Annex I

Report of the Working Group on Stock Definition

Members: Bravington (Convenor), Amaral, Baird, Bickham, Cipriano, Donovan, Double, Goto, Hoelzel, Kanda, Lyrholm, Palsbøll, Punt, Silva, Skaug, Tiedemann, Verborgh, Waples.

1. INTRODUCTORY ITEMS

1.1 Election of Chair and appointment of rapporteurs

Bravington welcomed participants, was elected as Chair, and appointed himself as rapporteur.

1.2 Adoption of agenda

The agenda adopted is given as Appendix 1.

1.3 Review of documents

The primary document considered was SC/61/SD1.

2. STATISTICAL AND GENETIC ISSUES RELATING TO STOCK DEFINITION

2.1 Guidelines for DNA data quality

In recent years, the Scientific Committee has engaged in several in-depth discussions centred on the genetic data that form part of the delineation of stock structure hypotheses, for example in the bowhead whale *Implementation Review*. The Committee's experiences have underlined that a clear understanding of the reliability of each genetic dataset is essential for correct interpretation in terms of stock structure, and have re-emphasised the importance of developing suitable quality protocols for genetic data used in providing management advice, including associated issues with the Data Availability Agreement. In recent years, the SDWG has been developing such guidelines to cover marker validation and systematic quality control throughout the entirety of a genetic study. These guidelines may currently be found in IWC (2008, [SC60 Report], pp252-256). They were endorsed in that meeting by the SC, which stated: "*Although accordance to the guidelines in Appendix 2 is highly desirable, this does not preclude consideration of genetic work failing to fully meet these standards (though all studies should endeavour at least to report on whether the guidelines in Appendix 2 have been met). The guidelines, if met, should assist IWC SC members in judging the respective reliability of information from genetic studies. In addition, for studies explicitly carried out to give stock definition advice to the IWC, adherence to the guidelines is strongly recommended.*" The guidelines constitute a 'living document' that will be updated by the SC to keep up with the rapid progress in genetic techniques.

This year, the SDWG had planned to include some actual numerical guidelines for measure of data quality (noting that generic numerical guidelines may not be appropriate for all measures of data quality). However, there was no opportunity to start this intersessionally, so the topic was deferred until next year. To fuel discussion next year, Tiedemann offered to lead an intersessional email group in conducting a literature review of the range of error rates and other quality measures associated with studies published in peer reviewed journals.

2.2 Guidelines for genetic analyses

In parallel with the development of data quality guidelines, the Committee has asked the SDWG to provide guidelines for some of the more common types of statistical analyses of genetic data that are employed in IWC management contexts. These will update and considerably expand the initial attempt made during the 2nd TOSSM workshop (JCRM 2007, SC/58/Rep5). The guidelines should cover two aspects: comments on general statistical usage (e.g. multiple comparison tests in the context of stock structure), plus more general summaries on the appropriate domains of application of different stock structure tools such as STRUCTURE, BayesAss, etc.

SC/61/SD1 proposed an outline for the analytical guidelines document. The exercise is complex, since numerous types of genetic analyses are possible and users have a wide range of choices of software programs for implementing the analyses. Furthermore, to maximize practical usefulness, the document should help to clarify the link between collection and analysis of genetic data and use of this information in a management context. The stated goal is to produce a document that will be useful both for geneticists (in helping to provide rigorous and consistent standards for data analysis) and non-geneticists (in explaining the background, key assumptions, and advantages and limitations of various approaches). SC/61/SD1 included a draft outline for the guidelines document, and three worked examples to provide a flavour of the level of detail that might be provided for specific sections. Presentation of the draft document at an early stage in its development was intended to elicit comments and suggestions by non-geneticists as well as geneticists to facilitate development of a final product that would be of maximal practical usefulness.

The SDWG thanked Waples and the other authors of SC/61/SD1 for their efforts, and agreed to develop the document along the lines proposed, with some modifications. The structure is shown in Appendix 2, and summarized here.

