

## Annex D

### Report of the Sub-Committee on the Revised Management Procedure

Members: Allison, An, Baba, Bannister, Berggren, Best, Bjørge, Brandon, Breiwick, Brownell, Butterworth, Charrissin, Childerhouse, Cooke, Danielsdóttir, DeMaster, Dereksdottir, Donovan, Findlay, Fortuna, Fujise, Galletti, Goto, Gronvik, Gunnlaugsson, Hakamada, Hammond, Hatanaka, Holloway, Hyugaji, Ilyasenko, Iñiguez, Jackson, Kanda, Kato, Kawahara, Kell, Kim, Kitakado, Leaper (Rebecca), Leaper (Russell), Lens, Leslie, Lovell, Miller, C., Miyashita, Mori, Morishita, Nakamura, Ohsumi, Øien, Okamura, Ólafsdóttir, Palka, Panigada, Pastene, Peel, Perrin, Pike, Pinto de Lima, Polacheck, Punnett, Punt, Robbins, Rosenbaum, Schweder, Shimada, Simmonds, J., Skaug, Tanaka, Tominaga, Van Waerebeek, Vikingsson, Wade, Walloe, Winship, Witting, Yamakage, Yasokawa, Yoshida, Young.

#### 1. INTRODUCTORY ITEMS

##### 1.1 Pre-meeting meeting

Prior to the main meeting, a pre-meeting meeting was held over the two days 24-25 May. Chaired by Donovan, it agreed an agenda (attached as Appendix 1), and discussed items 2.1, 2.2, 3.1.1-3.1.3.

##### 1.2. Election of Chair, appointment of Rapporteurs

As above, Donovan chaired the pre-meeting meeting, and Bannister chaired the sub-committee during the main meeting. Rapporteurial duties were undertaken by Butterworth, Donovan, Hammond, Perrin, Pike, Punt and Wade.

##### 1.3. Adoption of Agenda

See item 1. 1.

##### 1.4. Review of Documents

Documents available included SC/58/RMP 1-8; PFI 1-9; O21; Rep 1

#### 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*

IWC (2005) developed a set of guidelines and requirements for the Scientific Committee when it attempts to Implement the RMP for a given species and Region. These guidelines relate to: a) the information needed to initiate a pre-implementation assessment, b) the nature and outcomes from a pre-implementation assessment, and c) the steps when conducting an Implementation. An Implementation will normally be completed two years after the Scientific Committee recommends that the pre-implementation assessment is complete and the Implementation can commence. The Implementation occurs over two intersessional workshops and two Annual meetings. The objective of the second of these intersessional workshops is to review the results of trials and develop recommendations to the Scientific Committee regarding management areas, RMP variants, operational constraints, future research to narrow the range of plausible hypotheses/eliminate some hypotheses, and 'less conservative' RMP variants (and their associated research programmes and duration).

The review of the results of the trials involves *inter alia* assessing conservation performance for each trial and RMP variant using pre-determined guidelines and hence classifying each combination of RMP variant and trial as 'acceptable', 'borderline' or 'unacceptable'. IWC (2005) developed some generic guidelines for evaluating performance, but did not specify the exact numerical values for the thresholds that define 'acceptable' and 'borderline'. It indicated that these values would be based on the values for the relevant performance measures for the single stock trials and that the values for 'acceptable' performance would be based on the results for the D1 and R1 base-case trials for the 0.72 tuning of the CLA, while those for 'borderline' performance would be based on the results for these trials for the 0.60 tuning of the CLA.

A previous attempt to calculate values for the thresholds that define 'acceptable', 'borderline' and 'unacceptable' conservation performance based on these specifications had not been completely successful (Punt 2005). IWC (2006) therefore proposed the following way 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.
- (4) 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'.

At last year's 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 given in Appendix 2.

The sub-committee agreed that the work completed intersessionally was sufficient to define the threshold levels for defining 'acceptable' and 'borderline' performance for classifying the performance of RMP variants for *Implementation Simulation Trials*. An amendment to the Requirements and Guidelines for Implementations (IWC 2005) is given in Appendix 3. The sub-committee recommended that this amended text be adopted.

#### 2.1.2 Develop a list of agreed stock structure archetypes

Pastene presented SC/58/RMP3, which reviewed the process to define stocks during RMP *Implementations* of North Pacific common minke and Bryde's whales. The *Implementation* process was completed for NP common minke whales during the 2003 Scientific Committee meeting and an *Implementation Review* is scheduled for the 2008 meeting. The 'First Intersessional Workshop' for North Pacific Bryde's whale was held in October 2005. Some general aspects of stock structure in these two North Pacific assessment cases were summarized, namely; (i) the kind of data available and analyses conducted and, (ii) how the information on stock structure was used in the *Implementation* process. SC/58/RMP3 identified some problematic issues inherent to the analysis and interpretation of stock structure, which made it difficult to reach agreement on this topic in the North Pacific cases, particularly in the case of the common minke whale: (i) different views on how to define stocks, (ii) stock delineation based on samples from migratory corridors or feeding grounds, (iii) use of novel and yet to be validated analytical techniques during assessments. Some suggestions were provided on how these issues can be addressed. The Committee has not agreed on a methodology to evaluate quantitatively the plausibility of different stock scenarios. It is argued in SC/58/RMP3 that all these issues, although complex, should not be over-emphasized and used as an argument to delay a reasonable agreement on stock structure scenarios. In reality, there is no way to address completely all of the issues listed above. Stock scenarios should be defined taking into account the best available scientific data recognizing that these could be improved during the *Implementation Review*. Furthermore, SC/58/RMP3 argued that discussion on stock structure for different *Implementations* should be carried out on a consistent basis, as a scientific approach demands, and that the same criteria should be used to define stocks in different *Implementations*.

In discussion, the sub-committee agreed that peer-review of new methods for the analysis of genetic data was important and that this was best achieved through the Committee itself. It was desirable for this review to be done by a Working Group that was separate from the Working Group undertaking the *Implementation*. The sub-committee noted that the purpose of the TOSSM project being developed under the Sub-committee on Stock Definition (see Annex I) was to test such methods; however, this will not be completed for a number of years.

In the absence of a framework such as TOSSM, the Committee has developed a number of stock structure hypotheses or archetypes in *Implementations* conducted to date. Details of these can be found for North Atlantic (IWC, 2004b), Southern Hemisphere (IWC, 1993) and western North Pacific minke whales (IWC, 2004a), western North Pacific Bryde's whales (SC/58/Rep1) and North Atlantic fin whales (SC/58/Rep3).

## 2.2 Proposal for revision of the CLA

### 2.2.1 General approach to such considerations

IWC (1994) specifies the protocol for evaluating proposed amendments to the RMP. For a proposal for an amendment to be considered, there needs to be some evidence, in the form of simulation trial results or otherwise, that the proposed amendment would result in improved performance in at least some respect. The protocol consists of three steps:

- (i) Adequate notice shall be given to the Commission and the Scientific Committee of any proposal for amendment to the RMP.
- (ii) Given the time it will take of the Committee to evaluate such proposals, suitable evidence shall be presented to indicate that the proposed amendment would indeed represent an improvement. In this context, an amended procedure that allowed higher catches or lower catch limit variability will only be considered an improvement by the Committee if it performs adequately on all risk-related performance statistics, and better than the current version of the RMP on at least some catch- or risk-related performance statistics. This evidence shall take the form of results from appropriate, fully specified and programmed simulation trials, a list of which needs to be developed by the Scientific Committee (based *inter alia* on Table 3 in IWC 1993). These trials shall have been carried out by the proposer.
- (iii) The Committee may then specify further simulation trials and/or modification of trials already carried out along with criteria for the evaluation of the results. Advice to the Commission may then be given at its next Annual Meeting, subject to the completion of the work specified.

### 2.2.2 Consideration of the Norwegian proposal

In the notification given by Norway at the Sorrento meeting in 2004 four possible changes to the 'base-case' and robustness trial and to the CLA were indicated. Only two of these have been explored in SC/58/RMP7, which was introduced by Walløe:

- (1) To run the simulations for more than a hundred years and until equilibrium abundance is reached. Exploratory simulations showed that equilibrium was nearly reached after 300 years, which was used as the period for all trials. The resulting long term (after 300 years) Depletion Level (DL) was chosen as the relevant parameter to characterize the trials instead of the Tuning Level (TL - depletion level after 100 years), which is usually wrongly interpreted as the long term depletion;
- (2) The MSYR should refer to the 1+ component of the population instead of the mature component. The Committee has used the 1+ component and specific choices of MSYR with  $MSYR(1+) = 1\%$  as the minimum in its development of the SLA of the AWMP 'because it considered they more closely correspond to biological reality' (IWC 1998).

When the 'base case' simulations were run for 300 years for the three TL options of 0.60, 0.66 and 0.72, chosen when the CLA was developed in the late 1980s, the final depletion ends up at 0.73, 0.74 and 0.76, respectively. Three parameters in the CLA can potentially be used for tuning: PPROB, IPL and PSLOPE. PPROB was used as the tuning parameter in the original CLA. Values of PPROB = 0.4015, 0.4629, 0.5222, gave TL = 0.72, 0.66 and 0.60, respectively, when IPL = 0.54 and PSLOPE = 3.0. However, it is not possible to obtain a DL lower than 0.70 by increasing PPROB. In addition, it is theoretically unwise to increase PPROB above 0.5.

In SC/58/RMP7, DL has been tuned to lower levels by increasing PSLOPE while keeping PPROB at 0.5 and IPL at 0.54. Two values of DL were chosen: 0.66 and 0.69. For these two DLs the six base case simulations and the 24 most difficult robustness simulation trials from the development around 1990 have been carried out. In addition three base case simulations have been carried out with MSYR(mature)=1% for comparison. For each of the 33 simulation trials the results are displayed as depletion levels, lowest population, average catch, average annual variation in catch, average population as a function of time, 5% lowest population as a function of time, and average catch per year as a function of time. Three sample trajectories are also shown. The trial results clearly show that both new tunings give CLAs which perform adequately on all risk-related performance statistics, one somewhat better than the other, and that they both perform better than the current versions on catch-related performance statistics.

SC/58/RMP7 had used IWC (1993) as a basis for the set of simulation trials conducted, on the interpretation that this was a list of trials developed by the Scientific Committee and thus satisfied point (ii) under item 2.2.1. The authors of SC/58/RMP7 noted that their Table 3 lists robustness trials not performed. To run trials over 300 years takes considerable time and SC/58/RMP7 had concentrated on the more demanding trials. The T5 and T7 trials are not very challenging and the 'tent' model is difficult to implement in the age-structured model now used in all simulation trials so the T10 and T11 trials were also not conducted. The performance of the T13 trials, in which MSYR changes abruptly ever 33 years, was deemed likely to be spanned by the corresponding MSYR=1% and MSYR=4% trials. The T14 trials are irrelevant because regular surveys are mandatory, and the T15 trials are not very relevant because surveys must be carried out every 5 years.

The sub-committee did not have time to fully review the results of SC/58/RMP7. The main point raised in discussion related to the behaviour exhibited in some cases in which catches initially increase and then decrease quite sharply followed by reduced catches in the long term to achieve acceptable final depletion levels after 300 years. The population trajectories show a corresponding strong initial dip followed by a long slow recovery. Some members pointed out that this was not a good generic feature of a procedure for sustainable management of a resource. The authors of SC/58/RMP7 pointed out that the current CLA exhibits similar properties in that relatively high catches are linked to an initial dip in the population trajectory and when there is no dip in the trajectory there are poor catches. They believed that making possible the highest continuing yield should be linked to the population achieving equilibrium in the long term.

As a note to the authors of SC/58/RMP7, it was suggested that the results of the episodic events trials appeared to show that an internal correction factor in the trials had not been recalculated to account for the change from MSYR(mature) to MSYR(1+).

