Annex D1

Report of the Working Group on the *Implementation Review* for Western North Pacific Common Minke Whales

Members: Hammond (Convenor), Allison, An, Baba, Baker, Baulch, Bjørge, Brownell, Butterworth, Chilvers, Cipriano, Cooke, de la Mare, de Moor, Deimer-Schüette, Di Guardo, Donovan, Double, Elvarsson, Funahashi, Gaggiotti, Gunnlaugsson, Hakamada, Hiruma, Hoelzel, Iñíguez, Jaramillo-Legorreta, Kanda, Kasuya, Kato, Katsuyama, Kelly, Kim, G.H., Kim, H.W., Kishiro, Kitakado, Lang, Leaper, Leslie, Miyashita, Murase, Øien, Okamura, Palka, Palsbøll, Pampoulie, Park, J.Y., Park, K.J., Pastene, Punt, Sakamoto, Skaug, Sohn, Tajima, Tiedemann, Uozumi, Wade, Walløe, Waples, Yamada, Yasokawa.

1. INTRODUCTORY ITEMS

1.1 Election of Chair and appointment of rapporteurs
Hammond was elected Chair. Punt and Waples were appointed as rapporteurs.

1.2 Chair's opening remarks
Hammond reminded members that the Working Group was following the schedule for an *Implementation (Review)* (IWC, 2012c). Last year's meeting was the scheduled First Annual Meeting but the Working Group had been unable to complete the necessary business. It had not been possible to assign plausibility to stock structure hypotheses, primarily because it had not been possible to complete conditioning of the trials. This meant that the two-year schedule for the *Implementation Review* had been disrupted.

This year's meeting is effectively a repeat of the First Annual Meeting with the same list of tasks that had been initiated last year. There had been another intersessional meeting in December 2011 to facilitate the work necessary to ensure that all relevant tasks could be completed at this year's meeting.

1.3 Adoption of Agenda
The adopted Agenda is given in Appendix 1.

1.4 Review of documents
The documents considered by the sub-committee were SC/64/NPM1-11; SC/64/O9-10; SC/64/Rep2, and relevant extracts from past reports of the Committee.

2. OBJECTIVES OF THE FIRST ANNUAL MEETING

The Chair drew attention to the Committee's Requirements and Guidelines for *Implementations* relevant to the tasks to be undertaken at the First Annual Meeting (IWC, 2012b, p.102) and to Appendix 2 to last year's report that describes the remaining steps necessary to complete the *Implementation Review* (IWC, 2012b, pp.117-20).

3. REPORT OF THE DECEMBER 2011 INTERSESSIONAL WORKSHOP

Donovan presented a summary of the report of the second 'First Intersessional Workshop' for the *Implementation Review* of western North Pacific common minke whales held 12-16 December 2012, kindly hosted by the Government of Japan (SC/64/Rep2). The primary objective of the Workshop was to ensure completion of the conditioning of trials by the 2012 Annual Meeting, although a number of other topics were addressed to assist the Committee in its work to complete the *Implementation Review*. Conditioning is the process of selecting the values for the parameters of the operating models such that the predictions from these models are consistent with the available data.

3.1 Stock structure and mixing matrices
A major part of the work undertaken related to determining the stock structure hypotheses to be incorporated into the *Review* and then how to parameterise them. The Workshop agreed to a request from the computing team to rename existing hypotheses I, II and III as A, B and C for logistical reasons.

The G3 genetics sub-group reported that they considered that the three existing hypotheses were either consistent with the genetic data, or the samples sizes were too small to allow confidence in reaching any definite conclusion (Gaggiotti *et al.*, 2011).

The Workshop spent a considerable period of time developing mixing matrices for the three hypotheses. Given the temporal/geographical complexities of the hypotheses this is a difficult and iterative task. The work was advanced greatly at the Workshop and was finalised after the Workshop. The final matrices can be seen in Annex D of SC/64/Rep2. As agreed at the 2011 Annual Meeting (IWC, 2012b), further work was undertaken to define 'pure' stocks for the stock structure hypotheses and the updated information can be found in Annex F of SC/64/Rep2.

3.2 Conditioning
Despite improvements to the minimisation algorithm for the conditioning program (e.g. see Punt and Elvarsson, 2011), some convergence problems remained. To understand the possible consequences and implications of convergence difficulties, the Workshop recommended that: (a) population trajectories be produced when the minimisation algorithm is close to, but not finally at, completion as well as for the 'best estimates' for these trajectories; and (b) the operating model be conditioned for a few trials from multiple starting values.

The Workshop reviewed input data for conditioning the trials:

(a) abundance (no changes to the information agreed at the 2011 Annual Meeting - see IWC, 2012b);
(b) direct catches (an agreed 'Best' and 'High' catch series was agreed – see Annex D of SC/64/Rep2); and
(c) indirect catches (the approach agreed at the 2011 Annual Meeting was incorporated and the Workshop agreed to additional sensitivity tests - see Annex D of SC/64/Rep2).

The Workshop examined a number of diagnostic tables and plots, and agreed that adequate conditioning had been achieved for all six baseline models, although further refinements and improvements might need to be made prior to the 2012 Annual Meeting.

The Workshop identified a number of conditioning-related issues that required further work including: (a) estimated bycatch numbers; (b) abundance in sub-area 2C; (c) the proportion of juveniles compared to adults in sub-area 9; and (d) abundance of Y-stock under hypothesis B. An intersessional working group was established to further this work.

3.3 Projection issues

The Workshop also considered a number of issues related to simulation of the *CLA*. Many of these centred on issues related to abundance surveys. With respect to whether or not to use minimum abundance estimates, the Workshop agreed not to use such estimates apart from for sub-areas 5 and 6W, where despite their limitations these were the only abundance estimates available. They will be treated as unbiased estimates of abundance for the sub-areas concerned in simulating *CLA* applications.

The Workshop agreed that whether and how to use estimates with low coverage or design concerns and the treatment of JARPN and any JARPN II surveys that did not have Committee oversight raised policy issues which would require a decision by the full Scientific Committee. To assist such discussion at the 2012 Annual Meeting, the Workshop requested Japan and Korea to prepare documents for that meeting containing specified information on the surveys whose results were accepted for conditioning (see table 6 in Annex D of SC/64/Rep2).

The Workshop also identified a number of surveys for which results were not available in time for them to be considered for conditioning. It agreed that abundance estimates from these should be presented for review at the 2012 Annual Meeting.

The RMP specifications (IWC, 2012d, footnote 21a) allow for appropriate statistical procedures to be used to extrapolate to areas not covered in some surveys in a time series, provided that they were covered in others, and that allowance is made for additional variance. The Workshop identified cases where this might be appropriate and agreed that a document presenting results for these extrapolations should be presented for review at the 2012 Annual Meeting, to inform the process of finalising the specification of *CLA* application simulations.

3.4 Specification of *Implementation Simulation Trials*

With respect to future whaling operations, no changes were proposed to the specifications given at the 2011 Annual Meeting (IWC, 2012b). Given the inter-relationship between future survey plans and management variants, it was agreed that Japan and Korea should take into account Workshop discussions and present their proposed survey plans to the 2012 Annual Meeting.

The Workshop endorsed the statistics and plots related to the performance of RMP variants developed for previous recent *Implementations* (IWC, 2012b). These statistics and plots, along with the results of applying the approach for evaluating conservation and utilisation performance outlined in the Requirements and Guidelines for *Implementations* will need to be provided to the 2013 Annual Meeting.

The Workshop agreed to near-final trials specifications, noting that some modifications might be necessary depending on the work of the Steering Group guiding post-Workshop work (see table 8 in Annex D of SC/64/Rep2). The changes agreed reflect the changes made to the specifications to ensure that: (a) the hypotheses match the intent of the underlying stock structure hypotheses; and (b) conditioning can be achieved satisfactorily. The Workshop extended the set of trials to include new sensitivity tests:

(a) the number of animals bycaught is proportional to the square-root of abundance rather than to abundance - this sensitivity test examines the impact of possible saturation effects;
(b) a substantially larger fraction of ages 1-4 animals from the O- and OE-stocks are found in sub-areas 2R, 3 and 4 year-round so that the proportion of these animals in sub-area 9 is closer to expectations given the length-frequencies of the catches from sub-area 9;
(c) set the proportion of animals of ages 1-4 in sub-area 9 to zero and allow the abundance in sub-areas 7CS and 7CN to exceed the abundance estimates for these sub-areas - projections for this sub-area will need to account for the implied survey bias; and
(d) place an upper limit of 200 on the number of 1+ animals in 2009 in sub-area 2C.

In response to the Committee's agreement at the 2011 Annual Meeting that a version of hypothesis C that did not assume multiple J stocks would be valuable, the Workshop agreed that should time be available (see Item 9), options should be investigated in which there was one J stock and two O stocks, and two J stocks and one O stock.

3.5 Consideration of data/analyses to reduce hypotheses in future

The Workshop noted the inherent complexity of the western North Pacific common minke whale *Implementation* and the lack of data for certain temporal and geographical cells. It agreed that it was important to begin considering ways to try to improve this data-deficient situation prior to the next *Implementation Review* as early as possible. The initial discussions focused on matters related to stock structure and abundance. A number of suggestions were made including: the use of telemetry and the surveying of southern areas to identify breeding grounds to facilitate identification of 'pure' stocks and thus assignment of individuals to putative stocks. In many areas, even the abundance estimates that are available have large CVs and improved surveys are needed; in coastal areas, aerial surveys should be considered.

3.6 Submissions relevant to the plausibility of hypotheses

The Workshop noted that a decision will be taken at the 2012 Annual Meeting as to whether analyses of CPUE data (or sightings per unit effort data, SPUE) can be used qualitatively to inform assignment of plausibility weights to the hypotheses on which the trials are based. The Workshop provided suggestions for the content of a document to be presented to assist in this discussion and recommended that analyses of the CPUE/SPUE data consider a variety of assumptions so that the robustness of any conclusions from these data can be evaluated. The Workshop established an Advisory Group that would be available to advise the Japanese scientists developing the summary document and conducting GLM analyses of the data should they wish, although the group was not used.

Following discussions at the 2011 meeting, the Workshop reiterated the value of a summary table to summarise the evidence about key questions related to hypotheses which can then be used to inform the assignment of plausibility to hypotheses. It requested Waples to work with Gaggioti and Hoelzel, as well as the proponents of the hypotheses, to develop such a table.

3.7 Schedule of work required prior to the 2012 Annual Meeting

The Workshop developed a work plan (and Steering Groups) to try to ensure that the necessary results and documents were available for the 2012 Annual Meeting (Table 9 of SC/64/Rep2).

In concluding his summary, Donovan thanked the participants for the constructive atmosphere in which discussions were held. He particularly thanked Allison and de Moor (assisted by Punt) for their tremendous work before and during the Workshop, and the rapporteurs.

The Working Group noted that considerable progress had been made during the Intersessional Workshop, and expressed its appreciation to the Workshop participants for their hard work. The Working Group especially thanked Donovan for his excellent Chairmanship of this Workshop.

4. REVIEW RESULTS OF CONDITIONING

4.1 Review intersessional work and issues arising

Allison reported that the baseline trials for stock structure hypotheses A, B and C had been conditioned. Allison and de Moor had identified a number of problems with the fits of the operating model to the data during the intersessional period. They suggested a number of potential changes to the trial specifications developed during the December 2011 Intersessional Workshop to address these problems.