To set out the context for a non-IWC audience, the document will start with an example: a description of a real IWC management conundrum involving alternative stock structure scenarios. It was agreed that Western North Pacific Bryde's whales would be suitable, since the situation is not so complicated as to obscure the generic points. A draft of this example is given in Appendix 3.

The guidelines themselves will address five main categories:
(1) Species ID and delimitation.
(2) Analysis of diversity within populations.
(3) Estimation of population size (census, effective, and historical).
(4) Analysis of stock structure, i.e. diversity across populations, including discussion of particular algorithms/software. Descriptions of 1-2 page length, with comments on domain of applicability, pitfalls, and appropriate interpretation.
(5) Generic issues in analysis (e.g. multiple testing, MCMC issues, influence of selection, interpretation of negative results).

The document will end by returning to the original example (and perhaps others, to be added). It will first consider what might be possible given an ideal sampling scheme, then how best to proceed with the kind of samples available now, and then whether there are implications for appropriate data collection as well as analysis. The document will work through the analytical steps that might usefully be taken, and will point out some that should not be.

While the SDWG's work mainly concerns guideline category (4) (which is also the focus of the introductory example), all five categories are of current importance to the Committee, and it should be possible to make progress on several fronts simultaneously. With respect to guidelines on specific software packages, it was noted that software in this field evolves quickly, so that the guidelines will sometimes be dealing with a moving target; nevertheless, it would be very valuable to give software-specific comments where feasible. Again, the guidelines will be a 'living document' subject to ongoing revision.

To progress development of the guidelines, Waples agreed to convene an intersessional email group (Annex Q). Individual members of the SDWG offered to address particular sub-categories, as listed below. Sections of particular relevance to North Pacific Minke IST will be identified and it is hoped to have an initial draft ready for these by 1 Oct, so that the sections can be reviewed amongst the email group and then circulated to those involved in the NPM intersessional workshop. It is hoped to bring a draft paper with most sections addressed to SC62 meeting.

2.2.1 Relationship between biological populations and management stocks
In discussion of SC/61/SD1, the SDWG observed, as it often has before, that the stock-related demograph parameters of relevance to management (e.g. mixing proportions, per capita dispersal rates per year) are quite different to the genetic differentiation parameters typically considered in studies of population genetics (e.g. Fst, absolute numbers of migrants per generation). Particularly in the context of guidelines meant for geneticists as well as IWC stalwarts, it is important to try to bridge the gap between concepts. To this end, the SDWG considered a simple hypothetical example corresponding to Archetype I in TOSSM, in which two breeding populations are linked by dispersal, and the catch level is proportional to the combined abundance, but all the catch is taken from Population 1. Because the catch is calculated for the 'wrong' animals, it is possible that P1 will become over-exploited. Whether this actually occurs depends on three key parameters: the overall catch rate, the proportion of the total abundance coming from Population 1, and the dispersal rates. If P1 is much larger than P2, then it is able to withstand the catches provided the overall catch rate is not too high, regardless of the dispersal rate. Also, if P1 is small but the dispersal rate is high enough, then migrants from P2 can replenish P1 fast enough to compensate for the excessive catch. This is illustrated in Figure 1, for a fixed catch rate; see Appendix 4 for the underlying population dynamic equations.