Wade presented SC/58/RMP8, which briefly reviewed the process by which the Committee has previously agreed to specific calculations for the CLA, and discussed implications of changes to the CLA proposed by the Norwegian re-tuning in SC/58/RMP7. A number of candidate CLAs were considered when the Committee originally developed the CLA as part of the RMP. To compare these candidate CLAs in a fair manner, each candidate CLA was 'tuned' to achieve depletion levels of 0.60, 0.66, and 0.72K after 100 years from the start of the simulation. These depletion levels and the choice of 100 years were not considered specific management objectives – they were simply specified to cover a range of levels that were consistent with management goals, and allowed for the evaluation of the performance of the different candidate procedures in a consistent fashion. After the 'C' procedure was selected as the CLA, these different tunings were then additionally used to examine the tradeoffs between catch and risk within the 'C' procedure itself. The performance of the 'C' procedure for these three different tunings was presented to the Commission to provide a range of tradeoffs between catch and risk for the Commission to consider. When the Commission adopted the CLA it specified that the 0.72 tuning should be used in the CLA, as they considered that particular trade-off of catch and risk to be preferable.

SC/58/RMP8 noted that the authors of SC/58/RMP7 have run simulations to modify and 're-tune' the CLA. Specifically, the time horizon for examining final depletion level was changed from 100 to 300 years, two new lower target depletion levels were examined, and a higher population growth rate was used by changing MSYR(mature)=1% to MSYR(1+)=1%. Thus there are two separate items that have been changed in SC/58/RMP7. In addition to re-tuning the CLA to create new candidate CLAs, they also changed the trial scenarios used to evaluate candidate CLAs, and this makes it difficult to compare the performance of these new candidate CLAs with the existing CLA using the results presented in SC/58/RMP7. The guidelines for considering modifications to the CLA require that new candidate CLAs are demonstrated to perform better than the agreed CLA on the same performance trials. This is not possible given the trials run in SC/58/RMP8, which, with regards to performance, shows that, as expected, the two new tunings to lower depletion levels resulted in better performance on catch statistics, but poorer performance on risk statistics, illustrating the clear trade-off between catch and risk. SC/58/RMP8 suggests, therefore, that the two new lower depletion level tunings in SC/58/RMP7 do not appear to be acceptable on risk. With regard to MSYR, SC/58/RMP8 noted that the existing CLA had been previously examined in trials where MSYR(mature)=4%, and its performance was found to be acceptable by the Committee. Further, the CLA includes a uniform prior distribution from 0% to 5% on MSYR(1+), and thus allows for higher growth rates.

Walløe responded that the performance on risk statistics of one of the new tuning levels (gamma 0.69) in SC/58/RMP7 was very similar to the old Tuning Level of 0.60 if one considers MSYR(1+) an appropriate choice for a minimum MSY rate, and that examination of the results of further simulation trials and response curve plots would be required to show if there really is a difference in risk-related performance.

The sub-committee briefly discussed the relevance of running simulations over 300 years vs the 100 years used during the development of the RMP. The sub-committee agreed that the acceptability of a proposed amended CLA would *inter alia* be judged on performance over 100 years. Further discussion on this point is recorded below.

During discussions of the proposals in SC/58/RMP7, a number of issues were raised concerning how the sub-committee should proceed in evaluating proposed amendments to the CLA. To avoid confusion when considering the steps that needed to be taken, the sub-committee recalled how the process by which the current CLA had been evaluated and selected.

#### THE ROLE OF TUNING IN THE EVALUATION AND SELECTION OF THE CLA

When it came to recommending a single management procedure to the Commission for possible adoption, the Committee noted that comparison of alternative procedures is not entirely trivial, because they have to meet competing objectives. A procedure can always be modified to reduce the risk of depletion of stocks, but at the cost of allowing less catch. Likewise, higher catches can be achieved, but at the cost of a greater risk of depletion of the stock. The Committee recognised that in the presence of this trade-off, it could be difficult to compare the underlying performance of two candidate procedures if they are tuned to achieve different trade-offs between the two main objectives of catch and risk, and if only one tuning is presented.

However, if results from several alternative tunings of each procedure are available, then it may be possible, by interpolating the results if necessary, to draw conclusions about the relative performance of the two procedures.

In 1990 (IWC, 1991) the Committee therefore requested developers to present results a wide range of tunings to ensure that there would be at least some overlap in the range of risk-related performances of the different candidate procedures. In preparation for a final selection between procedures, the fourth and final Comprehensive Assessment Workshop on Management Procedures recommended that selection of a procedure be based on results for three specific tunings of each of the five candidate procedures. These three tunings should be such so as to achieve a median final depletion (ratio of current to unexploited population) of the mature female population after 100 years of 0.60, 0.66 and 0.72, in a specific reference trial. The reference trial chosen was the so-called D1 trial. "The D1 trial was chosen because it reflected the greatest discrimination between the preferred tunings of the various developers; the highest and lowest final population values adopted for tuning purposes reflected the range covered by these preferred tunings" (IWC 1992a).

A set of 12 trials was selected for performance comparison, in addition to a number of robustness trials for which it was merely required that a procedure perform "acceptably". Fourteen basic performance statistics were provided for each trial, but comparison between procedures was to be made on the basis of just 18 composite statistics calculated from the core set of 12 trials (IWC 1992b). The Committee selected one of the five candidate procedures on this basis, for recommendation to the Commission (IWC 1992c). The Commission accepted the recommendation and selected one of the three tunings presented, the 0.72 tuning (IWC 1992d).

#### DISTINCTION BETWEEN TUNING TARGETS AND PERFORMANCE CRITERIA

The denomination of the selected tuning in terms of a median final depletion to 0.72K in the D1 trial has caused considerable confusion within the secondary literature on the RMP. The tuning level of 0.72 has frequently been erroneously interpreted as a management target, when in fact it was only for comparative purposes for assessing the relative performance of alternative procedures. No special weight was placed by the Scientific Committee on either the D1 trial, nor on the median final depletion as a performance measure.

Of the eight risk-related performance measure used for management procedure selection, only two related to final depletion and six related to the lowest depletion over the 100-year period. The D1 trial was chosen because of the discrimination it offered between procedures, not because it was a typical, central or average trial in any sense. On the contrary, it was a relatively extreme trial in that it was based on the lower extreme of the range of MSY rates (1%, 4% and 7%) considered across the trials. For MSY rates in the middle of the range (i.e. the 4% trials), the median 100-year depletion was considerably higher.

The Committee based its selection of the current *CLA* on the consideration of a number of performance measures (18 in all). The actual final depletion level achieved by any catch limit algorithm, whether after 100 or 300 years, depends on the details of the specific trial. The current *CLA*, which has a nominal tuning level of 0.72, will only actually exhibit a median final depletion of 0.72 in the precise reference trial used for tuning purposes.

#### IMPLICATIONS FOR EVALUATION OF AMENDMENTS TO THE *CLA*

The considerations outlined above have two implications for the evaluation of amendments to the *CLA*, such as proposed in SC/58/RMP7. First, achievement of a specific tuning level is not itself a measure of conservation performance. In comparing proposed amendments with the current *CLA*, performance should be compared across the full range of performance measures. Second, comparison of the performance of two procedures can only be validly evaluated their performance over the same trial. Whether or not MSY rate is defined in terms of the recruited or mature population, and whether or not the tuning is to a given population size after 100 or 300 years are not very important provided that alternative procedures are evaluated over the same trials and the performance measures are appropriate.

The sub-committee highlighted three steps that should be taken to evaluate any amendment to the current *CLA*, including those presented in SC/58/RMP7:

- (1) Agreement on the range of MSY rates to be used in the trials;
- (2) Identification of an appropriate standard set of simulation trials;
- (3) Definition of an appropriate set of performance statistics to be used in evaluating alternative procedures;

There was some initial discussion of these steps. The sub-committee agreed that the development of proposed amendments to the *CLA* could proceed in parallel with consideration of these points.

#### RANGE OF MSY RATES TO INCLUDE IN TRIALS

To address point (1) above, the sub-committee established a Working Group under Cooke with terms of reference: (a) to make proposals on how to structure a review of the plausible range of MSY rates for use in management procedure evaluation; (b) to determine what interim range of MSY rates should be used in trials pending the results of that review. The report of the Working Group is given in Appendix 4.

The sub-committee agreed to the report and accordingly recommended:

- (a) The interim range of MSY rates to be used in further development of amendments to the *CLA*, pending review in 2007, shall be from  $MSYR(\text{mature})=1\%$  to  $MSYR(1+)=7\%$ , corresponding approximately to the range 0.66-7% for  $MSYR(1+)$  or 1-10% for  $MSYR(\text{mature})$ ;
- (b) A review of MSY rates should be conducted at next year's meeting and completed by the 2008 Annual Meeting at the latest. The review should be limited to baleen whales;
- (c) The review should not simply take published estimates of rates of increase or MSY rates but should examine them critically with respect to, *inter alia*, the nature of the data and the analysis methods used. Papers considered should be made available as For Information documents;
- (d) 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 sub-committees is drawn to this request;

- (e) A list of cases to be considered in the review should be drawn up in advance, by the end of 2006. The list should include existing estimates and corresponding primary papers, analyses in preparation intended for submission to the review, and other relevant datasets of which analyses would be desirable.

The sub-committee received a brief presentation of SC/58/RMP6, which considered data on all landed minke whales from the northeastern Atlantic (small areas ES, EB, EW and EN) from 1952. The purpose is to construct a relative abundance series for the stock, and later to estimate productivity in the population. Due to changes in the management regime, with individual boat quotas from 1984 and a resumption of commercial whaling from 1993 under conditions rather different from in previous years, two unconnected series are constructed, one for 1952-1983 and one from 1993-2004. The simple and generally applicable net catcher day method of Cooke (1984) is employed, despite its shortcomings as discussed previously. The estimated relative change declines to 63% of its 1952 level in 1983, and increases to 186% of its 1993 level in 2004. Since autocorrelations have not been accounted for, the large estimated increase in recent years is rather uncertain.

In response to a question about how the decline in the CPUE series for the Lofoten area was handled in analysis, it was clarified that SC/58/RMP6 generated a CPUE series for the entire Norwegian fishery, including the Lofoten area. It was noted that spatial effects can be difficult to model in analyses of CPUE data, especially when there are closed areas into which results are extrapolated. These and other technical issues were important to consider in these types of analysis. There was insufficient time to discuss this paper in any detail. The sub-committee looked forward to receiving further results next year in the context of its review of MSY rates.

#### SET OF SIMULATION TRIALS FOR TESTING AMENDMENTS TO THE CLA

To address points (2) and (3) above, the sub-committee established another Working Group, under Allison, with terms of reference: (a) to review simulations trials conducted previously to evaluate catch limit algorithms and to specify a list of trials that should now be used to evaluate any proposed amendments to the *CLA*; (b) to revise the set of performance statistics to be used for comparison of catch limit algorithms; (c) to consider the specification of tuning levels for comparing catch limit algorithms.

The sub-committee thanked the Working Group for its work, the report of which is given in Appendix 5 in the form of a set of requirements for simulation trials to be conducted to evaluate proposed amendments to the *CLA*. The sub-committee recommended that these requirements be integrated into the RMP specification.

One topic discussed by the Working Group required further work. Additional trials that modelled possible environmental degradation should be developed in addition or to replace the trials in which  $K$ , perhaps together 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. Specification of these additional trials should be done before next year's meeting. The resulting trials will then be included in the full trial set in Appendix 5, but not in the list of those required to demonstrate an improvement in behaviour of a candidate procedure.

The sub-committee noted that some of the original trials used to test the *CLA* have been deleted from the list in Appendix 5. The tent model trials have been replaced by trials based on the Pella-Tomlinson recruitment function in which  $MSYL$  is varied; in addition, the trials proposed 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 Pella-Tomlinson recruitment function. Trials that 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 phase-out rule makes them unnecessary. In addition, the minimum data standards trials (IWC, 1994) are not included in the full set of trials because the response curves were considered sufficient in this regard.

