made to apply a different value for R_0 (stock-recruitment scale parameter), which would be estimated in the model-fitting process. Results of this conditioning trial resulted in an estimate of *h*=0.57 and a value of R_0 about half the value estimated for the earlier years. The workshop decided to maintain this run as a robustness test. Parameter values related to MSY, depletion, etc. will be computed using the set of parameters estimated for the most recent period.

Projections

Use stock recruitment parameters estimated for the most recent period and values of ρ and σ_R as specified for the baseline sets.

CPUE-abundance relationship

Catchability Model

Conditioning

The following model was proposed at the 7th SC meeting to link abundance with expected CPUE.

$$CPUE_y = q_y \tilde{N}_y^{\varpi} \left(1 + \beta \left(\frac{E_y - E_{2000}}{E_{2000}} \right) + \gamma \left(\frac{E_y - E_{2000}}{E_{2000}} \right)^2 \right)$$

$$\text{where} \quad \tilde{N}_y = \sum_a \left(\frac{s_{LL1,y,a}}{\frac{1}{(a_2 - a_1 + 1)} \sum_{j=a_1}^{j=a_2} s_{LL1,y,j}} \right)^{\psi} N_{y,a} \qquad (1)$$

$$\text{and} \quad E_y = \frac{C_{LL1,y}}{CPUE_y}$$

In this model, parameters $\beta, \gamma, \varpi, \psi, q_y$ and a_1 and a_2 are specified by the user. Current default values are: $\beta = 0$, $\gamma = 0$, $\omega = 1$, $\psi = 1$, $(a_1, a_2) = (4,18)$ or $(8,12)$.

Parameters β and γ : changing the values of β and γ had little or no effect in the conditioning (CCSBT-MP/0304/07).

Parameter ω : Is one of the axes in the grid, with values 1 and 0.75.

Parameters a_1 and a_2 (age range to standardize selectivity for CPUE predictions) are included as one grid axes with two alternative ranges: (1) a_1=4 and a_2=18 (2) a_1=8 and a_2=12. The rational for changing a_2 from 30 to 18 was that selectivities estimated for ages 19-30 are very low.

Projections

Same as in conditioning.

Trends in efficiency

Conditioning

The analyses looking at historical CPUE trend based on a linear increase (CCSBT-MP/0304/07) showed that no improvement was obtained by imposing this relationship. The CPUE working group recommended to include a test assuming a linear increase in catchability of 1% per year throughout the whole time series. This test was later dropped but an increase in *q* of 0.5 % a year (half way between Q0 and Q1) was kept in both the conditioning and in the projections in the core set.

Suggestions were made that catchability might best be modeled as a break-point (two periods, pre and post GPS/plotting). This was supported somewhat with the residual pattern.

Projections

The 0.5% annual increase in *q* is also applied in conditioning. CPUE is generated using autocorrelated trends in catchability, as estimated from conditioning. The empirical estimate of autocorrelation based on the entire time series (1969-2003) is used. For the sigma: use a value of 0.2 or the empirical estimate for the entire time series, whichever is largest. Alternatively, the user can select a value as a command option to the projection code by typing

```
-cpuestd xx
```

where for xx>=0, the value xx is the standard deviation of the log of the cpue residuals.

A sensitivity test assuming a 20% up and down change in q in year 2006 is included.

NB!: first int in *.grid file needs to be set by hand to 1 (20% increase in q) or –1 (20% decrease in q) instead of the default 0.

Growth

Projections:

Size and weight at age assumed constant over time but different for the four fisheries. Data are input in fixed_quants04.

Errors in catches

Conditioning:

Core set assumes no errors in catches. A robustness test is done with actual catches (5% for 1969-1990 and 15% for $y \geq 1991$) larger than those in the input file (controlled by a switch in *.dat file).

Projections:

- Core set assumes TAC=catch.
- When underreporting of catches is assumed in the conditioning, it is also assumed in projections (actual catch is 15% larger than TAC). Only reported catches are assumed known by the MPs. In other words, the MP does not know the "true" historical catch vectors used for conditioning and the simulated actual catches.

