IWC/SC/58/AWMP3
May 3, 2006

A sex ratio based assessment of common minke whales off West Greenland

LARS WITTING, *email: lawi@natur.gl*
Greenland Institute of Natural Resources, Box 570, DK-3900 Nuuk, Greenland

ABSTRACT
The sex ratio in the West Greenland catch history of the common minke whale (*Balaenoptera acutorostrata*) is used to assess the current status of the common minke whale population that supplies the West Greenland hunt. The female fraction in common minke whale foetuses is around 1/2, but the fraction in the West Greenland catch has varied around 3/4 since the beginning of the hunt in 1948. This difference is likely to reflect sex specific behaviour, where females tend to occur in other areas than males, but it may also reflect a female selective hunt and/or a female bias in the sex ratio at birth. These hypotheses were examined by trial simulations, where an age- and sex-structured population model with density regulated dynamics were set to cover a maximum sustainable yield rates between 1% and 7%, a current abundance between 800 and 50,000 females, different degrees of female bias in the sex specific dispersal, a sex specific hunt, a female bias in the sex ratio at birth, increasing trends in the female bias of a sex specific dispersal and a sex specific hunt, and a uniform, increasing and decreasing age-selectivity in the hunt. Given the trials and the data is it concluded that a current abundance in the order of 20,000 individuals is a conservative estimate, and that a current catch of 175 individuals most likely is sustainable.

KEYWORDS: SEX RATIO, MODELLING, ATLANTIC OCEAN, WHALING - ABORIGINAL

INTRODUCTION

Common minke whales (*Balaenoptera acutorostrata*) in the North Atlantic have a segregational behaviour where females tend to occur further to the north than males (Jonsgård 1962; Larsen and Øien 1988; Øien 1988; Horwood 1989). In the eastern North Atlantic, females are found to dominate the catches in the Barents Sea, while males predominate the catches around the British Isles and on the Norwegian coast including Finnmark (Øien 1988). The same pattern is found in the Norwegian catches in the western North Atlantic, where males dominate the catches in the southern areas with the percentage of females increasing going northwards along East Greenland and West Greenland (Larsen and Øien 1988). Here it is also observed that females dominate the catches early in the season with their proportion decreasing thereafter, indicating that females arrive earlier in the season migrating farther north than males (Larsen and Øien 1988).

Sexual segregation by time and geographical area is usually not the main data input for assessments of large whales. Assessment models have instead been fitting population dynamic models to abundance data, with the historical catches being used to describe the dynamic development of the stock. But when abundance data are either limited or relating to some unknown sub-fraction of the total population, and the catch history is well documented it might be more informative to fit the population models to the sex ratio of the catch history than to the available abundance data. This is especially true if the sex ratio in the catches differs from the overall sex ratio of the population. The catch data should then carry a signal on the exploitation level of the population, in the sense that the catch sex ratio should change if the population is becoming over-utilised to the extend that the sex ratio of the population is becoming biased relative to the pristine sex ratio.

The common minke whale off West Greenland is a case where the sex ratio of the catches may provide a data signal on the level of exploitation. Catches of common minke whales off West Greenland has occurred regularly since the end of the 1940s, with the annual take having a maximum of four to five hundred individuals in the early 1970s, and a current take of approximately 175 individuals per year. Throughout the period the catch has been predominately of females. The average proportion of females in the catch from 1948 to 2004 is 0.74, and this should be compared with a foetal sex ratio that is not significantly different from even [40% females among 43 foetuses from the Norwegian hunt (Larsen and Kapel 1982), and 54% females among 339 foetuses from the Greenland hunt (Witting 2000)].

As the sex ratio in the foetuses of pregnant females caught off West Greenland is even (Larsen and Kapel 1982; Witting 2000), it is most likely the geographical sub-structuring of the two sexes during summer (Larsen and Øien 1988; Øien 1988; Horwood 1989; Andersen et al. 2003) that determines the female bias in the West Greenland catch of common minke whales. An alternative explanation is sex specific harvest selectivity combined with an even or uneven dispersal of males and females. This latter model may also explain the female biased catch, but is less likely to be true as female common minke whales cannot generally be distinguished from males at distance.

While considering both of these hypotheses in this paper I use the sex ratio information of the historical catches to perform a conservative assessment of the common minke whale population that supplies the West Greenland harvest. This is done by incorporating the behaviour of sex specific dispersal and sex specific catch selectivity into an age- and sex-structured population model with density-regulated dynamics, and by fitting this model to the sex ratio information of the catch history.

Owing to the continued female bias of the West Greenland catch there is some lack of a data signal on the upper bounds of the population. This is because larger stocks can more easily maintain a continued female biased catch. Other things being equal, this implies a likelihood that is monotonically increasing with the overall abundance having a diminishing return where the likelihood approaches an asymptotic value as the overall abundance increases towards infinity. In order to obtain upper bounds on the population in an earlier assessment Witting (2005) applied an abundance-based penalty to the likelihood function, with the tuning of the penalty function being determined by the diminishing return in the likelihood with increased abundance.

While the method in Witting (2005) is likely unproblematic from a scientific point of view, it provides no information on a preferred tuning of the penalty function. And as it may be difficult, or even impossible, to obtain consensus on an assessment tuning level in a management organisation like the IWC, I have chosen a different approach in the current paper. Here, I simulate trials that capture a realistic range of the dynamics of the common minke whale population that supplies the West Greenland harvest. For each trial I generate hundred simulated catch histories and performs Bayesian assessments on all the catch histories in the aim of identifying an assessment approach that is conservative relative to the true population status of the trials.

Then, when a final assessment is run on the original data it should provide an assessment that is conservative in the sense that the real population should be doing better than indicated by the assessment results.

METHOD

Catch statistics

The IWC database on catches of common minke whales off West Greenland starts in 1948, with reported catches of males and females falling into three major sets: Catches taken by Greenlandic whalers from 1955 to 1978, and from 1985 to the present, and catches taken by the Norwegian whalers from 1968 to 1985. Greenlandic whalers also took common minke whales from 1948 to 1954 and again from 1979 to 1984, but sex specific reporting is almost absent in these years.

The three time series of sex specified catches are listed in Table 1, with the top figure in Fig. 1 showing the proportion of the reported catches that were reported with sex. Nearly all the Norwegian catches were reported with sex, while the proportion reported with sex was generally below 50% for the Greenlandic catches from 1955 to 1978, with the proportion declining to approximately 10% toward the end of the period. The absolute number of sex reports remained relatively stable over the period, with the decline in the proportion reflecting mainly an increase in the absolute number of catches. From 1985 and onwards sex specific reporting was generally high in the Greenlandic catches, with the fraction of sex specific reporting being above 90% in 10 out of 12 years since 1993.

Although the sex ratio of the sex specific reporting has fluctuated over the years there is no apparent trend in the sex ratio (Fig. 1, middle). The average of the yearly sex ratio ($r = \hat{C}^m/\hat{C}^f$) of reported caught males ($\hat{C}^m$) over reported caught females ($\hat{C}^f$) varies only little between the three data sets (geometric mean of 0.30 for Greenlandic whalers from 1955 to 1978, 0.34 for Greenlandic whalers from 1968 to 2004, and 0.44 for Norwegian whalers from 1968 to 1985), while the three sets differ more substantial in the variation (cv for $\ln r$ of 0.62 for Greenlandic whalers from 1955 to 1978, 0.23 for Greenlandic whalers from 1968 to 2004, and 0.96 for Norwegian whalers from 1968 to 1985).

When, for each of the three sets, it is assumed that there is no trend and $\ln r_t$ in year t is normalised [$\ln \hat{r}_t = (\ln r_r - \overline{\ln r})/\sigma$] by subtracted by the mean ($\overline{\ln r}$) and dividing by the standard deviation (σ) of the set, it follows that the joint distribution of $\ln \hat{r}_t$ over all years with sex specific reporting is very close to normal (Fig. 1, bottom).

Reconstructing sex specific catches

A single sex specific time series of the total removal of male and female common minke whales off West Greenland was constructed. The sex ratio of the sex specific reporting in any year t from a specific fishery (Greenlandic whalers or Norwegian whalers) was assumed to apply to the total number of whales landed and struck and loss by that fishery in that year. And 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 of the two fisheries were then added to provide a time series of total sex specific removal (Table 2). The r-ratio for this series has a geometric mean of 2.9 (95% CI:2.5-3.4).

Sampling and process variation

The *cv* of the sampling variance on the sex ratio estimate r in year t was set to the sampling *cv* for the number of reported males given the binominal reporting of males and females. The sampling *cv* of the male/female ratio in year t is then

$$cv_{r,t} = \sqrt{1/\hat{C}_t^m - 1/\hat{C}_t} \tag{1}$$

where $\hat{C} = \hat{C}^m + \hat{C}^f$ is total sex specific reporting.

The total variation in the estimates of r is larger than the sampling variation indicating that there is process variation in the sex ratio of the catches. Assuming, as indicated by the data, that there has been no directional change in the underlying mean of the catch sex ratio over the sample period, the reported r_t ratio in any year (t) may be seen as a reflection of random binominal sampling of males of females from a log normal distributed $\tilde{r}_t$ that reflects, e.g., random variation in the sex ration among the common minke whales on the whaling grounds, or random variation in a sex specific selection preference among the hunters with the sex ratio on the whaling grounds being constant.