### 2.3 Work plan

As its work plan for next year, the sub-committee referred to the work specified under Item 2.2.2 relating to the review of plausible MSY rates and the specification of additional trials for testing amendments to the *CLA*. Two Intersessional Working Groups were established to facilitate this work:

- (1) to plan the review of MSY rates: Cooke (Convenor), Butterworth, Gunnlaugsson, Hatanaka, Polacheck, Punt, Schweder, Tanaka, Wade;
- (2) to specify additional trials for testing amendments to the *CLA*: Allison (Convenor), Butterworth, Cooke, Donovan, Kawahara, Punt, Walløe.

The sub-committee noted that the review of MSY rates may take up considerable time at next year's meeting and that it may be prudent to plan for a pre-meeting workshop on this subject.

## 3. RMP – PREPARATIONS 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

Donovan introduced SC/58/Rep 2, the report of the intersessional workshop. The Workshop took place at the National Research Institute of Far Seas Fisheries, Shizuoka, Japan from 25-29 October 2005. The objective of the meeting was to develop an appropriate *Implementation Simulation Trials* structure and to specify the associated conditioning so that it can be carried out before the following Annual Meeting. Full details can be found in the 'Requirements and Guidelines for *Implementations*' agreed by the Committee in 2004 (IWC, 2005).

A considerable part of the Workshop was dedicated to addressing issue related to stock structure hypotheses with emphasis on their role in specifying an appropriate trial structure. The Workshop reviewed the evidence for each of the five hypotheses agreed by the *pre-Implementation Assessment*. It noted and welcomed the considerable additional data that have become available since 1999. It noted that the assignment of plausibility in a trial context would not take place until the First Annual meeting in 2006. The Workshop considered information from several approaches, both genetic and non-genetic.

After considerable discussion, it agreed that four stock hypotheses were sufficient for the trials. These are given in Appendix 6, Fig. 2.

*Hypothesis 1:* Only one stock of Bryde's whales is found in the area from 130°E to 160°W (excluding the area in which East China Stock is found) and there are no sub-stocks.

*Hypotheses 2 and 3:* Different stocks in sub-areas 1 and 2. Hypothesis 3 differs from hypothesis 2 in that the stock found in sub-area 1 is also found in sub-area 2.

*Hypothesis 4:* Identical to hypotheses 2 but with two sub-stocks that mix in sub-area 1.

Most of the discussion at the Workshop focussed on hypothesis 4. It agreed that at this stage, a hypothesis that included sub-stocks was necessary to implement one of the possible explanations for the differences in the age distributions for JARPN II and for sub-areas 1E and 2, recognising that further work may eliminate this explanation by the next annual meeting. Results from a simple model filter analyses do not provide evidence for or against sub-stocks but suggest that *if* there is more than one sub-stock in sub-area 1, then there is likely to be considerable mixing.

It also considered a number of other issues related to trials structure and recommended that:

- (1) all trials would assume that  $g(0)=1$ ;
- (2) there was insufficient information to develop an estimate of the lower bound for MSYR;
- (3) alternative catch series be developed based on the agreements reached in SC/58/Rep1;
- (4) experimental work be undertaken to try to determine whether observed differences in age distributions are (a) related to age reading and/or sampling issues in the commercial data (this has implications for calculation of  $t_m$  and  $M$ ); (b) real and reflect age-segregated distribution (this implies that trials should allow for smaller proportions of older animals in 1W than in 1E+2); or (c) differences are real and may indicate some degree of stock structure between sub-area 1W and 1E+2 (this has implications for stock structure hypothesis 4).
- (5) the three previously agreed sub-areas be retained (see Fig. 2) but moving the 1W/1E boundary to 165°E (and undertaking sensitivity trials for 155°E);
- (6) trials should assume an age at recruitment as applied to past coastal whaling (i.e. assuming no length limits) although it was agreed to investigate the implications of retaining or removing such limits;
- (7) the trials structure given in SC/58/Rep2 (Annex G) adequately captures the full range of uncertainty for the western North Pacific Bryde's whales;
- (8) a control program that implements these trials be developed and that the trials be conditioned and results reported to the 2006 Annual Meeting using values for abundance, biological parameters, small area definitions and management variants detailed in the report;
- (9) The results be presented according to the manner described in SC/58/Rep2.

A number of recommendations were made to try to ensure that this work was completed before the 2006 meeting.

In discussion, the sub-committee thanked Donovan and the participants for their work. It endorsed the recommendations. Progress on the recommendations is discussed under the relevant Items below.

### *3.1.2 Objectives of the First Annual Meeting*

The primary purpose of the first Annual Meeting is to review the results of conditioning and to finalise the *ISTs*.

New analyses of data available up to the time of the previous workshop are acceptable but new data may not be introduced at this stage. After reviewing the results of the conditioning, the *Trials* themselves may be changed, but the overall structure can not be changed.

The primary output will be the detailed specifications of the final *ISTs*. These will be determined on the basis of:

- (1) final consideration of the plausibility of the various hypotheses and hence the weight assigned to each of the trials (the overall balance of the *ISTs* will be accounted for when weights are assigned);
- (2) discussion of what data/research may reduce the number of hypotheses and possible time-frames for this research/data collection;
- (3) updates/improvements to standard data sets (i.e. abundance, catches, bycatches) for use by the *CLA* in final trials and when evaluating the plausibility of hypotheses and hence assigning weights to trials (new data would not be used when conditioning the trials); and
- (4) specification of operational features (geographical and temporal) and management variants.

The specification of final trials:

- (1) may include trials to examine effects of using one RMP variant over an initial period (up to 10 years) followed, after a 5-year phase-out period, by a more conservative variant (see discussion below);
- (2) will exclude potential 'low' weight trials (e.g. those where at least one factor is considered to have 'low' plausibility); and
- (3) will assign weights to the remaining trials of 'high', 'medium', or 'no agreement'.

A timetable for the remaining work (including circulation of trial results and format) will be developed – the timetable will be determined so that there is a reasonable expectation that the results of the trials will be available well before the second intersessional workshop.

The Committee will also commence discussions related to defining the inputs for actual application of the *CLA* (catches, bycatches, estimates of abundance and projected future anthropogenic removals).

### 3.1.3 Review results of conditioning

Allison introduced the modifications to the trial specifications (SC/58/Rep2, Annex G) that had been developed since the Workshop. The sub-committee agreed to the revised specifications as now included in Appendix 6<sup>1</sup>.

### 3.1.4 Updates to standard data sets

#### 3.1.4.1 ABUNDANCE

The sub-committee agreed with the conclusions of the intersessional workshop. The pooled abundance estimates obtained for the complete survey area considered (blocks F-M: see Fig. 1 of Annex H) was 21,826 (CV = 0.295), to apply to 1995, with an additional standard deviation estimate of  $\sigma_A=0.673$ . These results (and corresponding estimates for smaller areas as reflected in Annex H) were adopted for use in the *Implementation Simulation Trials*. It was noted that this estimate did not include contributions from blocks A-E (see Fig. 1 of Annex H); this was not of concern for conditioning purposes because the 1998-2002 surveys indicated these contributions to be relatively small, but the methodology to be used to obtain abundance estimates for input on application of the RMP will need to take account of these blocks as well. With respect to using abundance estimates in the RMP, to avoid coding problems it was agreed that two separate estimates and their CVs should be used.

#### 3.1.4.2 CATCHES AND OTHER ANTHROPOGENIC REMOVALS

SC/58/RMP1 provides a description of the Philippines commercial whaling operation for Bryde's whales in the 1980s. It confirmed, from official documents obtained in Manila, that the whaling was pelagic, not coastal, and that the whales were not taken in Philippine EEZ waters but near the Ogasawara Islands and in Micronesia. Some of the whales taken in Micronesia may have been of the pygmy species (known currently as *Balaenoptera omurai*) but as yet still included in the IWC list of recognized taxa with the "ordinary" Bryde's whale under *B. edeni*. These points have been taken into consideration when constructing the final catch series and alternatives given as Appendix 7.

### 3.1.5 Final consideration of plausibility (including weighting of trials in terms of overall balance)

#### 3.1.5.1 STOCK STRUCTURE HYPOTHESES

Pastene presented SC/58/PFI2 which provided an examination of the plausibility of the four stock structure hypotheses set out in the Report of the First Intersessional Workshop on the Western North Pacific Bryde's Whale *Implementation* (SC/58/Rep1). The paper considered information available from allozymes, mtDNA, microsatellites, sighting distributions, catch distributions, external body proportions, biological parameters, mark-recapture and age distributions of catches.

The Sub-Committee reviewed the information summarized in Table 2 of SC/58/PFI2 regarding the plausibility of different stock structure hypothesis. After lengthy discussions, which included recognition of the importance of including a NIW (Not Inconsistent With) designation to cover cases where the absence of any identified signal in the available data precluded any distinction being drawn between the relative plausibilities of hypotheses, a summary of this information as provided in Table 1 was agreed. It was emphasized that the purpose of this table was not to provide a form of scoring system with "totals" of some form dictating plausibility verdicts, but rather to provide an aid towards an integrative appraisal of all the potential sources of information.

The agreed process for interpreting the results from a set of RMP *Implementation Simulation Trials* requires that a weight of either high, medium or low be accorded to each trial (in cases where there is wide range of views and hence no agreement, the weight of medium applies) (Ref). It is important to recognize that where there are several hypotheses for a particular factor (such as stock structure) in the trials, the aim of the plausibility exercise is not to rank these hypotheses in order, but rather to classify them into one of these three broad categories. The Sub-Committee recognized that this necessitated some distinction between "plausibility" and "weight" of specific trials (for which the plausibility of several factors in combination has also to be addressed). While there might be differences in plausibility between certain hypotheses, the necessity that the RMP trials process provide an adequate check that a recommended RMP variant manifest robustness to uncertainty, means that it may appropriate to have two hypotheses feature in, say, the high weight category, even though they differ somewhat in plausibility. There are only two weight designations (high and medium) that result in trials being retained in the testing process, which means that considerations of balance have also to be taken into account, with weights accorded also factoring in aspects of "importance" in this overall context.

Following lengthy deliberation, the Sub-Committee developed the following views about the four stock structure hypotheses.

*Hypothesis 1:* There was general agreement that this had high plausibility (relative to the other hypotheses), and at baseline level should be accorded a high weight in trials.

*Hypotheses 2 and 3:* It was agreed that these two hypotheses could be treated equivalently for the purposes of plausibility and weight designation. There was agreement that the plausibility of these hypotheses was less than that of Hypothesis 1, but there were varying views as to whether this difference was sufficient to warrant a medium plausibility ranking when only three categories were available for such designation. There was eventual agreement that a high-medium plausibility ranking was appropriate. In the context of weighting for trials, most members of the Sub-Committee were of the view that considerations of balance dictated that a high weight be accorded to the associated trials at the baseline level. However some members believed that according a medium weight would have been more appropriate.

*Hypothesis 4:* It was agreed that this hypothesis did not merit a high plausibility ranking. Some members considered that the ranking should be low. They noted the indications provided by genetic information in Table 1, and commented on the considerable amount of data and extent of associated analyses to support this viewpoint that had become available since the 1998 review (Table 1 in SC/58/Rep1); they thus believed that this hypothesis should be ranked as of low plausibility. Others, while welcoming the provision of these extra data, noted that the absence of genetic evidence for distinction did not preclude the existence of some stock structure. They also referred to the unresolved reasons for apparent differences in age-structure in catches, and accordingly believed that the hypothesis should be accorded medium plausibility. Given this absence of agreement, at the baseline level trials based on

---

<sup>1</sup> the sub-committee was unable to consider this matter in detail in its sessions and it was agreed if any problems arose in the final results to be presented in a working paper by Allison a decision on how to handle this will need to be made accordingly in plenary.

this hypothesis would be accorded medium weight. It was agreed that once the investigations of the ageing data recommended in (Ref) had been conducted, this decision should be revisited.