(NB!: a switch needs to be set by hand in first line of *.grid file to run projections for this robustness test: set second int to 1)

Fishing Mortality Specifications

The fishing mortality specifications in the original model were based on:

Note: year subscripts omitted for simplicity

$$C = \sum_f \sum_a s_{f,a} F_f N_a \qquad (1)$$

$$C = \sum_f C_f \quad \text{so that} \quad C_f = (\sum_a s_{f,a} N_a) F_f \qquad (2)$$

$$\text{and} \quad F_f = \frac{C_f}{\sum_a s_{f,a} N_a} \qquad (3)$$

Note also:

$$C_{f,a} = s_{f,a} F_f N_a \qquad (4)$$

$$C_a = (\sum_f s_{f,a} F_f) N_a \qquad (5)$$

Problem arises if $(\sum_f s_{f,a} F_f) > 1$ so that $C_a > N_a$.

Age-specific exploitation rates $(\sum_f s_{f,a} F_f)$ were bounded at 0.99. When the bound was exceeded, the catch at age for the fisheries involved was reduced to meet the bound but the exploitation rates of the other ages were not adjusted. This may lead to unnecessary reductions of catches in cases when the TAC could have been taken if selectivities of the other ages had been increased.

A formulation proposed by Doug Butterworth was used to improve performance. The formulation uses finite harvest rates and adjusts selectivities in the projection model to try to meet fishery-specific TACs without exceeding the bounds on harvest rates.

Case of one fleet (or non-overlapping selectivities):

Consider the single-fleet case, so omit *f* subscript:

Compute F using equation (3) above; if $F \leq 0.9$, no change

If $F > 0.9$, then:

$$C = \sum_a g(s_a F) N_a \quad (6)$$

$$C = \sum_a s_a^* F \; N_a \quad \text{where modified selectivity} \quad s_a^* = \frac{g(s_a F)}{F} \quad (7)$$

$$\text{Propose} \quad g(x) = \begin{cases} x & x \leq 0.9 \\ 0.9 + 0.1[1 - \exp(-10(x-0.9))] & 0.9 < x \leq \infty \end{cases} \quad (8)$$

Note:

(i) $g(x) < 1$, hence: $C_a = g(s_a F) N_a < N_a$ as required.

(ii) $g(x)$ is continuous and derivative-continuous at x=0.9

A process such as Newton-Raphson is used to solve equation (6) for F and hence compute $C_a = g(s_a F) N_a$.

Extension to more than one fleet

If from equation (3) $\sum_f s_{f,a} F_f < 0.9$ for all ages, then equations (3)-(5) remain. If $\sum_f s_{f,a} F_f > 0.9$ for any age, then

$$C = \sum_a g(\sum_f s_{f,a} F_f) N_a \quad (9)$$

where $g(x)$ as above, so that $C_a = g(\sum_f s_{f,a} F_f) N_a < N_a$ as required.

Assume farther that effective proportional reduction of selectivity for each fleet at a certain age *a* is the same for each fleet (but differs by age). Then the modified selectivity $s^*_{f,a}$ is given by:

$$s^*_{f,a} = s_{f,a} \left[\frac{g(\sum_{f'} s_{f',a} F_{f'})}{\sum_{f'} s_{f',a} F_{f'}} \right] \quad (10)$$

Then
$$C_{f,a} = s^*_{f,a} F_f N_a = s_{f,a} \left[\frac{g(\sum_{f'} s_{f',a} F_{f'})}{\sum_{f'} s_{f',a} F_{f'}} \right] N_a \quad (11)$$

Then
$$C_a = \sum_f C_{f,a} = \sum_f s_{f,a} F_f \left[\frac{g(\sum_{f'} s_{f',a} F_{f'})}{\sum_{f'} s_{f',a} F_{f'}} \right] N_a \quad \text{as required by (9)}$$

$$= g(\sum_f s_{f,a} F_f) N_a$$

Thus, a multivariate root finding process (e.g. extended Newton-Raphson) is needed to solve for F_f in the following coupled non-linear differential equations for $f = f_1, f_2, f_3 \ldots$:

$$C_f = \sum_a C_{f,a} = \sum_a s^*_{f,a} F_f N_a$$

i.e. $$C_f = \sum_a s_{f,a} F_f \left[\frac{g(\sum_{f'} s_{f',a} F_{f'})}{\sum_{f'} s_{f',a} F_{f'}} \right] N_a$$

Data available to Management Procedures

The following data are assumed to be available to the MPs at the time *t* when a TAC is computed (*t* is the "decision year" in Table 1):

Catch data from OM up to year *t*-1

TAC up to year *t*

CPUE up to year *t*-1

Age-composition up to year *t*-1

The TAC computed in year *t* is implemented in year *t*+2.

