Report of the Third AWMP Workshop on Greenlandic Hunts

Report of the Third AWMP Workshop on Greenlandic Hunts

1. INTRODUCTORY ITEMS

The meeting was held at DTU-Riso, Roskilde, Denmark from 14–17 December 2009. The participants were Donovan (Convenor), Allison, Apostolaki, Brandão, Butterworth, Punt, Schweder and Witting.

1.1 Convenor's opening remarks

Donovan welcomed the participants. He noted that this was primarily a technical workshop to try to ensure that sufficient progress was made to enable the Scientific Committee to make a decision at its next meeting as to whether the sex ratio method for West Greenland common minke whales could be used for management purposes. In addition, it would provide time for initial discussions on the development of long-term *SLA*s for the Greenlandic stocks.

1.2 Election of Chair

Donovan was elected Chair.

1.3 Appointment of rapporteurs

Donovan and Punt acted as rapporteurs.

1.4 Adoption of Agenda

The adopted Agenda is shown in Annex A.

1.5 Documents available

The documents available to the Workshop were relevant extracts of previous meeting reports.

2. PROGRESS WITH INTERSESSIONAL WORK ON THE SEX RATIO METHOD

2.1 Development of code

Allison reported that progress had been made since the 2009 Annual Meeting towards coding the method for calculating lower confidence intervals for carrying capacity based on the sex ratio method. However, there had not been sufficient time to finalise the coding and run the robustness tests.

The Workshop spent considerable time modifying and testing the code used to implement the method for computing lower confidence limits for model outputs such as carrying capacity, current population size and current depletion when assessments are based on the sex ratio method. In addition, code was developed to implement the model-based robustness tests. Modifications to the method led to fewer instances where there was evidence for a lack of convergence of the minimisation method.

3. FINALISING WORK ON THE SEX RATIO METHOD

3.1 Review the specifications for the sex ratio method

The estimates of carrying capacity, *K*, for some of the simulated datasets are smaller than the value of *K* used to generate those datasets. Annex B shows that including the deviance for such cases when computing lower confidence intervals leads to biased estimates of lower confidence intervals. Annex B also shows that setting the deviance to zero in this instance removes this source of bias. The Workshop **agreed** to revise the specifications for how lower confidence intervals are computed to impose the constraint outlined in Annex B. The full specifications for the sex ratio method, the data used for parameter estimation, and method for computing confidence limits for model outputs are given in Annex C.

The Workshop **agreed** that applications of the sex ratio method to estimate lower confidence intervals for model outputs would be conducted for:

(1) assessments based on different models (Closed [three versions – see Item 3.4], Site Fidelity, Influx) and MSYR rates (1% and 2%); and
(2) restrospective analyses (dropping off years of data one by one starting from the most recent year for the 10 most recent years).

3.2 Review the specifications for the existing robustness test

The Workshop revised and clarified the specifications for the robustness tests based on population dynamics models. The output statistics from the robustness tests were expanded (see Annex D) and it was **agreed** that the deviance for any simulated datasets for which the estimate of *K* exceeds the value of *K* used to generate the dataset concerned, would be set to zero.

3.3 Determination of need for any further robustness test

The Workshop **agreed** to add four additional robustness tests. Two of these tests (17 and 18) explore the performance of the sex ratio method when $\text{MSYR}_{1+} = 4\%$ while the other two robustness tests (4 and 5) explore the impact of using the standard version of the Closed model (see Annex C for specifications) when conducting assessments when the Closed-a and Closed-b models are the 'true' models. The final set of robustness trials are listed in Tables 1 and 2.

3.4 Evaluation of new specification of the closed model

The Workshop reviewed an alternative parameterisation of the Closed model (Witting and Brandão, 2010). The Workshop noted that problems had been encountered with finding the maximum likelihood estimates for the parameters for this version of the Closed model due to possible over-parameterisation. It was noted that more parsimonious versions of the Closed model exist which fit the data almost as well as the standard model (a difference of <0.2 likelihood units). The Workshop agreed that two additional variants of the Closed model should be considered for assessments and the robustness tests in addition to the current (standard) Closed model in order to encompass a range of assumptions regarding how the sex ratio changes over time and by area. These variants are:

(1) Closed-a. As for the standard Closed model, except that the proportion of females in the northern and southern area does not change over time (implemented by setting the parameter $\beta^f = 0$).
(2) Closed-b. The proportion of females in the southern area is a time-invariant proportion of the number of females off west Greenland and the sex ratio for the southern area is a logistic function of time.

4. CONSIDERATION OF WORK REQUIRED TO DEVELOP *SLAS* FOR ALL GREENLAND FISHERIES BEFORE THE END OF THE INTERIM PERIOD

In Greenland, a multispecies hunt occurs. The expressed 'need' is for 670 tonnes of edible products from large whales for West Greenland; at present this involves catches of common minke whales, fin whales and bowhead whales – Greenland has also requested a catch of humpback whales from the Commission. The flexibility among species is important to the hunters.

The issue of what is the 'correct' level of need itself is outside the scope of the Scientific Committee. In generic terms, the relevant Governments submit a 'need statement' to the Commission and it is then a Commission decision as to whether to accept that need request. Once that is agreed then the task of the Scientific Committee is to evaluate whether that need request can be achieved within the agreed conservation objectives of the Commission.

Where need is expressed as a number of animals of a particular species/stock this can be a relatively straightforward exercise. However, in developing long-term *SLA*s in the context of a 100-year simulation period, then the Committee (and the Commission) has agreed that it is important to bound the likely levels of future need for testing purposes in order to avoid having to re-evaluate the *SLA* itself every time an increased need request is accepted (should that occur). This bound is termed the 'need envelope' and has been developed by the Chair of the AWMP in conjunction with the hunters. It is important to note that this is a hypothetical upper bound in terms of the robustness of the *SLA* and neither commits the Commission to accepting increased need requests should these be presented nor indeed prevents the submission of need requests greater than the bound at some time in the future. In the latter case, the *SLA* would have to be re-evaluated as the circumstances would be outside the tested parameter space (this could be undertaken in the context of an *Implementation Review* in the same way that other new information might be obtained that led to the conclusion that further Robustness Trials were needed).

For both the bowhead and gray whale *SLA* development process, the need envelope took the form of a linear increase in need from 'current' in year zero to three times that value by the 100th year (see Fig. 1).

However, in the case of a multispecies fishery where need is expressed in terms of numbers of tonnes of edible products, the issue can be more complex. One approach is that conversion factors of tonnes of edible products (per strike) could be used for each species so that conservation implications can be examined in terms of numbers of animals by species/stock. However, the need can be met with many combinations of different catches of the different species. One possible approach would be to develop separate *SLA*s for each species and then develop an algorithm to determine flexibility within safe limits among the species. Further consideration is needed with respect to the issue of a combined approach to catch limits for more than one species which has up until now been the norm, and of course, this will require consultation with both the hunters and the Commission.