In a sex specified catch sample of $\hat{C}_t = \hat{C}_t^m + \hat{C}_t^f$ individuals in year t, the number of reported females may thus reflect

$$\hat{C}_t^f = Bin(\hat{C}_t, \tilde{\vartheta}_t) \tag{2}$$

where $\tilde{\vartheta}_t$ is the true fraction of females, given as $\tilde{\vartheta}_t = 1/(1 + \tilde{r}_t)$, with the true male / female ratio $\tilde{r}_t$ being a realisation of a fishery specific random process

$$\begin{aligned} \tilde{r}_t &= \bar{r}_t e^{\epsilon_t} \\ \epsilon_t &\sim N(0; \sigma^2) \end{aligned} \tag{3}$$

where $\bar{r}_t$ is the true mean of that year and σ^2 the process variance. The true year specific $\bar{r}_t$-ratio was given by Eq. (7) during trial iterations.

Population dynamic model

A similar age- and sex-structured model with density-regulated dynamics was applied to the trial iterations and to the Bayesian assessments. The number of animals in age classes larger than zero was

$$\begin{array}{ll} N_{t+1,a+1}^{m/f} = (N_{t,a}^{m/f} - C_{t,a}^{m/f}) s_a & 0 \le a \le x-2 \\ N_{t+1,x}^{m/f} = (N_{t,x}^{m/f} - C_{t,x}^{m/f}) s_x + (N_{t,x-1}^{m/f} - C_{t,x-1}^{m/f}) s_{x-1} & \end{array} \tag{4}$$

where s_a is age specific annual survival, $N_{t,a}^{m/f}$ is the number of males/females of age a at the start of year t, $C_{t,a}^{m/f}$ is the catch of males/females of age a during year t, and $x = 15$ is a lumped age-class. The age distribution of the catches was sex specific and proportional to the age-structured abundance, except that no catches were taken from age-class zero.

The harvest was taken from a sub-component H of the population, with the sex ratio in the harvested component reflecting the sex ratio of the total population and a sex specific aggregation between the total and the harvested component. The relative number of males and females at age a in the harvested component was given as

$$\begin{aligned} H_a^m &= N_a^m \\ H_a^f &= \vartheta_h N_a^f \end{aligned} \tag{5}$$

where $0 \leq \vartheta_h \leq \infty$ determines the female fraction in the harvested component relative to the female fraction in the total component. Eq. (5) implicitly assumes a source-sink type of dynamic dispersal pattern where the harvest area in West Greenland can act as a sink in the sense that a relative depletion in the harvest area would induce some inflow of common minke whales from other areas.

The catch was allowed to be both age and sex selective so that the catch of males and females of age a in year t can be given as

$$\begin{aligned} C^m_{t,a} &= c_a c_t H^m_a \\ C^f_{t,a} &= \vartheta_c c_a c_t H^f_a \end{aligned} \tag{6}$$

where c_t reflect the year-specific harvest effort, c_a the age-specific selectivity, and $0 \leq \vartheta_c \leq \infty$ the sex specific selectivity with $\vartheta_c = 1$ implying no selectivity on the sexes. The sex ratio r in the catches is then

$$r = \frac{\sum_{a=0}^{x} c_a N^m_a}{\vartheta_c \vartheta_h \sum_{a=0}^{x} c_a N^f_a} \tag{7}$$

Let the annual survival rate s_a of animals of age a be

$$s_a = \begin{cases} s_{juv} s_{ad} & \text{if } a = 0 \\ s_{juv} & \text{if } 1 \leq a \leq a_{ad} \\ s_{ad} & \text{if } a > a_{ad} \end{cases} \tag{8}$$

where s_{juv} is the survival rate for 'juveniles', s_{ad} is the survival rate for adults, and $a_{ad} = 1$ is the greatest age at which the 'juvenile' survival rate applies.

The number of births at the start of year t, B_t, is

$$B_t = \sum_{a=a_m}^{x} B_{t,a} \tag{9}$$

where a_m is the age of reproductive maturity, and $B_{t,a}$, the number of births in age class a, is

$$B_{t,a} = b_t M^f_{t,a} \tag{10}$$

where b_t is the fecundity rate for mature females at time t, and $M^f_{t,a}$ is the number of mature females in age class a at the start of year t, defined as

$$M^f_{t,a} = \begin{cases} 0 & \text{if } a_m > a \\ N^f_{t,a} & \text{if } a_m \leq a \end{cases} \tag{11}$$

Let the component of the population that imposes density-regulation be the one plus component

$$\hat{N} = \sum_{a=1}^{x} N^f_a + N^m_a \tag{12}$$

and let the density-regulation on the fecundity rate b_t take the Pella-Tomlinson form

$$b_t = b^* + [b_{max} - b^*][1 - (\hat{N}_t / \hat{N}^*)^z] \tag{13}$$

where b^* is the birth rate at carrying capacity N^*, b_{max} is the maximal birth rate, and z the level of density dependence.

Although not explicit parameters of the model, the maximum sustainable yield level (msyl) and the maximum sustainable yield rate (msyr) were treated as parameters in the analysis, with both parameters relating to the harvested component, i.e., to the one plus component of the population. The msyl depends mainly to the compensation parameter z, with the relationship between z and the msyl being solved numerically.

A measure of a sustainable harvest in year t was defined as

$$sh_t = \begin{cases} ry_t & \text{if } d_t < \text{msyl} \wedge ry_t < 175 \\ 0.9\text{msy} & \text{if } d_t \geq \text{msyl} \wedge 0.9\text{msy} < 175 \\ 175 & \text{if } d_t < \text{msyl} \wedge ry_t > 175 \text{ or } d_t \geq \text{msyl} \wedge 0.9\text{msy} > 175 \end{cases} \tag{14}$$

where ry_t is the replacement yield in year t, and 175 is the current quota.

Trials

Trial catch histories

Population dynamic trials were simulated, with the historical catches of the trials having absolute catches that were similar to those originally reported in the West Greenland harvest, and simulated sex ratios that were based on the population dynamics of the trials. The sex ratio model of Eq. (7) and the sex ratio variation model of Eqs. (2) and (3) were used to simulate both the historical times series of sex specific catches and the historical times series of sex specific reporting of catches.

The sex ratio of the historical catch of a trial in year t was calculated from the true sex ratio ($\bar{r}_t$) of the trial in that year, the process variation of Eq. (3), and the sampling variation of Eq. (2) with the absolute number of catches in that year representing the sample. This catch was the catch that was subtracted from the population during a trial simulation, and the sex ratio of this catch differed from the reported catch sex ratio where the sampling variation of Eq. (2) was based on the number of sex specific reporting in the original catch history. The sex sample for the reported catch history was set to be a sub-sample of the total harvest, thus implicitly assuming that all reporting of sex was correct.

To obtain estimates of time transitions in the process variation, the original catch series were divided into three periods: 1955 to 1967 where only Greenlandic whalers were catching common minke whales, 1968 to 1985 where both Greenlandic whalers and Norwegian whalers were catching whales, and 1986 to the present where again only Greenlandic whalers were catching whales. The period from 1948 to 1954, where there is no sex specific reporting, was simulated as the period from 1955 to 1967.

The process variance σ^2 was estimated by numerical iteration for each of the three time periods, with the geometric mean of the r ratios for each time period being used as an estimate of the true $\bar{r}$ ratio. The σ^2 parameter of the simulation model of Eqs. (2) and (3) was then adjusted so that the simulated variation in the reported sex ratio was similar to the actual variation in the reported sex ratio. For each period, n_d-random data sets of r_t over all the years in the period were generated from the geometric mean of the r ratio in the true data set, the number of sexed individuals in each year and a value for σ^2, with the number of generated data sets n_d being set so that the error cv on the average standard deviation in r across the n_d-data sets were below 0.1 percent. The σ^2 parameter was then adjusted so that the absolute difference between the standard deviation in the r ratio of the original data set and the average standard deviation in the r ratio across the simulated data sets was less than 0.001. This gave σ estimates of 0.55 for 1955 to 1967, 0.57 for 1968 to 1985, and 0.097 for 1986 to 2005. The process variation was also estimated for the Greenlandic whaling from 1955 to 1978 and 1985 to 2005, and the Norwegian whaling from 1968 to 1985. This gave σ estimates of 0.59, 0.10 and 0.75, which were used for the error bars in the middle figure in Fig. 1, assuming a total cv of $\sqrt{cv_b^2 + cv_p^2}$ with cv_b reflecting binominal sampling variation and cv_p reflecting process variation.

Trial parameters

Being based on the age and sex structured population model with density regulated dynamics, the trials were generally defined by the total abundance of females in year 2005 (N), the msyr, the msyl, the female/male ratio of the harvested component of the population relative to the female/male ratio in the total component (ϑ_h), the sex selectivity factor of the harvest (ϑ_c), and the female fraction at birth (ϑ).

Three trial sets were defined (Table 3). The first (**wga**) was based on variation in the abundance and msyr, assuming a standard set for the other trial parameters. The second set (**wgb**) was incorporating variation in the sex ratio parameters ϑ_h, ϑ_c and ϑ. And the third set was incorporating time-trends in the sex ratio parameters ϑ_h and ϑ_c, as well as variation in the age-structured selectivity of the harvest.

The msyr was set to 0.01, 0.04 and 0.07 across the **wga**-trials, and the abundance in 2005 was set to have a *log* uniform distribution across the trials ranging from 800 to $50,000$ females. The 800 females resemble a lower 95% confident limit of a ship born survey off West Greenland in 2005 (Heide-Jørgensen et al. 2006), thus representing a pessimistic view of the abundance of the common minke whale population that supplies the West Greenland harvest. The $50,000$ females, on the other hand, was set to represent the optimistic view that West Greenland waters are visited by common minke whales from large areas of the Central North Atlantic and the Western North Atlantic. Together, the point estimates of the 1997 estimate from the CM area (Skaug et al. 2002), the 2001 estimate from the CIC area (Borchers et al. 2003), and the 2001 estimate from the CG and CIP areas (Gunnlaugsson et al. 2003) of the Central North Atlantic sums to approximately 90.000 common minke whales.

