

Annex L

Report of the Working Group on Ecosystem Modelling

Members: Kitakado (Convenor), Baba, Belchier, Bell, Burkhardt, Butterworth, Cañadas, Collins, Cooke, Cunen, Currey, de la Mare, de Moor, Diallo, Donovan, Double, Enmynkau, Fortuna, Frey, Friedlaender, Gunnlaugsson, Hakamada, Haug, Herr, Hielscher, Hinke, Hjort, Isoda, Ivashchenko, Kelkar, Konishi, Lang, Lee, Lundquist, Mate, Mallette, McKinlay, Moore, Morita, H., Morita, Y., Moronuki, Murase, Funahashi, New, Øien, Palka, Phay, Pierce, Punt, Redfern, Reeves, Reyes, Rogers, Santos, Simmonds, Skaug, Slugina, Solvang, Tamura, Tulloch, Víkingsson, Von Duyke, Wade, Walløe, Walters, Weinrich, Yasokawa, Yasunaga, Zerbini.

1. INTRODUCTORY ITEMS

1.1 Convenor's opening remarks

Kitakado welcomed the members of the Ecosystem Modelling Working Group (hereafter Working Group).

1.2 Election of Chair

Kitakado was elected Chair.

1.3 Appointment of Rapporteurs

Butterworth, McKinlay, New and Skaug were appointed as rapporteurs with assistance of the Chair.

1.4 Adoption of Agenda

The adopted Agenda is included as Appendix 1.

1.5 Documents available

The documents available to the Working Group were identified as SC/67a/EM01-16, Redfern *et al.* (2017), Solvang *et al.* (2017), Weinstein *et al.* (2017) and Mate *et al.* (2016).

2. BODY CONDITION ANALYSIS FOR THE ANTARCTIC MINKE WHALE

2.1 Review results of analyses

2.1.1 Review of analyses

Based on an analysis of length-weight relationships, SC/67a/EM02 reported on trends in minke whale body condition as determined from data collected during the JARPA sampling program, 1989-2005. Penalised regression splines were used to model total body weight as non-linear functions of body length, time within season, foetus length and long-term trend over year. Four discrete subsets of the JARPA data were examined after exploratory analyses revealed differences in the length-weight relationship between sexes and between those animals considered to have a high or low diatom load. For a majority of the data (83%, comprising all males, and females with low diatom load) the authors found no evidence for a decline in total body weight over the 17 years of the JARPA program. For females with high diatom load there existed some signal to indicate a decline in body weight, however the long-term trend was not linear, was not consistently in decline for all animals within the group, and was based on small sample sizes (average 37 samples/year). The authors concluded these results provided little evidence for a widespread decline in food availability.

De la Mare presented SC/67a/EM01, which sets out analyses on the subset of the data for whales with measured fat weight (measured for the first whale taken each day from 1988/89 onwards, although a substantial number of days were not sampled). Konishi and Walløe (2015) considered that fat weight is the most appropriate measure of body condition. The analyses in SC/67a/EM01 indicated that total body weight (also included in the JARPA data) was a more complete measure of body condition. The JARPA data showed that the seasonal gain in lean weight (total weight minus blubber weight) exceeded the seasonal gain in blubber weight. A range of models fitted to total weight data were not consistent with any significant long term decline in body condition. Analyses of fat weight with a fixed effects model resulted in an apparent decline significant with $p=0.015$, but this probability value overstated the significance by ignoring the effects of model selection, pseudo-replication and random effects. Their mixed effects models did not show a significant decline in fat weight. Systematic trends in the segment of the population being sampled was evidenced by changes in ages and sex ratios. Overall it difficult to determine whether the apparent changes in body condition in a subset of the models reflected real changes in the population, or whether they were an artefact due to variability in the segment of the population being sampled.

SC/67a/EM04 presented an analysis of the Antarctic minke whale data from the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA). Six response variables were considered, which were all potential proxies for body condition: blubber thickness at two sites, half girth at two sites, fat weight, and an index based on total weight. A large, biologically plausible linear mixed effect model intended to incorporate all effects influencing body condition was proposed and analysed. Model selection was carried out by means of the focused information criterion (FIC) with the goal of increasing the precision of the estimate of the linear effect of year. Both parts of the analysis supported the conclusion that there had been a decrease in body condition over the 18 years under study. Five out of six proxies for body condition had clear, negative, significant estimates for the linear effect of year. Also, for these five responses the FIC procedure selected models with similar conclusions, while a baseline model without a linear effect of year was not favoured. With the last of these six proxies, an index based on total weight, the estimated effect of year was negative, but not significant. Also, in this case, the FIC procedure preferred a model not containing the linear effect of year. However, total weight was considered to be less clearly linked to body condition compared to the other responses.

SC/67a/EM03 was drafted as a response to the draft of SC/67a/EM04 that was circulated as part of the exchange of papers two months before the meeting. The paper provided a constructive critique of the analyses in SC/67a/EM02 including a number of suggestions or requests for clarification. Results were presented that show that on about 28% of days no whales were measured for fat and the proportion of the catch that was weighed changed over the JARPA period, which may add to bias in trends in fat weight.

SC/67a/EM07 addressed a question originating from the analyses of minke whale body condition data collected under the JARPA program. Two teams investigated the hypothesis of yearly decline in body condition and found somewhat conflicting results. A key disagreement concerned whether total weight (suitably standardised) was an equally good (or better) proxy for body condition than measures directly related to blubber (for example fat weight). Both teams found that the total weight of the minke whales appeared to be constant over the JARPA period, while other response variables (for example fat weight) seemed to experience significant decline. These results were considered paradoxical by the authors of SC/67a/EM03. SC/67a/EM07 aimed to explain that the seeming conflict of fat weight and total weight results was not necessarily a paradox. Some simplified examples and a small simulation study were provided to support the case that disagreeing results could be explained by the increased residual variance when using total weight compared to using fat weight. The potential causes of the large residual variance in the total weight were also briefly discussed.

SC/67a/EM08 addressed some issues originating from the analyses of minke whale body condition data collected under the JARPA program. Two teams investigated the hypothesis of yearly decline in body condition and found somewhat conflicting results (see SC/67a/EM01 and SC/67a/EM02 from the first team, and SC/67a/EM04 from the second). Both teams proposed several linear mixed effects models for the response variable fat weight, and this note investigated whether the conclusions in SC/67a/EM04 were affected by incorporating some of the variables and interaction terms suggested in SC/67a/EM01. A total of nine new models were considered and all the models had a significant, negative, linear effect of year, and achieved better AIC values than the winning linear mixed effect model in SC/67a/EM01. Using one of the new models as a new wide model, a full FIC (focused information criterion) analysis of a set of 27 candidate models was performed. The main conclusions in SC/67a/EM04 remained unchanged.

De la Mare explained that the paradox referred to in the commentary in SC/67a/EM03 was the proposition arising from the analyses in SC/67a/EM04 that a substantial decline in nutritive condition purportedly measured by indicators such as fat weight could be evidence for important changes in feeding conditions without there being a corresponding decline in body weight. The problem was not statistical but rather it was the biotic implications of the results in SC/67a/EM04 that were paradoxical. The main results presented in SC/67a/EM01 led to no paradox. The lowest AIC and BIC models for body weight did not include a year trend. However, the trend from the lowest AIC model that did include a year trend gave an estimated rate of change in total body weight (using only the data that include fat weights) of -0.0048 tonnes per year with a standard error of 0.012 , $t=-0.40$, and was clearly not statistically significant. The corresponding estimate of mean body weight at the start of the 17-year JARPA period was 6.98 tonnes (both sexes combined). This estimated decline was -0.068 %/year, leading to a total decline on body weight over the 17-year period of 1.16 %. Running the same model using all the body weights (not just those from animals who also had their fat weights measured) gave a trend of -0.0032 tonnes/year with a standard error of 0.0078 , $t=-0.4$. In this case the percentage decline was -0.045 per year and a total 17-year decline of 0.77 %. The results from the mixed effects model for fat weight was a trend of -0.0035 tonnes per year with

a standard error of 0.00213 , $t=-1.66$, and was shown to be not significant by bootstrap. The estimated mean fat weight for females in year 2 was 1.873 tonnes, giving a percentage decline per year of 0.189 , and a total over 17 years of 3.23 %. Thus there was no paradox in the main results of SC/67a/EM01; the point estimates of the rates of changes in body weight and fat weight are similar, and neither is statistically significant or substantial.

In response to SC/67a/EM01-03, SC/67a/EM16 argued that the basis of 'body weight is better proxy than fat weight or blubber thickness' used the data incorrectly. This suggested that it was reasonable that models with fat weight and blubber thickness showed significant yearly decline with no decline in total body weight, which was also confirmed by SC/67a/EM04 and SC/67a/EM07. These did not admit other factors exist such as sampling design. SC/67a/EM16 also explained unfavourable use of highly correlated close variables (fat weight and body weight; $r=0.9$ in JARPA data) for both response and independent variables in de la Mare's models. SC/67a/EM16 also argued that there was no need of data separation by diatom load level by SC/67a/EM02 by showing a correlation between foetus length and diatom load. SC/67a/EM02 finally concluded that the important declines of nutritional condition in minke whales over the JARPA period, which were significant at the 5% level, remained valid.

2.1.2 Discussion

Several participants argued against the use of total body weight as an indicator of body condition, because total weight was accumulated over many years. An additional problem was that bone weight increases with the age of individuals, and that lipid in bones can be replaced by water and lead to an increase in the total weight. It was argued that blubber thickness or total fat weight were the most appropriate indicators for detecting inter-annual changes in feeding conditions. The same conclusion had been reached for North Atlantic minke whales. Aguilar highlighted that, although indicative of body condition, variation in blubber thickness is not proportional to the evolution of body fat reserves and actually tends to underestimate changes if other morphometric and biochemical variables were not incorporated into the energetic model (Aguilar *et al.*, 2007). For this reason, the actual decline in body condition was likely to be more severe than the observed decline in blubber thickness.

Aguilar pointed out that the relationship between body weight and body length is not commonly considered a reliable proxy of body condition because it is affected by a number of biases: in females it is influenced by pregnancy, it requires incorporation of age in any condition determination because bone grows continuously along life, and it is strongly determined by muscle mass. The latter bias was particularly influential because protein incorporates large quantities of water and its energetic density is much lower than that of fat, so significant variation in muscle mass does not reflect parallel variation in condition. The assumption put forward by document SC/67a/EM01 that fat stores in bone and muscle were as important as those in blubber was inconsistent with previous results on cetacean bioenergetics, the reason being that the data used for the calculations were wrongly selected from the bibliography. Appendix 2 details more accurate values for lipid content that confirm blubber as the main body depot for lipid reserves in baleen whales. In response, the authors of SC/67a/EM01-03 referred to their results that show that fat weight and body weight, conditioned on suitable covariates, showed no substantial

or significant declines over the period of JARPA. They also felt that, according to their analyses, the concordance between the two measures – fat weight and total weight – lent considerable support to the proposal that total weight was an appropriate measure of body condition.