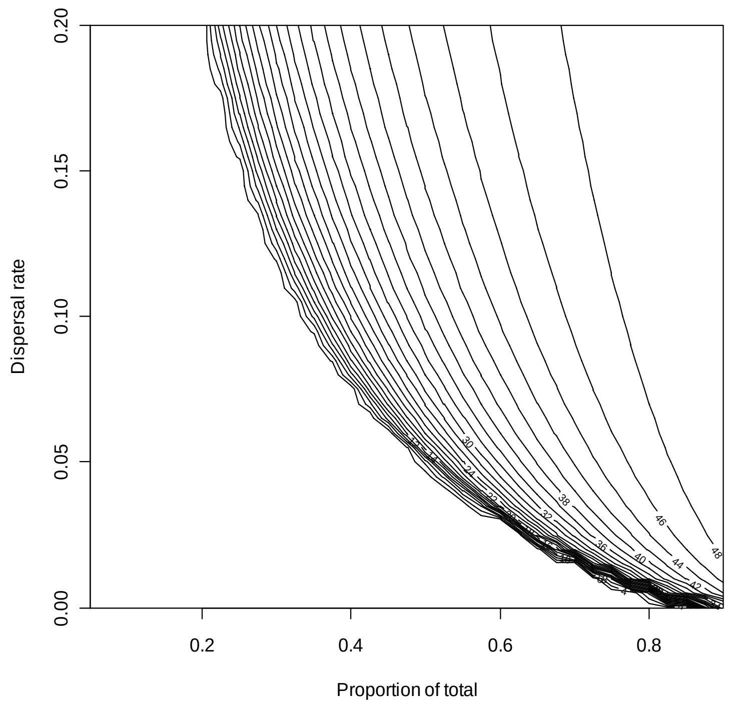

Figure 1. Behaviour of one of a pair of linked breeding stocks when only one is exploited. X-axis shows pre-exploitation proportion of the combined abundance coming from the exploited population. Y-axis shows the per capita annual dispersal rate from the exploited population. Contour lines show relative abundance of exploited population after 200 years; the highest (rightmost) contour is close to 50%, the lowest is 0%.

In this scenario, given a particular catch rate and a particular abundance ratio P1:P2, it is possible to calculate a threshold dispersal rate above which over-exploitation will not occur, corresponding to one of the contour lines in Figure 1 and depending on the definition of over-exploitation. For management purposes, the 'one stock or two stocks' question corresponds to whether the true dispersal rate is above or below the threshold, regardless of whether the biological reality is two breeding populations or just one. Given a few other demographic parameters, which typically would be estimable for whales, a threshold demographic dispersal rate can in principle be translated into thresholds for corresponding genetic quantities such as migrants-per-generation or Fst (see Appendix 4). Note that no actual genetic data is required for this step, since the threshold can be calculated for population dynamic quantities alone. While this translation is not necessarily accurate, it is straightforward, and does at least provide a reference point for considering the appropriateness of various genetic methods. If the threshold Fst turns out to be exceedingly low, for example, then assignment methods will have very low power (unless perhaps 1000s of SNPs are available) and close-kin methods might be the only option.

The SDWG noted that it would not be hard to develop similar illustrative models for other Archetypes of stock structure, e.g. in order to illustrate how mixing affects matters. Such examples may be useful both in developing the guidelines above, and in considering how to 'tune' genetic methods for testing in TOSSM.

3. TESTING OF SPATIAL STRUCTURE METHODS

The general aim of the TOSSM project is to facilitate comparative performance testing of population structure methods intended for use in conservation planning (Martien *et al.*, in press). From an IWC perspective, the TOSSM software package, written in R, allows evaluation of methods for detecting genetic structure, in terms of how successfully they can be used to set spatial boundaries for management. Following intersessional work this year, the basic TOSSM framework is now complete, and the software is available to all as a fully-documented R package on CRAN (*http://www.cran.r-project.org*), and a paper describing TOSSM has been accepted in *Molecular Ecology Resources*. An MS describing the results of some of the testing that has already happened (e.g. of STRUCTURE; SC/59/SD3) is in preparation. Maintenance and development of TOSSM rests with Karen Martien at SWFSC, where the underlying TOSSM code is being used in various genetics and management projects. Martien expressed to the SDWG her willingness to help anyone wishing to use or extend TOSSM to explore genetic methods in a management context, whether in the IWC or beyond.

Development of reference datasets for Archetypes I (single stock), II (two stocks with dispersal), and IV (dispersal + mixing) is complete. Generating these reference datasets takes a little work - more than is required to test particular population genetics methods using the reference datasets - because it is necessary to tune some of the genetic parameters in order to achieve plausible distributions of present-day allele frequencies. SWFSC itself has no current plans to develop Archetype III (isolation-by-distance) or V (persistent feeding stocks) specifically, but the software for making reference datasets is available to all from the TOSSM website.

