Annex I

Report of the Working Group on Stock Definition

Members: Bravington (Chair), Brandon, Daníelsdóttir, Donovan, Dueck, Edwards, Engel, Galletti, Givens, Goto, Hucke-Gaete, Jackson, Kitakado, Lauriano, Lens, Leslie, Peel, Ripley, Taylor, Waples.

1. INTRODUCTORY ITEMS

1.1 Opening remarks

Bravington welcomed participants and noted that the focus of the Stock Definition Working Group this year was an intersessional Workshop in Potsdam, March 2006, to progress the TOSSM (Testing of Spatial Structure Models) project. With the absence of several key personnel from this year's meeting, the agenda would be shorter than usual.

1.2 Election of Chair and appointment of rapporteurs

Bravington was elected Chair and also served as rapporteur.

1.3 Adoption of Agenda

The adopted Agenda is given as Appendix 1.

1.4 Review of documents

Documents considered were SC/58/SD1-3, SC/58/BRG13 and SC/58/Rep6(Draft).

2. STATISTICAL AND GENETIC ISSUES RELATING TO STOCK DEFINITION

SC/58/SD3 described a general model for pairwise microsatellite allele matching probabilities. The model can be used for analysis of population substructure, and is particularly focused on relating genetic correlation to measurable covariates. The approach is intended for cases when the existence of subpopulations is uncertain and *a priori* assignment of samples to hypothesised subpopulations is difficult or unappealing. The model estimates genetic correlation associated with population substructure and with other variables that may be related to the temporal, spatial, or other aspects of the samples. Hypothesis tests for population substructure and for covariate effects can be carried out using permutation methods. Simulated examples in SC/58/SD3 illustrated the effectiveness, power and reliability of the approach. The authors also demonstrated successful application of the model to real data. Finally, the authors reviewed the philosophy and utility of the method in comparison to other familiar approaches including the program STRUCTURE (Pritchard *et al.*, 2000) and permutation chi-square tests for allele frequency differences.

In discussion, it was noted that testing on real populations where structure is known to be absent might be informative, if such can indeed be found. The power might be increased further by using associations across multiple loci, and the author intends to pursue this. Since several different biologies can give rise to the exact same statistical model in SC/58/SD3, interpretation of any structure that is detected will require application of specific 'domain knowledge' for the population(s) being studied. In particular, since the idea behind the paper is to think of covariate effects as reflecting something different from population structure, and then attributing any residual excess homozygosity to population structure, it is important to use covariates that biologically cannot be linked to population structure.

The Working Group welcomed the work in SC/58/SD3, which completed the first phase of testing recommended last year. It recalled the Committee's Data Availability Arrangement (DAA), which seeks to ensure that analyses used in helping to formulate important management advice should have had the opportunity for a full scientific review; to ensure this, Procedure A of the DAA establishes distinct 2- and 3- month deadlines for analyses using 'standard' and 'novel' methods respectively. The Working Group has reviewed the analysis methods in SC/58/SD3 for two years, and has agreed it to be an appropriate method for analysis of population structure, especially in situations where a confounding covariate is suspected of contributing to genetic correlation; consequently, the Working Group recommended that the Data Availability Group no longer classify the methods of SC/58/SD3 as 'novel'.

The author of SC/58/SD3 plans to release *R* software for the single continuous covariate case, and extensions to multiple covariates are possible in principle. It was noted that the TOSSM datasets will provide a standard reference set which could be helpful in demonstrating the performance of methods such as that in SC/58/SD3. In order to test the particular methods of SC/58/SD3 in TOSSM, though, it would of course be necessary to provide an appropriate covariate. It was **agreed** to consider these matters further intersessionally.

SC/58/BRG13 (see Annex F) included simulation tests of the distribution of p-values of allele frequency differences between cohorts from the same population, in a case where a genetic bottleneck might lead to differences between the cohorts. Since there is no real 'population structure' in the simulations, a significant result would in a sense be a 'false' positive. Among several somewhat puzzling phenomena, it was found that the *p*-values were generally left-skewed with fewer false positives than expected in the null case of completely independent samples.