The authors of SC/67a/EM04 were asked about the role played by the spatial covariates. In response, they said that spatial effects are modelled separately as a latitude effect, which is statistically significant and included in all models, and longitude effect (referred to as ‘region’). The longitudinal covariate was not always selected by the FIC criterion, and did not seem to have a large effect on the estimated decline in blubber thickness.

The authors of SC/67a/EM04 were further asked about the method used to estimate the standard deviation in the estimated slope of decline, and in particular if the model selection process had been taken into account. They answered that a cross-validation method, which involved splitting the data into two parts, had been used to ensure the validity of results. Later discussions identified that this data splitting process, while done for generally desirable reasons (i.e. to test the best model selected on a hold-out set of data), as described, introduced a stochastic element to results of model selection that had not been described or summarised. Some members of the Working Group felt this introduced some doubt about the general validity of the results of model selection. The authors of SC/67a/EM04 answered that while the comments above may have somewhat influenced the results after model selection with FIC (the second half of SC/67a/EM04), they in no way influenced the first part of SC/67a/EM04, which contained no stochastic data-splitting, and thus in no way influenced the main conclusions of SC/67a/EM04. Some members of the Working Group thought it would be useful to consider an assessment of the relative stability of parameter confidence intervals in light of the realised JARPA sample sizes and model complexity.

In relation to the estimation of standard deviations, the authors of SC/67a/EM04 explained that these had been evaluated under the selected model, rather than the more conservative approach of using the wide model, but it was expected that using the wide model instead would not make a huge difference. Additional analyses undertaken during the meeting confirmed this (see Appendix 3).

It was argued that in SC/67a/EM02 a substantial part of the data had been left out. The authors of SC/67a/EM02 replied that in response to some concerns expressed by Cunen, Walloe, and Hjort early in the collaboration, most of the data that was originally omitted (samples without measured stomach weight) were reintroduced to the analyses, with updated analyses presented in Appendix D of SC/67a/EM02. The results of the analysis using the more complete data were not appreciably different from the results from the analysis using the reduced data, which excluded cases missing stomach weight.

The Chair noted that it had previously been agreed that there had been a statistically significant decline in blubber thickness and fat weight. It was suggested that when viewed in the wider context of the Committee’s interest, especially from a management perspective, significance at the 5% level is probably not the most important related issue. The Committee’s interest would likely focus on the use of such information in multi-species models. Such use could take two forms: either qualitative confirmation (or otherwise) of trends suggested by a model fitted to other data, or as a component of the likelihood used in fitting such ‘standardised’ blubber thickness values to multi-species models. In either case, the

inputs desired would be annual ‘standardised’ estimates for blubber thickness with associated coefficients of variation. For this, a demonstration of 5% statistical significance, although desirable, would not be essential; rather in the case, say, of fitting a model to information which included these estimates, the model would be the more influenced by those annual estimates with greater precision. The primary utility of variance estimates would be to ensure that they were no greater than the variance of the residuals for the model fit to the ‘standardised’ blubber estimates, to guard against an over-parametrised model overfitting to the information available.

The Working Group **agreed** that, thanks to the collaborative effort, considerable progress had been made in achieving convergence on the question of how to analyse for trends in body condition and/or blubber thickness in the JARPA data. Both teams and the Working Group **agreed** that the estimation of changes over time is more complex than had originally been assumed, because of the need to take account of additional components of variance which are partially confounded with the realised sampling design, and which had not been taken into account on the initial analysis.

2.1.2.1 POSITION OF DE LA MARE AND MCKINLAY (HEREAFTER DM)

DM’s position was that the relevant accumulation of weight was not confined to the blubber and visceral fat, because the accumulation of lean weight (which also includes fat) within a season exceeds the accumulation of weight in blubber and visceral fat. This was consistent with the observations of Lockyer (1981), who in relation to blue and fin whales stated that ‘...the greatest observed increases in weight occur in the internal musculature ...’. DM stated that while some details provided by Aguilar may be improvements, the general conclusion was not; all inferences drawn about fat weight and body weight in the analyses were drawn from the data, and the precise details of the illustration in SC/67a/EM01 were not critical to those results. SC/67a/EM01 and SC/67a/EM03 showed that the additions to both forms of weight (lean + fat weight) reflected feeding, and the analyses showed that there was no significant or substantial trend in total weight. Although DM consider that total weight was the appropriate measure of feeding success, the analyses in SC/67a/EM01 of fat weight also showed no substantial or significant trend (-0.2% per year). The results for blubber thickness (variable BT11) were also not significant. In SC/67a/EM03, DM noted several concerns regarding the FIC analyses that remained unaddressed. FIC results seemed to be single realisations of an approach to model selection that has some random elements. A comparison of the two analyses of fat weight in SC/67a/EM04 and SC/67a/EM08, which used different wide models, had different estimated coefficients, FIC values and confidence intervals, even though the selected final model was the same in both cases. If the coefficients and confidence intervals are instances from distributions of these statistics then those distributions will be necessary to understand the method. It was not clear why analyses that selected the same final model should have had such different results; for example, confidence regions of -0.0126, -0.0039 for SC/67a/EM04 versus -0.0141, -0.0002 for SC/67a/EM08. If the results in SC/67a/EM08 were driven primarily by using a different wide model (with more interaction terms) then the fat weight results in SC/67a/EM08 became consistent with the results of DM; the year trend was not significant because the confidence intervals in SC/67a/EM08 included zero when the wide model variance was used in the calculations in SC/67a/EM07. Finally,

DM drew attention to the results in SC/67a/EM03 which showed that the FIC fat weight model in SC/67a/EM04 was a considerably poorer fit in terms of AIC and BIC compared to their 'best' models that do not include fixed year effects.

2.1.2.2 POSITION OF CUNEN, WALLØE, HJORT AND KONISHI (HEREAFTER CWHK)

The position of CWHK was the following. Despite the arguments from DM, CWHK remained convinced that fat weight, and the related blubber and girth measurements, were better proxies for body condition than total weight. On this point they agreed with the biological arguments made by Aguilar and pointed out that the total weight of any animal is clearly a function of many long-term variables, many of which were not measured (and may be impossible to measure). One example given was the weight of the skeleton, which is known to increase with age in many mammalian species. Using the total weight must thus be expected to obscure any pattern between fat weight and year (SC/67a/EM08). Their choices concerning the statistical modelling had been sound, reasonably robust and consistent with previous recommendations by this Working Group and by earlier recommendations by the authors of SC/67a/EM01 and SC/67a/EM02. The findings supported and confirmed the conclusions reached earlier by Konishi and Walløe (2015): for five of the six body condition proxies, there was a clearly significant linear decline over year. There was basic agreement with DM regarding parts of the basic statistical modelling and model fitting tools. However, CWHK did not agree on the necessity of splitting the data into four parts and then analysing these separately (as has been done in SC/67a/EM02). They claimed that potential differential patterns due to Sex or Diatom coverage could be taken care of by carefully chosen interaction terms. Concerning the claims made by DM in SC/67a/EM03 about the preferred model in terms of FIC having 'considerably poorer fit in terms of AIC and BIC' compared with their 'best' models, CWHK claimed this was solely a consequence of different size-controlling variables (BWt vs BLm), and that they demonstrated in SC/67a/EM08 that their wide model had a better AIC value than the 'best' model in SC/67a/EM02. Also, CWHK considered that the model selected by FIC based on that wide model would also have had a better AIC value than the 'best' model in SC/67a/EM02. The JARPA data set remains a rich source of biologically important information and there is scope for further statistical work of interest regarding some of the finer issues and details of the data. CWHK stood by the main conclusions they had reached regarding the decline in body condition over the JARPA time period.

2.1.3 Conclusion

The Scientific Committee agreed by consensus at its 2014 meeting that there had been a statistically significant (5% level) decline in blubber thickness and fat weight (IWC, 2015). In subsequent years, analyses challenging (as well as supporting) that agreement have been presented. In the Working Group, there was no consensus to recommend a change to the past agreement.

2.2 Review approached used in body condition analyses for other stocks

Solvang *et al.* (2017) presented a study of North Atlantic minke whales regarding the energy deposited at high productive arctic latitudes in summer. It was expected that the whales' body condition on the summer grounds would reflect food availability during their most intensive feeding period, and thus indicate how well the high-latitude ecosystems

can support the populations. During the commercial catch operations on feeding grounds in Norwegian waters, body condition data (blubber thickness and girth) were collected from 10,556 common minke whales caught from 1993 to 2013. To investigate associations between condition and time/area, the authors applied the following three models: (1) multiple regression models with covariates, sex, year, latitude and longitude, to find significant coefficients of the covariate; (2) random effect models involving the random effects of variations by year or area and with sex as a fixed variable; (3) varying coefficient models, which were applied to investigate variation with year/area and to interpret covariate effects by visualisations. The significance of the estimated coefficients were assessed by the authors' proposed statistical tests. In conclusion, the total trend over the two decades of data available suggested a decrease in minke whale condition. However, this trend was most pronounced during the high summer season when the seasonal effect over the annual sampling periods from April to September was considered.

The results in Solvang *et al.* (2017) were of relevance to the *Implementation Review* of North Atlantic minke whales. It was noted that it would be interesting to include prey resources and competing species (cod and harp seals) as explanatory variables in the analyses. It was also suggested that date-of-capture be included as a covariate, in addition to 'season' which was already included. The authors indicated that they have included date-of-capture as a covariate to the model, and have considered day effect in the varying coefficients model in canonical correlation analysis (Yamamura *et al.*, 2016), which indicated a significant positive effect to body condition.

3. REVIEW ISSUES RELEVANT TO ECOSYSTEM MODELLING WITHIN THE COMMITTEE

3.1 Individual-based energetic models

De la Mare indicated that work was on-going on the relevant models (SC/67a/RMP02), but that they did not yet include competition between species.

3.2 Effects of long-term environmental variability on whale populations

The issue of variability in baleen whale demographics was examined at a Workshop held in 2010 (IWC, 2011). Although data were limited, an average coefficient of variation of 0.40 for the inter-annual variability in the rate of successful reproduction was estimated for the stocks examined. There were insufficient data to estimate the long-term (for example decadal-scale) variability of reproduction. Cooke assessed the implications of this level of variability for trajectories of recovering baleen whale populations, with and without the assumption of additional variability at the decadal scale (see Appendix 4).