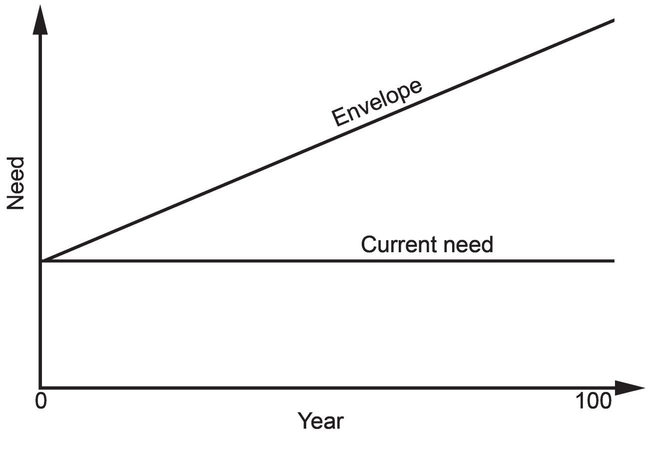

Fig. 1. Need envelope.

At present for the Greenland hunt, the Committee has been asked by the Commission to provide advice by species/stock and the Commission has set strike limits by species/stock. For common minke whales, the Committee had been unable to provide unequivocal advice until the 2009 Annual Meeting when it used the most recent aerial survey abundance and the agreed 'safe interim advice algorithm' adopted two years ago (IWC, 2010, p.139); this advice (178 per year) was less than the limits set by the Commission in 2008 (200 per year in a five-year block quota; IWC, 2008). In the case of fin whales, the situation was similar. Until recently, the Committee was unable to provide advice on strike limits for fin whales but the Commission set a limit of 19 (IWC, 2008). Two years ago, the Committee, using the 'safe interim advice algorithm', agreed that an annual strike limit of 19 would not harm the stock (IWC, 2009). In the case of bowhead and humpback whales, the 'safe interim advice algorithm' was used to determine that strike limits proposed by Denmark on behalf of Greenland (2 and 10 animals per year, respectively) would not harm the stocks (IWC, 2009).

4.1 Future work

The Workshop considered this briefly but agreed that full consideration for common minke whales should await the results of work on the sex-ratio method. The present assessment method for fin whales could prove a starting point for consideration of an *SLA* for fin whales. The question of flexibility amongst species in terms of meeting a total yield of edible products requires considerable further thought as well as consultation with hunters and the Commission. In the meantime, the Workshop **agreed** that it would be useful for a working paper to be developed for the 2010 Annual Meeting that summarises for both fin whales and common minke whales in terms of:

(1) the existing stock structure hypotheses used in the RMP *Implementations*;
(2) the available information and data on stock structure, with an emphasis on the western and central North Atlantic;
(3) the available information on abundance.

5. WORK PLAN

The Workshop **agreed** to the following work plan in the period before the Annual Meeting:

(1) undertake the computing work and running of the revised trials on the sex-ratio method outlined under Item 3 (Allison, Schweder, Witting); and
(2) provide a short working paper on *inter alia* the stock structure hypotheses for North Atlantic fin and common minke whales used in the RMP *Implementations* as suggested under Item 4 (Donovan and Punt).

6. ADOPTION OF REPORT

The report was adopted by e-mail.

REFERENCES

International Whaling Commission. 2008. Chair's report of the 59th Annual Meeting. Annex L. Amendments to the Schedule adopted at the 59th Annual Meeting. *Ann. Rep. int. Whal. Comm.* 2007: 131.

Table 1

List of robustness tests based on population models (unless specified otherwise the estimator should match the population model used to generate the data).

Case	Population model	K[1]	Over-dispersion	MSYR	Other
1	Influx	150,000	Estimated	2%	
2	Closed	150,000	Estimated	2%	
3	Model 5	150,000	Estimated	2%	
4	Closed-b	150,000	Estimated	2%	Closed estimator
5	Closed-a	150,000	Estimated	2%	Closed estimator
6	Influx	75,000	Estimated	2%	
7	Influx	50,000	Estimated	2%	
8	Influx	20,000[1]	Estimated	2%	
9	Influx	150,000	Estimated	2%	Closed estimator
10	Closed	150,000	Estimated	2%	Influx estimator
11	Influx	150,000	2 × estimated	2%	
12	Influx	150,000	1	2%	
13	Influx	150,000	Estimated	1%	
14	Closed	150,000	Estimated	1%	
15	Influx	150,000	Estimated	2%	+ 20-yrs extra data
16	Closed	150,000	Estimated	2%	+ 20-yrs extra data
17	Influx	150,000	Estimated	4%	
18	Closed	150,000	Estimated	4%	

Table 2

Specifications for the model-free robustness tests (separately for each dataset and for all datasets at the same time).

Case	Slope	Mean
1	Unchanged	Unchanged
2	+0.05	Unchanged
3	–0.05	Unchanged
4	Unchanged	+0.05
5	Unchanged	–0.05

International Whaling Commission. 2009. Report of the Scientific Committee. Annex E. Report of the standing working group on the Aboriginal Whaling Management Procedure. *J. Cetacean Res. Manage. (Suppl.)* 11: 145–68.

International Whaling Commission. 2010. Report of the Scientific Committee. Annex E. Report of the standing working group on the Aboriginal Whaling Management Procedure. *J. Cetacean. Res. Manage. (Suppl.)* 11(2): 135–53.

Witting, L. and Brandão, A. 2010. Report of the Scientific Committee. Annex E. Report of the standing working group on the Aboriginal Whaling Mangement Procedure. Appendix 5. Closed models for West Greenland minke whales. *J. Cetacean Res. Manage. (Suppl)* 11(2): 151–52.

Annex A

Agenda

1. Introductory items
    1.1 Convenor's opening remarks
    1.2 Arrangements for the meeting
    1.3 Election of Chair
    1.4 Appointment of rapporteurs
    1.5 Adoption of Agenda
    1.6 Documents available

2. Progress with intersessional work on the sex ratio method
    2.1 Development of code
    2.2 Results from first runs of the robustness test

3. Finalising work on the sex ratio method
    3.1 Review the specifications for the sex ratio method
    3.2 Review the specifications for the existing robustness test
    3.3 Determination of need for any further robustness test
    3.4 Evaluation of new specification of the closed model

4. Consideration of work required to develop *SLA*s for all Greenland fisheries before the end of the interim period

5. Work plan

6. Adoption of Report

Annex B

One-Sided Confidence Limits from Likelihood Functions

Tore Schweder

One-sided confidence intervals might be required for subject matter reasons, or simply because data are insufficient to provide ordinary two-sided confidence intervals. In case of the abundance of minke whales subject to aboriginal subsistence whaling off West Greenland, both reasons apply. To enable safe management, lower confidence bounds are required. Also, the data on sex ratios in the catches, to be considered below, yield a likelihood function with maximum for infinite abundance.

In general terms, let θ be a scalar parameter of interest, and let ζ be a vector of nuisance parameters. The data X yield a likelihood function $L(\theta;\zeta;X)$, profile likelihood $L(\theta;X) = L(\theta;\hat{\xi}(\theta);X)$ where $\hat{\xi}(\theta) = \arg\max L(\theta;\zeta;X)$ and with the maximum likelihood estimate $\hat{\theta}$.

The profile deviance function is $D(\theta;X) = 2\log(L(\hat{\theta};X)/L(\theta,X))$ is assumed well behaved, with a left branch declining continuously to zero at $\hat{\theta}$ and a right branch increasing continuously from this point. Traditional confidence intervals are obtained by probability scaling the profile deviance function by the chi-square distribution. Thus the left limit is obtained from the left branch, and the right endpoint of the confidence interval is found from the right branch. Lower one-sided confidence limits with degree above 50% are found from the left branch of the deviance function, while upper one-sided confidence limits are found from the right branch.

