

Appendix 1

AGENDA

1. Introductory items
 - 1.1 Pre-meeting meeting
 - 1.2 Election of Chair, appointment of rapporteurs
 - 1.3 Adoption of Agenda
 - 1.4 Review of documents
2. Revised management procedure (RMP) – general issues
 - 2.1 Finalise the guidelines and requirements for implementing the RMP
 - 2.1.1. Develop the thresholds for defining 'acceptable' and borderline performance for classifying the performance of RMP variants for *Implementation Simulation Trials*
 - 2.1.2. Develop a list of agreed stock structure archetypes
 - 2.2 Proposal for revision of the CLA
 - 2.2.1 General approach to such considerations
 - 2.2.2 Consideration of Norwegian proposal
 - 2.3 Work plan
3. RMP – preparation for implementation
 - 3.1 Western North Pacific Bryde's whales
 - 3.1.1 Report from the First Intersessional *Implementation* Workshop for western North Pacific Bryde's whales
 - 3.1.2 Objectives of First Annual Meeting
 - 3.1.3 Review results of conditioning
 - 3.1.4 Updates to standard data sets
 - 3.1.4.1 Abundance
 - 3.1.4.2 Catches and other anthropogenic removals
 - 3.1.5 Final consideration of plausibility (incl. weighting of trials in terms of overall balance)
 - 3.1.5.1 Stock structure hypotheses
 - 3.1.5.2 Other (e.g. MSYR)
 - 3.1.6 Data/research to reduce hypotheses
 - 3.1.7 Specification of operational features and management variants
 - 3.1.8 Specification and classification of final trials
 - 3.1.9 Work plan
 - 3.2 North Atlantic fin whales
 - 3.2.1 Report of the joint NAMMCO/IWC workshop
 - 3.2.2 Report of the intersessional working group
 - 3.2.3 Progress on completion of the pre-implementation assessment
 - 3.2.3.1 Plausible stock structure hypotheses
 - 3.2.3.2 Preliminary investigation of experimental ways to distinguish among competing hypotheses
 - 3.2.3.3 Dispersal rates
 - 3.2.3.4 Abundance estimates (incl. $g(0)$ issues and plans for future surveys)
 - 3.2.3.5 Catch data (including alternative series)
 - 3.2.3.6 Future whaling operations
 - 3.2.3.7 Other anthropogenic removals
 - 3.2.3.8 Conditioning (incl. biological parameters)
 - 3.2.4 Work plan
4. Consideration of surveys under the guidelines and requirements
 - 4.1 Norwegian surveys
 - 4.2 Other
5. Other business
6. Adoption of report

Appendix 2

Evaluating criteria for defining conservation performance for *Implementation Simulation Trials*

ANDRE E. PUNT¹ AND CHERRY ALLISON²

IWC (2006) proposed the following steps to determine the conservation performance of an RMP variant for each stock in an *Implementation Simulation Trial* for which $MSYR=1\%$:

- (1) Construct a single stock trial, which is 'equivalent' to the *Implementation Simulation Trial*. For example, if a particular *Implementation Simulation Trial* involved carrying capacity halving over the 100-year projection period, the 'equivalent single stock trial' will also involve carrying capacity halving over the next 100 years.
- (2) Conduct two sets of 100 simulations based on this single stock trial in which future catch limits are set by the CLA. The two sets of simulations correspond to the 0.6 and 0.72 tunings of the CLA. Rather than basing these calculations on a single initial depletion, the simulations for each stock to be conducted for the set of initial depletions for the stock concerned in the *Implementation Simulation Trial*.
- (3) The cumulative distributions for the final depletion and for the depletion ratio (the minimum over each of the 100 years of the ratio of the population size to that when there are only incidental catches) to be constructed for each tuning of the CLA.

The lower 5%-ile of these distributions to form the basis for determining whether the performance of the RMP for the *Implementation Simulation Trial* is 'acceptable', 'borderline' or 'unacceptable'.

Fig. 1 outlines these steps for a hypothetical case. Fig. 1(a) shows the cumulative distribution of the initial depletion in the *Implementation Simulation Trial*. Figs 1(b) and (c) show cumulative distributions for the final depletion and the depletion ratio. These figures are constructed by running single-stock trials where the initial depletions are set to those values on which Fig. 1(a) was based. The horizontal lines in Figs 1(b) and (1c) are used to compute the thresholds that define 'acceptable', 'borderline' or 'unacceptable' performance. Note that the performance of an RMP variant for a specific *Implementation Simulation Trial* is 'acceptable' if either the final depletion or the depletion ratio satisfies the criteria for 'acceptable', and similarly for 'borderline' (Table 1).

At its 2005 meeting, the Committee requested that three analyses be conducted to evaluate the approach outlined above. The results of this work done intersessionally are reported in SC/58/RMP2 and described below.

CUMULATIVE DISTRIBUTIONS FOR SINGLE-STOCK TRIALS

Fig. 2 shows cumulative distributions for the initial depletion, the final depletion, and the depletion ratio for six initial depletion distributions (two cases for mean initial depletions of 0.3, 0.5 and 0.7) to illustrate the sensitivity of the thresholds. The difference between the two cases is the inter-simulation variation in the initial depletion (standard deviations of 0 and 0.05; 0.05 was chosen because it is large enough to differ from 0 and small enough that the impact of the mean initial depletion is still evident). Table 2 lists the values for the thresholds. Note that the initial depletion for the trial in which there is only a single initial depletion is set equal to the mean of the initial depletions for the trial in which initial depletion varied among simulations.

