Report of 2nd AWMP Workshop on Greenlandic Fisheries

Contents

The Workshop was held at the National Institute of Aquatic Resources (DTU-Aqua), Charlottenlund Castle from 24-27 March 2009. The participants were Allison, Brandão, Butterworth, Donovan, Givens, Heide-Jørgensen, Punt, Schweder and Witting.

1. INTRODUCTORY ITEMS

1.1 Welcoming remarks
Donovan (convenor) welcomed the participants. He drew attention to the following quotation from the 2008 Scientific Committee report (IWC, 2009a):

The Committee has been unable to provide satisfactory management advice on common minke whales off West Greenland. However, it has made considerable progress in developing an assessment method for common minke whales using sex ratio data and established an ambitious work plan to complete this work by the 2009 Annual Meeting. Although a safe method for providing interim advice has been developed for fin, bowhead and humpback whales (see Item 8.1), it is important that work on developing an SLA for fin whales begins immediately. An intersessional Workshop is essential to maintain momentum on both these important tasks.

He stressed that the Workshop must make sufficient progress to complete its work with respect to evaluating the use of the sex ratio approach to assessment of common minke whales by Madeira. In addition, it was important to begin to look at the development of an *SLA* approach for fin whales.

1.2 Election of Chair
Donovan was elected as Chair.

1.3 Appointment of rapporteurs
Butterworth, Givens and Punt acted as rapporteurs.

1.4 Adoption of Agenda
The adopted agenda is given as Annex A.

1.5 Review of documents
The documents available to the meeting were SC/M09/AWMP1-5 (Annex B) and relevant extracts from past reports.

2. COMMON MINKE WHALES

2.1 Updated abundance estimate
SC/M09/AWMP3 reported on a visual aerial line transect survey for common minke whales conducted off West Greenland in August-September 2007. This survey has already been discussed by the Scientific Committee in the context of abundance estimates for fin whales and humpback whales (Heide-Jørgensen *et al.*, 2008a; Heide-Jørgensen *et al.*, 2008b). A total of 8,670 km of survey effort covered 14 strata in sea states <5 with a total stratum area of 213,996 km^2. The 27 sightings of minke whales were all within a strip width of 300m and the average time from first detection to when the sighting passed abeam was 1.7s. Due to the narrow strip width of the distribution of detections strip census methods were used to analyse the survey. Two methods were deployed to correct the strip census estimates for whales missed by the observers and whales that were submerged during the passage of the plane. Method 1 included all detections of minke whales (n=27) and corrected for an instantaneous availability that included submergence of whales. The 'at surface' abundance of minke whales was 1,216 (CV=0.23) and a correction for whales missed by the observers with a simple mark-recapture estimator resulted in a corrected abundance of 1266 (0.21) whales. Adjusting for the availability bias resulted in a fully corrected estimate of 11,894 (95% 7,387-19,152) minke whales. Method 2 used only detections of minke whales that were observed to break the surface (n=19). Applying this method gives an 'at surface' abundance of minke whales of 847 (CV=0.30) and correcting for whales missed by the observers increases the abundance to 920 (0.30) whales. Adjusting for the availability bias resulted in a fully corrected estimate of 24,260 (95% CI 12,507-47,058) minke whales.

The Workshop welcomed these important new data. It noted that the two estimation methods explored in the paper produced substantially different results, and discussion focussed on understanding and reconciling the differences. The primary cause of the difference relates to alternative choices for adjusting for availability bias (i.e. the negative bias associated with the fact that animals are below the surface and thus not available to be seen by observers). In 'Method 1', the period of time the whale is detectable is estimated from photography, and the overall availability is obtained by multiplying this time by the frequency of surfacings (as estimated from a previous cue-counting survey). In 'Method 2', the proportion of time available for detection is estimated directly from telemetry data from five whales.

The Workshop considered a number of aspects of the data used for availability estimates. Although the cue counting data were obtained in West Greenland, the photographic data were obtained in Iceland and the telemetry data were obtained in Norway, Iceland and Greenland. The telemetry data were also collected during both day and night.

There was some discussion as to whether behavioural differences among these times/locations might make the use of some data less appropriate for use in the current analyses. However, surfacing times are strongly dependent on physiological requirements. Further, with respect to the telemetry data, particular behavioural variations can be partially integrated out by measuring availability over time periods of several hours or more. The Workshop **agreed** that for the telemetry data only daytime data should be used, and it was recommended that the data be partitioned into a few 'windows' of several hours each. These improvements could facilitate better estimation of availability.

The telemetry data provide times at the surface over sampling windows of various temporal lengths (but not the number of surfacings). Given the nature of the satellite communication and the possibility of multiple surfacing in a short time span being mistaken for a single surfacing event, telemetry-based estimates of proportions of time spent at the surface were judged to be positively biased when they were derived from short temporal windows. Therefore, the estimated mean surface time of 3.2% was believed to be too great. The Workshop recommended two possible improvements. First, the telemetry data could be truncated to exclude sampling periods of less than, say, 500 or 750 seconds. The mean of the remaining data would be lower, perhaps in the range of 2.0 to 2.5%. Alternatively, for computing the average time at the surface each data record could be weighted by its temporal span; this corresponds to integrating across the accumulated record for each whale. Neither option is complex and the Workshop **agreed** that that both should be used in any further analyses.

In broader terms, the Workshop considered further how best to estimate availability bias. To calculate and use an estimate of availability, the term must first be defined in relation to whale behaviour and survey protocol.

Minke whales are usually detectable when they are at or near the water surface i.e. the period derived from the Method 1 photography, aside from some minor approximations related to partitioning states.

