

**Report of the Second  
Workshop on the  
Development of an  
Aboriginal Subsistence  
Whaling Management  
Procedure**

## REPORT OF THE SECOND WORKSHOP: CONTENTS

1. Introductory items .....	417
2. Review of progress on pre-workshop tasks identified last meeting.....	417
2.1 Secretariat computing.....	417
2.2 Continued development of <i>Strike Limit Algorithms (SLAs)</i> .....	417
3. Stock structure.....	417
3.1 Bowhead whales.....	417
3.2 Gray whales.....	418
4. Implementation trial structure for bowhead whales .....	419
4.1 Classes of trials, their identification and relationships between them.....	419
4.2 Review of Annex E (IWC, 2000b), appendix 3B, sections A-C .....	420
4.3 Principles of trial design .....	420
4.4 Specification of trials .....	420
4.5 <i>SLA</i> tuning.....	425
4.6 Review of performance statistics.....	425
4.7 <i>SLA</i> optimisation .....	426
5. Implementation trial structure for gray whales.....	426
5.1 Simulation scenarios .....	426
5.2 Review of SC/N99/AWMP1 sections A-C.....	426
5.3 Specification of trials .....	426
5.4 Performance statistics.....	427
6. Other matters .....	427
6.1 Work plan and priority tasks.....	427
7. Adoption of report.....	427

# Report of the Second Workshop on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP)

**Members:** Donovan (Chair), Allison, Breiwick, Butterworth, DeMaster, Givens, Huang, Magnússon, Moore, Punt, Rugh, Schweder, Zeh.

## 1. INTRODUCTORY ITEMS

### 1.1 Convenor's opening remarks

The Workshop was held in Seattle, Washington in November 1999. Donovan welcomed the participants. He reminded the members that the main purpose of the meeting was to finalise the single stock implementation trials for the Bering-Chukchi-Beaufort Seas stock of bowhead whales and to initiate development of similar trials for the eastern North Pacific stock of gray whales. He emphasised that the Commission had endorsed the Standing Working Group's (SWG's) work plan at the last meeting and that it wished the SWG to keep to the 'fast track' work plan option to the greatest extent possible.

### 1.2 Election of Chair

Donovan was elected Chair.

### 1.3 Appointment of rapporteurs

Allison, Givens and Punt acted as rapporteurs, with assistance from the Chair.

### 1.4 Adoption of agenda

The adopted agenda is given as Annex A.

### 1.5 Review of documents

The documents available to the SWG were SC/N99/AWMP1-6 and these are listed in Annex B. The glossary of terms is given as Annex C and full details of the latest trial scenarios, statistics and assumptions as Annex D.

For ease of reading, unless another reference is given, 'Last meeting' refers to Annex E (IWC, 2000b), the report of the Standing Working Group on the Development of the AWMP during the 1999 meeting of the Scientific Committee.

## 2. REVIEW OF PROGRESS ON PRE-WORKSHOP TASKS IDENTIFIED LAST MEETING

### 2.1 Secretariat computing

Allison reported that she had completed tasks 1-4 as listed in table 4 of Annex E (IWC, 2000b). The trials based on the deterministic and stochastic population dynamics models have been implemented and routines for calculating performance statistics coded. The input files for each of the *Initial Exploration Trials* were produced and parameter files for each trial created by applying the conditioning option of

the simulation software. All of this software had been distributed by e-mail during the intersessional period. In addition, Givens updated the software archive on the AWMP SWG web page at <http://www.stat.colostate.edu/~geof/iwcawmp.html>. Zeh reported that she had completed a preliminary version of an *S-plus* program to plot the performance statistics produced by the simulation software. The SWG thanked Allison, Zeh and Givens for their work on these matters.

### 2.2 Continued development of Strike Limit Algorithms (SLAs)

SC/N99/AWMP3 reported on the application of variants of two families of *SLA* to the 18 *Initial Exploration Trials* for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. None of the variants considered were able to adequately distinguish between trials in which  $MSYR = 1\%$  and in which  $MSYR = 4\%$  because the data available for assessment purposes differed systematically from the true underlying population trajectory when  $MSYR$  differed substantially from about 2.5%. H-optimisation was implemented by defining the strike limit as a weighted sum of covariates where each weight is constrained to be non-negative. The covariates considered included strike limits from two catch control laws and two data downweighting schemes. The SWG thanked Punt for his work in this regard and strongly encourages the submission of other candidate *SLAs*.

## 3. STOCK STRUCTURE

In discussing this item, it was stressed that consideration should focus on issues related to the specific fisheries in question, rather than a broad discussion of stock structure. Consideration began with a discussion of situations that might cause particular problems for *SLAs*. As a general rule when there is mixing between stocks, there are few problems if the catch is taken in proportion to the abundance of each stock in the area. Problems arise if segregation (spatial or temporal) leads to surveys being performed in an area containing two stocks but where the catch is (in the extreme) taken from one stock only. This is equivalent to a constant absolute rather than a relative bias. The inverse case is that of harvesting a stock which is found outside the survey area. This may lead to under utilisation but is not a problem in terms of risk.

### 3.1 Bowhead whales

The issue of possible sub-stock structure in the Bering-Chukchi-Beaufort stock of bowhead whales had been raised at last year's Scientific Committee meeting (IWC, 2000b).

In this context, information on the historic catches of bowhead whales of the Bering, Chukchi and Beaufort Seas was considered. Table 1 of Annex D (IWC, 2000a) lists the annual catches. Bockstoe and Botkin (1980) reviewed the available information on commercial catches from 1848-1914. Very few catches of bowheads were taken prior to 1848 when the whaling grounds near the Bering Strait were discovered by the Yankee whalers.

The initial catches were very high and the bowhead population in the Bering Strait area quickly reduced; it is also possible that in addition to over-exploitation, the whales responded to catching by leaving the hunting grounds. By 1853, catches had become so poor that, between 1855 and 57, the fleet virtually abandoned the Bering Strait region and moved into the Okhotsk Sea. This area too was soon overhunted and, in 1858, the whalers returned to the northern Bering Strait where they took regular catches for the next 50 years. This more northerly hunting became possible with the advent of steam-auxiliary whaling vessels in around 1880, which enabled whalers to penetrate further into the ice. In 1889, steamers reached the bowheads' summer feeding grounds off the MacKenzie River delta, and from then until 1914, the focus of the industry was concentrated largely on those waters.

Commercial catches ceased around 1914, after which only aboriginal catches have been taken.

In the present Alaskan fishery, the largest catches are taken at Barrow. Other significant catches are taken at Wainwright and Point Hope, with occasional catches being taken off St Lawrence Island in the Bering Strait. Braham (1995) presents catch data by village for the Alaskan fishery from 1973-1992.

Rugh summarised present knowledge on the distribution and migration of bowhead whales. They winter in the Bering Sea and in spring migrate through the Bering Strait. Most move east to Barrow and summer in the region between Barrow and the MacKenzie River. In the autumn they move back to the west, some moving past Wrangle Island and then back round along the coast of Chukotka. It is also possible that in spring some whales move to the west after leaving the Bering Strait and spend the summer in the Chukchi Sea. There are recent observations of whales in the Chukchi Sea in the summer.

The bowhead census occurs as whales migrate past Point Barrow in the spring. They migrate through a relatively narrow (maximum *ca* 15 miles) 'passage' through the ice and the census occurs at the same time and in the same area as the spring hunt. During the autumn migration (when hunting also occurs) there is considerably more open water and the concept of 'gauntlet hunting' sometimes applied to the spring hunt is less applicable.

The importance of the catch distribution was considered, for example in the hypothetical case of a heavily depleted stock mixing with a healthy stock. This could be the situation if two stocks are mixing in the Bering Strait and then separating, with the bigger stock going east. This would mean that catches off St Lawrence Island might include animals from the smaller stock, which is undesirable. However, even if this represented a true scenario, it was agreed that, given the low catches off St Lawrence Island (approx. seven whales), this would not be a major issue.

Treating the Bering-Chukchi-Beaufort bowheads as a single stock could be problematic if in fact different stocks pass Barrow at different times and so are not taken in proportion to their relative abundances. However, it is known that young whales generally migrate past Barrow before older whales, which is difficult to reconcile with such

a multiple stock hypothesis. The SWG requested that a map showing catches by sex and size landed by village be compiled for the next Scientific Committee meeting.

There is evidence from harpoons that suggests that, in summer in very low ice years, there may be some movements from this stock into the Davis Strait. Thus, the possibility that, as the Bering-Chukchi-Beaufort stock recovers, it might help to repopulate other depleted stocks of bowhead whales cannot be ruled out. However, it was noted that any catches allowed under an AWMP would still allow the Bering-Chukchi-Beaufort stock to increase and thus, even if re-population is possible, catches would slow down rather than halt such a process.