The values for the thresholds that define 'acceptable' and 'borderline' performance for the two performance statistics are, as expected, most sensitive to the initial depletion level. However, the values of these thresholds are fairly insensitive to variation in the initial depletion, except when the initial depletion is low (~0.3). This result is not unexpected given that the Standard Deviation of the initial depletion distribution was pre-specified when constructing the trials for which results are reported in Fig. 2 and Table 2.

APPLICATION TO IMPLEMENTATION SIMULATION TRIALS DEVELOPED FOR NORTH ATLANTIC MINKE WHALES

Table 3 lists the initial depletions for all the stocks in all the North Atlantic minke whale trials for which $MSYR=1\%$, the resulting thresholds for the final depletion and depletion ratio statistics, the lower 5th percentiles for the final depletion and depletion ratio based on the actual trials, and the net result of using the criteria developed by IWC (2006) to evaluate conservation performance. The catch limits for the North Atlantic minke whale trials are set by catch-cascading over the Central and Eastern *Medium Areas* (IWC, 1994). These results of the North Atlantic minke trials are not identical to those in IWC (1993) because, for consistency with the single-stock trials, all of the calculations are based on the CATCHLIMIT program. IWC (2006) agreed that this program would be used in simulation trials owing to improvements in computing speed. However, for completeness, the results based on the old *CLA* program are included in parentheses in Table 3. As expected, differences between the results based on old *CLA* program and on CATCHLIMIT are small.

The results in Table 3 indicate that the conservation performance of the RMP variant selected for the North Atlantic minke whales would be classified as 'acceptable' for the *Implementation Simulation Trials* considered in Table 3 (the conservation performance of a trial is taken to be that for the stock for which conservation performance is evaluated to be poorest). There were two cases in which performance on one of the two performance statistics was below the threshold that defines 'acceptable' performance, but overall performance was 'acceptable'.

APPLICATION TO IMPLEMENTATION SIMULATION TRIALS DEVELOPED FOR NORTH PACIFIC MINKE WHALES

IWC (2006) suggested calculating the thresholds for the four baseline trials for the North Pacific minke whales (IWC 2003). Table 4 lists the initial depletions for the 'J', 'O' and 'W' stocks, the resulting thresholds for the final depletion and depletion ratio statistics, the lower 5th percentiles for the final depletion and depletion ratio based on the actual trials, and the net result of using the criteria developed by IWC (2006) to evaluate conservation performance. Note that 'zero catch' trajectories on which the depletion ratio is based assume no RMP catches, although there are incidental catches, as specified for the trials concerned. The catch limits for the North Pacific minke whales are set by treating sub-areas 1, 2, 3, 4, 5, 6, 9N, 10 and 13 as Residual Areas, and sub-areas 7+8+11+12 and sub-area 9 as combination areas with catch limits cascaded to sub-areas within each combination area. The results for the *Implementation Simulation Trials* differ from those reported in IWC (2003) because they are based on the CATCHLIMIT program. However, for completeness, the results based on the old *CLA* program are included in a parenthesis in Table 4. As expected, there are few differences between the results based on old *CLA* program and on CATCHLIMIT.

The results in Table 4 indicate that the conservation performance of the RMP variant selected for North Pacific minke whales would be classified as 'acceptable' for the baseline trials based on stock structure hypotheses A and B and 'borderline' for the baseline trials based on stock structure hypotheses C and D. The conservation performance for the 'J' stock is 'acceptable' even though the final depletion is low. This is because there are few RMP catches from this stock so that depletion ratio is close to 1 (i.e. the population trajectory with RMP catches is similar to that without RMP catches). The primary reason for the low final depletion values for the 'J' stock in Table 4 is that the thresholds are calculated from single stock trials that do not include bycatch, whereas the multi-stock trials do include bycatch. It is not clear how to include bycatch in the single stock trials because bycatch is taken from multiple stocks but the use of the depletion ratio as a performance measure obviates the need for further consideration of this.

CONCLUSION

The sub-committee noted that there is still a fair amount of inter-simulation variability in the values for the thresholds. It may be appropriate therefore to conduct more than one simulation for each initial depletion when initial depletion varies among simulations in a trial. The sub-committee also noted, while conducting these calculations, the authors had discovered that MSYR had been specified in several different ways in previous *Implementation Simulation Trials*. Standardizing this to specifications that are already implemented in the single-stock control program would reduce the workload needed to apply the criteria for evaluating conservation performance.

REFERENCES

International Whaling Commission. 1993. Report of the Scientific Committee. Annex I. Report of the Working Group on Implementation Trials. *Rep. int. Whal. Commn.* 43: 153-96.

International Whaling Commission. 1994. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on Management Procedures. *Rep. int. Whal. Commn.* 44: 74-92.

International Whaling Commission. 2003. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* 5 (Suppl.): 107-53.

International Whaling Commission. 2006. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* 8 (Suppl.): 78-90.