The telemetry data (Method 2) provide information about the time span when the tag was above water. This does not correspond directly to availability for several reasons. First, the whale can be available for detection for a period of time when it is below the surface, (i.e. the tag is underwater but parts of the body are not and/or the animals can be seen through the water or the whale is entirely below the surface at a shallow depth). Method 2 addresses this by analysing only the sightings where the observers recorded that some part of the whale was above the surface. However, the head may be above the surface while the tag is still underwater. Furthermore, the opportunity for sighting is quite brief (due to the plane speed, brief surfacing periods, and narrow strip width). As a result, observers may have detected some sub-surface whales in front of the plane but postponed recording the sightings until they were abeam of the plane, at which time the whale may have broken the surface. In this case, the number of whales included for analysis using Method 2 would be too large.

From the photographic data, it is possible to partition the availability period into portions below the surface, breaking the surface, and on the surface. As a rough approximation, the portion corresponding to when a tag would be above water was about 1/3 of the overall availability. However, about 2/3 of the observed whales were used for Method 2 based on its definition of availability (namely, tag above water). These fractions differ by a factor of 2, which is roughly the same factor by which the final abundance estimates differ. In addition, there was some discussion over the most appropriate way to estimate the window of time a whale may be within visual range of the observers (and hence the average probability of detecting a whale at the surface) used in both methods. An alternative approach that should be explored is given in Annex C.

In conclusion, the Workshop **recommends** further investigation of the most appropriate way to address availability bias and to adjust for it in the abundance estimate (including an appropriate estimate of variance). At this point, there is no reason to eliminate either Method 1 or Method 2 from consideration. Both methods could have biases in either direction. However, the apparent challenges faced with Method 2 for appropriately matching a definition of availability with a (sub)set of sightings to produce a suitable availability adjustment may be greater than obtaining an availability estimate and a final abundance estimate using Method 1.

2.2 Progress with respect to the use of sex ratio data and modelling approaches

SC/M09/AWMP2 reported on a meeting between Brandão, Butterworth and Witting in Cape Town in December 2008 in relation to modelling approaches using sex ratio data. Useful progress was made on a number of fronts; the most important of these was confirming that the computer codes for a common age-structured model to be used for the analyses planned, which had been developed separately by Brandão and by Witting, provided identical outputs.

2.2.1 Witting-Schweder

SC/M09/AWMP5 provided updated methodology and results for the Witting –Schweder approach to obtain a lower confidence bound on the population size of the West Greenland minke population. Changes made since the 2008 Scientific Committee meeting included a refined procedure for finding the global minimum of the negative log likelihood function and taking account of additional variance (to the binomial variance) by estimating this variance when fitting to the sex-ratio data. SC/M09/AWMP5 also extended the methodology to take account of population trends indicated by a time series of [relative indices/absolute] of population abundance, which the authors indicated to be their preferred approach.

Some difficulties were noted with the Witting estimator, particularly problems in defining the likelihood in terms the observed sex ratio when this could be zero or infinity. Witting had dealt with this through amending the ratio to reflect, e.g., number of males +1/number of females +1. Some reservations were raised about the possibility that this introduces bias, but in the interests of time results were accepted as they were. A suggestion for an improved approach to this is given under Item 2.2.4.

An independent check was made during the meeting to confirm that Witting's likelihood had been coded correctly, and that estimation with an alternative minimiser led to unchanged results.

Witting reported that his approach of minimising by starting from a number of different values of K, and then selecting the parameter set leading to the lowest negative log-likelihood, had been applied consistently and worked reasonably. A problem however was instances of negative deviance in a small proportion of cases. It was noted that these negative deviances tended to be associated with higher true values of deviance. This led the Workshop to conclude that if, say, a 5%-ile lower confidence bound was required, it would be sufficient to reduce the proportion of negative deviances close to zero and treat any remaining instances of negative deviance for a given value for K as ones with the actual deviances all higher than the highest deviance for that value of K. The Workshop **agreed** that this would lead to a lesser (and hence more conservative) estimate of the lower 5% confidence bound than for an estimator without such an associated difficulty.

2.2.2 Brandão-Butterworth

SC/M09/AWMP4 provided updated methodology and results for the Brandão-Butterworth approach. Adjustments made since the 2008 Scientific Committee meeting included increasing the number of replicate data sets generated to 1000, and estimating parameter values for the model with fixed K before doing so for the case where K is also estimable, and using the estimates from the fixed-K analysis as starting values for the estimated-K analysis search. This guaranteed that no instances of negative deviance would occur.

An independent check was made during the meeting to confirm that Brandão's likelihood had been coded correctly, and that estimation with an alternative minimiser led to unchanged results.

Although the Brandão-Butterworth approach excluded instances of negative deviances by construction, it manifested the difficulty that the AD Model Builder package (http://admb-foundation.org/) used indicated a relatively larger proportion of cases where non-convergence was indicated because the approach's maximum derivative condition was not met.

Two approaches were pursued during the meeting to reduce the proportion of non-converged solutions, which were shown to tend to have larger deviance values, so that their exclusion from the process of determining deviance percentiles would lead to biased estimates of lower confidence limits for abundance:

(a) restarting the minimisation for instances indicated as non-converged;

(b) starting the minimisation search at multiple values of K as for Witting's estimator (see above).

Annex D shows the results of applying (a) and (b) to a case where the true K was fixed at 17,000 for the Brandão and Butterworth approach. Under (a), 76% of the simulations converged for both fixed and estimated K with no restarts and this increased to 96% after allowing for up to four restarts. The upper 5th percentile of the deviance increased appreciably when allowing for a single restart but tended to stabilise with more restarts. Under the implementation of (b) considered, for the situation where K is estimated, successful convergence increased from 88% to 98%.

2.2.3 Overall results to date

Whilst progress had been made during the intersessional period, especially after the Cape Town authors meeting in December 2008 (SC/M09/AWMP2), considerable work remains. Results presented for the lower 5% and 10% confidence bounds on current abundance in SC/M09/AWMP4 and 5 for the two approaches showed wide differences; furthermore the directions of the differences were opposite for models 3 and 4b. Despite the work undertaken at the Workshop on minimisation, at this stage it remains unclear whether these differences reflect minimisation problems and/or sensitivity to the different features of the approaches.