The simulations suggested that the trajectories of recovering stocks would be expected to show very little signal of environmental variability, and would be well approximated by deterministic models, until the stocks have recovered to about 0.5K (where K is carrying capacity) or more. The fact that many populations have shown smooth exponential increase as they have recovered from low levels, does not imply that they will continue to show smooth trends. Particularly in the case of the Southern Hemisphere, where for several decades baleen whale stocks had been recovering from low levels, higher variability might be expected from now onwards, as stocks recover above 0.5K.

The Working Group took note of the predictions and agreed to keep the item on its agenda, to be discussed if new analyses are forthcoming. The Working Group suggested that efforts be made to include effects of environmental variability in population models, including the individual-based energetic models that are being developed.

3.3 Modelling of competition among baleen whales

Friedlaender presented results from three studies on the foraging ecology of humpback and Antarctic minke whales from satellite tagging studies in the waters off the western side of the Antarctic Peninsula. This research is part of the IWC-SORP supported research programme on the foraging ecology of baleen whales in the Antarctic. In the first paper (SC/67a/EM10), state-space models were employed to satellite tag data from both species to understand the influence of environmental parameters (e.g. sea ice) on the foraging behaviour of each species. Comparisons were also made to understand how the foraging ranges of each species was defined and affected by environmental variables. The authors found that humpback whales spent greater periods of time in area-restricted search and were less likely to switch behavioral states than minke whales, forage in open water and close to shore whereas minke whales were more tied to seasonal sea ice regardless of its location. There was overlap in the core foraging areas of humpback and minke whales, but minke whales had to search far broader areas in order to find suitable habitat for foraging and predator avoidance. There was no current indication that prey was limiting in this ecosystem, but the potential for competition may exist as climate-driven changes decrease the amount of available foraging habitat for minke whales while concurrently increasing open water for humpback whales in which to forage. The results provide the first quantitative estimates for the foraging behaviour of both krill predators in Antarctic waters and provide insights as to the potential effects of a rapidly changing environment on the structure and function of a polar ecosystem. In the other study (SC/67a/EM11) movement was partitioned into seasonal changes in geography, composition, and characteristics using a multi-state mixture movement model. Whales later in the austral fall spent more time in movements associated with foraging, travelled at lower speeds between foraging areas, and shifted their distribution northward and inshore. Seasonal changes in movement were likely due to a combination of sea ice advance and regional shifts in the primary prey source. This study presented an important step towards mechanistic models of movement in the marine environment at broad scales. In the final study (Weinstein *et al.*, 2017) the spatial distribution of satellite-tagged humpback whales and krill fishery effort were analysed within the small-scale management units defined by the Convention for the Conservation of Antarctic Marine Living Resources (CCAMLR). Using a Bayesian movement model to partition whale movement into traveling and area-restricted search states, it was found that both whale behaviour and krill catch effort were spatially clustered, with distinct hotspots of the whale activity in the Gerlache and southern Bransfield Straits. These areas aligned with increases in krill fishing effort, and present potential areas of current and future conflict. The study recommended that the Antarctic West and Bransfield Strait West management units merit particular attention when setting fine-scale catch limits and, more broadly, consideration as critical areas for krill predator foraging.

A question was raised regarding the temporal scale over which the tagging took place. Friedlaender indicated that all minke whales were tagged on the same day, while all

humpback whales were tagged within four days on either side of the minke whales. The acknowledged difficulty in tagging minke whales was overcome, in part, by approaching the whales when they are in groups and by assessing their behaviour to minimise the probability the minke whale will move away. Friedlaender offered to discuss the authors' approach with interested parties.

3.4 Stable isotope analyses

Aguilar presented SC/67a/EM05 and SC/67a/EM06, which both focused on the validation of stable isotope sampling techniques. SC/67a/EM05 addressed the reliability of using faecal material to assess short-term diet composition in baleen whales. It investigated whether stable isotope values of the prey were affected by digestive enzymes and bacteria during the passage along the digestive tract by analysing faeces from Icelandic fin whales and comparing results with those from krill and a variety of fish prey. Results showed that stable isotope values of krill remained unaltered. Also, the stable isotope values of faeces, which under visual inspection appeared to only contain fish remains, revealed that contribution of krill in the digested food was indeed substantial. This demonstrated that: (i) results from macroscopic gross analysis of faeces may be misleading because less digestible components, such as fish bones, may be overrepresented; and (ii) that faecal stable isotope values contribute significant information to the assessment of short-term diet.

SC/67a/EM06 focused on the use of baleen plates, which are composed of inert tissue that grows incrementally and that therefore archives in a sequential manner the stable isotopic values of the whale body pool during a time span of several years. Baleen plates differ in size and sometimes in coloration between different segments of the filtering row or between sides of the mouth, so concern has been raised on the effect of such variation on structural composition and growth rates of the plates that might be affecting the stable isotope values and their oscillations. The paper examined the replicability of patterns between baleen plates occupying different positions in the mouth of a fin whale. Results showed that all baleen plates, independently of their position in the filtering apparatus, size or coloration, grow at the same rate and display similar stable isotope values and oscillations. Therefore, position of sampling along the baleen plate row should not be a reason of concern when conducting stable isotope studies.

The discussion of SC/67a/EM05 focused on how long the intestine contents keep their isotopic concentration. The main conclusion was that, in principle, stable isotopes are not degraded and can be measured in decomposed material. However, there was some concern that in highly degraded tissue, microbial activity may change the composition, for example, the enrichment of nitrogen ranges. This was highlighted as an area in need of investigation.

A question was raised as to whether the results of SC/67a/EM06 could be extended to similar analyses for Antarctic minke whales, where the baleen is regularly sampled from a central position, as there was a desire to ensure that best sampling practice was being followed. The authors stated that their main conclusion was that any baleen plate will produce the same information, while the longest plate contains the most information because of the extended time period over which it has grown. Therefore, the authors concluded that if the samples were always being drawn from the same location, then the results of SC/67a/EM06 would also be applicable to minke whales.

3.5 Review 'regime shift' component of NEWREP-NP

Annex 17 of SC/67a/SCSP10 reported on the review of the 'regime shift' component of NEWREP-NP. In the original plan of NEWREP-NP, 'regime shift' issues were one of the secondary objectives. The original aim of this point was to contribute to the understanding of the implications of environmental change in terms of whale stock management, rather than detection of a major environment change itself. It is difficult to predict whether a major environmental change (categorised as a 'regime shift') would occur within twelve years of NEWREP-NP. These discussions during the Expert Panel's meeting led to the recognition that the original wording of this proposal, 'regime shift', may have been too specific. Hence, the proponents amended this wording as 'environmental change' in order to better reflect their intention. Because 'regime shift' has been identified as one of the types 'environmental changes', the amendment was consistent with the original proposal. Furthermore, given the Panel's recommendation, the original Secondary Objectives I(v) and II(v) were treated as Ancillary Objectives rather than Secondary Objectives in the revised plan. The proponents will monitor spatial distribution of whales, compositions of prey species and body conditions of target whales. They will then investigate potential influential factors, such as the availability of prey resources, which may explain temporal changes in the indexes mentioned above if observed. Such monitoring and investigation will contribute to future in-depth assessment of whales as in the case of Antarctic minke whales.

3.6 Others

Reyes presented SC/67a/EM13 which took note of IWC Resolution 2016-3 'Cetaceans and Their Contribution to Ecosystem Functioning'¹ highlighting the important role that whales play in cycling nutrients through the oceans, in the sequestration of carbon and in enhancing ecosystem productivity, as well as the importance of 'whale falls' as microhabitats in the deep sea. In the resolution, the Commission asked the Scientific Committee 'to screen the existing research studies on the contribution of cetaceans to ecosystem functioning to develop a gap analysis regarding research and to develop a plan for remaining research needs'. SC/67a/EM13 was intended to help this process and provided a comprehensive bibliography of relevant scientific publications to date and suggestions for further research to help fill knowledge gaps. Furthermore, the authors recommended holding an international workshop to further develop research into the roles of cetaceans as ecosystems engineers.

Discussion revolved around the testability of the interesting hypotheses laid out in SC/67a/EM13, especially given the potential complexity of the models that would ensue. The authors clarified that they had brought the item to EM for advice and guidance on the best way forward with regards to building their hypotheses into quantitative models that can be fit to data. Advice was offered regarding the use of tools such as EcoSim, as well as other papers and projects on animal movement and habitat use that speak to how and where animals can be part of ecosystem models using data, rather than simulations. The Working Group welcomed the presentation of SC/67a/EM13 and **agreed** that the sub-committee was the proper place to bring such work. In addition, the Working Group **encouraged** relevant submissions in the future, especially in light of Resolution 16.3.

4. ECOSYSTEM MODELLING IN THE ANTARCTIC OCEAN

4.1 Review progress of modelling

SC/67a/EM14 presented a revision of the Mori and Butterworth (2006) model, incorporating model improvements and updates of abundance and trend information in krill and predator species (SC/67a/EM14). Key additions to the model were the inclusion of a dispensatory effect for Antarctic fur seals in the krill and predator dynamics, and the imposition of bounds on K_a (the carrying capacity for krill in Region *a*, in the absence of its predators); these led to a better fit to the data overall. A particular difference in results compared to those from the Mori-Butterworth model was more oscillatory behaviour in the trajectories for krill and some of its main predators. This likely resulted from the different approach to modelling natural mortality for krill (which decreases the residual mortality remaining after taking account of consumption by the main predators) and warrants further investigation. That may in turn resolve a key mismatch in the model, which predicts minke whale oscillations in the Indo-Pacific region to be out of phase with results from a SCAA assessment of these whales.

Tulloch presented SC/67a/EM12, which described a focused spatial 'Model of Intermediate Complexity for Ecosystem Assessments' (MICE) for phytoplankton, krill, copepods and five baleen whale species for the Southern Hemisphere. The model included predator-prey interactions, and estimated whale population trajectories from 1890 to present. Forward projections to 2100 coupled the predator-prey model to a global climate model. The model predicted Antarctic blue, fin, and southern right whale populations at <50% pre-exploitation numbers (K) in 2100, even given 100 years without catches, because of slow growth rates. Southern right whales were estimated to currently be <11% of their carrying capacity, while humpback whales were predicted to recover to K by 2050. Spatial differences in the recovery of whale species between oceanic regions were highlighted, with slower recovery of blue and fin whales in the Atlantic/Indian region, and slower recovery of southern right whales in the Pacific. Minke population trajectories tracked future expected increases in primary productivity. By using the most up-to-date corrected whaling records, accounting for some key uncertainties (e.g. through model calibration, sensitivity analyses) and fitting to all available survey data, the model presented an updated assessment for blue, fin, humpback, right and minke whales in the Southern Hemisphere as a basis for exploring ecosystem dynamics in the Southern Hemisphere. Results demonstrated key differences in population trajectories and estimates between models that account for, or ignore, predator-prey linkages. This is a strategic model that provides a platform for exploring additional hypotheses and management strategies, and is being modified in a step-wise fashion to explore predator-prey interactions and the effects of future environmental change on krill and whales.