(1) Generate abundance estimates for sub-areas 5 and 6W by sampling from the uniform distribution over the minimum and maximum abundance estimates and fit to these data assuming a low CV (0.1). This change was made because the specifications from the December 2011 Intersessional Workshop led to population size trajectories which hit the minimum abundance estimates even though there is no direct information on the abundances in sub-areas 5 and 6W between the maximum and minimum abundance estimates.

(2) Impose a maximum abundance in sub-area 2R of 500 in August-September of 2009 for the trials based on stock structure hypothesis C. This maximum was imposed to avoid undesirably high number of animals in this area.

(3) Decrease the proportion of J-stock (stock structure hypotheses A and B) and JW-stock (stock structure hypothesis C) in sub-area 12SW in June 1984-99 from 25% to 20%. This change allowed the operating model to fit the abundance estimates for sub-area 12SW better, without leading to poorer fits to the abundance estimates for sub-area 11.

(4) Introduce γ_{24} for O-stock juveniles in April in sub-area 12NE for the trials based on stock structure hypothesis A to match the specifications for the trials based on stock structure hypotheses B and C.

(5) Modify the multipliers for γ_{24} for O-stock adults in sub-area 12NE for stock structure hypothesis B to match the specifications for the trials based on stock structure hypotheses A and C.

(6) Change the entry for OE-stock females in sub-area 11 in September from γ_{22} to $2\gamma_{22}$ for consistency with the baseline trials for stock structure hypotheses A and B.

(7) Modify the specifications for sensitivity tests 10 and 11 so that the proportion of J-/JW-stock in sub-area 12SW in June 1984-99 is 10% and 30% respectively. This change was made given the change made to the value for this proportion in the baseline trials (see (3) above).

(8) Conduct sensitivity tests 21 and 22 for all three stock structure hypotheses. Previously these sensitivity tests were not to be conducted for stock structure hypothesis A.

(9) Introduce two new sensitivity tests (29 and 30) in which the abundance in sub-area 6W is set to the minimum (sensitivity test 29) and to the maximum (sensitivity test 30). This change was made given the change to the way abundance is to be determined for sub-area 6W (see (1) above).

(10) Introduce a new gamma parameter in sub-area 2C for sensitivity test 28. This change was made to satisfy the specification that the number of 1+ whales in 2009 in sub-area 2C in any month is less than 200.

The Working Group **endorsed** these suggestions, and thanked Allison and de Moor for their considerable intersessional work.

4.2 Review results

Appendix 2 summarises the results for the six baseline trials (three stock structure hypotheses and two MSY rates). The Working Group reviewed the results in Appendix 2 and **agreed** that the conditioning for these trials had been acceptably achieved. There was insufficient time to evaluate the results of the conditioning of all the sensitivity tests. The Working Group **agreed** that the results for trials for which 100 simulations were available suggested that it is possible to determine whether conditioning has been achieved successfully based on the fit of the operating model to the actual data (estimates of abundance, proportions of J-/JE-/JW-/OE-/OW-stocks, and penalties). The Working Group appointed a small group (Allison, Butterworth, de Moor, Punt) to review the results of all remaining trials and to prepare a summary document for the full Scientific Committee so that a final decision on the satisfactory completion of conditioning could be made.

5. UPDATES TO STANDARD DATASETS

5.1 Abundance estimates

Abundance estimates play three roles in the *Implementation* process: (a) for use in conditioning trials; (b) for use when applying the *CLA* during *Implementation Simulation Trials*; and (c) for actual application of the *CLA*. The abundance estimates for use during conditioning were selected during the First Intersessional Workshop in December 2010 (IWC, 2012a). The primary objective of this meeting was to select which abundance to use when applying the *CLA* during *Implementation Simulation Trials*. The abundance estimates for use in actual application of the *CLA* will be finalised during the Second Annual Meeting (the 2013 Annual Meeting of the Scientific Committee).

5.1.1 Review new estimates

SC/64/NPM6 provided the cruise report for a sighting survey designed to obtain information on the distribution and abundance of minke whales and other cetaceans in the Yellow Sea. Many sighting surveys have been conducted by Japanese and Korean researchers to assess J-stock in the Yellow, East, and Okhotsk Seas. However only three surveys (2001, 2004 and 2008) have been conducted in the Yellow Sea, although more data are needed for the assessment of J-stock. Furthermore, survey coverage is still very low in this area and the IO passing mode for $g(0)$ estimation has never been attempted. The 2011 survey was conducted using the research vessel *Tamgu 3* (360G/T) from 2 to 30 May 2011. The research vessel searched 1,227.7 n.miles during the survey period. Fourteen minke whales in 14 primary sightings and 95 finless porpoises in 57 sightings were observed during the survey.

The Working Group expressed its appreciation to the Government of Korea for its continued commitment to surveys for minke whales in Korean waters, and to An for his role of oversight on behalf of the Committee.

SC/64/NPM7 presented an estimate of abundance for minke whales in sub-area 5 based on the 2011 survey. Distance data collected during the survey were used to fit three detection models (half-normal, hazard-rate, uniform). The uniform model with cosine adjustment was chosen based on Akaike's Information Criterion (AIC). The abundance of minke whales was estimated at 587 (95% CI=261-1,371) based on the assumption $g(0)=1$.

The Working Group noted that SC/64/NPM7 did not include information on, for example, distances and angles to sightings, which had been requested in the past. There also seemed to be a problem with estimation of expected group size. In addition, the estimates of abundance were sensitive to the choice of the detection function. It was suggested that the data from past surveys could be analysed together to obtain more reliable estimates of the detection function. Coverage of blocks A2 and A4 was poor during the survey. The Working Group suggested conducting a multi-year analysis of the sighting data for the Yellow Sea.

The Working Group did not accept this estimate for use in *Implementation* of the RMP but looked forward to the presentation of a revised estimate in the future.

5.1.2 Acceptability for running trials

SC/64/NPM2 provided an updated summary of the information on survey procedures for the Japanese dedicated sighting surveys conducted by the Institute of Cetacean Research (ICR) and the National Research Institute of Far Seas Fisheries (NRIFSF), in response to a recommendation from the December 2011 Intersessional Workshop. These data were used to estimate abundance (see, e.g. Hakamada and Kitakado, 2011; Miyashita and Okamura, 2011). The selection of the starting point and survey design for the ICR surveys followed past practice. The starting point of the survey was chosen randomly on a line of latitude or longitude so that sighting surveys could be conducted efficiently. Regular zig-zag lines were established systematically within each sub-area so that coverage was uniform. Two survey modes (closing mode and passing mode with abeam closing) were used during the ICR surveys. The total research distance was determined according to the number of research days. Surveying did not take place when the weather was bad. The vessel drifted when the weather was expected to improve quickly, but proceeded along the trackline without surveying when the weather was not expected to improve quickly. Coverage was poor in some blocks in some years because of this. The research distance for the NRIFSF surveys was determined from the research period, the research area, past results (expected number of sightings) and expected weather conditions. A zig-zag form of trackline was used to cover each block uniformly. The starting point of the trackline was tentatively at the corner of the each block. However, the actual starting point was randomly selected on the border of the block and the tracklines shifted in a parallel fashion. Three modes (normal closing mode, independent observer passing mode and independent observer with abeam closing mode) were used since 2002 during the NRIFSF surveys. Before starting, the time in each block was selected given the pre-determined research distance. The vessel drifted when sightings had to stop due to weather and the weather was expected to improve quickly. The surveys attempted to cover the block as uniformly as possible during the allocated period, but some sub-blocks have low coverage owing to weather. The number of sightings in some areas in both types of surveys was low because of reasons such as the difference between the timing of survey and the migration of whales. SC/64/NPM2 concluded that sighting procedures for the ICR surveys follow the RMP Requirements and Guidelines for Surveys (except that the surveys were not subject to Committee oversight) and that the survey procedures for the NRIFSF surveys met all these requirements and guidelines.

SC/64/NPM3 analysed JARPN II sightings data for common minke whales in sub-areas 7CS, 7CN, 7WR, 7E, 8 and 9 collected during 2008 and 2009. Sighting data were collected by dedicated sighting vessels using standard methods. The survey area in 2009 was the same as in 2008, except that the area north of 45°N in sub-area 9 was not surveyed. The data for closing and passing mode were combined because it is reasonable to assume that the difference in survey mode would not bias abundance estimates because there are few very high density areas and the school size for common minke whales is nearly always 1. Analytical procedures were similar to those in Hakamada and Kitakado (2011). It was assumed that $g(0)=1$. Selection of the detection function considered covariates such as sea state (Beaufort scale), sub-area and year. AIC was used to select the best model to estimate effective search half width. The selected detection model was the hazard-rate model with Beaufort scale and year as covariates, and the fit seemed adequate. The abundance estimate for 2008 for sub-area 9 was 1,840 (CV=0.576). The abundance estimates for 2009 for sub-areas 8 and 9 were 507 (CV=0.830) and 1,693 (CV=0.701), respectively. SC/64/NPM3 stated that these estimates could be used in the *CLA*. The sightings data for sub-area 7CS, 7CN, 7WR and 7E during 2008 and 2009 and for sub-area 8 during 2008 may not be suitable to provide abundance estimates for use in the *CLA* because of survey coverage and timing.

The Working Group noted that the estimates of abundance for 2008 and 2009 use information from other years. It therefore **recommended** that variance-covariance matrices be computed for the entire time-series of abundance estimates for sub-areas 7CS, 7CN, 8, and 9. It also noted that coverage in sub-areas 7CS and 7CN was low (and hence that there were few sightings in these sub-areas). Moreover, unlike some previous surveys, there was no coastal stratum and the trackline went north-south rather than east-west. In response, Hakamada noted that the surveys were designed and conducted before sub-areas 7CS and 7CN were defined by the Scientific Committee for use in the *Implementation*.

As noted under Item 3, the December 2011 intersessional Workshop had noted that whether and how to use estimates with low coverage or design concerns and the treatment of JARPN and JARPN II surveys (i.e. surveys that had not originally been intended to produce estimates for use in the RMP) that did not have Committee oversight raised policy issues which would require discussion by the full Scientific Committee (SC/64/Rep2). It also suggested that documents outlining certain aspects of the Japanese and Korean surveys be presented to assist in that discussion.

The Working Group concurred with this view and agreed that the general issues be passed onto the sub-committee on the RMP for advice before completing its discussions on appropriate abundance estimates for use when applying the *CLA* during trials. To assist the RMP discussions it had a preliminary discussion of the information presented on the Japanese surveys provided in SC/64/NPM2.

During this initial discussion, several issues related to the surveys were noted by some participants as summarised below.

- Issues related to design and equal coverage probability: it was commented that many surveys appear to start in the corner of each stratum implying that the start point was not always random with consequent implications for coverage probability.
- Survey areas apparently without a systematic design, or areas with an incomplete pattern with no explanation given for the design.
- Dealing with low coverage: there is low coverage of some of the survey blocks.
- Issues related to migration direction: some of the surveys were conducted in the direction of expected migration;
- Timing and design of surveys: the zig-zag pattern/timing of the surveys in some blocks increases the chance of sighting the same animals more than once and leads to an unequal probability of area coverage;
- Survey strategy when conditions preclude surveying: survey vessel drifts at the position where it stopped due to weather if weather is expected to recover, but proceeds along the track line when this is not the case. The decision whether to drift or move seems subjective and affects coverage probability and may lead to bias.