Most of the work undertaken at the Workshop focussed on minimisation issues and the results given under Annex D suggest that the approaches considered there will resolve most (possibly approaching all) of the convergence problems encountered with at least the Brandão-Butterworth approach. Ideally a zero convergence failure rate is desired but it was recognised that this might prove difficult to achieve. Failures could be dealt with by assigning a deviance higher than the highest deviance for converged cases to them, thus providing lesser estimates of the lower confidence bounds for abundance.

The Workshop discussed what might be an appropriate target level for the rate of successful convergences for reliable (from the perspective of providing management advice) estimation of the lower 5th percentile of abundance. Although no consensus 'target' was agreed, views expressed ranged from 97%-100%. Those favouring the higher levels within the range indicated that lower levels could be acceptable if upper deviance percentiles showed sufficient tendency to asymptote as the number of minimisation restarts allowed was increased. Recommendations for further work are given under Item 2.2.4.

Additional work on sensitivities is considered under Item 2.2.4.

With respect to improving understanding of abundance estimation based on sex ratio data, the Workshop noted that there were a variety of additional analyses that could be helpful to further examine the performance of such estimation procedures but which the group had not had time to carry out. These include further examination of:

(1) model residuals for signs of autocorrelation to check the appropriateness of distributional assumptions inherent in the estimation methods;
(2) plots of profile likelihood functions to better understand the nature of the likelihood surface and to potentially identify reasons for optimization difficulties;
(3) box plots of estimates of *K* from simulated data for each particular *K* value fixed within the procedure used to formulate a simulation-based estimate of the deviance surface.

Such approaches would be useful diagnostics for checking the reliability and suitability of methods and for understanding potential problems.

Beyond such details, the Workshop further identified broad simulation testing as a potentially critical component in confirming the reliability of these novel methods for estimating abundance on the basis of sex ratio data. Such testing would probably consist of generating pseudo-datasets from a variety of stochastic mechanisms (under a variety of scenarios) including those assumed in the estimation model and, more importantly, from mechanisms that differ substantially from the model assumptions.

2.2.4 Recommendations

In developing its recommendations, the Workshop recognised that priorities would have to be set that took into account a number of factors, including importance to the overall aim of deciding whether or not sex ratio methods are sufficient to provide management advice and practical difficulties in the tasks proposed. In cases where the availability of computing resources represented a challenge then participants in the Workshop, especially Allison and Punt, offered to assist in ensuring that runs could be carried out on fast machines during the pre-Annual Meeting period.

PRESENTATION OF RESULTS

Annex E provides detailed **agreements** regarding the format for reporting results from the approaches. Schweder undertook to develop software to facilitate reporting in this manner that would be of great assistance not only to the Workshop but to the developers themselves. This is the first part of a wider exercise to standardise the reporting of results for all the components of the process leading to the development of an AWMP *SLA* and thus facilitate interpretation and understanding.

MINIMISATION

The Workshop **recommends** that the minimisation procedures for both the Brandão-Butterworth and Witting-Schweder approaches use both mechanisms (a) and (b) in Item 2.2.2 to reduce the proportion of non-converged solutions towards zero. Should zero not be reached then developers should present results for upper deviance percentiles with number of minimisation restarts and results for assigning a deviance higher than the highest deviance for converged cases to the non-converged cases.

SENSITIVITIES

As noted above, in order to understand the wide differences in the results for the two approaches discussed under Item 2.2.3 a number of sensitivity analyses were considered.

Highest priority was accorded to tests involving variation of the assumptions associated with the formulation of the likelihood and consequently the generation of simulated catches-by-sex data consistent with these assumptions. The Workshop **requests** that the likelihood given in Annex F should be the base case in such investigations; the primary alternative variants to consider are the overdispersed Poisson (as used in SC/M09/AWMP4) and the logit ratio (as used in SC/M09/AWMP5) or the similarly overdispersed binomial distribution.

Medium priority was accorded to approaches for matching the variances of data generated for the simulations and those for the actual data on the sex composition of the catches (including approaches for reflecting that not all of the catch was sampled for sex every year). There was insufficient time to discuss this in detail. These aspects will be explored to the extent that developers' time allowed and the Workshop encourages developers to consult with the wider AWMP group on this work.

The Workshop further **agrees** that baseline approaches should be run for no more than two alternative catch series to be developed by Givens, to explore there impact on model outputs. It was **agreed** to leave the choice to Givens with further discussion on the choices for these alternatives and the interpretation of the results to be undertaken at the 2009 Annual Meeting.

The Workshop made a number of other suggestions for work on sensitivities including:

(a) adjustments to the 'Witting approach' to handle difficulties caused by zero catch of one or other sex in a year for ratio-based likelihoods (e.g. working with the proportion of females in the catch);

(b) alternative weightings of the log-likelihood to reflect more or less informative data.

However, the Workshop recognised that developers were unlikely to have sufficient time to address these aspects before the 2009 Annual Meeting.

2.2.5 Conclusion

The Workshop **agreed** that considerable progress had been made on a number of issues relating to obtaining a fuller understanding of the two proposed methods and their properties, as well as beginning the validation process in terms of *inter alia* code-checking. However, the primary objective of the work over the last three years has been to ascertain whether or not a sex ratio approach can be developed that is sufficiently robust to provide management advice. It had been hoped that the intersessional work would be sufficient to enable the Scientific Committee to finalise its views on this matter.