In considering the ecosystem models presented in SC/67a/EM14 (MB model) and SC/67a/EM12 (MICE model), the Working Group noted the differences in objectives, trophic interactions captured and scales of the models, but also some of the synergies in key data requirements. Both are krill-based predator-prey multispecies models, and are naturally underpinned by similar data requirements (though at different scales) and a requirement for a sound understanding of ecosystem function. As ever, models and data are imperfect. A discussion of data deficiencies, or 'missing links' in the

¹<https://archive.iwc.int/pages/search.php?search=%21collection72&k=>.

models, identified cephalopods and salps as two potentially important species about which relatively little is known. More broadly, it was acknowledged that little was known about dynamics, abundances or trends in mesopelagic species.

The need for better data for describing population dynamics of individual species, and for more quantitative information about energy transfer between related trophic levels was emphasised. While MB is macro scale, and MICE tends towards mesoscale, both have the potential to derive useful input from studies of small-scale processes. Telemetry-based studies of individual animal energetics might provide one such example, with the additional possibility that such studies can help to quantitatively determine the nature of functional responses. Other advances in technology also open new opportunities, such as cameras on drones, fishing gear, and land-based remote monitoring.

The implications for hypothesis generating or testing from ecosystem models was briefly discussed. The MB model is well suited for assessing large-scale, whole-of-ecosystem response to broad-scale change (e.g. the differential recovery of baleen whales after the cessation of long-term whaling). MICE, on the other hand, is able to utilise covariates at various scales to provide short- to medium-term predictions for defined ecosystem responses to hypothesised changes in covariate values. Ecosystem response to changing resource availability is an obvious example, which has seen application in monitoring CCAMLR Small Scale Management Units (SSMUs).

4.2 Cooperation with CCAMLR on multi-species modelling

The Working Group was pleased to be able to welcome several CCAMLR members to participate in discussions, in particular Mark Belchier, current Chair of the CCAMLR Scientific Committee. A previous IWC-CCAMLR workshop on data requirements for ecosystem models was held in 2008, so it was timely that another workshop is in the planning stages for 2019. The Working Group **agreed** that data sharing, data quality control, and identifying data gaps were key issues to be resolved at an institutional level between the IWC and CCAMLR. ‘Data first, models second’ was the flavour of the discussion. It was recognised that CCAMLR and IWC both share similar goals in terms of developing whole-of-ecosystem modelling approaches, and that this similarity could be leveraged to the advantage of both organisations. A defined area for collaborative modelling between CCAMLR and IWC members was suggested, with the Antarctic Peninsular as one possibility. This would perhaps be a worthwhile topic for an IWC-CCAMLR workshop on ecosystem processes and models.

4.3 Plan for joint SC-CCAMLR – IWC SC workshops

Plans for two joint SC-IWC/SC CCAMLR workshops to develop multi-species models of the Antarctic marine ecosystem were discussed. The chairman of CCAMLR’s SC noted that there had been no agreement reached at CCAMLR XXXV to fund attendance of CCAMLR scientists at the proposed 1.5 day ‘plenary’ workshop scheduled for 2017. However, it was recalled that the CCAMLR Commission anticipated that a workshop would be held in 2018.

Noting that there was still a clear need to progress the development of multi-species models, several options to advance the work were considered. It was **agreed** that this work would best be facilitated through workshops, noting that the breadth of topics was likely to make the use of an

Intersessional Contact Group difficult. The TORs agreed in 2015 were reviewed and were still considered valid. However, it was suggested that there should be a greater emphasis on the Western Antarctic Peninsula region (CCAMLR subarea 48.1) and especially in the regions in which the krill fishery has, in recent years, become much more spatially constrained. Outputs could help to inform the ‘risk assessment’ approach to spatial management in the region and assist in the development of the feedback management of the krill fishery.

Two workshops could be run in conjunction with the IWC-SC and WG-EMM (or CCAMLR intersessional meetings) in 2018 and 2019 respectively (see Appendix 5). The Working Group endorsed the revised plan for the workshops with CCAMLR. It was noted that progress may be facilitated if the attendance at the first workshop was limited (to a maximum of about 4 attendees each as for primary goal would be planning future initiatives rather than reviewing research). The broad ecological scope the 2008 IWC/SC meeting had been ambitious but this may have reduced its overall impact.

The Working Group **recommended** that collaboration between IWC-SC/SC-CCAMLR be ongoing, and that the revised plan for the workshops be implemented.

5. APPLICATION OF SPECIES DISTRIBUTION MODELS (SDMS) AND ENSEMBLE AVERAGING

5.1 Review progress of guideline for SDMs

An intersessional Correspondence Group (CG) has been established since SC/65b to develop guidelines and recommendations for best modelling practices of species distribution models (SDMs). The group conducted a preliminary review of SDMs applied to baleen whales during the first intersessional period between SC/65b and SC/66a (Murase *et al.*, 2015). Subsequently, the group conducted preliminary reviews of machine learning methods, which are commonly used as SDMs, and during the subsequent intersessional period between SC/66a and SC/66b developed general guidelines for their application (Murase *et al.*, 2016). SC/67a/EM15 reported on progress made by the CG between SC/66b and SC/67a. During the period, the group updated Murase *et al.* (2015) by adding information from Murase *et al.* (2016) as well as integrating a further 12 reviews of SDM papers published between March 2015 and December 2016. The CG plans to complete the tasks assigned to the group during the intersessional period from SC/67a to SC/67b. The work plan includes the following tasks: (1) revising descriptions of each machine learning method; (2) adding short methods descriptions for boosted regression trees (BRT) and generalised additive models (GAM); (3) adding a short guideline for GAM, with appropriate citations; and (4) final preparation for journal publication.

The Working Group thanked the CG for work during the intersessional period, and invited interested members of the group to provide feedback. It was noted that while the focus of the review had been on machine learning methods for species distribution models, GAMs were becoming an increasingly useful framework for these kinds of analyses and that section could potentially be expanded. It was noted that many machine learning methods are relatively opaque when it comes to model checking diagnostics, but that modern GAM’s are firmly grounded in statistical theory and have available a rich suite of GLM-related diagnostic tools. The Chair noted that the guidelines for SDMs is intended

to be a living document, and so areas could be expanded as interest and time allowed. Finally, it was noted that an explanatory application of the guidelines to some real or simulated data would be useful.

5.2 Review progress of works on SDMs and ensemble modelling

A joint IWC-National Marine Fisheries (NMFS) Workshop titled 'Towards Ensemble Averaging of Cetacean Distribution Models' was held in San Diego, USA, prior to SC/66a. The objective of the Workshop was to convene a group of experts in modelling, statistics, and marine ecology to identify methods to compare and combine model predictions using existing species distribution models (SDMs) for Eastern North Pacific blue whales as a case study.

In 2016, the US E-mail Correspondence Group (US-CG) (Elizabeth Becker, Monica DeAngelis, Daniel Palacios, and Jessica Redfern) determined that a scaled-down version of the original work plan was necessary. The US-CG decided to focus on the risk of ships striking blue whales on the USA west coast and to use only those models that covered the entire USA west coast. Preliminary work to create an ensemble of the predictions was conducted and raised a number of questions.

In 2017, the US-CG, together with Karin Forney and Elliott Hazen, made significant progress in addressing these questions. In particular, the predictions from each model were developed using unique grids and spatial resolutions. The group created a unified grid for all predictions and identified areas where model predictions were similar and where they were different. They also developed methods to scale the different predictions (e.g. density versus probability of occurrence). Finally, the authors used the area under the receiver operating characteristic curve (AUC) and related metrics to explore different methods for weighting the predictions in the ensemble. This work is expected to be completed in the coming year and the plan is to submit a manuscript to a peer-reviewed journal.

In considering Redfern *et al.* (2017), the Working Group noted that a summary of this work had also been presented during SH (see Annex H), but that the focus of discussion here would be on the methodological aspects of the study. The goal of the study was to predict cetacean distributions in data poor ecosystems. Blue whales (*Balaenoptera musculus*) were used as a case study because they are an example of a species that have well-defined habitat and are subject to anthropogenic threats. The highest blue whale densities in the eastern Pacific Ocean are associated with upwelling-modified waters that are highly productive and support dense aggregations of krill. Consequently, habitat variables that identify variations in upwelling, circulation, and water column stratification that may affect forage availability were used in the models. The study used 377 sightings of one or more blue whales and approximately 225,400km of effort from surveys conducted by NOAA Fisheries' Southwest Fisheries Science Center from August through November (California Current: 1991, 1993, 1996, 2001, 2005, 2008, and 2009; eastern tropical Pacific: 1998, 1999, 2000, 2003, and 2006). Generalised additive models (GAM) (Wood, 2006) were used to relate the number of blue whales in each transect segment to the habitat variables, largely following the methods of Becker *et al.* (2016). Four measures of model performance (AUC, TSS, and the percentage of sightings in the highest 2 and 10% of predicted densities) identified a single model that provides the best match to the blue whale sightings in each ecosystem. This model was used to predict

blue whale distributions, rather than using an ensemble of predictions from GAMs with different habitat variables.

The Working Group queried why ensemble averaging hadn't been performed for competing models. The authors explained that model assessment metrics and independent experts identified a single best model that performed better than ensemble. The chosen models performed consistently well on both quantitative metrics and qualitative expectations. There was some discussion which pointed out that the methods performed well for these particular data, but that the good performance may be specific to the case in question. Therefore, there was interest in whether picking a best model may result in uncertainty being under-represented should the method be applied more generally. It was suggested that the broad geographic area of the study region would likely capture several distinct behavioural states (e.g. transiting and foraging), and that different models may be capturing different aspects of behaviour unequally. In discussion of the on-going ensemble modelling, it was noted that the methods for combining uncertainty when averaging an ensemble of models were not yet well developed. The Working Group **encourages** an update on the progress of this work at a future meeting of the Scientific Committee.

6. OTHER MATTERS

6.1 Review information on krill distribution and abundance by NEWREP-A

SC/67a/EM09 reported on krill and oceanographic surveys in Antarctic Areas V-W during the 2016/17 austral summer season as a part of second New Scientific Whale Research Program in the Antarctic Ocean (NEWREP-A) dedicated sighting survey. Two research vessels, one of which is a trawler-type vessel since this year, were engaged with krill acoustic survey and net samplings by small ring nets and an Issak-Kid Midwater Trawl (IKMT) for species identification and size compositions of plankton at 32 stations and 13 stations, respectively. Oceanographic observations using CTDs and water sampling were also conducted coincidentally. Krill and oceanographic data are currently being examined, and results obtained in the 2016/17 season will be presented to a CCAMLR specialists' workshop. Feedback from the specialists will be reflected in the planning of the 2017/18 survey.