Table 1
Categories of conservation performance (reproduced from IWC, 2005)

| Acceptable | Borderline | Unacceptable |
|--|---|--|
| Either: lower v%-ile of final depletion $>\alpha K$; and/or the lower x%-ile of the minimum over each of the 100 years of the ratio of the population size to that in the same scenario but there are only incidental catches, remains $>\gamma$. | Either: lower w%-ile of final depletion $>\beta K$; and/or the lower y%-ile of the minimum over each of the 100 years of the ratio of the population size to that in the same scenario but there are only incidental catches, remains $>\delta$. | Anything that cannot be classified as 'acceptable' or 'borderline' |

Table 2

Values for the thresholds that define 'borderline' and 'acceptable' performance for three average initial depletion levels and two values for the standard deviation of the initial depletion level. 'Borderline' thresholds are calculated using the 0.6 tuning of the *CLA*, 'Acceptable' thresholds are calculated using the 0.72 tuning of the *CLA*. MSYR=1% for all trials.

| Average initial depletion | SD initial depletion | 'Borderline' thresholds | | 'Acceptable' thresholds | |
|---------------------------|----------------------|------------------------------|-----------------|------------------------------|-----------------|
| | | Lower 5%-ile Final depletion | Depletion ratio | Lower 5%-ile Final depletion | Depletion ratio |
| 0.3019 | 0 | 0.425 | 0.607 | 0.547 | 0.795 |
| 0.3019 | 0.05 | 0.388 | 0.578 | 0.436 | 0.763 |
| 0.5018 | 0 | 0.437 | 0.500 | 0.606 | 0.693 |
| 0.5018 | 0.05 | 0.435 | 0.500 | 0.601 | 0.682 |
| 0.7051 | 0 | 0.448 | 0.468 | 0.621 | 0.649 |
| 0.7051 | 0.05 | 0.439 | 0.466 | 0.617 | 0.641 |

Table 3

Summary of the application of the performance criteria for the North Atlantic minke trials. The two numbers for each threshold are respectively those based on the 0.6 and 0.72 tunings of the *CLA*. The values in parenthesis in the "Lower 5%-ile" columns are the results based on the using the *CLA* program on which the actual trials were based. Values indicated in bold-underline are not 'acceptable' for one of the two performance statistics.

| Trials/Stock | Initial depletion | Final depletion | | Depletion ratio | | Conservation performance |
|--------------|-------------------|-----------------|----------------------|-----------------|----------------------|--------------------------|
| | | Thresholds | Lower 5%-ile | Thresholds | Lower 5%-ile | |
| Trial NO-1 | | | | | | |
| CIC | 0.7628 | 0.396/0.562 | 0.642 (0.642) | 0.418/0.593 | 0.685 (0.685) | Acceptable |
| CM | 0.8195 | 0.403/0.560 | 0.622 (0.623) | 0.417/0.582 | 0.664 (0.664) | Acceptable |
| EN | 0.5269 | 0.383/0.543 | 0.574 (0.592) | 0.459/0.653 | 0.741 (0.763) | Acceptable |
| EC | 0.5009 | 0.385/0.541 | 0.543 (0.569) | 0.469/0.665 | 0.666 (0.681) | Acceptable |
| ES | 0.4690 | 0.383/0.536 | 0.585 (0.598) | 0.483/0.679 | 0.754 (0.771) | Acceptable |
| EB | 0.4759 | 0.385/0.536 | 0.583 (0.597) | 0.480/0.676 | 0.751 (0.771) | Acceptable |
| Trial NO-3 | | | | | | |
| EN | 0.4442 | 0.385/0.530 | 0.542 (0.560) | 0.500/0.691 | 0.793 (0.817) | Acceptable |
| EC | 0.4168 | 0.386/0.523 | 0.523 (0.545) | 0.510/0.704 | 0.736 (0.759) | Acceptable |
| ES | 0.3809 | 0.380/0.515 | 0.546 (0.561) | 0.534/0.727 | 0.799 (0.820) | Acceptable |
| EB | 0.3877 | 0.383/0.512 | 0.545 (0.563) | 0.529/0.721 | 0.797 (0.822) | Acceptable |
| Trial NO-4 | | | | | | |
| CIC | 0.7858 | 0.400/0.564 | 0.627 (0.627) | 0.419/0.591 | 0.672 (0.672) | Acceptable |
| CM | 0.8084 | 0.403/0.564 | 0.616 (0.620) | 0.417/0.586 | 0.658 (0.662) | Acceptable |
| Trial NO-5 | | | | | | |
| EN | 0.6102 | 0.384/0.550 | 0.596 (0.605) | 0.434/0.623 | 0.704 (0.715) | Acceptable |
| EC | 0.5865 | 0.381/0.549 | 0.552 (0.564) | 0.437/0.631 | 0.602 (0.616) | Acceptable |
| ES | 0.5595 | 0.383/0.545 | 0.610 (0.619) | 0.448/0.638 | 0.720 (0.731) | Acceptable |
| EB | 0.5659 | 0.383/0.546 | 0.606 (0.615) | 0.446/0.636 | 0.716 (0.727) | Acceptable |
| Trial NO-6 | | | | | | |
| CIC | 0.7271 | 0.393/0.561 | 0.611 (0.616) | 0.420/0.600 | 0.654 (0.658) | Acceptable |
| CM | 0.8090 | 0.404/0.564 | 0.561 (0.567) | 0.417/0.587 | 0.591 (0.585) | Acceptable |

Table 4

Summary of the application of the performance criteria for the four baseline trials for North Pacific minke trials. The two numbers for each threshold are respectively those based on the 0.6 ('borderline') and 0.72 ('acceptable') tunings of the *CLA*. Thresholds are calculated using single stock trials, which do not include bycatch. The values in parenthesis in the 'Lower 5%-ile' columns are the results based on using the *CLA* program on which the actual trials were based. Values indicated in bold-underline are not 'acceptable' for one of the two performance statistics.

