Annex G

Report of the Sub-Committee on In-Depth Assessments

Members: Walløe (Chair), Skaug (Co-chair), An, Baba, Bannister, Best, Brandon, Bravington, Brownell, Burt, Butterworth, Campbell, Charrassin, Childerhouse, Chilvers, Cooke, Elvarsson, Ensor, Fujise, Gales, Gallego, Gedamke, Goodman, Gunnlaugsson, Hakamada, Hammond, Hatanaka, Hedley, Holloway, Hughes, Jaramillo-Legorreta, Jérémie, Kanda, Kasuya, Kitakado, Kelly, Kock, Lauriano, Leaper, Liebschner, Lockyer, Luna, Lyrholm, Matsuoka, Miyashita, Morishita, Muller, Murase, Øien, Okada, Okamura, Palka, Pastene, Punt, Roel, Sekiguchi, Uoya, Uozumi, Williams, Yamakage, Yasokawa, Yoshida, Young.

1. ELECTION OF CHAIR AND CO-CHAIR

Walløe welcomed the participants and was elected Chair. Due to the high workload this year, Skaug was asked to co-chair the sub-committee, being mainly responsible for items related to abundance estimation of Antarctic minke whales.

2. APPOINTMENT OF RAPPORTEURS

Burt, Butterworth, Cooke and Hedley agreed to act as rapporteurs.

3. ADOPTION OF THE AGENDA

The adopted Agenda is given in Appendix 1.

4. DOCUMENTS AVAILABLE

The documents relevant to the work of the sub-committee were SC/62/IA1-15; SC/62/O15-17 and SC/62/Rep3 and Rep6.

5. ANTARCTIC MINKE WHALES

5.1 Abundance

5.1.1 Report from intersessional e-mail working group

Skaug reported on work conducted by the Abundance Estimation Intersessional Working Group. Tasks to be considered by the group were listed in Appendix 3 of Annex G (IWC, 2010), most of which were directed towards elucidating possible causes for the difference in abundance estimates for Antarctic minke whales from the IDCR/SOWER data from the recent OK (Okamura and Kitakado, 2009) and SPLINTR (Bravington and Hedley, 2009) models. In completing most of these tasks, substantial progress had been made towards this in two regards: (i) a reference dataset, which did not require any further processing for the two models to be applied, had been developed for model comparisons; and (ii) Bravington had completed a non-spatial version of the SPLINTR model (see Item 5.1.3). For (i), a number of internal inconsistencies in the 'standardised' dataset were identified, and as noted at last year's meeting, it was essential that when comparing models, the data were identical.

Due to lack of documentation on the reference dataset during the initial discussion, there were some concerns expressed that the reference dataset might not be the most appropriate for abundance estimation. However, since the purpose of this dataset is for valid comparisons between the models, it was **agreed** that this dataset was suitable for this purpose. Model developers were free to use and post-process alternative versions of the data, (such as the standardised data), for their preferred estimates of abundance from the IDCR/SOWER surveys. Because of differences in the way the data are processed by the two models, it was not likely that there would be an agreed 'best' dataset for the analysis.

Bravington indicated what changes had been made to the standard dataset to produce the reference dataset. Changes made were principally minor, but without them direct comparisons of the two models would not have been possible. The main change was to the boundary files defining the strata. In principle, it should be possible to use these boundaries to assign any effort record to a physical stratum. However, the effort and boundary datasets are often inconsistent, and in fact some effort falls outside the strata altogether. In those cases where such an inconsistency occurred, the stratum boundaries were slightly adjusted, so that each effort record fell in into the 'obvious' stratum. It was noted that these stratification changes are not changes to the data *per se*. Therefore they do not influence the spatial-SPLINTR abundance estimates at all, because that method does not use strata. They only affect the comparison between OK and stratified-SPLINTR, because the latter needs consistent strata in order to run at all. The effect of the stratification changes on OK is very small, because the revisions make very little difference to the physical area of the strata and the OK model does not use precise location data relative to stratum boundaries. The sub-committee **agreed** that it would be useful if Bravington prepared a paper intersessionally to formally document the differences between the datasets, and record apparent inconsistencies in the IDCR/SOWER data.

There was some discussion on how or whether to initiate a process to correct inconsistencies in the data stored within DESS. No conclusion was reached.

5.1.2 Results from simulated datasets

SC/62/IA14 provided results from applying the IWC 'standard' method (Branch, 2006), the OK and SPLINTR models to simulated data, focusing on the latter two. In general, both of these models performed quite well, although when bias did occur, it tended to be positive for the OK model and negative for SPLINTR. For the most complex scenarios, in which density, school size and weather gradients interacted (but excluding those for which duplicate sightings were mis-classified), estimates from SPLINTR were slightly less biased than those from the OK model. However, measurement errors caused positive bias in the SPLINTR estimates, but not in the OK estimates. Whilst non-synchronised diving positively biased the OK estimates in the most complex scenarios, its effect on the SPLINTR estimates was not totally clear; although it was not significant in the complex scenarios, it did cause significant negative bias in simpler cases. The reason for this was not known.

In relation to the patterns of surfacing, it was pointed out that there are empirical data from dive time experiments conducted on IDCR/SOWER cruises; these provide information on surfacing intervals and, for schools of size two or more, surfacing synchronicity. Despite efforts to obtain dive time data across a range of school sizes, it should be noted that there are only limited observations from schools of size 1, as singletons proved difficult to follow. A preliminary analysis of these data is in Hedley and Ensor (2006).

The sub-committee expressed its thanks to Palka for co-ordinating this extensive simulation study. It has been extremely valuable in helping to develop and refine the models, and is now enormously helpful in examining the differences between them. None of the scenarios show the level of difference between the OK and SPLINTR estimates as is currently seen in the real data analyses. This suggests either that the magnitudes of factors currently in the simulations do not cover the ranges found in the real data (either singly or in combination), or that there are additional factors not currently in the simulations that are important for modelling the real data.

Palka indicated that she was willing to continue to work on the simulation study, including providing new scenarios if necessary, but these would need to be specified without delay. At this point, suggestions for new scenarios were referred to the Abundance Estimation Intersessional Working Group, to be re-established this year (see the work plan). If the work of that Group identified specific factors that should be examined to elucidate reasons for differences in the estimates, then new scenarios could be helpful.

5.1.3 Comparison of OK and SPLINTR using the reference dataset

As noted above, different post-processing of data for use by the OK and SPLINTR models clouds investigations into the differences between their estimates. The reference set, agreed mutually by the modellers concerned, was established to compare the models on the basis that no such post-processing would be required to run the models, and hence they would use exactly the same data.

Furthermore, it would be difficult to compare the OK and SPLINTR models directly, even using the reference dataset, as it may not be clear whether any differences between the two resulted from differences in the sighting probability components of the models (the cue-based hazard probability model and the trackline conditional independence model), the school size distribution and school size error models, or differences between the stratified Horvitz-Thompson estimation and spatial modelling of school density. As a first step, it was agreed to compare the OK model with the non-spatial version of SPLINTR, using the reference dataset. A chain of comparisons depicting how this approach fits into an overall comparison of the two models as applied to their preferred data is shown in Fig. 1.

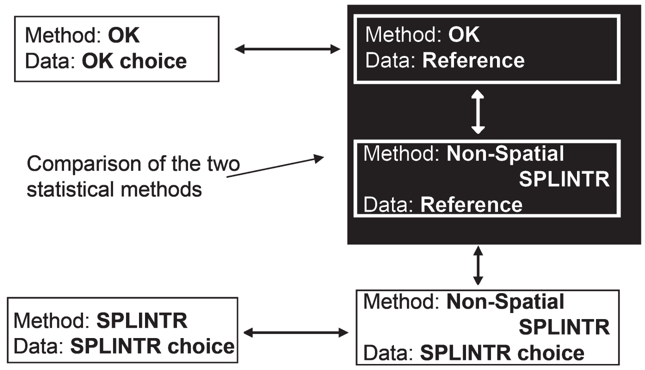

Fig. 1. Chain of comparisons needed to understand the difference between OK and SPLINTR models.

For the OK model, the preferred dataset for analysis includes some records that have been removed in the reference dataset, e.g. because of missing covariates. For the SPLINTR model, the preferred dataset slightly increases the transect effort in a way that is intended to accommodate areas searched (and sightings made within those areas) before and after breaks in effort. There are also some small adjustments in timing to ensure sightings fall into effort legs.

During a two-day pre-meeting immediately prior to this meeting and using the reference dataset, the OK and non-spatial SPLINTR output were compared. Consistency checks revealed that the basic data (numbers of sightings and amount of effort) were the same for each. Estimated mean school sizes (E(s)), effective strip half-widths (eshw), and encounter rates (n/L) were combined using the simple line transect formula for estimating abundance (N=n·E(s) A/2·L·eshw). The resulting estimated abundances for each model were consistent with those reported from the model. This simple check ensured that these estimated quantities from each model were being combined correctly to estimate abundance. Further diagnostic checks were as follows (all by stratum): (i) plots of E(s), eshw, and abundance estimated by OK against corresponding values estimated by SPLINTR; (ii) plots of eshw from OK against eshw from SPLINTR, separately by school size category (1, 2, 3-4, 5-9 and 10+); and (iii) observed against model-predicted number of sightings by school size category × platform combination.

The mean school size plots revealed some variation between the models, but this was not considered to be sufficient to be causing the difference in abundance estimates. However, the eshws for OK were about half of those of SPLINTR, and estimated abundances were approximately doubled, highlighting a need for further investigation. Disaggregating eshws by school size category showed some pattern. For smaller schools (of size 1, 2 and 3-4), estimated strip widths by stratum were consistently higher for SPLINTR than for OK. For schools of size category 5-9, the estimates were comparable, whilst for the largest schools, eshws for OK were higher than those for SPLINTR. In the IDCR/SOWER data, the majority of schools are of size 1 and 2, so the effects of the variation in these plots would not be expected to ‘cancel out’. It was **agreed** that these plots, together with similar ones disaggregated by school size category and platform, were useful in identifying factors causing the difference in estimation from the two models.

The conclusion from these comparisons was that the difference seen in the results from the two models was not due to the data, and was probably not due to differences in mean school size. The question was: had sufficient progress been made to suggest whether further investigations would elucidate a reason for the difference? It was **agreed** that with sufficient commitment to the further work outlined in the Work Plan (Item 5.1.9), including an intersessional workshop, there was a reasonable prospect that the reasons for the differences in the estimates from the two methods would be elucidated. The sub-committee therefore **agreed** to proceed with these investigations until next year’s meeting.

Contingency plans, including producing model-averaged estimates of abundance may also need to be considered in the event that it was not possible to resolve the difference in the estimates. An investigation by Skaug comparing estimates from OK, SPLINTR and a model-averaged estimate on the simulated data showed that for these data, the model-averaged estimator had smaller bias than either of the two individual models (Appendix 2). There was some discussion on the appropriateness of model-averaged estimates on the real data, but given the progress made this year, it is anticipated that a perhaps more satisfactory outcome can be achieved as a result of the planned intersessional work to resolve the reasons for the differences in estimates.

5.1.4 Results from each method using their preferred dataset

SC/62/IA3 and SC/62/IA12 presented 'survey-once' estimates (Branch and Butterworth, 2006) of abundance for the CPII and CPIII surveys from the OK and SPLINTR models respectively (Table 1).