The Working Group thanked the authors for their work. It was also clarified that the departure from the expected krill survey design was due to the data being collected in conjunction with a survey for whales, and would still meet the needs of the study.

6.2 Other

Mate *et al.* (2016) described new archival Advanced Dive Behaviour (ADB) tag technology that has documented dive profiles (depth and duration) as well as foraging effort by blue, fin and sperm whales. ADB tags need to be recovered to obtain all of the detailed data for subsequent analysis, so there is a practical limit of how long they are allowed to stay attached for convenient surface recovery in the study area. For example, 7 ABD tags on blue whales provided data on >17,300 dives, during average attachments of 23 days, documenting high daily variability in diel dive depths and foraging effort. The time between intense foraging bouts (with ARS type movements) was frequently characterised by 'transiting' (more linear) travel that can last 8-16 days and cover long distances with continued diving without foraging proxies to search for dense prey patches. The

Table 1
Summary of the work plan for the EM working group.

Item	Intersessional 2017/18	2018 Annual Meeting (SC/67b)
(1) Cooperation with CCAMLR on multispecies modelling	Prepare a pre-meeting Workshop under a Steering Group (see Table 2).	Hold a pre-meeting Workshop to review the status of multispecies models and available data series (see Appendix 5).
(2) Applications of species distribution models (SDMs)	Intersessional Working Group activity (see Annex W).	Review progress by SDM working group.
(3) Effects of long-term environmental variability on whale populations	Continue further analyses.	Review progress by working group.
(4) Further investigation of individual-based energetics models	Continue further analyses.	Review results of further analyses.
(5) Modelling of competition among whales	Continue further analyses	Review results of further analyses.
(6) Update of information on krill distribution and abundance by NEWREP-A	Conduct a survey by consultation of CCAMLR specialists.	Review results of the survey and analysis.

details of the foraging effort and variability had never been recorded for such long periods. On sperm whales, ADB tags documented how foraging bouts at different depths could change without diel influence to reveal details of habitat and prey preferences. Because researchers traditionally study foraging in ‘good habitats’, where whales are abundant, scientists have come to think whales do not have problems finding food. The foraging proxies reported by ADB tags now quantify the extent of changed foraging predicted by state-space analyses and give a much more detailed understanding of the patchy prey distribution whales encounter and which aspects of habitat types they find most productive. The authors deployed more advanced tags last year on blue and fin whales that no longer require recovery to acquire similar information for >100 days.

The Working Group thanked the authors for the work and encouraged the continuation of the research and expressed the desire that future results continue to be presented to the Working Group. The authors further clarified that they have used click reflection intervals to help estimate whale lengths.

7. WORK PLAN

See Table 1 for the work plan and see Annex W for a list of intersessional correspondence groups.

8. ADOPTION OF REPORT

The report was adopted on 17 May 2017 at 17:35. The Chair expressed his sincere appreciation to the rapporteurs, Butterworth, McKinlay, New and Skaug, for their excellent work. The Working Group thanked Kitakado for his leadership and gratefully accepted his offer to convene the Group next year.

REFERENCES

- Aguilar, A., Borrell, A. and Gómez-Campos, E. 2007. The reliability of blubber thickness as a measure of body condition in large whales. 11pp. Paper SC/59/O17 presented to the IWC Scientific Committee, May 2007, Anchorage, USA (unpublished). 11pp. [Paper available from the Office of this Journal].
- Becker, E.A., Forney, K.A., Fiedler, P.C., Barlow, J., Chivers, S.J., Edwards, C.A., Moore, A.M. and Redfern, J.V. 2016. Moving towards dynamic ocean management: how well do modelled ocean products predict species distribution? *Remote Sensing* 8: 149.
- International Whaling Commission. 2011. Report of the Third Intersessional Workshop on the Review of MSYR for Baleen Whales, Seattle, 20-24 April 2010. *J. Cetacean Res. Manage. (Suppl.)* 12:399-411.
- International Whaling Commission. 2015. Report of the Scientific Committee. Annex K1. Report of the Working Group on Ecosystem Modelling. *J. Cetacean Res. Manage. (Suppl.)* 16:277-90.
- Konishi, K. and Walloe, L. 2015. Substantial decline in energy storage and stomach fullness in Antarctic minke whales (*Balaenoptera bonaerensis*) during the 1990s. *J. Cetacean Res. Manage* 15: 77-92.
- Lockyer, C. 1981. Growth and energy budgets of large baleen whales from the Southern Hemisphere. *FAO Fisheries Series No. 5 (Mammals in the Sea)* 3: 379-487.
- Mate, B.R., Irvine, L.M. and Palacios, D.M. 2016. The development of an intermediate-duration tag to characterize the diving behaviour of large whales. *Ecol. Evol.*: 1-11. DOI: 10.1002/ece3.2649.
- Mori, M. and Butterworth, D.S. 2006. A first step towards modelling the krill-predator dynamics of the Antarctic ecosystem. *CCAMLR Science* 13: 217-77.
- Murase, H., Friedlaender, A., Kelly, N., Kitakado, T., McKinlay, J., Palacios, D.M. and Palka, D. 2016. Progress report of the intersessional corresponding group ‘Applications of species distribution models (SDMs)’ since 66a IWC/SC. Paper SC/66b/EM04rev1 presented to the IWC Scientific Committee, June 2016, Bled, Slovenia (unpublished). 31pp. [Paper available from the Office of this Journal].
- Murase, H., Friedlaender, A., Kelly, N., Palacios, D.M. and Palka, D. 2015. A preliminary review of species distribution models (SDMs) applied to baleen whales. Paper SC/66a/EM03 presented to the IWC Scientific Committee, May 2015, San Diego, CA, USA (unpublished). 18pp. [Paper available from the Office of this Journal].
- Redfern, J.V., Moore, T.J., Fiedler, P.C., De Vos, A., Brownell Jr, R.L., Forney, K.A., Becker, E.A. and Ballance, L.T. 2017. Predicting cetacean distributions in data-poor marine ecosystems. *Divers. Distrib.* 23: 394-408. [Available at: <http://wileyonlinelibrary.com>].
- Solvang, H.K., Yanagihara, H., Oien, N. and Haug, T. 2017. Temporal and geographical variation in body condition of minke whales (*Balaenoptera acutorostrata acutorostrata*) in the Northeast Atlantic. *Polar Biol.* 40: 667-83.
- Weinstein, B.G., Double, M., Gales, N., Johnston, D.W. and Friedlaender, A.S. 2017. Identifying overlap between humpback whale foraging grounds and the Antarctic krill fishery. *Biol. Cons.* 210: 184-91.
- Wood, S.N. 2006. *Generalized Additive Models: an Introduction with R*. Chapman and Hall, Boca Raton, Florida. 391pp.
- Yamamura, M., Yanagihara, H., Solvang, H.K., Oien, N. and Haug, T. 2016. Canonical correlation analysis for geographical and chronological responses. *Procedia Comput. Sci.* 96: 1,351-60.

Appendix 1

AGENDA

1. Introductory items
 - 1.1 Convenor's opening remarks
 - 1.2 Election of Chair
 - 1.3 Appointment of rapporteurs
 - 1.4 Adoption of Agenda
 - 1.5 Documents available
2. Body condition analyses
 - 2.1 Antarctic minke whales
 - 2.1.1 Review results of analyses
 - 2.1.2 Discussion
 - 2.1.3 Conclusion
 - 2.2 Review approaches used in body condition analyses for other stocks
3. Review issues relevant to ecosystem modelling within the Committee
 - 3.1 Individual-based energetic models
 - 3.2 Effects of long-term environmental variability on whale populations
 - 3.3 Modelling of competition among baleen whales
 - 3.4 Stable isotope analyses
 - 3.5 Review 'regime shift' component of NEWREP-NP
 - 3.6 Others
4. Ecosystem modelling in the Antarctic Ocean
 - 4.1 Review progress of modelling
 - 4.2 Cooperation with CCAMLR on multi-species modelling
 - 4.3 Plan of joint SC-CAMLR-IWC SC workshops
 - 4.4 Work plan
5. Application of species distribution models (SDMs) and ensemble averaging
 - 5.1 Review progress of guideline for SDMs
 - 5.2 Review progress of works on SDMs and ensemble modelling
6. Other matters
 - 6.1 Review information on krill distribution and abundance by NEWREP-A
 - 6.2 Others
7. Work plan
8. Adoption of report

Appendix 2

RELATIVE CONTRIBUTION OF BLUBBER AND OTHER BODY COMPARTMENTS TO THE LIPID BODY POOL OF LARGE WHALES AND THEIR ROLE AS PROXIES OF BODY CONDITION - COMMENTS ON SC/67a/EM01

Alex Aguilar

SC/67a/EM01 presented to this meeting proposes that fat stores in bone and muscle are as important as blubber stores and, consequently, that blubber thickness, or by extension any index based on blubber characteristics, would not be a good proxy of body condition. However, this assumption appears inconsistent with previous results on energetics and lipid metabolism, which do show that blubber indeed constitutes the main lipid depot in baleen whales (Aguilar and Borrell, 1994; Aguilar *et al.*, 2007; Lockyer, 1981; 1987b).

The reason for this apparent inconsistency appears to be that document SC/67a/EM01 was based on extreme values or non-representative data on muscle and bone lipid content extracted from two papers by Lockyer (1981) and Lockyer (1987b) North Atlantic (Iceland) fin whales. These papers are re-examined here to extract from them truly representative values, and incorporate other values obtained from a comparable study also on North Atlantic (Spain) fin whales (Aguilar and Borrell, 1994). As it can be seen in Table A2.1, which depicts the ranges of lipid content values determined in these studies, results were quite consistent in both cases.

Unfortunately, the sample size was not detailed in Lockyer (1987a), so an overall value for each tissue could not be calculated. However, when combining the ranges of the percent lipid values from Table A2.1 with the mean tissue weights by sex calculated for minke whales in document SC/67a/EM01, the resulting figures (Table A2.2) show that the contribution of the blubber to the total lipid body pool is about 5 to 10 times larger than that of bone and about 4 to 8 times larger than that of muscle.

Also, it should be pointed out that the lipid contained in the bone is not necessarily all available for energy because its main function is to lighten the bone density and thus provide buoyancy to the whale (Lockyer, 1987a). For this reason, the sensitivity of bone lipid content to changes in body condition should be expected to be very limited.

All this confirms previous findings that blubber is the main depot for lipid reserves in the body of baleen whales and, as a consequence, that thickness and lipid content of blubber are the most sensitive proxies of body condition in baleen whales.