Assumptions made for generating the data are detailed below:

Catch

Historic catch statistics as used for conditioning and all previous TACs set by the MP.

CPUE

Historic median CPUE (even when other CPUE series are used for conditioning) and CPUE simulated according to specifications above (see CPUE-abundance relationships).

Age-composition data

Age composition data are simulated using cohort-slicing. Cohort-sliced data is also provided for the historical period to be used by the MPs.

Attachment 6

Robustness MP trials

Trials		Number of trials	Full integration	Grid cells	Number of replicates
Tagging data	Exclude tagging data	1		Sqrt grid with 3 M10 levels (i.e including 0.07)	2000
Recruitment	No AC	1	×		2000
	Additional years of low recruitment (a)2002-2003, (b) 2002-2005 (details below)	2	×		2000
	Increase $N_{2004,3}$ and $N_{2004,4}$ by a factor of 3	1	×		2000
Maximum exploitation rate for Surface fishery	See details below	1	×		2000
Indonesian selectivity	Max estimated age = 18	2		M10 high M0 central h low and central Omega =1 CPUE median Sqrt sample sizes, _a 4-18	500
CPUE	Catchability up	2		M10 central M0 central h low and central Omega =1 CPUE median Sqrt sample sizes, _a 4-18	500
Carrying capacity	Carrying capacity change	2		M10 central M0 central h low and central Omega =1 CPUE median Sqrt sample sizes, _a 4-18	500

Uncertainty in catches	Uncertainty in catches	2		M10 central M0 central h low and central Omega =1 CPUE median Sqrt sample sizes, _a 4-18	500

For comparison with the robustness trials, the following sets will be provided:

- a Sqrt grid run with tagging, i.e. a subset of the reference. Note that M10=0.07 was excluded from the reference set because it had a very low plausibility when tagging data are included.
- Single cells matching those in the 5th column.

Additional years of low recruitment

Let R_{low}= average of $\hat{N}_{2000,0}$ and $\hat{N}_{2001,0}$, then two robustness trials are defined as:

(a) Assume R for 2002-2003 = R_{low} , then autocorrelated stochastic starting in 2005 (to run set switch to 2 in mycontrol.dat).

(b) Assume R for 2002-2005 = $R_{\text{low},}$ then autocorrelated stochastic starting in 2006 (to run set switch to 4 in mycontrol.dat).

Note: averaging of 2000 and 2001 recruitments can be done in the projection code as full vector of recruitments is passed.

$$N_{2004,3} = \text{ unchanged}$$

$$N_{2004,2} = \frac{\hat{N}_{2004,2}}{\hat{N}_{2002,0}} R_{\text{low}}$$

$$N_{2004,1} = \frac{\hat{N}_{2004,1}}{\hat{N}_{2003,0}} R_{\text{low}}$$

$N_{2004,0} = R_{\text{low}}$, same for $N_{2005,0}$, etc. depending on number of years.

This formulation involves holding constant the fishing mortality of one year olds in 2003. The code is general and the number of years with low recruitment can be changed by the user. Details about the code are shown below in the section on Code Details.

Maximum exploitation rate for surface fishery

Current exploitation rates for the surface fishery simulated by the reference set were thought to be possibly too high relative to maximum exploitation rates estimated for 1984-1988. A new robustness trial was designed to strongly penalize the exploitation rate for age 3 in 2003 and 2004 when it exceeds 80% of the average exploitation rate for ages 2 to 3 in 1984-1988. To implement this trial, the conditioning code needs to be modified by including a penalty into the objective function. Let

$$H_{s,\text{high}} = \frac{1}{5}\sum_{y=1984}^{y=1988} \frac{\sum_{a=2}^{a=3} C_{s,y,a}}{\sum_{a=2}^{a=3} N_{y,a}} \qquad \text{and} \qquad H_{s,y,3} = \frac{C_{s,y,3}}{N_{y,3}},$$

then the penalty added to the objective function would be:

$$\text{Penalty} = \begin{cases} 0 & \text{for} \quad H_{s,2004,3} < 0.80\, H_{s,\text{high}} \\ \displaystyle\sum_{y=2003}^{y=2004} \frac{(H_{s,y,3} - 0.8\, H_{s,\text{high}})^2}{0.02^2} & \text{for} \quad H_{s,2004,3} \geq 0.80\, H_{s,\text{high}} \end{cases}$$