Table 1

Comparison of 'survey-once' estimates of abundance, by *Management Area*, from the OK and SPLINTR models. Estimates shown have been extracted from the papers SC/62/IA3 and SC/62/IA12 and rounded, with CVs incorporating additional variance given in parentheses. CVs given in Table 4 of SC/62/IA12 did not incorporate additional variance but for ease of comparison, these were calculated at last year's meeting.

	Area I	Area II	Area III	Area IV	Area V	Area VI	TOTAL
CPII							
OK	209,000	261,000	187,000	104,000	635,000	90,000	1,486,000
	(0.35)	(0.38)	(0.42)	(0.37)	(0.29)	(0.39)	(0.17)
SPLINTR	117,000	141,000	87,000	61,000	282,000	59,000	747,000
	(0.38)	(0.39)	(0.55)	(0.36)	(0.34)	(0.40)	(0.19)
CPIII							
OK	65,000	93,000	126,000	79,000	244,000	105,000	712,000
	(0.34)	(0.37)	(0.33)	(0.45)	(0.33)	(0.34)	(0.17)
SPLINTR	35,000	56,000	59,000	36,000	140,000	57,000	382,000
	(0.33)	(0.35)	(0.31)	(0.33)	(0.31)	(0.33)	(0.17)

The authors of SC/62/IA3 pointed out that the model used this year was similar to that presented last year but that with slight a modification to the probability distributions for school size bias related to confirmation status. Analyses were presented which examined the sensitivity on abundance to this modification. Two further sensitivity analyses were also presented, one which examined the form of the Q function in the hazard probability model, and one which looked at different covariates affecting detectability (Sightability and Beaufort Class). The AIC best model had a truncated negative binomial model for IO mode and a truncated Poisson distribution for CL mode as a probability model of school size bias, a logistic form as a Q function, and Sightability for CPII and Beaufort Class for CPIII as a weather covariate. In addition, SC/62/IA3 investigated the effect of Platform C (Upper Bridge) on radial distance and abundance. When excluding Platform C, the authors reported that the fit of radial distances was somewhat improved and the abundance estimates decreased by 15% on average.

In discussion, it was suggested that the reduction in the abundance estimates when Platform C was removed may indicate that something is not quite right with the model. The authors agreed, noting a lack of fit at small radial distances was potentially the cause. In addition, there was some evidence of lack-of-fit in the school size distributions particularly for size 1, in CPIII. In response, the authors noted that such lack-of-fit had not been found last year, indicating that this may be due to the change in the school size bias model in relation to confirmation status. It was also commented that the general form of the model had been improved; it no longer included any *Management Area*-specific parameters.

The model presented in SC/62/IA12 did not differ from that presented at last year's meeting. At that meeting, the spatial models had generally appeared to fit the data well, but some lack-of-fit was evident in the analysis of the 2003/04 survey in Area V, where the predicted numbers of sightings were higher than those observed. In SC/62/IA12 therefore, they had adopted a different approach for modelling that year's data, fitting two separate spatial models to accommodate the spatial and temporal discontinuities evident in both survey effort and ice coverage. They reported that these models appeared to fit much better. In an effort to gain greater understanding of the effect of uneven survey coverage during the surveys, a new stratified-version of SPLINTR had been developed, which removed all elements of the spatial version of SPLINTR. SC/62/IA12 reported that the difference between the model-based and stratified SPLINTR estimates were quite small: about 10% lower for CPII and 7% lower for CPIII. By comparing results from applying stratified-SPLINTR to the reference dataset and the authors' preferred dataset (with slight lengthening of transects as mentioned in 5.1.3 above), it was concluded that such modification of the data made only a small difference (about 2-3%) in estimated abundance. Finally, the authors also pointed out that since there had been no change to the model presented last year, the lack-of-fit seen in the perpendicular distance plots, where the fits failed to capture observed spikes in the data at very small distances, would still exist.

In discussion, some concerns were raised regarding the data omitted by the authors in producing these estimates. Spatial modelling demands more precise data collection and data errors – even in validated data – are frequently discovered when applying these models. The authors responded that the tiny proportion of data excluded were such cases where the data were internally inconsistent; the small differences in estimates from the OK model using the reference dataset and the standardised dataset suggested that the omissions would also only make a small difference to the estimates for SPLINTR too. In further discussion, it was commented that the stratified SPLINTR model had proved to be a very useful halfway-house for examining differences between spatial SPLINTR and OK estimates. It was suggested that consideration be given to proceed with these comparisons, as they may also be useful for comparisons of variance estimates between the two modelling approaches.

The sub-committee expressed its thanks to both sets of authors for producing estimates and presenting a substantial amount of new work this year. Work had progressed well collaboratively by e-mail in order to get to this point. As discussed at last year's meeting, the sub-committee was now in a position where with one set of estimates alone, neither of the models' performance in the simulations and the diagnostics would raise sufficient concerns to fail to accept the estimates, but the fact that the estimates themselves were so different was problematic.

5.1.5 Difference in abundance

The comparison of results from the reference dataset had not yet revealed why the two models were yielding such different estimates. Furthermore, the estimates from the SPLINTR model were slightly lower than those from the IWC 'standard' method (Branch, 2006). This result is somewhat surprising since the expectation was that by producing estimates which did not assume $g(0)=1$, the estimates would increase. However, the IDCR/SOWER Antarctic minke whale data are complex so that the interpretation of comparisons may not be straightforward, as there are several confounding effects which affect the abundance estimators in different ways (as was illustrated to a degree in the simulation results presented in SC/62/IA14). Nonetheless, it was **agreed** that comparing not only OK and (stratified) SPLINTR results, but also results from the 'standard' method may help to understand the reasons for the differences in the estimates and to develop new diagnostics to test the models.

In the light of the results in SC/62/IA3 and SC/62/IA12 and from the reference dataset comparisons, it was thought that the differences between the estimates were probably caused by differences in estimation of the sighting probability component of the models. This represented some progress in elucidating the reasons for the differences in the estimates; results from SC/62/IA12 suggested that spatial modelling would only explain a difference of about 10% for example. The models adopt different approaches to estimating mean school size. The OK model uses a parametric form based on stratum and distance from ice, whilst SPLINTR estimates a spatial school size surface. As a result, there was some variation in the stratum-by-stratum estimates of mean school size between the two models but overall, mean school size estimation was not thought to be the root cause of the observed differences. Based on a preliminary examination of *g*(0) estimates in the models and rough empirical estimates from the data, the sub-committee **agreed** that *g*(0) is one key area on which to focus intersessional investigations. It was noted that the BT-NSP data collected on recent SOWER cruises were directly relevant (Burt *et al.*, 2009).

Considering the difference in estimates between the OK and SPLINTR models, there are three – not necessarily unrelated – issues: the scientific question of pursuing the work to model the data and finding reasons for the difference in the estimates; the implications, if any, for future survey programmes; and the procedural question of what the Scientific Committee should do in the interim (or if a solution to the first question proved impossible). The sub-committee has been instructed to plan for a RMP *Implementation* of Antarctic minke whales in 2015 (see 5.1.8), therefore it was important to have agreed absolute abundance estimates as well as indicators of change.

5.1.6 Reasons for differences between estimates from CPII and CPIII

Estimates from the OK, SPLINTR and standard method (Branch, 2006) were consistent in that they showed a decline from CP2 to CP3. Conclusions reached about the reasons for this change should integrate information from other sources such as changes in ice coverage during the survey periods concerned. Until recently, there was little quantitative information on the number of Antarctic minke whales that might be present within the pack ice but the sub-committee was pleased this year to receive several papers reporting on and analysing data from surveys of whales within this region.

5.1.6.1 REPORTS ON DISTRIBUTION OF SEA ICE

SC/62/IA4 investigated trends of sea ice in the period of IWC IDCR/SOWER circumpolar surveys from CPI to CPIII (1978-2004). The sea ice trends are fundamental information to understand the year-to-year sea ice variability. Trends in sea-ice-extent (sea-ice-field) in the IWC *Management Areas* were estimated by Murase and Shimada (2004) for data up until 2002. SC/62/IA4 extends the temporal coverage to also include 2003 and 2004. In addition, trends of sea-ice-area-in-the-sea-ice-field and open-sea-area-in-the-sea-ice-field are also considered. The trends in each IWC *Management Area*, western and eastern half of each IWC *Management Area* (e.g. Area I West and East) and each 10° degree longitudinal sector in January are investigated. Region-specific year-to-year sea ice variabilities were detected. The variabilities were not consistent even in the same *Management Area*. For example, differences of open-sea-area-in-the-sea-ice-field in two 10° degree longitudinal sectors in Area V (170°E-180° and 180°-170°W) between CPII and CPIII were large in comparison with the rest of Area V. These sectors correspond to the Ross Sea region where the difference in Antarctic minke whale abundance could be large. The number of Antarctic minke whales in the sea-ice-field is expected to be large if the open-sea-area-in-the-sea-ice-field is large. The difference in abundance estimates between the CPII and CPIII surveys can be partly explained by the change in open-sea-area-in-the-sea-ice-field. As recommended by Scientific Committee in the past, the authors recommend that further region-specific investigation is necessary to understand the reason why the abundance estimates are different between CPII and CPIII.

In welcoming this work, some members commented on slight difficulties interpreting the plots in SC/62/IA4, suggesting that polynyas might be included in the future. Murase noted problems defining polynyas from satellite data, and although it may be possible in areas where the vessels had surveyed in a particular year, this would not lead to a consistent series. Aside from potential measurement errors, definition of which were beyond the scope of the work undertaken, the material in SC/62/IA4 was considered valuable, and it was **agreed** that further investigations along these lines should continue.

Following the re-establishment of an intersessional Working Group on Abundance Estimates and Sea Ice Extent Changes at last year's meeting, SC/62/IA5 reported progress in preparing the sea ice data required to investigate the relationship between sea ice characteristics and Antarctic minke whale abundance estimates. The authors have made progress importing satellite sea ice data from Area II into GIS. The data include: the coastline of Antarctica; satellite sea ice data; IDCR/SOWER stratum boundaries, sighting data and effort data; days after sea ice melting; and days covered by sea ice. Imported sea ice data in Area II suggested that realisation of sea ice conditions in geographically complex regions such as the Weddell Sea in Area II and the Ross Sea in Area V is difficult because surveys proceeded following retreating ice to the south, in addition to longitudinal directions. In such cases, the authors consider that the use of average sea ice data during the survey periods is an alternative solution. They also report that the work of the intersessional Working Group established last year is now not expected to be completed until next year's meeting.

A review of papers submitted to the Scientific Committee since 2001 relating minke whale densities to sea ice characteristics included in SC/62/IA5 was welcomed. The sub-committee **recommended** that the important exercise of ice data preparation be continued in time for next year's meeting. The exact nature of any models relating minke whales densities in open water to those in the ice was not discussed, but it was **agreed** that investigation of the relationships between whale density and ice characteristics was an area worth pursuing.

5.1.6.2 REPORTS AND ANALYSES FROM AERIAL SURVEYS

This year, the sub-committee was pleased to receive reports (SC/62/IA8 and SC/62/O15) from two aerial survey programmes: the Australian East Antarctic programme (which co-ordinated in 2009/10 with the SOWER survey) using a fixed wing plane; and the German programme surveying the area in the Weddell Sea from a helicopter launched from the ice breaker vessel, the *Polarstern* (which was also used as a platform of opportunity for cetacean sightings). These programmes are some of the first attempts to gather quantitative data to estimate densities of minke whales in the pack ice. Preliminary analyses from each programme are also now available (SC/62/IA9 and SC/62/IA13).