In response, the following points were noted by some participants.

- Not all surveys started at corner points.
- Other major surveys (e.g. SOWER) had routinely started in the corner of blocks – the important issue is whether this in practice would result in appreciably biased estimates.
- Low achieved coverage is a result of conditions outside the control of the survey and incomplete coverage of some lines is almost inevitable – decisions as to whether estimates based on incomplete coverage can be accepted must be undertaken on a case-by-case basis.
- Surveying with the direction of migration can be problematic but in this case satellite information suggested the rate of migration was low and conducting surveys north to south early during the season would lead to under-estimation of abundance.
- Practical decisions have to be made concerning whether to drift or to continue along tracklines in the case of unacceptable weather – the decisions are based on weather forecasts and survey time left, not sighting rates.

After this initial exchange of ideas, the Working Group requested the sub-committee on the RMP to provide advice on the following issues.

(a) Under what, if any, circumstances can abundance estimates from surveys which were not designed for use in the RMP be considered for use in trials and the *CLA*? Issues to consider include: lack of Committee oversight, data availability, etc.?
(b) Under what, if any, circumstances can abundance estimates be accepted for use in trials and the RMP from surveys that do not strictly follow all of the guidelines, e.g. with respect to random start points, migration direction, etc.?
(c) Generic guidance is needed on matters such as acceptable levels of completed coverage, decisions on whether to drift or continue in unacceptable weather conditions.

The report of these discussions is given in Annex D, Item 2.8.

A small group (Butterworth (Chair), An, Cipriano, de la Mare, Funahashi, Miyashita, Hakamada, Kasuya, Kelly, Matsuoka, Skaug, Sohn, Slooten, Wade) reviewed all of the available abundance estimates and categorised each estimate into ‘Yes’, ‘No’, ‘No agreement’, and ‘Yes*’ (see Appendix 3). The category Yes* indicates that further analysis needs to be considered for an estimate to become acceptable for application of the RMP. Surveys which had been accepted for use in the trials during the 2003 *Implementation* were automatically deemed acceptable. The Working Group thanked the small group for their work and **endorsed** their recommendations.

The Working Group noted that further evaluations of this type would be facilitated by the side-by-side provision of planned tracklines and achieved coverage, without showing transit lines, together with sub-area boundaries.

Regarding those estimates for which no agreement had been reached on whether or not they were acceptable for use in trials, the Working Group **agreed** that the baseline trials should be conducted for the least and most aggressive RMP variants both using and not using the ‘No agreement’ estimates when applying the *CLA*. If the results of the trials are sensitive to the inclusion of the ‘No agreement’ estimates, the proponents would be requested to justify how the ‘No agreement’ estimates could become acceptable with further analysis. The final decision on whether further analysis is likely to allow ‘No agreement’ estimates to be acceptable will be made by the Intersessional Steering Group (see Item 12).

SC/64/NPM5 extrapolated abundance estimates to parts of sub-areas 8, 11, and 12NE which were not covered during some past surveys, to eliminate the bias in estimated abundance trend which arises due to variable coverage. SC/64/NPM5 analysed the block-wise abundance estimates from the surveys in each of the sub-areas using linear models, with some sub-area-specific assumptions. The resultant extrapolated estimates of abundance for sub-areas 8 and 11 did not differ substantially from the estimates based on the original coverage, while those in sub-area 12NE differed substantially from the original estimates. The authors of SC/64/NPM4 suggested that the extrapolated estimates should be considered for use when applying the *CLA*.

The Working Group thanked the authors of SC/64/NPM5 for responding to the recommendation from the December 2011 Intersessional Workshop, and noted that annotation 21A to the RMP specifications (IWC, 2012a) states that ‘A part of an *Area* which is unsurveyed in a single year may count as surveyed when the data from several years are combined, provided that an appropriate multi-year regression analysis is used, and additional variance is taken into account’. The Working Group noted that blocks B11-2 and B12NE-2 had only been surveyed once which meant that there are insufficient data to inform additional variance. The Working Group **agreed** that the information for sub-area 8 satisfied the requirements for applying annotation 21A.

5.1.3 Acceptability for application of the CLA

The Working Group had insufficient time to consider this item at this meeting and referred it for consideration at the forthcoming Second Intersessional Workshop (see Item 12).

5.2 Other

5.2.1 Best catch series for running trials

The Working Group **agreed** with the recommendation of the December 2011 Intersessional Workshop that the 'Best' catch series (outlined in Annex D of SC/64/Rep2) was appropriate for the direct catches.

The Working Group noted that a single series of bycatches would be used for all of the trials when applying the RMP, irrespective of the true values for the bycatches, which differ among trials, and simulations within trials. The Working Group **agreed** that the bycatches would be set to the averages of the predicted bycatches based on the fit to the actual data of the operating model for the six baseline trials (Appendix 4).

It is necessary to specify how future incidental catches will be generated. The Working Group had insufficient time to agree this but appointed a small group (Allison, Butterworth, Miyashita, Punt, Wade) to bring a proposal to the full Scientific Committee.[1]

6. FINAL CONSIDERATION OF PLAUSIBILITY (INCLUDING WEIGHTING OF TRIALS IN TERMS OF OVERALL BALANCE)

6.1 Procedure for assigning plausibility

The Committee's Requirements and Guidelines for *Implementations* (IWC, 2012c) summarise the final steps in an *Implementation*. A key step is assigning plausibility to hypotheses and, by extension, to all of the trials. Trials are assigned 'low', 'medium' or 'high' weights, or are categorised as 'no agreement', which are treated as 'medium' weighted trials. Trials with 'low' weights are not considered further in the *Implementation* while failure of a management variant on any 'high' weight trial leads to that variant being eliminated from further consideration, including with respect to the 'with research' option. 'Acceptable' conservation performance is required on all 'high' weight trials but 'borderline' or 'unacceptable' conservation performance on a number of 'medium' weight trials, leads to further consideration, as detailed in IWC (2012c).

Appendix 5 lists the factors considered in the trials and the final plausibilities assigned to each factor.

6.1.1 Use of catch or sightings per unit effort data (CPUE/SPUE)

The December 2011 Intersessional Workshop noted that the present meeting would decide whether analyses of CPUE data (or sighting per unit effort data, SPUE) could be used qualitatively to inform assignment of plausibility weights to the hypotheses (stock structure and MSYR) on which the trials are based (see Item 3.6). The Workshop had noted that a document outlining relevant operational factors needed to be developed for the Committee to make a decision in this regard, and it had made a number of recommendations regarding such a document.

SC/64/NPM4 summarised information pertaining to catch, sightings and effort data from Japanese small-type whaling during 1977-87 in relation to common minke whales. The basic information available includes catch positions and related information such as distance from port to catch position, operation time, and operation pattern by season. Operation time includes all the activities of catcher boats from the start of searching to completion of the whaling operation, but does not include stopping and floating. It is not possible to split operation time into searching time, chasing time, and towing time because this information is not recorded in the dataset. Yearly changes in distances from port to catch position were not observed and most whaling occurred close (<60 n.miles) to the coast. Three typical operation patterns were identified, and those have changed only slightly since 1982. Factors which affect operation efficiency, such as the use of motorboats for chasing, and additional payments to crew for finding whales had been introduced before 1977. Sighting information was recorded even when minke whales were not targeted or caught. The authors of SC/64/NPM4 concluded that catch or sighting per unit effort data can be useful as an index of population trend if standardised.

The Working Group thanked the authors of SC/64/NPM4, which covered most of the factors identified during the December 2011 Intersessional Workshop. The Working Group noted that there was considerable variation in where individual vessels operated during the year, and that if vessel movement reflects availability of whales, CPUE or SPUE may be biased as an index of relative abundance. It was suggested that focusing on April-May only may provide more consistency.

Miyashita provided the Working Group with results of further GLM-based analyses of the sightings and effort data. The Working Group considered that further analysis and model diagnostics would need to be provided before the resultant SPUE trends could be used to assist the assignment of plausibility to hypotheses related to stock structure and MSYR. Given the time available, this was not feasible this year. It was noted that these data could be re-analysed and presented to the next *Implementation Review*, although some members considered that use of whaling SPUE data was problematic and that no analyses of these data would lead to information which could inform plausibility.

6.2 Stock structure

To assist the Working Group in assigning plausibility to hypotheses on stock structure in western North Pacific common minke whale, documents from the proponents of Hypotheses A/B and of Hypothesis C (see fig.1 in last year's report, IWC, 2012b, p.103) summarising their main points, had been requested.

SC/64/NPM1 summarised two kinds of information: (a) work conducted since 2003 to refine previous hypotheses based on O and J stocks, which evolved into current Hypotheses A and B; and (b) rebuttals to the main arguments for Hypothesis C. Two major points of disagreement separate these hypotheses: (1) whether one or two J-like and one or two O-like stocks exist; and (2) whether a different stock (Y) occurs in the Yellow Sea. The first component separates Hypotheses A and B from Hypothesis C; the second separates Hypothesis A from Hypotheses B and C. Current Hypotheses A and B had been refined based on new genetic analyses that followed a process and recommendations from the Committee over the years (e.g. genetic separation of J/O stocks, hypothesis testing on separated J and O stock samples, examination of statistical power, etc.). No evidence for genetic heterogeneity was found within the J and O stock animals. In contrast, Hypothesis C, which proposes division of J stock into JW and JE and O stock into OW and OE, is based on genetic analyses conducted on total samples (pooled samples of the O and J stocks); no analytical effort had been made to account for mixture in these samples. Following advice from the Committee, several types of non-genetic

[1]See Appendix 10 for the proposal which was agreed by the Scientific Committee.

data were examined by the proponents of Hypotheses A and B, and comparative analyses were conducted separately for J and O stock samples. Again, no significant differences were found within the O or within the J stock animals. Although proponents of Hypothesis C had argued initially that several types of non-genetic data supported their division into JW/JE and OW/OE, at this point the only such evidence cited by those proponents is conception date, which is not a strong argument as explained in SC/64/NPM1. The authors of SC/64/NPM1 believed that responding to recommendations from the Committee for specific analyses was important; one of the most relevant such recommendations was to conduct analyses separately for J and O stock samples. The authors of SC/64/NPM1 noted that the proponents of Hypotheses A and B, and other members of the Committee had done this but the proponents of Hypothesis C had not. The possibility of additional structure in the Yellow Sea (presence of a Y stock) cannot be ruled out; however, current scientific evidence (conception date and microsatellite data) is weak and much more work is required to evaluate the possibility of additional structure. Therefore, the authors of SC/64/NPM1 believed that the plausibility of Hypothesis A should be high, that of Hypothesis B should be medium or low, and that of Hypothesis C should be low.