Table 1: General summary of the information useful to assess plausibility of alternative stock-structure hypotheses. A '+' indicates evidence in favour of a hypothesis, '-' indicates evidence against a hypothesis, '(+)' indicates weak evidence in favour of a hypothesis, '(-)' indicate weak evidence against a hypothesis, a '(())' indicates very weak evidence, and 'NIW' indicates that the evidence is not inconsistent with the hypothesis. Note that the designation NIW often reflects the asymmetrical nature of information on stock structure (i.e., existence of differences can be viewed as positive evidence for multiple stocks, but absence of differences provides no information, and cannot be viewed as positive evidence for a single stock). Notations at the bottom of this table were extracted in part from the report of the 'First Intersessional Workshop' (IWC, 2005).

Evidence	Hypothesis 1	Hypothesis 2	Hypothesis 3	Hypothesis 4
Allozymes	(+) <sup>a</sup>	(-) <sup>a</sup>	(-) <sup>a</sup>	(-) <sup>a</sup>
mtDNA	((+)) <sup>b</sup>	NIW <sup>c</sup>	NIW <sup>c</sup>	_d
Microsatellites	NIW <sup>b</sup>	NIW <sup>c</sup>	NIW <sup>c</sup>	_d
Sighting distribution	NIW <sup>e</sup>	NIW <sup>e</sup>	NIW <sup>e</sup>	NIW <sup>e</sup>
Catch distribution	NIW <sup>e</sup>	NIW <sup>e</sup>	NIW <sup>e</sup>	NIW <sup>e</sup>
External body proportion	NIW <sup>f</sup>	NIW <sup>f</sup>	NIW <sup>f</sup>	NIW <sup>f</sup>
Biological parameters	NIW <sup>f</sup>	NIW <sup>f</sup>	NIW <sup>f</sup>	NIW <sup>f</sup>
Mark-recapture	NIW <sup>g</sup>	NIW <sup>h</sup>	NIW <sup>h</sup>	NIW <sup>g</sup>
Age distribution	NIW <sup>i</sup>	NIW <sup>i</sup>	NIW <sup>i</sup>	(+)/NIW <sup>i,j</sup>

<sup>a</sup>The Sub-Committee agreed that while the power to detect genetic structure might be considered to be low as only a single locus was analysed, the fact that this had been sufficient to detect structure in the Pacific Ocean suggest that if there are multiple stocks differentiated at a level similar to that between the western North stock and other stocks, examination of this single locus should be sufficient to detect this.

<sup>b</sup> Clustering and hypothesis testing analyses of mtDNA and microsatellites were conducted, and as noted by the Committee, none of these analyses revealed any significant heterogeneity within sub-area 1. Only 6 mtDNA samples are available from sub-area 2 – these provide very weak evidence for hypothesis 1, but this is not sufficiently strong to be considered inconsistent with hypotheses 2 and 3.

<sup>d</sup> *Mixing of two sub-stocks*: There are no genetic data for the breeding grounds so the possibility of multiple sub-stocks cannot be excluded. Furthermore if two breeding stocks mix almost completely, it will be difficult to detect differences using, for example, genetics tests based on comparisons between data for the west and east of sub-area 1. However while complete mixing may lead to all methods of detecting stock structure having low power, the plausibility of this was considered fairly low given the behaviour of most large whales. Hypothesis tests based on comparisons for approximately the same area in sub-area 1 found no significant differences among years which suggest that if two sub-stocks mix in sub-area 1, there is little difference in the distribution proportion among years. Although the data set encompasses only four years, such a lack of variation in distribution proportions among years seems unlikely given the known behaviour of whales. In principle evidence for hypotheses 4 could be obtained by testing for deviation from Hardy-Weinberg equilibrium within sub-areas 1W and 1E because such deviations provide evidence of non-random mating as well as selection or migration, i.e. when genetically two different populations are being sampled. Analyses of nuclear markers for Bryde's whales in sub-area 1 have been conducted and these analyses provide no evidence for the significant deviations in Hardy-Weinberg equilibrium within sub-areas 1W and 1E which would provide support for multiple sub-stocks.

*Statistical power*: The Workshop received a paper, which evaluated the power to detect population structure using the chi-square test and Fisher Exact test under an island model in which population differentiation is controlled using a single parameter, Fst. Statistical power of the genetic analysis in sub-area 1 was found to be high for moderate sample sizes and quite small values for Fst, while it was higher for microsatellite data than for mtDNA.

<sup>e</sup>The Sub-Committee examined the sightings data and agreed that it revealed no evidence of a discontinuity in distribution within sub-areas 1 and 2. A similar conclusion was reached with respect to catch distribution. Discontinuity in the commercial catches identified in earlier meetings merely reflected operational constraints. Evidence of spatial discontinuity would be viewed as positive evidence for multiple stocks, but absence of discontinuity is viewed as neutral information, as there is no expectation that their prey would have a discontinuous distribution. Therefore, multiple stocks could exist without spatial discontinuity.

<sup>f</sup>External body proportion data (three features) and several biological parameters (body length, pregnancy rate, length at sexual maturity, seasonality in breeding) had been examined using data from past commercial whaling. Although operational differences (e.g. different minimum length limits for coastal and pelagic whaling) meant that some comparisons could not be made, the authors of these analyses concluded that there were no differences that could not be attributed to operational factors. Evidence of differences in biological parameters would be viewed as positive evidence for multiple stocks, but absence of differences does not provide positive evidence for a single stock, because it is not necessary for separate populations to diverge in their biological parameters, as these are often constrained by their ecology.

<sup>g</sup>Mark-recapture analysis revealed movement of animals within sub-area 1. A very limited number of marks were placed in sub-area 2 and while none has been recovered in sub-area 1, the sample sizes for this sub-area are sufficiently small that even if there is mixing between sub-area 1 and 2, zero recaptures would not be highly unlikely.

<sup>h</sup>Mark-recapture data are available mainly for sub-area 1.

<sup>i</sup>One of the possible explanation for the differences observed in age distribution between sub-area 1W and 1E+2 is that these differences are real. Other possible explanations were geographical segregation by age or non-representativeness of the samples, perhaps as the result of unreadability of certain earplugs, inaccuracy in length determination, or inter-reader differences in ageing.

<sup>j</sup>There was not consensus on this verdict. The different views hinged on different interpretations on an appropriate entry given that there were alternative explanations for the observation in question.

### 3.1.5.2. Other (e.g. MSYR)

The Sub-committee reviewed the *Implementation Simulation Trials* specified in Table 6 of the Report of the Bryde's whale Intersessional Workshop. All 22 trials were retained, and an additional six trials were specified. Four of these new trials used mixing matrix B (instead of D) in combination with the High catch series and with Additional process error (at both MSYR rates). It was agreed that the additional process error would be specified using values from Annex H of the Intersessional Workshop report. The final two additional trials will also use mixing matrix B with the High catch series in combination with Age-specific mixing (at both MSYR rates).

Each trial will be assigned a weight that is equal to the lowest plausibility of any component of that trial. The plausibility of the components of each trial was discussed and the sub-committee agreed to the following assignments.

#### *MSYR*

It was agreed that MSYR=4% would be given high plausibility and MSYR=1% would be given medium plausibility.

#### *Age-dependent mixing*

The age-dependent mixing accounts for the observations of different age distributions in different areas. It was agreed that both the inclusion of age-dependent mixing and the exclusion of it would be given high plausibility.

#### *Additional process error*

It was agreed that the inclusion of additional process error would be given high plausibility.

#### *Stochastic mixing*

It was agreed that stochastic mixing of two stocks in sub-area 1 would be given medium plausibility.

#### *Alternative boundary*

The default in trials is that the boundary between 1-W and 1-E is at the survey boundary. Alternative boundary allows for the stock boundary to be in locations other than the survey boundary. These alternative boundaries were all given medium plausibility.

#### *Low and High catch series*

The low catch series and the high catch series are specified to account for uncertainty in the catch series, mainly due to uncertain species identification between Bryde's and sei whales, along with the possibility that the true catches are higher than the reported catches (Kasuya and Brownell, 2001). In discussion, it was suggested by Perrin that the high catch series was equally plausible as the reported catch series, and both should have high plausibility. As reported previously in IWC 2006, Kasuya and Brownell (2001) had noted the total actual catches off the Ogasawara Islands might be 1.6 times the reported catches but the original records have been lost and are not available. Therefore, it is not possible to examine this issue any further to elucidate which catch series is more likely, and consequently both catch records should be accorded high plausibility. This ensures that the performance of both catch series (Best and High) will be held to the same standards in trials. The Sub-committee agreed that both the Best and High catch series would be given high plausibility.

### 3.1.6 Data/research to reduce hypotheses

Under this item the sub-committee noted its earlier discussions under item 3.1.1, and in particular under item 3.1.5.1, Hypothesis 4.

### 3.1.7 Specification of operational features and management variants

The following four management options will be considered: Management options based on calculating limits by *Small Area*:

- (1) Sub-areas 1W, 1E<sup>2</sup> and 2 are *Small Areas* and catch limits are set by *Small Area*.
- (2) Sub-area 2 is taken to be a *Small Area* and the complete sub-area 1 is treated as a *Small Area*. For this management option, all of the future catches in sub-area 1 are taken from sub-area 1W.

Management options based on applying catch cascading:

- (3) Sub-area 2 is taken to be a *Small Area* and sub-area 1 is taken to be a *Combination area*. Sub-areas 1W and 1E are *Small Areas*, with *catch-cascading* applied.
- (4) Sub-areas 1 and 2 (combined) are taken to be a *Combination area*, and sub-area 2 and sub-areas in 1W and 1E are *Small Areas*, with *catch-cascading* applied.

The simulation application of the RMP is based on using the "best" catch series (see Table 2).

### 3.1.8 Specification and classification of final trials

The full list of specified trials, including the weights of each trial given by the lowest plausibility of any component, is found in Appendix 6 table 6.

### 3.1.9 Work plan

In accordance with the 'Requirements and Guidelines for *Implementations*' (IWC, 2005) plans were made for the Second Intersessional Workshop for the Western North Pacific Bryde's whale *Implementation*. Japan offered to host the meeting, as for the first Intersessional Workshop, to be convened by Kawahara. The cost will be the same as for the first workshop.

---

<sup>2</sup> Defined to be 140°E-165°E and 165°-180°irrespective of the true boundary used to define the structure of the populations in the operating model.

## 3.2 North Atlantic fin whales

### 3.2.1 Report of the joint NAMMCO/IWC workshop

Walløe presented the report of the Joint NAMMCO/IWC Scientific Workshop on the Catch History, Stock Structure and Abundance of North Atlantic Fin Whales (SC/58/Rep3). The main objective of the Workshop was to consider the available information on stock structure, catch history, biological parameters and abundance and trends in order to advance the fin whale assessments ongoing in the two organizations.

Several papers on stock structure, based on both genetic and non-genetic data, were presented at the workshop. A number of key factors emerged that require further work before a full understanding of the contribution of the genetic work to the elaboration of stock structure in the North Atlantic fin whales can be completed, and these were given as recommendations for work to be completed before the IWC meeting in St Kitts. The Workshop then went on to consider the hypotheses with respect to feeding areas, using the schematic figures of Danielsdóttir *et al.* (IWC 2005) as a guide. It is important to stress that the figures are schematic and the location of the 'breeding stocks' is not intended to suggest any specific geographical location. The Workshop considered each of the figures in turn and modified them where appropriate. The Workshop noted that in many cases the discriminatory evidence is weak. The results of these discussions are given in SC/58/Rep3, Fig. 1. The Workshop agreed that pressures of time meant that it had not been possible to fully consider the need for possible further scenarios (*e.g.* incorporating possible north-south structure, alternative links and/or strength of links between breeding stocks and feeding areas, or finer structure within feeding areas).