| Case/Stock | Initial depletion (median) | Final depletion | | Depletion ratio | | Conservation performance |
|------------|----------------------------|-----------------|----------------------|-----------------|----------------------|--------------------------|
| | | Thresholds | Lower 5%-ile | Thresholds | Lower 5%-ile | |
| Baseline A | | | | | | |
| J | 0.300 | 0.354/0.462 | 0.111 (0.111) | 0.587/0.764 | 0.964 (0.964) | Acceptable |
| O | 0.704 | 0.391/0.555 | 0.719 (0.723) | 0.407/0.593 | 0.751 (0.744) | Acceptable |
| W | 0.993 | 0.423/0.558 | 0.919 (0.920) | 0.409/0.554 | 0.908 (0.907) | Acceptable |
| Baseline B | | | | | | |
| J | 0.300 | 0.354/0.462 | 0.111 (0.111) | 0.587/0.764 | 0.964 (0.964) | Acceptable |
| O | 0.728 | 0.394/0.557 | 0.728 (0.729) | 0.409/0.589 | 0.752 (0.757) | Acceptable |
| Baseline C | | | | | | |
| J | 0.300 | 0.354/0.462 | 0.108 (0.108) | 0.587/0.764 | 0.968 (0.967) | Acceptable |
| Ow | 0.253 | 0.347/0.457 | 0.430 (0.427) | 0.554/0.740 | 0.693 (0.696) | Borderline |
| Oe | 0.662 | 0.387/0.555 | 0.571 (0.565) | 0.406/0.595 | 0.723 (0.731) | Acceptable |
| W | 0.990 | 0.408/0.555 | 0.819 (0.824) | 0.398/0.552 | 0.817 (0.822) | Acceptable |
| Baseline D | | | | | | |
| J | 0.300 | 0.354/0.462 | 0.109 (0.109) | 0.587/0.764 | 0.969 (0.969) | Acceptable |
| O | 0.290 | 0.356/0.482 | 0.381 (0.385) | 0.510/0.699 | 0.692 (0.688) | Borderline |
| W | 0.987 | 0.408/0.554 | 0.804 (0.806) | 0.398/0.554 | 0.797 (0.800) | Acceptable |

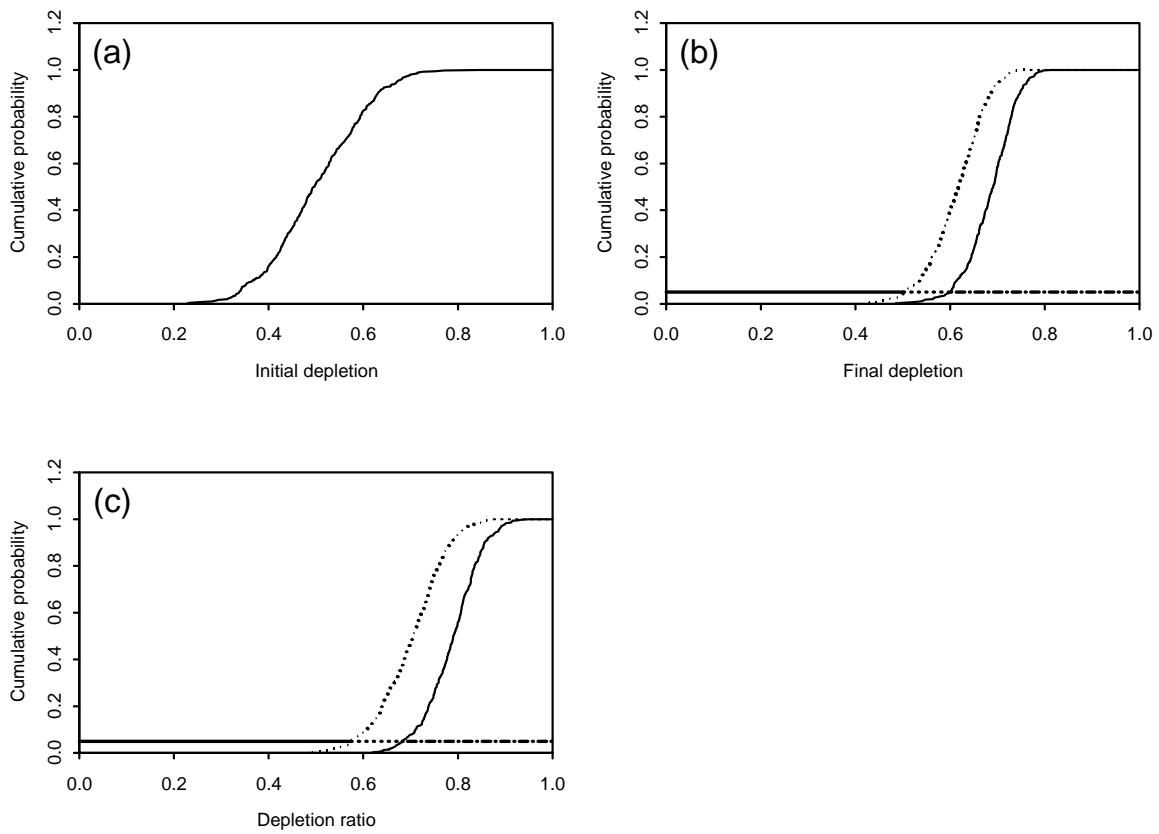


Figure 1. Application of the performance criteria evaluated in this paper to a hypothetical *Implementation Simulation Trial*. Panel (a) plots the distribution for the initial depletions in the trial. The solid and dotted lines in panels (b) and (c) denote the cumulative distributions for the 0.72 and 0.60 tunings of the *CLA*. The solid, dotted and dashed horizontal lines denote the ranges for the two performance statistics for which performance would be considered to be 'unacceptable', 'borderline' and 'acceptable' respectively.