SC/62/IA8 detailed a full-scale, double-platform aerial survey for Antarctic minke whales which was conducted in East Antarctica in the 2009/10 austral summer. The survey targeted polynyas within pack-ice between 93° and 113°E between mid-December, 2009 and early February, 2010. The aim of the aerial survey was to collaborate with a concurrent IWC-SOWER voyage surveying north of the ice edge, and to collect environmental information to study the distribution of minke whales within pack-ice environments. The 2009/10 aerial survey was conducted in three phases: the first phase repeated a survey design from the previous summer period, based in and around Vincennes Bay; the second phase moved survey effort over to the Shackleton Ice Shelf and the Davis Sea; and the final phase repeated the Vincennes Bay survey, but also extended transects around 40n.miles north of the sea ice boundary. In total, 4,923n.miles of effort was achieved, covering around 55,559n.mile^2 of survey area. Across the entire survey period there were 24 on-effort sightings (34 individuals) of minke whales; 5 sightings (5 individuals) of 'like' minke whales; and 5 sightings (5 individuals) of minke whales observed off-effort. Other species sighted were killer whales, southern right whales, sperm whales, southern bottlenose whales and a number of sightings of unknown species. Of note was the absence of humpback whale sightings north of the sea ice edge, despite the concurrent IWC-SOWER voyage (SC/62/IA1) counting many such whales. Given humpback whales are generally conspicuous from the air, it is likely that such animals were not missed, but that there were not present on the dates the aerial survey targeted areas north of the sea ice edge. Although no direct overlap in space and time was achieved between the aerial survey and the IWC-SOWER voyage, there was around 11,900n.mile^2 overlap where both programmes surveyed within 14 days of each other.

Although no synoptic coverage with the SOWER vessel had been achieved, Bravington commended the fact that the aerial and shipboard surveys had surveyed part of the same area within a short period of time, during two weeks when the ice conditions did not change substantially. The sub-committee **agreed** that the collaboration had been highly successful, both in collection of data, and in regular communications and data exchanges during the surveys.

A preliminary analysis of data collected during the 2009/10 aerial survey in east Antarctica (SC/62/IA8) was presented in SC/62/IA9. This analysis also included minke whale sighting data from a smaller-scale aerial survey undertaken in the austral summer of 2008/09 (Kelly *et al.*, 2009b). A basic MRDS analysis yielded estimates of relative densities for areas within both aerial surveys. A proper left-truncation of the sighting data was not applied in this preliminary analysis due to software limitations; the authors intend to investigate alternative truncation options in future analyses. In the Vincennes Bay area, relative density of minke whales in December 2008 was around 10 times that of densities observed in December 2009. There was also an intra-season increase in relative density of minke whales in Vincennes Bay: estimated relative density of minke whales in the Vincennes Bay in late January-early February 2010 was 2-4 times higher than in December 2009 (based on point estimates). Densities of minke whales were higher in the north of the Davis Sea as compared to the south. It may be that pack-ice dynamics and the relative position of the shelf-break (krill habitat) are influencing inter- and intra-summer densities of minke whales across the aerial survey study area; as these analyses are preliminary, such inferences are highly speculative at this time.

Although preliminary in nature, the authors also commented that the figures shown in SC/62/IA9 suggested a fairly high and consistent 'recapture' probability out to 600m. They anticipated that the data would therefore provide a good idea of the proportion of whales available to be seen but were in fact missed. The sub-committee welcomed this work and looked forward to receiving an updated analysis next year.

A brief discussion was held concerning the plans for future analyses of minke whale sightings from the aerial survey described in SC/62/IA8. The basic research question upon which these analyses will be based is whether, using such aerial survey data, a number of minke whales can be found in pack-ice in Area IV-East that is able to account for decreases in numbers estimated by either SPLINTR (SC/62/IA12) or OK (SC/62/IA3) abundance estimation methods. Presently, there is no satisfactory data to estimate availability of Antarctic minke whales for an aerial survey, but some estimates do exist for common minke whales that could, at least, provide a lower bound. A more informal or indirect way of estimating availability would be to compare the uncorrected estimates of abundance from the aerial survey data to IWC-SOWER abundance estimates from the same region. The aerial survey data will be analysed within a MRDS framework and combined with the rough estimates of availability. This distance analysis will then form the foundation of a spatial model of minke whale density based on geographical and sea ice variables. It is hoped such a model will aid in model-based abundance estimates of minke whales in pack-ice in East Antarctica.

SC/62/O15 reported on two helicopter and shipboard cetacean surveys conducted in the Weddell Sea from the German research vessel *Polarstern*, in 2006/07 and 2008/09. Helicopter tracklines covered a total of 13,124km and 13,417km in 2006 and 2008 respectively, while the ship survey covered 1,171km and 2,011km respectively. Minke whales were primarily observed in the ice. Killer and southern bottlenose whales were also seen in the ice while all baleen whales (other than minke whales and sperm whales) were only observed in open water. Humpback whales were the most frequently sighted species on the shipboard survey in 2006/07 and the helicopter survey in 2008/09. Environmental information, including the proportion of ice coverage, was collected continuously. One striking finding was a much higher encounter rate for all cetaceans from the ship than from the helicopter. The authors consider two possible explanations for this difference: (1) observers on the helicopter missed more animals on the trackline than those from the ship (i.e. $g(0) < 1$); and (2) whales were drawn to leads created while *Polarstern* broke the ice.

SC/62/IA13 reported preliminary results from helicopter surveys for Antarctic minke whales and sea ice, conducted from *Polarstern* in 2006/07 and 2008/09. The cruise plan (described in SC/62/O15) was designed to achieve other research and logistical objectives in conjunction with reprovisioning the German Antarctic base at Neumayer; therefore placement of helicopter tracklines around the ship's cruise track was designed to sample across as wide a range of ice conditions as possible. In the first year of the study, the survey covered the ice edge in the Weddell Sea, the areas formerly occupied by Larsen A and B, and western Antarctic Peninsula, and only a line from Cape Town to Neumayer in the second. The survey design was further constrained by a maximum flight time of 2 hours at 80 knots, and the majority of surveys were therefore squares of 40n.miles (74km) on a side and placed on a diagonal to sample as far as possible from this ship's track. The resulting survey yielded over 13,000km of dedicated trackline effort from each of two years, and 94 on-effort sightings of Antarctic minke whales in the two years combined. The helicopter served as an excellent, highly manoeuvrable platform, which allowed observers to hover immediately after each sighting to confirm species identity, take photographs, and to confirm school size. Only single platform data collection are available, so data were analysed assuming $g(0)=1$. Available environmental covariates included ice concentration (observed visually along the trackline, and inferred from satellite imagery) and distance to ice edge (defined as a smooth line joining all grid cells in which ice concentration was greater than or equal to 15%). Ice covariate data were calculated for each segment and each sighting on the corresponding day of the survey. A classification tree indicated that the first split occurred in the data at 143km from the ice edge: i.e. most whales that were seen farther from the ice edge than this were humpback and fin whales. The second split occurred at 5% ice concentration (as observed along the trackline): almost all whales seen in ice concentrations >5% were Antarctic minke whales. A tentative analysis was reported that used the Density Surface Modelling engine in Distance (Thomas *et al.*, 2010) using the 'count' method (Hedley *et al.*, 1999). The density surface model was used to predict animal density throughout the study area based on ice concentrations on three days representing the beginning, middle and end of the two survey periods. The model showed highest density of minke whales in a narrow band of modest ice concentration (approximately 5-20%), but reanalysis is required to put robust bounds on this band to infer habitat preference. Next steps include plans to reanalyse the data using soap-film smoothers (Wood *et al.*, 2007), error distributions that are robust to unmodelled overdispersion in the data, and new methods developed by Hedley and Bravington to propagate the variance from the model through to the resulting abundance estimate (Hedley *et al.*, in press). The authors emphasised that they invite collaboration with their Australian colleagues doing aerial survey work.

Kelly agreed that these data, in combination with those from the Australian aerial survey, would help to make useful inroads into questions related to densities of minke whales in the ice, expressing interest in collaborative analyses. The sub-committee thanked all the authors for their work, and extended their appreciation to the governments of Australia, Germany and the Netherlands for supporting this research.

5.1.7 Planning for a RMP Implementation *as proposed in the SAG report*

The sub-committee Chair appointed a small group under Butterworth to consider this item. Members were Bravington, Cooke, Hakamada, Hatanaka, Kelly, Pastene and Walløe (Observer). As there was no time for the group to meet during the sub-committee period, it was **agreed** that they would report their conclusions during the Scientific Committee Plenary.

After some discussion of the relative priorities of In-depth and *pre-Implementation assessments* of North Pacific sei whale and of Antarctic minke whales, the timeline given in the SAG report was **agreed**.

5.1.8 Work plan

The sub-committee proposed that the following work be completed intersessionally:

(1) continue the work to evaluate the reasons for difference in estimates from the OK and SPLINTR models; and
(2) continue to address reasons for the differences between CPII and CPIII Antarctic minke whale estimates, by investigating:
    (a) the relationship between whale density and days after sea-ice melt; and
    (b) the relationship between estimates of abundance and sea ice characteristics.

The Intersessional Working Group on Abundance Estimates was re-established in order to make progress with item (1) above. The sub-committee **recommended** that the programme of work detailed in Appendix 3 be completed. Last year, plans had been made to hold an intersessional workshop to expedite progress on this item; in the event, a suitable date for the workshop could not be agreed and so work was carried forward to a pre-meeting at SC/62. Both sets of modellers have committed to the timeline set out in Appendix 3; note that a workshop, to be held by February 2011 at the latest, is considered to be essential for satisfactory completion of item (1). For comparative purposes, it was also **agreed** that estimates – and as applicable, diagnostics – from the IWC 'standard' method should be included in the model evaluations. The simpler formulation of this method made its behaviour easier to understand; this was expected to be useful when considering the model output from OK and SPLINTR.

With regard to putative relationships between minke whale density and sea ice, the sub-committee has received several papers over the course of the past decade. The work identified in item (2) above represents exploratory work to examine these relationships in a quantitative framework. It requires the preparation of sea ice and other environmental data for model input, as well as estimates of minke whale abundance by 10° longitudinal slice. Bravington undertook to provide the latter using the SPLINTR estimates (since SPLINTR is a spatial model, it can more readily provide estimates by sub-area than stratified models). The data preparation and exploratory investigations would be carried out by Murase and Kitakado (see section 8 below). The sub-committee noted that there appeared to be some value in continuing to discuss the details of satellite sea ice data processing. Kelly agreed to cooperate with this.

5.2 Catch-at-age analyses

5.2.1 Report from the intersessional working group

SC/62/IA7 reported on activities during the past year. These are elaborated in the documents summarised in sections 5.2.2 and 5.2.3 below.