Focusing on one-sided confidence intervals of the form $(\theta^L(\alpha;x),\infty)$ based on observed data $X = x$ with degree α not too far below 0 (>0.5). Lower confidence limits are obtained from the left deviance branch:

$$LD(\theta;X) = D(\theta;X)I(\hat{\theta} > \theta)$$

From probability calculation or simulation, the cumulative null distribution of the left deviance branch is

$$F^{LF}(\upsilon;\theta) = P_\theta(LD(\theta;X) \leq \upsilon)$$

The left confidence curve for an observation $X = x$ is defined as the probability re-scaled left deviance branch re-scaled left deviance branch:

$$LC(\theta;x) = F^{LD}(LD(\theta;x);\theta)$$

for $LD(\theta;x)>0$, i.e. the $(\theta, LD(\theta;x))$ on the left deviance branch,

$$P_\theta(LD(\theta;X) \leq LD(\theta;x)) = LC(\theta;x)$$

Since the left confidence curve is decreasing for $\theta<\hat{\theta}$, $LC(\theta;x)<LC(\theta^L;x)$, is equivalent to $\theta\geq\theta^L$. With $\theta^L = \theta^L(\alpha;x)$ solving $\alpha = LC(\theta^L;x)$ the statement $\theta^L\geq\theta^L(\alpha;x)$ has confidence α and $\theta^L(\alpha;x)$ is indeed a lower confidence limit for a one-sided confidence interval of degree α.

Illustration

To illustrate the basic argument consider the simplest of cases with a sufficient statistic X being normally distributed with mean θ and unit variance. Here a lower confidence limit

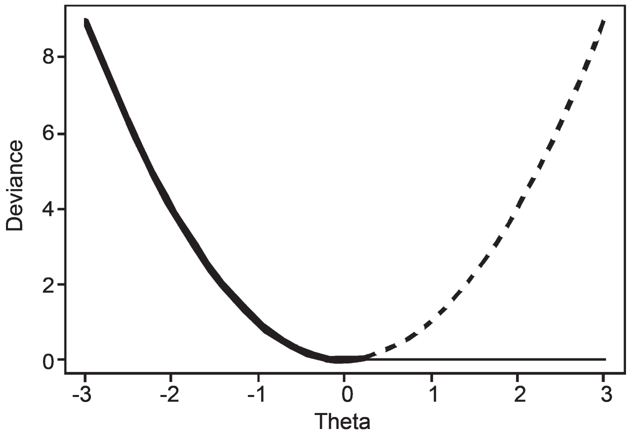

Fig. 1. The normal deviance function (dashed) and its left branch (solid line) for an observation $x = 0$.

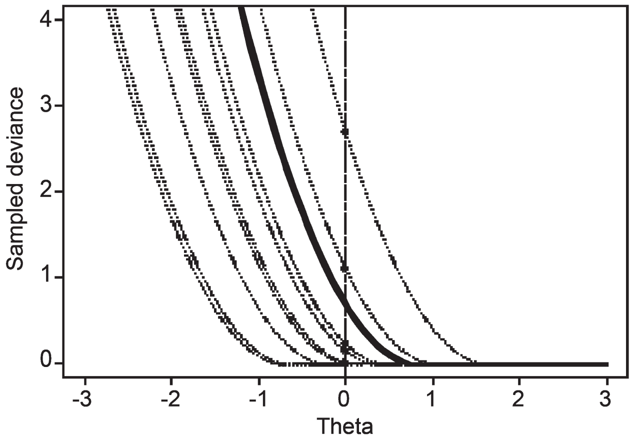

Fig. 2. One observed left deviance branch (solid line) and 10 simulated left deviance branches (dotted). Normal likelihood, $\theta = 0$.

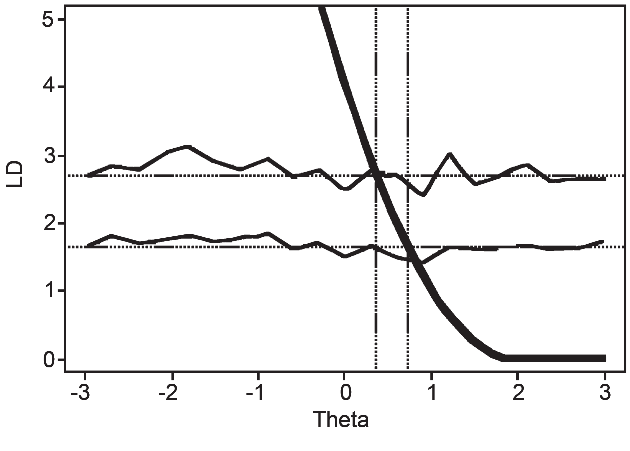

Fig. 3. The left deviance branch from an observation $x = 2$ from a $N(\theta,1)$ distribution (solid line). The dashed curves are simulated 95% and 90% quantiles (1,000 replicates) of $LD(\theta;X)$ and the dotted lines are the respective theoretical quantiles $\Phi\text{–}1(\alpha)$; $\alpha = 0.9$; 0.95. The vertical dotted lines are at the lower α confidence limit $\theta L(\alpha;x) = x - \Phi\text{–}1(\alpha)$.

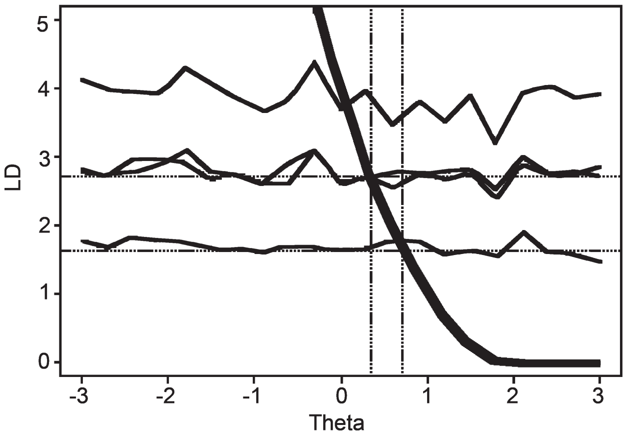

Fig. 4. Left deviance branch for $x = 2$ with correctly simulated quantile curves (95% and 90%), theoretical quantile curves, and quantiles 95% and 90% quantiles obtained by simulating the deviance rather than the left deviance branch. The correct 95% quantile coincides with the 90% quantile for the deviance.

based on an observation $X = x$ is $\theta^L(\alpha;x) = x - \Phi^{-1}(\alpha)$. This is easily argued from the pivotal quantity θ–X.

The same result is obtained from considering the monotonous left branch of the deviance function:

$$LD(\theta;x) = D(\theta;x)I(\theta - x) = (\theta < x)^2\ I(\theta < x)$$

where I is the indicator function and D is the deviance. Fig. 1 shows the normal deviance as a dashed curve, and its left branch as a solid line. To probability scale the observed left branch in Fig. 2 at $\theta(= 0)$, one might simulate $X \sim N(\theta,1)$ to obtain an empirical distribution of $LD(\theta;X)$. Ten simulated left deviance branches are shown in Fig. 2, and the sample of $LD(\theta;X)$ are shown as the intersections with the vertical line at $\theta = 0$. The frequency of $LD(\theta;X) \leq LD(0;X)$ is 8/10 in this illustration. The confidence of $\theta \geq 0$ is thus simulation estimated to 8/10, and the lower confidence limit of degree $\alpha > 0.8$ is below 0. Fig. 3 shows the left deviance branch and quantiles based on simulating 1,000 replicates for each of 21 values of θ. Fig. 4 shows the same, and also the result of simulating the deviance rather than the left deviance branch. As expected, the 90% quantile for the deviance coincides with the 95% quantile for the left deviance. This is so because the deviance function is symmetric.