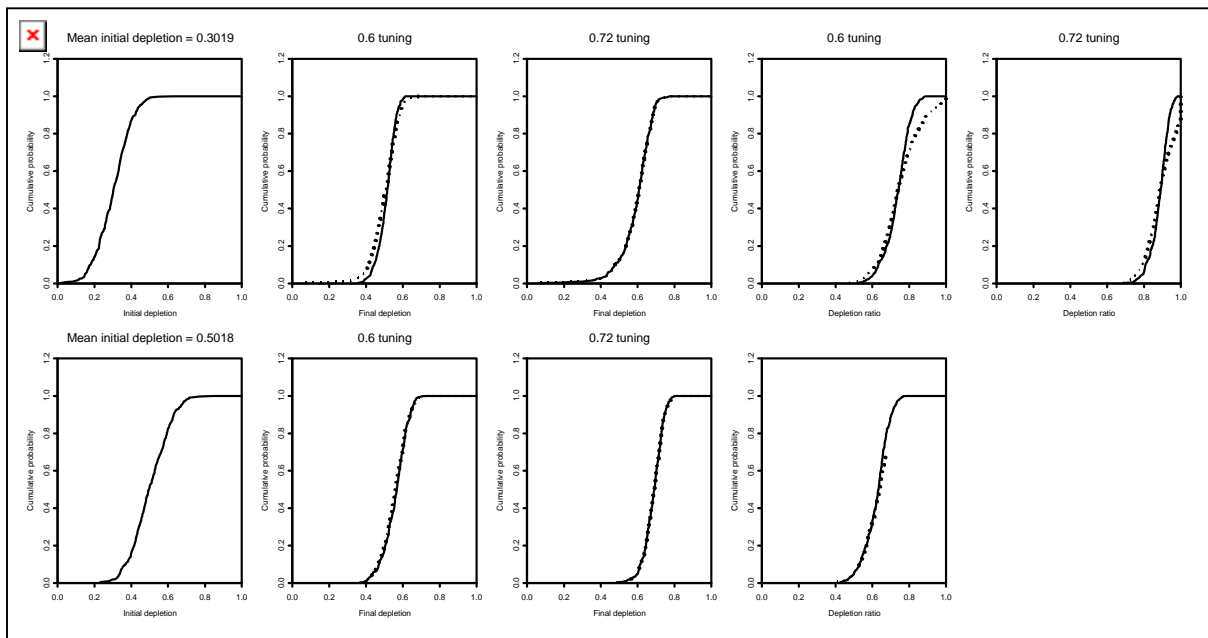


Figure 2. Initial depletion distribution, distributions for the final depletion and the depletion ratio for the 0.6 tuning of *CLA*, and distributions for the final depletion and the depletion ratio for the 0.72 tuning of *CLA*. Results are shown for three mean initial depletions. The solid and dotted lines in columns 2-4 denote respectively results when there is no inter-simulation variation in the initial depletion and when the inter-simulation standard deviation of the initial depletion is 0.05.

Appendix 3

AMENDMENT TO THE REQUIREMENTS AND GUIDELINES FOR IMPLEMENTATIONS

The Requirements and Guidelines for *Implementations* are given in JCRM 7 (suppl.): 84-92.

The following amendments are proposed (note that 'variant' refers to RMP variants, i.e. specifications or *Small Areas, Catch cascading*, etc):

Page 87, Item 4.1

Replace the first sentence under point (1) with the following text:

The conservation performance (given the highest priority by the Commission) for each trial and variant shall be examined using the following guidelines to determine whether each combination of variant and trial will be classified as 'acceptable', 'borderline' or 'unacceptable' (see box 1 of Fig. 2).

For each stock in an *Implementation Simulation Trial* for which $MSYR=1\%$:

- (4) Construct a single stock trial, which is 'equivalent' to the *Implementation Simulation Trial*. For example, if a particular *Implementation Simulation Trial* involved carrying capacity halving over the 100-year projection period, the 'equivalent single stock trial' will also involve carrying capacity halving over the next 100 years;
- (5) Conduct two sets of 100 simulations based on this single stock trial in which future catch limits are set by the *CLA*. The two sets of simulations correspond to the 0.60 and 0.72 tunings of the *CLA*. Rather than basing these calculations on a single initial depletion, the simulations for each stock shall be conducted for the set of initial depletions for the stock concerned in the *Implementation Simulation Trial* under consideration;
- (6) The cumulative distributions for the final depletion and for the depletion ratio (the minimum over each of the 100 year projection of a trial of the ratio of the population size to that when there are only incidental catches) shall be constructed for each of these two tunings of the *CLA*;
- (7) The lower 5%-ile of these distributions shall form the basis for determining whether the performance of the RMP for the *Implementation Simulation Trial* is 'acceptable', 'borderline' or 'unacceptable';
- (8) If the 5%-ile of the final depletion or the 5%-ile of the depletion ratio for the *Implementation Simulation Trial* that shows better performance is less than for the equivalent single stock trial with 0.60 tuning of the *CLA*, the performance of the RMP shall be classified as 'unacceptable';
- (9) If the 5%-ile of the final depletion or the 5%-ile of the depletion ratio for the *Implementation Simulation Trial* that shows better performance is greater than for the equivalent single stock trial with 0.60 tuning of the *CLA* but less than for the equivalent single stock trial with 0.72 tuning of the *CLA*, the performance of the RMP shall be classified as 'borderline';
- (10) If the 5%-ile of the final depletion or the 5%-ile of the depletion ratio for the *Implementation Simulation Trial* that shows better performance is greater than for the equivalent single stock trial with 0.72 tuning of the *CLA*, the performance of the RMP shall be classified as 'acceptable'.