5.2.2 Age estimation

Lockyer enlarged on the Antarctic minke whale ageing exercise (SC/62/IA11) which she had carried out during the intersessional period in terms of the experimental design agreed by the Scientific Committee (IWC, 2009, p.209). This had involved readings of ear plugs from minke whales taken in the period 1974/75-2005/06, including both Antarctic commercial and JARPA samples. The primary aim of the work was to determine whether evidence exists of a drift in reader performance, and, if so, to quantify it. A secondary aim was to quantify age-reading error variability. Left ear plugs were selected only from females, and only from samples where a useable age had been achieved by

Japanese researchers. The experimental sample comprised 50 randomly taken ear plugs from each of 5 sub-sets totalling 250 ear plugs in all. The sub-sets were taken from Area IV in the periods 1974/75-1976/77, 1982/83-1984/85, 1989/90-1991/92, 1997/98-1999/2000 and 2003/04-2005/06, thus encompassing a 25-year time span. The ear plugs were selected by staff from the laboratory at the Tokyo University of Marine Science and Technology, under the supervision of Kitakado. The sample was numbered independently of all existing identifying marks for the first reading. The numbering was random for the entire set of 250 plugs, but the plugs were read in numerical order from 1-250. After completion of the first reading, the sample was reassigned new identifying numbers and re-ordered randomly. The ear plugs were then read again in numerical order 1-250. After the second reading was completed, a sub-set was randomly selected from the 250 set, but this time choosing 10 plugs from each time period, totalling 50 plugs in all. These were then read again.

During the reading procedure, the reader had no input or access to actual data pertaining to the sample, i.e. the plugs were read 'blind'.

The readings were undertaken within a three week period with approximately 50 ear plugs read daily using a Nikon binocular microscope to examine all ear plugs with an eye objective 10×B22 and zoom magnification ×0.8-×8 facility. A break period of two days was scheduled between the first and second, and then second and third readings, to minimise possible recognition of certain samples. The oldest whale examined was >60 GLGs and the youngest had no GLG visible (young of year). The impression was that ear plug size in general was very variable, and not always correlated with age. In addition, the early-forming GLGs were the most problematic to interpret, the pattern of deposition frequently appearing distorted and irregular, especially in old animals. For this reason, the source of error in ageing in old animals was thought likely to be mainly due to problems in the early GLGs. The late-forming GLGs were much easier to interpret, despite becoming more narrowly packed together, because they were usually regular in form. In addition, accessory laminae were sometimes present and confusing in young ear plugs. For this reason occasionally two possible alternative readings were provided because the reader could not be certain which to choose. Normally – though not in an experimental situation such as this – readers might refer to biological data to help resolve such issues.

Age readings were recorded in Excel data format, and the following data were recorded for each ear plug specimen at each of the three reading sessions to facilitate the subsequent analysis of the data: sequential specimen number, age readings including both weighted and simple mean ages of 5-10 counts, comments including whether the ear plug was intact, whether the neonatal line was present, whether the plug was cut centrally, colour and general size and appearance, and a readability rating from excellent, good, poor to unreadable when the age reported should be disregarded.

In discussion appreciation was expressed at the manner in which the experiment had been carried out to maintain independence of and a blind approach to readings as specified in the protocol; thanks were also expressed to Japanese graduate students who had assisted in the conduct of the experiment.

A recommendation by Lockyer that a standard reference set of minke earplugs be maintained for age-reading training purposes received support.

SC/62/IA2 explored the impact of period/reader on age-determination by three Japanese readers by comparing age-estimates from earplugs from a control reader (Lockyer) with age-estimates by the Japanese readers (Masaki, Kato and Zenitani). A total of 250 plugs selected according to the predetermined protocol (IWC, 2009) were used in the analyses (see SC/62/IA11 for details). A conditional distribution of an observed age given a true age was defined to estimate the extent of ageing error for two groups of readers. Parameters determining ageing error matrices were estimated using a maximum likelihood method under several scenarios regarding the bias of the control reader. The analysis showed that incorporating a reader effect into the variance component to quantify the extent of random age-reading error improved the goodness of fit substantially (in terms of model selection criteria) compared to incorporating these effects into the mean structure. A model with reader effects in both the mean and variance structures provided the most parsimonious fit to the data among the models investigated. The period effect models tended not to fit the data as well as expected because of two readers within one period.

Overall, the results demonstrated that the Japanese readers and the control reader differed in terms of both expected age given a true age and variance in age-estimates. The results also suggested that the expected age and random uncertainty in age-estimates differed among the Japanese readers although the differences were not severe. This work could contribute to how catch-at-age data are used in the statistical catch-at-age analyses and in future virtual population analyses. The authors of SC/62/IA2 expressed their appreciation to Japanese scientists for allowing them to access past Japanese age-estimates through Procedure B.

This study was welcomed by members and seen as an important step forward. The comparisons indicated a few outliers at large ages, but it was noted that the models fitted assumed variance to increase with age so that such instances would not impact estimates of inter-reader bias greatly, and further that in practice age readings also took account of auxiliary information about the animals which would tend to diminish the proportion of such outliers. It was noted that Lockyer tended to report greater ages than the Japanese readers, but that differences amongst the Japanese readers was slight, and that there was no indication of a trend in bias in Japanese readings over the period from the commencement of commercial takes of Antarctic minke whales to recent years of whaling under research permit.

It was pointed out that while SC/62/IA11 makes a valuable contribution, it does not provide any information about the accuracy of the age readings in absolute terms, given that none of the ear plugs come from known-aged individuals. In response, comments were made that the absence of known-aged individuals was the norm for fish populations generally. However, for a number of fish populations there were indications from seasonal studies that layers were seasonal. Similarly, studies of fin whales, as well as corpora counts and animals with known histories indicated that the growth layers counted to age whales were laid down annually. Best pointed out that in the absence of known-aged individuals, the use of corpora counts from the same whales would provide an independent estimate of relative age, from which possible age-related biases in ear plug reading could be investigated.

It was **recommended** that guidelines for dealing with stranded animals include encouragement to obtain samples which could provide information on the animal's age.

5.2.3 Analyses using modified catch-at-age data

In SC/62/IA6, Punt examined the impact of allowing for ageing error based on the analyses of the age-reading experiment (SC/62/IA2) when conducting assessments for Antarctic minke whales in Areas III-E, IV, V and V-W using statistical catch-at-age analysis (SCAA) by means of sensitivity tests. These sensitivity tests explored three scenarios: (a) no ageing error; (b) ageing error is modelled as in previous base-models; and (c) ageing error is based on the results from SC/62/IA2. Time-trajectories of total (1+) population size and recruitment were qualitatively the same, irrespective of how age-reading error was modelled.

In discussion it was noted that while estimates of recruitment and abundance for the three different assessments were close over recent years, absolute values showed relatively large differences over the 1930s and 1940s, though also noted that estimation variance would be expected to be much higher over this period.

The question was raised of whether the issue of age estimation error had now been adequately addressed, or rather more investigation was needed through further analyses of readings made by a group of readers during the 1983 minke whale ageing workshop. The sub-committee noted comments that the experimental reading exercise conducted recently by Lockyer was far more rigorous and reliable than the 1983 comparisons, and further that introducing age-reading error into the SCAA evaluation did not change population trends qualitatively. It consequently **decided** that no further experiments or analyses on age reading errors were necessary to resolve problems raised in the JARPA review.

This decision did not however imply that other issues associated with the data and analyses, such as reasons for the different length distributions at age for younger-aged commercial and JARPA catches, had been resolved. Further work needed is discussed in the following section.

5.2.4 Work plan

The following issues were identified as requiring attention before investigation of catch-at-age based assessments of Antarctic minke whales using SCAA might be considered to have been completed.

(1) Confirm satisfactory convergence of the SCAA estimator with the inclusion of the ageing error matrix now developed.
(2) Check whether the SCAA model together with its various estimated selectivity functions can account satisfactorily for the different length-at-age distributions for younger animals in the commercial and JARPA catches.
(3) Check the impact of possible misreporting of the length distribution by the USSR commercial fleet on the SCAA results, possibly by assuming these catches to have the same length distribution as contemporaneous Japanese commercial catches.
(4) Investigate the effect of alternative assumptions in regard to Lockyer's possible bias in the age reading experiment.
(5) Explore how useful it would be for the models to have independent age estimates from corpora counts, to investigate possible age-related biases.

It was noted that these investigations would require an extension of permission from Japan for use of their minke whale catch-at-age data, and also that the investigations would be improved if data from the most recent JARPA cruises could also be made available. The sub-committee **recommended** that such an approach be made to Japan under Procedure B. Kato indicated that corpora count data were available, and that these data would be provided if necessary.

The following intersessional steering group was nominated to co-ordinate this data application and also to oversee progress on the outstanding analysis issues identified above: Punt (convenor); Butterworth, Kitakado and Polacheck.

6. CRUISES

6.1 Results from the 2009/10 IDCR/SOWER field studies

The planning meeting for the 2009/10 IDCR/SOWER cruise was held in Tokyo, Japan in September 2009 (SC/62/Rep6). The meeting reviewed the Scientific Committee discussions at last year's meeting and noted that highest priority had been assigned to collaboration with the proposed Australian aerial survey (Kelly *et al.*, 2009a) and, in case the aerial survey could not continue as planned, priority should be given to humpback whale biopsy sampling and photo-id image collection. At the meeting the Australian Antarctic Division confirmed that the aerial survey was to continue as planned. The meeting welcomed this information and agreed the SOWER survey should be synchronised with the aerial survey and so the region between 100°E-115°E was selected as the research area for SOWER. This research area was similar to the two most recent IDCR/SOWER cruises. The meeting agreed that, as a contingency plan, humpback whale photo-id and biopsy work should take place in the southern stratum between 120°E-135°E.

The 2009/10 SOWER cruise was conducted in Area IV, aboard the Japanese research vessel *Kaiko Maru* (SC/62/IA1) and had two main objectives: to undertake a sightings survey in collaboration with an Australian Antarctic Division aerial survey, and to continue research on the priority species (southern right, blue, fin, and humpback whales) including biopsy/photo-id as well as identification of sub-species for blue whales. A total of 1,072n.miles were covered during two repeat surveys of the region (100°E-115°E and extending from the pack ice to 60n.miles north of the ice edge) and in two survey modes (SS-II and BT-Option II). A further 92n.miles of SS-II and BT-Option II effort was conducted between 100°E-102°E and then the vessel continued eastwards along the ice edge in BB mode.

The total number of minke whales sighted during the entire coverage of the research area was 83 groups, comprising 152 animals. Two concentrations of minke whales were encountered along the ice edge during BB mode. Humpback whales were the most frequently sighted species in the research area (174 groups comprising 322 animals). Biopsy samples and individual identification photographs were taken from 21 and 45 humpback whales, respectively. No blue whales were observed but five fin whales in three groups were sighted, two of these groups near the ice edge. Twenty-eight groups of southern right whales were sighted (comprising 38 animals) with biopsy samples from 22 animals and identification photos of 26 individuals. One mixed-species group, consisting of one southern right whale and one humpback whale, was photographed and biopsy samples were taken. Nine groups of killer whales (78 animals) were sighted, however, most groups did not show strong characteristics for any type, except one group of 20 animals that were identified as Type A. Experiments using a photogrammetric system, to measure angle and sighting distances, were planned but, due to missing equipment, were not performed.

Collaboration with the Australian Antarctic Division aerial survey had the highest priority for the survey. Some flights were carried out during the SOWER survey but the plane was never seen in the vicinity of the SOWER vessel. The weather was poor for most of the survey period and the SOWER vessel had 68% and 56% of off-effort time during the two repeat surveys.

The *Kaiko Maru* had not been used on SOWER cruises before and the bowdeck was lower and smaller, and the vessel sides were higher than the *Shonan-Maru No.2* which had been used in recent years. However, no difficulties were encountered when taking biopsy samples or identification photos of the target species (humpbacks and right whales). The Cruise Leader expressed her appreciation to the Captain and the crew of the *Kaiko Maru* for their cooperation throughout the survey.