Table A2.1

Lipid content of various tissue types in fin whales.

Tissue	Tissue/species	Lipid content %	Reference
Blubber	Fin whale	65.5-81.1	Lockyer (1987a)
	Fin whale	67.3-73.3	Aguilar and Borrell (1994)
Muscle	Fin whale	3.4-5-1	Lockyer (1987a)
	Fin whale	5.1-8.9	Aguilar and Borrell (1994)
Bone	Fin whale	18.37	Lockyer (1987a)
	Fin whale	8.2-10.4	Aguilar and Borrell (1994)

Table A2.2

Weight of the various tissue types in minke whales (extracted from SC/67a/EM01) and total lipid weight for these tissue types calculated from the ranges of values detailed in Table A2.1.

Tissue type	Tissue weight (tonnes)		Total lipid weight (tonnes)	
	Females	Males	Females	Males
Bone	1.39	1.17	0.11-0.25	0.09-0.21
Muscle	4.04	3.51	0.14-0.36	0.12-0.31
Blubber	1.76	1.38	1,15-1.43	0.90-1.12

REFERENCES

- Aguilar, A. and Borrell, A. 1994. Reproductive transfer and variation of body load of organochlorine pollutants with age in fin whales (*Balaenoptera physalus*). *Archives of Environmental Contamination and Toxicology* 27(4): 546-54.
- Aguilar, A., Borrell, A. and Gómez-Campos, E. 2007. The reliability of blubber thickness as a measure of body condition in large whales. 11pp. Paper SC/59/O17 presented to the IWC Scientific Committee, May 2007, Anchorage, USA (unpublished). 11pp. [Paper available from the office of this Journal].
- Lockyer, C. 1981. Growth and energy budgets of large baleen whales from the Southern Hemisphere. *FAO Fisheries Series No. 5 (Mammals in the Sea)* 3: 379-487.
- Lockyer, C. 1987a. Cetacean bioenergetics. In: Huntly, A.C., Costa, D.P., Worthy, G.A.J. and Castellini, M.A. (eds). *Marine Mammal Energetics*. Society for Marine Mammalogy, Special Publication 1.
- Lockyer, C.H. 1987b. The relationship of body fat, food resource and reproductive energy costs in North Atlantic fin whales (*Balaenoptera physalus*). *Symposia of the Zoological Society, London* 57: 343-61.

Appendix 3

SUPPLEMENTARY ANALYSIS

Celine Cunen

During discussions in the Working Group, the authors of SC/67a/EM04 (Cunen, Walløe, Hjort and Konishi, but specifically Cunen) were asked to provide some additional analyses. They were asked to investigate the effect of diatom-coverage by splitting up the dataset according to the dichotomous diatom variable proposed in SC/67a/EM02, thereby separately analysing animals of low and high diatom coverage. Later, it was also suggested to split the data by sex. Thus, here Cunen presents separate models for four groups of minke whales: males with low diatom coverage, females with low diatoms coverage, males with high diatom coverage and females with high diatom coverage. Two response variables are used, fat weight and BT11 (for definition see SC/67a/EM04).

Further suggestions concerned removing observations from the Ross sea, which was done here. In addition, the authors of SC/67a/EM04 were asked to produce plots showing the effect of each year, i.e. allowing each year to have a different intercept. The new model is referred to as the ‘categorical model’:

$$Y \sim \text{YearCat} + \text{BLm} + \text{DateNumS} + I(\text{DateNumS}^2) + \text{LatNumS} + (\text{Fetus.length}) + \text{LatNumS} * \text{DateNumS} + \text{LatNumS} * I(\text{DateNumS}^2) + \text{Age} + (0 + \text{DateNumS} + I(\text{DateNumS}^2) | \text{YearCat})$$

Note that:

- ‘region’ was removed, as were all interaction terms with this categorical variable (not possible to use with year as categorical);
- ‘age’ was added (clearly significant and was mentioned in the meeting with interest);
- ‘ice’ was removed;
- the random effects of Year on the effect of date were retained, but the random effect on the intercept was removed, since this is in a way the same as the categorical year-term now included; and
- foetus length only applies to the female groups.

The categorical model was applied to the four groups of animals defined above, and for each group figures displaying the categorical effect of each year, along with error bars corresponding to one standard error are provided. In all cases, year 1 is set as the reference value. The figures must thus be interpreted as the *change* in intercept (compared to year 1) for each year. For fat weight the scale of measurements is tonnes, for BT11 it is centimetres.

The author also applied versions of our original wide model (see SC/67a/EM04) to each group. In this note, this model is referred to as the ‘linear model’. The linear model had to be simplified in some cases due to the reduced sample sizes. For this model, all the estimated coefficients are provided.

In conclusion, the four groups of animals displayed somewhat differing patterns for some of the explanatory variables. The author claims that these differences can be taken care of by carefully chosen interaction terms and thus do not necessitate the data-splitting which has been undertaken here. Splitting up the dataset is generally considered unfortunate, leading to reduced power.

As a whole, the plots of the year effects from the categorical model indicated decrease in body condition for most of the animal groups. Generally, the negative trend was clearer with fat weight than with BT11 (which seems to display more year-to-year variation). An exception is the plot with fat weight for females with low diatom coverage. There the last years seem to indicate an increase in body condition. Note, however, that females with low diatom coverage is the smallest group among the four considered (for fat weight this group only includes 96 animals).

The results from the linear model are summarised in Table A3.1. The estimated linear effect of year was negative in all groups, and for several was clearly significant (at a 5% level). Groups with low diatom levels generally had less clear negative patterns than the groups with high diatom loads. This could have a biological reason.

Table A3.1
Summary results from the linear models.

Sex/diatom	Fat weight		BT11	
	<i>n</i>	Est. (<i>t</i> statistic)	<i>n</i>	Est. (<i>t</i> statistic)
Males, low	162	-0.013 (-4.2)	940	-0.008 (-0.8)
Female, low	96	-0.004 (-0.9)	573	-0.013 (-1.5)
Male, high	281	-0.010 (-3.6)	1,793	-0.016 (-2.0)
Female, high	103	-0.015 (-2.1)	657	-0.0028 (-2.1)

All in all, the results presented here do not change the authors’ original conclusions from SC/67a/EM04. They are consistent with the hypothesis that there has been a decline in body condition during the 18 years of the JARPA study.

Fat weight results – Males with low diatom coverage (162 observations)

Table A3.2 provides the estimated coefficients for the linear model for the fat weight of males with low diatom coverage, and Fig. A3.1 displays the year effects from this categorical model.

Table A3.2

The estimated coefficients for the linear model for the fat weight of males with low diatom coverage.

	Estimate	Std. error	<i>t</i> value
(Intercept)	-1.370	0.334	-4.100
YearNum	-0.013	0.003	-4.244
BLm	0.322	0.041	7.915
Ice	0.052	0.028	1.842
DateNumS	0.069	0.018	3.762
I(DateNumS^2)	0.018	0.015	1.178
LatNumS	-0.02	0.021	-1.090
Age	0.003	0.001	2.218
DateNumS:LatNumS	-0.03	0.018	-2.155
I(DateNumS^2):LatNumS	0.003	0.019	0.173

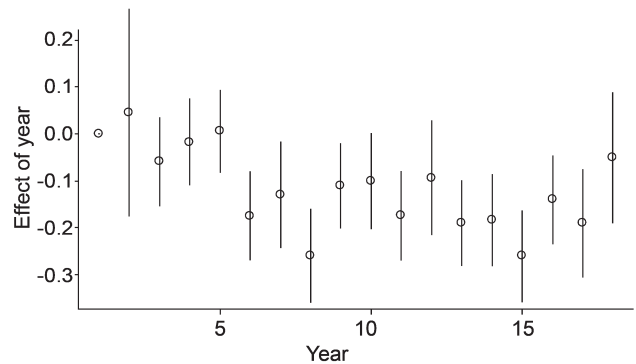


Fig. A3.1. The categorical effect of each year on fat weight (in tonnes) for males with low diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for the other years must be interpreted as the change in intercept compared to year 1.

Fat weight results – Females with low diatom coverage (96 observations)

Table A3.3 provides the estimated coefficients for the linear model for the fat weight of females with low diatom coverage, and Fig. A3.2 displays the year effects from this categorical model.

Table A3.3

The estimated coefficients for the linear model for the fat weight of females with low diatom coverage.

	Estimate	Std. error	<i>t</i> value
(Intercept)	-2.267	0.529	-4.285
YearNum	-0.004	0.005	-0.881
BLm	0.412	0.062	6.653
Ice	-0.057	0.053	-1.078
DateNumS	0.002	0.047	0.041
I(DateNumS^2)	0.072	0.032	2.269
LatNumS	-0.014	0.035	-0.404
Age	0.007	0.003	2.308
Fetus.length	0.002	0.001	1.794
DateNumS:LatNumS	-0.039	0.025	-1.544
I(DateNumS^2):LatNumS	-0.015	0.026	-0.582

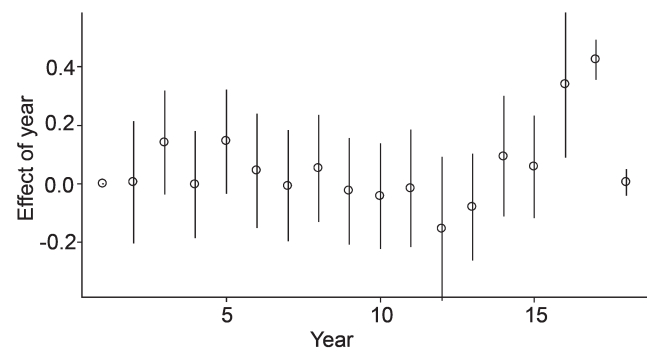


Fig. A3.2. The categorical effect of each year on fat weight (in tonnes) for females with low diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

Fat weight results – Males with high diatom coverage (281 observations)

Table A3.4 provides the estimated coefficients for the linear model for the fat weight of males with high diatom coverage, and Fig. A3.3 displays the year effects from this categorical model.

Table A3.4

The estimated coefficients for the linear model for the fat weight of males with high diatom coverage.