The female fraction at birth was set to 0.5 for all **wga**-trials, as indicated by data. The fraction of females in the foetuses of pregnant common minke whale females caught off West Greenland has been estimated to 0.41 (Larsen and Kapel 1982; Larsen 1984) and 0.54 (Witting 2000), and a rather similar fraction of 0.48 have been found for East Canadian common minke whales (Mitchell 1974). None of these values differs significantly from the even sex ratio 0.50.

The female/male ratio of the harvested component of the population relative to that of the total population (ϑ_h) was set to represent the sole cause for the female biased harvest in the **wga** trials, i.e., ϑ_h was set to 2.9 and ϑ_c to one. The msyl was set to 0.60 for all **wga**-trials.

The msyr and the msyl were set to 0.04 and 0.60 for all **wgb**- and **wgc**-trials, with the female abundance in 2005 being either low $(1,000)$, medium $(5,000)$ or high $(25,000)$. For the **wgb**-trials, the cause of the sex ratio biased harvest was set to resemble sex specific harvest selectivity for trial one to three, while the female bias of the harvested population component was set to be large ($\vartheta_h = 3.4$) for trial four to six, and small ($\vartheta_h = 2.4$) for trial seven to nine, covering the 95% confidence interval around the point estimate of 2.9. For trial ten to 12 the cause for a sex-biased harvest was distributed evenly between a sex bias of the harvested component and a sex biased harvest selectivity ($\vartheta_h = \vartheta_c = 1.7$), and for trial 13 to 15 the cause was split between a sex bias harvest component and 60% females at birth.

For the **wgc**-trials were the sex ratio parameters ϑ_h, ϑ_c and ϑ set to the values in the **wga**-trials, except that time trends were added to the ϑ_h and ϑ_c parameters in trial one to nine; with the trends reflecting stable values ($\vartheta_h = 2.9$ & $\vartheta_c = 1$) from 1948 to 1970, where after one or both of the parameters would increase linearly to 2005 (see Table 3 for detail). Although not all that likely, an increase in the female bias of these factors with time can be problematic as it can potentially compensate an expected data signal of overexploitation; leaving the male fraction in the harvest stable while it would have increased in the absence of a trend in the female bias. The trends were initialised in 1970 because the 1970s were the period with the largest catches and, thus, the period where we might expect the first signs of an overexploitation to show up. The opposite trend, with an increase in the male bias over time, is less problematic as it would induce a false data signal of overexploitation and, thus, it is not included in the trials. Such a trend would also be extremely unlikely as there is no sign in the catch history of an increase in the male fraction with time.

In the previous trials have all individuals in age-classes larger than zero had the same probability of being caught, while no individuals were taken from age-class zero. In order to test the assumption, was the relative age-structured selectivity in trial ten to 12 in trial set **wgc** set to increase linearly from 0.07 for individuals in age-class one to one for individuals in age-class $x = 15$. In trial 13 to 15 was the relative selectivity set to decline from one for individuals in age-class one to 0.07 for individuals in age-class x. In all trials were no individuals taken from age-class zero.

Life history parameters

The msyr is the only growth related parameter in the trials because it is the most crucial single parameter that determines the growth potential in the population dynamic model. Each msyr have infinitely many possible life history combinations associated with it, and to capture this variation was each trial described by hundred iterations with each iteration having it's own randomly chosen combination of life history parameters.

Prior distributions were set on the life history parameters in order to solve the life history parameterisations for the trials (Table 4). Larsen (1991) summarised estimates of biological parameters in North Atlantic common minke whales. An annual natural survival rate of 0.90 was estimated by Horwood (1989) for the central North Atlantic, and a rate of 0.91 for the eastern North Atlantic was given by Ugland (1977). The prior on adult survival (s_{ad}) was set to cover the range from 0.92 to 0.99, and juvenile survival (s_{juv}) was set to cover the range from 0.10 to 0.99, with the additional constraint that juvenile survival is smaller than adult survival.

Various studies have found annual pregnancy rates between 0.86 and 0.99 for North Atlantic common minke whales (Sergeant 1963; Mitchell and Kozicki 1975; Christensen 1981; Larsen and Kapel 1983; Sigurjonsson 1988), and the prior on the maximal birth rate (b_{max}) was set to cover the range from 0.50 to 1.00 indicating birth every year or every other year. The age of sexual maturity was estimated to lie between five and eight years (Mitchell and Kozicki 1975; Christensen 1981; Sigurjonsson 1988), and the prior on the age of reproductive maturity (a_m) was set to cover the range between three and ten year.

Later studies (Øien, pers. comm.) have found that age determination in North Atlantic common minke whales is problematic, questioning some of the life history estimates given above. The ranges covered by the priors, however, are likely to be broad enough to capture realistic parameter values. Here it is also important to keep in mind that it is mainly the prior on the msyr that is driving the prior on the relative productivity of the population.

To parameterise a trial iteration with a given randomly drawn msyr, were the msyl and all life history parameters except juvenile survival drawn randomly from their priors. It was then by numerical iterations attempted to solve for a juvenile survival rate, within the upper and the lower bounds, that would generate the selected msyr. If no solution was possible, was the selected msyr maintained while a new set of life history parameters and msyl were drawn from their priors searching for a parameter combination that would allow for a solution where the juvenile survival rate would match the selected msyr. If no solution was found after 1000 randomly selected parameter sets, would a new msyr be drawn from the prior.

Trial averaging

Each trial is a more or less likely model hypothesis on the dynamics of the common minke whale population that supplies the West Greenland harvest, and in order to guide the evaluation of the relative likelihood of the different trials I applied the model selection framework of the Akaike information criterion (Akaike 1973; Johnson and Omland 2004).

Following the log-normal distribution of Fig. 1, the likelihood of a trial iteration was calculated under the assumption that the estimation errors of the male/female ratio were log-normally distributed

$$L = \prod_t \exp\left(-\frac{[\ln(r_t^i/r_t)]^2}{2cv_t^2}\right)/cv_t \tag{15}$$

where $1955 \leq t \leq 2005$, r_t^i is the male/female ratio of the catch of the trial iteration in year t, r_t is the reported male/female ratio of the West Greenland catches, and cv_t the coefficient of variation of the reported male/female ratio incorporating both the sampling and process variation described earlier.

The average likelihood $\bar{L}$ of a trial across the hundred trial iterations was then used to calculate the Akaike weight of a trial. As all trial models had the same number of parameters and were fit to the same catch history, the Akaike weight of the ith trial in trial set **wgi** was

$$w_i = \frac{\bar{L}_i}{\sum_{i \in wgi} \bar{L}_i} \tag{16}$$

These weights can be interpreted as the probability that trial i is the best model for the observed data, given the data and the set of trials. Thus, when the Akaike weights are used to produce a weighted average of a management related statistics across the trials, we obtain an average estimate that is weighted according to the relative likelihood of the different trials.

Bayesian assessments

Sets of Bayesian assessments were conducted for all trial iterations and the original data, with all assessments in a set being based on the same priors and population dynamic model. In each assessment was the population projected by subtracting the reported catches from the abundance, starting from a dynamic equilibrium in 1948.

Three density regulated population dynamic models were tried out in initial assessments, one being the age and sex structured model, and the two others being discrete sex structured models, one with yearly catches being taken before reproduction and the other with yearly catches being taken after reproduction. The age and sex structured model was selected for all subsequent assessments as it was found to be better to predict the true trajectory of the trials, especially for the later years.

The priors on the msyr and the female bias of the harvested population components (ϑ_h) were then used as tuning parameters, in order to tune the assessment set towards a conservative assessment where the estimates of abundance and sustainable harvest in 2005 were smaller than the true values of the trials.

Prior distributions

The prior distributions for the life history parameters (s_{ad}, s_{juv}, a_m and b_{max}) in the Bayesian assessments were set to the priors of the trials. The prior on the carrying capacity was uniform between $2,000$ and $200,000$ individuals, and the prior on the $msyl$ was uniform from 0.5 to 0.7. The female fraction at birth was fixed at 0.5, and there were no female biased selection in the harvest, i.e., $\vartheta_c = 1$.

For the final tuned assessment set was the msyr set to be uniform from 0.01 to 0.02, and the female bias of the harvested population component (ϑ_h) was set to 3.3. The female bias of the harvested component is an efficient tuning parameter to scale the estimates of absolute abundance for the high-abundance trials, while it has a much smaller effect on the low-abundance trials. The chosen ϑ_h value of 3.3 resulted in median abundance estimates that were generally smaller than the true abundance of the trials while, given $0.01 \leq \text{msyr} \leq 0.02$, most of the median abundance estimates were higher than the true abundance for a ϑ_h value of 2.9, which is the expected value from the original catch history.

Bayesian integration

The Bayesian integration for each assessment was obtained by the sampling-importance-resampling routine (Berger 1985; Rubin 1988), where n_1 random parameterisations θ_i ($1 \leq i \leq n_1$) are sampled from an importance function $h(\theta)$. This function is a probability distribution function from which a large number, n_1, of independent and identically distributed draws of θ can be taken. For each drawn parameter set θ_i the population was projected from 1948 until 2005. For each draw an importance weight, or ratio, was then calculated

$$w(\theta_i) = \frac{L(\theta_i)p(\theta_i)}{h(\theta_i)} \tag{17}$$

where $L(\theta_i)$ is the likelihood given the data, and $h(\theta_i)$ and $p(\theta_i)$ are the importance and prior functions evaluated at θ_i. In the present study the importance function is set to the joint prior, so that the importance weight is given simply by the likelihood. The n_1 parameter sets were then re-sampled n_2 times with replacement, with the sampling probability of the ith parameter set being

$$q_i = \frac{w(\theta_i)}{\sum_{j=1}^{n_1} w(\theta_j)} \tag{18}$$

This generates a random sample of the posterior distribution of size n_2. The resample of the posterior distribution was set to $n_2 = 1000$, and the sample from the joint prior distribution was set to $n_1 = 500,000$.