In this regard, the Workshop **agreed** that the complexity and importance of the work meant that by Madeira, the Committee would rather be in a position to agree that the work undertaken, including the sensitivity tests outlined under Item 2.2.4, either:

(1) was sufficient to determine that the sex ratio methods, even with considerable further development, were unlikely to be able to supply sufficiently robust information to inform management advice; or

(2) one or both were sufficiently promising to warrant further work in terms of development and more complete robustness testing.

With respect to (2), there was insufficient time available to have a comprehensive discussion of the nature of any further robustness testing, although it was noted that development of a full simulation approach would be a major undertaking.

2.3 Progress with *SLA* development

SLA development for minke whales off West Greenland requires the specification of an operating model. This cannot commence until the Committee has completed its evaluation of the potential for methods based on sex ratio data to estimate population size and, in particular, the lower bound for current population size.

3. FIN WHALES

3.1 Updated abundance estimate

SC/M09/AWMP1 reported on an aerial line transect survey of fin whales conducted off West Greenland in 2007 which was used to estimate the current abundance of fin whales on this summer feeding ground. A total 24 sightings of fin whale groups were collected during 9,433 km of survey effort in sea states <5. Based on conventional distance sampling techniques an abundance of 4,359 whales (95% CI 1,879 – 10,114) was estimated. The survey was conducted as a double platform survey and mark recapture distance sampling techniques could be used for correcting for perception bias which gave an estimate of 4,127 whales (95% CI 1,329 – 12,816). The lower abundance for the corrected estimate may be due to the different truncation for the CDS analysis (250m vs. 800 m). The authors indicated that both estimates are negatively biased because no corrections were applied for whales that were submerged during the passage of the survey plane. The abundance estimate is furthermore only representing the coastal areas of West Greenland and sightings at the westernmost areas of the strata, so the authors note that the entire Baffin Bay-Davis Strait summer abundance of fin whales could be considerably larger.

The Workshop appreciated the revised information given in this paper that had been provided in response to comments made previously (International Whaling Commission, 2008). The Scientific Committee had previously recommended the CDS method pending the clarifications it had requested. Discussion at the Workshop focused on comparison of the CDS and MRDS approaches and which estimate to be recommended for use in assessment.

One difference between the two methods is how pod size is incorporated. In the CDS analysis, the mean pod size is estimated by averaging across all strata. In the MRDS analysis, mean pod sizes differ between strata. This is important for the MRDS analyses because there was a single very large aggregation observed in stratum 9. About 50 whales were photographically recorded overall in this aggregation of smaller groups. The two observers reported only 15 and 25 of these (with some of the rest being off-effort). A pod size of 20 was used in these analyses; this was by far the largest pod size included.

The CDS and MRDS analyses also used different truncation distances and therefore had different total sample sizes. The MRDS used a greater truncation distance because it required more duplicates. Despite the greater sample size for MRDS, the Workshop concurred with the earlier view of the Committee that the amount of data available to support the approach was too limited.

For the MRDS analysis, the stratum contributing the most to the overall abundance estimate was stratum 9, where the large group was observed. For the CDS analysis, it was stratum 4 which had relatively sparse effort over a large area. However, for both analyses sighting rate dominated the variance in the abundance estimate. With respect to stratum 4 in the CDS analysis, a re-analysis presented last year based on pooling the data across all strata (i.e., removing stratification) produced an abundance estimate that was somewhat—but not substantially—lower. This indicates that the CDS analysis is not inappropriately dominated by the contribution from stratum 4.

The results from the CDS and MRDS analyses are fairly similar. Considering the above discussion and the reasons listed last year (International Whaling Commission, 2008), the Working Group preferred the CDS analysis. The Working Group recommended that the estimate of 4,359 whales with 95% confidence interval 1,879 – 10,114 should be used for the purpose of assessment. This estimate is negatively biased because no correction was applied for whales that were submerged during the passage of the survey plane.

3.2 SLA development

Last year, the Committee had **agreed** an approach for providing interim advice for fin whales off West Greenland that is valid for up to two five year blocks (IWC, 2009b). It also **agreed** that safe long-term management is best accomplished under an agreed *Strike Limit Algorithm* (one candidate might be based on the approach for developing interim advice that was agreed last year). The Workshop noted the benefits of having multiple development teams and encouraged members of the Committee to identify potential *SLA*s for fin whales off West Greenland.

Simulation evaluation of *SLA*s requires the development and parameterisation of a set of operating models. The Workshop **agreed** that the set of trials developed to evaluate variants of the *CLA* for North Atlantic fin whales and those used last year to compare alternative methods for providing interim management advice are an appropriate starting point for developing trials for this case. The trials for the North Atlantic fin whales were focused on the areas likely to be subject to whaling off Iceland. These trials will need to be modified to focus more on the uncertainties pertinent to West Greenland if they are to form the basis for evaluation *SLA*s for fin whales. The Workshop **recommended** that a short working paper on appropriate operating models for West Greenland fin whales be presented to the AWMP SWG in Madeira.

4. OTHER ISSUES

No other issues were raised

5. WORKPLAN

The Workshop **agreed** that the priority work given under Item 2.2.4 be carried out in time for the pre-meeting at Madeira.

It also noted that should the Committee agree that one or more methods are appropriate for providing advice on catch limits, then it would be necessary for the computer program(s) for the two approaches and all scenarios to be validated by the Secretariat. It was agreed that, since the programs are written in different languages and since the best form of validation is replication of the results, the validation would be performed by rewriting the models in Fortran into a single program. Allison outlined a number of the validation steps that would be required, noting that some work had already been at least partially completed (see Table 1).