Therefore, the SWG agreed that consideration of other than a single stock scenario was unnecessary for the Bering-Chukchi-Beaufort bowhead whale. However, this issue may be reconsidered if new evidence becomes available. DeMaster indicated that a comprehensive review of the stock identity of all bowhead whale stocks would be submitted to the next Scientific Committee meeting.

### 3.2 Gray whales

DeMaster gave an overview of the distribution of gray whales in the North Pacific. There are two stocks, which are distinct both on their breeding grounds and on their feeding grounds. The western stock breeds in the South China Sea and feeds in the Okhotsk Sea while the eastern stock breeds in the lagoons of Baja and feeds in the area from Northern California to the Chukchi Peninsula. Feeding gray whales are only found in shallow waters and their summer distribution largely reflects the position of the coastal shelf. Studies of parasites and morphology have not proved useful in stock differentiation. Significant differences in the  $F_{ST}$  statistic have been found between the two stocks although the small sample size ( $n=9$ ) from the western region may be too small to characterise the genetic diversity of the stock.

There is a long history of whaling of gray whales on both sides of the Pacific (see table 1 of SC/52/AWMP1) but the catch history for this species is very poorly documented. Commercial whaling ceased around 1946.

In the present aboriginal fishery, the Russians harvest along the coast of the Chukchi Peninsula. The gray whale population has been increasing since 1967 in the presence of catches of up to 180 per year. The nature of the Russian aboriginal catch has changed since 1994. Catches prior to 1992 were taken by a modern catcher vessel on behalf of the native people. Catches are now taken using small boats and traditional methods. The absence of a catch in 1992 and 1993 was not due to an absence of whales, but an absence of successful effort. One whale was also taken off the northern coast of the USA in 1999 by the Makah Indians.

The SWG agreed that the scenario of a single eastern stock was in accordance with the data and that single stock trials without spatial sub-structure should be started immediately. These trials are developed under Item 5.

However, the SWG agreed that there is some evidence that might suggest the existence of sub-stocks which exhibit feeding site fidelity (photo-identification data for some areas is not conclusive but suggests some site fidelity). There is no evidence for discrete feeding areas that are important every year. In addition, there is also some evidence for segregation by sex; catch data from the Chukchi Peninsula show significantly more females close to shore (Bogoslovskaya, 1986), but this may be confounded by selectivity for larger animals by whalers, particularly prior to 1992.

The SWG agreed that development of an exploratory multi sub-stock model should begin in parallel with single stock trials. Levels of mixing rate parameters consistent with the increasing survey estimates should be investigated. Only segregation in feeding areas should be considered, as there is no evidence for breeding segregation. This scenario is similar to that used for the North Pacific Bryde's whale trials (IWC, 2000a).

The SWG discussed information that might be collected to help to discriminate between different hypotheses and/or to set bounds on model parameters. Genetic samples for individual identification would be useful and the SWG agreed to encourage collection of samples. Funding by the IWC might facilitate this.

DeMaster indicated that a comprehensive review of the stock identity of gray whale stocks would be submitted to the next Scientific Committee meeting.

#### 4. IMPLEMENTATION TRIAL STRUCTURE FOR BOWHEAD WHALES

Last year, the SWG identified trial classes for post-*Initial Exploration Trials*: *Evaluation Trials*, *Robustness Trials* and *Cross-validation Trials*. *Evaluation Trials* would be used for formal comparison of candidate SLAs, limited in number compared to *Robustness Trials* and initiated prior to them. *Robustness Trials* examine SLA performance for a full range of plausible scenarios and would be applied to a restricted set of SLAs found to perform well in *Evaluation trials*. *Cross-validation Trials* would be case-specific trials held aside from SLA developers so that resulting SLAs can be subjected to a subsequent independent test.

##### 4.1 Classes of trials, their identification and relationships between them

SC/N99/AWMP5 proposed that certain relationships should hold between the regions of parameter space spanned by different classes of trials. Specifically, *Evaluation Trials* should resemble *Initial Exploration Trials* though perhaps be more diverse. *Robustness Trials* should be the broadest class of trials. The space of *Evaluation Trials* should be wholly contained within the broader *Robustness Trial* space and should consist of a balanced set of realistic scenarios. *Cross-validation Trials* should be trials which fall within the space of *Evaluation Trials* but not within the larger *Robustness Trials* space because they are intended to be used to confirm the comparative evaluation and selection of SLAs that is based on *Evaluation Trial* performance.

The general issue of the relationship between *Robustness Trials* and *Evaluation Trials* was discussed, as well as the criteria for including a trial in one class or the other. Issues considered included: (a) whether the primary distinguishing factor between *Evaluation Trials* and *Robustness Trials* should be the level of plausibility; and (b) whether trials which led to a significant change in stock status, or ideal management, should be used for SLA evaluation even if they were not among the most plausible. The SWG agreed that none of the trial classes, including *Robustness Trials*, should include any implausible scenarios. It reiterated that the function of *Evaluation Trials* is to allow comparison of candidate SLAs. In this context such trials should be designed to provide relevant information for comparison, and therefore not restricted to the most plausible trials only.

The SWG reconfirmed that *Robustness Trials* should be used to examine SLA performance for a full range of plausible scenarios in order to identify whether it behaves as

expected, and not necessarily to see whether the SLA provides satisfactory strike limits. Although *Evaluation Trials* are the primary method for comparing candidate SLAs, performance on the *Robustness Trials* might be used in a secondary role to distinguish between SLAs which performed similarly well on *Evaluation Trials*.

The SWG then discussed the nature and role of *Cross-validation Trials*. Although all SLAs will be subjected to this independent test, the original motivation for such trials were concerns that an H-optimised SLA might reflect over-fitting. The group further noted two types of such concerns: (1) unpredictable behaviour in the interior of a tested region of parameter space due to over-fitting; and (2) unpredictable behaviour somewhat outside a tested region of parameter space due to extrapolation. Concerns about the latter behaviour led to discussion as to whether *Cross-validation Trials* should be allowed to fall in either the space of *Evaluation Trials* or of *Robustness Trials*. The authors of SC/N99/AWMP5 had argued that as *Cross-validation Trials* were intended to cross-validate SLA comparison and selection, such comparison should primarily occur in the *Evaluation Trial* space (i.e. *Cross-validation Trials* should be limited to the *Evaluation Trial* space but be different from any other trials in that space). The SWG agreed that: (1) *Cross-validation Trials* within the *Evaluation Trial* space are highly desirable; and (2) each *Robustness Trial* also serves as a *Cross-validation Trial* in the broader sense desired by some members. Overall, any candidate SLA would be subjected to a broad spectrum of trials whose results might be interpreted for different purposes by different SWG members.

The SWG agreed to a two-pronged approach to cross-validation: some such trials should be designed during the 2000 Scientific Committee meeting, and some trials developed at the Seattle workshop would be designated in Adelaide as being useful for cross-validation. For the first approach, the SWG believed that such trials should not be more extreme than *Evaluation Trials* and could likely be developed quickly by choosing values of several key parameters such as  $MSYR_{1+}$  and Final Need that were internal to the ranges spanned by *Evaluation Trials*. For the second approach, the SWG planned to examine the results of the analyses Huang planned to complete for the 2000 meeting (see below) in order to identify existing *Robustness Trials* that are shown to be inside or nearby the scenario space spanned by *Evaluation Trials*.

The SWG reiterated the importance of the concept implicit in *Cross-validation Trials* of holding aside some trials from developers. In discussion, the SWG agreed that all but a small subset of the *Robustness Trials* will be withheld from the developers. The Secretariat will run all the *Robustness Trials* and thus, to a large extent, these will also function as *Cross-validation Trials*. However, it was agreed that, pending the results of the visualisation of scenario space analysis discussed below, additional *Cross-validation Trials* may be specified at the next meeting.

Another major topic of SC/N99/AWMP5 was the identification of trial class boundaries or outlying scenarios. All trials can be uniquely identified by the value of a high-dimensional parameter vector. The authors noted that the identification of scenario space boundaries or outlying trials would be extremely difficult given the high-dimensionality of the parameter space; therefore they proposed assessing trials on the basis of the shapes of population trajectories produced by the trial. SC/N99/AWMP5 proposed five statistics to measure trajectory shape, including measures of population

growth/decline, oscillation, absolute size and within-scenario uncertainty. These statistics were assessed under several different catch regimes. The authors considered that such statistics could be used as part of an iterative process to construct and delimit a scenario space: after first measuring the apparent boundaries of a space based on a smaller collection of key trials, additional proposed members of the same trial space could be assessed to confirm that they exhibited trajectory behaviour that was not atypical of that exhibited by existing members. Such statistics can also be used to identify outlying members of a collection of trials by examining a pairwise scatterplot of the five statistics for all trials. Trials with unusual characteristics or an unusual combination of characteristics could be identified from either approach for possible elimination, reconsideration or alteration. Such methods are useful for ensuring that desired relationships between trial classes hold, and for ensuring that unrealistic trials are not specified unintentionally by the SWG.