Simulation is not necessary in this sample case. For $\upsilon > 0$:

$$LD(\theta;X) \leq \upsilon \Leftrightarrow LD^{-1}\ (\upsilon;X)$$

Thus, $LD^{-1}(\upsilon;x)$ is a lower confidence limit of degree $P_\theta(LD(\theta;X) \leq \upsilon) = \alpha$. In the sample normal case:

$$P_\theta(LD(\theta;X) \leq \upsilon) = \Phi(\sqrt{\upsilon}) = \alpha$$

For $\upsilon > 0$, and $\theta^L(\alpha;x) = LD^{-1}(\upsilon;x) = x - \sqrt{\upsilon} = x - \Phi^{-1}(\alpha)$ is indeed the required one-sided lower confidence limit.

Annex C

Specifications of West Greenland Minke Whale Sex Ratio Method Robustness Trials

Population dynamics

A sex-structured age-structured model is used:

$$N^g_{t+1,0} = 0.5\rho_{t+1}N^{mat(f)}_{t+1} \qquad \text{(P1)}$$

$$N^g_{t+1,a+1} = (N^g_{t,a} - C^g_{t,a})\ s_a \qquad 0 \leq a \leq x\text{–}2 \qquad \text{(P2)}$$

$$N^g_{t+1,x} = (N^g_{t,x} - C^g_{t,x})\ s_x + (N^g_{t,x-1} - C^g_{t,x-1})\ s_{x-1} \qquad \text{(P3)}$$

where:

g is the whale gender (male/female),

$N^g_{t,a}$ is the total number of minke whales of gender g of age a in year t,

$N_{t,a}$ is the total number of minke whales of age a in year t, which is given by: $N_{t,a} = N^m_{t,a} + N^f_{t,a}$,

$C^g_{t,a}$ is the number of West Greenland minke whales of gender g of age a caught in year t,

ρ_{t+1} is the fecundity rate for year t+1, which takes the Pella-Tomlinson form:

$$\rho_{t+1} = b_k + [b_{\max} - b_k]\left[1 - (N^{1+}_{t+1}/K^{1+})^z\right] \quad \text{where} \qquad \text{(P4)}$$

b_k is the birth rate at carrying capacity K,

b_{max} is the maximal birth rate,

z relates to the strength of density dependence (the degree of compensation parameter).

N^{1+}_{t+1} is the total number of 1+ minke whales in year t+1, given by:

$$N^{1+}_{t+1} = \sum_{a=1}^{x}(N^f_{t+1,a} + N^m_{t+1,a}) \qquad \text{(P5)}$$

K^{1+} is the carrying capacity of 1+ minke whales, given by:

$$K^{1+} = \sum_{a=1}^{x}(N^f_{1948,a} + N^m_{1948,a}) = N^{1+}_{1948} \qquad \text{(P6)}$$

$N^{mat(f)}_{t+1}$ is the number of mature females at the start of year t+1, given by:

$$N^{mat(f)}_{t+1} = \sum \gamma_a N^f_{t+1,a}, \text{ where} \qquad \text{(P7)}$$

$$\gamma_a = \begin{cases} 0 & a < a_{mat} \\ 1 & a \geq a_{mat} \end{cases} \quad \text{and } a_{mat} \text{ is the age of reproductive maturity;} \qquad \text{(P8)}$$

x is the maximum age considered (i.e. the 'plus group').

s_a is the age specific annual survival rate, given by:

$$s_a = \begin{cases} s_{juv}s_{ad} & a = 0 \\ s_{juv} & 1 \leq a \leq a_{ad}, \\ s_{ad} & a > a_{ad} \end{cases} \quad \text{where:} \qquad \text{(P9)}$$

s_{juv} is the survival rate for juveniles, s_{ad} is the survival rate for adults, and

a_{ad} is the greatest age at which the 'juvenile' survival rate applies (taken as a_{ad}= 1).

The number of whales of gender g of age a caught in year t is given by:

$$C^g_{t,a} = C^g_t R^g_{t,a} / \sum_{a'=0}^{x} R^g_{t,a'} \qquad \text{(P10)}$$

where:

C^g_t is the total catch of minke whales of gender g in year t,

$R^g_{t,a}$ is the number of recruited minke whales of age a and gender g in year t, given by

$$R^g_{t,a} = G_a N^g_{t,a} \quad \text{where} \qquad \text{(P11)}$$

G_a is the age-specific differentiation of the catch relative to the age composition of the overall population, given by:

$$G_a = \begin{cases} 0 & a = 0 \\ a / a_c & 1 \leq a < a_c, \\ 1 & a \geq a_c \end{cases} \qquad \text{(P12)}$$

where a_c is the age at full recruitment.

Parameters

The parameter values used are given in Table 1.

Table 1

The parameters used.

Parameter			
Age of adult a_{ad}	1		
Age at maturity a_{mat}	7		
Age at first capture a_c	6		
Maximum age considered (the 'plus group') x	6		
Adult survival s_{ad}	0.91		
	MSYR = 1%	**MSYR = 2%**	**MSYR = 4%**
Juvenile survival s_{juv}	0.646321	0.724728	0.886542
Maximal birthrate b_{max}	0.94	0.94	0.94
Birthrate at equilibrium b_k	0.886542	0.603496	0.403297
Strength of density dependence z	1.164367	2.052816	1.724999

Initial setup

The numbers for each gender g at each age a in the pristine population are set up as follows:

$$N^g_{1948,a} = 0.5\eta N^*_a \quad \text{where} \quad \eta = \mathrm{K} \Big/ \sum_{a=0}^{x} N^*_a \quad \text{and} \qquad \text{(P13)}$$

$$N^*_a = \begin{cases} 1 & a = 0 \\ s_{a-1}N^*_{a-1} & 1 \leq a < x - 1 \\ s_{x-1}N^*_{x-1}/(1 - s_x) & a = x_c \end{cases} \qquad \text{(P14)}$$

Fishery models

Three (of an original six) models form the focus of this work (Fig. 1): (IWC, 2009, pp. 21)

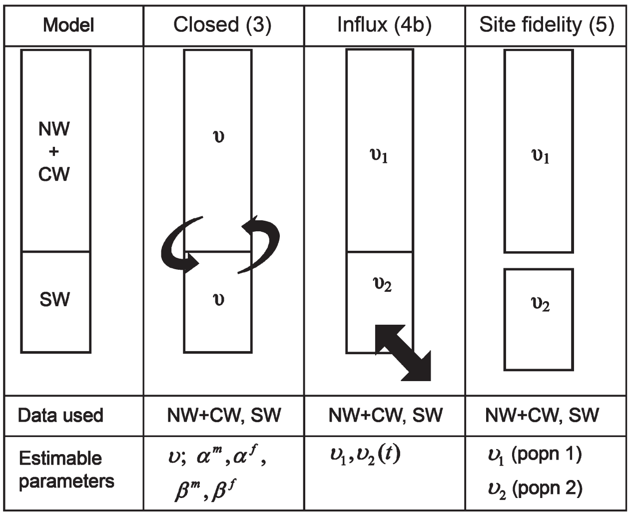

Fig. 1. Sex-ratio models for common minke whales off West Greenland. ν determines the degree of sex imbalance; α and β determines how ν changes over time.