In discussion, it was noted that this effect would be consistent with a reduction in difference due to kinship effects between the cohorts; see e.g. Waples (1989). Results of further simulation tests are awaited.

3. TOSSM (TESTING OF SPATIAL STRUCTURE MODELS)

3.1 Report from intersessional Workshop

Donovan introduced the draft report (SC/58/Rep6) of the second TOSSM Workshop held at Potsdam University from St. Patrick's Day to 21 March 2006. The report will be finalised by email with the participants.

The primary achievements of the Workshop can be summarised as follows.

(1) Considerable progress was made in the detailed computing work needed to:

  (a) identify and fix problems in the linking of the coalescent (SIMCOAL) and individual based model (RMETASIM) to allow the generation of simulated datasets;
  (b) complete the run.tossm module that can be used to generate genetic samples from the datasets developed by RMETASIM, link this into the testing of candidate boundary setting methods and the running of management algorithms and ultimately to generate appropriate performance statistics.

(2) The technical specifications for the initial TOSSM trials (demographic structure, genetic structure, initialising the population matrix, harvesting and catch control, sampling and trials) were completed; the numbers of scenarios to be covered is 70 [5 archetypes $\times$ 7 dispersal rates $\times$ 2 (sample sizes for the genetic data)].
(3) An initial set of methods to be tested within the framework were identified as were issues related to automating these to determine boundaries and the persons who would 'champion' these (see Item 8).
(4) Preliminary results were available from two methods (MIXPROP and GENELAND), showing example boundary-setting algorithms in use through complete simulations of TOSSM.

A summary of the framework is given as Fig. 1.

Donovan thanked Tiedemann for hosting the Workshop and particularly Strand, Bravington, Punt and (remotely) Martien for their work with the computing issues and generation of datasets. The Workshop made a number of recommendations for future work and these are discussed under Item 3.3.

In discussion, the Working Group thanked the Workshop Chair and participants, and strongly endorsed the contents and general recommendations of the Workshop. It was noted that the considerable size of the simulated datasets made it more practical to store them at point of source (Southwest Fisheries Science Center, La Jolla, USA) rather than with the IWC. Comments on details of the simulations can be found in discussion of paper SC/58/SD2 below.

The possibility of also testing coalescent-based approaches such as MIGRATE (Beerli, 2006) was raised. MIGRATE itself was discussed at the first TOSSM workshop, but seemed too awkward to adapt to TOSSM. However, the program LAMARC (Beerli and Felsenstein, 2001) might be more suitable. LAMARC quantifies gene flow between multiple sub-populations under different conditions of population growth. Although it is computationally intensive, the time burden may not be prohibitive when the number of populations is small, as for most of the TOSSM archetypes. It was agreed that LAMARC should be investigated further for TOSSM. The program 'Isolation with Migration' (Hey and Nielsen, 2004) was also suggested for further investigation when time permits.

3.2 Update on progress since the Workshop

SC/58/SD1 presents preliminary results of running simulation tools developed by TOSSM, using a sequential hypothesis test procedure to set boundaries. Four stock structure scenarios are considered. These scenarios differ in the number of true stocks, and also in the underlying dispersal rates between stocks. The time horizon considered is 100 years. During the first five years of the projections, a catch of 300 animals is removed. Microsatellite frequency data are then sampled for 18 loci from 50 randomly selected