The SWG welcomed the ideas and approaches outlined in SC/N99/AWMP5, recognising the importance of the problems addressed. It noted that the proposed statistics related to within-scenario variation might need to be revised in consideration of stochastic model simulations. A feature of the methods in SC/N99/AWMP5 is that the definitions of outlying or unusual scenarios do not depend on any candidate SLAs under consideration by the SWG even though presumed catch regimes are needed in order to compute the statistics. This had been achieved by assuming three catch regimes: zero catch, constant catch of 100 and 'Potential Biological Removal' (Givens, 1998). The SWG recommends that Huang repeats the example analyses in SC/N99/AWMP5 for the full set of trials (and each class) established at this meeting, so that the results can be examined at the 2000 meeting of the Scientific Committee.

The SWG further agreed that it would be important to examine the stock trajectories resulting from all trials in order to confirm that these trajectories were not inconsistent with the available data; for the *Evaluation Trials*, informal inspection of trajectory plots will suffice. Statistical approaches for detecting or quantifying any mismatch will be discussed during the next meeting.

## 4.2 Review of Annex E (IWC, 2000b), appendix 3B, sections A-C

The SWG reviewed Annex E, appendix 3B, sections A-C from last year's report and agreed that the population projections for the stochastic model would start in 1748. The revised trial specifications are listed in Annex D. A glossary is also included to detail the symbols used in the trial specifications.

## 4.3 Principles of trial design

SC/N99/AWMP4 contained a large collection of proposed trials whose development adhered to several principles that the authors believed necessary for fair and effective AWMP testing. First, the set of *Evaluation Trials* should not be strongly tilted towards either pessimistic or optimistic scenarios (with respect to stock conservation). The SWG agreed that this was a reasonable goal, noting that it will ensure some degree of balance without attempting to define, or require any strict adherence to, 'perfect balance'. It also agreed that balance was not as great a concern in *Robustness Trials*.

SC/N99/AWMP4 suggested that *Evaluation Trials* and *Robustness Trials* emphasise uncertainty in model structure. They believed that it is likely that several models with different dynamical structures may fit observed data reasonably well, yet lead to different future predictions in some cases. They noted that the available data have been used once to develop the 'conventional wisdom' about model structure, then again to assess uncertainty of estimation and prediction given that model structure. They believed that this double-use of the data tends to understate the actual uncertainty inherent in prediction and decision making; ambiguity about model structure should be allowed to dilute prediction precision. They felt that the population dynamics models generally used by the SWG represent only a narrow and simplified view of how whale dynamics may actually occur, and the AWMP will be chosen on the basis of computer simulations using these models, with absolutely no real-world empirical testing. They noted that excessive faith in a popular model is risky; it is safer to make decisions after reviewing several fundamentally different sets of assumptions.

There was considerable discussion of this issue within the SWG and this is reflected under Item 4.4.2.

In discussion, the SWG also emphasised that the development of trial scenarios must be based on the premise that the available data are fixed external information. Scenario diversity should be created by choosing models that interpret the existing data differently and not by constructing trials that change the data themselves.

## 4.4 Specification of trials

### 4.4.1 Conditioning

The BALEEN II population model (Punt, 1999a) underlies the simulation trials. This model has seven parameters (see Table 1). Conditioning a trial involves selecting values for these parameters that are as consistent with available data as possible, given the specifications of the trial. A Bayesian approach is taken when assessing the status of the Bering-Chukchi-Beaufort bowhead stock. The conditioning of trials follows the same logic and uses the same data and prior distributions, unless specified otherwise. The output of the Bayesian assessment is a joint posterior distribution for the free parameters of the model. In practice, the joint posterior distribution is represented numerically by a random sample from the posterior distribution. A conditioned trial is thus represented by a sample of 100 parameter vectors drawn from the (conditional) posterior distribution of the trial given its specifications.

Table 1

Summary of parameters used in the BALEEN II population model.

Parameter	Notes
MSYL	Can be defined on mature, exploitable or 1+
MSYR	Can be defined on mature, exploitable or 1+
Density dependent component	Mature, exploitable or 1+
Female age at maturity	Defined by 50% and 95% values
Age at recruitment	Defined by 50% and 95% values
Mortality rate	(Age specific)
Carrying Capacity	

When a trial specifies that a subset of the parameters have specific values, the conditioning results in a joint posterior distribution for the remaining free parameters. If, for example, a trial sets values for  $MSYR_{1+}$  and  $MSYL_{1+}$ , the

free parameters are  $K$ ,  $a_m$ ,  $a_{adult}$ ,  $S_{adult}$  and  $c_{min}$ . When the structure of the model is changed (e.g. time-trends are assumed for historic survey bias), conditioning is essentially the same as conducting an 'assessment' based on the changed model. Similarly, if the prior distribution is changed or stochastic dynamics are allowed for, this also essentially results in an alternative 'assessment'. Conditioning a trial is a computer intensive operation.

The SWG noted that the conditioning of trials for *SLA* simulation testing was based on the most recent bowhead assessment. If significant new data became available and a new assessment was done, the conditioning of the trials might also need to be revised. It agreed that such action could be addressed in the broader context of an Aboriginal Whaling Management Scheme through Implementation Reviews.

#### 4.4.2 Selection of factors to consider

The SWG considered the trials and factors proposed in SC/N99/AWMP2 and SC/N99/AWMP4 and used these to identify the range of factors (and their levels) that should be included in the trials (Table 2). For each factor, one or more reference levels were selected. The reference level indicates the most important value, to be used when crossing factors.

The SWG agreed that the length of each simulation would remain at 100 years, noting that visual examination of plots of population trajectories (e.g. Annex E) could be used to identify undesirable *SLA* behaviour over the longer term. Such behaviour might be more likely with strongly serially

dependent dynamics or trends in parameters such as  $K$ . The SWG agreed that it was not necessary to consider the case in which the abundance estimates include some calves because the bias arising from this would be small relative to that for the factors that explicitly consider survey bias. It was also agreed that the level of temporal correlation in the estimates of abundance (see table 3 of Annex D) was sufficiently small not to impact performance noticeably. The implications of density-dependent natural mortality are not considered in the trials. This is because Punt (1999b) found little difference in the results of trials assuming density-dependence acts on juvenile natural mortality and those assuming that it acts on fecundity. It was agreed not to change the priors used when conditioning the trials because it is unlikely that reasonable changes to the priors will impact performance noticeably. The SWG agreed that survey estimates will be generated based on the 1+ population size four years before the next strike limit comes into effect, i.e. if the strike limit is to be set for year  $t$ , the abundance estimate pertains to year  $t-4$ . This largely reflects the need to consider the 'strategic' survey issue discussed below.

The SWG discussed whether it was more appropriate to consider a range of fixed values for  $MSYR_{1+}$  or to conduct trials in which a prior is placed on  $MSYR_{1+}$ , and  $MSYR_{1+}$  is then treated as a 'free' parameter in the conditioning. The latter approach is more consistent with recent Bayesian assessments of the Bering-Chukchi-Beaufort bowhead stock whereas the former makes interpretation of results easier, as  $MSYR_{1+}$  can have a major impact on performance. For example, the lower 5th percentile of the lowest depletion

Table 2  
Factors to be included in bowhead trials.