SC/64/NPM11 reviewed and summarised evidence for the three stock structure hypotheses for western North Pacific common minke whales that were considered plausible at the first intersessional Workshop: Hypothesis A - a stock in the Sea of Japan and Yellow Sea and one in the Pacific Ocean; Hypothesis B - one stock each in the Yellow Sea, Sea of Japan, and Pacific Ocean; and Hypothesis C - one stock in the Yellow Sea, one in the Sea of Japan, a 'J-like' stock along the Pacific coast of Japan, and two 'O-like' stocks in Pacific nearshore and offshore waters. The authors of SC/64/NPM11 reiterated that finding conclusive evidence to fully resolve the stock structure is unlikely given that no samples have been collected on the putative breeding grounds in winter when presumably 'pure' stocks would exist. Instead, they pointed out the primary information on population structure comes from biological information on conception dates and genetic data collected from year-round coastal bycatch and commercial and Special Permit hunting during migration. They summarised the available evidence as follows. Whales in the Yellow Sea have only autumn conception dates, whales in the Sea of Japan and along the Pacific coast of Japan have a mix of autumn and winter conception dates, and whales from the rest of the Pacific only have winter conception dates. Hypotheses B and C are both equally consistent with data on conception dates, but Hypothesis A is not, and so the authors of SC/64/NPM11 considered it to have lower plausibility. Results from both mtDNA and microsatellite genotypes show significant differences in most pairwise comparisons between spatial areas for western North Pacific minke whales. Of primary importance for distinguishing Hypotheses B and C are the significant differences seen between three regions in the Pacific Ocean – the coast of Japan, nearshore waters >10n. miles from the coast, and offshore waters, as well as the significant differences seen between the east and west coasts of Japan. One explanation proposed for these significant differences is that there are differing proportions of just two stocks ('J-stock' and 'O-stock') in each of these four areas. However, allozyme and microsatellite allele frequencies only show strong evidence for mixing of stocks in other regions (i.e. along the Korean coast of the Sea of Japan, and north of Hokkaido). The four areas in question do not show strong evidence for mixing of stocks. Therefore, the authors of SC/64/NPM11 considered that Hypothesis C, which has differentiated stocks in each of these four locations, is in agreement with the genetic data, and therefore is considered to have high plausibility.

In addition to these summaries, two papers with new genetic analyses were presented. SC/64/NPM9 used computer simulations to examine the effect of different sample sizes on the distributions of the correlations between θ and F_{IS}. Waples (2011) had proposed that, in a sample that contains individuals only from two distinct stocks, the largest departures from equilibrium (quantified as F_{IS}) should be seen at the loci that show the largest allele frequency differences between the two stocks (quantified as θ). Waples (2011) had used this logic to test the two-stock hypothesis (i.e. only O and J stocks in Hypotheses A and B) proposed for the North Pacific minke whales around Japan. Because the observed relationships of these genetic indices in samples of minke whales from SA7-bycatch, SA7-Kushiro, SA7-Sanriku, SA2, and SA11 were weaker than the expected relationship estimated from artificial mixtures of only putative J and O individuals, Waples (2011) suggested that these samples might have contained individuals from more than two stocks. However, it was also indicated that the robustness of the analyses should be evaluated because this approach was novel and the behaviour of these indices was unclear under various situations such as different sample sizes.

SC/64/NPM9 followed this suggestion and evaluated sensitivity of the relationship between θ and F_{IS} to different samples sizes. Distributions of the correlation coefficients between θ and F_{IS} in 1:1 mixtures of individuals from the two distinct stocks that reflected the situation of the bycatch sample from SA7 (J=93, O=90, unclassified=29) were examined under three different sample sizes from each stock: 5,000 (same as simulated N_e, total sample size for both stocks=10,000), n=400 (2n=800, comparable to samples from SA6 and SA9), and n=100 (2n=200, comparable to SA7-bycatch). Genetic differences between the two populations in the simulation we comparable to those found between O and J stocks (Kanda *et al.*, 2009). In the simulation results, the correlation coefficients were distributed as expected from Waples (2011) in the cases of n=5,000 and n=400, but fluctuated quite widely and some of the values were quite low in the case of n=100. The results of SC/64/NPM9 suggest that the relatively low correlations observed in the SA2, SA7, and SA11 samples in Waples (2011) also could be due to the small sample size rather than due to mixture of the individuals from more than two stocks. Because of the considerable variability seen in the simulated data, the authors of SC/64/NPM9 suggested that further evaluation is required before the results of Waples (2011) could be used as evidence against Hypotheses A and B.

In discussion, it was noted that the correlation coefficient between θ and F_{IS} for the sample from SA7 bycatch was 0.3, as reported by Waples (2011), and 21% of the correlation coefficients in the simulated data for SC/64/NPM9 were this low or lower. This indicates that Hardy-Weinberg results for this sample are consistent with patterns expected in an equal mixture of just two stocks similar to O and J. In discussion, it was suggested that it would be useful to extend these analyses to the two-locus (linkage disequilibrium - LD) correlations that were also reported in Waples (2011). The LD correlations are likely to be more robust to sampling error than the single-locus correlations, because the latter are based on just 16 data points (one for each locus), while the LD correlations use data for 120 pairwise comparisons of loci. It would be useful to conduct comparable sensitivity analyses for uneven mixture fractions and, perhaps, variations in mutation rate across loci. Results of those analyses would provide a more solid basis for drawing conclusions about compatibility of the Hardy-Weinberg results with Hypotheses A and B. A question arose as to whether the results could also be sensitive to presumed allele frequencies for the two stocks in the mixture. That is theoretically possible, although it is not considered to be an important factor in this case because large samples of relatively pure O and J stock individuals are available and Waples (2011) had found that results differed little whether they used allele frequencies from all individuals collected in SA6 and SA9 or only those estimated to have a high probability of being J or O (respectively) based on STRUCTURE analyses.

SC/64/NPM10 responded to a request from last year's meeting for follow-up analyses comparing the performance of two Bayesian clustering programs (STRUCTURE and HWLER) for detecting the number of gene pools represented in a sample. Specifically, it was of interest to determine whether HWLER could provide greater resolution for mixtures of genetically similar stocks. SC/64/NPM10 simulated genetic data for 16 microsatellite-like loci (the number used in recent evaluations of North Pacific minke whales) and created artificial 1:1 mixtures of two populations having levels of genetic differentiation ranging from none (panmixia) to moderate (F_{ST}=0.06). Both programs only detected one population when true panmixia was modelled, but both also failed to detect a second population at the weakest level of differentiation (F_{ST}=0.007). STRUCTURE reliably detected two populations at F_{ST}=0.02 but HWLER did not, but HWLER was more consistent in resolving mixtures for F_{ST} >0.03. The paper discussed some diagnostic tests that could help determine when these programs are not producing reliable results.

The Working Group thanked the author of SC/64/NPM10 for responding to its request and noted that the results provide additional confirmation that these Bayesian clustering methods cannot detect the weakest levels of population structure, at least using currently available numbers of genetic markers. The Working Group noted that this limitation emphasises the potential value of efforts to characterise the lowest levels of gene flow that would be sufficient to reduce genetic differences to the point at which they cannot be detected without samples from breeding grounds. It then should be possible to evaluate whether demographic exchanges at this level would be substantial enough to lead to different management outcomes (see also Item 9). Several more technical aspects of the performance of STRUCTURE at moderate levels of population differentiation (F_{ST}=0.045-0.06) were also discussed; details are given in the report of the Working Group on Stock Definition (see Annex I).

As requested at the December 2011 Intersessional Workshop, the G3 group (Waples, Hoelzel, Gaggiotti) reformatted the information from Appendix 9 of last year's Working Group report (IWC, 2012b) and condensed it to a single-page table. The rows of this table represent various types of information (nuclear and mitochondrial DNA; morphology and life history) and the columns are key stock-structure hypotheses/questions. After some discussion, the G3 group updated the table based on comments from members of the Working Group. The updated table is given as Appendix 6.

A number of papers in recent years have considered the relevance to stock-structure hypotheses of departures from Hardy-Weinberg equilibrium at one and two gene loci; for example, SC/64/NPM9 presented new sensitivity analyses related to this issue. An attempt to synthesise this information and assess the net consequences for stock-structure hypotheses was presented to the Working Group (Appendix 7). Two main questions are involved:

(1) is the sample from a given area (e.g., SA2 or SA7CS), a mixture, or from just a single stock?
(2) Can all of these putative mixed samples be explained by mixtures in different fractions of the same two stocks (O and J, as proposed by Hypothesis B)?

Regarding the first question, Appendix 7 concluded that available data suggest that each of the samples from SA7W-K, SA7W-S, SA2, and SA7-bycatch contains individuals from more than one population. The results are more equivocal concerning whether or not all of the empirical data can be explained by a mixture of just O and J individuals. Appendix 7 concluded that, overall, evidence from Hardy-Weinberg departures for more than two O+J stocks is only weak to moderate.

In discussion, it was noted that explicitly modelling the effects of a third stock on Hardy-Weinberg results could be informative, but that such a project would be rather open-ended, given that the genetic characteristics of such a hypothetical third stock are not well defined. The question was raised as to whether the issue of Hardy-Weinberg departures is the lynchpin upon which decisions about the plausibility of the OE vs OW split depend. In response, it was noted that Hardy-Weinberg tests related to mixing are important for evaluating comparisons of allele frequencies that show significant differences between samples of putative OW and samples from further east, but that conclusions of the G5 group (Waples, Hoelzel, Gaggiotti, Palsbøll, Teidemann) regarding Hypothesis C were also influenced by the PCA analyses (Gaggiotti and Gascuel, 2011) that showed residual genetic heterogeneity in O-like individuals after removal of putative J-stock individuals. It was pointed out that Slikas and Baker (2011) simulated an artificial mixture of O and J individuals and found much higher levels of Hardy-Weinberg departures than found in actual samples from SA7CS and SA7CN, leading to the conclusion that O+J mixing was not a plausible explanation for genetic characteristics of these samples. However, it was also noted that the artificial mixture in Slikas and Baker (2011) used a very large sample size (almost 900 individuals) and represented a nearly 1:1 mixture of O and J stock, while the samples in question were much smaller and (based on estimates in Waples, 2011) involved mixtures with only about 20-25% J stock. Both of these features of the simulated mixture in Slikas and Baker (2011) would tend to increase power to detect Hardy-Weinberg departures compared to that expected for the actual samples.

Following presentation and discussion of all the primary papers, the G5 group summarised their interpretation of the relative support for and against the five hypothesised stocks (JE, JW, OE, OW, Y), based on the cumulative genetic information presented and discussed during the last several years (Appendix 8). In addition to reiterating strong support for the core J-O stock differences, the major conclusions were as follows. Y stock: evidence for=moderate; evidence against=low. Two J stocks: evidence for=low; evidence against=moderate to high. Two O stocks: evidence for= moderate; evidence against=low to moderate. The G5 group concluded that the genetic data suggest that the existence of a Y stock is more likely than not, the existence of two J stocks is unlikely, and the existence of two O stocks is somewhat more likely than not.

The Working Group thanked the G5 group for this valuable summary, which generated some lively discussion. In response to comments from proponents of the different stock structure hypotheses the G5 group explained that they had considered all the available information and arguments, and that their summary reflected what they believed were the conclusions best supported by the genetic data alone.

Pastene commented that although several types of data had been considered during the *Implementation* process thus far, he felt that the conclusions on plausibility were too heavily weighted to the genetic data, noting that the Committee has recognised previously the value of using data from a suite of techniques.