The SDWG recalled that a number of methods have already been tested with TOSSM (SC58, SC59, SC60), but to date only for the simplest Archetypes I and II. This year, the SDWG considered those stock structure hypotheses for large cetaceans that are currently being considered by the Committee (ENP grays, ENP bowheads, Antarctic minkes, WNP Bryde's, NA minkes, NA fins, WNP minkes, humpbacks) in the light of the TOSSM archetypes. Archetype IV is most common, often with some elements of Archetype V. The SDWG encouraged those involved in largely-Archetype IV cases (e.g. in NPM) to consider whether TOSSM can be used as an easy way to evaluate the power of e.g. STRUCTURE, noting that this had already been tackled in SC/59/SD3 for Archetypes I and II. It may or may not prove necessary to develop further reference datasets for specific stock structure hypotheses such as in NPM.

Although Archetype IV/V cases may be the primary focus for now, it was noted that the isolation-by-distance model of Archetype III remains highly relevant (including to Small Cetaceans sub-committee), and that development of this Archetype and testing of methods on it would be useful.

No new results on tests of population genetic methods were presented this year.

4. WORK PLAN

The SDWG agreed that its work plan for the 2010 Annual meeting would be as follows.

(1) Progress on TOSSM (new tests of methods new reference datasets)
(2) Update guidelines on DNA Data Quality
(3) Review proposed guidelines on analysis of genetic data for use in management
(4) Other statistical and genetic issues related to stock definition
(5) Consideration of possible definitions of 'unit to conserve' (noting that Appendix 2 this year represents significant progress)

Intersessional email groups will address (i) DNA quality (Tiedemann) and (ii) Guidelines on analysis of genetic data for use in management (Waples).

Appendix 1

AGENDA

1. Introductory matters
1.1 Convenor's opening remarks
1.2 Election of chair and appointment of rapporteurs
1.3 Adoption of agenda
1.4 Review of documents

2. Statistical and genetic issues relating to stock definition
2.1 DNA Data Quality
2.2 Guidelines for analysis methods

3. TOSSM (Testing of Spatial Structure Models)
3.1 Update on progress
3.2 Directions for further work

4. Overall work plan

5. Adoption of report

Appendix 2

PROPOSED STRUCTURE OF THE ANALYSIS GUIDELINES DOCUMENT

I. Species identification and delimitation
II. Analysis of diversity within populations
   A. Information related to tests of Hardy-Weinberg (HW) equilibrium. We assume that HW evaluations have been conducted as part of the DNA data quality control step; here, attention would focus on HW deviations that might provide insight into biological processes (such as inbreeding or population mixtures).
   B. Information related to tests of linkage disequilibrium (LD). As was the case for HW, the focus would be on signals that might provide insight into biological processes.
   C. Measures of genetic diversity, including rarefaction (controlling for sample size in estimating allelic richness).
III. Estimating population size
   A. Census size, *N*
      1. DNA mark-recapture
      2. Analysis of close relatives
      3. Identifying recent population bottlenecks
   B. Effective population size, N_e
      1. Historical N_e (including coalescent)
      2. Contemporary N_e
         a. Single-sample methods
         b. Two-sample (temporal) methods
      3. Recent vs. ancestral N_e using isolation with migration models
      4. Signals of historical population expansion/contraction
   C. N_e/N ratios
IV. Analysis of diversity among populations (aka stock structure)
   A. Putative populations defined *a priori*.
      1. Testing panmixia
      2. Describing population structure
         a. F_{ST}, genetic distance, and related measures
         b. Ordination
         c. Isolation by distance/Landscape genetics (units = samples)
      3. Estimating migration
         a. Methods that assume migration-drift equilibrium and estimate long-term patterns of gene flow (mN_e)
            i. Island model
            ii. Stepping-stone models
            iii. Coalescent methods and directional gene flow
         b. Assignment methods that estimate contemporary migration rate (*m*)
         c. Isolation with migration models to estimate splitting times and post-division migration rates
      4. Mixture analysis (e.g., resolving stock composition of samples from feeding grounds or migration pathways) Robin
      5. Kinship
   B. No *a priori* basis (or a questionable basis) for grouping individuals into populations
      1. Clustering programs
      2. Clustering based on ordination
      3. Landscape genetics (units = individuals) including temporal
      4. Kinship
V. Generic/cross-cutting issues
   A. Choice of markers (mtDNA, microsats, SNPs)
   B. Ascertainment bias
   C. Multiple testing
   D. Mutation rates
   E. Confidence intervals, sampling and experimental design
   F. Underlying assumptions and sensitivity to their violation
   G. Bayesian vs. maximum likelihood vs. frequentist methods
   H. MCMC issues (burnin, convergence)
   I. Integrating genetic and non-genetic data
   J. Possible influence of selection
   K. Interpreting negative results
VI. Return to introductory example, and work through
VII. Summary and conclusions