Symbol	Varying Factors	Other Levels (Reference levels shown <b><u>bold and underlined</u></b> )
	$MSYR_{1+}$	1%, <b><u>2.5%</u></b> , 4%
	<b>MSYL</b>	0.4, <b><u>0.6</u></b> , 0.8
A	Density dependent component	<b><u>1+</u></b> or mature
B	Population dynamics model	<b><u>Deterministic</u></b> , stochastic (demographic), <b><u>Stochastic (demographic + serially corr.env.var.)</u></b>
C	Need in final year	<b><u>67</u></b> , 134, <b><u>201</u></b>
D	Form of need time dependence	<b><u>Linear</u></b> , step function (changing from 67 to final need level in year 2053)
E	Survey frequency	<b><u>5yr</u></b> , 10 yr, 15 yr
F	Strategic surveys	<b><u>No</u></b> , Yes
	Survey bias i) constant	0.67, <b><u>1.0</u></b> , 1.5
G1	ii) historical time dependence	<b><u>None</u></b> , increase by 50% from 1978 to 2002 (0.67→1, 1→1.5)
G2	iii) future time dependence	<b><u>None</u></b> , decrease by 50% from year 0 to year 100, increase by 50% from year 0 to year 100, sinusoidal see Fig 1
H	Survey CV: (true, estimated)	(0.1, 0.1), (0.25, 0.1); ( <b><u>0.25</u></b> , <b><u>0.25</u></b> ); (0.34, 0.25)
	Age data: bias and CV	<b><u>None</u></b> , new version (Geof)
I	Historic (1848-1914) catch bias	0.5, <b><u>1.0</u></b> , 1.5
J	Carryover strategy	<b><u>Good (no bias or extra variance: <math>\gamma_{12.9}=0</math>, <math>\lambda_{12.9}=1</math>)</u></b> , Poor ( $\gamma_{12.9}=2$ , $\lambda_{12.9}=1.5$ )
K	Time dependence in $K$ *	<b><u>Constant</u></b> , halve linearly over 100yr, double linearly over 100yr, Sinusoidal from a base value in year 0 increasing to a maximum of 150% in year 40.
L	Time dependence in resilience *	<b><u>Constant</u></b> , halve linearly over 100yr (initially 4% / 2.5% $MSYR$ only), double linearly over 100yr (initially 1% / 2.5% $MSYR$ only), Step $MSYR$ 2½%→1%→2½% every 33 yrs, alone or in sync with halving / doubling $M$
M	Time dependence in natural mortality, $M$ *	<b><u>Constant</u></b> , halve linearly over 100yr, double linearly over 100yr.
N	Episodic events *	<b><u>None</u></b> , 1% probability each year that 50% of animals die
O	Integrated	<b><u>NA</u></b> , priors for $MSYR$ , $MSYL$ , historical catch bias, extent of environmental variation in the probability of birth.
P	Initial year of population projection	<b><u>1848</u></b> , 1940
Factors held fixed		
Length of simulation		
Priors for conditioning: $S_{adult}$		
Abundance estimates		
(i) contain calves?		
(ii) correlation in estimates		
(iii) delays in using results		
Density-dependence in $M$		

\* Effects of factors K, L, M and N begin in year 2003 (i.e. at start of management).

distribution for an 'integrative' trial may reflect simply the 5th lowest value generated for  $MSYR_{1+}$  rather than a case in which the *SLA* failed to achieve conservation goals to some extent. The posterior mean for  $MSYR_{1+}$  from the 1998 assessment (IWC, 1999a) is close to 2½% while the 1 and 4% approximate upper credibility limits for  $MSYR_{1+}$ . The SWG agreed to examine fixed values for  $MSYR_{1+}$  in most trials (1%, 2½% and 4%) but to use a posterior resulting from a  $U[1\%, 4\%]$  prior on  $MSYR_{1+}$  in some of the robustness trials. In selecting this approach, it was recognised that the assumptions  $MSYR_{1+}=1\%$  or  $=4\%$  are not comparable with the trend in the existing abundance data unless allowance is made for a change in survey bias ( $MSYR_{1+}=1\%$ ) or a higher  $MSYL$  ( $MSYR_{1+}=4\%$ ).

The SWG agreed to consider values for  $MSYL$  of 0.4, 0.6 and 0.8. Trials will be conducted in which these levels refer to the 1+ and mature female components of the population. The component to which  $MSYL$  refers is always the same as that on which density-dependence acts for numerical stability reasons. Values for  $MSYL_{1+}$  correspond to lower values for  $MSYL_{mat}$ . The reference level, 0.6, is the value conventionally assumed by the Scientific Committee when conducting assessments of baleen whale stocks (e.g. Fowler and Baker, 1991). Fowler (1991) also supports the assumption that  $MSYL_{1+}$  is larger than 0.5. The upper value of 0.8 is based on attempting to avoid oscillatory population size trajectories while the lower value of 0.4 was selected because of the belief that  $MSYL$  for baleen whales is not likely to be lower than those for fish species, for which  $MSYL$  is conventionally assumed to lie between about 0.3 and 0.5.

SC/N99/AWMP4 proposed that two alternative population dynamics models be included in trials and outlined three variants of the stochastic population dynamics model. The SWG agreed to base trials on the deterministic population dynamics model and the two variants of the stochastic population dynamics model that are likely to exhibit the most (allowing for serially correlated environmental variation) and least (demographic stochasticity only) variation.

Three levels for the need in the final year of the 100-year simulation period are considered. The two reference levels reflect the upper bound of the need envelope (201) and the current level of need (67). The latter is included to examine whether the current level of need is sustainable across all scenarios. The trials include a final need level of 134 to test the assumption that if performance is satisfactory on the bounds of the need envelope that this is also the case within the need envelope. For the reference case, need increases linearly from 67 in 2003 to its final level (67, 134 or 201) in 2103. An alternative scenario involves increasing need from 67 in 2003 to its final level in a single step in 2053; the total need over the 100-year period for this scenario is the same as that for the reference case. These two scenarios should provide an adequate test of whether the shape of the need trajectory impacts overall performance.

A factor 'strategic' examines the impact of varying survey timing to attempt to maximise strike limits. In simple terms, one might wish to carry out another survey quickly (i.e. it was too low) or one might postpone a new survey as long as possible if the strike limit from the previous survey was satisfactory. Thus, additional surveys are undertaken if the strike limit based on the survey four years earlier is lower than that strike limit or if it is less than need in the case when need was satisfied over the previous five-year period. If the strike limit is greater than the previous strike limit when need was previously unsatisfied or if it equals need, the

inter-survey period is increased from 5 to 15 years. The trials that involve changing survey frequency should be conducted both with and without 'phase-out rules' to allow the impact of different survey frequency on performance to be determined without the confounding effect of 'phase-out rules'.

The SWG agreed that levels of survey bias of 0.67 and 1.50 should be included in the trials. It discussed factors that are likely to impact the magnitude of possible bias in the surveys at Point Barrow. It was agreed that there are factors that might lead to positive or negative bias although these could not be quantified. However, it was believed that a lower bound of 0.67 was more appropriate than the 0.5 lower bound adopted for the RMP trials. This is to achieve multiplicative symmetry and because the scope for negative bias is likely to be less for the shore-based surveys of the Bering-Chukchi-Beaufort stock than for stocks surveyed on their feeding grounds, as was the focus for the RMP trials. The upper limit for the level of bias was set equal to the value for the RMP trials because the key source of positive bias considered when developing the 1.5 bias for the RMP trials, namely errors in estimating  $g(0)$  is implicitly relevant in the bowhead case. However, some members believed that 1.33 was a more appropriate upper bound than 1.5 because of the differences in the nature of the surveys off Point Barrow from those typically used to provide abundance estimates based on line-transect surveys under the RMP. The SWG considered several scenarios for time-dependence in survey bias. It decided to treat the issue of the time-trends in survey bias for the period 1978-2002 separately from the issue of time-trends in survey bias from 2003 (Annex E Appendix 3b, fig. 1). The SWG agreed to consider two scenarios regarding historical survey bias (no increase and an increase of 50%) and four scenarios regarding future survey bias (none, increase by 50%, decrease by 50% and a scenario in which survey bias varies sinusoidally). Changes in survey bias are assumed to start in 2003, the first year for which a strike limit is set using the *SLA*. The scenarios regarding time trends in survey bias are not meant to reflect changes in survey methodology alone but also changes in, for example, whale behaviour (possibly in response to changes in oceanographic conditions).

Four combinations for the true and estimated coefficients of variation for the estimates of abundance (based on a population size of 0.6K) are considered. The reference case examines an estimated CV of 0.25 for consistency with the most recent estimates of abundance (table 3 of Annex D) combined with the assumption of no 'additional variance'. This last assumption is based on the analysis of Zeh *et al.* (1995). Only one scenario for error in the age-composition data is considered because, to date, no *SLA* developer has made use of these data. If this changes, it may be necessary to develop trials based on additional assumptions regarding errors in these data. The particular scenario regarding error in the age-composition data differs from that for the *Initial Exploration Trials* because it assumes information about maturity-at-age is both in error and assumed (incorrectly) to be relatively precise.

Substantial errors in the historical (1848-1914) catch data are possible owing, *inter alia*, to the fact that documentation of catches is available for only a small fraction of logbooks for some years of this period. Halving and doubling the historical catches should provide an adequate test of robustness to this source of error.

The SWG considered how to evaluate potential carryover rules. It was agreed that carryover rules, such as that currently in force, that allow some carryover within 5-year



blocks (but not between 5-year blocks) have no adverse biological implications. The SWG was unable to develop trials to test carryover rules that allow the carryover of some strikes between blocks because of the wide variety of possible carryover rules. It recommended that Donovan draft a document outlining some possible carryover rules for presentation to the Commission. This document would be reviewed by the SWG at its next meeting. It could include the concept of a carryover envelope - the range of possible levels of carryover - for incorporation in future trials.