In discussion, there was some attempt to reduce the number of stock structure hypotheses for consideration in the *Implementation Simulation Trials*. It was noted that the conclusion in Appendix 8 regarding Y stock did not depend on data on conception date, which some consider the strongest evidence for Y stock. Some members suggested that as a consequence, Hypothesis A be assigned 'Low' plausibility. This was not agreed to by the proponents of that hypothesis, who pointed out that reliability of the conception date data has been questioned (IWC, 2012b, pp.122-23) and who argued that the genetic data are too limited to be considered strong support for existence of Y stock. Similarly, assigning 'High' plausibility to a 4-stock version of Hypothesis C that includes two O stocks but only one J stock, and 'Medium' plausibility to Hypothesis C did not receive agreement. In the end, however, it was not possible to reach agreement on any of these alternatives and, as a consequence, all three main stock structure hypotheses (A, B and C) were 'no agreement' and were therefore treated as if they had been assigned 'Medium' plausibility.

6.3 MSYR and other factors

6.3.1 MSYR

The previous *Implementation* assigned 'high' plausibility to $MSYR_{mat}$=4% and 'medium' plausibility to $MSYR_{mat}$=1% (IWC, 2005). It was noted that these whales are found in a region in which there are very large fisheries which might impact the prey base. However, the size of any such an effect on MSYR cannot be quantified at this time. In addition, the review of MSY rates will not be completed during the current meeting so there is effectively no new information related to MSYR for western North Pacific minke whales. The Working Group therefore assigned 'high' plausibility to $MSYR_{mat}$=4% and 'medium' plausibility to $MSYR_{mat}$=1%, as in the previous *Implementation*.

6.3.2 g(0)

The baseline trials are based on the hypothesis *g*(0)=0.8. This value is based on the estimate of *g*(0) by Okamura *et al.* (2010) for the combination of top barrel and upper bridge. The December 2010 First Intersessional Workshop (IWC, 2012a) had noted that this estimate is conservative because the *g*(0) value is to be applied identically to all surveys, including those by Korean vessels which have lower top barrels, and hence seem likely to miss a greater proportion of minke whales on the trackline. The Working Group assigned the hypothesis *g*(0)=0.8 'high' plausibility. The assumption *g*(0)=1 was assigned 'medium' plausibility.

6.3.3 Others factors

The Working Group assigned 'medium' plausibility to all of the trials except for the following three:

- Trial 24, which is based on stock structure hypothesis C, but with a single O-stock and two J-stocks. This trial was assigned 'low' plausibility given the results of the genetics analyses (see Appendix 8).
- Trials 21 and 29, which are based on the abundance in sub-areas 5 and 6W, respectively, being set to the 'minimum' values. These trials were assigned 'low' plausibility because the Korean surveys in sub-areas 5 and 6W only cover a small fraction of the overall area of these sub-areas.

The Working Group noted that results of trials 21 and 29 might provide useful information regarding the behaviour of the trials, and **recommended** that these trials be conducted if time is available.

7. SPECIFICATIONS OF OPERATIONAL FEATURES AND MANAGEMENT VARIANTS

Miyashita reported that Japan intended to conduct coastal whaling in sub-areas 7CS, 7CN and 11, and pelagic whaling in sub-areas 8 and 9. Coastal whaling will be restricted to 10 n.miles from the coast and during August-October in sub-area 11 to minimise catches of J-stock animals. Whaling in sub-areas 8 and 9 will take place during April-October.

An reported that Korea intended to conduct whaling using small-type catcher boats in sub-areas 5 and 6W from March to November. Operations will be conducted up to 60 n.miles from the coast in sub-area 5 and up to 30 n.miles from the coast in sub-area 6W.

The agreed RMP variants and the associated *Small* and *Medium Area* definitions are given in Appendix 9, which will be published on the IWC website.

Allison reported that the trials will take longer to run than in previous *Implementations* because the *CLA* will be implemented using the Norwegian 'CatchLimit' program rather than the Cooke version of the *CLA*. The Working Group **agreed** that priority should be given to running all RMP variants for the baseline trials as quickly as possible so that any of the RMP variants that are clearly likely to perform 'unacceptably' can be excluded from further consideration. The process of distributing and evaluating trials will be co-ordinated by the Intersessional Steering Group (see Item 12).

8. SPECIFICATION AND CLASSIFICATION OF FINAL TRIALS

8.1 Final specifications

The final trial specifications are given in Appendix 9, which will be published on the IWC website.

8.2 Future surveys

SC/64/NPM8 reported that a survey in the Yellow Sea will be conducted in IO passing mode using the research vessel *Tamgu 3* during spring 2013. Although the Korean Cetacean Research Institute has conducted four sightings surveys in this area in the past, their coverage has been low, particularly in the offshore block, and no attempt has been made to estimate *g*(0) for Korean surveys. This survey will be the first of a series of four which will survey a larger fraction of sub-areas 5 and 6W. The primary objective of the survey is to obtain information on the distribution and abundance of common minke whales, while a secondary objective is to collect information on the distribution of other cetaceans. Tracklines covering 933.7 n.miles will be searched using IO mode with binoculars. Other research activities including biopsy sampling and photo-identification may be conducted during the survey.

The Working Group was pleased to hear that additional surveys would continue to be conducted in the waters off Korea and appointed An to provide oversight on its behalf. In relation to survey design, the Working Group **recommended** that the survey be conducted north to south (block Y2 then block Y1). The Working Group supported surveys that covered a larger fraction of sub-areas 5 and 6W but noted that conducting the survey in two latitudinal bands will make comparisons with past surveys more difficult.

Following this discussion, An informed the Working Group that the design of the 2013 survey would be modified in response to the comments of the Working Group and will be conducted in blocks YS1 and YS2 in sub-area 5 (see Fig. 1).

The Working Group **agreed** that for trials purposes it will be assumed that the proportional coverage of sub-areas will be unchanged.

The planned future surveys and a proposal for how past surveys can be combined to calculate survey estimates for *Small Areas* are given in Appendix 9, which will be published on the IWC website.

9. CONSIDERATION OF DATA/ANALYSES TO REDUCE HYPOTHESES IN FUTURE

SC/64/O9 described the results of a sightings and satellite tagging survey for common minke whales in sub-area 7, north of 35°N and west of 150°E. The survey was carried out from 25 April to 8 June 2011, using the research vessel *Shunyo-maru* in closing mode along predetermined track lines. Two schools (2 animals) of common minke whales, eight schools (11 animals) of Bryde's whales, one fin whale, two schools (3 animals) of humpback whales, and seven schools (13 animals) of sperm whales were sighted on 1,160.8 n.miles of trackline. One of the two minke whales was targeted for tagging with a satellite tag. However, the tag could not be attached successfully. The few encounters with minke whales could be due to low effort in coastal waters and the possible effects of the 2011 earthquake and tsunami.

SC/64/O10 described results of a sighting and biopsy sampling survey for common minke whales in the Okhotsk Sea, including the Russian EEZ, during spring 2011. The survey was carried out from 14 May to 26 June 2011, using the research vessel *Kaiyo-maru No.8*. The primary objective of the survey was to estimate the mixing rate of the J- and O-stocks in the Okhotsk Sea during spring. A total of 1,295.3 n.miles was searched in restricted closing mode. Five schools (10 animals) of common minke whales, 30 schools (37 animals) of fin whales, and one North Pacific right whale were sighted. Three schools of common minke whales were targeted for biopsy skin sampling using the Larsen gun, but no samples were obtained because of difficulties closing on the animals. The few sightings of common minke whales indicated that the migration into the Okhotsk Sea had only just started. Bad weather and the inability of the vessel, *Kaiyo-maru No.8*, to chase the whales also impacted the success of the biopsy sampling aspect of the cruise.

The Working Group expressed its support for continued efforts to collect telemetry and biopsy data to help elucidate stock structure for minke whales in this region. It concurred that a vessel capable of successfully chasing minke whales should be chartered for future cruises of this type.

It is clear from material discussed in Item 6.2 that, in spite of many years of concerted efforts and a great deal of genetic and non-genetic data, considerable uncertainties remain regarding stock structure of western North Pacific minke whales. This issue is particularly difficult because

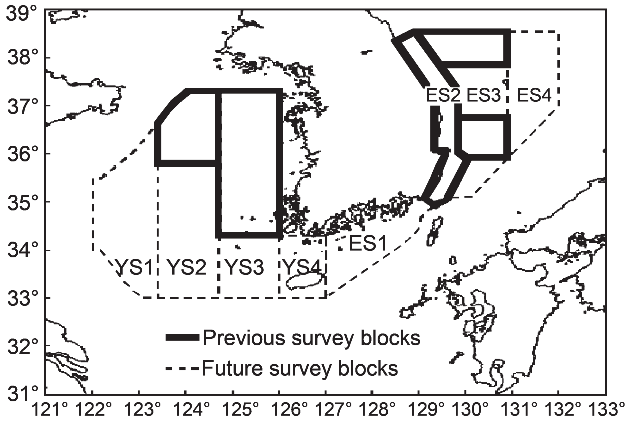

Fig. 1. Survey blocks for previous and future surveys in Korean waters.

of the lack of any samples from breeding grounds. The Working Group considered a number of types of genetic analyses that might help to reduce these uncertainties in the future, including the following.

Sensitivity analyses of recently-used methods

(a) The effects of purging on PCA results. Assessment using known mixtures of real and simulated data to assess whether significant heterogeneity within O-type individuals (Gaggiotti and Gascuel, 2011) could be caused by residual J stock individuals or by family structure.

(b) The effect of sample size and mixing proportions on statistical power in one- and two-locus tests of Hardy Weinberg deviations. These analyses would extend the approach used in SC/64/NPM9 to include a broader range of sample sizes and mixture fractions, and would extend the evaluations to two-locus tests of linkage disequilibrium. In addition, similar analyses could be conducted to develop an empirical distribution of the expected correlation between F_{IS} and likelihood disequilibrium (LD) values for the same loci in different samples.

Development and application of new analyses

(a) Assessment of the temporal-spatial distribution of genetically identified pairs of closely related individuals (e.g. SC/64/RMP1). If multiple J or O stocks exist, this should be evident in the detection probability and spatio-temporal distribution of pairs of individuals that are closely related. This kinship-based approach has the advantage that it does not depend on equilibrium assumptions and instead reflects contemporary movements of individuals. Power depends on sample size as well as population size.

(b) Development of an approximate-Bayesian-computation (ABC) estimation of the relative probabilities of the existing stock-structure hypotheses and intermediate scenarios. This research was proposed and endorsed by the Working Group last year (IWC, 2012b, p.113) but was not supported by the Committee because of the limited funding available. Models representing each hypothesis would be based on the assumption that the O-type population is ancestral, and that environmental change has driven the generation of population divisions over time.

The importance of considering further work on non-genetic data was also noted but no specific analyses were proposed.

In addition to these proposed analyses specifically related to North Pacific minke whales, the Working Group considered an approach that would more broadly address core stock-structure problems that recur for many species in many areas. This general approach has two parts: (1) determining what levels of demographic mixing between breeding populations do and do not make a difference in terms of conservation goals or management outcomes; and (2) using genetic and other methods to determine whether actual levels of connectivity are above or below this threshold.