Appendix 3

WESTERN NORTH PACIFIC BRYDE'S WHALES:
A SUMMARY FOR USE IN THE ANALYSIS GUIDELINES DOCUMENT

Bryde's whales feed during the Northern summer in mid-latitudes across the North Pacific, eastwards from at least 130°E. They migrate there from breeding ground(s) further south; the location of these ground(s) is unknown. Separate genetically-distinct population(s) with mirror-image behaviour exist in the Southern Hemisphere. Taxonomically, Bryde's whales form a complex, with for example distinct and small coastal forms in several places, but all the discussion here concerns the offshore form only. There may or may not be stock substructure within the feeding grounds, and/or on the breeding grounds. The IWC is considering proposals for an ongoing catch of Bryde's whales on the feeding grounds between 130°E (W Japan) and 155°W/205°E (Hawaii) and about 10°N to 40°N.

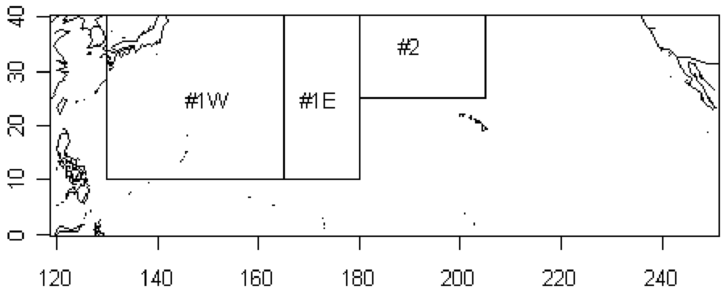

Figure 1. Areas of the North Pacific where the IWC is considering catches of Bryde's whales. See main text for meaning of boxes.

Some data on basic biological and population dynamic parameters is available.

- Bryde's whales mature around age 7, and median age is around 16.
- The abundance for areas #1W, #1E and #2 combined was estimated to be about 22000 in 1995. No reliable estimates are available outside these areas, so if these areas only constitute part of the feeding grounds of the underlying stocks, the population abundance will be higher.
- Catches have been taken since at least 1906 primarily from #1E and #1W. Annual takes were in the low hundreds until the 1970s, with roughly 1000 animals per year taken until 1987 when catching stopped. Catches of around 50 per year resumed in 2000.

In terms of non-genetic data that might inform stock structure: there is no discontinuity in distribution within the catch area, nor any evidence of morphometric differences. Differences in age composition have been found between the three sub-areas; although interpretation is unclear, this provides part of the motivation for considering the possibility stock structure. Mark-recapture shows that movements do occur across the two western sub-areas; there has not been enough effort in the easternmost sub-area, #2, to tell how much movement there is to/from #2. Two satellite tags have now been deployed, for short durations.

In terms of genetic data:

- single-locus allozyme data has shown differences between WNP Bryde's' and other Bryde's' whales in the Pacific, but not within the catch zones;
- mitochondrial and microsatellite DNA has been collected from the two western sub-areas in the 1980s and 2000s, with no significant heterogeneity revealed;
- no genetic samples have been collected from the easternmost sub-area;
- no genetic samples are available from the breeding ground(s), and there is little prospect of getting any.