Table 1

The different aspects of the program that will require validation and steps already performed

	Aspect	Checks already performed
1	Biological model	2 models agree
2	Likelihood calculation	Punt has checked likelihoods for some models.
3	Estimation	Partially checked by Punt
4	Generation of data	
5	Calculation of confidence interval	

6. ADOPTION OF REPORT

In concluding the Workshop, Donovan thanked Finn Larsen and Lotte Kindt-Larsen for their assistance. He stressed the importance of trying to reach agreement in Madeira on whether the sex ratio approach was a practical way forward for managing the West Greenland common minke whale fishery. This would require considerable efforts from the developers and he looked forward to receiving the results in Madeira. The participants thanked Donovan for steering them through another complex technical workshop and Jette Donovan Jensen for restaurant advice. The report was adopted at 6pm on 27 March 2009 although additional work was required to finalise Annex E for the report.

REFERENCES

Heide-Jørgensen, M.P., Laidre, K., Hansen, R.G., Rasmussen, M.H., Burt, M.L. and Borchers, D.L. 2008a. Rate of increase and current abundance of humpback whales in West Greenland. Paper SC/M08/AWMP6 presented to the Workshop on Developing Assessment Methods and a Management Procedure for Greenlandic Fisheries, 26-29 March 2008, Copenhagen (unpublished). 22pp. [Paper available at the Office of this Journal].

Heide-Jørgensen, M.P., Laidre, K., Simon, M., Rasmussen, M.H., Burt, M.L. and Borchers, D.L. 2008b. Abundance of fin whales in West Greenland in 2007. Paper SC/M08/AWMP7 presented to the Workshop on Developing Assessment Methods and a Management Procedure for Greenlandic Fisheries, 26-29 March 2008, Copenhagen (unpublished). 10pp. [Paper available at the Office of this Journal].

International Whaling Commission. 2008. Report of the AWMP Workshop on Greenlandic Fisheries, 26-29 March 2008, National Institute of Aquatic Resources (DTU-Aqua), Dept. of Marine Fisheries, Charlottenlund Castle, Copenhagen, Denmark. 19pp. Paper SC/60/Rep2 presented to the IWC Scientific Committee, June 2008, Santiago, Chile. 19pp.

International Whaling Commission. 2009a. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 11:In press.

International Whaling Commission. 2009b. Report of the Scientific Committee. Annex E. Report of the standing working group on the Aboriginal Whaling Management Procedures. *J. Cetacean Res. Manage. (Suppl.)* 11:In press.

Annex A

Agenda

1. INTRODUCTORY ITEMS

1.1 Welcoming remarks
1.2 Election of Chair
1.3 Appointment of rapporteurs
1.4 Adoption of Agenda
1.5 Review of documents

2. COMMON MINKE WHALES

2.1 Updated abundance estimate

2.2 Progress with respect to the use of sex ratio data and modelling approaches

2.2.1 Witting-Schweder
2.2.2 Brandão-Butterworth
2.2.3 Recommendations

2.3 Progress with SLA development

3. FIN WHALES

3.1 Updated abundance estimate

3.2 SLA development

4. OTHER ISSUES

5. WORKPLAN

6. ADOPTION OF REPORT

Annex B

LIST OF DOCUMENTS

SC/M09/AWMP1 Heide-Jørgensen, M.P, Laidre, K.L., Simon, M., Rasmussen, M., Burt, M.L. and D.L. Borchers. Revised abundance estimates of fin whales in West Greenland in 2007

SC/M09/AWMP2 Witting, L., Brandão, A. and D.S. Butterworth. Sex ratio assessment of West Greenland minke whales: Improvements and agreements at the 2008 authors meeting in Cape Town

SC/M09/AWMP3 Heide-Jørgensen, M.P, Witting, L., Laidre, K.L., Hansem R.G. and M. Rusmussen. Fully corrected estimates of minke whale abundance in West Greenland in 2007.

SC/M09/AWMP4 Brandão, A. and D.S. Butterworth. Further results for lower confidence limits for the abundance of West Greenland minke whales

SC/M09/AWMP5 Witting, L. and T. Schweder. Lower confidence bound on population status from catch sex ratio: applied to minke whales off West Greenland

Annex C

Forward distance correction in aerial line transecting for minke whales in West Greenland

TORE SCHWEDER

In SC/M09/AWMP3 a rectangular model is used for forward availability for sighting in aerial line transecting for minke whales in West Greenland. A rectangular window of exposure leads to a correction for forward exposure as that given by Laake *et al.* (1997). As for shipboard surveys of North Atlantic common minke whales, a bivariate curved hazard probability model might be a possibility. Fig. 3 of SC/M09/AWMP3 is consistent with truncation at 300m perpendicular distance, and with uniformly distributed perpendicular distances. This motivates a hazard probability function of the form

$$Q(x,y) = h\,(y) \quad -300 < x < 300,\ 0 < y$$

where y is measured in meters. A possibility is to have Q flat for a stretch forward, and then curve down to zero. With a Gaussian tail forward from the plateau the model is

$$Q(x,y) = \begin{cases} \rho & -300 < x < 300, 0 < y < \beta \\ \rho \exp\left(-\frac{1}{2}\left(\frac{y-\beta}{\sigma}\right)^2\right) & -300 < x < 300, \beta < y. \end{cases}$$

There are two problems to be solved for this model. First, the parameters p, β, σ must be estimated from the observed forward distances and also the duplicate sightings. The other problem is to derive the detection function that follows from this model.