(1) **'Closed' model (previously model 3)**. West Greenland is divided into two strata (NW+CW and SW); data for the two strata are included separately in the likelihood function. Separate (and time invariant) values for the parameters determining the degree of sex imbalance are estimated for each stratum. Allowance is made for time-dependent exchange of females and males between the two strata.

(2) **'Influx' model (4b).** As for the 'closed' model, except that there in no time-dependent exchange; rather the parameter determining the fraction of males in the SW stratum is assumed to change over time (or as a function of temperature).

(3) **'Site Fidelity' model (5)**. The animals in the NW+CW and SW strata exhibit site fidelity. For computational simplicity, this model is implemented by treating the animals in the NW+CW and SW strata as separate populations.

The number of minke whales caught in the whaling fishery/period *i* in year *t* **which are known by sex** is given by:

$$C_t^i = C_t^{m(i)} + C_t^{f(i)} \quad \text{where} \tag{M1}$$

$C_t^{m(i)}$ is the number of males caught in fishery/period *i* in year *t* and is $C_t^{f(i)}$ the corresponding catch of females.

'Site Fidelity' Model (5)

In the **'Site Fidelity' model** (5), the expected number of female minke whales caught by each fishery/period *i* in year *t* is given by:

$$\hat{C}_t^{f(i)} = C_t^i \frac{R_t^f}{R_t^f + \lambda^i R_t^m}, \quad \text{where } R_t^g = \sum_{a=0}^{x} R_{t,a}^g \tag{M2}$$

$R_{t,a}^g$ is defined by equation P11 and

λ^i is the selectivity of males relative to females for the fishery and period concerned, which is assumed to remain constant over that period, with equation (M1) following from the associated assumptions that:

$$\hat{C}_t^{f(i)} = F_t^{(i)} R_t^f; \quad \hat{C}_t^{m(i)} = \lambda^i F_t^{(i)} R_t^m \text{ , and} \tag{M3}$$

In this model the period/fishery North and South strata are run separately (using either the total Northern or the total Southern catch series) and so *i* represents:

$$i = \begin{cases} I & \text{the early Northern or the early Southern Greenlandic fishery (1955–1978)} \\ II & \text{the Northern or the Southern Norwegian fishery (1968–1985)} \\ III & \text{the late Northern or the late Southern Greenlandic fishery (1987–2007)} \end{cases}$$

'Influx' Model (4b)

In this model the fishery/period *i* represents:

$$i = \begin{cases} I & \text{the early Greenlandic fishery (1955–1978)} \\ II & \text{the late Southern Greenlandic fishery (1987–2007)} \\ III & \text{the late Northern Greenlandic fishery (1987–2007)} \end{cases}$$

The number of male whales in the SW stratum is assumed to be influenced by whales moving in from other areas in recent years. To effect this, the λ parameter in fishery *III* is assumed to vary linearly over the period from 1987 to 2007 when these Greenlandic catches were sampled for sex i.e. expected number of female minke whales caught in fishery/period *III* in year *t* is given by:

$$\hat{C}_t^{f(III)} = C_t^{III}\frac{R_t^f}{R_t^f + \lambda_t^{III}R_t^m} \tag{M4}$$

where

$$\lambda_t^{III} = \frac{\lambda^{III}}{1 + (t - 1987)\left((\lambda^{III}/\gamma) - 1\right)/(2007\text{–}1987)} \quad \text{so} \quad \lambda_{2007}^{III} = \frac{\lambda^{III}}{1 + \left((\lambda^{III}/\gamma) - 1\right)} = \gamma \tag{M5}$$

The expected number of female minke whales caught in fishery/periods *I* and *II* in year *t* ($\hat{C}_t^{f(I)}$ and $\hat{C}_t^{f(II)}$) are set as in equation M2 above.

‘Closed’ Model (3)

In this model the fishery/period *i* represents:

$$i = \begin{cases} I & \text{the early Greenlandic fishery (1955–1978)} \\ II & \text{the late Southern Greenlandic fishery (1987–2007)} \\ III & \text{the late Northern Greenlandic fishery (1987–2007)} \end{cases}$$

Model 3 is a closed population model with time dependence of the fractions of females and of males distributed in the NW + CW and SW strata in the later period of Greenlandic whaling commencing in 1987. The expected number of female minke whales caught in fishery/period *I* in year *t* ($\hat{C}_t^{f(I)}$) is set as in equation M2 above.

The proportion of whales of gender *g* in the northern area is assumed to change with time during the recent period of Greenlandic whaling as:

$$p_t^g = R_t^{n,g} / R_t^{WG,g} = e^{\alpha^g + \beta^g\tilde{t}} / \left[1 + e^{\alpha^g + \beta^g\tilde{t}}\right] \tag{M7}$$

where $R_t^{n,g}$ is the number of recruited whales in the northern area of gender *g* in year *t*.

Since the overall West Greenland (WG) area is closed, it follows that $R_t^{n,g} = R_t^{WG,g} - R_t^{s,g}$ where *n* denotes the northern and *s* the southern area. So the proportion of whales of gender *g* in the SW stratum is given by $1 - p_t^g$. With two genders, there are four parameters ($\alpha^m,\beta^m,\alpha^f,\beta^f$) to describe the allocation of the two sexes between the two areas.

Thus, the expected number of female minke whales caught in fisheries *II* and *III* in year *t* is given by

$$\hat{C}_t^{f(II)} = C_t^{II}\frac{R_t^f p_t^f}{R_t^f p_t^f + \lambda_t^{III}R_t^m\, p_t^m} \tag{M8}$$

and

$$\hat{C}_t^{f(III)} = C_t^{III}\frac{R_t^f(1 - p_t^f)}{R_t^f(1 - p_t^f) + \lambda_t^{III}R_t^m\,(1 - p_t^m)} \tag{M8}$$

Note: in the ‘Closed’ model λ^{II} and λ^{III} are identical by definition.

The time *t* is specified by $\tilde{t} = t - 1987$.

In the code the α and β parameters are re-parameterised in terms of the proportions in years 1987 and 2007 (p_{87}^g and p_{07}^g):

$$\alpha^g = \frac{\tilde{t}_{07}\ 1\mathrm{n}\left(p_{87}^g/(1 - p_{87}^g)\right) - \tilde{t}_{87}\ 1\mathrm{n}\left(p_{07}^g/(1 - p_{07}^g)\right)}{(\tilde{t}_{07} - \tilde{t}_{87})} \quad \text{and} \tag{M9}$$

$$\beta^g = \frac{1\mathrm{n}\left(p_{07}^g/(1 - p_{07}^g)\right)\ 1\mathrm{n}\left(p_{87}^g/(1 - p_{87}^g)\right)}{(\tilde{t}_{07} - \tilde{t}_{87})} = \frac{1\mathrm{n}\left(p_{07}^g/(1 - p_{07}^g)\right)\ 1\mathrm{n}\left(p_{87}^g/(1 - p_{87}^g)\right)}{(2007\text{–}1987)} \tag{M10}$$