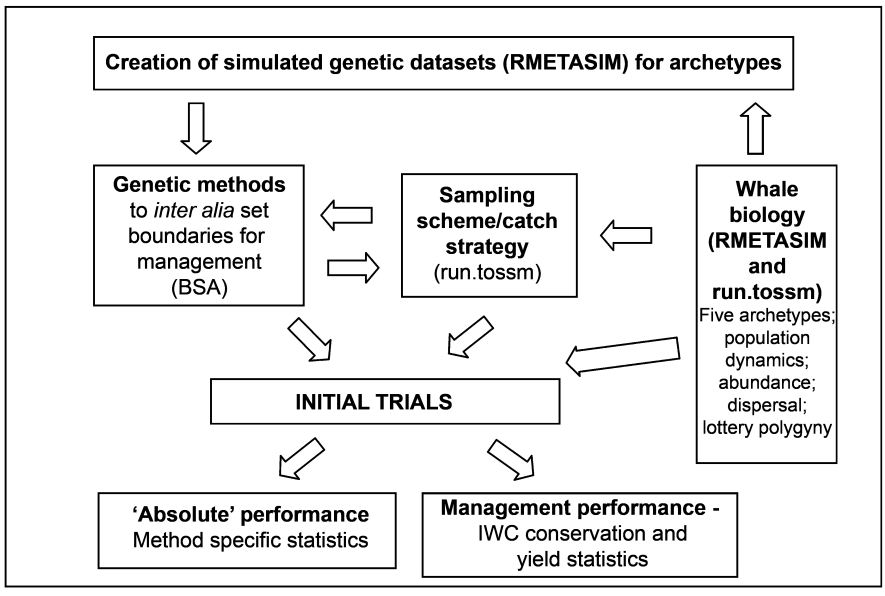

Fig. 1. How the initial trials in TOSSM are set up and run. Items in brackets are pieces of software where various options need to be set.

animals taken in each of the three FIMAs (Fully Internally Mixed Areas). These frequencies are input for the Boundary Setting Algorithm (BSA). The BSA uses sequential chi-square hypothesis testing to select among alternative stock boundaries. The resulting boundaries are used as the basis for splitting or lumping adjacent FIMAs into small areas for which catch limits are set. After the boundaries have been dictated, the model is then projected 95 years into the future during which time catches are subtracted from the small areas according to the *Catch Limit Algorithm* (*CLA*) of the Revised Management Procedure (RMP). The projections are repeated for each stock structure scenario over a range of alpha levels (all else being equal, given the chi-square null hypothesis of panmixia, larger alphas correspond to a greater tendency for the BSA to lump areas). In general, the results for this BSA are better than the authors expected. The results suggest that this method can be tuned to avoid unintentionally depleting stocks in multi-stock situations without major losses in catch. However, these results are preliminary, and additional analyses based on a larger number of scenarios (e.g. those with higher dispersal rates) are needed to determine the extent to which this conclusion is robust.

In discussion, the Working Group thanked the authors for their prompt work in the short time since the intersessional workshop. There are now 'proof-of-concept' results from three methods (MIXPROP [SC/58/Rep6], GENELAND [Guillot *et al.*, 2005a; Guillot *et al.*, 2005b]; and sequential hypothesis testing [SC/58/SD1]) showing that TOSSM can indeed be run from start to finish, and which clearly lay out the steps needed to turn a genetic analysis into a testable management tool. There are of course many different choices that could be made about exactly how to turn any of these methods into a boundary-setting tool (α-levels, selection of number of populations, etc.), and more work will be needed to consider the effects of these different choices. With respect to the specific results in SC/58/SD1, it was emphasized that the dispersal rates in the TOSSM datasets examined in SC/58/SD1 are very low compared to many situations considered by the Scientific Committee; these particular sets of simulated data are intentionally 'easy', so that results cannot be extrapolated to more realistic scenarios.

SC/58/SD2 outlined progress made in simulation of datasets for TOSSM, including changes made to the life history matrices being used in the simulations and changes made to RMETASIM (Strand, 2002), the programme being used to run the simulations. During the TOSSM Workshop held in La Jolla, CA, in January of 2003, a total of 90 different population structure scenarios were chosen for simulation during Phase I of the TOSSM project (IWC, 2004), which are recapped in the paper. Prior to the meeting, 12 scenarios had been completed, most of which have carrying capacities of K=7,500 individuals. With current computing capabilities, completion of all scenarios listed in the paper is expected to take an additional 6 months of computer time. Suggestions were made for changes in some parameter values and a more focused approach to sensitivity tests in order to speed the completion of the simulations. The most important of these included limiting K=30,000 runs to specific scenarios where performance of methods was not yet clear, introducing K=2,500 runs for their utility in areas of more serious conservation concern, and reducing the number of mutation rates used for microsatellites to a single value, which produced a sufficient range in number of alleles to examine how allele number affects performance.