The likelihood was calculated by Eq. (15), with the r_t^i ratio being the reported sex ratio, the r_t ratio begin the theoretically expected value from Eq. (7), and the coefficient of variation cv_t incorporating both the sampling and the process variation described earlier.

If the importance function is adequately specified, the mean of the importance sample for each parameter should approach the mean from the true posterior distribution, given a sufficiently large sample. To examine if the sampled posterior quantities can be assumed to be representative of the true posterior distribution, convergence diagnostics were calculated. One such diagnostic is the maximum importance weight of a parameter set relative to the total summed importance weight over all n_1 draws. McAllister et al. (2001) suggest that the maximum importance weight needs to have dropped below 1% of the total sum; in the present study it dropped below 0.1% in all assessments.

RESULTS

Akaike weights

Large differences were found in the Akaike weights of the different trials (Table 5), indicating that some trials may be much more likely to reflect the dynamics of the common minke whale population than other trials are. For the **wga** trials, which were based on variation in the abundance and msyr, was it found that trials with a female abundance in 2005 at and below 6,300 have basically no Akaike weight. For the **wgb** trials, which were based on variation in the sex ratio related parameters, was it especially the high-abundance trials three, 12 and 15 that have a large weight, with the remaining trials having almost no Akaike weight. For the **wgc** trials it is especially trial nine, with a high abundance and an increasing trend in both of the ϑ_h and ϑ_c parameters, that has a high weight, although trial 15, with a high abundance and a decline in catch selectivity with age, has approximately half the weight of trial nine. Apart from these two trials do three other **wgc**-trials have a small weight, while the remaining trials have almost no Akaike weight.

Trial results

A realisation of the abundance projection for each trial, together with the median and the 90% credibility interval of the corresponding assessment are shown in Fig. 2 for all **wga** trials, in Fig. 3 for all **wgb** trials, and in Fig 4 for all **wgc** trials. These projections and associated estimates are typical for the trials; illustrating the tendency towards conservative abundance estimates for the high-abundance trials (caused mainly by the ϑ_h parameter value of 3.3) and the general absence of too positive production estimates even for the low production trials (caused mainly by the 0.01 to 0.02 constraint on the msyr prior).

The generally conservative abundance estimates of the assessments are also shown in Table 6 that, for the hundred iterations of each trial, gives the average, minimum and maximum values of the median abundance estimate for 2005 relative to the true trial-abundance in 2005 (median estimate divided by true abundance). The un-weighted average and the Akaike weighted average of the 2005 abundance across the **wga** trials were respectively 70% and 32% of the true abundance, with corresponding maximum values being 140% and 98%. The un-weighted and the Akaike weighted average for the 2005 abundance were less conservative for the **wgb** and **wgc** trials; being 97% and 46% of the true abundance for the **wgb** trials (with the corresponding maximum estimates being 200% and 120%), and 93% and 51% of the true abundance for the **wgc** trials (with the corresponding maximum estimates being 230% and 190%).

More importantly in terms of exploitation is the *sh* estimate of the sustainable harvest in 2005, also given in Table 6. Here the un-weighted and the Akaike weighted average of the median estimate across the **wga** trials were a 2005 sustainable harvest of 69% and 41% of the true value, with corresponding maximum values being 130% and 72%. For the **wgb** and **wgc** trials, the un-weighted and the Akaike weighted average were more conservative; being 33% and 18% of the true value for the **wgb** trials (with the corresponding maximum estimates being 63% and 36%), and 34% and 20% of the true value for the **wgc** trials (with the corresponding maximum estimates being 65% and 36%). Thus, the relative lack of conservatism in the abundance estimates for the **wgb** and **wgc** trials are not affecting the estimate of sustainability which is actually more conservative for the **wgb** and **wgc** trials.

The risk of overestimating the 2005 abundance and the 2005 sustainable harvest level, however, is even better integrated and described by the fraction of the hundred trial iterations that have abundance and sustainable harvest estimates below the true values of the trials, letting this fraction be described separately for each trial and as an un-weighted and an Akaike weighted average across all trials. This information is given in Table 7 for the **wga** trials, in Table 8 for the **wgb** trials, and in Table 9 for the **wgc** trials, with all tables showing results for seven percentiles between the 2.5 and the 50*th* percentiles.

Looking first at the **wgb** and **wgc** trials is it found that nearly all sustainable harvest estimates are below the true levels of sustainable harvest, with the un-weighted and the Akaike weighted probability of underestimating the sustainable harvest being 98% and 100% for the 50*th* percentile of the **wgb** trials, and 97% and 100% for the 50*th* percentile of the **wgc** trials. More variation is found in the estimates of the 2005 abundance, with the un-weighted and the Akaike weighted probability of underestimating the abundance declining from 93% and 100% for the 2.5*th* percentile to 69% and 97% for the 50*th* percentile for the **wgb** trials, and from 98% and 100% for the 2.5*th* percentile to 70% and 92% for the 50*th* percentile for the **wgc** trials.

The **wga** trials are slightly more conservative for the abundance estimates, and less so for the sustainability estimates. Table 7 shows, that the un-weighted and the Akaike weighted probability of underestimating the 2005 abundance decline from 97% and 100% for the 2.5*th* percentile, to 81% and 100% for the 50*th* percentile. And that the un-weighted and the Akaike weighted probability of underestimating the sustainable harvest in 2005 decline from 92% and 100% for the 2.5*th* percentile, to 80% and 98% for the 50*th* percentile. Note here the variation between trials, where the low production (msyr = 1%) trials with a 2005 female abundance equal to or below $6,300$ (trial 9, 12 and 15) have probabilities of 91%, 75% and 13% of underestimating the sustainable harvest for the 2.5-percentile. These trials, however, have basically no Akaike weight (Table 5). If instead the probability of underestimating the sustainable harvest is checked across all **wga**, **wgb** and **wgc** trials with a non zero Akaike weight, for the 20*th* percentile, we find that it is 100% except for trial six in **wga** where it is 99%, and for trial two in **wgc** where it is 94%.

Assessment results

The parameter estimates for the assessment based on the original harvest data are given in Table 10. Given the data and the trials, for the 10*th* percentile is it estimated that the un-weighted probability that the 2005 abundance exceeds 15,600 is 93%, and that the corresponding Akaike weighted probability in 100%. Furthermore, the un-weighted probability that the sustainable harvest in 2005 exceeds 162 is 96%, while the corresponding Akaike weighted probability is a 100%. For the 20*th* percentile is it estimated that the un-weighted probability that the 2005 abundance exceeds 19,600 is 88%, and that the corresponding Akaike weighted probability is

100%. Furthermore, the un-weighted probability that the sustainable harvest in 2005 exceeds 175 is 95%, while the corresponding Akaike weighted probability is a 100%.

DISCUSSION

The result that there is an un-weighted probability of 88%, and a Akaike weighted probability of 100%, that the current abundance exceeds 19,600 common minke whales, can be compared with a negatively biased abundance estimate of 4090 (95% CI: 1650–10200) common minke whales from the harvest area in 2005 (Heide-Jørgensen et al. 2006). Depending on the extend of the negative bias, this suggests that the harvest in West Greenland is supplied by a population that is approximately five, or more than five, times larger than the number of common minke whales that occur just off West Greenland. Although this difference might at first seem large, an abundance of 20,000 common minke whales is not large compared with the likely abundance of approximately 90.000 common minke whales in the Central North Atlantic (Skaug et al. 2002; Borchers et al. 2003; Gunnlaugsson et al. 2003) and the largely unknown abundance of common minke whales in the Western North Atlantic.

But could the assessment results reflect, not a true data signal, but only an artefact that is being introduced by the specifications of the trials and/or the specifications of the priors in the Bayesian assessments. The applied method would clearly fail if it is possible to get it to produce "conservative" abundance estimates that are unrealistically high compared with the known abundance and distribution of common minke whales in the North Atlantic.

The upper limit on the prior on the carrying capacity in the Bayesian assessments was increased from $200,000$ to two millions, in order to examine if the conservative abundance estimate of the assessment was an artefacts of the prior on the carrying capacity. This had only a minor impact on the assessment results. Where the $5th$ and the $50th$ percentiles of the 2005 abundance for the original assessment on the real data were $12,000$ and $34,000$, the corresponding estimates for the assessment with an upper limit of two million on the carrying capacity were $14,000$ and $36,000$. Clearly, the conservative abundance estimates of the assessment are not controlled by the abundance prior.

But if the assessment tuning parameter ϑ_h is changed from 3.3 to the expected point estimate of 2.9, would it then be possible to generate a so called "conservative" abundance estimate that is unrealistically high. An upper limit on the carrying capacity prior in the Bayesian assessments of two millions and a fixed ϑ_h at 2.9, gave unrealistically high 2005 abundance estimates of 130.000 and 910.000 for the $5th$ and the $50th$ percentiles. But these estimates could not be classified as being conservative. Across nine trials that covered the most optimistic range of plausible abundance estimates (2005 female abundance of $100,000$, $50,000$ and $10,000$, each with a msyr of 0.01, 0.04 and 0.07), did the latter assessment produce median 2005 abundance estimates that were on average 4.3 times the true abundance, with a maximal overestimate that was 46 time the true abundance. It may therefore be concluded that the applied method has an inherent trade-off where abundance estimates that cannot be guaranteed conservative can be unrealistically high, while conservative abundance estimates tend to fall within a plausible range.