The Workshop received a complete review of estimates of biological parameters for fin whales, including age and length at sexual maturity, asymptotic length, length at age 5, age at recruitment, mortality rate, ovulation rate and interval and the proportion pregnant in the mature female catch. It was agreed that there was nothing in the review to necessitate change to the parameter values used previously by both the IWC (IWC, 1992) and NAMMCO (NAMMCO 2000, 2001, 2004, 2006) Scientific Committees.

A number of papers detailing catch series for the Northeast and Central Atlantic were presented to the Workshop. It was agreed that there was sufficient uncertainty in the catches, in particular in years when the fin whale catch was estimated from the total catch by species proportion and in years when the struck and lost rate was thought to be appreciable, to warrant development of alternative catch series. It was agreed that the information in the catch series will be used as a basis to develop a 'high' and a 'low' series containing the maximum and minimum catches, however this work could not be completed at the Workshop. In addition the Workshop considered papers detailing various CPUE indices for the Icelandic, Norwegian and Faroese fin whaling operations, and provided a series of recommendations for improvement and better documentation of these indices. It was considered particularly important that papers proposing CPUE series provide adequate documentation of the rationale behind any assumptions made and values chosen and consideration of alternative values and assumptions to capture uncertainty/possible bias. It was recommended that priority be given to investigating whether appropriate CPUE series could be developed for the 'early' (pre-1915) Icelandic whaling operations and Faroese whaling after the 1st World War.

Several papers detailing abundance estimates from international and Norwegian surveys carried out in the Northeast and Central North Atlantic since 1987, as well as recent Canadian and Greenlandic surveys, were presented at the Workshop. The Workshop found the estimates from the NASS and Norwegian surveys for the Central and Northeast Atlantic to be acceptable for use in assessments and agreed that for general purposes the best estimate of current abundance in the Central North Atlantic (including the Faroes) is 25,800 (CV=0.125) for the year 2001. The best estimate for the eastern North Atlantic is 4,100 (CV=0.210) from the 1996-2001 survey series. These estimates are based on the assumption that  $g(0)=1$ . Estimates of  $g(0)$  from recent NASS and Norwegian surveys were presented and fall in the range of 0.7 to 0.9 depending on whether the estimate is for the single or combined platforms. It was considered that for the purposes of assessment the assumption of  $g(0)=1$  was adequate.

The Workshop noted that estimated abundance in the area west and southwest of Iceland increased at an annual rate of 10% (95% CL: 6% - 14%) between 1987 and 2001. This is the area where nearly all fin whaling has been conducted since 1915. Estimated abundance in the whole EGI area has increased at 3% (95% CL: -1% - 7%) per year, *i.e.* this rate of increase is not significant at the 5% level. There was no evidence of any trend in abundance in the eastern North Atlantic.

A new assessment model of the EGI fin whale population, modelled as four subpopulations with movement between areas was presented. The model is sex- and age-structured, and is fitted to CPUE, sightings survey abundance split by area, and mark-recapture data using both maximum likelihood and Bayesian approaches. For the base case and most sensitivity tests, the overall recruited population is increasing and above 80% (base case 84%) of pre-exploitation abundance. The Workshop could not draw firm conclusions from this modelling exercise, but noted that the more complex models involving two or more spatial components did fit the historical and modern CPUE and abundance data better than single homogeneous stock models.

The Workshop provided a series of recommendations for future work that are detailed in its report (SC/58/Rep3) Section 10. It was agreed that all documents submitted to the respective Scientific Committees pertaining to the assessment of fin whales in the North Atlantic, and the Reports of the respective Committees, would be exchanged in the future. The first joint meeting between the NAMMCO and IWC Scientific Committees was considered successful, efficient and productive, and it was hoped that this level of cooperation on issues of common importance could be continued.

In the sub-committee's discussion of the Joint Workshop report, Aguilar noted that most evidence suggested that Mediterranean fin whales were a separate population with little or no exchange with other North Atlantic groups. Records from past whaling operations as well as more recent studies indicate that fin whale densities in the Strait of Gibraltar area were low historically and are so now. Nevertheless evidence from satellite tagging and stable isotope ratios (Guinet *et al.* 2005) indicates that some exchange between the Mediterranean and the Atlantic through the Strait does occur.

### 3.2.2 Report of the intersessional workshop

Vikingsson presented the Report of the Intersessional Working Group on North Atlantic Fin Whales. The Working Group met briefly in St Kitts, but based its report mainly on a meeting held in connection with the Joint NAMMCO/IWC Workshop (see 3.2.1). The conclusions of the Working Group are included where relevant under item 3.2.3 below.

### 3.2.3 Progress on completion of the pre-implementation assessment

### 3.2.3.1 PLAUSIBLE STOCK STRUCTURE HYPOTHESES

The Sub-committee noted that plausible stock hypotheses need only to be specified in broad detail at this stage, that they should be consistent with the data and inclusive enough that it is deemed unlikely that new data collected during the *Implementation* process will suggest a major new hypothesis. New analyses conducted subsequent to the Joint Workshop are detailed below.

SC/58/PFI6 presented results from the population genetic structure analysis of two datasets that were calibrated and combined by Bérubé and Danielsdóttir and contain genetic data of six microsatellite loci (genotypes) and one mtDNA locus (control region sequences). The main objective of this study was to assess further the population genetic structure of North Atlantic fin whales at their feeding locations by combining datasets and therefore enlarge the dataset presented in Bérubé *et al.* (1998). The sample sizes were increased for West Greenland, Iceland and Spain and the geographical coverage was extended by adding new locations, such as Norway and the Faroe Islands. The combined datasets consist of a total of 649 samples from eight North Atlantic fin whale feeding locations (Gulf of Maine (n=31), Gulf of St. Lawrence (n=109), West Greenland (n=56), Iceland (n=129), Faroe Island (19), Norway (38) and Spain (n=92) and Mediterranean Sea (n=74) and in addition as reference samples, samples from the Sea of Cortez (n=75), in the Gulf of California and the North Pacific Ocean (n=13). The combined samples from the North Atlantic and the Mediterranean Sea as well as the combined samples from the North Atlantic (excluding the Mediterranean Sea for the microsatellite loci), deviated significantly from Hardy-Weinberg genotypic proportions due to heterozygote deficiency. Deviations were also observed among the samples collected in the Gulf of St. Lawrence and off Faroe Islands but they were not statistically significant after applying the sequential Bonferroni test. The Slatkin linearized  $F_{ST}$  homogeneity tests, based on the microsatellite loci data within and among the North Atlantic and Mediterranean Sea samples, revealed significant between the different sample years of the Faroe Islands samples, but it should be noted that sample sizes were low. The Faroe Islands samples were the most divergent samples from the remaining North Atlantic Ocean sampling localities (estimates of  $F_{ST}$  ranged from 0.005 to 0.113 among the North Atlantic localities). One possible explanation may be some level of pod structure in fin whales, but each pod would have to have large geographic ranges to explain the overall low level of genetic structure across the North Atlantic. The Faroe Island samples differ from the samples collected in other areas by being collected more or less at the same position in and in a very short time (over a few hours). In all other areas only one or two samples may have been collected at the same time and place. Future data analyses and sample collection need take this possibility into consideration. Significant heterogeneity was also detected between sampling localities in the North Atlantic and Mediterranean Sea against the Sea of Cortez (estimates of  $F_{ST}$ s ranged from 0.325 to 0.528). The BayesAss estimation (Wilson and Rannala 2003) that was undertaken as the rate of current dispersal may be of more immediate management concern than the average rate of gene flow over evolutionary time. The probability that an individual was an immigrant indicated that between 17% and 33% of the individuals in each population are immigrants which suggested large amounts of gene flow among all putative populations/stocks. The amount of migration rates between locations suggested that movement of animals between locations were predominantly unidirectional. The point estimates revealed 20 percent or more of individuals in a single area are immigrants. Such exchange rates are in a range where populations are likely to be demographically correlated, and perhaps should not be viewed as demographically or genetically independent populations. A total of 35 polymorphic sites were detected defining 78 haplotypes of the mtDNA control region sequence data. The estimates of the nucleotide diversity at all North Atlantic sampling localities were all within the same range (0.012-0.014). The nucleotide diversity observed in the Mediterranean Sea and in the North Pacific Ocean samples were significantly lower than any of the observed values at North Atlantic sampling localities. In addition, the nucleotide diversity of 0.001 estimated in the samples from the Sea of Cortez was exceptionally low, and significantly lower than in any other sampling localities. Based on the mtDNA data, significant levels of heterogeneity were detected among some of the North Atlantic samples and also between the Mediterranean Sea and all North Atlantic samples. Significant levels of heterogeneity were detected between Sea of Cortez and all the North Atlantic/Mediterranean Sea sampling localities as well as between North Pacific and all the North Atlantic/Mediterranean Sea sampling localities. However, estimates of genetic divergence among the North Atlantic (and Mediterranean Sea) sampling localities were all very low and [as the nuclear DNA] suggesting high levels of exchange between sampling areas.

SC/58/PFI6 concluded that the genetic analyses based upon nuclear as well as mitochondrial loci all suggested high levels of gene flow among all North Atlantic sampling areas; although both allele and haplotype frequencies were statistically different among the majority of the sampling areas, the actual level of divergence is *very* low. The estimated migration rates were in a range where populations are likely to be demographically correlated, and perhaps should not be viewed as demographically or genetically independent populations. The analyses ignored the signal of exponential population expansions detected in the North Atlantic samples by Bérubé *et al.* (1998). Hence, the high degree of genetic similarity among the North Atlantic sampling areas may be due to recent divergence rather than high gene flow. However, the BayesAss analyses suggest differently. The number of migrants estimated from the  $F_{ST}$  estimates is on the order of 30 migrants per generation. While this rate may initially seem much lower, it should be kept in mind that the estimate is an estimate of the effective number of migrants and hence should be related to the effective population size, which may be lower than the census population size. In addition, the ~30 migrants per generation is between each pair of populations, which results in a much larger number of immigrants once summed for all connected populations.

In discussion the Sub-committee noted that the methodology used in the BayesAss software is relatively new and has not been tested on a wide range of scenarios, particularly ones where the level of differentiation is as low as that found in North Atlantic fin whales. The estimated levels of immigration may be upwardly biased under this circumstance. Donovan noted that this program is being tested with simulated datasets under the TOSSM project.

All samples came from feeding grounds, not from the breeding grounds, the location of which is uncertain. A degree of mixing of different stocks on the feeding grounds might be expected and is indicated by other genetic analyses. The methodology used in BayesAss has not been tested under these circumstances, and it was therefore not clear if dispersal between feeding areas or alternatively the degree of mixing of discrete stocks on the feeding areas was being estimated. Samples from the breeding grounds would be useful in this regard but are not yet available.