The Working Group noted the considerable resources required to undertake these simulations, and agreed that it would be appropriate to revise the specifications for the first generation of datasets. Specifically:

(1) the case K=30,000 is not crucial in the first instance;
(2) a single mutation rate will suffice for now;
(3) 200-500 replicates would suffice for now.

3.3 Directions for further work

The following methods and champions are proposed for the first round of investigations in TOSSM:

(a) MIXPROP (Kitakado)
(b) BAYESASS (Gaggiotti)
(c) GENELAND/STRUCTURE (Martien/Tiedemann)
(d) Sequential hypothesis testing (Punt)
(e) Boundary Rank (Martien)
(f) F_{ST}-based estimate of dispersal (Waples)
(g) LAMARC (Jackson)
(h) Pre-defined boundary (Bravington)

There is clearly considerable work to be done in refining the computational issues surrounding generation of the simulated datasets and in automating the boundary setting methods. The Working Group agreed to establish an intersessional Steering Group of Champions, consisting of the above plus Taylor and Daníelsdóttir, with terms of reference to oversee the development of simulations and testing of methods. Given the workload, the Working Group agreed that it would not be productive to have an intersessional workshop in the coming year but rather to review the results of simulations for the above methods at the 2007 Annual Meeting; at that time, it will be appropriate to consider whether there is a need for a further workshop. In addition, Bravington and Donovan agreed to produce a short worked example that would demonstrate the steps that might be followed in making management decisions similar to those considered by TOSSM in order to assist geneticists unfamiliar with an IWC management context.

This year a number of sub-committees have received papers in which STRUCTURE has been applied, and considerable discussion has ensued about how to interpret the results in an IWC context. It is therefore particularly timely to use TOSSM to investigate STRUCTURE. In fact, as noted under Item 3.1, this has already begun, with an adaption of the GENELAND package into a boundary-setting algorithm. GENELAND is very similar to STRUCTURE with the primary exception that GENELAND can make explicit use of spatial location data, constraining the clustering so that the resultant clusters are geographically contiguous. The geographic constraint can be turned off, in which case GENELAND implements STRUCTURE. Before the 2007 meeting, it is hoped that GENELAND (and its STRUCTURE) can be applied to some of the more challenging TOSSM archetypes.

The Working Group noted the major human resources needed for TOSSM dataset generation work, and for testing of methods such as STRUCTURE. Although TOSSM advanced quickly in the eighteen months following its first workshop, and notwithstanding the progress that has been made since in the overall project, progress with the simulation side of TOSSM has been slow in the last two years. The main reason has been a lack of personnel, exacerbated by unforeseen technical difficulties with RMETASIM, the model being used to generate the datasets. Though many of these difficulties have been overcome, it is clear that it will be necessary to hire someone to work full-

time on TOSSM if the project is to function in a timely and efficient manner. Recalling that the TOSSM project addresses one of the hardest scientific problems in IWC, and one which cuts across many sub-committees, the Working Group therefore strongly endorsed the proposal in Appendix 2, which seeks matching funds for one year to hire a computing technician. It was noted that governments with an interest in any cetacean population where stock structure uncertainty is an issue for management, may wish to contribute towards the funding.

4. WORK PLAN

Most of the issues to be addressed intersessionally in TOSSM are described in this report and in the TOSSM Workshop draft report. In addition, it was noted that the TOSSM simulated datasets, with both mtDNA and microsatellites, provide the base data that could be used to address issues of data quality. Several sub-committees face decisions on whether certain markers are of sufficient quality to be used in analyses that may affect management. TOSSM datasets can be sampled, and corrupted, with the exact sample sizes of strata for specific cases, to assess the power of available programmes for checking data quality (e.g. Microchecker) to detect faulty markers that would influence inferences about structure. Submission of papers on this topic for the 2007 meeting was encouraged.

Provisional items for the 2007 Agenda are:

(1) statistical and genetic issues relating to stock definition;
(2) review progress with TOSSM (in particular with respect to STRUCTURE);
(3) genetic data quality;
(4) potential definitions of 'unit-to-conserve' and their implications for management.