Code Details

When the number of years with low recruitment is set to 4, (year classes 2002-2005), the initial abundances are set without random error using the equations in the text. Random autocorrelated error starts in 2006. For the general code where the number of years *n* is specified by the user, autocorrelation always starts the year after the block of low recruitments. As long as *n*>=3 the formulation is the same. In order to get the autocorrelated residuals we need to compute the recruitment dev (tau) to be used as predictor, given by

Tau(y) = ln(Rlow) –ln(R predicted from S)

If n=3 and thus R_{2004} = Rlow, the tau(first_yr) needs to be set based on the recruitment dev of 2004,

Tau(first_yr) = ln(Rlow) – $\ln(\hat{N}_{2004,0})$

For n=1 and 2 the low recruitment only affects some of the initial abundances set in get_init_pop, so it is messy.

For n=2,

$$N_{2004,2} = \frac{\hat{N}_{2004,2}}{\hat{N}_{2002,0}} R_{\text{low}}$$

$$N_{2004,1} = \frac{\hat{N}_{2004,1}}{\hat{N}_{2003,0}} R_{\text{low}}$$

and stochastic $N_{2004,0} = \hat{N}_{2004,0} \exp\left\{\hat{\rho}\tau_{2003} + \varepsilon_{2004}\right\}$ where $\tau_{2003} = \ln(R_{low}) - \ln(\hat{N}_{2003,0})$

For n = 1

$$N_{2004,2} = \frac{\hat{N}_{2004,2}}{\hat{N}_{2002,0}} R_{\text{low}}$$

and stochastic $N_{2004,1} = \hat{N}_{2004,1} \exp\left\{\hat{\rho}\tau_{2002} + \varepsilon_{2003}\right\}$ where $\tau_{2002} = \ln(R_{low}) - \ln(\hat{N}_{2002,0})$

and stochastic $N_{2004,0} = \hat{N}_{2004,0} \exp\left\{\hat{\rho}\tau_{2003} + \varepsilon_{2004}\right\}$ where $\tau_{2003} = \hat{\rho}\tau_{2002} + \varepsilon_{2003}$

Tau(first_yr) needs to be set using the deviate for $N_{2004,0}$.

Attachment 7

List of performance statistics

Let Cy be the total catch in year y and $C_{\text{surf},y}$ the surface fishery catch in year y

(1) Average catch:

(1.1) $\frac{\sum_{2004}^{2004+4} C_y}{5}$ (1.2) $\frac{\sum_{2004}^{2004+9} C_y}{10}$ (1.3) $\frac{\sum_{2004}^{2004+19} C_y}{20}$ (1.4) $\frac{\sum_{2004}^{2031} C_y}{28}$

(2) Average ratio of surface to total catch:

$$\frac{1}{28}\sum_{2004}^{2031}\frac{C_{\text{surf},y}}{C_y}$$

(3) Ratio of biomass (S: spawning biomass) to initial biomass

(3.1) $\frac{S_{2004+5}}{S_{2004}}$ (3.2) $\frac{S_{2004+10}}{S_{2004}}$ (3.3) $\frac{S_{2004+20}}{S_{2004}}$ (3.4) $\frac{S_{2032}}{S_{2004}}$

(4) proportion of projections in which spawning biomass drops below 90% of 2004 level any time in the projection.

(5) proportion of projections in which spawning biomass drops below 80% of 2004 level any time in the projection.

(6) proportion of projections in which spawning biomass is below 90% of 2004 at the end of the projection period in 2032.

(7) proportion of projections in which spawning biomass is below 80% of 2004 at the end of the projection period in 2032.