Allowing the carrying capacity,  $K$ , the recruitment rate (through the value of the resilience parameter  $A$ ) and natural mortality to vary with time allows both the impacts of changes over time in environmental conditions through forcing of the population dynamics and different population dynamics processes to be examined. The impact of these effects is first felt in year 0. The results for trials based on time-varying  $K$  and natural mortality may be difficult to interpret, as  $K$  is a key parameter used when defining the performance statistics. The SWG therefore agreed that the results for these trials should be scaled to the population trajectory for the case in which all future strike limits are set to zero. A similar approach was used for the RMP trials in which  $K$  varied over time (IWC, 1992). Trials are conducted in which the value of  $K$ ,  $A$  and survival rate change linearly over time and in which  $MSYR$  changes from 1% to 2½% every 33 years. A trial that involves  $A$  and natural mortality changing in 33 year steps/halving to doubling in sync is conducted to examine the impact of environmental variation (for example in ice coverage), which should effect both birth and death rates. An episodic events factor examines the implications of a 50% reduction in population size, which occurs with 1% probability each year. The natural mortality rate will have to be increased for trials in which episodic events can occur so that the population will remain at its carrying capacity on average in the absence of exploitation.

SC/N99/AWMP4 highlighted the importance of uncertainty in model structure (see Item 4.3) and argued that it was important that a wide variety of population dynamics models be considered in the trials. The SWG noted that all of the existing trials are based on a single underlying population dynamics equation. For illustrative purposes, Givens presented the results of two quite different tentative population dynamics models from that used in the simulation code; however, these included birth-death processes, constrained the birth rate to plausible values, mimicked the recent rate of increase and matched the current abundance. Some members believed that changing the parameterisation of the existing density-dependence function, in combination with allowing for time-dependence in  $K$ , resilience and natural mortality could be used to mimic dynamics caused by other underlying models. The SWG agreed to develop some trials along these lines. Trials based on factors  $I$ ,  $K$ ,  $L$  and  $M$  should permit a preliminary exploration of the effectiveness of using these factors to capture alternative dynamics. Further consideration of model structure uncertainty will be pursued as a high priority issue at the next meeting.

Allowance for stock and/or spatial structure can also lead to very complicated dynamical behaviour but the information on stock structure available did not suggest any clear hypotheses that warranted consideration at this stage. Butterworth noted that all population dynamics models include the processes of births and deaths and, in general, the results of population model fits were not particularly sensitive to the choice of the way density-dependence is modelled. He highlighted several alternative population

dynamics models: density-dependence on the age-at-maturity, only modelling the population after 1970 (i.e. ignoring the historical catch data), changes over time in genetic structure and selection against certain genetic traits (e.g. faster growing animals) over time as a result of harvesting.

The SWG agreed to allow for more flexibility in the model structure by conducting trials which eliminate one of the most constraining aspects of the bowhead assessment: the early historic catch record. Accordingly, it agreed to conduct three trials in which the population projections start with a steady state age structure in 1940 rather than in 1848. The procedure for conditioning the trials involves specifying a prior for  $K_{1+}$  and applying the algorithm described in Sections A.6 and A.7 of SC/N99/AWMP1. The SWG agreed that the prior for  $\log(K_{1+})$  would be  $U[\log(7000), \log(31000)]$  (IWC, 1999b).

#### 4.4.3 Evaluation trials

In developing the *Evaluation and Robustness Trials*, it was agreed that some trials should be conducted in which  $MSYR_{1+}$  equals 1% and equals 4%. However, as shown in SC/N99/AWMP3, applying the conditioning algorithm when  $MSYR_{1+} = 1\%$  or 4% often leads to population trajectories that are incompatible with the observed abundance estimates for the Bering-Chukchi-Beaufort bowhead stock. Therefore  $MSYR_{1+} = 1\%$  is combined with an assumption of changes in survey bias and  $MSYR_{1+} = 4\%$  is combined with the assumption that  $MSYL_{1+} = 0.8$ . Combining factors in this way should permit the model predictions of the abundance estimates to be comparable with the actual data.

Table 3 lists the 30 *Evaluation Trials*. These trials are based on a base-case in which  $MSYR_{1+} = 2.5\%$  and need increases from 67 to 201 over the 100-year simulation period (trial BE01). Trials BE02 to BE10 involve changing one feature of the base-case trial. The *Evaluation Trials* examine the implications of different levels of need, future survey bias, future survey CVs,  $MSYL_{1+}$ ,  $MSYR_{1+}$  and survey frequency. Multi-factor evaluation trials consider the implications of poor data (positively biased surveys and high additional variance), low  $MSY$  rate and different levels of future need. All of the trial scenarios are conducted for the deterministic variant of the operating model but some of the *Evaluation Trials* are also conducted using the stochastic population dynamics model. The trials consider inter-survey periods of 5 and 10 years in order to be able to provide the Commission with advice on the consequences of different inter-survey intervals.

#### 4.4.4 Robustness trials

Table 4 lists the *Robustness Trials*. These trials consider variants of the *Evaluation Trials* that examine additional assumption violations. The trials generally examine the implications of the robustness factors combined with different values for  $MSYR$  and survey CV. This is because previous experience suggests that  $MSYR$  is the major factor impacting performance while a scenario in which the true survey CV is 0.1 when the actual CV is 0.25 was one in which candidate CLAs performed relatively poorly.

In prioritising the work for the next meeting, it was agreed that seven trials, in addition to the *Evaluation Trials*, would be set up by the Secretariat to enable SLA developers to construct an SLA that might be robust to some factors considered in the *Robustness Trials*.

Table 3

The thirty evaluation trials for the Bering-Chukchi-Beaufort Seas stock of bowhead whales. The symbols 'D', 'S<sub>D</sub>', and 'S<sub>E</sub>' denote respectively the deterministic population dynamics model, the stochastic population dynamics model with demographic stochasticity, and the stochastic population dynamics model with stochasticity and serially-correlated environmental variation.

Trial no.	Model	Description	MSYR <sub>1+</sub>	MSYL <sub>1+</sub>	Final need	Survey frequency	Historical survey bias	Future survey bias	Survey CV (true, estimated)	Age data
BE01*	D,S <sub>D</sub> ,S <sub>E</sub>	Base case	0.025	0.6	201	5	1	1	0.25, 0.25	Good
BE02	D	Constant need	0.025	0.6	67	5	1	1	0.25, 0.25	Good
BE03	D,S <sub>E</sub>	Future +ve bias	0.025	0.6	201	5	1	1.5	0.25, 0.25	Good
BE04	D	Future -ve bias	0.025	0.6	201	5	1	0.67	0.25, 0.25	Good
BE05	D	Underestimated CVs	0.025	0.6	201	5	1	1	0.25, 0.10	Good
BE06*	D	MSYL <sub>1+</sub> = 0.4	0.025	0.4	201	5	1	1	0.25, 0.25	Good
BE07*	D,S <sub>D</sub> ,S <sub>E</sub>	MSYL <sub>1+</sub> = 0.8	0.025	0.8	201	5	1	1	0.25, 0.25	Good
BE08	D	10 yr surveys	0.025	0.6	201	10	1	1	0.25, 0.25	Good
BE09*	D,S <sub>E</sub>	MSYR <sub>1+</sub> = 1%	0.01	0.6	201	5	0.67 → 1	1	0.25, 0.25	Good
BE10*	D,S <sub>E</sub>	MSYR <sub>1+</sub> = 4%	0.04	0.8	201	5	1	1	0.25, 0.25	Good
BE11	D,S <sub>E</sub>	Bad data	0.025	0.6	201	5	1	1.5	0.25, 0.10	Poor
BE12*	D,S <sub>E</sub>	Difficult 1%	0.01	0.6	134	5	1 → 1.5	1.5	0.25, 0.10	Poor
BE13	D	Difficult 1%; const need	0.01	0.6	67	5	1 → 1.5	1.5	0.25, 0.10	Poor
BE14	D	Need increases to 134	0.025	0.6	134	5	1	1	0.25, 0.25	Good
BE15	D,S <sub>E</sub>	Future +ve bias; 134 need	0.025	0.6	134	5	1	1.5	0.25, 0.25	Good
BE16	D,S <sub>E</sub>	MSYR <sub>1+</sub> = 1%; 134 need	0.01	0.6	134	5	0.67 → 1	1	0.25, 0.25	Good
BE17	D	MSYR <sub>1+</sub> = 4%; 10 yr surveys	0.04	0.8	201	10	1	1	0.25, 0.25	Good
BE18	D	MSYR <sub>1+</sub> = 1%; 10 yr surveys	0.01	0.6	201	10	0.67 → 1	1	0.25, 0.25	Good
BE19	D	Difficult 1%; const need; 10yr surveys	0.01	0.6	67	10	1 → 1.5	1.5	0.25, 0.10	Poor

\* Requires conditioning.