Fig. 1 outlines these steps for a hypothetical case. Fig. 1(a) shows the cumulative distribution of the initial depletion in the *Implementation Simulation Trial*. Figs 1(b) and (c) show cumulative distributions for the final depletion and the depletion ratio. These figures are constructed by running single-stock trials where the initial depletions are set to those values on which Fig. 1(a) was based. The horizontal lines in Figs 1(b) and (c) are used to compute the thresholds that define 'acceptable', 'borderline' or 'unacceptable' performance.

Page 88

Delete Table 1.

Pages 88 and 89

Renumber Figures 1 and 2 as Figures 2 and 3, respectively.

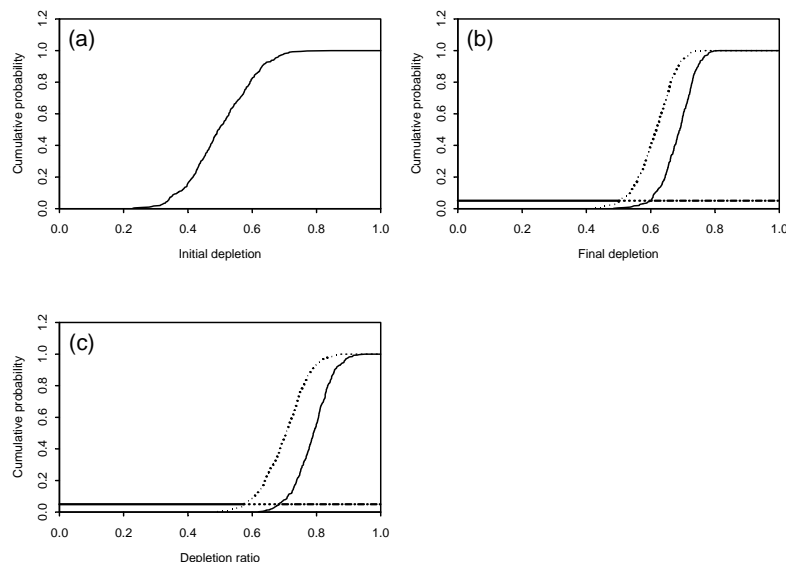


Figure 1. Application of the performance criteria to a hypothetical *Implementation Simulation Trial*. Panel (a) shows the distribution for the initial depletions in the trial. The solid and dotted curves in panels (b) final depletion and (c) depletion ratio denote the cumulative distributions for the 0.72 and 0.60 tunings of the *CLA*. The solid, dotted and dashed horizontal lines denote the ranges for the two performance statistics for which performance would be considered to be 'unacceptable', 'borderline' and 'acceptable', respectively, for that statistic in isolation. The final classification depends on consideration of both criteria together (see text).

Appendix 4

REPORT OF SUBGROUP ON REVIEW OF MSY RATES

Members: Cooke, Butterworth, Gunnlaugsson, Hatanaka, Polacheck, Schweder, Tanaka, Wade.

The group had two terms of reference:

- (i) determine what interim range of MSY rates should be used in trials in the meantime pending a review of the range of plausible MSY rates;
- (ii) make proposals for how to structure a review of the plausible range of MSY rates for use in management procedure evaluation.

1. Interim range of MSY rates

The group agreed that if further work on RMP variants is conducted before completion of the review described below, the full range of MSY rates used to date should be used in trials, namely from 1% (mature) to 7% (1+). This corresponds roughly to the range 0.66%-7% for MSYR(1+), or 1%-10% for MSYR(mature). This range will be revisited following the review.

2. Review of the plausible range of MSY rates

The last comprehensive review by the Scientific Committee of the plausible range of MSY rates for baleen whales was conducted in 1993 (*R/WC 44:181-189*). At the time the Committee concluded that there was no need to change the plausible range of MSY rates from the range of 1% - 7% (mature) that had been used in the RMP trials until then.

The group considered that sufficient new information had become available since 1993 to justify a new review, which should be conducted at the 2007 Annual Meeting. It should be completed by the 2008 Annual Meeting at the latest. The group recommended that review be limited to baleen whales.

2.1. Structure of the review

The structure of the review can be based on that of the 1993 review. The items are:

- (1) estimates from recovering populations
- (2) estimates from changes in biological parameters
- (3) estimates from population dynamic models of exploited populations
- (4) estimates from catch at age data
- (5) estimates from interspecific comparisons
- (6) generic issues, specifically the issue of yield curve shapes, and the effects of changing environments.