SC/58/PFI7 presented results of the genetic analyses of two fin whale datasets that aimed to study further the temporal and micro- and macrogeographical population structure of North Atlantic fin whales sampled at different feeding grounds. At the Joint Workshop the interpretation of  $P$ -values, when values of  $F_{ST}$  were very small, was discussed. The Workshop agreed that this topic required further investigation and discussion and referred the matter to a working group consisting of Skaug, Kitakado and Butterworth, in consultation with Palsbøll, Danielsdóttir and Pastene. In particular it noted that it was important when presenting results of  $F_{ST}$  values that Confidence Intervals (CI) to be calculated (*e.g.* using bootstrapping) for new and previously published data where significant differences have been reported (Árnason *et al.* 1992, Danielsdóttir *et al.* 1991; 1992a; 2005 & 2006). The first dataset was on the genetic variation at nine microsatellite loci in 1022 fin whales sampled at five North Atlantic areas; *i.e.* West off Iceland (n=900), Spain (n=39), Norway (n=54), West off Greenland (n=16) and East off Canada (n=13). The data were based on further statistical analysis of

data presented in Daniëlsdóttir *et al.* (2005 & 2006). The new statistical programs applied, included Confidence Intervals (CI's) of  $F_{ST}$  as suggested at the Joint Workshop, Nm and graphical illustration of potential "populations" as in Waples and Gaggiotti (2006). There was significant heterogeneity within all samples, except the West Greenland samples, all due to heterozygote deficiency. Various genetic analyses resulted in significant genetic heterogeneity among the Icelandic samples, revealing temporal and seasonal differences in the samples from the years 1981-1989. The level of genetic differentiation was low among the 13 samples ( $F_{ST} \sim 0.004$ ,  $p < 0.05$ ). As before greatest significant differences were observed between the Newfoundland Canadian samples and the other locations. The CI's of the pairwise  $F_{ST}$  comparisons varied considerably. When looking at the lower value of the 95% CI of the  $F_{ST}$  values, it reduced the number of the significant pair-wise comparisons down from 26 to 10 of 78 comparisons made, two were comparisons between Icelandic sample years and eight were involving the Newfoundland Canadian and Icelandic samples. However, if the higher value was considered, all comparisons were significant. According to the graphical illustration of potential "populations" as in Waples and Gaggiotti (2006) the number of groups should be five. The first consisting of Icelandic sample years (IC82, IC84, IC85, IC86, IC87, IC88), Norway (NO) Spain (SP) and West Greenland (WG). Then each of IC81, IC83, IC89 and Newfoundland Canada (NFL) as separate groups. As previously suggested, there might be variations in herds at feeding locations in the different sample years of Iceland, contributing to the heterogeneity within and among years and feeding locations or there could be disproportional mixing of different breeding units at the feeding grounds both in time and space.

The second dataset is on the microsatellite genetic variation in 226 fin whales (*Balaenoptera physalus*) sampled in five North Atlantic areas; i.e. off West Iceland, Norway, Spain and West Greenland, a total of: 129, 54, 39 and 16 samples, respectively. The analyses are based on genotypes at 16 microsatellite loci. More number of samples is to be screened so only preliminary statistical analysis of this dataset was presented.

The Sub-committee noted in discussion that the significant heterogeneity between sampling years seen at Iceland and the Faroes suggested either temporal variation in the summer distribution of stocks and/or that different proportional mixtures of stocks were being sampled on the feeding grounds. Such year to year differences may exist in other areas as well but the temporal resolution of the sampling is not sufficient to detect them. The situation at the Faroes was exceptional in that the samples were collected in single events over a short period of time. Therefore it is possible that single "herds" of related individuals may have been sampled in each event.

There was some discussion over the issue of the use of simple  $P$  values vs bootstrap confidence intervals for pairwise  $F_{ST}$  comparisons. Although it was generally agreed that providing confidence intervals for  $F_{ST}$  values is to be encouraged, it was pointed out that the  $P$  value from a standard test of heterogeneity provides a more direct test of the null hypothesis that all samples have been drawn from the same population.

SC/14/PF18 presented results of new statistical analyses of three old fin whale allozyme and carbonic anhydrase datasets previously published in Árnason *et al.* (1992) Daniëlsdóttir *et al.* (1991 & 1992a). At the Joint Workshop the interpretation of  $P$ -values, when values of  $F_{ST}$  were very small, was discussed. The Workshop agreed that this topic required further investigation and discussion and referred the matter to a working group consisting of Skaug, Kitakado and Butterworth, in consultation with Palsbøll, Daniëlsdóttir and Pastene. In particular it noted that it was important when presenting results of  $F_{ST}$  values that Confidence Intervals (CI's) be calculated (*e.g.* using bootstrapping) for new and previously published data where significant differences have been reported (Árnason *et al.* 1992, Daniëlsdóttir *et al.* 1991; 1992a; 2005 & 2006). The first dataset is on the genetic variation at 11 variable allozyme loci in 328 fin whales (*Balaenoptera physalus*) sampled in two North Atlantic areas; i.e. off West Iceland and Spain, a total of: 283 (Iceland 1985, 1986, 1987 and 1988) and 46 (Spain 1985) samples, respectively. The data is based on further analysis of genotypes at 11 allozyme loci: *Ada*, *Ak-1*, *Gpd*, *Ldh-A*, *Mdh-S*, *Mpi-1*, *Pep-A*, *Pgm-1*, *Pgi* and *Sod-A*. The new statistical programs applied included  $F_{ST}$ , Confidence Intervals (CI's) of  $F_{ST}$ , Nm, PCA, MDS, STRUCTURE and graphical illustration of potential "populations" as in Waples and Gaggiotti (2006). There was significant heterogeneity within each sample and overall samples, all due to heterozygote deficiency. Various genetic analyses on the dataset resulted in high significant genetic heterogeneity among the Icelandic and Spanish samples, and temporal differences in the Icelandic samples from the years 1981-1988 ( $F_{ST} \sim 0.078$ ,  $p < 0.0001$ ) as well as differences between Icelandic and Spanish samples ( $F_{ST} \sim 0.094$ ,  $p < 0.0001$ ). Divergence between Icelandic sample years was less than between Icelandic and Spanish samples. The Spanish samples were overall the most divergent of the samples and with lower Nm than among the Icelandic sample years. The CI's of the pairwise  $F_{ST}$  varied considerably and were larger in comparisons of larger  $F_{ST}$  i.e. between Icelandic and Spanish samples than between Icelandic sample years. This could indicate that the observed level genetic divergence between groups is less than concluded before from the previous results of high  $F_{ST}$  and significance between groups, however when considering the lower CI values, all comparisons remained significant. Based on Nei's (1978) genetic distances, multidimensional scaling (MDS) analysis revealed genetic divergence with a stress value of 0.0229 between axes 1 and 2. The STRUCTURE analysis indicated two groups among the samples. As previously suggested, there might be variations in herds at feeding locations in the different sample years of Iceland, contributing to the heterogeneity within and among years and feeding locations. The genetic heterogeneity within and between temporal Icelandic samples and geographical samples in addition with the likelihood of number of breeding units estimated in STRUCTURE both are in agreement with the hypothesis that the North Atlantic fin whale is genetically structured on the feeding grounds and different breeding units might mix disproportionately at the feeding grounds both in time and space. The heterogeneity among Icelandic samples years could therefore also be disproportional mixing of two breeding units represented in the different years.

The second dataset (II) is on the genetic variation at five allozyme loci (*Ak-2*, *Est*, *Ldh-A*, *Mdh-S* and *Pgi*) of 67 fin whales from three North Atlantic areas; i.e. Newfoundland Canada ( $n=24$ ), Norway ( $n=19$ ) and off West Iceland ( $n=24$ ). There were statistical significant heterogeneity within the pooled samples and genetic divergences were found among Newfoundland Canadian, Norwegian and Icelandic samples ( $F_{ST} \sim 0.343$ ,  $P < 0.0001$ ). All three  $F_{ST}$  pair-wise comparisons varied considerably in their CI's indicating variation in the estimation of significant comparisons from two to all significant comparisons of three made and reducing the earlier level of observed genetic differentiation. The STRUCTURE analysis indicated three clusters among the samples. The results from this study based on allozyme loci indicate that the fin whale samples from the feeding grounds off Iceland, Norway, Spain and Newfoundland Canada are significantly heterogenous, but genetic divergence is high between groups. The genetic heterogeneity within and between the samples in addition with the likelihood of number of breeding units estimated in STRUCTURE both are in agreement with the hypothesis that the North Atlantic fin whale is genetically structured on the feeding grounds and different breeding units might mix disproportionately at the feeding grounds.

The level of genetic differentiation was higher for the allozyme than the microsatellite data (Daniëlsdóttir *et al.* 2005; 2006; Bérubé *et al.* 1998; SC/58/PF16). To better estimate the true unbiased geneflow among the samples, further analyses need to include estimation on effective population size ( $N_e$ ) to have a better estimate of effective number of migrants per generations ( $N_m$ ), BayesAss to estimate current migration rates and MIGRATE to estimate historical migration rates.

The third dataset (III) is on the *Ca* locus genotype variation in 1,159 fin whales sampled in two North Atlantic areas; i.e. off Spain (1983 & 1984) and West Iceland (1971 and 1981-1989), a total of 26 and 1133 samples, respectively. There was heterogeneity in two Icelandic sample years, but no overall significant heterogeneity was observed at this locus within or among samples ( $F_{ST} \sim 0.003$ ,  $P > 0.05$ ). The Spanish 1983 samples showed greatest overall  $F_{ST}$  values compared to all other locations and the least number of migrants,  $N_m$ .

Skaug presented SC/58/PFI9 which detailed the screening of a dataset consisting of 15 microsatellite loci from 226 fin whales from several North Atlantic locations for closely related individuals. Five pairs of individuals were identified as being closely related, four of which were consistent with a parent-offspring relationship. Two of these parent-offspring pairs had been conjectured to be mother-calf pairs when the biopsies were obtained. Of the two pairs that were not sampled at the same location, one showed a linkage between North Norway and the area west of Svalbard, and the other between North Norway and West Iceland.

The Sub-committee agreed that this method shows promise for identifying relatives in genetic samples and can assist in the interpretation of analyses related to stock boundaries and degree of mixing among areas. It was noted that the number of matches of closely related individuals would increase with sample size, and the certainty of detection with the number of microsatellite loci included in the analysis. In response to a query Skaug noted that it would be possible to determine if the number of related individuals in a sample was greater or less than expected given the assumption of random mixing.

Kitakado presented the results of a preliminary analysis using a new method aimed at simultaneous estimation of mixing proportions and genetic differentiation for stocks for North Atlantic fin whales. The method uses individual genotypes at multiple loci and does not require presence of baseline stocks. An integrated likelihood function with elimination of nuisance parameters was employed to estimate the parameters, and then the maximum values under one and two stock scenarios were compared to determine the likely number of stocks. To investigate model performance, a simulation study was conducted under the one stock scenario and with two simulated stocks with  $F_{ST} = 0.1$ . The model successfully discriminated the 1 and 2 stock situations. For analysis for fin whales, the same data were used as in SC/58/PFI7 (226 individual's genotypes at 17 microsatellite loci). The result was consistent with one stock in the area. Kitakado emphasized the preliminary nature of this analysis and that further investigations are required.

In response to a query Kitakado noted that while the method is similar to STRUCTURE in that it estimates the number of genetic stocks present in a sample without reference to sample origin, the estimation methods employed are different. In addition this method provides explicit parameters for mixing proportions of the putative stocks identified.

The  $F_{ST}$  used in the simulation study was much higher than that commonly observed between fin whale sampling areas. The ability of the model to detect population structure when the populations are so weakly differentiated has not been tested. The Sub-committee considered this method promising and recommended that Kitakado continue simulation testing of the model under levels of differentiation observed in whale populations. In addition the method should be applied to other species and stocks for which stock structure is better known, such as bowhead and gray whales.

### Conclusions

The Report of the Joint Workshop (SC/58/Rep3) provided a range of stock hypotheses and recommendations for further genetic work to refine or suggest new ones. The sub-committee was gratified to see that much of this work had been accomplished and thanked the authors of PFI6-8 and of the Report of the Intersessional Working Group for their hard work in fulfilling these recommendations in the short period since the Joint Workshop. Considering the new information brought forward, the Sub-committee found no reason to modify any of the existing stock hypotheses, or to suggest new ones.

#### 3.2.3.2 PRELIMINARY INVESTIGATION OF EXPERIMENTAL WAYS TO DISTINGUISH AMONG COMPETING HYPOTHESES

It was noted that recommendations for further analytical work to distinguish among competing stock hypotheses should be provided at this meeting so that the work could be carried out in time for the first intersessional meeting for the *Implementation assessment* if it proceeds.

Photo-ID work was identified as a potentially useful method, but it was noted that there are only two catalogues available, for the Gulf of Maine and the Mediterranean. Robbins agreed to compile all information on photo ID for the Sub-committee.