Table 4

The robustness trials.

Trial no.	Factor	Basic trials (Table 3)	Factor Level
BR01*	A: Density-dependence	1, 1 <sup>s</sup> , 6, 7, 7 <sup>s</sup>	Density-dependence on mature (BE trials use 1+)
BR02	B: Stochastic dynamics	2 <sup>s</sup> , 8 <sup>s</sup> , 13 <sup>s</sup> , 18 <sup>s</sup>	Stochastic dynamics (with serially-correlated environmental variation)
BR03	D: Form of need time dependence	1, 9, 9 <sup>s</sup> , 10	Step function in year 2053
BR04	E: Survey frequency	1, 9, 10	15 yrs
BR05	F: Strategic surveys	1, 9	Yes + CV = (0.25, 0.25)
		1, 9	Yes + CV = (0.34, 0.25)
BR06*	G: Survey bias time dependence	1	Historic bias: 1.5 constant; Future bias: decreasing (1.5→1)
		1	Historic bias: 0.67 constant; Future bias: increasing (0.67→1)
		9, 9 <sup>s</sup>	Historic bias: 1.0 constant; Future bias: sinusoidal (see Fig 1)
		12, 12 <sup>s</sup>	Historic bias: 1.0 constant; Future bias: decreasing (1.0→0.67)
BR07	H: Future survey CV	1	CV = (0.1, 0.1)
		1, 1 <sup>s</sup>	CV = (0.34, 0.25)
		9	(0.1, 0.1) + sinusoidal survey bias (factor G)
BR08*	I: Historic catch bias	1, 7, 7 <sup>s</sup> , 9, 9 <sup>s</sup> , 10	0.5
		1, 7, 7 <sup>s</sup> , 9, 9 <sup>s</sup>	1.5
BR09	K: Time dependence in K	1, 2, 9, 10	K halves linearly over 100 years
		1, 2, 9, 10	K doubles linearly over 100 years
		1	K varies sinusoidally with maximum in year 40
BR10	L: Time dependence in resilience	10	Resilience halves linearly over 100yr,
		9	Resilience doubles linearly over 100yr
		1, 2, 8	Step MSYR 2½%→1%→2½% every 33 yrs (alone)
		1, 2, 8	Step MSYR 2½%→1%→2½% every 33 yrs in sync with M
		10	Resilience halves linearly over 100yr + CV = (0.1, 0.1)
		9	Resilience doubles linearly over 100yr + CV = (0.1, 0.1)
		1	Step MSYR 2½%→1%→2½% every 33 yrs (alone) + CV = (0.1, 0.1)
		1	Step MSYR 2½%→1%→2½% every 33 yrs in sync with M + CV = (0.1, 0.1)
BR11	M: Time dependence in M	1, 2, 9, 10	M halves linearly over 100 years
		1, 2, 9, 10	M doubles linearly over 100 years
		1, 9, 10	M halves linearly over 100 years + CV = (0.1, 0.1)
		1, 9, 10	M doubles linearly over 100 years + CV = (0.1, 0.1)
BR12	N: Episodic events	1, 1 <sup>s</sup> , 2, 9, 9 <sup>s</sup> , 10, 10 <sup>s</sup>	1% probability each year that 50% of animals die
BR13*	O: Integrated	1, 11	MSYR <sub>1+</sub> ~U[0.01, 0.04]; fixed MSYL <sub>1+</sub> =0.6
		1, 11	MSYR <sub>1+</sub> ~U[0.01, 0.04]; MSYL <sub>1+</sub> ~U[0.4, 0.8]
		1, 1 <sup>s</sup> , 2, 11, 11 <sup>s</sup>	MSYR <sub>1+</sub> ~U[0.01, 0.04]; MSYL <sub>1+</sub> ~U[0.4, 0.8];
			historical catch bias ~U[0.5, 1.5]; Serial correlation ~ U[0.47, 0.95]
BR14*	P: Initial year of population projection	1, 9, 10	1940 (reference or base case level is 1848)

\* - Requires additional conditioning.

The SWG recommended that Huang apply the assessment scheme in SC/N99/AWMP5 to identify which *Robustness Trials* behaved in a similar manner to the *Evaluation Trials* and which exhibit notably different behaviour. This will be useful for determining the extent to which *Robustness Trials* can be used for cross-validation.

#### 4.5 SLA tuning

The SWG agreed that depletion tuning of candidate SLAs should be conducted to match the 1+ median final depletion achieved by the idealised SLA 'H' for trials BE9 and BE1 (using the deterministic population dynamics model). Section E of Annex D provides additional details on this approach.

#### 4.6 Review of performance statistics

The SWG agreed to revise the list of performance evaluation statistics as shown in Table 5. Additional details are given in section F of Annex D. These changes resulted from several topics that arose during review of appendix 3B of Annex E (IWC, 2000b).

First, *Robustness Trials* dealing with time dependence of certain parameters (BR9, BR10 and BR11 in Table 4) require reconsideration of some risk and recovery statistics. Statistic D1 (Final Depletion) should be replaced with D8 = PT/PT\* for such trials, although D1 should also be retained as an optional statistic in these cases. D8 should be optional for all other trials. Statistics R1, R3, and R4 should be omitted on these trials. The reason for all these changes is that statistics related to carrying capacity and MSYL are not

well-defined when the trial assumes that carrying capacity, resilience, or natural mortality varies over time. Aside from adding D8, the SWG agreed that inspection of trajectory plots (D6 and D7) should play a more prominent role in such trials to examine risk and recovery performance. A new statistic (D9) was also added to give the minimum population level.

The *Robustness Trials* related to episodic catastrophes (BR12 in Table 4) also motivated consideration of another statistic: the probability that the stock is reduced to extinction or very near to extinction. Such a statistic might be less sensitive to the occurrence of a catastrophe near the end of the 100-year simulation than would a statistic like D1, final depletion. Although the group saw merit in the proposed new statistic, it felt it would be easier to formulate a suitable formal definition after the 2000 Scientific Committee meeting at which time results from these trials would be available for inspection. The new D9 statistic also gives information on this point.

The SWG noted that its intention was to formulate all performance statistics so that higher values corresponded to more desirable performance. With this in mind, revised statistics N10 and N11 are given in Annex D. The SWG further noted that it intended to determine a preference between N10 and N11. Givens and Huang noted that they had used both statistics in a different context while conducting the research summarised in SC/N99/AWMP5 and had settled on an analogue to N11. The group agreed to determine a preference at the 2000 Scientific Committee meeting.

Table 5  
Performance statistics.

ID	Mandatory	Optional	Time periods	Can be shown using Zeh graph program	Used to explain performance to layperson	Used to evaluate performance for Scientific Committee	Name	Comment
D1	1+, mature		100, <del>50, 20</del>	Yes	Yes	Yes	Final Depletion	Delete 20, 50 years (1999)
D2	<del>1+</del> , mature		100, <del>50, 20</del>	Yes	Yes	Yes	Lowest Depletion	Delete 1+ and 20, 50 years (1999)
<del>D3</del>							<del>Lowest Stock Level under Exploitation</del>	Delete: misleading (1997)
<del>D4</del>							<del>Lowest Depletion under Exploitation</del>	Delete: misleading (1997)
<del>D5</del>							<del>Relative Survival Probability</del>	Delete: not useful (1997)
D6		1+, mature	100	Yes	Yes	No	Trajectories 1 and 2	Demote to optional (1999)
D7		1+, mature	100	Yes	Yes	No	Pointwise Quantile Trajectories	Demote to optional (1999)
D8	BR9, 10, 11		100	Yes		No	Rescaled final Depletion	Defined at 1999 workshop
D9		1+, mature	100	Yes		No	Minimum number of whales	Defined at 1999 workshop
N1		Yes	20, <del>50</del> , 100	Yes	Yes	Yes	Total Need Satisfaction	Delete 50 years (1999)
N2		Yes	20, <del>50</del> , 100	Yes	Yes, after rescaling	Yes	Longest Shortfall	Delete 50 years (1999)
<del>N3</del>							<del>Severe Shortfall Frequency</del>	Delete: not useful (1999)
N4		Yes	20, <del>50</del> , 100	No	Yes	Yes	Shortfall Frequency	Delete 50 years (1999)
<del>N5</del>		Yes	20, <del>50</del> , 100	Yes	Yes	Yes	Block Need Satisfaction	Delete (1999 workshop)
<del>N6</del>							<del>Adapted Utility</del>	Delete: undesirable properties (1997)
N7		Yes	100	Yes	No	Yes	Percent Need Satisfaction Pointwise Quantile Trajectory Plot	1996. Redefined in 1997
N8		Yes	100	Yes	No	Yes	Percent Need Satisfaction Trajectories 1 and 2 Plot	1996
N9	Yes		20, <del>50</del> , 100	Will be	Yes	Yes	Average need satisfaction	Delete 50 years (1999)
N10	Yes		100	Will be	No	Yes	Average Annual Variation in Catch	Modified (1999 workshop)
N11	Yes		100	Will be	No	Yes	Anti-curvature Catch Variation Statistic	Modified (1999 workshop)
R1	1+, mature		100	Yes	Yes	Yes	Relative Recovery	Redefined in 1997
<del>R2</del>							<del>Relative Probability of Recovery</del>	Delete (1999)
R3		1+, mature	100, <del>50, 20</del>	No	Yes	Yes	Time Frequency in Recovered State after Recovery	Delete 20, 50 years (1999)
R4		1+, mature	100, <del>50, 20</del>	No	Yes	Yes	Relative Time to Recovery	Delete 20, 50 years (1999)