The Working Group welcomed all these suggestions but noted that because relative plausibility has now been assigned to the different stock-structure hypotheses, analyses specific to North Pacific minke whales would not be high priority for the Committee. However, the more general, cross-cutting approach was of considerable interest, and the Working Group **agreed** that this work should receive high priority. The following suggestions were made to facilitate implementation of this idea:

(1) review the data in IWC files to determine the range of migration rates that have made a difference in management outcomes under RMP trials;
(2) conduct additional evaluations under RMP-lite to assess the relationships between population size, stock productivity, migration rate, and sustainable harvest rate; and
(3) use TOSSM or related approaches to make an explicit link between genetic indices detectable with genetic markers and levels of connectivity between populations.

It was noted that the *Implementation Review* for North Atlantic common minke whales will undertake some evaluations related to Item 2 above (Annex D, Item 3.3) and therefore it would be desirable to coordinate efforts in that regard. It was also noted that similar work was being undertaken by scientists at the US Southwest Fisheries Science Center. Cumulative results of these analyses should make it apparent whether general rules of thumb about 'tipping point' levels of migration can be identified, or whether the outcomes are so diverse that every situation must be evaluated as a special case.

As noted during the December 2011 Intersessional Workshop (SC/64/Rep2), in addition to issues of stock structure, other difficulties in conducting the present *Implementation Review* centred on abundance estimates, including their unavailability in some areas and the large CVs for some of the estimates that were available. The difficulties faced by the small group considering abundance estimates (see Item 5.1.2) amplify this concern.

The Working Group **agreed** that, to avoid such difficulties in future *Implementation Reviews*, the Committee should consider taking a more active and collaborative approach to this issue. Examination of trial results will assist in identifying the key temporal and geographical areas where new/improved abundance estimates would be most valuable. The Committee should consider developing, in conjunction with the appropriate range states, a short-medium term survey strategy (including design and required effort) and analytical approach that would improve the availability of satisfactory abundance estimates with reasonable CVs at the appropriate geographical and temporal scale to facilitate future *Implementation Reviews*. This could follow a similar process to that used to develop the IWC-POWER programme (Annex G, Item 6.2).

10. INPUTS FOR ACTUAL APPLICATIONS OF THE *CLA*

The Working Group **agreed** that the best estimates of the direct catches and the average predicted bycatch from the six baseline trials would be used for applications of the *CLA*. The Working Group did not have sufficient time to select abundance estimates for use in application of the *CLA*. This issue will need to be addressed at the Second Intersessional Workshop (see Item 12).

11. OTHER MATTERS

There were no other matters raised for discussion.

12. WORK PLAN

The Working Group **agreed** that, with the exception of those issues highlighted above in its report, it had completed the work required at the First Annual Meeting. The remaining work required to complete the *Implementation Review* for western North Pacific minke whales is structured around the Second Intersessional Workshop and the Second Annual Meeting (next year), as described in the Committee's Requirements and Guidelines for *Implementations* (IWC, 2012c).

A detailed work plan, including a full list of tasks with associated timelines, will be presented to the Scientific Committee.[2]

The Working Group reappointed the Intersessional Steering Group (Butterworth (Convenor), Allison, An, Baker, de Moor, Donovan, Double, Hammond, Kanda, Kelly, Kitakado, Park, Pastene, Punt, Wade and Waples) to guide the work prior to next year's meeting.

13. ADOPTION OF REPORT

The report was adopted at 10:45 on 20 June 2012.

REFERENCES

Gaggiotti, O., Hoelzel, R. and Waples, R. 2011. Report of the G3 Review Group. Paper SC/D11/NPM2 presented to the First Intersessional Workshop for the *Implementation Review* of western North Pacific common minke whales, 12-16 December 2011, Tokyo, Japan (unpublished). 1pp. [Paper available from the Office of this Journal].

Gaggiotti, O.E. and Gascuel, F. 2011. Stock structure of North Pacific minke whales as revealed by Principal Component Analyses of microsatellite data. Paper SC/63/RMP23 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 7pp. [Paper available from the Office of this Journal].

Hakamada, T. and Kitakado, T. 2011. Abundance estimation for the western North Pacific common minke whales based on sighting information from JARPN and JARPN II. Paper SC/63/RMP18 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 25pp. [Paper available from the Office of this Journal].

International Whaling Commission. 2005. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 7:77-113.

International Whaling Commission. 2012a. Report of the first RMP intersessional workshop for western North Pacific common minke whales. *J. Cetacean Res. Manage. (Suppl.)* 13:411-60.

International Whaling Commission. 2012b. Report of the Scientific Committee. Annex D1. Report of the Working Group on the *Implementation Review* for western North Pacific common minke whales. *J. Cetacean Res. Manage. (Suppl.)* 13:102-29.

International Whaling Commission. 2012c. Requirements and Guidelines for *Implementations* under the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 13:495-506.

International Whaling Commission. 2012d. The Revised Management Procedure (RMP) for Baleen Whales. *J. Cetacean Res. Manage. (Suppl.)* 13:483-94.

[2]The work plan agreed by the Scientific Committee is given as Appendix 11.

Kanda, N., Goto, M., Kishiro, T., Yoshida, H., Kato, H. and Pastene, L.A. 2009. Microsatellite analysis of minke whales in the western North Pacific. Paper SC/J09/JR30 presented to the Expert Workshop to Review Results of JARPN II, 26-30 January 2009, Tokyo, Japan (unpublished). 14pp. [Paper available from the Office of this Journal].
Miyashita, T. and Okamura, H. 2011. Abundance estimates of common minke whales using the Japanese dedicated sighting survey data for RMP *Implementation* and CLA - Sea of Japan and Sea of Okhotsk. Paper SC/63/RMP11 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 34pp. [Paper available from the Office of this Journal].
Okamura, H., Miyashita, T. and Kitakado, T. 2010. *g*(0) estimates for western North Pacific common minke whales. Paper SC/62/NPM9 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 7pp. [Paper available from the Office of this Journal].
Punt, A.E. and Elvarsson, B.T. 2011. Improving the performance of the algorithm for conditioning *Implementation Simulation Trials*, with application ot North Atlantic fin whales. Paper SC/D11/NPM1 presented to the First Intersessional Workshop for the *Implementation Review* of western North Pacific common minke whales, 12-16 December 2011, Tokyo, Japan (unpublished). 7pp. [Paper available from the Office of this Journal].
Slikas, B. and Baker, C.S. 2011. Investigation of Hardy-Weinberg equilibria and population differentiation in North Pacific minke whales, based on microsatellite genotypes of Japanese 'bycatch' and scientific whaling. Paper SC/63/RMP17 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 6pp. [Paper available from the Office of this Journal].
Waples, R.S. 2011. Can evidence for spatial and/or temporal genetic heterogeniety of North Pacific minke whales be explained by different mixture fractions of the same two core stocks, or is it necessary to postulate an additional stock(s)? Paper SC/63/RMP7 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 7pp. [Paper available from the Office of this Journal].

Appendix 1

AGENDA

1. Introductory items
    - 1.1 Election of chair and appointment of rapporteurs
    - 1.2 Chair's opening remarks
    - 1.3 Adoption of Agenda
    - 1.4 Review of documents
2. Objectives of the First Annual Meeting
3. Report of the December 2011 intersessional Workshop
4. Review results of conditioning
    - 4.1 Review intersessional work and issues arising
    - 4.2 Review results
5. Updates to standard datasets
    - 5.1 Abundance estimates
        - 5.1.1 Review new estimates
        - 5.1.2 Acceptability for running trials
        - 5.1.3 Acceptability for application of the *CLA*
    - 5.2 Other
        - 5.2.1 Best catch estimates for running trials
6. Final consideration of plausibility (including weighting of trials in terms of overall balance)
    - 6.1 Procedure for assigning plausibility
    - 6.2 Stock structure
    - 6.3 MSYR and other factors
7. Specification of operational features and management variants
8. Specification and classification of final trials
    - 8.1 Final specifications
    - 8.2 Future surveys
9. Consideration of data/analyses to reduce hypotheses in future
10. Inputs for actual application of the *CLA*
11. Other matters
12. Work plan
13. Adoption of Report

Appendix 2

BASELINE CONDITIONING RESULTS

These will be made available on the IWC website.

Appendix 3

REPORT OF THE GROUP TO ADVISE ON THE ACCEPTABILITY OF NORTH PACIFIC MINKE ABUNDANCE ESTIMATES FOR USE IN RMP *IMPLEMENTATION SIMULATION TRIALS*

Members: Butterworth (Chair), An, Cipriano, De la Mare, Funahashi, Miyashita, Hakamada, Kasuya, Kelly, Matsuoka, Skaug, Sohn, Slooten, Wade

The Group noted that further considerations of this nature would be facilitated by the side-by-side provision of planned tracklines and achieved coverage, without showing transit lines in addition, together with an indication of sub-area boundaries.

The Group's recommendations on acceptance are reflected in the final two columns of the Table below in the form of yes/no agreement/no, followed by a brief rationale for any disagreement. The notation Yes* indicates that further analysis needs to be considered for an estimate to become acceptable. In cases where surveys had been accepted for use in Trials in the 2003 *IST* process, these surveys were deemed automatically acceptable for the current process.

Table 1

Acceptability of abundance estimates.

Sub-area	Year	Season	Areal coverage (%)	STD estimate[1]	CV[2]	Current conditioning	Used in 2003 trials?	2010 comment related to decision on acceptability for conditioning	Use in current trials?	Rationale
5	2001	Apr.-May	13.0	1,534	0.523	Min	-	Low area coverage	Yes*	Only area completed. Needs further analysis.
	2004	Apr.-May	13.0	799	0.321	Min	-	Low area coverage	Yes*	Only area completed. Needs further analysis.
	2008	Apr.-May	13.0	680	0.372	Min	-	Low area coverage	Yes*	Only area completed. Needs further analysis.
	2011						-		Yes*	Only area completed. Needs further analysis.
6W	2000	Apr.-May	14.3	549	0.419	Min	-	Low area coverage	Yes*	Inshore segment only with adjustment for differential extent of inshore coverage (no extrapolation)
	2002	Apr.-May	14.3	391	0.614	Min	-	Low area coverage	Yes*	As above
	2003	Apr.-May	14.3	485	0.343	Min	-	Low area coverage	Yes*	As above
	2005	Apr.-May	14.3	336	0.317	Min	-	Low area coverage	Yes*	As above
	2006	Apr.-May	14.3	459	0.516	Min	-	Low area coverage	Yes*	As above
	2007	Apr.-May	14.3	574	0.437	Min	-	Low area coverage	Yes*	As above
	2009	Apr.-May	14.3	884	0.286	Min	-	Low area coverage	Yes*	As above
	2010						-		Yes*	As above
6E	2002	May-Jun.	79.1	891	0.608	Yes(#1)	-	Only use northern part due to possible double counting	**No agreement**	Poor coverage and analysis difficulties. Poor availability. Only use northern part. Note: 6E is combined with 10E in most variants.
	2003	May-Jun.	79.1	935	0.357	Yes(#1)	-	As above	Yes	
	2004	May-Jun.	79.1	727	0.372	Yes(#1)	-	As above	Yes	(Incomplete coverage). Only N offshore block used.
10W	2006	May-Jun.	59.9	2,476	0.312	Yes	-		Yes	
10E	2002	May-Jun.	100.0	816	0.658	Yes	-		Yes	Check issue of coverage
	2003	May-Jun.	100.0	405	0.566	Yes	-		Yes	
	2004	May-Jun.	100.0	474	0.537	Yes	-		**No agreement**	Design question: (most sightings in concentration near coast).
	2005	May-Jun.	100.0	666	0.444	Yes	-		Yes	May have analytical difficulties.