The sub-committee agreed that obtaining samples for genetic analysis from the breeding grounds would be very valuable for identifying fin whale stocks, which might then be distinguishable on the feeding grounds. However the location(s) of the breeding grounds are presently unknown. Satellite tags applied late in the season on the feeding ground might be very useful in identifying fin whale breeding areas. In addition such applications could provide information on migration routes and movements between feeding areas. However it was recognized that the rate of success in fin whale tagging, in terms of tag functioning and duration, had been low in previous attempts, and that these issues must be resolved before tagging could proceed on a large scale.

The new analytical methods for genetic data used in SC/58/PFI 6 (BayesAss), 7 and 8 (STRUCTURE), and in the Report of the Intersessional Workshop, were considered promising, but in all cases they should be tested using simulated datasets showing similar levels of genetic variation to that observed in fin whales. It was hoped that this might occur under the TOSSM project but the sub-committee recognized that the results might not be available for the prospective *Implementation Assessment*.

#### 3.2.3.3 DISPERSAL RATES

$F_{ST}$  values and other data provided by genetic studies will need to be used to produce realistic ranges of dispersal rates for input into trials. The sub-committee noted that dispersal rates refer to permanent movement of individuals between breeding stocks and differ from mixing proportions which refer to situations where more than one breeding stock feeds in a particular area; trials need estimates of both these quantities. This work requires careful consideration and must be completed before the First Intersessional Workshop once an *Implementation* has begun. The issue was referred to an intersessional working group consisting of Palsbøll, Skaug and Waples.

### 3.2.3.4 ABUNDANCE ESTIMATES (INCLUDING G(0) ISSUES AND PLANS FOR FUTURE SURVEYS)

Vikingsson confirmed that all survey data will become available in the IWC database as per the data availability guidelines. It was concluded that these were of sufficient temporal and spatial resolution to calculate abundance estimates both on the scale of the sub-areas that would be likely to be used in conditioning simulation trials and for use in the *CLA*.

It had been noted at the Joint Workshop (SC/58/Rep3) that no recent abundance data are available for the American coastline. However Palka informed the Intersessional Working Group that there are estimates available for the Gulf of Maine – Bay of Fundy region and that these could be made available to the Scientific Committee.

Abundance estimates from the 1987, 1989, 1995 and 2001 North Atlantic Sightings Surveys (NASS) as well as the 1996-2001 Norwegian surveys were reviewed by the Joint Workshop and found to be of sufficient quality for use in simulation trials and in the *CLA*. The issue of  $g(0)$  for fin whale ship surveys was also examined by the Joint Workshop, which concluded that  $g(0)$  was close to 1 and the assumption of  $g(0) = 1$  would be sufficient for trials.

The sub-committee concurred with these conclusions of the Joint Workshop.

Vikingsson informed the sub-committee that a new survey would be carried out in 2007, and that Iceland planned to continue large-scale surveys at 5-6 year intervals. Plans for the Trans North Atlantic Sightings Survey (TNASS, see 4.2), which will cover a large part of the northern North Atlantic and includes the participation of Norway, the Faroe Islands, the Russian Federation, Iceland, Greenland and Canada, were provided in SC/58O21. A synoptic redfish survey will take place in 2007, coordinated by the ICES Study Group on Redfish Stocks (SGRS), with international participation from several countries including Iceland. The Icelandic vessel will be used as a cetacean survey platform. The sub-committee endorsed this collaborative effort and recommended that the Commission encourage the relevant governments participating in the international redfish survey to include a cetacean component in the survey.

### 3.2.3.5 CATCH DATA (INCLUDING ALTERNATIVE SERIES)

Allison confirmed that the catch history for North Atlantic fin whales was sufficiently well known to allow a catch series to be developed for use in trials, and that work on this task was in progress. This compilation will include notations on data quality and available ancillary information. There are sufficient uncertainties in the catch series (allocation of catches between species, catch location, struck and lost) to anticipate the use of alternative catch series, and the compilation of catch data will be used to develop such alternative series.

SC/58/PFI4 presented CPUE series for fin whales during the early modern whaling operation in Iceland (1983-1915) that have been revised from SC/M06/FW13, based on the recommendations drawn up at the Joint Workshop. The series are split by the west (Vestfjord) and East coast operation. The FpBM series for the East coast (fin whales per boat month) is based on individual catch records from 1904 to 1913 restricted to the period 14<sup>th</sup> of May to 11<sup>th</sup> August. This series might be used only from 1907 when fin whale catches exceeded 50% of the total. Individual catch records are too limited for the Vestfjord operation so total fin whale catch per boat has to be used (FpB). Fin whale catches exceed 50% continuously from 1902, which is considered a reasonable starting point in this case. The operational range expanded over time and in particular in the last year in 1915 when a large part of the catches were taken on the Greenland side of the Iceland - Greenland midline, which suggests that this year should not be included. Two alternatives in correcting for handling times are given for both series. The season length is generally not known in the Vestfjord operation, but is needed for this correction and is assumed to have been 4 months based on all information available. An uncorrected catch per boat series (CpB) is also presented and is considered a reasonable index for the Vestfjord operation although not accounting for handling times dampens the trend therein. Operational changes over time and limitations of the data are believed to be more likely to mask the real trend in the population.

The Sub-committee thanked the authors for their work in responding to the recommendations of the Joint Workshop.

### 3.2.3.6 FUTURE WHALING OPERATIONS

This RMP implementation is requested by Iceland, and the only likely commercial whaling operation in the near future at least is on the traditional fin whaling grounds off west Iceland. The operational factors in this fishery have been described by Sigurjónsson (1988). Since the ban on whaling was imposed by the Icelandic parliament in 1915, Icelandic whaling for large whales has been limited to a single land station in Hvalfjörður, Faxaflói in west Iceland apart from limited catches from a single whaling station in Northwest Iceland during 1935-1939. The whaling station in Hvalfjörður operated during 1948-1989 and has been maintained so that future large whaling operations are likely to be restricted to this station. Company regulations restricted the whaling operation in various ways. Thus, the processing capacity of the factory limited the number of whaling vessels, operating at any one time, to four and further restrictions on catch rates were imposed temporarily during busy periods (Sigurjónsson 1988). The Hvalfjörður whaling station is the only one in Iceland and there are no plans to build whaling stations in other parts of Iceland. It can therefore be assumed that future whaling will only be conducted from this single land station and that the operating area of the whaling vessels will be restricted to the same area as during the commercial whaling 1948-1985. The whaling area off Iceland during 1951-1973 is shown in Fig. 2 of PFI5. Some of the catches were taken west of the present 200nm EEZ for Iceland, but any future catches would likely be within this 200nm limit. The whaling seasons from the whaling station in Hvalfjörður usually started in the latter half of May and ended near the end of September. Since 1968 the whaling season started in the first half of June (Rørvik *et al.* 1976). Future catches under the RMP would likely be confined to the period June-September.

The sub-committee noted that there are no plans to initiate commercial whaling for fin whales by other nations that caught fin whales in the North Atlantic prior to the onset of the moratorium such as Spain, the Faroes, Norway and Canada. The possibility of a very limited catch (5-10 animals) for scientific purposes off Norway and the Faroes in the medium future cannot be precluded and might thus be included as a sensitivity test (Bjarni Mikkelsen, Lars Walløe pers. commn).

### 3.2.3.7 OTHER ANTHROPOGENIC REMOVALS

SC/58/PFI5 summarized non-natural mortality of fin whales, other than direct catch, as compiled from National Progress reports covering the period 1997-2004 (IWC website). During this eight year period a total of 11 fin whales was reported killed from fisheries interactions, four in the North Atlantic and two in the Mediterranean (see Table 1 of SC/58/PFI5). In the same period nine fin whales were reported killed or likely killed by ship strikes. Eight of these were in the North Atlantic.

Read *et al.* (2004) used bycatch statistics from the USA to make crude extrapolations to estimate global bycatch of large cetaceans. Of 201 reported bycatches in US waters during 1990-1999, six were fin whales.

The sub-committee noted that since systematic monitoring of cetaceans was initiated in Iceland around 1980 there have been no reported bycatches of fin whales off Iceland, in contrast to many other species of cetaceans in this area. Only one stranding of a fin whale is reported for this period, a calf that stranded near a small boat harbour in 1994. The animal had wounds on the head that might have been caused by a ship strike (MRI unpublished information). This is the only known example for fin whales off Iceland of anthropogenic mortality other than direct catch. Apart from a single bycatch off Ireland and a ship strike in the Mid North Atlantic, it is the only example for the whole Northeastern Atlantic (east of W-Greenland). Somewhat higher rates in the Western North Atlantic may be due to the apparently more coastal distribution of the species in this region.

Berggren informed the Sub-committee that the Sub-committee on Bycatch had received some new information on ship strikes of fin whales in the Mediterranean (SC/58/For Info12 and SC/58/For Info37) and Strait of Gibraltar (SC/58/BC5) this year, which indicates that this source of mortality may be greater in this area than in others. However he emphasized that rate of reporting may be low in all areas. There is no information to suggest a high rate of lethal bycatch of fin whales in the North Atlantic.

The sub-committee recognised that although the efficiency of reporting schemes is difficult to evaluate with precision, it can be concluded from the available evidence that non-natural mortality of fin whales in the North Atlantic (outside the Mediterranean) is insignificant compared to abundance. The Sub-committee concluded that there is no need to model incidental catches in the *Implementation Simulation Trials*.

#### 3.2.3.8 CONDITIONING (INCL. BIOLOGICAL PARAMETERS)

SC/58/PFI3 presented some new investigations of life history parameters based on data from biological samples of fin whales caught off Iceland from 1967-1985 and scientific catches from 1986-1989, originally presented in Lockyer and Sigurjónsson (1991). The classification of animals with respect to maturity and pregnancy stage when foetuses and/or ovaries were lost or damaged, that then had to be based on other measurements, is validated. Samples were collected only for different parts of the season in the first years and in two years appear selective with respect to size. The scientific catches show some differences from the commercial catches. The proportion pregnant changes over the season, and with age, so the imbalance in the samples needs to be considered in estimates of pregnancy rate. The changes in fecundity with age are best demonstrated by comparing females with a different number of ovarian corpora. After the first pregnancy the females are estimated to rest for 1.64 years. Of these the younger animals appear to rest for longer which indicates that early maturity may come at some cost. The resting period decreases to 1.37 years after 10 to 15 ovulations but then increases again to 1.74 years in animals after 21 or more ovulations, indicating reproductive senescence (also reported by Kjeld *et al.* 2006). In addition, no foetuses were found in the oldest animals that appeared to be pregnant from inspection of the ovaries. Fewer of the younger females caught were in the later stages of pregnancy, which implies that they need a longer time feeding to build up the energy required for pregnancy and lactation. The earlier findings of changes in the age at maturity based on the proportion mature and proportion of first ovulators, are confirmed and are in line with estimates based on transition phase readings in the ear plugs (Konradsson *et al.* 1991). Estimates of mortality are derived based on an assumption of constant cohort size. The estimate of natural mortality is from 0 to 0.07. It is suggested that a high proportion of maturing animals in the most recent catches and changes in the age distribution of the catch can best be explained by an influx of immature animals.

The Sub-committee thanked the authors for this compilation and noted that it will be of use in the *implementation* process. It was agreed that the available biological data, including estimates of relevant biological parameters, was sufficient for *Implementation Simulation Trials*.

#### 3.2.4 Work plan

The sub-committee concluded that the only outstanding item to be completed before proceeding towards an *Implementation* was the development of a list of catches with associated ancillary information that would allow the development of a best and alternative catch series for use in simulation trials. Considerable progress towards this had been made at the Joint Workshop and it was anticipated that this could be completed in time to begin an *Implementation* this year. However, it was considered impractical to begin the *Implementation* this year because of a lack of resources and the high workload of the Committee, which must complete the North Pacific Bryde's whale *Implementation* as a priority. Therefore the Sub-committee proposed that the *Implementation* for North Atlantic fin whales be initiated in 2007. It was emphasized that this delay in initiation of the *Implementation* was due to the priorities and workload of the Scientific Committee rather than a lack of sufficient preparation on the part of the initiating Member State, Iceland.