The SWG agreed to remove block need satisfaction (N5) from the list of performance evaluation statistics because the imposition of 5-year block strike limits in *SLA* simulation testing (see appendix 3B, Annex E (IWC, 2000b)) had greatly reduced concern about assessing short-term cumulative need satisfaction. The remaining need statistics could now effectively address this concern.

Finally, the SWG noted that it may be necessary to reconsider some statistics and/or their interpretations on integrative *Robustness Trials* like BR13. This topic will be discussed at the 2000 Scientific Committee meeting.

#### 4.7 *SLA* optimisation

Several aspects of the simulation software which developers might use for optimising *SLAs* required changes due to the new trial structure. The SWG agreed that the software should be altered to calculate a sum of square deviations from an idealised *SLA* 'H' in the manner shown in equation G.2 of Annex D, as had been originally advocated by Givens, Punt and Bernstein (1999). The addition of equation G.2 to the Annex was necessary because it had been unintentionally omitted from past documentation despite being included in the simulation software. The SWG also noted that the definition of 'H' (and hence the calculation of the sum of squared deviations) might be unclear for *Robustness Trials* BR9, BR10 and BR11. *SLA* developers were very unlikely to want to use such trials in any optimisation; therefore this topic was not addressed further.

Several interesting optimisation strategies have been presented to the SWG in recent years (Givens, 1999; Witting, 1999) and some of these may lead to improved *SLA* performance. The SWG reiterated that a developer wishing to optimise *SLA* performance is encouraged to do so in whatever fashion she/he desires. The objective function need not be based on sums of squared deviations. The particular 'H' given in Annex D is not the required goal of any optimisation exercise. Moreover, a developer might choose to use an objective function which was related to any subset of *Evaluation Trials*, available *Robustness Trials*, or other trials individually crafted for the purpose.

### 5. IMPLEMENTATION TRIAL STRUCTURE FOR GRAY WHALES

#### 5.1 Simulation scenarios

The SWG believed that consideration of the classes of trials important for *SLA* development and testing - and the relationships between them - was largely a matter which is independent of any particular stock and therefore referred to its earlier discussion of the topic (see Item 4.1).

#### 5.2 Review of SC/N99/AWMP1 sections A-C

SC/N99/AWMP1 presented draft trial specifications for the Eastern North Pacific gray whales based on the *Initial Exploration Trials* for the Bering-Chukchi-Beaufort bowhead stock (IWC, 2000b). The trials involve modifying the specifications related to the age-at-recruitment, the age-specific survival rates, how the population projections are initialised and how the trials are conditioned. Unlike the trials for the Bering-Chukchi-Beaufort bowhead whales, the trials specified in SC/N99/AWMP1 require specification of priors for  $K_{1+}$  and the amount of 'additional variation' associated with the estimates of abundance.

The Workshop noted that the age-at-recruitment for the Russian harvest was less in recent years than in the past. However, prior to modifying the operating model to allow changes over time in the age-at-recruitment, assessments

based on ages-at-recruitment of 1 and 5 should be conducted to determine if a difference in the age-at-recruitment has a major impact on estimated trajectories of population size. The Workshop agreed that the population projections should start in 1930 but that trials should consider changes to this. SC/N99/AWMP1 only used the 1+ abundance data for conditioning purposes. The Workshop agreed that the data on calf counts should also be used when conditioning. The Workshop recommended that hypotheses regarding the relationship between the calf counts at Granite Canyon and the model estimates of the number of calves needed to be developed.

The initial level of need is 124 based on the current catch limit for the ENP gray whales. Initial trials should include a need envelope of the same shape as those for Bering-Chukchi-Beaufort bowheads and West Greenland minke whales (constant and increasing to three times the initial level over a 100-year period). The Workshop recommended that the Commission should be advised about this interim specification and requested to provide guidance regarding need envelopes.

#### 5.3 Specification of trials

Table 6 lists the factors to be considered in the single stock *Evaluation*, *Robustness* and *Cross-validation Trials* for the ENP gray whales. Any multi-sub-stock trials will require factors additional to those in Table 6 to capture uncertainty adequately. The factors specific to gray whales in Table 6 are sensitivity to the first year considered in the population projections, time-trends in amount of additional variability, and uncertainty about the magnitude and sex ratio of the historical catches.

Table 6

Factors to be included in trials for the Eastern North Pacific stock of gray whales.

MSYR <sub>1+</sub>		
MSYL		
Density dependent component		
Population dynamics model		
Age of recruitment		
First year considered in population projections		
Need in final year		
Form of need time dependence		
Survey frequency		
Strategic surveys		
Survey bias:	(i)	constant
	(ii)	historical time dependence
	(iii)	future time dependence
Survey CV: (true, estimated)		
Additional variability: time trends		
Historic catch uncertainty:	(i)	magnitude
	(ii)	sex ratio
Calf count data		
Carryover strategy		
Length of simulation		
Time dependence in:	(i)	$K$
	(ii)	resilience
	(iii)	natural mortality, $M$
Episodic events		
Integrated		
Alternative population models		

Table 7 lists a set of *Initial Exploration Trials* for the ENP gray whales. The choices for MSYR<sub>1+</sub> in Table 7 are based on the estimates from recent assessments (e.g. Wade and DeMaster, 1996; Punt and Butterworth, 2001). The

magnitude of the rate of increase in survey bias will be selected to achieve adequate fits to the abundance data.  $MSYL_{1+} = 0.6$  for all but trials GI02 and GI03.

Table 7

The nine *Initial Exploration Trials* for the Eastern North Pacific stock of gray whales. The symbols 'D', 'and 'S<sub>E</sub>' denote respectively the deterministic population dynamics model, and the stochastic population dynamics model with demographic stochasticity and serially correlated environmental variation.

Trial no.	Description	MSYR <sub>1+</sub>	MSYL <sub>1+</sub>	Historical survey bias	Model	Final need
GI01*	Base case	0.035	0.6	1	D,S <sub>E</sub>	372
GI02*	MSYL <sub>1+</sub> = 0.4	0.035	0.4	1	D	372
GI03*	MSYL <sub>1+</sub> = 0.8	0.035	0.8	1	D,S <sub>E</sub>	372
GI04*	MSYR <sub>1+</sub> = 1.5%	0.015	0.6	?? → 1	D,S <sub>E</sub>	372
GI05*	MSYR <sub>1+</sub> = 5.5%	0.055	0.6	1 → ??	D,S <sub>E</sub>	372

#### 5.4 Performance statistics

The SWG reaffirmed its intention that the same performance evaluation statistics should be used in the AWMP development of each aboriginal fishery. It noted, however, that for evaluation of management procedures for the Eastern North Pacific stock of gray whales, the intended simulation model and the nature of the available data do not facilitate precise estimation of carrying capacity. Therefore, performance evaluation statistics such as  $D1 (P_T/K)$  will not be as informative as for trials with other stocks. The SWG agreed that development of more appropriate evaluation statistics in this case (for example the ratio of  $P_T$  to the abundance in a recent year) should be discussed in detail at the 2000 Scientific Committee meeting.

## 6. OTHER MATTERS

### 6.1 Work plan and priority tasks

Allison noted that certain public-domain computer routines she used for generating random numbers and calculating inverse probability functions in the stochastic simulation model did not produce replicable results across computer platforms and FORTRAN compilers. The SWG believed

that this problem was likely caused by differing machine precisions (and/or program statements referring to these precisions). Allison believed she could fix this problem without excessive delay, through changing code, compilers, software, or some combination of such approaches. The SWG agreed that a solution to this problem was among the tasks of highest priority to be completed before the Adelaide meeting and urged that any reasonable level of necessary funding be provided to resolve the problem as quickly and easily as possible. The solution must work on any IBM-compatible PC as well as common Unix platforms.