The likelihood function

The **Schweder likelihood** function is used (see IWC, 2009). The negative of the approximate log-likelihood (ignoring constants) which is minimised in the fitting procedure is given by:

$$-1\mathrm{n}\,L = \sum_{i=I}^{III}\sum_{y=y_i^1}^{y_i^*}\left\{\frac{1}{\sigma_i^2}\left(C_y^{f(i)} - \hat{C}_y^{f(i)}\right)^2 \Big/ \left(2\,\hat{C}_y^{f(i)}\left(1 - \frac{\hat{C}_y^{f(i)}}{C_y^i}\right)\right) + 1\mathrm{n}\,\sigma_i + 1\mathrm{n}\sqrt{\hat{C}_y^{f(i)}\left(1 - \frac{\hat{C}_y^{f(i)}}{C_y^i}\right)}\right\} \tag{L1}$$

where

y_i^1 is the first year of catches for period i,

y_i^* is the last year of catches for period i,

σ_i measures overdispersion of the distribution of catches compared to a Poisson distribution for which the variance is equal to the expected catch for the period and fishery concerned, whose maximum likelihood estimate is given by:

$$\sigma_i = \sqrt{\frac{1}{n_i}\sum_{y=y_i^1}^{y_i^*}\left\{\left(C_y^{f(i)} - \hat{C}_y^{f(i)}\right)^2 \Big/ \left(2\,\hat{C}_y^{f(i)}\left(1 - \frac{\hat{C}_y^{f(i)}}{C_y^i}\right)\right)\right\}}. \tag{L1}$$

n_i is the total number of years in the summation of each whaling period.

Simulation algorithm

For the best estimate of virgin biomass (*K*) (here taken to be 200,000 as a surrogate for infinity), the models are fitted to the original data to obtain estimates for the overdispersion (σ’s) and the selectivity of males relative to females (λ’s) for the period and fishery concerned. Then, for a given value of the true virgin biomass (*K*), and the overdispersion as estimated for *K* =

200,000, the models are fitted to the original data to obtain estimates of the λ's. For each model the deviance as a function of K is obtained for the original data.

The following method is used to calculate the confidence intervals:

(a) Estimate the overdispersion (σ's) for $K = \infty$ (taken as $K = 200{,}000$) using the true catch data.
(b) For each of a specified set of values of K (denoted K_L):
  (c) Estimate the parameter values (denoted P_L) for fixed K ($= K_L$) using the true catch data and using P_{L-1} to initialise the run;
  (d) Generate the catch data using $K = K_L$, the parameters P_L (except for the dispersion) and the σ's from step a using the following method:
    (i) set up the pristine population (equation P13);
    (ii) generate the catch data (for each year generate the catch data and then project the population forward by one year removing the generated catches (equations P1–3))
  (e) Fit the model to the catch data generated in step (d) with for the case when K is fixed ($= K_L$) [initialised with P_L] and the case when K is estimated [initialised using the fixed K fit parameters just obtained] to give the deviance value for the generated data[1].
  (f) Steps (d) to (e) are repeated 1,000 times to get the distribution of the deviance values corresponding to $K = K_L$.

Data generation

The data generation takes into account that not all whales are sampled for sex, and that there is a period over which both Norwegian and Greenlandic catches occurred. The assumption is made that the Norwegian catch was always fully sampled, so the sampled Greenland catch is generated from the total Greenland catch each year.

The values of the selectivity parameters (λ^i and γ) used below are those estimated for the value of K under consideration whereas the overdispersion estimates (σ_i) are those estimated for $K = 200{,}000$.

Data generation for the Influx model (4b)

The catches by sex for the sampled animals (animals for which sex is known in the actual data) are generated under the assumption that they are governed by an overdispersed binomial distribution as detailed below. The catches by sex for the unsampled animals are then generated using the same approach. The removals by sex from the population are the sum (by sex) of the sampled and unsampled catches. In the data generation algorithm described below, in instances in which a negative catch is generated for one of the sexes, the catch for that sex is set to zero and consequently the catch for the opposite sex is set to the total number being sampled.

C_t^i is the catch from Fishery i catch in year t which **is known by sex** (see equation M1).

In the period 1948–86 the numbers of catches sampled by sex ($\tilde{S}_t^{f(I)}$ and $\tilde{S}_t^{m(I)}$) for the early Greenlandic fishery are generated from the normal distribution given by equation D.1.

$$\tilde{S}_t^{f(i)} = \mathrm{N}\left(\frac{R_t^f}{R_t^f + \lambda^i R_t^m} C_t^i,\quad \sigma_i^2 \frac{R_t^f}{R_t^f + \lambda^i R_t^m} C_t^i\right), \text{ and } \tilde{S}_t^{m(i)} = C_t^i - \tilde{S}_t^{f(i)} \tag{D1}$$

where i = I (the early Greenlandic fishery 1955–1978). In the years 1948–1954 and 1979–1986 there was no sampling so C_t^i and $\tilde{S}_t^{f(I)} = \tilde{S}_t^{m(I)} = 0$.

The numbers of unsampled catches by sex ($\tilde{U}_t^{f(I)}$ and $\tilde{U}_t^{m(I)}$) from 1948–86 are generated from the normal distribution given by equation D.2.

$$\tilde{U}_t^{f(i)} = \mathrm{N}\left(\frac{R_t^f}{R_t^f + \lambda^i R_t^m} C_t^{U(i)},\quad \sigma_i^2 \frac{R_t^f}{R_t^f + \lambda^i R_t^m} C_t^{U(i)}\right), \text{ and } \tilde{U}_t^{m(i)} = C_t^{U(i)} - \tilde{U}_t^{f(i)} \tag{D2}$$

where i = I, $C_t^{U(I)} = C_t^{WG} - C_t^i$ and C_t^{WG} is the total Greenlandic catch in year t. (D3)

The generated sampled and unsampled numbers are added to the (known) Norwegian catches by sex to give the total generated catch by sex

$$\text{e.g. } \tilde{C}_t^f = C_t^{f,Nwy} + \tilde{S}_t^{f(I)} + \tilde{U}_t^{f(I)} \tag{D4}$$

Note the total catch in year t, is the sum of the Greenlandic and Norwegian catches in that year, i.e

$$C_t = C_t^{WG} + C_t^{Nwy} \tag{D5}$$

In the period 1987–2007 the total Greenland catches (C_y) are split into the NW+CW and SW strata (C_t^{NW+CW} and C_t^{SW}), where these are taken to be the observed data as used for Model 5 (NW+CW) and Model 5 (SW).

The numbers of sampled and unsampled catches ($\tilde{S}_t^{f(II)}$, $\tilde{S}_t^{m(II)}$, $\tilde{U}_t^{f(II)}$ and $\tilde{U}_t^{m(II)}$) for the period 1987–2007 Northern strata are generated in a similar manner to those for 1946–86 using equations D1 and D2 with $i = II = NW{+}CW$ i.e. the λ^i, σ_i correspond to the Greenlandic (1987–2007) period and

$$C_t^{U(II)} = C_t^{NW+CW} - C_t^{II} \tag{D6}$$

---

[1] When fitting the model the approach of Brandão and Butterworth (2009) is used to overcome convergence problem, using a combination of several re-runs and various initialisation values of K. The procedure adopted consists of seven different K initialisation values, which span the range of K values as proportions of the true K, with one of these Ks being the true K. For each set, the one with the lowest negative log-likelihood is chosen and convergence checked. If the convergence criterion is not met, the model fitting procedure is re-run with initial parameter values set to be those obtained in the last run and a further set of K initialisation values as before. A total of five such re-runs takes place unless the convergence criterion is met.