Estimation

Due to the multiplicative nature of Q the estimation splits in two. The parameter p is estimated from the duplicate sightings from the two independent observer, while the two remaining parameters are estimated only from the forward distance data. The probability density for observed forward distances is

$$f(y;\beta,\sigma) = \begin{cases} \left(\beta + \sigma\sqrt{\pi/2}\right)^{-1} & 0 < y < \beta \\ \left(\beta + \sigma\sqrt{\pi/2}\right)^{-1} \exp\left(-\frac{1}{2}\left(\frac{y-\beta}{\sigma}\right)^2\right) & \beta < y. \end{cases}$$

There are two observers, each assumed to have hazard probability $Q(x,y)$. The union of observations from the two yields forward distances $y_1 < y_2 < \ldots < y_n$ where $y_1 < \ldots < y_m$ is a sample from the uniform distribution over $0 < y < \beta$ and the remaining observations has a half-normal distribution shifted to β. Since some of the observations are of the same surfacings, there will be some dependence between them, which is only of minor concern in the estimation of β and σ, and is neglected for now. The number m is binomially distributed with expectation $n\,\beta\left(\beta + \sigma\sqrt{\frac{\pi}{2}}\right)^{-1}$. The maximum likelihood estimator has the form

$$\begin{aligned} \widehat{\sigma}^2_m &= \frac{1}{n-m}\sum_{i>m}(y_i - \beta)^2 \\ \widehat{\beta}_m &= y_m, \end{aligned}$$

and m is determined by maximizing the log likelihood $\sum \log\left(f\left(y;\ \hat\beta_m, \hat\sigma_m\right)\right)$.

Detection function

Assuming the surfacings of minke whales to follow a Poisson point process with intensity α the detection function is

$$\begin{aligned} g(x) &= 1 - \exp\left(-\frac{\alpha}{v}\int_0^\infty Q(x,y)\,dy\right) \\ &= 1 - \exp\left(-\frac{\alpha}{v}\left(\rho\left(\beta + \sigma\sqrt{\pi/2}\right)\right)\right) \quad 0 < x < 300. \end{aligned}$$

Here, v is the speed of the observer. When the speed is high, as is the case for a twin otter airplane, some allowance might be made for the length of surfacings.

Estimation of p

The parameter p captures how able the observer is to see a surfacing. This depends on splash glare etc. and on the length of the surfacing. Let the number of sightings that were picked up by both observer at the time of initial sighting of the individual be k. This number is less than the total number of sightings n. The fraction k/n should be informative on p. The hazard probability of both observer A and B making joint initial sighting is $Q_{AB}(x,y) = Q(x,y)^2$. The detection function for the joint platform is thus

$$\begin{aligned} g_{AB}(x) &= 1-\exp\left(-\frac{\alpha}{v}\int_0^{\infty} Q^2(x,y)\,dy\right) \\ &= 1-\exp\left(-\frac{\alpha}{v}\left(\rho^2\left(\beta+\frac{\sigma}{2}\sqrt{\pi}\right)\right)\right) \quad 0<x<300. \end{aligned}$$

The detection function for the combined platform is

$$g_{A\cup B}(x) = g_A(x) + g_B(x) - g_{AB}(x)$$

The fraction $r(p) = g_{AB} \,/\, g_{A\cup B}$ is a function of p when estimates of the parameters α, β and σ have been used. An estimate of p is found by solving the equation

$$r(p) = \frac{k}{n}$$

Annex D

Results of analyses to examine the convergence of the implementation of the Brandão-Butterworth method

ANDRÉ E. PUNT

The Brandão-Butterworth method (SC/M09/AWMP4) involves fitting a population dynamics model to actual and simulated sex-ratio data. The criterion used to be determine whether the non-linear minimization method used to minimize the objective function (the negative of the logarithm of the likelihood function) converged is that the maximum (over all parameters) of the absolute value of the gradient of the objective function for each parameter is less than 0.001. However, many of the simulations for the closed model did not converge according to this criterion. This Annex explores two ways to increase the proportion of simulations which converge according to this criterion.

A) Restarting the minimization method

This method involves re-applying the minimization method from the parameter vector resulting from the previous application of this method if failure to converge occurs. This approach was tested by generating 1,200 data sets based on *K*=17,000 and the closed model (with the initial values for the remaining parameters set to their maximum likelihood estimates). The minimization method was applied when (a) *K* is fixed to its true value, and (b) *K* is treated as an estimable parameter. Table 1 lists the outcome from this approach by listing: (a) the number of data sets (out of 1,200) for which it was necessary to restart the minimization (e.g. the minimization method failed to converge for 217 of 1200 data sets when *K* was fixed at 17,000), (b) the number of data sets which converged at each stage (e.g. of the 217 attempts to find convergence for the *K*-fixed model on a first restart, 147 actually converged), (c) the number of simulations for which both the *K*-fixed and *K*-estimated models converged, and (d) the upper 5th percentage of the deviance distribution.

Table 1

Summary of the performance of the minimization restart method

	First Minimization	**First restart**	**Second restart**	**Third restart**	**Fourth restart**
(a) Number of minimization attempted					
K fixed case	1200	217	70	37	30
K estimated	1200	148	48	37	34
(b) Number converged at each step					
K fixed case	983	147	33	7	3
K estimated	1052	100	11	3	0
Both cases	915	179	40	7	3
(c) Cumulative number converged	915	1094	1134	1141	1144
(d) Upper 5th percentile of -LnL	1.58	1.96	1.99	1.99	2.03

B) Multiple starting values

This method involves generating several alternative parameter vectors, starting the minimization method for each parameter vector and selecting the parameter vector corresponding to the lower value for the maximum of absolute value of the gradient vector. In the example below, the alternative parameter vectors were constructed based on seven starting values for the *K* parameter (16,000; 16,500, 17,000 18,000; 20,000; 25,0000; 50,000), with the values for the remaining parameters set to their maximum likelihood estimates, Testing this method for 1,000 simulated data sets based on a "true" *K* of 17,000 and a closed model leads to 877 simulations for which convergence occurs if the minimization method is started at a value for *K* of 17,000 and 977 when multiple starting values are used.

Annex E

Format for Reporting Results from the Simulation Approach

A common format was developed for presentation of results from the simulation approach, in order to facilitate comparisons between the approaches.