The SWG reviewed its long-term workplan for completion of a bowhead AWMP (see item 8.1.1 of Annex E, IWC, 2000b). That plan included two schedules; the SWG reaffirms its intention to adhere to the faster schedule as long as that remained possible. Accordingly, the SWG compiled a list of research and computing tasks identified at this meeting which it believed were necessary for development to progress (Table 8). The deadlines listed in that table indicate the latest date by which each task would need to be completed if the faster track was to be maintained for bowhead AWMP development. The SWG strongly urges that the resources necessary to complete these tasks on schedule are made available.

Other tasks:

A map showing catches of bowhead whales by sex and size landed by village (see Item 3.1).

Encourage collection of genetic samples from gray whales for individual identification. Funding by the IWC might facilitate this.

Huang to repeat the example analyses in SC/N99/AWMP5 for the full set of trials (and each class) established at this meeting, so that the results can be examined at the 2000 meeting of the Scientific Committee.

Donovan to draft a document outlining some possible carryover rules for presentation to the Commission (see Item 4.4.2).

## 7. ADOPTION OF REPORT

The report was adopted at 17:00 on Saturday, 20th November 1999. Donovan thanked the National Marine Mammal Laboratory for their invitation to host the meeting and for their hospitality. The SWG expressed its appreciation to Donovan for his excellent chairmanship.

Table 8

Tasks.

Trial/subject	Task	Target date	Time estimate
BE deterministic trials	Amend program, prepare datasets and run conditioning	6 Jan.	3 weeks
BE stochastic trials	Sort out reproducibility problem	15 Jan.	? 1-4 weeks
	Prepare datasets and run conditioning	25 Jan.	
BR trials except 5, 9-12	Amend program, prepare datasets and run conditioning	14 June	1 month
	Check trajectories fit the data		
BR trials 5, 9-12	Changes to program.	1 May	2 months
Statistics	Amend trajectories and statistics	6 Jan.	2 weeks
Zeh plotting program	Add new statistics	1 Mar.	
H optimisation.	Provide partial output specifications	25 Dec.	
Trial visualisation		14 June	
GI trials	Finalise specifications. Run test trials	14 June	2 months
G multi stock model	Methods to potentially discriminate single/multi-sub-stock assumptions	14 June (NMML)	
	Potential <i>SLA</i> inputs related to possible multi-sub-stock dynamics		

## REFERENCES

- Bockstoce, J.R. and Botkin, D.B. 1980. *The Historical Status and Reduction of the Western Arctic Bowhead Whale (Balaena mysticetus) Population by the Pelagic Whaling Industry, 1848-1914*. New Bedford Whaling Museum, New Bedford, Mass. 160pp. [Final report to the National Marine Fisheries Service by the Old Dartford Historical Society (New Bedford Whaling Museum) under contract 03-78-MO2-0212].
- Bogoslovskaya, L.S. 1986. On the social behaviour of gray whales off Chukotka and Koryaka. *Rep. int. Whal. Commn* (special issue) 8:243-51.
- Braham, H.W. 1995. Sex and size composition of bowhead whales landed by Alaskan Eskimo whalers. pp. 281-313. In: A.P. McCartney (ed.) *Studies in Whaling. 3. Hunting the Largest Animals: Native Whaling in the Western Arctic and Subarctic*. The Canadian Circumpolar Institute, Edmonton, Alberta.
- Fowler, C.W. and Baker, J.D. 1991. A review of animal population dynamics at extremely reduced population levels. *Rep. int. Whal. Commn* 41:545-54.
- Givens, G.H. 1998. AWMP development and diverse prototypes. *Rep. int. Whal. Commn* 48:483-95.
- Givens, G.H. 1999. Multicriterion decision merging: competitive development of an aboriginal whaling management procedure. *J. Am. Stat. Assoc.* 94:1003-4.
- Givens, G.H., Punt, A.E. and Bernstein, T.A.O. 1999. Equivalence tuning of SLAs. *J. Cetacean Res. Manage.* 1(3):289-95.
- International Whaling Commission. 1992. Report of the Fourth Comprehensive Assessment Workshop on Management Procedures, Annex D. Single stock trials for finalised procedures. *Rep. int. Whal. Commn* 42:317-8.
- International Whaling Commission. 1999a. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 1:1-52.
- International Whaling Commission. 1999b. Report of the Scientific Committee. Annex F. Report of the Standing Working Group (SWG) on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage. (Suppl.)* 1:157-78.
- International Whaling Commission. 2000a. Report of the Scientific Committee. Annex D. Report of the Sub-Committee on the Revised Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 2:79-92.
- International Whaling Commission. 2000b. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 2:125-54.
- Punt, A.E. 1999a. Report of the Scientific Committee. Annex R. A full description of the standard BALEEN II model and some variants thereof. *J. Cetacean Res. Manage. (Suppl.)* 1:267-76.
- Punt, A.E. 1999b. Some aspects related to the specification of simulation trials for fishery type 2. Paper SC/51/AWMP1 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 15pp. [Paper available from the Office of this Journal].
- Punt, A.E. and Butterworth, D.S. 2001. An examination of certain assumptions made in the Bayesian approach used to assess the eastern North Pacific stock of Gray whales (*Eschrichtius robustus*). *J. Cetacean Res. Manage.* (special issue) 3:In press.
- Wade, P.R. and DeMaster, D.P. 1996. A Bayesian analysis of eastern Pacific gray whale population dynamics. Paper SC/48/AS3 presented to IWC Scientific Committee, June 1996, Aberdeen (unpublished). 21pp. [Paper available from the Office of this Journal].
- Witting, L. 1999. Total catch optimisation of catch control laws for subsistence whaling. Paper SC/51/AWMP6 presented to the IWC Scientific Committee, May 1999, Grenada, WI (unpublished). 7pp. [Paper available from the Office of this Journal].
- Zeh, J.E., Raftery, A.E. and Schaffner, A.A. 1995. Revised estimates of bowhead population size and rate of increase. Paper SC/47/AS10 presented to the IWC Scientific Committee, May 1995 (unpublished). 28pp. [Paper available from the Office of this Journal].

---

## Annex A

### Agenda

1. Introductory items
    - 1.1 Convenor's opening remarks
    - 1.2 Election of Chair
    - 1.3 Appointment of rapporteurs
    - 1.4 Adoption of agenda
    - 1.5 Review of documents
  2. Review of progress on pre-workshop tasks identified last meeting
    - 2.1 Secretariat computing
    - 2.2 Continued development of *Strike Limit Algorithms (SLA)*
  3. Stock structure
    - 3.1 Bowhead whales
    - 3.2 Gray whales
  4. Implementation trial structure for bowhead whales
    - 4.1 Classes of trials, their identification and relationships between them
  - 4.2 Review of Annex E (IWC, 2000b), appendix 3B, sections A-C
  - 4.3 Principles of trial design
  - 4.4 Specification of trials
  - 4.5 SLA tuning
  - 4.6 Review of performance statistics
  - 4.7 SLA optimisation
  5. Implementation trial structure for gray whales
    - 5.1 Simulation scenarios
    - 5.2 Review of SC/N99/AWMP1 sections A-C
    - 5.3 Specification of trials
    - 5.4 Performance statistics
  6. Other matters
    - 6.1 Work plan and priority tasks
  7. Adoption of report
-

## **Annex B**

### **List of Documents**

SC/N99/AWMP

1. Punt, A.E. Draft specifications for trials for the eastern North Pacific stock of gray whales for the case where stock structure is not uncertain.
2. Punt, A.E. Additional trials for the Bering-Chukchi-Beaufort Seas stock of bowhead whales.
3. Punt, A.E and Pribac, F. Further examination of an *SLA* for the Bering-Chukchi-Beaufort Seas stock of bowhead whales.
4. Givens, G.H. A complete collection of aboriginal whaling management procedure simulation trials proposed for the Bering-Chukchi-Beaufort Seas stock of bowhead whales.
5. Givens, G.H. and Huang, M. On the identification of outlying and extrapolative scenarios for AWMP simulation testing.
6. Zeh, J. Documentation of S functions to plot and print AWMP performance statistics.

---

## **Annex C**

### **Terminology**

[See This Volume, pp. 9-10]

---

## **Annex D**

### **Fishery Type 2: Implementation for Bering-Chukchi-Beaufort Bowhead Whales**

[See This Volume, pp. 17-27]

---

## **Annex E**

### **Trajectory Plots for Bowhead *Initial Exploration Trials***

[Trajectory plots on following pages]









