The number of sampled females in the SW Greenlandic catch $\tilde{S}_t^{III,f}$ is generated from:

$$N\left(\frac{R_t^f}{R_t^f + \lambda_t^{III} R_t^m}\, C_t^{III},\quad \sigma_{III}^2\, \frac{R_t^f}{R_t^f + \lambda_t^{III} R_t^m}\, C_t^{III}\right), \tag{D7}$$

and number of unsampled females in the SW Greenlandic catch $\tilde{U}_t^{SW,f}$ is generated from:

$$N\left(\frac{R_t^f}{R_t^f + \lambda_t^{III} R_t^m}\, C_t^{U_{III}},\quad \sigma_{III}^2\, \frac{R_t^f}{R_t^f + \lambda_t^{III} R_t^m}\, C_t^{U_{III}}\right), \tag{D7}$$

where λ_t^{III} is assumed to change linearly over time during the period and is defined as given in equation M5

and $C_t^{U_{III}} = C_t^{SW} - C_t^{III}$.

The generated catches from each strata are summed to give the total generated catch by sex, e.g.

$$\tilde{C}_t^f = \tilde{C}_t^{NW+CW,f} + \tilde{C}_t^{SW,f} = \tilde{S}_t^{II,f} + \tilde{U}_t^{NW+CW,f} + \tilde{S}_t^{III,f} + \tilde{U}_t^{SW,f} \tag{D9}$$

Data generation for the Closed Model (3)

The data generation algorithm is essentially the same as for the Influx Model (above), but with the following changes:

- When generating data for the NW+CW strata: $R_t^g \rightarrow R_t^g\, p_t^g$ where p_t^g is defined in equation M7.
- When generating data for the SW strata: $R_t^g \rightarrow R_t^g\, (1 - p_t^g)$

Robustness trials

The robustness tests are based on two classes of data sets: (a) data sets generated using a population dynamics model under alternative sets of assumptions, and (b) data sets generated using ad hoc algorithms to alter the observed data (model-free data sets). The model-based robustness tests are used to examine how the estimation results change in response to:

(1) different 'true' K (and hence N_{2007}) values [when the population dynamics model is known];
(2) different extents of overdispersion [when the population dynamics model is known];
(3) model mis-specification [when the model used to generate the data sets differs from that on which the estimation method is based, such as an intentional confusion between models 3 and 4b, but not necessarily limited to these two models]; and
(4) data sets that are longer than the current data set. The future data will be generated by projecting the population dynamics model forward under the average catch over the last 10 or 20 years.

The model-free data sets explore the behaviour of the estimation method when the trend of the sex-ratios is changed in a systematic manner.

'Adequate performance' will be evaluated under the principle that changing the data should lead to changes in the model output in the expected direction and of the expected magnitude or that it should be possible to provide a qualitative explanation for any discrepancies between the model output and *a priori* expectations.

(a) Data sets generated using a population dynamics model under alternative sets of assumptions

The initial set of robustness trials is listed in Table 2. The values for K in Table 2 are selected in order to allow a range of stock status levels to be examined in the simulations; the values are preliminary and may need to be refined once initial results are available.

The following process will be used in conducting and evaluating the results of model-based robustness trials:

(1) Estimate the overdispersion and other parameters (denoted P_{op}) by fitting the population dynamics model to the actual data set given the specified 'true' model and value of K.

Table 2

An initial list of robustness tests based on population models (unless specified otherwise the estimator should match the population model used to generate the data).

Case	Population model	K^1	Overdispersion	MSYR	Other
1	Influx	150,000	Estimated	2%	
2	Closed	150,000	Estimated	2%	
3	Model 5	150,000	Estimated	2%	
4	Closed-b	150,000	Estimated	2%	Closed estimator
5	Closed-a	150,000	Estimated	2%	Closed estimator
6	Influx	75,000	Estimated	2%	
7	Influx	50,000	Estimated	2%	
8	Influx	20,000[1]	Estimated	2%	
9	Influx	150,000	Estimated	2%	Closed estimator
10	Closed	150,000	Estimated	2%	Influx estimator
11	Influx	150,000	2 × Estimated	2%	
12	Influx	150,000	1	2%	
13	Influx	150,000	Estimated	1%	
14	Closed	150,000	Estimated	1%	
15	Influx	150,000	Estimated	2%	+20 yr extra data
16	Closed	150,000	Estimated	2%	+20 yr extra data
17	Influx	150,000	Estimated	4%	
18	Closed	150,000	Estimated	4%	

(2) Generate 1,000 data sets taking account of uncertainty in the sex ratios of the catches, using the 'true' K and 'true' model (i.e. using the P_{op} parameters).
(3) For each data set:
    (a) Calculate the deviance[2] for the data set corresponding to the 'true' value of K when the estimation is based on the 'estimation' model (i.e. the difference in log likelihoods between the fit with fixed *K* (= K_{true}) and the fit estimating *K*). The parameters corresponding to the fit with fixed *K* are denoted P_{est} .
    (b) Generate a large number (e.g. 1,000) of data sets based on the true value of *K* and the values for the other parameters of the model (the P_{est} parameters obtained in step (a)) and compute the deviance for each of these data sets. The model used to generate the data sets and compute the deviance is the 'estimation' model.
    (c) Record the percentile that the deviance from step (a) represents in the distribution generated from step (b).
    (d) Record if the deviance of the upper 5%- and 10%-iles of the distribution obtained from step (b) is smaller than the deviance obtained from step (a).
(4) Plot the collection of 1,000 percentiles (one for each simulated data set) and assess the frequency of the percentages being larger than 2%, 5%, 10%, etc. (Fig. 2b).
(5) Calculate the percentage of times for the 1,000 generated data sets that the deviance of the 95%- and 90%-iles is smaller than the deviance of the generated data.

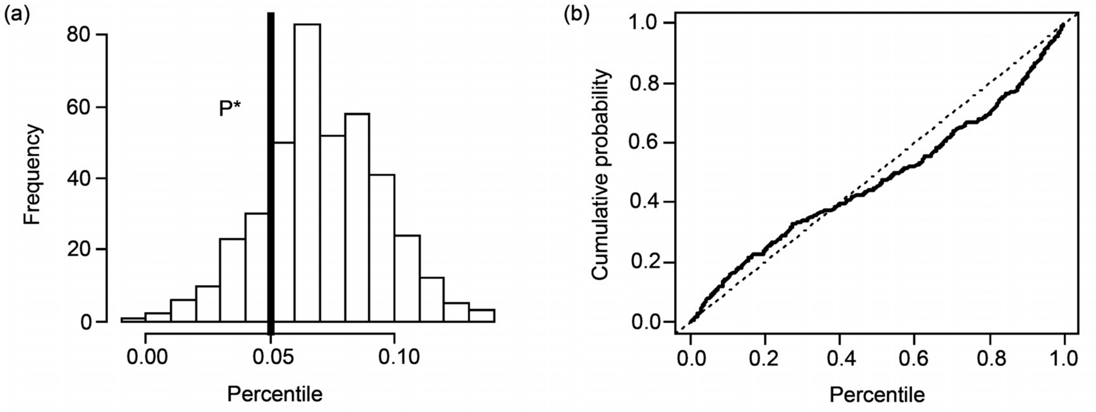

Fig. 2. (a) Histogram of percentiles from 400 simulated data sets along with the P* based on the actual data. (b) Cumulative distribution of percentiles.