In terms of proposed future catches: catch levels over time would be regulated as usual under the RMP, based on time series of abundance estimates collected across the catch area. The quota-setting rule has been shown to have minimal risk of causing unacceptable depletion if applied to a single freely-mixing population. However, if there was population structure on the catch grounds, and if the entire quota was taken from a sub-area containing mostly one subpopulation, then it would be possible to cause unacceptable local depletion in that sub-area over, say, a 100-year timeframe; any refill from subpopulations that primarily feed in other sub-areas would be too slow in management terms. The question before the Scientific Committee is whether it is necessary to divide the catch quota spatially across this large area, to avoid risk of local depletions. Currently, three sub-areas are under consideration (#1W=130°E-165°E, #1E=165°E-180°E, #2=180°E-205°E/155°W).

The boundaries of these sub-areas are fairly arbitrary, biologically speaking, and they simply represent a tractable way in which catches could be divided if necessary. The five stock structure hypotheses currently under consideration (ref. to add) concern the extent of mixing between different putative subpopulations in the different sub-areas. The range of hypotheses and indeed the sub-area boundaries are not set in stone, so there are maybe three distinct tasks to consider:

- to use genetic data to evaluate the relative plausibility of the various hypotheses currently under consideration;
- to use genetic data to help parameterise quantities such as mixing proportions (of breeding sub-populations in feeding sub-areas) and dispersal proportions (migration between sub-populations);
- to consider whether the range of hypotheses needs alteration/expansion.

Appendix 4

EQUATIONS FOR HARVESTING UNDER DISPERSAL EXAMPLE, AND LINK TO GENETIC DIFFERENTIATION PARAMETERS

The relationship between harvesting, dispersal rates and conservation goals can be explored using the following simple population dynamics model for two populations linked by dispersal.

$$
\begin{aligned}
N^1_{y+1} &= N^1_y + rN^1_y(1-N^1_y/K^1) - C_y - \phi^{1\to 2}N^1_y + \phi^{2\to 1}N^2_y \\
N^2_{y+1} &= N^2_y + rN^2_y(1-N^2_y/K^2) + \phi^{1\to 2}N^1_y - \phi^{2\to 1}N^2_y
\end{aligned}
\qquad (1)
$$

where N^i_y is the number of animals in stock *i* at the start of year *y*,

r is the intrinsic rate of growth (assumed to be 0.1),

K^i is the carrying capacity of stock *i*:

$$K^1 = \alpha K^{\text{tot}}; K^2 = (1-\alpha)K^{\text{tot}} \qquad (2)$$

K^{tot} is the total carrying capacity (assumed to be 1000),

α is the proportion which stock 1 is of the total at pre-exploitation equilibrium,

$\phi^{i\to j}$ is the proportion of animals of stock *i* which migrate (permanently) each year to stock *j*; by definition:

$$\phi^{2\to 1} = \phi^{1\to 2}K^1 / K^2 \qquad (3)$$

C_y is the catch during year *y*, assumed to be taken only from stock 1, and computed by applying an exploitation rate λ to the total population size (both stocks combined) at the start of year *y*, i.e.:

$$C_y = \lambda(N^1_y + N^2_y) \qquad (4)$$

The populations do not have the same carrying capacities and only one of the populations is subject to harvest, but the catch limit for population 1 is (somewhat unintentionally) based on the total numbers in both populations (because it is not known that there two populations). Note that the per capita dispersal rates at equilibrium must be different for the two populations, in order to maintain different population sizes.

For the calculations, the exploitation rate λ is set to 0.05, corresponding to *MSY* in the panmictic case (i.e. very high dispersal). In that case, the equilibrium abundance relative to unexploited would be 50%.

Translation between the dispersal rate phi and genetic differentiation parameters might be accomplished as follows. If plausible values for the N/Ne ratio and the generation time are available, per capita annual dispersal can be easily transformed into genetic dispersal rates (m*Ne). Under mutation-drift equilibrium and a specified migration model, there is a simple formula relating m*Ne to Fst or related quantities (Wright, 1978). However, under non-equilibrium scenarios or other violations of the rather stringent underlying assumptions, this formula is biassed, and more sophisticated methods (e.g. coalescence) would be recommended for analysis of real data. To what extent this extra sophistication would be worthwhile for purposes of setting a 'target Fst' will depend on the specific context being considered.