2.2. New information since 1993

Considerable new information is available or is expected to become available under each of these headings:

- (1) Not only are there data from more recovering stocks, but some may already be near or above MSY so that not merely an estimate of the maximum growth rate at low population sizes, is available but potentially also on the changes in growth rates as the populations increase. There is information from, for example, gray and right whales and North Atlantic humpbacks which should also be used. Information from SH humpbacks should arise from the current Comprehensive Assessment.
- (2) There is new information on changes in biological parameters at least of gray and northern humpback whales with increasing population level. The 1997 JARPA review resulted in revised estimates of changes in southern hemisphere minke whale parameters.
- (3) Schweder reported the intention to use the CPUE series developed in SC/58/RMP6 to fit a population model to determine confidence bounds for MSYR for Northeastern Atlantic minke whales. Similar analyses on fin whales were submitted to the joint IWC/NAMMCO meeting. The group encourages further submissions of this kind on other stocks where possible.
- (4) The extensive analyses on JARPA catch-at-age potentially provided information on MSY rates for SH minke whales. It was noted that changes in *K* or other parameters are needed to explain these data.
- (5) Some comparison tables of biological parameters and their changes for cetaceans have been published and should be made available as For Information papers but the group felt that the Committee would probably want to construct its own comparison table based on its review of the information.
- (6) The ubiquity of environmental changes has become more apparent since the last review. In addition to the SH minke data mentioned above, variability in parameters was evident in other data sets such as for right and gray whales and fin whales off Iceland. An important issue is how to define and measure MSYR in the context time-varying parameters.

2.3. Collation and use of the available information

The group recommended that the review not simply used published estimates of increase rates or MSY rates, but would want to categorize them at least according to the nature of data used and the analysis method. For an estimate to be considered in the review, the relevant primary paper or manuscript should be made available as a For Information paper.

Where relevant data exist, but have not been analysed with respect to the question of MSY rates, relevant analyses of these data should be actively solicited. The attention of other subcommittees should be drawn to this request.

The group recommended that a list of the cases to be considered in the review be drawn up in advance, say by the end of 2006. This should list:

- existing estimates and the corresponding primary papers;
- analyses in prep. which are intended for submission to the review;
- other relevant data sets of which analyses would be desirable.

3. Other matters

The group identified but did not discuss the following issues that may need to be addressed in addition to the review:

1. Revision of the standard set of simulation trials;
2. Revision of the set of performance statistics to be used for comparison (some of those used for selecting the first versions of the RMP are based on syntheses of results across a range of MSY rates);
3. Implications of a change in the plausible range of MSYR for tuning of the RMP or variants thereof.

| Description | Trials | | | | |
|---|---------|----------|---------|-------|-------|
| T1 Initial population 0.20K | T1-T1 | T1-T1.5 | T1-T2.5 | T1-T4 | T1-T7 |
| T1 Initial population 0.4K | T1-F1 | T1-F1.5 | T1-F2.5 | T1-F4 | T1-F7 |
| | T1-R2.5 | | | | |
| | T1-R7 | | | | |
| T16. MSYR declines linearly to half its initial value | T16-D1 | T16-D1.5 | T16-D4 | | |
| T17. K & MSYR decline linearly to half initial values | T17-D1 | T17-D1.5 | T17-D4 | | |

¹Formerly Tent model with MSYL=40%; replaced by standard age structured model with P-T recruitment function

In addition the following response curve plots should be produced for MSY rates of 1%, 1.5% and 4% and initial depletion levels of 0.3K (R trials), 0.6K (S trials) and 0.99K (D trials). Figure 1 shows an example of a response curve. Other parameters are kept at their base case values.

Biological parameters and models

| | |
|--------------------------|--|
| Parameter | Values |
| MSYL level | 0.3K; 0.45K; 0.6K*; 0.75K; 0.9K |
| MSYR level | 0.5% 1%*; 2.5%; 4%*; 7%* |
| Varying MSYR | a) 1%*; 14yr steps 1-4-1; 14yr steps 4-1-4; 33yr steps 1-4-1; 33yr steps 4-1-4; 4%* b) Constant 1%*; increases linearly ¹ from 1%-2%; 1%-4%; decreases linearly ¹ from 4%-2%; 4%-1%; Constant 4%* |
| Other recruitment models | Pella-Tomlinson (PT) (40% MSYL + rec lim); PT (60% MSYL); PT (80% MSYL); |
| Other models | Base case*; Age of maturity = 10; Standard Pella-Tomlinson (PT); PT + maximum recruitment limitation; P.T. + 25yr recruitment delay |

Stock condition and historic catches

| | |
|----------------------------|---------------------|
| Period of historic catch } | 10; 30*; 50 |
| Period of protection } | None*; 15yrs; 30yrs |

Environmental impacts

| | |
|----------------------------------|--|
| K and MSYR decrease ¹ | Constant*; decrease to: 80%; 60%; 40%; 20%; 1% of initial values |
| Varying K | Doubles linearly; constant*; halves linearly ¹ ; cyclic, starting min; cyclic, starting max |

Survey-related

| | |
|------------------------|---|
| Varying Bias | 0.5 constant; linear increasing ¹ 0.5-1.0; No bias (1.0)*; linear decreasing ¹ 1.5-1.0; 1.5 constant |
| CVobs | 0.1 2; 0.2*; 0.4; 0.6; 0.8; 1.0 (The corresponding values of CV(true) are 0.36; 0.40; 0.53 0.69; 0.87; 1.06) |
| Level of process error | Process error factor n = 0; 0.5; 1*; 2; 3; 4 (CVobs=0.2 in these trials) (The corresponding values of CV(true) are 0.2; 0.32 0.40; 0.53; 0.64; 0.73) |
| Survey frequency | 2 yrs; 5 yrs*; 7 yrs; 10 yrs; 20 yrs; determined by forecast quota |
| Surveys cease | after 0 yrs (i.e. initial survey only); 20 yrs; 40 yrs; 60 yrs; 80 yrs; continuous* |

¹ The parameter(s) vary linearly over the 100 years of management (and are constant prior to management).