In the assessment has it implicitly been assumed that the sex has been determined correctly and reported correctly. Most of the sex information has been provided by the hunters, and the reported female fraction has been estimated to be the same (67%) for animals caught by Greenlanders and animals caught by the Norwegian whaling vessels (Kapel 1980). An independent estimate for 1982 and 1996-1998 of 77% females is available from the genetic analysis of Andersen et al. (2003). For 166 samples from common minke whales taken off West Greenland they found that only three males were reported as females, and two females were reported as males. In result there appears to be no reason to question the general reliability of the sex information in the harvest data. However, should the catch sex ratio be included as an essential part of future assessments and a future AWMP for West Greenland common minke whales, might it be advisable to shift to a system where the sex ratio is genetically estimated from tissue samples of caught whales.

REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. pp. 267–281, *In*: B. N. Petrov and F. Csaki (eds.) *Second International Symposium on Information Theory.* Akademiai Kiado.

Andersen, L. W., E. W. Born, R. Dietz, T. Haug, N. Øien and C. Bendixen 2003. Genetic population structure of minke whales *Balaenoptera acutorostrata* from Greenland, the North East Atlantic and the North Sea probably reflects different ecological regions. *MEPS* 247:263–280.

Berger, J. O. 1985. *Statistical decision theory and Bayesian analysis.* Second ed. Springer-Verlag, New York.

Borchers, D. L., D. Pike, T. Gunnlaugsson and G. Vkingsson 2003. Analyses of the NASS 1987 and 2001 minke whale cue counting surveys taking account of distance estimation errors.. *IWC/SC/55/NAM3. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Christensen, I. 1981. Age determination of minke whales, *Balaenoptera acutorostrata*, from laminated structures in the tympanic bullae. *Ann. Rep. int. Whal. Commn.* 31:245–253.

Gunnlaugsson, T., D. Pike, G. Vikingsson, G. Desportes and B. Mikkelson 2003. An estimate of the abundance of minke whales (*Balaenoptera acutorostrata*) from the NASS-2001 shipboard survey. *NAMMCO/SC/11/AE/6* .

Heide-Jørgensen, M. P., M. S. Juul and K. L. Laidre 2006. Estimates of large whale abundance in Greenland waters from a ship-based survey in 2005. *IWC/SC/58/AWMPx. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Horwood, J. 1989. *Biology and exploitation of the minke whale.* CRC Press, Boca Raton Fl.

Johnson, J. B., and K. S. Omland 2004. Model selection in ecology and evolution. *Trends Ecol. Evol.* 19:101–108.

Jonsgård, Å. 1962. Population studies on the minke whale *Balaenoptera acutorostrata* Lacepede. pp. 159–167, *In*: E. D. Le Cren and M. W. Holdgate (eds.) *The Exploitation of Natural Animal Populations.* Blackwell Scientific Publications, Oxford.

Kapel, F. O. 1980. Sex ratio and seasonal distribution of catches of minke whales in West Greenland. *Rep. int. Whal. Commn.* 30:195–199.

Larsen, F. 1984. Preliminary report of a survey for large cetaceans off West Greenland, 1983. *IWC/SC/36/O8. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Larsen, F. 1991. Biological parameters of North Atlantic minke whales. *Ann. Rep. int. Whal. Commn.* 41:160.

Larsen, F., and F. O. Kapel 1982. Norwegian minke whaling off West Greenland, 1976–80 and biological studies of West Greenland minke whales. *Rep. int. Whal. Commn.* 32:263–274.

Larsen, F., and F. O. Kapel 1983. Further biological studies of the West Greenland minke whale. *Ann. Rep. int. Whal. Commn.* 33:329–332.

Larsen, F., and N. Øien 1988. On the Discreteness of Stocks of Minke Whales at East and West Greenland. *Rep. int. Whal. Commn.* 38:251–255.

McAllister, M. K., E. K. Pikitch and E. A. Babcock 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. *Can. J. Fish. Aquat. Sci.* 58:1871–1890.

Mitchell, E. 1974. Preliminary report on Newfoundland fishery for minke whales (*Balaenoptera acutorostrata*). *Ann. Rep. int. Whal. Commn.* 24:159–176.

Mitchell, E., and V. M. Kozicki 1975. Supplementary information on minke whale (*Balaenoptera acutorostrata*) from Newfoundland fishery. *J. Fish. Res. Board Can.* 32:985–994.

Øien, N. 1988. Length distributions in catches from the north-eastern Atlantic stock of minke whales. *Rep. int. Whal. Commn.* 38:289–295.

Rubin, D. B. 1988. Using the SIR algorithm to simulate posterior distributions. pp. 395–402, *In*: J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. M. Smith (eds.) *Bayesian Statistics 3: Proceedings of the Third Valencia International Meeting, 1–5 June 1987.* Clarendon Press, Oxford.

Sergeant, D. E. 1963. Minke whales, *Balaenoptera acutorostrata* Lacepede, of the western North Atlantic. *J. Fish. Res. Board Can.* 20:1489–1504.

Sigurjonsson, J. 1988. Studies on age and reproduction in minke whales (*Balaenoptera acutorostrata* Lacepede) in Icelandic waters. *IWC/SC/40/Mi22. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Skaug, H. J., N. Øien, G. Bothun and T. Schweder 2002. Abundance of northeastern Atlantic minke whales for the survey period 1996-2001. *IWC/SC/54/RMP5. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Ugland, K. I. 1977. Studies in an equilibrium model for whale stock assessments. *Ann. Rep. int. Whal. Commn.* 38:402–412.

Witting, L. 2000. Indices of biological parameters for large cetaceans caught in Greenland 1988–1999. *IWC/SC/52/AS3. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Witting, L. 2005. An assessment of minke whales off West Greenland. *IWC/SC/57/AWMP4. Available from the International Whaling Commission (http://www.iwcoffice.org/)* .

Greenlandic whalers 1955 to 1978

Year	m	f	Year	m	f	Year	m	f	Year	m	f
1955	7	8	1961	7	9	1967	7	42	1973	8	39
1956	5	15	1962	17	43	1968	10	47	1974	6	34
1957	6	18	1963	32	47	1969	14	42	1975	1	17
1958	5	6	1964	26	37	1970	12	20	1976	2	20
1959	2	17	1965	19	30	1971	6	25	1977	15	39
1960	2	15	1966	24	49	1972	6	40	1978	2	13
Greenlandic whalers 1985 to 2005											
1985	59	163	1991	22	66	1997	42	99	2003	58	117
1986	38	107	1992	18	72	1998	39	118	2004	44	129
1987	14	29	1993	25	74	1999	34	123	2005	34	130
1988	6	34	1994	22	78	2000	36	102	2006	-	-
1989	14	32	1995	44	103	2001	32	91	2007	-	-
1990	15	63	1996	36	120	2002	33	88	2008	-	-
Norwegian whalers 1968 to 1985											
1968	7	13	1973	67	154	1978	10	65	1983	25	42
1969	119	46	1974	43	209	1979	31	44	1984	20	49
1970	74	52	1975	11	91	1980	13	62	1985	28	24
1971	86	177	1976	38	149	1981	15	46	1986	-	-
1972	32	91	1977	21	54	1982	24	42	1987	-	-

Table 1: **Yearly reporting** of caught male (m) and female (f) common minke whales by Greenlandic whalers from 1955 to 1978, and from 1985 to 2005, and by Norwegian whalers from 1968 to 1985.

Year	m	f	Year	m	f	Year	m	f	Year	m	f	Year	m	f
1948	1	3	1960	7	49	1972	52	227	1984	77	228	1996	38	126
1949	1	4	1961	15	20	1973	114	383	1985	87	187	1997	44	104
1950	2	7	1962	20	52	1974	76	393	1986	38	107	1998	41	125
1951	4	12	1963	67	99	1975	23	301	1987	28	58	1999	37	133
1952	8	24	1964	67	95	1976	55	323	1988	16	93	2000	38	107
1953	8	24	1965	76	120	1977	100	260	1989	19	44	2001	36	103
1954	5	17	1966	74	151	1978	34	221	1990	17	72	2002	38	101
1955	10	12	1967	35	209	1979	91	234	1991	28	81	2003	61	124
1956	6	16	1968	62	273	1980	75	258	1992	21	82	2004	46	133
1957	6	18	1969	186	248	1981	117	148	1993	27	80	2005	34	132
1958	14	16	1970	152	181	1982	84	232	1994	23	81	2006	-	-
1959	6	49	1971	124	335	1983	90	246	1995	46	107	2007	-	-

Table 2: **Yearly catch** of male (m) and female (f) West Greenland common minke whales, as reconstructed from total reported catch and reporting on caught males and females.