The sub-committee thanked the Government of Japan for generously providing the vessel and crew for this survey, and also thanked the Cruise Leader for her efforts. Noting that this was the last IDCR/SOWER cruise, the sub-committee also extended its appreciation to all member nations who had contributed to this extensive programme, and particularly to the governments of Japan and the former Soviet Union, for providing the survey vessels. Furthermore, the sub-committee thanked all those who had been involved with the cruises, including the Steering Group, the Cruise Leaders, the researchers and the crews. The data collected during the program were an unparalleled source of information on Antarctic cetaceans. The experience gained from these surveys would also continue to be of use in planning future studies, in the Southern Ocean and elsewhere.

At the Scientific Committee meeting in Santiago in 2008, a Steering Group was formed to consider creating a Special Issue of the *Journal of Cetacean Research and Management* on the IDCR/SOWER surveys. No work has been reported from this Group to date; the sub-committee **agreed** that such a volume is still merited; Best agreed to discuss this with the Head of Science.

6.2 Plans for cetacean sighting surveys in the Antarctic in the 2010/11 season

SC/62/O17 described a dedicated, systematic cetacean sighting survey which was being planned to take place from December 2010 to February 2011 in order to obtain estimates of abundance for use in the RMP. The research area will be south of 60°S in Area V and the western part of Area VI (130°E-145°W), including the Ross Sea. This survey will be conducted in relation with the Japanese Whale Research Program under special permit in the Antarctic (JARPA II). Two dedicated, sighting survey vessels, *Shonan-Maru No.2* and *Yushin-Maru No.3*, will be used and the survey procedures are planned to be based on the standard SOWER search modes; closing (NSC) mode and passing with the independent observer (IO) mode. Distance and angle estimation training, as well as some experiments, will be conducted. Abundance of Antarctic minke whales will be estimated using analysis methods being developed by members of the sub-committee. Biopsy skin sampling of blue, fin, humpback, southern right, and sperm whales will be opportunistically collected for assessing stock structure. Photographs for identification studies of large cetaceans, such as blue, southern right and humpback whales, will also be taken. Researchers will record the data (weather, effort, sighting and experiments data) using the on-board computer during the survey. These data will be validated at the Institute of Cetacean Research and submitted to the IWC Secretariat based on the IWC Scientific Committee Guidelines. A planning report will be prepared by Japan and a Cruise Report, prepared by the Japanese researchers, will be submitted to next year's meeting.

During discussions, the sub-committee reflected on its current difficulties interpreting estimates of Antarctic minke whale abundance from the IDCR/SOWER surveys, and that as far as possible, the lessons learned from those surveys – and their ongoing analyses – should be used to improve surveys and data analyses in the future. The Ross Sea is a particularly difficult region to survey, since it is large in area and has a complex and rapidly changing ice configuration. Two potential issues relevant to surveys in this Area relate to the spatial and temporal coverage in the region. In terms of the spatial coverage, it is important to attempt to design tracklines which give approximately even probability of coverage within a stratum, particularly if the intent of the analysis is design-based rather than model-based. It was suggested that historical and predictive sea ice maps may be useful indicators of what the survey area might be in 2010/11, and furthermore, that the 'survey design' component of the *Distance* software could be used to examine different trackline placements and survey region definitions. In terms of the temporal coverage, the concern is that if this is not considered carefully then any resulting estimates from the survey would be subject to the criticism that whales may have been double-counted. To avoid this, it is important to either survey the whole area 'synoptically' (i.e. over a sufficiently short period of time that whale movement and changing ice conditions are not significant), or to survey the relevant strata multiple times.

Based on its utility in the analyses of the IDCR/SOWER data, the sub-committee also **recommended** that instead of normal Closing Mode, SSII mode (closing-when-abeam) be adopted on this cruise. Other considerations for change in data collection from SOWER were also made. These included the collection of data to allow duplicate identification algorithms and measurement error models to be applied. Accurate sighting times and independent estimates of group size may be helpful in this regard.

In order to minimise difficulties associated with survey design, an intersessional Working Group was established under Matsuoka (also comprising Bravington, Ensor, Hedley and Kitakado). Matsuoka would prepare an intersessional report from this group which would also form a planning report as no planning meeting was scheduled.

The sub-committee **agreed** that Matsuoka would be responsible for IWC oversight.

6.3 Plans for cetacean sighting surveys in the North Pacific

6.3.1 IWC organised sighting surveys

During the last year's Scientific Committee meeting, Japan presented a proposal for a medium- to long-term research programme involving sighting surveys to provide information for cetacean stock management in the North Pacific. The Scientific Committee welcomed the initiative and agreed the value of a large-scale, medium-long term integrated research programme in the North Pacific and encouraged this in the context of international collaboration under IWC auspices. The Scientific Committee recommended that the planning process should start with a review of the current discussions on North Pacific issues within the Committee and a careful examination of available information and identification of gaps in knowledge.

A meeting to discuss the North Pacific survey programme was held in Japan in September, 2009 (SC/62/Rep3). The meeting noted four terms of reference:

(1) review the Scientific Committee's issues in the North Pacific and circulate a paper before SC/62;
(2) review the past and ongoing survey activities and available data in range states from completed *pro formas*;
(3) consider possible line transect survey plans and additional data collection (e.g. photo-id and biopsy) for the 2010 season; and
(4) prepare a proposal for an intersessional Workshop (to be held between SC/62 and SC/63) on future surveys beyond 2011.

The meeting reviewed previous Scientific Committee discussions regarding the proposal for a research programme in the North Pacific to provide information for stock management. The meeting agreed the priorities and cruise plan of the survey to be held in 2010, as well as considering the medium- to long-term objectives of such a programme.

SC/62/IA15 was provided in response to the first term of reference from the meeting and provided a summary of the Scientific Committee issues relating to North Pacific sei, common minke, Bryde's, right and blue whales. The distributions of these whale species were described and requirements for further surveys, in order to estimate abundance and investigate stock structure, were considered.

SC/62/IA10 presented the research plan for an IWC/Japan whale sighting survey taking place in summer 2010. The plan had been drawn up following guidelines agreed at the North Pacific programme intersessional meeting. The research area (170°E-170°W) had been chosen because for some species it spans proposed stock boundaries and has been poorly covered by previous surveys, representing an important information gap for several large whale species. The cruise will collect line transect data, to estimate abundance, and biopsy/photo-id data contributing to the work of the Scientific Committee on the management and conservation of populations of large whales in the North Pacific. It will provide:

(1) information for the proposed future in-depth assessment of sei whales in terms of both abundance and stock structure;
(2) information relevant to *Implementation Reviews* of whales (e.g. common minke whales) in terms of both abundance and stock structure;
(3) baseline information on distribution and abundance for a poorly known area for several large whale species/populations, including those that were known to have been depleted in the past but whose status is unclear; and
(4) biopsy samples and photo-id images to contribute to discussions of stock structure for several large whale species/populations, including those that were known to have been depleted in the past but whose status is unclear.

The cruise will last a total of about 60 days (including transit time) between July and August. In order to adequately cover the longitudinal range, the latitudinal range is restricted between a southern boundary at 40°N and a northern boundary at the Aleutian Islands chain. This region will allow for sufficient coverage and be expected to incorporate the latitudinal range of sei whales at that time of the year. Based on experience elsewhere in the North Pacific, allowing for poor conditions and time for photo-id and biopsy sampling work, an average of 65n.miles is expected to be covered per day in primary searching effort (12 research hours per day). The research vessel *Kaiko Maru* will be used and is equipped with a top barrel platform, IO platform and upper bridge. Biopsy sampling/photo-id work will be undertaken on priority species (North Pacific sei, common minke, right, blue, humpback and fin whales, with higher priority to the first two species). The Institute of Cetacean Research (ICR) data recording system will be used along with the data forms used on the SOWER cruise. The rules for data availability, shipping and storage will be as for the present SOWER cruise and IWC equipment will be used, if required. Copies of data, photographs etc. will be sent by ICR to the IWC Secretariat upon completion of the cruise. Records of all the data taken in US waters will be made available for unrestricted scientific research, including photographs and one-third of the sample from each biopsy sample collected in US waters. Four researchers can be accommodated on this cruise and US and Korean scientists will participate. The cruise will follow the requirements for reports and documentation developed for cruises that could provide data for use under the RMP and will be the responsibility of the Japanese scientists.

The sub-committee thanked the Government of Japan for its generous offer of a vessel for this survey. It was noted that the start and end points of the cruise track in the northern strata coincided with the US EEZ but it was confirmed that the start point of the trackline had been randomly generated and this was a coincidence. It was also noted that although 350 biopsy sample permits had been applied for, it was anticipated that this number would not be collected. Matsuoka was confirmed as cruise leader and assigned responsibility for IWC oversight.

The sub-committee **recommended** that the research objectives stated in SC/62/IA10, and listed as items (1)-(4) above, form the basis for planning a North Pacific survey in 2011. A working group under Kato was formed to discuss logistical details of this survey. The sub-committee endorsed the working group's report (given as Appendix 4), and **recommended** that the investigations regarding the use of Institutional permits to exchange biopsy samples proceed as soon as possible, with the results of the investigations being reported to the Planning Meeting (scheduled for October 2010).

Furthermore, the sub-committee **recommended** that the research objectives for the 2011 survey be taken forward and a coherent multi-year plan be developed for the survey programme. A Steering Group to oversee the IWC North Pacific surveys was established, convened by Kato, with the following members: An, Brownell, Clapham, Donovan, Ensor, Matsuoka, Miyashita, Murase, Pastene and Wade. It was proposed that a meeting of the Steering Group should be scheduled immediately prior to the Planning Meeting for the 2011 cruise, in order to develop the programme of research to be undertaken over the next few years.

6.3.2 Japanese sighting surveys

SC/62/O16 described two sighting surveys for cetaceans, taking place in the North Pacific in 2010, to examine the distribution of sei, Bryde's and minke whales and to estimate abundance. Both surveys are in the middle part of the western North Pacific. The first survey will take place from June to July in the region 35°N-40°N and 157°E-170°W, and the second survey will take place from July to August in the region 32°N-37°N and 145°E-180°. The main target species are sei and minke whales for the first survey and Bryde's whale for the second survey. The research vessel *Yushin-maru No.3* will be used for each cruise. Distance and angle estimation training and experiments will be conducted for abundance estimation. Sighting data will be analysed to obtain estimates of abundance for use in the RMP. Biopsy skin samples from large whales, such as blue, fin, sei, Bryde's, minke, humpback, right and sperm whales, will be opportunistically collected for assessing stock structure. Photo-id of large cetaceans, such as blue, right and humpback whales, will be also conducted. The cruise report will be submitted to next year's meeting.

During discussion it was confirmed that there were no plans to estimate the probability of detection on the trackline ($g(0)$) as the main focus of the survey were sei and Bryde's whales and so $g(0)$ would be expected to be close to one.

The sub-committee assigned responsibility to Matsuoka for IWC oversight.

7. PROGRESS TOWARDS AN IN-DEPTH ASSESSMENT OF NORTH PACIFIC SEI WHALES

7.1 Review of information

The available information was summarised in last year's report (IWC, 2010, pp.196-97). There is no new processed information available this year, but field work has continued: scientific catches (100 sei whales sampled out of 386 sighted) and associated sighting vessels (120 sei whales sighted) under JARPN II; and sightings surveys conducted independently by NRIFSF/Japan, resulting in effort with no sightings (SC/62/ProgRepJapan).