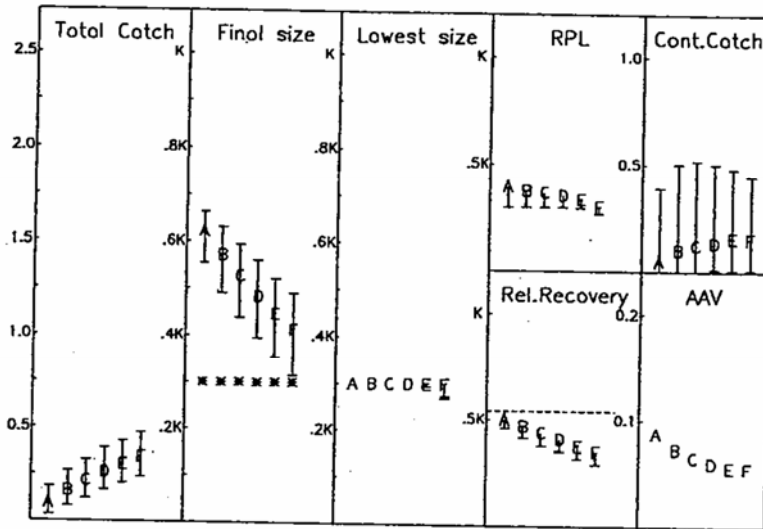


Figure 1 Sample Response curve (R1 trial, bias ranges from 1 (no bias) to 2.

3. Additional trials

The group agreed that additional trials which modelled possible environmental degradation should be developed in addition to or to replace the trials in which K, together perhaps with MSYR, varies over time. The current varying K trials have questionable behaviour when modelling population sizes above K and might better be modelled using an exponential model. The group suggested that specification of these additional trials required further work and recommends this be done before the next Scientific Committee meeting. The resulting trials would be included in the full trial set, but not in the list of those required to demonstrate an improvement in behaviour of a candidate procedure.

4. Trials not required

The tent model trials have been replaced by trials based on the P-T recruitment function in which MSYL is varied; in addition, the trials suggested above to use an exponential model will address the requirement for a trial using a population model other than the standard age structured model with a P-T recruitment function.

Trials which were designed to investigate minimum data standards (the T14-D1 and T14-R1 trials in which there was a survey in year 1 only, and other trials with inter survey intervals of 5,7 and 10 years in combination with various factors) were deleted as the phaseout rule makes them unnecessary. In addition the minimum data standards trials (ref RIWC 1994: 85-8) are not included in the full set of trials, as the response curves were considered sufficient in this regard.

5. Tuning Level

The group agreed that the current CLA which is tuned to 72% final depletion after 100 years, together with the 66% and 60% tunings provide a set of procedures with which to compare any candidate procedure. Hence it is unnecessary for any new procedure to be tuned to a particular level, as its performance will be compared to the existing (tuned) CLA by considering the full set of performance statistics.

This being the case, it is unnecessary to change the summary statistics (and in particular the continuing catch statistic which is defined over the years 90 to 99) or the time period of the trials. It is, however, open to developers to investigate properties of their procedure over longer time periods, if they so wish.

6. Summary statistics

Scaling: Catches are scaled by the initial carrying capacity. Population sizes are scaled by the carrying capacity except in trials where K varies. In these cases P_t and P_{min} are scaled by the population size resulting if no catch is taken in the management period.

1) Total catch (TC) distribution

- a) median of the 400 simulations (i.e. average of 200th and 201st values) b) 5th %ile (20th) c) 96th %ile (381st) d) mean

2) Final population size (P_f) distribution

- a) median b) 5th %ile c) 96th %ile

3) Lowest minimum population over 100 years distribution

- a) median b) 5th %ile c) 1st (i.e. lowest) value

4) Average annual catch variation (AAV)

$$AAV = \left\{ \sum_{i=1}^{100} \sum_{t=0}^{98} |C_{i,t+1} - C_{i,t}| \right\} / \left\{ \sum_{i=1}^{100} \sum_{t=0}^{98} C_{i,t} \right\}$$

5) Continuing catch (C_c) distribution, as defined below

- a) median b) 5th %ile c) 96th %ile

The 'sustainable yield' as a function of population size, $S_Y(P)$ is defined as: $S_Y(P) = \begin{cases} MSY & \text{for } P > MSYL \\ \text{long-term equilibrium } RY & \text{for } P < MSYL \end{cases}$

The 'continuing yield' realised by a catch C is defined as: $C_Y = \begin{cases} C & \text{for } C < S_Y \\ S_Y & \text{for } C > S_Y \end{cases}$

The 'continuing yield' or 'continuing catch' statistic (C_c) is the median over simulations of: $\frac{1}{10} \sum_{t=90}^{99} C_{Y,t}$

The statistic is computed by calculating $S_Y(P)$ for population sizes at fixed intervals from $P=0$ to $P=MSYL$, for each set of trials, and storing the results in an array. The statistic is calculated by interpolation from this array.

6) RPL = realised protection level = lowest stock size for which a catch is set (median & 5%ile) (Ref: RIWC 1993:222)

7) RR = relative recovery = stock level in the year when the zero catch trajectory reaches 54% (Median & 5%ile)

Graphical Presentation (ref RIWC 40:114)

Graphs showing the following are required (a sample of the graphs is given in Fig 2):

- (i) the mean population trajectory over the 400 simulations;
- (ii) the set of 5th and 96th %iles of the distribution of population sizes reached each year in the 400 simulations, on the same plot as (i);
- (iii) the mean catch trajectory over the 400 simulations; and
- (iv) the catches taken in the first two stochastic simulations, preferably on the same plot as (iii).