	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	N	50	50	50	18	18	18	6.3	6.3	6.3	2.2	2.2	2.2	.80	.80	.80
	msyr	.07	.04	.01	.07	.04	.01	.07	.04	.01	.07	.04	.01	.07	.04	.01
	msyl	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60
	ϑ_h	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
	ϑ_c	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
wga	ϑ	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
	N	1.0	5.0	25	1.0	5.0	25	1.0	5.0	25	1.0	5.0	25	1.0	5.0	25
	msyr	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
	msyl	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60
	ϑ_h	1.0	1.0	1.0	3.4	3.4	3.4	2.4	2.4	2.4	1.7	1.7	1.7	1.9	1.9	1.9
	ϑ_c	2.9	2.9	2.9	1.0	1.0	1.0	1.0	1.0	1.0	1.7	1.7	1.7	1.0	1.0	1.0
wgb	ϑ	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.60	.60	.60
	N	1.0	5.0	25	1.0	5.0	25	1.0	5.0	25	1.0	5.0	25	1.0	5.0	25
	msyr	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04	.04
	msyl	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60
	ϑ_h	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
	ϑ_c	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ϑ	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
wgc	p_{ex}	p_a	p_a	p_a	p_b	p_b	p_b	p_c	p_c	p_c	p_d	p_d	p_d	p_e	p_e	p_e

p_{ex}	Trial specification
p_a	ϑ_h scaling in time: $1x$ from 1948 to 1970, linear to $1.5x$ in 2005
p_b	ϑ_c scaling in time: $1x$ from 1948 to 1970, linear to $1.5x$ in 2005
p_c	ϑ_h & ϑ_c in time: both $1x$ from 1948 to 1970, linear to $1.2x$ in 2005
p_d	Catch selection over age: $c_0 = 0$, linear $c_1 = 0.07$ to $c_x = 1$
p_e	Catch selection over age: $c_0 = 0$, linear $c_1 = 1$ to $c_x = 0.07$

Table 3: Parameter values for the trials (T), of the **wga**, **wgb**, and **wgc** trial sets.

Par	s_{ad}	s_{juv}	b_{max}	a_m
min	0.92	0.10	0.50	3
max	0.99	0.99	1.00	10

Table 4: Minimum and maximum values of the uniform priors on the life history parameters.

T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
wga	1210	536	4437	17	78	1071	0	0	0	0	0	0	0	0	0
wgb	0	0	207	0	0	1	0	0	0	0	0	198	0	0	413
wgc	0	10	0	0	0	29	0	0	1189	0	0	67	0	0	539

Table 5: The Akaike weights of the different trials given a total weight of $10,000$. For details see the main text.

wga	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	p	$\tilde{p}$
	K	.29	.30	.31	.57	.59	.56	1.0	.98	.82	1.5	1.2	.97	1.5	1.2	1.0	.86	.35
	N	.21	.24	.30	.38	.43	.54	.51	.62	.87	.51	.78	1.4	.64	1.0	2.1	.70	.32
avg	sh	.05	.10	.45	.11	.21	.85	.18	.32	1.3	.13	.30	1.9	.21	.37	3.8	.69	.41
	K	.18	.19	.19	.40	.44	.42	.82	.80	.70	1.3	1.1	.83	1.4	1.1	.86	.71	.23
	N	.09	.11	.15	.17	.24	.29	.22	.32	.46	.19	.39	.72	.14	.45	1.1	.34	.16
min	sh	.03	.06	.29	.07	.15	.60	.07	.19	.89	.04	.14	1.0	.05	.17	1.9	.38	.28
	K	.85	.89	.84	.85	1.3	1.1	1.3	1.4	1.2	1.6	1.5	1.2	1.7	1.4	1.1	1.2	.89
	N	.75	.85	.94	.67	1.2	1.4	.86	1.3	1.5	.88	1.5	2.2	1.2	2.0	3.9	1.4	.98
max	sh	.10	.18	.74	.16	.47	1.7	.26	.52	2.2	.23	.53	3.9	1.0	.77	7.1	1.3	.72

wgb	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	p	$\tilde{p}$
	K	1.3	1.1	.48	1.3	1.4	1.6	1.2	.91	.29	1.3	1.1	.47	1.9	1.7	.69	1.1	.58
	N	.92	.64	.36	1.2	1.2	1.5	.66	.37	.15	.94	.64	.35	3.6	1.4	.56	.97	.46
avg	sh	.32	.33	.17	.46	.48	.32	.26	.20	.09	.32	.33	.16	.87	.50	.20	.33	.18
	K	1.1	.90	.35	1.1	.99	.52	1.1	.81	.26	1.1	.89	.34	1.6	1.4	.48	.86	.41
	N	.42	.34	.20	.54	.49	.37	.28	.21	.10	.39	.32	.20	2.0	.80	.30	.46	.25
min	sh	.16	.19	.12	.20	.27	.17	.11	.10	.07	.16	.18	.11	.50	.37	.13	.19	.12
	K	1.4	1.4	1.1	1.5	3.6	2.6	1.4	1.0	.35	1.4	1.4	1.1	2.4	2.2	1.5	1.6	1.3
	N	2.0	1.2	1.1	2.5	4.5	2.6	1.2	.57	.23	1.8	1.2	1.1	6.1	2.6	1.4	2.0	1.2
max	sh	.64	.52	.35	.94	1.3	.37	.94	.33	.13	.66	.52	.35	1.4	.69	.36	.63	.36

wgc	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	p	$\tilde{p}$
	K	1.4	2.3	2.5	1.2	.98	.36	1.3	1.1	.65	1.3	.87	.42	1.0	1.3	.56	1.1	.62
	N	1.7	2.5	2.4	.79	.50	.23	1.1	.76	.55	.91	.41	.30	.47	.98	.45	.93	.51
avg	sh	.69	.73	.33	.28	.27	.13	.38	.38	.21	.32	.20	.15	.34	.45	.19	.34	.20
	K	1.2	1.1	1.6	1.1	.86	.30	1.1	.92	.38	1.1	.76	.33	.85	1.0	.36	.86	.37
	N	.66	.66	1.4	.34	.28	.14	.53	.37	.23	.43	.22	.19	.15	.53	.22	.42	.22
min	sh	.26	.35	.30	.14	.15	.09	.16	.22	.12	.12	.10	.11	.06	.31	.12	.17	.12
	K	1.9	8.7	2.7	1.4	1.2	.57	1.5	1.7	2.1	1.4	.98	.85	1.2	2.1	1.6	2.0	1.9
	N	4.2	11	2.8	1.4	.84	.47	2.1	1.7	2.1	1.9	.64	.78	.90	2.3	1.5	2.3	1.9
max	sh	1.4	1.4	.37	.76	.42	.20	.77	.61	.36	.64	.31	.31	1.1	.75	.35	.65	.36

Table 6: The average, minimum and maximum of the median estimate over the true trial value for selected parameters, for the hundred iterations of the different **wga**, **wgb**, and **wgc** trials. The unweighted average and the Akaike weighted average across each trial set is given by p and $\tilde{p}$.

wga	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	p	$\tilde{p}$
	K	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.00	.09	1.0	.00	.02	1.0	.74	1.0
	N	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	.54	.97	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.88	1.0	1.0	.47	.96	1.0
	ry	.96	.99	.81	1.0	1.0	.91	1.0	1.0	.91	1.0	1.0	.75	1.0	1.0	.13	.90	.86
2.5%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.91	1.0	1.0	.75	1.0	1.0	.13	.92	1.0
	K	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.00	.07	1.0	.00	.02	1.0	.74	1.0
	N	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.95	1.0	1.0	.43	.96	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.87	1.0	1.0	.39	.95	1.0
	ry	.91	.98	.78	1.0	1.0	.87	1.0	1.0	.87	1.0	1.0	.62	1.0	1.0	.11	.88	.83
5.0%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.88	1.0	1.0	.62	1.0	1.0	.11	.91	1.0
	K	1.0	1.0	1.0	1.0	1.0	1.0	.98	.99	1.0	.00	.01	1.0	.00	.00	.94	.73	1.0
	N	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.88	1.0	1.0	.30	.95	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	.99	1.0	1.0	.99	1.0	1.0	1.0	1.0	1.0	.77	1.0	1.0	.31	.94	.99
	ry	.86	.92	.62	1.0	.99	.81	1.0	1.0	.80	1.0	1.0	.33	1.0	1.0	.03	.82	.71
10%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.80	1.0	1.0	.33	1.0	1.0	.03	.88	1.0
	K	1.0	1.0	1.0	1.0	1.0	1.0	.90	.99	1.0	.00	.01	.99	.00	.00	.79	.71	1.0
	N	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	.71	1.0	.95	.14	.92	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	.97	1.0	1.0	.97	1.0	1.0	.92	1.0	1.0	.57	1.0	1.0	.15	.91	.98
	ry	.66	.89	.46	.95	.99	.67	1.0	1.0	.67	1.0	1.0	.13	1.0	1.0	.00	.76	.56
20%	sh	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	.59	1.0	1.0	.13	1.0	1.0	.00	.85	1.0
	K	1.0	1.0	1.0	1.0	1.0	1.0	.72	.91	.99	.00	.00	.95	.00	.00	.64	.68	1.0
	N	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	.98	1.0	.99	.45	1.0	.94	.08	.90	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	.96	1.0	1.0	.92	1.0	1.0	.78	1.0	1.0	.38	1.0	.96	.05	.87	.96
	ry	.54	.81	.36	.90	.98	.47	1.0	1.0	.42	1.0	1.0	.06	1.0	1.0	.00	.70	.45
30%	sh	1.0	1.0	1.0	1.0	1.0	.94	1.0	1.0	.40	1.0	1.0	.06	1.0	1.0	.00	.83	.99
	K	1.0	1.0	1.0	1.0	.99	1.0	.64	.79	.99	.00	.00	.91	.00	.00	.51	.66	1.0
	N	1.0	1.0	1.0	1.0	.99	.98	1.0	.99	.90	1.0	.94	.29	.99	.77	.01	.86	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	.93	1.0	1.0	.86	1.0	1.0	.64	1.0	1.0	.25	1.0	.93	.01	.84	.94
	ry	.39	.70	.21	.87	.94	.27	1.0	1.0	.21	1.0	1.0	.01	1.0	1.0	.00	.64	.29
40%	sh	1.0	1.0	1.0	1.0	1.0	.89	1.0	1.0	.17	1.0	1.0	.01	1.0	1.0	.00	.80	.98
	K	1.0	1.0	1.0	1.0	.99	.98	.54	.67	.96	.00	.00	.78	.00	.00	.38	.62	1.0
	N	1.0	1.0	1.0	1.0	.99	.98	1.0	.99	.72	1.0	.83	.16	.97	.58	.00	.81	1.0
	$msyr$	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	1.0	1.0	.00	.67	.25
	d	1.0	1.0	.90	1.0	1.0	.82	1.0	1.0	.49	1.0	.98	.12	1.0	.79	.00	.81	.91
	ry	.25	.57	.12	.75	.92	.15	.99	1.0	.06	1.0	1.0	.00	.99	1.0	.00	.59	.19
50%	sh	1.0	1.0	1.0	1.0	1.0	.86	1.0	1.0	.08	1.0	1.0	.00	.99	1.0	.00	.80	.98