The sub-committee reviewed the available information with a view to assessing the likely amount of work involved in the In-depth Assessment and subsequent *pre-Implementation assessment*.

7.1.1 Stock structure

Analyses of 489 genetic samples from JARPN II catches and 301 samples from former Japanese commercial pelagic catches indicate no evidence for stock structure over the areas sampled (Kanda *et al.*, 2009). The eastern boundary of the JARPN II area is 170°E, but the commercial samples extend across the northern North Pacific to 135°W. If this result is confirmed, the area that has been genetically sampled would be regarded as containing a single stock, but the possibly of additional stocks in other areas cannot be excluded. The precautionary approach would be to entertain both hypotheses:

(1) whales in unsampled areas are from the same stock as the sampled area; and
(2) whales in unsampled areas are from different stocks to the sampled area.

The simplest way to cover both eventualities would be to designate the area with information as a single *Small Area*, but to apply catch capping at the *Medium Area* level. With the current state of knowledge, the entire North Pacific region would be a *Medium Area*, but if evidence of stock segregation is found in other areas later, the latter areas would be excluded from the *Medium Area*. Areas with no genetic information would be designated as *Residual Areas*.

7.1.2 Catch history

The sub-committee agreed last year to use the same division of past catches between sei and Bryde's whales as has been used for the western North Pacific Bryde's whale assessment. Allison reported that work is continuing on Soviet catch data in the North Pacific, and that the catch history should remain open until this work is completed.

7.1.3 Abundance estimates

In order to avoid potential double-counting arising from migrations, abundance estimates across the region should refer to a specific time of year. The sub-committee agreed that estimates should be prepared for two time periods: (i) May-June; and (ii) July-September, and that the decision as to which period to use for the primary abundance estimates would be taken later. Currently, abundance estimates are only available for the JARPN II area (North Pacific north of 35°N and west of 170°E). As noted last year, older abundance data are available for pelagic areas to the east of 170°E, but these would probably not be suitable for use in the RMP. The US and Canadian surveys conducted in the Aleutians and Alaskan waters, along the western coasts of Canada and the USA and in the ETP have yielded very few sei whale sightings. As discussed under item 6.3, new sightings surveys to be conducted in the North Pacific in 2010, and IWC-coordinated sightings surveys planned from 2011, may extend the area for abundance estimation into the offshore region east of 170°E over the next few years.

7.2 Plans for the assessment

The sub-committee agreed that unless new information obtained in the near future indicates a more complex picture, the options for RMP *Implementation* will remain relatively simple as outlined above, and that it is unlikely that *Implementation Simulation Trials* will be required. The sub-committee therefore agreed that the In-depth Assessment (IDA) and the *pre-Implementation assessment* (PIA) should be combined into a single exercise.

The timetable proposed in the SAG report and the Chair's proposal (IWC/62/7rev) envisaged an IDA in 2011 followed by a PIA in 2013 and an RMP *Implementation* in 2015, but this timetable was primarily motivated by the need to spread the workload in the event that the Assessment and *Implementation* involved substantial extra work such as *Implementation Simulation Trials*.

In light of the more modest workload now envisaged, a range of views were expressed regarding the timing of the combined Assessment. Cooke and others noted that the combined Assessment and *Implementation* could be carried out with current information with relatively little additional work, and could be accomplished in 2011 or 2012. The new data to be gathered over the next few years would be considered in an *Implementation Review* to be scheduled a few years after the initial combined assessment. Hatanaka and others expressed a preference for the assessment to be conducted at a later date, when more abundance data covering a wider area are available.

After some discussion, the sub-committee recommended that the combined IDA/PIA be scheduled for 2013. If no *Implementation Simulation Trials* are required, the RMP *Implementation* could be completed the following year.

8. WORK PLAN AND BUDGET REQUESTS

The sub-committee agreed that completing the In-depth Assessment of Antarctic minke whales was its primary objective. It identified the following priority topics for next year's meeting:

(1) to resolve the reasons for the differences between estimates of abundance of Antarctic minke whales between the OK and SPLINTR models, and thus provide agreed estimates of abundance at next year's meeting;

Table 2

Intersessional Working and Steering Groups, and their membership.

Group		Terms of reference	Membership
Abundance estimation methods (WG)	(i)	Run sensitivity tests on modified real datasets to understand differences between OK and SPLINTR (e.g. in terms of ESW, *g*(0), MSS, and the underlying sighting parameters).	Walløe (Convenor), Branch, Bravington, Butterworth, Cooke, Hedley, Kitakado, Okamura, Palka, Skaug, Wade.
	(ii)	As above, to understand differences between OK/SPLINTR and Standard.	
	(iii)	Following identification of the underlying statistics/parameters where big differences occur, develop ways to cross-check against empirical data.	
	(iv)	If necessary, design further simulation trials to test robustness.	
Catch-at-age analyses (WG)	(i)	Confirm satisfactory convergence of the SCAA estimator with the inclusion of the ageing error matrix.	Punt (Convenor), Butterworth, Kitakado, Polacheck.
	(ii)	Check whether the SCAA model together with its various estimated selectivity functions can account satisfactorily for the different length-at-age distributions for younger animals in the commercial and JARPA catches.	
	(iii)	Check the impact of possible misreporting of the length distribution by the USSR commercial fleet on the SCAA results, possibly by assuming these catches to have the same length distribution as contemporaneous Japanese commercial catches.	
	(iv)	Investigate the effect of alternative assumptions in regard to Lockyer's possible bias in the age reading experiment.	
	(v)	Request permission from Japan for continued use of the catch-at-age data, and for permission to use similar data from the most recent JARPA surveys.	
	(vi)	Explore how useful it would be for the models to have independent age estimates from corpora counts, to investigate possible age-related biases.	
IWC North Pacific Survey Planning (SG)	(i)	Identify medium and long term research objectives for the IWC-Japanese North Pacific surveys.	Kato (Convenor), An, Brownell, Clapham, Donovan, Ensor, Matsuoka, Miyashita, Murase, Pastene, Wade.
	(ii)	Develop a (multi-year) research plan to achieve these objectives.	
Survey design for 2010/11 Antarctic minke whale survey (WG)	(i)	Consider survey design with respect to spatial and temporal coverage, and ice extent, to reduce ambiguities in interpretation of abundance estimates.	Matsuoka (Convenor), Bravington, Ensor, Hedley, Kitakado.
	(ii)	Evaluate and consider adapting survey protocols to facilitate flexible and improved analyses.	

(2) to continue the development of the catch-at-age models of the Antarctic minke whales, including sensitivity tests to examine various assumptions regarding ageing errors and age-length keys; and
(3) to continue the examination of the differences between minke abundance estimated from CPII and CPIII, by further investigation of the relationship between sea ice and minke whale abundance.

Since the highest priority next year will be given to obtaining the abundance estimates of Antarctic minke whales using the IDCR/SOWER survey data, the sub-committee **recommended** that IWC funds be granted to support the work identified in (1) above (see also section 5.1.8; details are provided in Appendix 3). It was noted due to that the intersessional workshop planned for last year had not taken place, and that this had – to some extent – limited the progress that had been made. This year, no outstanding tasks were foreseen that would restrict progress in the same way, and it was expected that the proposed workplan represented the best possible solution for resolving the Antarctic minke whale abundance estimation issues by next year's meeting.

Considerable progress has been made intersessionally on analysing ageing errors with regard to the statistical catch-at-age modelling. The sub-committee **recommended** that the modelling work identified in item (2) continues (see also section 5.2.4), and proposed that the relatively small budget request for this be granted.

The sub-committee has received preliminary investigations suggesting that the apparent decline in Antarctic minke whale abundance from CPII to CPIII may be attributed, at least in part, to changes in sea ice conditions. Further work on this is planned in task (3) above. This work requires a substantial amount of data preparation before any analyses can proceed. The sub-committee supported the funding request for preparing these data, and **recommended** that the analyses outlined be conducted.

As for the preceding SOWER surveys, the sub-committee also **recommended** that funding be provided to enable the 2009/10 IDCR/SOWER survey data to be imported into DESS.

This year, the sub-committee received plans for an IWC-Japanese collaborative North Pacific survey, to form part of a multi-year programme in this region. The sub-committee **recommended** that the funds requested to support this survey be granted, as the survey will, *inter alia*: (i) aim to provide abundance estimates of, and collect genetic samples from, sei whales, which would inform the proposed In-depth Assessment; and (ii) collect data on North Pacific right whales in an area of known historical depletion.

Table 2 shows the intersessional Working and Steering Groups established by the sub-committee.

9. ADOPTION OF REPORT

On behalf of the sub-committee, Kato thanked the Chair and co-Chair for their faithful, thoughtful and calm chairmanship. The Chair expressed his thanks to the sub-committee members for their cooperation and to the rapporteurs for their efforts. The report was adopted at 12:00 on 7 June 2010.

REFERENCES

Branch, T.A. 2006. Abundance estimates for Antarctic minke whales from three completed sets of circumpolar surveys. Paper SC/58/IA18 presented to the IWC Scientific Committee, May 2006, St. Kitts and Nevis, West Indies (unpublished). 28pp. [Paper available from the Office of this Journal].

Branch, T.A. and Butterworth, D.S. 2006. Suggested options for the analysis of IDCR/SOWER minke whale data. Paper SC/58/IA19 presented to the IWC Scientific Committee, St. Kitts and Nevis, June 2006 (unpublished). 10pp. [Paper available from the Office of this Journal].

Bravington, M.V. and Hedley, S.L. 2009. Antarctic minke whale abundance estimates from the second and third circumpolar IDCR/SOWER surveys using the SPLINTR model. Paper SC/61/IA14 presented to the IWC Scientific Committee, June 2009, Madeira, Portugal (unpublished). 25pp. [Paper available from the Office of this Journal].

Burt, M.L., Ensor, P. and Borchers, D.L. 2009. Detection probability of Antarctic minke whales: analyses of the BT mode experiments conducted on the IWC-SOWER cruises 2005/06-2007/08. Paper SC/61/IA18 presented to the IWC Scientific Committee, June 2009, Madeira, Portugal (unpublished). 14pp. [Paper available from the Office of this Journal].

Hedley, S., Buckland, S.T. and Borchers, D.L. 1999. Spatial modelling from line transect data. *J. Cetacean Res. Manage.* 1(3): 255-64.

Hedley, S. and Ensor, P.H. 2006. A preliminary look at minke whale dive time data from the 2004/2005 IWC-SOWER survey. Paper SC/58/IA21 presented to the IWC Scientific Committee, May 2006, St, Kitts and Nevis, West Indies (unpublished). 7pp. [Paper available from the Office of this Journal].

Hedley, S.L., Bannister, J.L. and Dunlop, R.A. In review. Abundance estimates of Breeding Stock 'D' humpback whales from aerial and land-based surveys off Shark Bay, Western Australia, 2008. *J. Cetacean Res. Manage. (special issue)*.

IWC. 2009. Report of the Scientific Committee. Annex G. Report of the Sub-committee on In-depth Assessment (IA). *J. Cetacean Res. Manage. (Suppl.)* 11: 193-211.

International Whaling Commission. 2010. Report of the Scientific Committee. Annex G. Report of the Sub-Committee on In-Depth Assessments. *J. Cetacean Res. Manage (Suppl.)* 11(2):180-97.