(8) Minimum spawning biomass relative to current:

$$Min\left\{\frac{S_y}{S_{2004}}\right\} \text{ over 29-year projections}$$

(9) Spawning stock biomass in 2022 relative to what it would have been in the absence of fishing:

$$S_{2022}/S^*_{2022}$$

where S^*_{2022} is the spawning biomass in 2022 under a no catch scenario.

NOTE: This statistic will not be part of the summary output produced by the projection code, but will be computed as part of the new graphics package.

(10) Ratio of spawning biomass in 2020 and 2032 to 1980 biomass:

(10.1) $\frac{S_{2020}}{S_{1980}}$ (10.2) $\frac{S_{2032}}{S_{1980}}$

(11) Ratio of spawning biomass in 2020 and 2032 to biomass at MSY:

(11.1) $\frac{S_{2020}}{S_{MSY}}$ (11.2) $\frac{S_{2032}}{S_{MSY}}$

(12) Inter-annual variations in catches:

$$AAV = \frac{1}{26}\sum_{2005}^{2030} \frac{|C_{y+1} - C_y|}{C_y + 1^{-6}}$$

(13) Exploitation rate relative to MSY:

defined as Catch-to-total biomass ratio:

$$\frac{1}{5}\sum_{y=2027}^{2031} \frac{C_y}{\text{Total Biomass}_y} \quad \text{relative to} \quad \frac{C_{MSY}}{\text{Total Biomass at MSY}}$$

The latter is formulated in terms of the ratio between catch and total biomass (age 2 and older) over the last 5 years in the simulation versus the ratio of MSY catch to biomass (age 2 and older) to avoid the difficulties associated with the appropriate definition of fishing mortality when selectivities are changing. The group also noted that there are potential difficulties with regard to interpretation of these measures when selectivities change greatly, and/or if the split between the surface and longline catch changes from the values used in the MSY calculations. Note that the above implies computing the MSY and the total biomass (age 2 and older) at MSY for the different conditioning scenarios. This would be done using the most recent weights at age and selectivities at age.

(14) Whether the TAC trajectories change direction in the early years, with the notion that one did not want the TAC to first increase and then decrease or vice versa over the first 6 years. i.e., avoid situations where TAC_{2009} lies outside the range of TAC_{2006} and TAC_{2012} (low *A* desired). The statistics chosen reflects the probability of TAC going in the "wrong" direction as well as a measure of the extent of such changes.

If *n* replicates of a trial are conducted:

$$(14) \qquad A = \frac{1}{n}\sum_{i=1}^{n} (\Delta TAC)_i I_i$$

where $\Delta TAC_i = |TAC_{2009} - TAC_{2005}|$ and

$$I_i = \begin{cases} 0 & TAC_{2006} < TAC_{2009} < TAC_{2012} \\ 0 & TAC_{2006} > TAC_{2009} > TAC_{2012} \\ 1 & \text{otherwise} \end{cases}$$

low values of *A* are desired. This statistics is defined with respect to the first year when TAC can change. The specification above corresponds to option (a), in which 2006 is the first year when TAC can change. For option (b) the years are 2008, 2011 and 2014; for option (c) the years are 2008, 2013 and 2018. Furthermore, ΔTAC is calculated between the middle year (2009, 2011 or 2013) and 2005.

(15) Stability of TACs:

$$\text{Number of years when } \lfloor \Delta TAC_{y-1} \times \Delta TAC_y < 0 \rfloor \;\; (\Delta\, TAC_y = TAC_y - TAC_{y-1})$$

where all years with no change in TAC are ignored. This statistic evaluates the number of times TAC changes go in opposite direction in consecutive years.

(16) Consistency in the trends in biomass with those in the TACs:

$$\text{Number of years when } \lfloor \Delta S_y \times \Delta TAC_y < 0 \rfloor$$

It was acknowledged that this performance statistic would be difficult to interpret and would have to be viewed in conjunction with other performance statistics.

(17) Maximum decrease in TAC:

$$Min \lfloor \Delta TAC_y \rfloor$$

and for low *h* scenarios:

(18) Pr[(the slope of the regression of S_y versus time over the last five years)>0].

(19) To take into account industry concerns,

$$\text{Min}(CPUE_y)/CPUE_{2004}$$

NOTE:

Summary statistics (median, 10th and 90th quantiles) will be computed for statistics: (1)-(3), (8)-(13), (15)-(17) and (19).