	Estimate	Std. error	<i>t</i> value
(Intercept)	-1.891	0.196	-9.659
YearNum	-0.010	0.003	-3.622
BLm	0.395	0.024	16.753
Ice	0.030	0.023	1.303
DateNumS	0.096	0.020	4.900
I(DateNumS^2)	0.006	0.012	0.469
LatNumS	-0.032	0.015	-2.083
Age	0.004	0.001	4.156
DateNumS:LatNumS	-0.024	0.011	-2.119
I(DateNumS^2):LatNumS	-0.003	0.010	-0.293

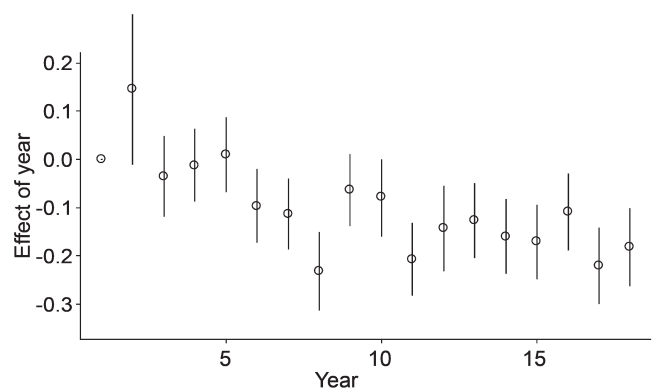


Fig. A3.3. The categorical effect of each year on fat weight (in tonnes) for males with high diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

Fat weight results – Females with high diatom coverage (103 observations)

Table A3.5 provides the estimated coefficients for the linear model for the fat weight of females with high diatom coverage, and Fig. A3.4 displays the year effects from this categorical model.

Table A3.5

The estimated coefficients for the linear model for the fat weight of females with low diatom coverage.

	Estimate	Std. error	<i>t</i> value
(Intercept)	-2.657	0.581	-4.571
YearNum	-0.015	0.007	-2.133
BLm	0.492	0.067	7.306
Ice	-0.128	0.092	-1.397
DateNumS	-0.003	0.047	-0.065
I(DateNumS^2)	0.027	0.035	0.764
LatNumS	0.002	0.050	0.048
Age	0.005	0.002	2.221
Fetus.length	0.001	0.001	2.065
DateNumS:LatNumS	0.004	0.040	0.092
I(DateNumS^2):LatNumS	-0.012	0.027	-0.460

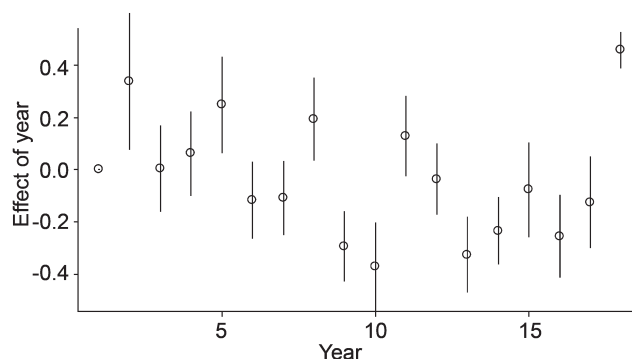


Fig. A3.4. The categorical effect of each year on fat weight (in tonnes) for females with high diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

BT11 results – Males with low diatom coverage (940 observations)

Table A3.6 provides the estimated coefficients for the linear model for BT11 for males with low diatom coverage, and Fig. A3.5 displays the year effects from this categorical model.

Table A3.6

The estimated coefficients for the linear model for BT11 for males with low diatom coverage.

	Estimate	Std. error	<i>t</i> value
(Intercept)	2.218	0.525	4.222
YearNum	-0.008	0.009	-0.814
BLm	0.107	0.063	1.700
Ice	-0.057	0.057	-0.987
DateNumS	0.327	0.037	8.949
I(DateNumS^2)	0.067	0.025	2.629
LatNumS	-0.005	0.036	-0.140
Age	0.000	0.002	0.084
Region1	-0.040	0.048	-0.831
Region2	-0.094	0.055	-1.722
DateNumS:LatNumS	-0.056	0.030	-1.885
I(DateNumS^2):LatNumS	-0.025	0.025	-1.009
LatNums: Region1	0.067	0.038	1.742
LatNums: Region2	-0.037	0.0447	-0.789

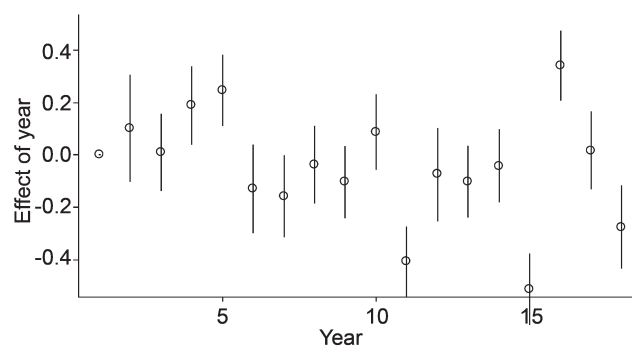


Fig. A3.5. The categorical effect of each year on BT11 (in cm) for males with low diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

BT11 results – Females with low diatom coverage (573 observations)

Table A3.7 provides the estimated coefficients for the linear model for BT11 for females with low diatom coverage, and Fig. A3.6 displays the year effects from this categorical model.

Table A3.7

The estimated coefficients for the linear model for BT11 for females with low diatom coverage.

	Estimate	Std. error	<i>t</i> value
(Intercept)	2.665	0.783	3.401
YearNum	-0.013	0.009	-1.464
BLm	0.051	0.091	0.562
Ice	0.131	0.089	1.472
DateNumS	0.065	0.051	1.289
I(DateNumS^2)	0.088	0.041	2.129
LatNumS	-0.129	0.061	-2.132
Age	0.001	0.004	0.132
Fetus.length	0.009	0.002	5.779
Region1	0.011	0.052	0.203
Region2	-0.082	0.059	-1.386
DateNumS:LatNumS	-0.096	0.047	-2.037
I(DateNumS^2):LatNumS	-0.024	0.043	-0.556
LatNums: Region1	0.022	0.057	0.384
LatNums: Region2	-0.206	0.085	-2.414

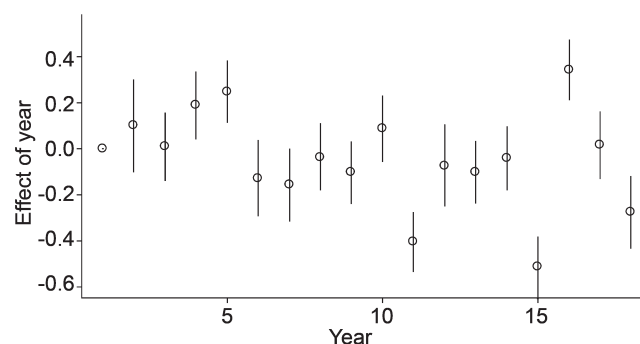


Fig. A3.6. The categorical effect of each year on BT11 (in cm) for females with low diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

BT11 results – Males with high diatom coverage (1793 observations)

Table A3.8 provides the estimated coefficients for the linear model for BT11 for males with high diatom coverage, and Fig. A3.7 displays the year effects from this categorical model.

Table A3.8

The estimated coefficients for the linear model for BT11 for males with high diatom coverage.

	Estimate	Std. error	t value
(Intercept)	3.076	0.420	7.316
YearNum	-0.016	0.008	-1.978
BLm	0.085	0.051	1.673
Ice	-0.063	0.052	-1.214
DateNumS	0.481	0.042	11.434
I(DateNumS^2)	0.050	0.034	1.464
LatNumS	-0.011	0.034	-0.318
Age	0.001	0.002	0.602
Region1	-0.144	0.042	-3.472
Region2	-0.057	0.050	-1.149
DateNumS:LatNumS	-0.021	0.025	-0.868
I(DateNumS^2):LatNumS	-0.025	0.022	-1.128
LatNums: Region1	0.044	0.033	1.354
LatNums: Region2	-0.044	0.043	-1.011

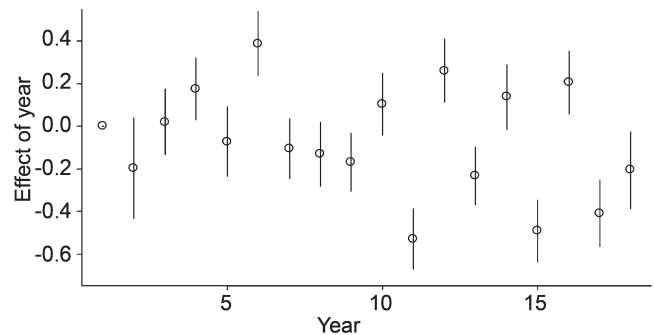


Fig. A3.7. The categorical effect of each year on BT11 (in cm) for males with high diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

BT11 results – Females with high diatom coverage (657 observations)

Table A3.9 provides the estimated coefficients for the linear model for BT11 for females with high diatom coverage, and Fig. A3.8 displays the year effects from this categorical model.

Table A3.9

The estimated coefficients for the linear model for BT11 for females with high diatom coverage.

	Estimate	Std. error	t value
(Intercept)	3.010	0.788	3.820
YearNum	-0.028	0.014	-2.073
BLm	0.049	0.089	0.548
Ice	0.013	0.119	0.108
DateNumS	0.134	0.083	1.608
I(DateNumS^2)	0.088	0.053	1.649
LatNumS	-0.165	0.075	-2.217
Age	0.005	0.003	1.733
Fetus.length	0.009	0.001	12.279
Region1	0.049	0.069	0.707
Region2	-0.013	0.075	-0.174
DateNumS:LatNumS	0.077	0.061	1.254
I(DateNumS^2):LatNumS	-0.046	0.045	-1.023
LatNums: Region1	0.141	0.073	1.942
LatNums: Region2	-0.186	0.114	-1.634

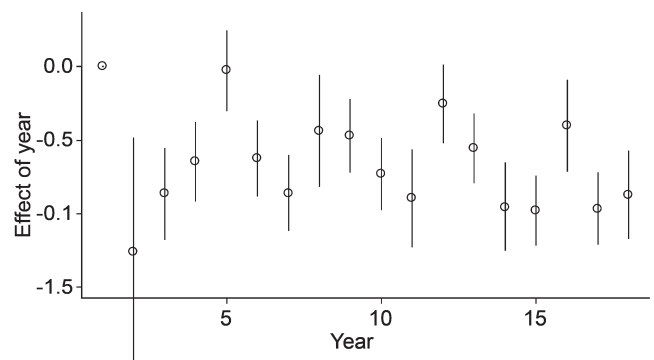


Fig. A3.8. The categorical effect of each year on BT11 (in cm) for females with high diatom coverage. The error bars correspond to one standard error. Year 1 is set as the reference value, and the values for other years must be interpreted as the change in the intercept compared to year 1.

SUPPLEMENTARY ANALYSIS 2: A QUESTION CONCERNING THE CORRECT VARIANCE ESTIMATES AFTER MODEL SELECTION WITH FIC

The authors of SC/67a/EM04 were asked an interesting question about how to correctly estimate the variance of the coefficient estimates after model selection with FIC. The authors computed the variances under the winning model (selected by the procedure), but Cooke pointed out that it may more natural to compute the variance estimates under the wide model (which is the assumed true model). In the figure below, the authors display the effect of these two choices (on fat weight, with the original model M0, and model M4 which was chosen by the FIC procedure – SC/67a/EM04). The black confidence curve uses the standard-error from M4 (computed by bootstrapping), while the blue curve uses the standard-error from M0 (the wide model, also by bootstrapping). As expected, the width of the confidence intervals increased slightly.