The key outputs for each simulation should be stored for later processing The minimal information that should be stored for each simulated data set and set of initial parameter values is: (a) the true value of K, (b) the estimated parameters (e.g. K and the parameters that determine the likelihood function), and (c) a measure of whether convergence occurred. The stored information should be used to estimate the rate at which convergence occurs as the number of, for example, initial parameter values, is increased. Plots of convergence rates could be produced for a small number of true K values, in particular for the value of K corresponding to the lower 5th confidence bound for K.

1. Model fit diagnostic plots

A plot of log(sex ratio) against year for scenarios 3, 4b and 5 jointly, for each MSYR as shown in the example below (Fig 1). The scale, tick marks and labels on the sex ratio axis and the scale for a single year should be the same in all plots. The observed sex data should also be shown with an indication of the relative weight, preferably by making the diameter of the point proportional to the relative weight (=√(sample size)). All six plots should be shown on the same page (i.e. 3 plots for MSYR 1% and 3 plots for MSYR=2%).

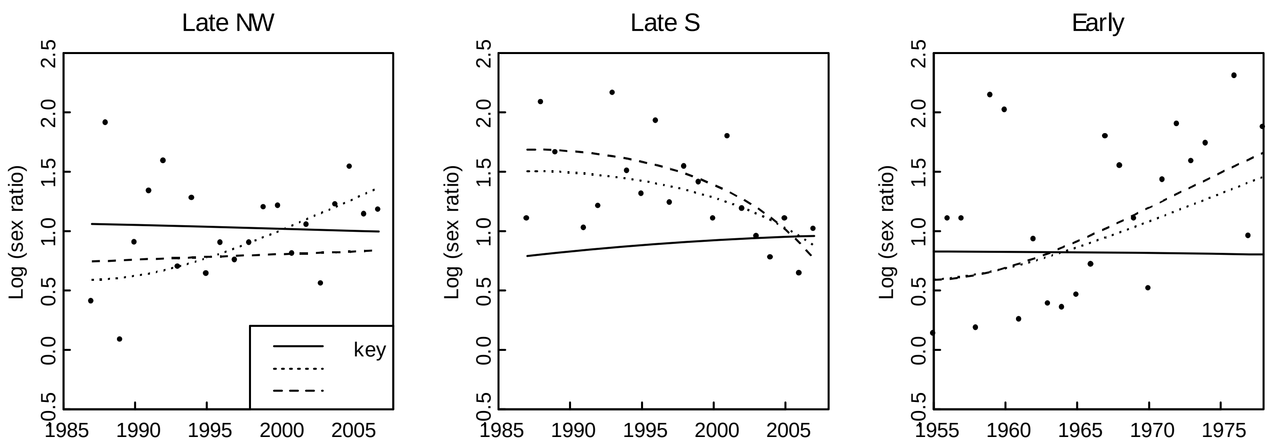

Fig 1. Log(sex ratio) against year for scenarios 3, 4b and 5 jointly, for MSYR=x%.

2. Distributional Assumption diagnostic plots.

A plot of QQnorm of all the [standardised?] residuals (pooled across all three fisheries) for each scenario (3, 4b and 5) and each MSYR (1% and 2%) separately, as shown in the example below (Fig 2). The residuals are from the same log scale as the fitted data. All six plots should be shown on the same page.

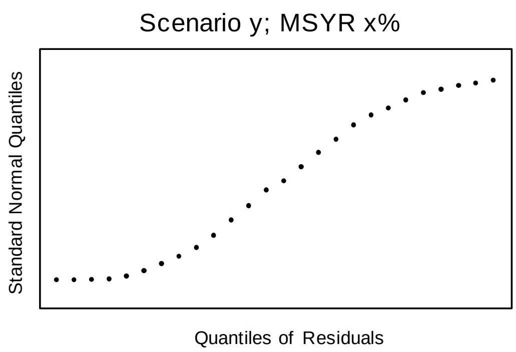

Fig 2. QQ norm of all the residuals (pooled across all 3 fisheries).

3. Results: Model Selection.

An AICC table should be presented listing the negative AICC for every scenario and MSYR as shown (see Table 1 for an example).

Table 1.

Example AICC table

Model Scenario & MSYR		Neg AICC
4b,	2%	112.6
3,	2%	108.7
5,	2%	101.0
4b,	1%	99.9
3,	1%	66.5
5,	1%	62.2
etc		Etc

4. Results: Deviance Profile Plot for the Best/Selected Model.

A single large deviance profile plot should be presented for the selected model i.e. the model with the best AICC, as in the example in Fig 3. The deviance based on the actual data is shown by the solid line and the 2.5%, 5%, 10% and 50% quantiles for the deviance function based on simulated data are shown as dashed lines. d* is the 5%ile. A density curve should be shown on the same plot.

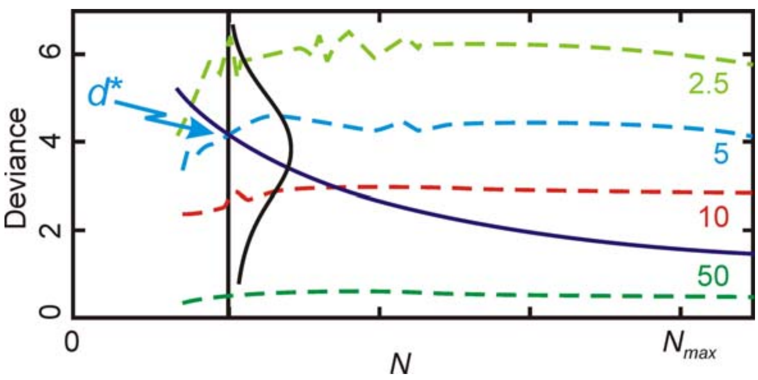

Fig 3. Deviance Profile of the Selected Model: Scenario y, MSYR x%.