Statistics on Risk:

An additional measure of risk from low spawning biomass proposed at the MPW3 , the following measure of risk from low spawning biomass was proposed. The statistic will be implemented as part of a the graphics package with parameters to be explored.

1. Calculate the spawning stock biomass as a fraction of a reference value, either estimated virgin biomass, B1980, B2004 or BMSY. We call this fraction value for year *y* F_y.

2. Specify a "risk threshold" that is a value below which we have concern about risk of possible recruitment collapse. Call this threshold T.

3. Calculate the average annual risk as

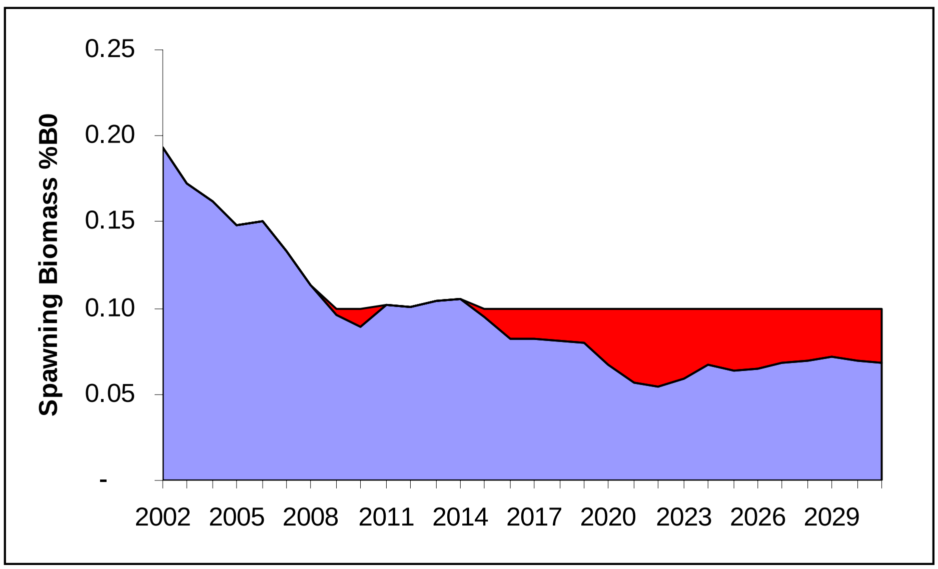
$$R = \frac{1}{n}\sum_{y=1}^{y=n}\left[u_y \left(\frac{T - F_y}{T} \right)^{\gamma} \right]$$

$$u_y = 0 \quad \text{if } T < F_y$$

$$u_y = 1 \quad \text{if } T \geq F_y$$

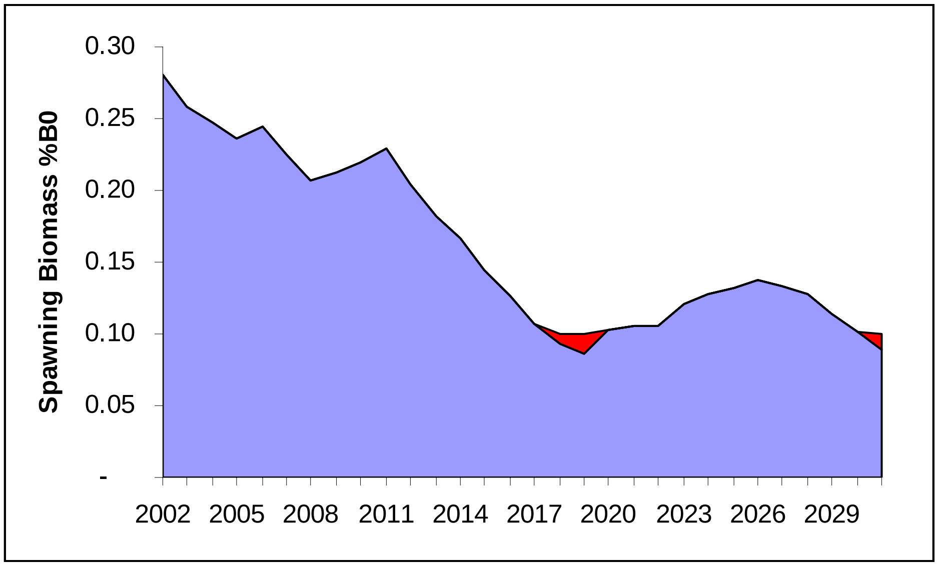
Thus the value of R can be interpreted as the fraction of the total possible risk that would occur if there was zero spawning biomass over the entire trajectory. The parameter γ controls the weight that different levels of biomass reduction have on the average, e.g. values larger than 1 weigh the larger reductions more heavily.