Kanda, N., Goto, M., Yoshida, H. and Pastene, L.A. 2009. Stock structure of sei whales in the North Pacific as revealed by microsatellite and mitochondrial DNA analyses. Paper SC/J09/JR32 presented to the Expert Workshop to Review Results of JARPN II, 26-30 January 2009, Tokyo, Japan (unpublished). 14pp. [Paper available from the Office of this Journal].

Kelly, N., Peel, D., Bravington, M.V. and Gales, N. 2009a. A planned aerial survey for minke whales in east Antarctica during summer 2009/10. Paper SC/61/IA4 presented to the IWC Scientific Committee, June 2009, Madeira, Portugal (unpublished). 5pp. [Paper available from the Office of this Journal].

Kelly, N., Peel, D., Pike, D., Bravington, M.V. and Gales, N. 2009b. An aerial survey for Antarctic minke whales in sea ice off east Antarctica: a pilot study. Paper SC/61/IA3 presented to the IWC Scientific Committee, June 2009, Madeira, Portugal (unpublished). 14pp. [Paper available from the Office of this Journal].

Murase, H. and Shimada, H. 2004. Possible impact due to variability of sea ice condition on minke whale abundance estimation in the Antarctic, in 1978 to 2004. Paper SC/56/IA10 presented to the IWC Scientific Committee, July 2004, Sorrento, Italy (unpublished). 15pp. [Paper available from the Office of this Journal].

Okamura, H. and Kitakado, T. 2009. Abundance estimates and diagnostics of Antarctic minke whales from the historical IDCR/SOWER survey data using the OK method. Paper SC/61/IA6 presented to the IWC Scientific Committee, June 2009, Madeira, Portugal (unpublished). 58pp. [Paper available from the Office of this Journal].

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A. and Burnham, K.P. 2010. Distance software design and analysis of distance sampling surveys for estimating population size. *J. Appl. Ecol.* 47: 5-14.

Wood, S.N., Bravington, M.V. and Hedley, S.L. 2007. Soap film smoothing. *J. Roy. Stat. Soc.* 70: 931-55.

Appendix 1

AGENDA

1. Election of Chair and co-Chair
2. Appointment of rapporteurs
3. Adoption of Agenda
4. Documents available
5. Antarctic minke whales
   - 5.1 Abundance
     - 5.1.1 Report from the intersessional e-mail Working Group
     - 5.1.2 Results from simulated datasets
     - 5.1.3 Comparison of OK and SPLINTR using the reference dataset
     - 5.1.4 Results from each method using their preferred dataset
     - 5.1.5 Abundance
     - 5.1.6 Reasons for differences between estimates from CPII and CPIII
       - 5.1.6.1 Reports on distribution of sea ice
       - 5.1.6.2 Reports from aerial surveys
     - 5.1.7 Planning for a RMP *Implementation* as proposed in SAG report
     - 5.1.8 Work plan
   - 5.2 Catch-at-age analyses
     - 5.2.1 Report from the intersessional working group
     - 5.2.2 Age estimation
     - 5.2.3 Analyses using modified catch-at-age data
     - 5.2.4 Work plan
6. Cruises
   - 6.1 Results from the 2009/10 IDCR/SOWER field studies
   - 6.2 Plans for cetacean sighting surveys in the Antarctic in the 2010/11 season
   - 6.3 Plans for cetacean sighting surveys in the North Pacific
     - 6.3.1 IWC organised sighting surveys
     - 6.3.2 Japanese sighting survey
7. Progress towards an In-depth Assessment on North Pacific sei whales
8. Work plan and budget requests
9. Adoption of Report

Appendix 2

STANDING ON THE SHOULDERS OF GIANTS - THE COMBINED OK-SPLINTR ESTIMATOR

Hans J. Skaug

Summary: It is shown that the average of OK and SPLINTR performs better on the simulated data than any of the estimators individually.

Introduction

At the current Scientific Committee meeting we are considering OK and SPLINTR as competing estimators of Antarctic minke whale abundance. Both estimators account for a large number of features of the IDCR/SOWER data, but nevertheless they give very different results.

Both methods have been subject to extensive simulation testing (SC/62/IA14), and it was found that there is much random variation between the estimators on individual simulation replica, in addition to some systematic difference. This suggests that averaging the two estimators may yield improved statistical properties.

Methods and results

The OK and SPLINTR results for the simulated datasets (school density, 100 replica of each of 54 simulation scenarios) were obtained from the author of SC/62/IA14. Let $d_{ij}^{(OK)}$ and $d_{ij}^{(SPL)}$ be the estimated school density for OK and SPLINTR, respectively, in simulation replica *j* of scenario *i*. Fig. 1 shows the within-scenario correlation r_i between $d_{ij}^{(OK)}$ and $d_{ij}^{(SPL)}$. The correlations range from 0.2 to 0.9 with a mean correlation of 0.6. The fact that the correlation is substantially less than 1 suggests that something can be gained by taking the average of the two estimators. I thus propose a new estimator of school size density (or equivalently whale abundance since OK and SPLINTR agrees on mean school size):

$$d^{(os)} = \frac{d^{(OK)} + d^{(SPL)}}{2}$$

The performance of $d^{(OS)}$ was compared to that of $d^{(OK)}$ and $d^{(SPL)}$ using the mean square error as the criterion:

$$mse(d) = \sqrt{\frac{1}{100}\sum_{j=1}^{100}\left(\frac{d_j - d_{true}}{d_{true}}\right)^2}$$

The main result of this working paper is that $mse_{(OS)}$ = 1.37, $mse_{(OK)}$ = 1.38 and $mse_{(OS)}$ = 1.78, i.e. $d^{(OS)}$ beats both estimators. Fig. 1 shows the results split into scenario.

Discussion and conclusion

Because MSE is a standard measure of performance of statistical estimators, the finding that OS has the lowest MSE suggests that averaging OK and SPLINTR is scientifically defendable. The reason that it is possible to improve on both estimators may be that OK and SPLINTR use different aspects of the IDCR/SOWER data (in some complicated way). The correlations in Fig. 1 support this. The combined estimator makes use of all aspects of data, and is thus to be preferred.

R code

```
d=read.csv("tohans.csv")
mse=function(x) sqrt(sum(x^2))
ind=(!is.na(d$ds.ok))&(!is.na(d$ds.spl))
d = d[ind,]                                          # Remove NA's
d$ds.os=0.5*(d$ds.ok+d$ds.spl)

d$spl=(d$ds.spl-d$ds.km2.act)/d$ds.km2.act
d$ok=(d$ds.ok-d$ds.km2.act)/d$ds.km2.act
d$os=(d$ds.os-d$ds.km2.act)/d$ds.km2.act

plot(tapply(d$os,d$scenario.factor,mse),
     tapply(d$ok,d$scenario.factor,mse))
plot(tapply(d$os,d$scenario.factor,mse),
     tapply(d$spl,d$scenario.factor,mse))
```

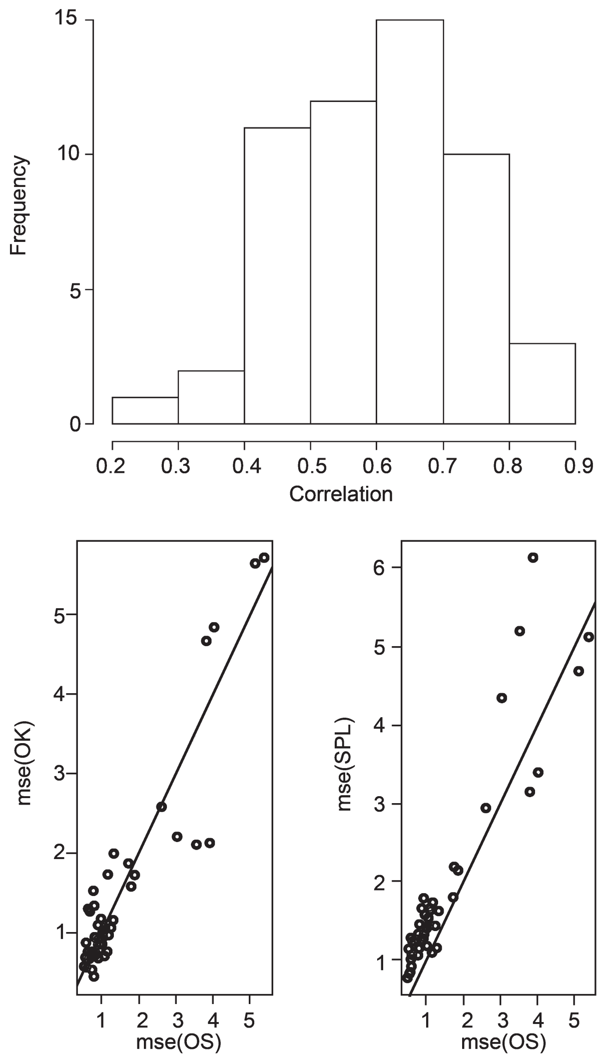

Fig 1. Correlation between OK and SPLINTR within each of 54 simulation scenarios.

Appendix 3

INTERSESSIONAL PROCESS FOR RESOLVING DIFFERENCES IN MINKE WHALE ABUNDANCE

Mark Bravington, Sharon Hedley, Toshihide Kitakado, Hiroshi Okamura and Hans Skaug

Over the past two years, OK and SPLINTR have presented estimates of Antarctic minke whale abundance from the CPII and CPIII IDCR/SOWER cruise data. There are big differences between the estimates, by almost a factor of 2. Such differences are much bigger than the statistical uncertainty, and much bigger than generally seen in the simulated datasets. The extent of differences between each method and the 'Standard' estimates are also perhaps unexpected.

As yet there is no clear explanation for the differences; there is no glaringly obvious deficiency for either OK or SPLINTR in the goodness-of-fit diagnostics considered *so far*. However, the fact is that the estimates are currently incompatible, and there must be some reason. It is important to continue the investigations, not just for understanding the historical abundance and adopting a 'current' abundance estimate, but also for analysing further Antarctic abundance estimates produced by similar survey protocols (e.g. paper SC/62/O16 submitted this year).

Our intersessional work during 2010 has helped to eliminate several possible sources of difference, but has not solved the mystery. We believe that it is important and feasible to continue the investigations, and that there is a reasonable prospect of a timely resolution; the investigations so far have helped focus attention on what the next steps should be, so fewer diagnostics will probably be needed, and the tedious dataset issues which dominated last year's intersessional work have been resolved. This Appendix proposes an intersessional work plan for further investigation, with the expectation of resolving matters for the 2011 meeting.

We will establish an intersessional Group containing the developers, plus other experts in complex abundance estimation problems. The role of the other experts will be to suggest investigations and examine diagnostics. We have proposed a sequence of checks (see below), but the process will be iterative: depending on the results from each check, we may either decide to change what is studied next, or to continue with the next planned check, or to stop because we have worked out the answer.

A timeline is proposed (Table 1). One intersessional Workshop will be required, and there will be email correspondence before and probably after the Workshop. The Workshop should take place preferably this calendar year, and in any case no later than February 2011. Assistance will be required to run some Standard-method analyses on modified datasets. Some further work on Palka's simulated datasets may be required.

Composition

Walløe (Chair), [Branch] Butterworth, Cooke, Palka, Skaug, Wade (advisory experts), Bravington, Hedley, Kitakado and Okamura (developers).