Table 7: The probabilities that the percentiles of parameter estimates are below the parameter values for the **wga**-trials (T), with p giving the average probability and $\tilde{p}$ the Akaike weighted probability across trials.

wgb	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	p	$\tilde{p}$
	K	.02	1.0	1.0	.01	.64	1.0	.06	1.0	1.0	.02	.99	1.0	.00	.00	1.0	.58	1.0
	N	1.0	1.0	1.0	1.0	.98	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.02	.99	1.0	.93	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.57	1.0	1.0	.97	1.0
	ry	1.0	1.0	.99	1.0	1.0	.94	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.94	.99	.97
2.5%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.01	.98	1.0	.00	.51	1.0	.04	1.0	1.0	.01	.98	1.0	.00	.00	1.0	.57	1.0
	N	1.0	1.0	1.0	.99	.98	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.02	.99	1.0	.93	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.49	1.0	1.0	.97	1.0
	ry	1.0	1.0	.99	1.0	1.0	.88	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	.92	.99	.95
5.0%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.00	.94	1.0	.00	.34	.82	.02	1.0	1.0	.00	.94	1.0	.00	.00	1.0	.54	1.0
	N	1.0	1.0	1.0	.96	.95	.90	1.0	1.0	1.0	1.0	1.0	1.0	.01	.95	1.0	.92	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	1.0	1.0	1.0	1.0	.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.24	1.0	1.0	.95	1.0
	ry	1.0	1.0	.99	1.0	1.0	.79	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	.90	.98	.94
10%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.00	.82	1.0	.00	.11	.54	.00	1.0	1.0	.00	.81	1.0	.00	.00	1.0	.49	1.0
	N	.96	1.0	1.0	.89	.84	.63	1.0	1.0	1.0	.96	1.0	1.0	.00	.70	1.0	.87	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	1.0	1.0	1.0	.95	.97	.97	1.0	1.0	1.0	1.0	1.0	1.0	.06	1.0	1.0	.93	1.0
	ry	1.0	1.0	.96	1.0	1.0	.60	1.0	1.0	1.0	1.0	1.0	.97	.99	1.0	.78	.95	.87
20%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0	1.0	1.0	1.0
	K	.00	.62	1.0	.00	.03	.41	.00	1.0	1.0	.00	.63	1.0	.00	.00	.98	.44	.99
	N	.94	1.0	1.0	.72	.71	.49	1.0	1.0	1.0	.95	1.0	1.0	.00	.42	1.0	.82	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.99	1.0	1.0	.87	.95	.92	1.0	1.0	1.0	.98	1.0	1.0	.01	1.0	1.0	.91	1.0
	ry	1.0	1.0	.92	1.0	1.0	.47	1.0	1.0	1.0	1.0	1.0	.93	.96	1.0	.65	.93	.79
30%	sh	1.0	1.0	1.0	1.0	.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.97	1.0	1.0	1.0	1.0
	K	.00	.45	1.0	.00	.01	.31	.00	1.0	1.0	.00	.46	1.0	.00	.00	.96	.41	.98
	N	.85	.99	1.0	.52	.55	.38	.96	1.0	1.0	.83	.99	1.0	.00	.19	.97	.75	.98
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.93	1.0	1.0	.72	.91	.87	1.0	1.0	1.0	.94	1.0	1.0	.00	.99	1.0	.89	1.0
	ry	1.0	1.0	.88	1.0	1.0	.32	1.0	1.0	1.0	1.0	1.0	.90	.90	1.0	.54	.90	.71
40%	sh	1.0	1.0	1.0	1.0	.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.90	1.0	1.0	.99	1.0
	K	.00	.30	.99	.00	.01	.23	.00	.98	1.0	.00	.28	.99	.00	.00	.94	.38	.96
	N	.65	.98	.99	.41	.49	.29	.94	1.0	1.0	.64	.98	.99	.00	.07	.96	.69	.97
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.87	1.0	1.0	.55	.83	.76	.99	1.0	1.0	.84	1.0	1.0	.00	.98	1.0	.85	1.0
	ry	1.0	1.0	.81	1.0	1.0	.20	1.0	1.0	.99	1.0	1.0	.82	.74	1.0	.45	.87	.63
50%	sh	1.0	1.0	1.0	1.0	.98	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.79	1.0	1.0	.98	1.0

Table 8: The probabilities that the percentiles of parameter estimates are below the parameter values for the **wgb**-trials (T), with p giving the average probability and $\tilde{p}$ the Akaike weighted probability across trials.

wgc	T	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	p	$\tilde{p}$
	K	.00	.12	1.0	.04	1.0	1.0	.01	.98	1.0	.02	1.0	1.0	.93	.79	1.0	.66	1.0
	N	.91	.72	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.98	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.97	.98	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ry	1.0	1.0	.92	1.0	1.0	1.0	1.0	1.0	.98	1.0	1.0	1.0	1.0	1.0	.98	.99	.98
2.5%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.00	.06	1.0	.03	1.0	1.0	.01	.90	1.0	.01	1.0	1.0	.82	.59	1.0	.63	1.0
	N	.88	.64	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0	.97	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.92	.92	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0
	ry	1.0	1.0	.87	1.0	1.0	1.0	1.0	1.0	.98	1.0	1.0	1.0	1.0	1.0	.98	.99	.98
5.0%	sh	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.00	.02	.54	.01	1.0	1.0	.00	.82	1.0	.00	1.0	1.0	.73	.30	1.0	.56	.99
	N	.76	.55	.80	1.0	1.0	1.0	.96	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0	.94	1.0
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.91	.79	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.98	1.0
	ry	1.0	1.0	.73	1.0	1.0	1.0	1.0	1.0	.93	1.0	1.0	.99	1.0	1.0	.92	.97	.93
10%	sh	1.0	.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.00	.01	.08	.00	.98	1.0	.00	.58	.99	.00	1.0	1.0	.65	.12	1.0	.49	.99
	N	.55	.41	.11	1.0	1.0	1.0	.94	.99	.99	.96	1.0	1.0	1.0	.99	1.0	.86	.99
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.76	.71	.91	1.0	1.0	1.0	.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.96	1.0
	ry	1.0	1.0	.57	1.0	1.0	.99	1.0	1.0	.89	1.0	1.0	.99	1.0	1.0	.90	.96	.90
20%	sh	1.0	.94	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	K	.00	.00	.00	.00	.93	1.0	.00	.35	.95	.00	1.0	1.0	.57	.04	.99	.46	.96
	N	.41	.25	.01	.97	1.0	1.0	.80	.99	.98	.94	1.0	1.0	1.0	.92	.99	.82	.98
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.60	.65	.76	1.0	1.0	1.0	.93	1.0	1.0	.98	1.0	1.0	1.0	.99	1.0	.93	1.0
	ry	.98	.99	.38	1.0	1.0	.99	1.0	1.0	.82	1.0	1.0	.99	1.0	1.0	.78	.93	.82
30%	sh	.98	.88	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.99	1.0
	K	.00	.00	.00	.00	.84	1.0	.00	.18	.93	.00	1.0	1.0	.43	.00	.98	.42	.94
	N	.23	.20	.00	.95	1.0	1.0	.65	.96	.94	.86	1.0	1.0	1.0	.81	.99	.77	.95
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.49	.54	.69	.99	1.0	1.0	.85	1.0	1.0	.93	1.0	1.0	1.0	.98	1.0	.90	1.0
	ry	.92	.99	.25	1.0	1.0	.98	1.0	1.0	.67	1.0	1.0	.98	.98	.99	.67	.90	.69
40%	sh	.92	.85	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.98	1.0	1.0	.98	1.0
	K	.00	.00	.00	.00	.68	1.0	.00	.07	.90	.00	1.0	1.0	.33	.00	.95	.40	.91
	N	.10	.10	.00	.86	1.0	1.0	.47	.88	.90	.67	1.0	1.0	1.0	.62	.97	.70	.92
	$msyr$	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	d	.35	.44	.54	.97	1.0	1.0	.72	.99	1.0	.88	1.0	1.0	1.0	.98	1.0	.86	1.0
	ry	.88	.98	.07	1.0	1.0	.93	1.0	1.0	.57	1.0	1.0	.93	.96	.99	.55	.86	.59
50%	sh	.88	.77	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.96	1.0	1.0	.97	1.0

Table 9: The probabilities that the percentiles of parameter estimates are below the parameter values for the **wgc**-trials (T), with p giving the average probability and $\tilde{p}$ the Akaike weighted probability across trials.