(b) Model-free data sets

The initial set of robustness trials is listed in Table 3 to be combined with the values in Table 2. The values in Table 3 (particularly those for the slopes in mean sex-ratio) are preliminary and may need to be refined once initial results are available.

The process for evaluating model-free robustness tests is similar. Specifically:–

(1) Select a set of values for *K*. For each value of *K*:
    (a) Calculate the deviance for the original data set.
    (b) Generate a large number (e.g. 1,000) of data sets based on the true value of *K* and the values for the other parameters of the model (these would be obtained from step (a)) based on an overdispersed binomial distribution for both the sexed and unsexed component of the catch.
    (c) Record the percentile, *P**, that the deviance from step (a) represents in the distribution generated from step (b).
(2) Repeat steps a–c above for each data set and value of *K*.
(3) For each value of *K* plot the percentiles obtained at step 2 using a histogram and *P** using a vertical line.

Table 3

Specifications for the model-free robustness tests (separately for each data set and for all data sets at the same time).

Case	Slope	Mean
1	Unchanged	Unchanged
2	+0.05	Unchanged
3	–0.05	Unchanged
4	Unchanged	+0.05
5	Unchanged	–0.05

REFERENCES

International Whaling Commission. 2009. Report of the Scientific Committee. *J. Cetacean Res. Manage.* (*Suppl.*) 11: 21.

Brandão, A. and Butterworth, D.S. 2009. Yet further results for lower confidence limits for the abundance of West Greenland minke whales. Paper SC/61/AWMP7 presented to the IWC Scientific Committee, June 2009. Madeira, Portugal (unpublished). 28pp. [Paper available from the Office of this Journal]

---

[2] The deviance is set to zero if the estimated *K* is less than the true value.

Appendix 1

TOTAL CATCH SERIES BY SEX

The total catch series by sex is listed in Table 1. To generate these data the sex ratio of the sex specific reporting in any year from the Greenland or Norwegian fishery was set to apply to the total number of whales landed and struck and loss by that fishery in that year. For years with no or almost no sex information on the removals by Greenlandic whalers (1948–54; 1979–84), the sex specific removals were estimated from the sex ratio of the reported removals in that fishery over all years with sex specific reporting. The estimated sex specific removals were then added to provide a time series of total sex specific removal.

Table 1

Total catches:

Year	Female	Male	Year	Female	Male	Year	Female	Male
1948	3	1	1968	273	62	1988	95	14
1949	4	1	1969	252	184	1989	43	20
1950	7	2	1970	181	152	1990	72	17
1951	12	4	1971	340	127	1991	81	28
1952	24	8	1972	278	114	1992	89	21
1953	24	8	1973	383	114	1993	84	28
1954	16	6	1974	393	76	1994	81	23
1955	12	10	1975	301	23	1995	108	47
1956	16	6	1976	323	55	1996	130	40
1957	18	6	1977	260	100	1997	105	43
1958	16	14	1978	221	34	1998	126	43
1959	49	6	1979	230	95	1999	134	38
1960	49	7	1980	257	80	2000	107	38
1961	20	15	1981	198	67	2001	103	36
1962	52	20	1982	228	88	2002	103	36
1963	99	67	1983	242	94	2003	124	62
1964	95	67	1984	225	80	2004	133	46
1965	120	76	1985	186	87	2005	141	35
1966	151	74	1986	107	38	2006	134	47
1967	209	35	1987	65	21	2007	127	40

Catches by period:

Period 1 (Greenlandic 1955–78)			Period 2 (late N Greenland)			Period 2 (late S Greenland)		
Year	Female	Male	Year	Female	Male	Year	Female	Male
1955	8	7	1987	9	6	1987	3	1
1956	15	5	1988	27	4	1988	8	1
1957	18	6	1989	13	12	1989	21	4
1958	6	5	1990	32	13	1990	27	1
1959	17	2	1991	38	10	1991	25	9
1960	15	2	1992	44	9	1992	30	9
1961	9	7	1993	44	22	1993	26	3
1962	43	17	1994	50	14	1994	27	6
1963	47	32	1995	68	36	1995	37	10
1964	37	26	1996	76	31	1996	48	7
1965	30	19	1997	70	33	1997	31	9
1966	49	24	1998	81	33	1998	42	9
1967	42	7	1999	86	26	1999	45	11
1968	47	10	2000	57	17	2000	24	8
1969	42	14	2001	56	25	2001	30	5
1970	20	12	2002	60	21	2002	36	11
1971	25	6	2003	59	34	2003	57	22
1972	40	6	2004	88	26	2004	39	18
1973	39	8	2005	93	20	2005	42	14
1974	34	6	2006	106	34	2006	19	10
1975	17	1	2007	97	30	2007	22	8
1976	20	2						
1977	39	15						
1978	13	2						

Annex D

Model-Based Robustness Tests for the Sex Ratio Method

The model-based robustness tests are used to examine how the estimation results change in response to:

(1) different 'true' K (and hence N_{2007}) values [when the population dynamics model is known];
(2) different extents of overdispersion [when the population dynamics model is known];
(3) model mis-specification [when the model used to generate the data sets differs from that on which the estimation method is based]; and
(4) data sets that are longer than the current data set. The future data will be generated by projecting the population dynamics model forward under the average catch over the last 10 or 20 years.

The following is the process for conducting and evaluating the results:

(1) Set a new 'true' value for K and select a 'true' model. Estimate the overdispersion parameters and any other parameters by fitting the population dynamics model to the actual data set given the chosen model and value of K.
(2) Generate 1,000 data sets taking account of uncertainty in the sex ratios of the catches, using the 'true' K and 'true' model.
(3) For each data set:
    (a) Calculate the deviance[1] for the data set corresponding to the 'true' value of K when the estimation is based on the 'estimation' model.
    (b) Generate a large number (e.g. 1,000) of data sets based on the true value of K and the values for the other parameters of the model (these would be obtained from step (a) and compute the deviance for each of these data sets. The model used to generate the data sets and compute the deviance is the 'estimation' model.
    (c) Record the percentile that the deviance from step (a) represents in the distribution generated from step (b).
    (d) Record if the deviance of the upper 5%- and 10%-iles of the distribution obtained from step (b) is smaller than the deviance obtained from step (a).
(4) Plot the collection of 1,000 percentiles (one for each simulated data set) and assess the frequency of the percentages being larger than 2%, 5%, 10%, etc.
(5) Calculate the percentage of times for the 1,000 generated data sets that the deviance of the 95%- and 90%-iles is smaller than the deviance of the generated data.

The process for evaluating model-free robustness tests is similar.

(1) Select a set of values for K. For each value of K:
    (a) Calculate the deviance for the original data set.
    (b) Generate a large number (e.g. 1,000) of data sets based on the true value of K and the values for the other parameters of the model – these would be obtained from step (a) based on an overdispersed binomial distribution for both the sexed and unsexed component of the catch.
    (c) Record the percentile, P^*, that the deviance from step (a) represents in the distribution generated from step (b).
(2) Repeat steps a–c above for each data set and value of K.
(3) For each value of K plot the percentiles obtained at step 2 using a histogram and P^* using a vertical line.

---

[1] The deviance is set to zero if the estimated K is less than the value.