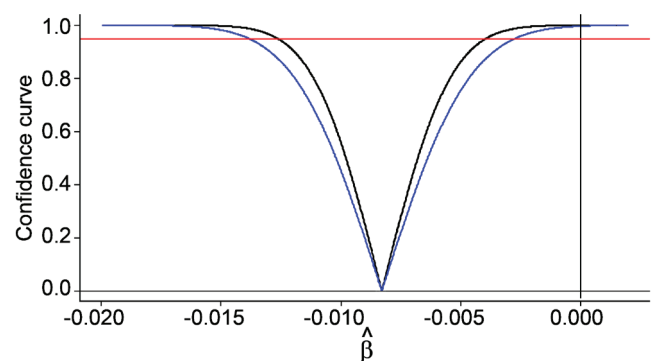


Fig. A3.9. Two confidence curves for the linear effect of year. Confidence curves point to the point estimate and display confidence intervals at all levels. The red line indicates the 95% level. The black curve is computed with the standard error from the selected model M4, while the blue curve uses the standard error from the wide model M0.

Appendix 4

EFFECTS OF LONG-TERM ENVIRONMENTAL VARIABILITY ON THE WHALE POPULATIONS

Justin G. Cooke

A Workshop held in 2010 (IWC, 2011), examined the issue of variability in baleen whale population dynamics. Data that could be used to estimate demographic variability was only available from a few populations, and are summarised in table 4 of the Workshop report. The mean observed coefficient of variation (CV) of the reproductive rate was about 0.4, but it was not possible to estimate long-term variability, which, if present, would have contributed only to a partial extent to the observed CV's in these relatively short series.

Predictions were made here using the environmental variability model of Cooke (2007) for recovering stocks of baleen whales in regions of low, medium and high habitat quality (see that paper and the discussion in IWC (2009) for explanation and definition).

For each habitat, there was assumed to be either only short-term variation in recruitment, with a CV of 0.4, or short and long-term variability each with a CV of 0.3, scaled to the rate at 0.25K. Long-term variability was modelled by assuming an inter-annual correlation of 0.9, which corresponds to variability on a decadal time scale.

A random sample of trajectories was plotted for each case, relative to K (mean carrying capacity) along with the deterministic trajectory for comparison (Figs. A4.1a-f).

It is notable that for medium and high habitat quality, the recovery was predicted to be very close to the deterministic trajectory until the population reaches about 0.5K. Observations of recovering populations over the last few decades, particularly in the Southern Hemisphere, have tended to be of populations below 0.5K, hence it is perhaps not surprising that population trajectories have not yet shown much variability. Now that populations of some species have reached or exceeded 0.5K, more variability can be expected.

It is not yet clear over what timescales the effects of long-term variability would become qualitatively different from shorter term variations.

REFERENCES

- Cooke, J.G. 2007. The influence of environmental variability on baleen whale sustainable yield curves. Paper SC/N07/MSYR1 presented to the MSYR Workshop, Seattle, USA, 16-19 November 2007 (unpublished). 19pp. [Paper available from the Office of this Journal].
- International Whaling Commission. 2009. Report of the MSYR Workshop, 16-19 November 2007, National Marine Mammal Laboratory, Alaska Fisheries Science Center, Seattle, WA, USA. *J. Cetacean Res. Manage. (Suppl.)* 11:467-80.
- International Whaling Commission. 2011. Report of the Third Intersessional Workshop on the Review of MSYR for Baleen Whales, Seattle, 20-24 April 2010. *J. Cetacean Res. Manage. (Suppl.)* 12:399-411.

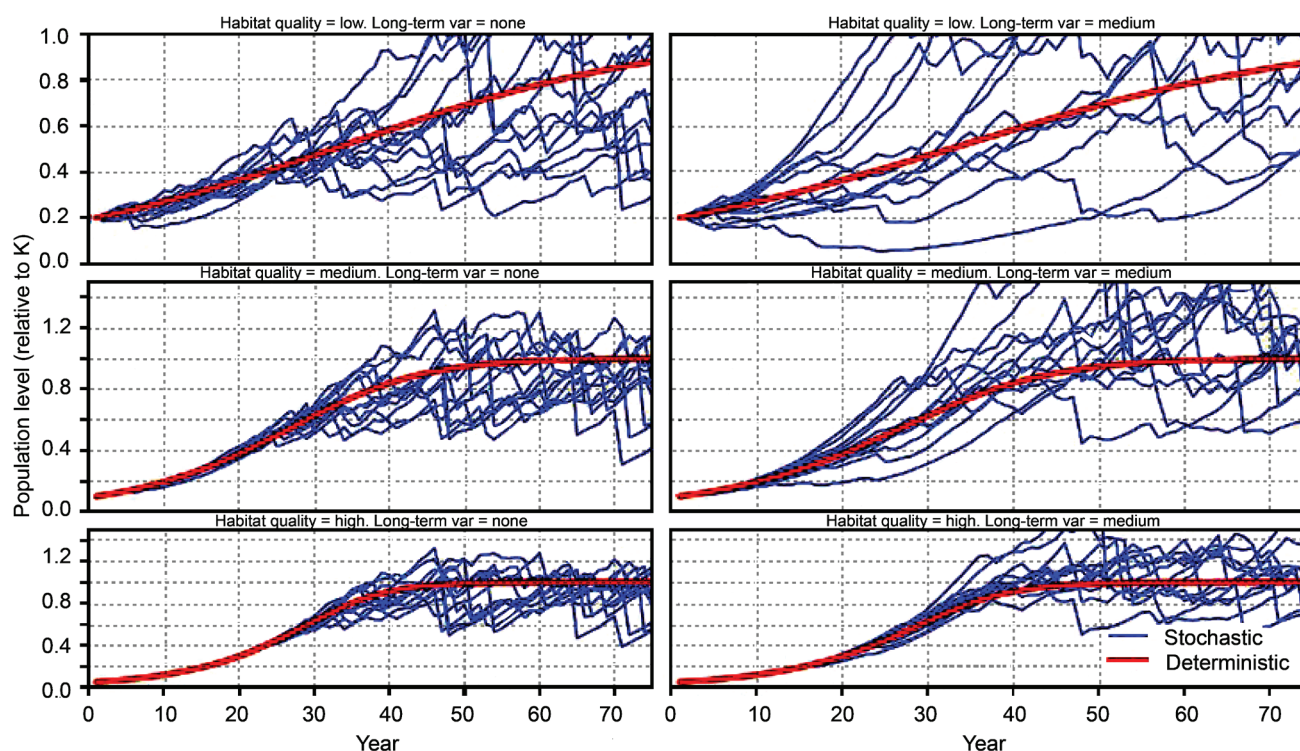


Fig. A4.1. Sample of stochastic population trajectories for recovering baleen whale populations using the environmental variability model of Cooke (2007), for three different habitat qualities, with and without decadal environmental variation.

Appendix 5

REVISED PLANS FOR THE JOINT SC-CAMLR AND IWC-SC WORKSHOP 2018-2019

A proposal for a Joint SC-CAMLR and IWC-SC two-day Workshop to develop multi-species models of the Antarctic marine ecosystem was discussed at the SC-CAMLR 2014, and a steering group to progress a Joint IWC-CCAMLR Workshop was formed (SC-CAMLR 2014 Paragraph 10.25). The joint workshop was perceived as an opportunity to increase knowledge on specific species and their interactions in different management areas, possibly initially focussing on the Antarctic Peninsula given it is a high-priority area for both CCAMLR and IWC (IWC, 2015). The steering group developed a paper identifying draft terms of reference (SC-CAMLR-XXXIV/BG/33). This was tabled to and endorsed by the SC-CAMLR 2015.

Terms of Reference (ToR) endorsed by SC-CCAMLR to guide the two CCAMLR-IWC Modelling Workshops in 2018 and 2019 are:

- (1) foster collaboration between SC-IWC and SC-CAMLR;
- (2) review outcomes from the joint workshop in 2008, assess progress since then including information on species interactions for species of interest to CCAMLR and IWC;
- (3) initial discussion on multispecies models of the Antarctic marine ecosystem and develop work plans toward the second workshop; and
- (4) consider multispecies models of the Antarctic marine ecosystem, at a scale that is able to inform strategic management advice, mainly focussing on the Antarctic Peninsula area as a test-case area, and set directions for future collaborative research activities that would be of mutual interest.

The 1st workshop (two days) in 2018 should briefly review outcomes from the joint workshop held in 2008 (assess progress since then and highlight information on species interactions that are of mutual interest to CCAMLR and IWC). It should initiate discussion on the purpose and the types of multispecies models that are needed by both organisations, and develop work plans towards the 2nd workshop in 2019. The ToR for the 2nd workshop will be updated following the 1st workshop.

After consideration, the steering group suggests the following draft agenda for the 1st workshop in 2018.

DRAFT AGENDA

1. Introduction
 - 1.1 Terms of reference
 - 1.2 Agenda and organisation of the meeting
 - 1.3 Background
2. Review the status of multispecies models and available data series
 - 2.1 Outcomes from the 2008 joint workshop and progress since then
 - 2.2 Key questions to be addressed by multispecies ecosystem models
 - 2.3 Purpose, status of, and suggestions regarding, relevant multispecies models
 - 2.4 Abundance and trends of species relevant for developing and fitting multispecies models
 - 2.5 Outstanding questions
3. Workplan for the 2nd workshop
 - 3.1 Review priority questions of mutual interest into the future
 - 3.2 The scale and the types of model to be developed
 - 3.3 Geographic areas and ecological issues of mutual interest
 - 3.4 Tasks and milestones
4. Report adoption
5. Close of the meeting.

WORKSHOP PREPARATION

The steering group will identify a list of potential participants and presenters by January 2018, and prepare a call for papers to be submitted to the workshop, with a deadline at least two weeks prior to the workshop. The call for papers will highlight the purpose of the workshop and identify the level of information sought including the purpose of existing models, the data required and data available for such models. The CCAMLR Observer is requested to liaise with CCAMLR Secretariat to discuss available from the CCAMLR Ecosystem Monitoring Program (CEMP) and krill fishery data and how that might be prepared and summarised ahead of the workshop.

REFERENCE

International Whaling Commission. 2015. Report of the Scientific Committee. Annex K1. Report of the Working Group on Ecosystem Modelling. *J. Cetacean Res. Manage. (Suppl.)* 16:277-90.