In the interim period the sub-committee recommended the following priorities for further research:

1. Completion of a comprehensive list of catches as noted above;
2. Refinement and extension of genetic and other analyses to discriminate between existing stock hypotheses and to estimate mixing proportions and dispersal rates.

## 4. CONSIDERATION OF SURVEYS UNDER THE GUIDELINES AND REQUIREMENTS

### 4.1 Norwegian surveys

Walløe presented results from SC/58/RMP4, which was a report of a Norwegian 2005 survey for minke whales in *Small Management Area* CM around Jan Mayen. As part of a six-year program over the period 2002-2007 with the aim to get a new estimate of minke whale abundance in the Northeast Atlantic, the areas around Jan Mayen in the Greenland Sea, the *Small Management Area* CM, was surveyed with two vessels during the summer 2005. There were three planned blocks with a planned basic coverage of about 2,000 nautical miles. In total about 2,100 nautical miles were searched with primary effort. The most common species sighted were minke whales (67 groups seen from the primary platform), humpback whales (33 groups), fin whales (27 groups) and sperm whales (18 groups). In addition sightings were made of killer whales, Northern bottlenose whales, blue whales and unspecified dolphins. Opportunistic collections were made of biopsy samples from one minke whale and two fin whales, and photo IDs were collected from 16 humpback whales.

The sub-committee welcomed the report of this new cruise. Polacheck raised the issue of whether new information on dive time had been collected, and noted that this issue had been raised in the past. The sub-committee recommended that new data on dive time should be collected during future surveys. Such data had been requested in the past, and it had been stated in cruise plans that such data would be collected. Walløe indicated that dive time data would be collected in 2006. With regards to review of the 2005 survey, Walløe suggested that the results of these new surveys do not need to

be reviewed now, as full results, including data analysis, will be presented in 2008, and that would be the appropriate time to review and consider approval of these surveys for the RMP. Official Scientific Committee oversight for the 2006 Norwegian survey was designated to Øien.

#### 4.2 Other

Víkingsson raised the issue of planning for a joint survey in 2007 in late June. It was noted that any part of a survey planned to provide data for use in the RMP needed to have its cruise plans submitted to the Scientific Committee for consideration prior to the survey being carried out. The sub-committee agreed it would review in detail the cruise plans for this survey in 2007. The planners of the survey were reminded that there are specific guidelines to be followed in submitting cruise plans for survey data to be used in an RMP implementation.

The next North Atlantic Sightings Survey (NASS) will be conducted in the summer of 2007. To date four NASS have been conducted, in 1987, 1989, 1995 and 2001. It is hoped that by participation of Greenland and Canada in addition to Iceland, Faroes and Norway and coordination with a simultaneous survey planned off the western coast of Europe (CODA), this survey will cover a larger area than any of the earlier surveys (SC/58/O21). The first planning meeting for the 2007 survey (TNASS) was held in Reykjavik in March 2006. The sub-committee noted that the NASS have used standard methodology (line-transect and cue counting) and covered large areas of the central and eastern North Atlantic. The methodology is in accordance with the requirements and guidelines for conducting surveys and analysing data within the RMP (IWC 2005) and abundance estimates from NASS for common minke whales and fin whales have been accepted for use in the RMP. Official Scientific Committee oversight for the 2007 survey was designated to Hammond and Donovan.

#### 5. OTHER BUSINESS

No items were raised under this item.

#### 6. ADOPTION OF REPORT

The report was adopted at 0835hrs, 3 June 2006. The sub-committee expressed its appreciation of the work of the Rapporteurs and Chair.

#### REFERENCES

- Árnason, A., Daníelsdóttir, A.K., Spilliaert, R., Sigurðsson, J.H., Jónsdóttir, S., Pálsdóttir, A., Duke, E.J., Joyce, P., Groves, J. and Trowsdale, J. 1992. A brief review of protein and DNA marker studies in relation to the stock identity of fin whales from Iceland and Spain. *Rep. int. whal. Commn* 42:701-05. (SC/F91/F16).
- Groves, V. and Trowsdale, J. 1992. A brief review of protein and DNA marker studies in relation to the stock identity of fin whales (*Balaenoptera physalus*) from Iceland and Spain. *Rep. int. Whal. Commn* 42: 701-705.
- Bérubé, M., Aguilar, A., Dendanto, D., Larsen, F., Notarbartolo di Sciarra, G., Sears, R., Sigurjónsson, J., Urban-R J. and Palsbøll, P.J. 1998. Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): analysis of mitochondrial and nuclear loci. *Molecular Ecology* 7: 585-599.
- Daníelsdóttir, A. K. Duke, E.J. Joyce, P. and Árnason, A. 1991. Preliminary studies on the genetic variation at enzyme loci in fin whales (*Balaenoptera physalus*) and sei whales (*Balaenoptera borealis*) from the North Atlantic. *Rep. int. Whal. Commn* (Special issue 13):115-124.
- Daníelsdóttir, A.K., Bérubé, M., Palsbøll, P.J., Stefánsson, M.Ö., Thorgilsson, B., Jörundsdóttir, Th.D. Ragnarsdóttir, A., Árnason, A., Gunnlaugsson, Th., Ólafsdóttir, D., Øien, N., Witting, L., Pampoulie C. and Víkingsson, G.A. 2006. Genetic stock delineation of fin whales. Paper presented at the joint North Atlantic Marine Mammal Commission (NAMMCO) and International Whaling Commission (IWC) Scientific workshop on the catch history, stock structure and abundance of North Atlantic fin whales, 23<sup>rd</sup> March – 26<sup>th</sup> March 2006, Reykjavík, Iceland. SC/58/FW/3 – SC/M06/FW5. 25pp.
- Daníelsdóttir, A.K., Sigurjónsson, J., Mitchell, E. and Árnason, A. 1992a. Report on a pilot study of genetic variation in skin samples of North Atlantic fin whales (*Balaenoptera physalus*). Paper SC/44/NAB16 presented to the IWC Scientific Committee, June 1992, 8pp.
- Daníelsdóttir, A.K., Stefánsson, M.Ö., Thorgilsson, B., Jörundsdóttir, Th.D. Ragnarsdóttir, A., Árnason, A., Gunnlaugsson, Th., Víkingsson, G.A. Ólafsdóttir, D., Bérubé, M., Palsbøll, P.J., Øien, N., Witting, L. and Pampoulie, C. 2005. Genetic analysis of North Atlantic fin whales (*Balaenoptera physalus*): Is there more than one breeding unit at the feeding ground west off Iceland? Paper presented at the *International Whaling Commission (IWC) Scientific Meeting*, 30<sup>th</sup> May – 8<sup>th</sup> June 2005, Ulsan, Republic of Korea. IWC SC/57/PFI4, 12pp.
- Guinet, C., Mate, B., Benteleb, I., André, J.-M., Mayzaud, P., Stephanis De, R. 2005. Where are the Mediterranean fin whales when the summer is over? *Proceedings from the 19th annual conference of the European Cetacean society and associated workshops* April 2-7, 2005, La Rochelle, France (abstract): 24.
- International Whaling Commission. 1991. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on Management Procedures. *Rep. int. Whal. Commn.* 41: 109.
- International Whaling Commission. 1992a. Report of the Fourth Comprehensive Assessment Workshop on Management Procedures. *Rep. int. Whal. Commn.* 42: 312.
- International Whaling Commission. 1992b. Report of the Fourth Comprehensive Assessment Workshop on Management Procedures. *Rep. int. Whal. Commn.* 42: 334.
- International Whaling Commission. 1992c. Report of the Scientific Committee. *Rep. int. Whal. Commn.* 42: 55.
- International Whaling Commission. 1992d. Chairman's Report of the Forty-Third Meeting. Resolution on the Revised Management Procedure. *Rep. int. Whal. Commn.* 42: 47-48.
- International Whaling Commission 1992. International Whaling Commission 1992. Report of the comprehensive assessment special meeting on North Atlantic fin whales. *Rep.int.Whal.Commn* 42:595-651

- International Whaling Commission. 1993. Report of the Special Meeting of the Scientific Committee on the Revised Management Procedure. *Rep. int. Whal. Commn* 43: 223-39.
- International Whaling Commission. 1993. Report of the Scientific Committee. Annex E. Report of the sub-committee on Southern Hemisphere baleen whales. *Rep. int. Whal. Commn.* 43: 104-114.
- International Whaling Commission. 1994. Report of the Scientific Committee. *Rep. int. Whal. Commn* 44: 47.
- International Whaling Commission. 1998. Report of the Scientific Committee. *Rep. int. Whal. Commn* 48: 87.
- International Whaling Commission. 2004a. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage.* 6 (Suppl.): 75-184.
- International Whaling Commission. 2004b. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. Appendix 14. Report of the Working Group on North Atlantic minke whales Revised Management Procedure Implementation Review. *J. Cetacean Res. Manage.* 6 (Suppl.): 171-179.
- International Whaling Commission. 2005. Requirements and Guidelines for Implementations. *J. Cetacean Res. Manage.* 7 (Suppl.): 84-92.
- 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.
- Kasuya, T. and Brownell, R.B. Jr. 2001. Illegal Japanese coastal whaling and other manipulation of catch records. Paper SC/53/RMP24 presented to the IWC Scientific Committee, July 2001, London. 44pp.
- Kjeld, M., Ólafsson, Ö. Víkingsson, G.A. and Sigurjónsson, J. 2006. Sex hormones and reproductive status of the North Atlantic fin whales (*Balaenoptera physalus*) during the feeding season. *Aquatic Mammals* 32(1), 75-84.
- Konradsson, A. Sigurjónsson, J. and Gunnlaugsson, Th. 1991. Trends in age at sexual maturity in fin whales off Iceland based on transition phase in ear plugs. IWC SC/F91/F19 (unpublished) 14pp.
- Lockyer, C. and Sigurjónsson, J. 1991 The Icelandic fin whale (*Balaenoptera physalus*): biological parameters and their trend over time. IWC SC/F91/F8 (unpublished) 51pp.
- Nei, M. 1978 Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.
- North Atlantic Marine Mammal Commission. 2000. Report of the NAMMCO Scientific Committee Working Group on North Atlantic Fin Whales. Report of the Scientific Committee, Annex 2. In: *NAMMCO Annual Report 1999*, NAMMCO, Tromsø, pp. 189-211.
- North Atlantic Marine Mammal Commission. 2001. Report of the NAMMCO Scientific Committee Working Group on North Atlantic Fin Whales. Report of the Scientific Committee, Annex 4. In: *NAMMCO Annual Report 2000*, NAMMCO, Tromsø, pp. 274-294.
- North Atlantic Marine Mammal Commission. 2004. Report of the NAMMCO Scientific Committee Working Group on Minke and Fin Whales. Report of the Scientific Committee, Annex 1. In: *NAMMCO Annual Report 2003*, NAMMCO, Tromsø, pp. 187-229.
- North Atlantic Marine Mammal Commission. 2006. Report of the NAMMCO Scientific Committee Working Group on North Atlantic Fin Whales. Report of the Scientific Committee, Annex 2. In: *NAMMCO Annual Report 2005*, NAMMCO, Tromsø, In Prep.
- Punt, A.E. 2005. Some thoughts related to defining criteria for evaluating the performance for *Implementation Simulation Trials*. Paper SC/57/RMP1 presented to the IWC Scientific Committee, May 2005 (unpublished) 5pp.
- Read, A.J., Drinker, P., Northridge, S. 2004. By-catches of large whales in US fisheries from 1990-1999. 7pp. Paper SC/56/BC8 presented to the IWC Scientific Committee July 2004, Sorrento, Italy
- Rorvik, C.J., Jonsson, J., Mathiesen, O.A. and Jonsgard, A. 1976. Fin whales off the west coast of Iceland-distance, segregation by length and exploitation. *Rit Fisk.* 5(5):1-30.
- Waples, R.S., Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15(6):1419-1439.
- Wilson GA, and Rannala, B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163(3): 1177-1191.