Cont.

Sub-area	Year	Season	Areal coverage (%)	STD estimate[1]	CV[2]	Current conditioning	Used in 2003 trials?	2010 comment related to decision on acceptability for conditioning	Use in current trials?	Rationale
7CS	2004	May	100.0	886	0.502	Yes	-		Yes	Northern part only.
	2006	Jun.- Jul.	100.0	3,690	1.199	Yes	-		Yes*	Analysis for non-random start. Note 2004 and 6 are in at different times.
7CN	2003	May	75.4	184	0.805	Yes	-		**No agreement**	Inadequate and heterogeneous coverage
7W	1991	Aug.-Sep.		1,164	0.183	2003 only	Yes		7CN: Yes* 7CS: Yes*	Not listed in IWC (2012), from 1990+, 1, 2 surveys, not split to new areas. OK for model based estimate. With analysis.
7WR	2003	May-Jun.	54.2	524	0.700	Min	-	Low area coverage	Yes*	Upper portion only with analysis for non-random starts.
	2004	May-Jun.	88.8	863	0.648	Yes	-		Yes	
	2007	Jun.-Jul.	88.8	546	0.953	Yes	-		Yes*	With analysis if start is non-random.
7E	1990	Aug-Sep.		791	1.848	2003 only	Yes	CV too high to be meaningful	No	CV too high.
7E	2004	May-Jun.	57.1	440	0.779	Yes	-		Yes	
	2006	May-Jun.	57.1	247	0.892	Yes	-		Yes	
8	1990	Aug.-Sep.	61.8	1,057	0.705	Yes	Yes	In other years, no whales observed in area not covered	Yes	Agreed in 2003.
	2002	Jun.- Jul.	65.0	0	482[3]	Yes	-		Yes	Note different survey timings.
	2004	Jun.	40.5	1,093	0.576	Yes	-	In other years, no whales observed in area not covered	Yes	
	2005	May- Jul.	65.0	132	1.047	Yes	-		Yes*	With analysis: non-random start; no planned coverage in upper left (Russian EEZ), 2 sets of lines in lower blocks.
	2006	May-Jul.	65.0	309	0.677	Yes	-		Yes	
7E+8	2007	Jun.-Jul.	65.0	391[4]	1.013	Yes	-		Yes*	With analysis: non-random start; no planned coverage in upper left (Russian EEZ).
9	1990	Aug.-Sep.	35.0	8,264	0.396	Yes	Yes		Yes	Agreed in 2003.
	2003	Jul.-Sep.	33.2	2,546	0.276	Min	-	Survey not co-incident with density peak in Aug-Sept	Yes	
9N	2005	Aug.-Sep	67.8	420	0.969	Yes	-		Yes	
11	1990	Aug.-Sep.	100.0	2,120	0.449	Yes	Yes		Yes	Agreed in 2003.
	1999	Aug.-Sep.	100.0	1,456	0.565	Yes	Yes	*Check map to make sure	Yes	Agreed in 2003.
	2003	Aug.-Sep.	33.9	882	0.820	Yes	-		**No agreement**	Potentially biased due to weather induced coverage omission to North. Agreed: not acceptable to include coastal transect in analysis.
	2007	Aug.-Sep.	20.2	377	0.389	Min	-	Low area coverage	Yes*	May need alternative calculation. Results must come from transect lines only.
12SW	1990	Aug.-Sep.	100.0	5,244	0.806	Yes	Yes		Yes	Agreed in 2003.
	2003	Aug.-Sep.	100.0	3,401	0.409	Yes	-		**No agreement**	Insufficient coverage.
12NE	1990	Aug.-Sep.	100.0	10,397	0.364	Yes	Yes		Yes	Agreed in 2003.
	1992	Aug.-Sep.	89.4	11,544	0.380	2003 only	Yes		Yes	Agreed in 2003. Year wrong in IWC (2012).
	1999	Aug.-Sep.	63.8	9,366	0.377	Yes	-	*Check map to make sure	Yes*	Omit E block – inadequate coverage. Limit N block to area surveyed.
	2003	Aug.-Sep.	46.0	13,067	0.287	Yes	-		**No agreement**	Agreed: 2 blocks should be omitted on basis of inadequate coverage. No agreement re the 3 other blocks (2 NW and one E): issue is whether there is adequate coverage in the most NW block (off Kamchatka).

[1]Standard (STD) estimate based on 'Top and Upper bridge', which will be corrected by estimate of g(0) for the combined platform 'Top and Upper bridge'. [2] CV does not consider any process errors. [3]Average of the SEs for the non-zero estimates. [4]The estimate of 0 from sub-area 7E was combined with the estimate of 391 from sub-area 8.

REFERENCE

International Whaling Commission. 2012. Report of the first RMP intersessional workshop for western North Pacific common minke whales. *J. Cetacean Res. Manage. (Suppl.)* 13:411-60.

Appendix 4

CONDITIONING 'BEST FIT' BYCATCH TRAJECTORIES FOR THE 6 BASELINES

These will be made available on the IWC website.

Appendix 5

FACTORS IN THE *IMPLEMENTATION SIMULATION TRIALS*

Table 1

Factors in the *Implementation Simulation Trials.*

Stock structure hypothesis	
Stock structure hypothesis A	M
Stock structure hypothesis B	M
Stock structure hypothesis C	M
$MSYR_{mat}$	
1%	M
4%	H
g(0)	
0.8	H
1.00 (Trial 3)	M
Other stock structure issues	
With a 'C' stock (Trial 2)	M
Some 'O' or "O/W' animals in sub-area 10E (Trial 5)	M
10% J (/JW) – stock in sub-area 12SE in June (Trial 10)	M
30% J (/JW) – stock in sub-area 12SE in June (Trial 11)	M
No 'C' animals in sub-area 12NE (Trial 12)	M
No 'OW' in 11 and 12SW (Trial 13)	M
No 'OE' in 11 or 12SW (Trial 14)	M
No 'OE' in 7WR (Trial 15)	M
Single J-stock (Trial 23)	M
Single O-stock (Trial 24)	L
Catches and bycatches	
High direct catches + alternative Korean + Japanese bycatch level (Trial 4) (Total direct catch = 40,224 cf baseline value = 38,174)	M
More Korean catches in sub-area 5 (and fewer in 6W) (Trial 8)	M
More Korean catches in sub-area 6W (and fewer in 5) (Trial 9)	M
Chinese incidental catch = 0 (Trial 18) (Baseline value = 2* Korean bycatch in subarea 5)	M
Number of bycaught animals is proportional to square root of abundance (Trial 25)	M
Mixing and dispersion	
Mixing proportion in 7Cs and 7CN calculated using 2/60 weight for bycatch (Trial 6)	M
Mixing proportion in 7Cs and 7CN calculated using 10/60 weight for bycatch (Trial 7)	M
Dispersal rate of 0.005 (Trial 16)	M
Dispersal rate of 0.02 (Trial 17)	M
A substantially larger fraction of whales 1-4 from O-/OE-stock are found in sub-areas 2R, 3 and 4 year round (Trial 26)	M
Set the proportion of O/OE animals of ages 1-4 in sub-area 9 and 9N to zero (Trial 27)	M
Abundance estimates	
Alternative abundance estimates in 6E (Trial 19)	M
Alternative abundance estimates in 10E in 2007 (Trial 20)	M
Abundance estimate in 5 = 'minimum' (Trial 21)	L
Abundance estimate in 5 = 'maximum' (Trial 22)	M
The number of 1+ whales in 2009 in sub-area 2C in any month < 200 (Trial 28)	M
Abundance estimate in 6W = 'minimum' (Trial 29)	L
Abundance estimate in 6W = 'maximum' (Trial 30)	M

Appendix 6

SUMMARY TABLE SHOWING SOURCES OF SUPPORT FROM THREE DATA CATEGORIES FOR DIFFERENT HYPOTHESES

R. Hoelzel, R. Waples and O. Gaggiotti

Summary Table showing sources of support from three data categories for different hypotheses (further to Appendix D1: proposed framework for scenarios with illustrative evidence – see for detail and citations). 'Incomplete' means no data currently available, but may be found following further literature review. 'None reported' means no data found following review at SC/63 in Tromsø.

Table 1

Sources of support for the different hypotheses.

Data	There are two clearly differentiated stocks, O and J that sometimes mix.	Mixing (in various proportions) of O and J is a sufficient explanation for all proposed stocks	There is a coastal migratory population in Subarea 7 (OW) that is different from a putative offshore 'OE' population.	There is a distinction between J-stock-derived whales east and west of Japan (JE vs JW).	There is a separate stock of whales in the Yellow Sea (Y) that may seasonally migrate into the Sea of Japan.
Genetic data based on mtDNA (matriline inference).	Clear haplotype frequency differences.	None reported.	None reported.	Significant differences between 6E and 2C; also the case for haplotypes rare in areas 8 and 9.	Further data needed (no significant difference found but some sample sizes were small).
Genetic data based on nuclear DNA (bi-parental inheritance reflecting both male- and female-mediated gene flow).	Strong evidence from both allozyme and microsatellite DNA data using various methods.	Analyses in STRUCTURE indicate that two populations (O and J) best fit the data, and analysis using 'HWLER' found only weak support for more than 2 stocks. However, in areas with putative mixing, 1- and 2-locus Wahlund effects were weaker than expected for simple O/J mixing.	Tests based on HWE deviation indicate mixing of O and J stocks and perhaps one or more additional stocks, and PCA analyses in areas 7-9 indicated genetic substructuring within O-stock individuals. However, for F_{st}, no significant differences among O-type whales from areas 2C,7Cs, 7CN, 7W, 7E, 8 and 9.	Small significant differences between areas 6E and 2C based on F_{st}, but not among J-type whales from 6E, 2C, 7CS and 7CN. Evidence of mixing of J and O stocks and perhaps or more stocks based on HWE deviations. However, individual-based PCA for areas 2C and 6 consistent with panmixia.	Significant but small F_{st} values found for comparisons between SA5 and various Sea of Japan samples, and significant deviations found for seasonal samples in the East Sea suggesting the migration of a Yellow Sea stock into the Sea of Japan. However, no substructure detected using the program STRUCTURE.
Morphological and life history data.	Some data on fluke and flipper coloration, but summary incomplete.	Incomplete.	Incomplete.	Incomplete.	Incomplete.

Appendix 7

SYNTHESIS OF RESULTS OF TESTS FOR ONE- AND TWO-LOCUS HARDY-WEINBERG EQUILIBRIUM

Robin Waples

A number of recent IWC papers have discussed tests of one- and two-locus Hardy-Weinberg equilibrium in North Pacific minke whales. As these results feature prominently in arguments for and against different stock-structure hypotheses, I briefly summarise what I believe are the most important results and discuss their relevance for assessing plausibility.