Terms of reference

- Run sensitivity tests on modified real datasets to understand differences between OK and SPLINTR (e.g. in terms of ESW, $g(0)$, MSS, and the underlying sighting parameters).
- Run tests as above to understand differences between OK/SPLINTR and Standard.
- Once we have identified the underlying statistics/ parameters where big differences occur, develop ways to cross-check against empirical data.
- If necessary, design further simulation trials to test robustness.

Statistics/diagnostics to be looked at as soon as possible

Investigations during this meeting suggested that estimates of $g(0)$ at size 1 are very different between the methods, and may be a big contributor to the difference in abundance estimates. There are at least two ways that the estimates might be ground-truthed: by using empirical summaries of duplicate frequency in the real data (but some care is needed to select appropriate subsets of the data), and consideration of SOWER BT-NSP results from recent years. This work should be done as soon as possible, since it may redirect our subsequent investigation.

Statistics to focus on

- Empirical perpendicular dist in CL mode (where school size is known), combined with empirical $g(0)$ from recent BT trials.

Sensitivity runs for all 3 methods on real data

These tests have been chosen to be feasible for all methods (except where stated) without major modifications to the *code*, so that we are testing the same model used for the full data; the plan is rather to modify the *data* used. For all of these, the Standard method should be run *without* using encounter rates from Closing mode, for compatibility with OK and SPLINTR. These suggestions are intended to be run in the order given: results from the earlier runs may make the later runs unnecessary. Further tests may be added if necessary.

(1) SPLINTR-like confirmation treatment: set Conf=Yes for all CL. Set Conf = No for all IO.
   - Aim: confirmation has subtle implications for school size issues. Can't change in SPLINTR. Can change in other methods.
   - Feasibility: do-able for all 3 methods.

(2) No SS error
   - Aim: SS error makes diagnosis very confusing.
   - Feasibility: Do-able for all 3. OK just set Conf=Yes for all sightings. SPLINTR: create artificial SSX data with all SSobs=SStrue.

(3) Fix SchoolSize = 1
   - Aim: investigate impact of SS effects (as opposed to $g(0)$/ESW effects).
   - Feasibility: Do-able for all 3 methods.

(4) SPLINTR without Platform C
- Aim: already done for OK, where it had a moderate but unexpected effect on the abundance estimate. Worth checking whether SPLINTR responds similarly.
- Feasibility: do-able for SPLINTR, done already for OK, not so relevant for Standard.

(5) Fix $g(0)=1$ for SPLINTR
- Aim: for comparison with Standard method.
- Feasibility: probably do-able (set all sightings to AB duplicate).

Table 1
Timeline.

Date		Task
01/08/10	(1)	Revised specification of statistics for empirical checks of $g(0)$/esw using SOWER data.
	(2)	Specification of what to report from sensitivity runs (e.g. abundance by stratum; school size frequency by weather…).
	(3)	Specification of any tedious details of sensitivity runs.
01/11/10	(4)	Circulation of results (and developers' comments on them).
By 02/11*	(5)	Intersessional Workshop, with specifications for any further work. (*preferably earlier than February).
15/04/11	(6)	Circulation of results from item (5).

Further work on simulated data

As yet, we have not identified any specific factors that need further testing through simulation. However, if our sequence of checks does identify factors in the real data that have not been tested severely enough in the simulation trials, then a few further trials may be necessary to assess robustness against those factors.

In addition, we have not yet tested the variance-estimation aspects of either method; these will subsequently be of importance to the Scientific Committee regardless of which estimates are ultimately used. This can be done using a single set of 100 scenarios. For purposes of checking variance estimates, it is probably not of critical importance which scenario is used, but the scenario should be complex enough to test all the aspects of the models that can contribute to estimation uncertainty. None of the scenarios tested so far simultaneously include *all* of what we currently consider to be important factors, so one further scenario should be developed that includes all important factors.

Further, it is desirable to have the simulated datasets presented in the same format as the SOWER data itself. The datasets are inevitably complex, and the formats of real and simulated data are currently very different. Using the same format would provide a guard against any differences between performance on real and simulated data that might arise through differences in the reading-in process.

Finally, in the light of our intersessional checks on the real data, it may be necessary to re-process some of the simulation output in order to report other statistics that we discover to be of significance.

Appendix 4

REPORT OF THE SMALL GROUP PLANNING THE 2011 IWC/JAPAN NORTH PACIFIC CRUISE

Members: Kato (Chair), An, Borodin, Brownell, Clapham, Donovan, Ensor, Matsuoka, Miyashita, Murase, Okada, Pastene, Saramillo (Interpreter), Sekiguchi, Yasokawa (Interpreter) and Uoya.

1. CHAIR'S OPENING REMARKS AND APPOINTMENT OF RAPPORTEUR

Kato was appointed as Chair. Ensor acted as rapporteur.

2. TERMS OF REFERENCE

The terms of reference for the group were to undertake preliminary logistic planning for the 2011 cruise. The research objectives were as follows:

(1) to collect data relevant to the proposed future In-depth Assessment of sei whales in terms of both abundance and stock structure;
(2) to collect data relevant to *Implementation Reviews* of whales (e.g. common minke whales) in terms of both abundance and stock structure; and
(3) to collect baseline distribution and abundance data, biopsy samples and photo-id images, for several large whale species/populations, including those that were known to have been depleted in the past but whose status is unclear, in a poorly known area.

These terms of reference do not include the identification of mid- and long-term research objectives for the IWC-Japan North Pacific cruise series. These longer-term objectives would be clearly formulated as part of a coherent multi-year plan developed intersessionally. If possible this could be undertaken at a Steering Group Meeting scheduled immediately prior to the Planning Meeting for the 2011 cruise (see agenda Item 5)

3. ADOPTION OF AGENDA

The agenda was adopted, and forms the basis of this report.

4. CRUISE LOGISTICS

4.1 Availability of vessel

The meeting was informed that the Government of Japan had made the generous offer of a research vessel and crew for the cruise. The actual vessel to be used has not yet been determined but it may be a vessel previously used in the IDCR/SOWER programme; the vessel will certainly have suitable characteristics to be able to undertake the plans outlined in this report and will have space for three or four researchers. Details were also uncertain of the Certification status of the vessel: Japanese domestic vessel or International

Table 1

Preliminary cruise budget for 2011. Figures in UK £ sterling.

Item	Grant	Travel	Insurance	Shipboard	Shore	Bank charges	Total
Cruise							
Cruise Leader	10,310	1,700	100	831	550	30	13,521
Scientist 1	6,200	1,700	100	831	550	30	9,411
Scientist 2	6,200	1,700	100	831	550	30	9,411
Japan	6,200	1,700	100	831	550	30	9,411
Sub-total							41,754
Equipment/communications							
Sighting							
Modification of ICR data logging system							3,000
Biopsy							
Repairs/maintenance Larsen guns							3,000
Darts x 50 @ 31 each							
Plugs x 1,000 @ 1.4 each							
Ammunition x 500							
Photo-id							
Repair/maintenance/transportation IWC cameras							200
Camera batteries (3)							300
External hard drive (2)							300
Communications: Inmarsat time for reception of visibility forecast and sea temperature data and communication with steering group							500
Transportation of IWC equipment and data							1,400
Planning meeting for 2011(2 days)							
Travel and subsistence for 3 participants: 3 x 1,500							4,500
Annual Meeting							
Cruise Leader travel and subsistence							2,500
Total							**57,454**

vessel. The latter could be advantageous to enable the vessel to enter a US port if such a situation arose.

4.2 Length of cruise
The cruise is scheduled for July and August 2011. The total duration of the cruise will be approximately 60 days, comprising approximately 46 days of research time and 14 days of transit between the homeport in Japan and the research area.

4.3 Number of participants
The vessel will have accommodation for a total of three or four researchers. The researchers will include appropriately qualified personnel from the US and Japan.

4.4 Cruise track design and research mode
The research area for the 2011 cruise was defined as the area bounded by longitudes 170°W and 150°W, and extending north from latitude 40°N to the Aleutian Island chain. It was noted that a survey in this area in particular, represented a valuable opportunity to gain information on the status of right whales, as there had been little recent systematic research in this region. Furthermore, the research area included the region where there had been substantial Soviet catches of right whales in the early 1960s.

A preliminary cruise track with a zigzag design was proposed, and a survey protocol using established IWC-SOWER survey methods. Precise details of stratification of the research area, cruisetrack design and survey methods will be finalised at the Planning Meeting.

Given the fundamental importance of accurate distance and angle data, an estimated angle and distance training exercise and associated experiment will be undertaken during the cruise.

4.5 Experiments other than sightings
Biopsy sampling is planned for the cruise and target species will include North Pacific sei, common minke, right, blue, humpback, grey, bowhead and fin whales. Priority species for biopsy sampling will include right whales, North Pacific sei and common minke whales (in regard to potential targets for biopsy sampling, it was noted that detections of Bryde's whales were not expected due to their distribution south of the southern boundary of the research area (on latitude 40°N)). Biopsy of other species, including killer and sperm whales will be attempted on an opportunistic basis.

Photo-id studies and/or video recording of right, blue and humpback whales will be undertaken.

Details of other experiments would be discussed at the Planning Meeting.

4.6 Other
It was noted that difficulties with CITES issues between Japan and Russia had been experienced last year when biopsy samples had been collected by a Japanese vessel inside the Russian 200n.mile EEZ. Furthermore, any IWC-Japan North Pacific cruise that operates in the US EEZ will encounter the same types of CITES problems.

It was **recommended** that to avoid this regrettable situation, CITES permit issues should be resolved as soon as possible and on a long-term basis rather than on an annual basis. It was also **recommended** that the CITES solution would be the establishment of an Institutional permit (for example on behalf of the Institute of Cetacean Research or The National Research Institute of Far Seas Fisheries). Institutional permits were frequently used for transfer of samples between the US and other countries. This would greatly facilitate import/export and future exchange of cetacean samples between institutions in Japan and the US. It was noted that analysis of samples at sea (thus avoiding

CITES issues) was not a valid scientific option as archival of biopsy samples was essential due to the rapid development of the scope of analyses, notwithstanding the difficulties in collection of samples. The Government of Japan **agreed** to investigate the option of establishing an Institutional permit, and would report the results of its investigations to the Planning Meeting later this year (see Item 5).

Regarding the intersessional development of mid- to long-term objectives for the IWC-Japan North Pacific cruises it was noted that Matsuoka, Miyashita and Clapham will provide an updated summary of North Pacific sighting survey data from US and Japanese cruises to the Long-term Planning Meeting (proposed to precede the cruise Planning Meeting).

5. PLANNING MEETING

It was proposed that a Planning Meeting for the 2011 cruise be held during two days in early October 2010. The Planning Meeting will be held in Tokyo, and Kato agreed to be convenor. Participants will include 2 or 3 non-Japanese participants (including Donovan). It was noted that convening a Steering Group meeting to decide on mid- long-term research objectives and formulate a multi-year work plan was vital and it was **recommended** (to help minimise costs) this could occur in conjunction with the Planning Meeting. Three days were suggested for this meeting of the Steering Group and participants would include 3-4 non-Japanese participants (including Brownell, who it was noted would be able to contribute funds for his participation).

6. BUDGET

The plans given above assume the availability of the same level of Japanese funding as for the 2009/2010 IWC-SOWER Antarctic cruise and the 2010 IWC-Japan North Pacific cruise. A budget request to the IWC of £57,454 is requested (Table 1). Brownell expressed his view that Scientist/Cruise Leader grants should not be provided to researchers who have a normal/full-time salary.