The two figures below illustrate graphically this area of "risk" and how it might differ in two scenarios of spawning biomass. The red area is the area of risk and the risk statistic R is a simply the average size of the area. In this scenario the value of R is 0.17.



In the scenario below the value of R is 0.011.

NOTE: This statistic will not be part of the summary output produced by the projection code, but will be computed as part of the graphics package.



The above definition involves two decisions, what reference value to use, and what threshold value.

Attachment 8

Options for Schedule of TAC changes

Decision rules are tuned to achieve three median rebuilding targets in year 2022: 0.9, 1.1 and 1.3 of current spawning biomass. The control file provides three options for how frequently TACs can be changed:

Option (a): first TAC for 2006, then every 3 years
Option (b): first TAC in 2008, then every 3 years.
Option (c): first TAC in 2008, then every five years.

These options include one extra year of lag between the year of decision when a TAC is computed and the year of implementation as requested by CCSBT after October 2003. Option (b) will be run for all three tuning levels (median 0.9, 1.1 and 1.3) while (a) and (c) will be run only for 1.1 tuning.

The maximum and minimum changes in TAC are:

options (a) and (b): 5000t and 100t;

option (c): 8000t and 100t.

Maximum TAC changes are not hardwired so users have the flexibility to explore other options in addition to the ones agreed upon. In the past, it was found that the maximum allowed changes appear to be real constraints so that for high tuning levels, there was little option but to cut TAC near to their maximum.

Decision year	Year of availability			Option a		Option b		Option c	
	Catch Data from operating model	Anticipated catches from TACs	CPUE Data from operating model	MP TAC Year	TAC Change allowed?	MP TAC Year	TAC Change allowed?	MP TAC Year	TAC Change allowed?
2004	2003	2004	2003*	2006	Yes	2006	no	2006	no
2005	2004	2005	2004	2007	no	2007	no	2007	no
2006	2005	2006	2005	2008	no	2008	Yes	2008	Yes
2007	2006	2007	2006	2009	Yes	2009	no	2009	No
2008	2007	2008	2007	2010	no	2010	no	2010	No
2009	2008	2009	2008	2011	no	2011	Yes	2011	No
2010	2009	2010	2009	2012	Yes	2012	no	2012	No
2011	2010	2011	2010	2013	no	2013	no	2013	Yes
2012	2011	2012	2011	2014	no	2014	Yes	2014	No
2013	2012	2013	2012	2015	Yes	2015	no	2015	No
2014	2013	2014	2013	2016	no	2016	no	2016	No
2015	2014	2015	2014	2017	no	2017	Yes	2017	No
2016	2015	2016	2015	2018	Yes	2018	no	2018	Yes
2017	2016	2017	2016	2019	no	2019	no	2019	No
2018	2017	2018	2017	2020	no	2020	Yes	2020	No
2019	2018	2019	2018	2021	Yes	2021	no	2021	No
2020	2019	2020	2019	2022	no	2022	no	2022	No
2021	2020	2021	2020	2023	no	2023	Yes	2023	Yes
2022	2021	2022	2021	2024	Yes	2024	no	2024	No
2023	2022	2023	2022	2025	no	2025	no	2025	No
2024	2023	2024	2023	2026	no	2026	Yes	2026	No
2025	2024	2025	2024	2027	Yes	2027	no	2027	No
2026	2025	2026	2025	2028	no	2028	no	2028	Yes
2027	2026	2027	2026	2029	no	2029	Yes	2029	No
2028	2027	2028	2027	2030	Yes	2030	no	2030	No
2029	2028	2029	2028	2031	no	2031	no	2031	No
2030	2029	2030	2029						No
2031	2030	2031	2030						

* For 2003, the actual CPUE value is used, not the CPUE value simulated by the operating model.