5. Results: Deviance Profile Plots

A deviance profile plot should be presented of deviance vs N for each scenario (3, 4b and 5) and each MSYR (1% and 2%) separately, as shown in the example below (Fig 4). The N-axis scale should be the same in all plots with a maximum value N* defined as follows: if $\hat{N}_{10}$ is the point where the 10%ile crossing point occurs, then N* is largest value of $\hat{N}_{10}$. In addition the plot should show two more x-axis scales, corresponding to depletion and K (since N, depletion and K are re-scalings of each other).

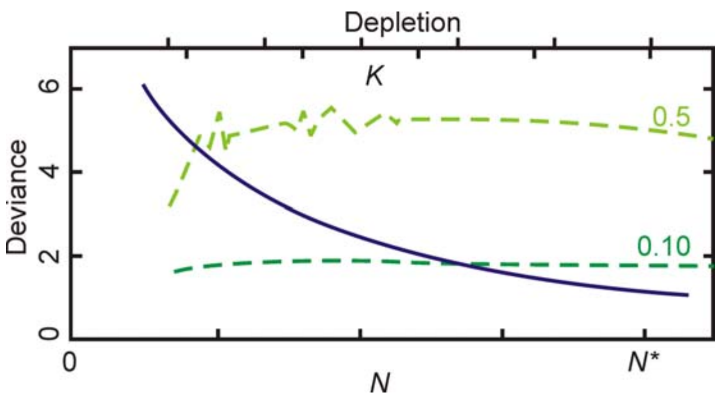

Fig 4. Deviance plot for Scenario y, MSYR x%.

6. Results: Table

A table should be presented listing $\hat{N}_{.05}(\text{low})$ (defined below), $\hat{N}_{.05}$, $\hat{N}_{.05}(\text{high})$, $\hat{N}_{.10}$ and $\hat{N}_{.50}$ for each scenario (3, 4b and 5) and each MSYR (1% and 2%) as shown in the example in Table 2. The model/MSYR combinations should be ranked by – AICC. Similar tables should be presented for Depletion and K.

Table 2.
Example results table for N

Model Scenario & MSYR	$\hat{N}_{.05}(\text{low})$	$\hat{N}_{.05}$	$\hat{N}_{.05}(\text{high})$	$\hat{N}_{.10}$	$\hat{N}_{.50}$
e.g. 4b, 2% Rank by –AICC					

For a quantity Q (where $Q = N$, depletion or K) $\hat{Q}_{.05}(\text{low})$ is, conceptually, the value of Q whose (97.5%) upper confidence limit (with respect to the sampling distribution of a sample quantile just reaches the solid line from below. The definition of $\hat{Q}_{.05}(\text{high})$ is analogous. The concepts are illustrated in Fig 5 below. For exactly 1000 simulations $\hat{Q}_{.05}(\text{low})$ corresponds to the .937$^{\text{th}}$ quantile of the deviance distribution, and $\hat{Q}_{.05}(\text{high})$ to the .964$^{\text{th}}$. This strategy represents both estimation uncertainty and simulation uncertainty.

[For a different number of simulations, Schweder has a formula – to be added].

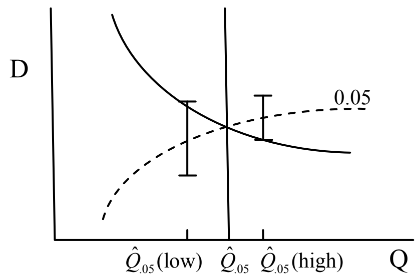

Fig 5. Illustration of the $\hat{Q}_{.05}(\text{low})$ and $\hat{Q}_{.05}(\text{high})$ concept

Note: after the Workshop on 21 May, Tore Schweder noted the following:

I have just now completed the S-script for the figures I find appropriate - somewhat a modification of what we agreed in Copenhagen. Two modifications:

Annex, Figure 3: the deviance distribution is very skewed (nearly a chi-square with df=1). A density is then not as illustrative as a cumulative distribution plot.

Annex, Figure 5: The purpose of this figure is to show the simulation uncertainty. Since the simulations yields one-sided confidence intervals that incorporate both simulation variation and sampling variation in the observed deviance, I find it better to show these intervals.

Annex F

Likelihood for Witting and Schweder and Brandão and Butterworth

TORE SCHWEDER

Witting and Schweder use the likelihood (9) in SC/M09/AWMP5, while Brandão and Butterworth use the likelihood (6) in SC/M09/AWMP4. Taking that of Brandão and Butterworth is motivated by an overdispersed Poisson model, and a normal approximation to that. In my view we should condition on the total catch $C = C^f + C^m$ for each data point (year, area). If anything, the likelihood should be an overdispersed binomial rather than an overdispersed Poisson. One possibility is to assume that C^f is distributed as a normal variate rounded to the nearest integer between 0 and C. To obtain a simple model of over-dispersion, let the normal variate have mean $\mu = Cp$ and variance $\sigma^2 Cp\ (1 - p)$. The degree of overdispersion is given by $\sigma \geq 1$. The modelled proportion of females is $p = ø / (1 + ø)$ where $ø$ is the modelled sex ratio. The log likelihood will then read:

$$\log(L) = \sum_{i=1}^{n}[-\frac{1}{2}\ [\frac{\left(C_i^f - C_i p_i\right)}{\sigma^2 C_i p_i (1-p_i)} + \log(\,C_i p_i\ (1-p_i))] - n log(\sigma)]$$

where the p_i are functions of the appropriate parameters constituting the various models. These are estimated along with $\sigma \geq 1$ by maximizing the log likelihood.

For estimated parameters catches are simulated by

$$\tilde{C}_i^{\,f} = round\ (C_i\widehat{p_i} + Z\hat{\sigma}\sqrt{C_i\widehat{p_i}\ (1 - \widehat{p_i}})$$

$$\tilde{C}_i^m = C_i - \tilde{C}_i^f$$

where $Z \sim N\,(0,1)$