5. ADOPTION OF REPORT

The report was adopted at 12:42pm on Friday 2 May 2006.

REFERENCES

Beerli, P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22(3):341-45.

Beerli, P. and Felsenstein, J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in subpopulations by using a coalescent approach. *Proc. Natl Acad. Sci. USA* 98(8):4563-68.

Guillot, G., Estoup, A., Mortier, F. and Cosson, J. 2005a. A spatial statistical model for landscape genetics. *Genetics* 170:1261-80.

Guillot, G., Mortier, F. and Estoup, A. 2005b. GENELAND: a computer package for landscape genetics. *Mol. Ecol. Notes* 5:945-59.

Hey, J. and Nielsen, R. 2004. Multilocus methods for estimating population sizes, migration rates, and divergence time, with applications to the divergence of *Drosophila pseudobscura* and *D. persimilis*. *Genetics* 167:747-60.

International Whaling Commission. 2004. Report of the Workshop to Design Simulation-based Performance Tests for Evaluating Methods Used to Infer Population Structure from Genetic Data, 21-24 January 2003, La Jolla, USA. *J. Cetacean Res. Manage. (Suppl.)* 6:469-85.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-59.

Strand, A.E. 2002. METASIM 1.0: an individual-based environment for simulating population genetics of complex population dynamics. *Mol. Ecol. Notes* 3:373-76.

Waples, R.S. 1989. Temporal variation in allele frequencies: testing the right hypothesis. *Evolution* 43(6):1236-51.

Appendix 1

AGENDA

1. Introductory Items
   - 1.1 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
3. TOSSM (Testing of Spatial Structure Models)
   - 3.1 Report from intersessional Workshop
   - 3.2 Update on progress since the Workshop
   - 3.3 Directions for further work
4. Work plan
5. Adoption of Report

Appendix 2

FUNDING PROPOSAL FOR TOSSM

Karen Martien

An adequate understanding of population structure is frequently a crucial element in enabling informed conservation and management decisions. Advancing our ability to detect population structure using genetic data has been identified as a research priority for both the Scientific Committee of the International Whaling Commission and the US National Marine Fisheries Service (NMFS). Though a number of analytical techniques are available that use genetic data to address population structure questions, none have been thoroughly tested to quantify their performance in a conservation and management context. To address this lack of reliable performance tests, the IWC Scientific Committee and NMFS jointly initiated the Testing of Spatial Structure Methods (TOSSM) project, the purpose of which is to design and conduct a large, iterative performance-testing study. The project was officially launched in January of 2003 with a workshop that brought together experts in whale behaviour and biology, genetic analysis, genetic modelling and simulation-based performance testing for the purpose of laying out the logical foundation for the project. Funding for the workshop was secured from both NMFS and the IWC.

Though the project advanced quickly in the eighteen months following that meeting, progress in the last two years has been slow. The primary reason for the slow progress has been a lack of personnel, exacerbated by unforeseen technical difficulties with RMETASIM, the model being used to generate the datasets. Though many of these difficulties have been overcome, it is clear that it will be necessary to hire someone to work full-time on TOSSM if the project is to progress in a timely and efficient manner. We therefore request funding to hire a technician to assist with the TOSSM project. The specific duties of the person would include:

(1) Collaborate with Allan Strand, RMETASIM's creator, to make the necessary revisions so that the package can be run reliably on Windows-based computers.
(2) Finish generating all datasets specified by the TOSSM steering committee.
(3) Complete and maintain the R package 'TOSSM' with which performance tests will be conducted. This includes providing all documentation necessary for publishing the package on the Comprehensive R Archive Network website.
(4) Assist in adapting existing analytical methods to work with the TOSSM package.

To complete these tasks, we propose hiring a full-time specialist programmer for one year, at a cost of $60,000US. The person hired for this position must possess strong computer and programming skills and have experience writing complex programs in R and in C or C++. This person will be hired through the contracting agent used by Southwest Fisheries Science Center and will be supervised by Dr Karen Martien. We have requested half of the funding for this position from NMFS, and request that the other half be provided by the IWC.

Budget request: US $30,000.

Time line: One year from start of funding.