Departures from Hardy-Weinberg equilibrium at single gene loci are typically indexed by the statistic F_{IS}, which indicates whether heterozygotes are less frequent than expected (positive F_{IS}), more frequent than expected (negative F_{IS}), or exactly as frequent as expected (F_{IS}=0). Most attention has focused on samples that produce a positive F_{IS}, because a deficiency of heterozygotes is expected when a sample includes individuals from 2 or more populations with different allele frequencies. Heterozygote deficiencies can also be caused by null alleles or some types of scoring errors in the laboratory, but the fact that samples from putatively pure J or O samples do not show Hardy-Weinberg departures, e.g. SA6 and SA9 (see Appendix 6 in IWC, 2012) suggests that the positive F_{is} values found in other areas are not due to this factor and reflect actual mixing.

One argument in support of two O-like stocks is that samples from subareas 7CS and 7CN do not show evidence of Hardy-Weinberg departures suggestive of mixing, yet allele frequencies in these samples differ from those taken farther east – see Slikas and Baker (2011); SC/64/

NPM11. However, the analyses this conclusion is based upon considered individual loci and used a Bonferroni correction for multiple testing, which will tend to produce a conservative result. A more informative approach is to use an overall test of F_{IS} across all 16 loci. This test found a statistically significant heterozygote deficiency in both 7W-K and 7W-S, as well as in SA2, SA7bycatch, and the combined sample of 7W-K and 7W-S (see Appendix 6 in IWC, 2012). I believe the most plausible explanation for these results is that individuals from two or more stocks have contributed to the samples in each of these areas.

Proponents of Hypotheses A and B argue that genetic differences found between samples from several different areas can all be explained by different mixture fractions of the same two stocks (J and O). If this hypothesis is true, a consistent pattern should be found in each area: the loci with the most positive F_{IS} values should be those for which allele frequencies differ the most between O and J stocks. Waples (2011) evaluated this hypothesis by comparing the correlation between F_{IS} and F_{ST} (a measure of allele frequency difference) in samples from putative mixtures with correlations found in artificial mixtures of 'pure' O and J individuals. In general, the correlations from the actual samples were weaker than found in the artificial mixtures, which was interpreted as evidence that at least some of the mixtures might include more than just two populations. However, this conclusion was tempered by the caveat that the method was novel and robustness of the results to a number of variables had not been evaluated. SC/64/NPM9 performed some sensitivity analyses of this idea using simulated data and found that, in artificial 1:1 mixtures of exactly two populations, the correlation between F_{IS} and F_{ST} was consistently very high as long as very large samples were used, but when a total of 200 individuals were used (100 from each population; comparable to sample sizes from some of the putative NPM mixtures), the correlations varied widely across replicates. The empirical correlation for the putative mixture most similar to the modeled conditions (SA7bycatch) fell within the lower 15% of the broad distribution found for simulated data involving a mix of just two populations.

The above analyses all considered correlations within individual samples. IWC (2012) also evaluated the correlations of F_{IS} values among pairs of samples from putative mixtures. If all these samples are mixtures of just two stocks, one would expect that the loci with high and low F_{IS} values would be consistent across samples. However, the actual correlations among pairs of the samples from subareas 2, 7bycatch, 7W-S, 7W-K, and 11 were distributed approximately randomly around 0. It seems likely that this cross-sample analysis could be affected by the same sensitivity to sample size shown in SC/64/NPM9, but this has not been evaluated.

Waples (2011) also considered two-locus departures from Hardy-Weinberg equilibrium (aka linkage disequilibrium = LD) and examined the correlation between LD at a pair of loci and the product of F_{ST} values for the two loci. As in the single-locus analyses, artificial mixtures of O and J individuals produced correlations that were higher than found in most empirical samples of putative O-J mixtures. SC/64/NPM9 did not evaluate effects of sample size on these two-locus correlations, but one might expect them to be more robust to small samples because the F_{IS} correlations are based on just 16 data points (one for each locus), while the LD correlations are based on 120 different pairwise comparisons of loci. Further simulation analyses and statistical tests of empirical data against the simulated distribution, controlling for sample size and mixture fraction, would help resolve these uncertainties.

In summary, significant overall heterozygote deficits are consistently found in subareas 2, 7bycatch, 7W-S, and 7W-K, and I believe this is the result of mixing of individuals from more than one population. As to whether the empirical patterns can all be explained by a mixture of just O and J individuals, the results are more equivocal. SC/64/NPM9 shows that the high correlations theoretically expected from a mixture of just two well-characterised populations can, under some circumstances at least, be much smaller as a result of random sampling error. But, it is not clear whether this factor alone explains all the discrepancy noted in Waples (2011) and Appendix 6 between expected and empirical correlations. The degree to which the sensitivity of correlations within samples to sample size also applies to the LD correlations and those across pairs of samples has not been evaluated. Overall, at this point and pending further sensitivity analyses, I believe the evidence from one- and two-locus Hardy Weinberg departures for more than 2 stocks is only weak to moderate.

REFERENCES

International Whaling Commission. 2012. Report of the Scientific Committee. Annex D1. Report of the Working Group on the *Implementation Review* for western North Pacific common minke whales. *J. Cetacean Res. Manage. (Suppl.)* 13:102-29.

Slikas, B. and Baker, C.S. 2011. Investigation of Hardy-Weinberg equilibria and population differentiation in North Pacific minke whales, based on microsatellite genotypes of Japanese 'bycatch' and scientific whaling. Paper SC/63/RMP17 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 6pp. [Paper available from the Office of this Journal].

Waples, R.S. 2011. Can evidence for spatial and/or temporal genetic heterogeneity of North Pacific minke whales be explained by different mixture fractions of the same two core stocks, or is it necessary to postulate an additional stock(s)? Paper SC/63/RMP7 presented to the IWC Scientific Committee, June 2011, Tromsø, Norway (unpublished). 7pp. [Paper available from the Office of this Journal].

Appendix 8

CONSIDERATION OF STOCK PLAUSIBILITY RELEVANT TO HYPOTHESES B & C BASED ONLY ON GENETIC DATA

R. Hoelzel, R. Waples, O. Gaggiotti, P. Palsbøll and R. Tiedemann

Short summary: The genetic analyses run to date all have uncertainties associated with them, however based on available data we cannot exclude the possible existence of a Y stock in the Yellow Sea or an OW stock in area 7, in addition to the core J and O stocks. Although these data cannot completely exclude the existence of a JE stock, we feel the balance of evidence is against it.

Table 1

Stock plausibility.

Stock	Evidence for:	Evidence against:
Y	**Moderate**: Significant microsatellite DNA F_{st} values between SA5 and Sea of Japan samples, and seasonal evidence for mixing based on HW deviation.	**Low**: No significant F_{st} based on mtDNA, but small sample sizes. Microsatellite DNA F_{st} values small.
JW (J)	**High**: Case for a core J stock is strong based on various data.	N/A
JE	**Low**: Significant mtDNA F_{st} comparing 6E and 2 based on non-purged dataset. Differential haplotype frequencies for two most common haplotypes. Weak suggestion of one or more additional stocks based on initial one and two locus Wahlund effects.	**Moderate or High**: F_{st} values were very small. Some mtDNA haplotype data suggest mixing between J and O (e.g. the number of haplotypes per individual is 0.19 in 2BC compared to 0.08 in 6BC and 0.10 in 8 and 9). Possible considerations of temporal aspects of comparisons (with bycatch not representing the same time period as the hunt). PCA found no evidence for differentiation between 2C and 6E. One and two locus Wahlund effect method requires further trials, as evidenced by preliminary simulation data presented in SC/64/NPM9. Fis results for SA2 indicate a mixture (not a new pure stock) and appear consistent with a mixture of just O and J.
OW	**Moderate**: PCA results using J-purged O stock sample provided support for an additional stock in OW compared to OE. Significant F_{st} differentiation comparing non-purged samples. Suggestion of one or more additional stocks based on initial one and two locus Wahlund effects. Haplogroup data suggests different frequencies in 7CN and 7CS regions compared to O-stock (8 and 9).	**Low or Moderate**: PCA work requires simulation analysis to evaluate effects of purging. PCA regression against length was not significant. Small but positive Fis considering all loci together suggests mixing in 7W-K and 7W-S regions. One and two locus Wahlund effect method requires further trials, as evidenced by preliminary simulation data presented in SC/64/NPM9. Haplogroup data (based on two SNPs) not clearly inconsistent with mixing.
OE (O)	**High**: Case for a core O stock is strong based on various data.	N/A

Appendix 9

NORTH PACIFIC MINKE WHALE *IMPLEMENTATION SIMULATION TRIAL* SPECIFICATIONS

The latest version of these will be made available on the IWC website.

Appendix 10

ADDITIONAL SPECIFICATIONS RE: CATCHES TO BE USED IN THE NORTH PACIFIC MINKE WHALE *IMPLEMENTATION SIMULATION TRIALS*

(a) Future bycatches

The future projections need to generate future bycatches by sub-area. The most parsimonious way to achieve this is to assume that the exploitation rate due to bycatch in the future equals that estimated for the trial in question for the most recent five-years, i.e.:

$$C^k_{B,t} = \overline{F}^k \, P^k_t$$

Where $C^k_{B,t}$ is the by-catch in sub-area k in year t, P^k_t is the total population (including calves) in sub-area k in year t averaged over all 8 time periods (March-October), and $\overline{F}^k$ is the average exploitation rate (sum over years of bycatch divided by the sum over years of P^k_t) over the last five years of the period used for conditioning.

(b) Reported catches

The magnitude of the past commercial catches when applying the RMP in trials should be set to those for the baseline trials. Consequently, the RMP will use what are in effect incorrect catches for trials 4, 8 and 9 which examine the implications of uncertainty about historical catches.

The by-catches used by the RMP variants should be the true by-catches, except for: (i) subarea 6W in which the reported bycatches are assumed in the base case trials to be in error by a factor of 2; and (ii) trial 4 in which the reported by-catches should be in error to reflect the under-estimation of bycatch inherent in these trials.

Appendix 11

WORK PLAN TO COMPLETE THE *IMPLEMENTATION REVIEW*

Step	Activity	Deadline	Who
1	Final trial specifications distributed to Steering Group	15 Jul. 2012	CA
2	Finalise abundance estimates for the use in the *CLA* in trials	31 Jul. 2012	TM and ISG
3	Review of final trial specifications	31 Jul. 2012	ISG
4	All 'best fit' conditioning results provided to the Steering Group	31 Jul. 2012	ISG
5	Base-case trials conditioning (100 simulations) completed	15 Aug. 2012	CA and CD
6	Run base-case trials with RMP variants1 and 11 in which the 'No agreement' abundance estimates are used or not used	31 Aug. 2012	CA
7	Decision to use or adjust or not use 'No agreement' abundance estimates	30 Sep. 2012	ISG
8	Complete all conditioning	31 Dec. 2012	CA and CD
9	Develop 'equivalent single stock trials'	31 Jan. 2013	CA, CD and AP
10	Complete all projections	Before Workshop	CA, CD and AP
11	Apply criteria for evaluating conservation performance	Before Workshop	CA, CD and AP
12	Workshop (late Feb.-early Apr. 2013) Review trial results Make recommendations on management areas, RMP variants, associated operational constraints Make suggestions for future research Identify 'less conservative' variants which may be 'acceptable with research'		
13	Conduct projections to evaluate variants that may be 'acceptable with research'	Before SC/65	CA and CD