		K	N	$msyr$	d	ry	sh
	e	21400	11600	0.01	0.62	61.2	127
	p	0.66	0.96	0.89	0.97	0.96	0.97
2.5%	$\tilde{p}$	1.00	1.00	0.75	1.00	0.94	1.00
	e	23200	13100	0.01	0.66	72.2	144
	p	0.65	0.95	0.89	0.97	0.95	0.97
5.0%	$\tilde{p}$	1.00	1.00	0.75	1.00	0.92	1.00
	e	25900	15600	0.01	0.71	84.4	162
	p	0.61	0.93	0.89	0.96	0.92	0.96
10%	$\tilde{p}$	1.00	1.00	0.75	1.00	0.86	1.00
	e	29600	19600	0.01	0.77	101	175
	p	0.56	0.88	0.89	0.93	0.89	0.95
20%	$\tilde{p}$	1.00	1.00	0.75	0.99	0.78	1.00
	e	33700	24200	0.01	0.82	113	175
	p	0.53	0.84	0.89	0.90	0.85	0.94
30%	$\tilde{p}$	0.98	0.99	0.75	0.99	0.68	1.00
	e	39000	29200	0.01	0.85	124	175
	p	0.50	0.79	0.89	0.88	0.81	0.93
40%	$\tilde{p}$	0.97	0.98	0.75	0.98	0.56	0.99
	e	44500	34000	0.02	0.87	134	175
	p	0.47	0.74	0.89	0.84	0.77	0.92
50%	$\tilde{p}$	0.96	0.96	0.75	0.97	0.47	0.99

Table 10: Percentiles of parameter estimates (e), and the probabilities that the estimates are below the true values given the **wga**, **wgb**, and **wgc** trial sets, with p being the unweighted probability and $\tilde{p}$ the Arkike weighted probability.

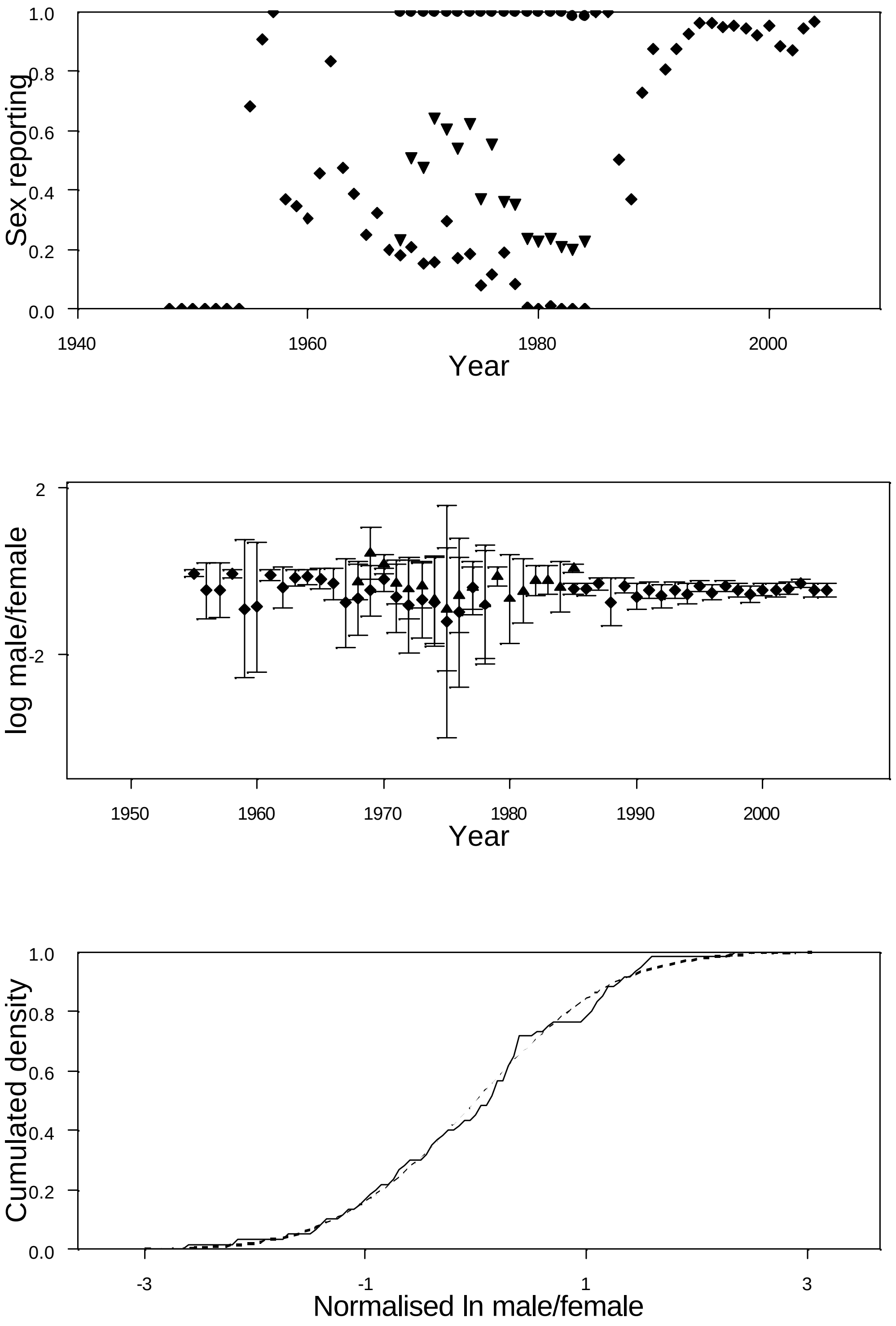


Figure 1: **Top:** The proportion of catches with reported sex; diamonds: Greenlandic whalers, dots: Norwegian whalers, triangles: combined data. **Middle:** The log of the reported male / female ratio; including 95% CI. **Bottom:** The cumulated density of the $N(0;1)$ normal distribution (dashed curve), and the normalised distribution of the logarithm of the reported male/female ratio in the catches from 1955 to 2004 (solid curve).

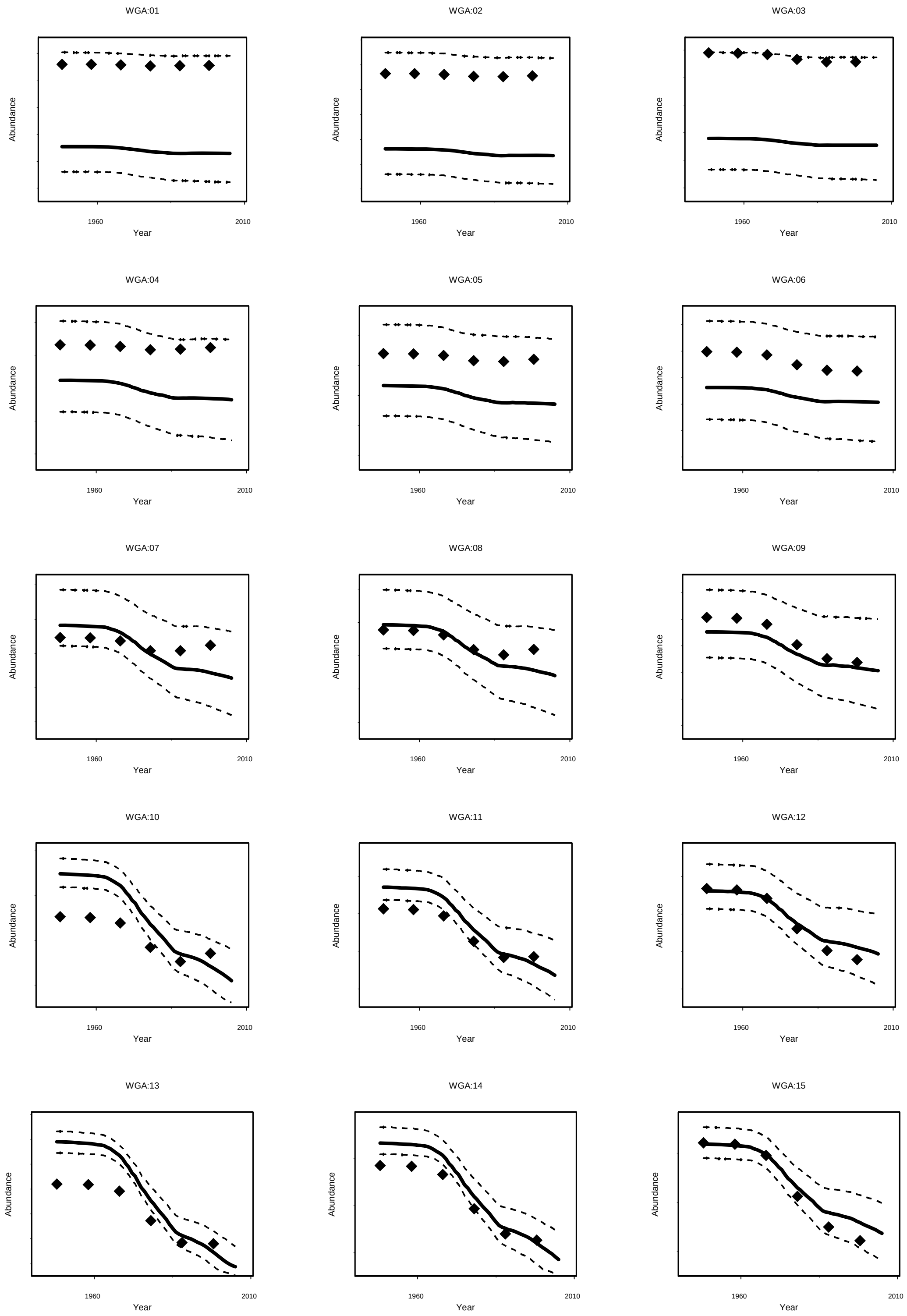


Figure 2: The projection of a single realisation of the **wga** trials, with the true abundance given by the diamonds, and the median (solid curve) and the 90% credibility interval (dashed curves) for the corresponding Bayesian assessment.

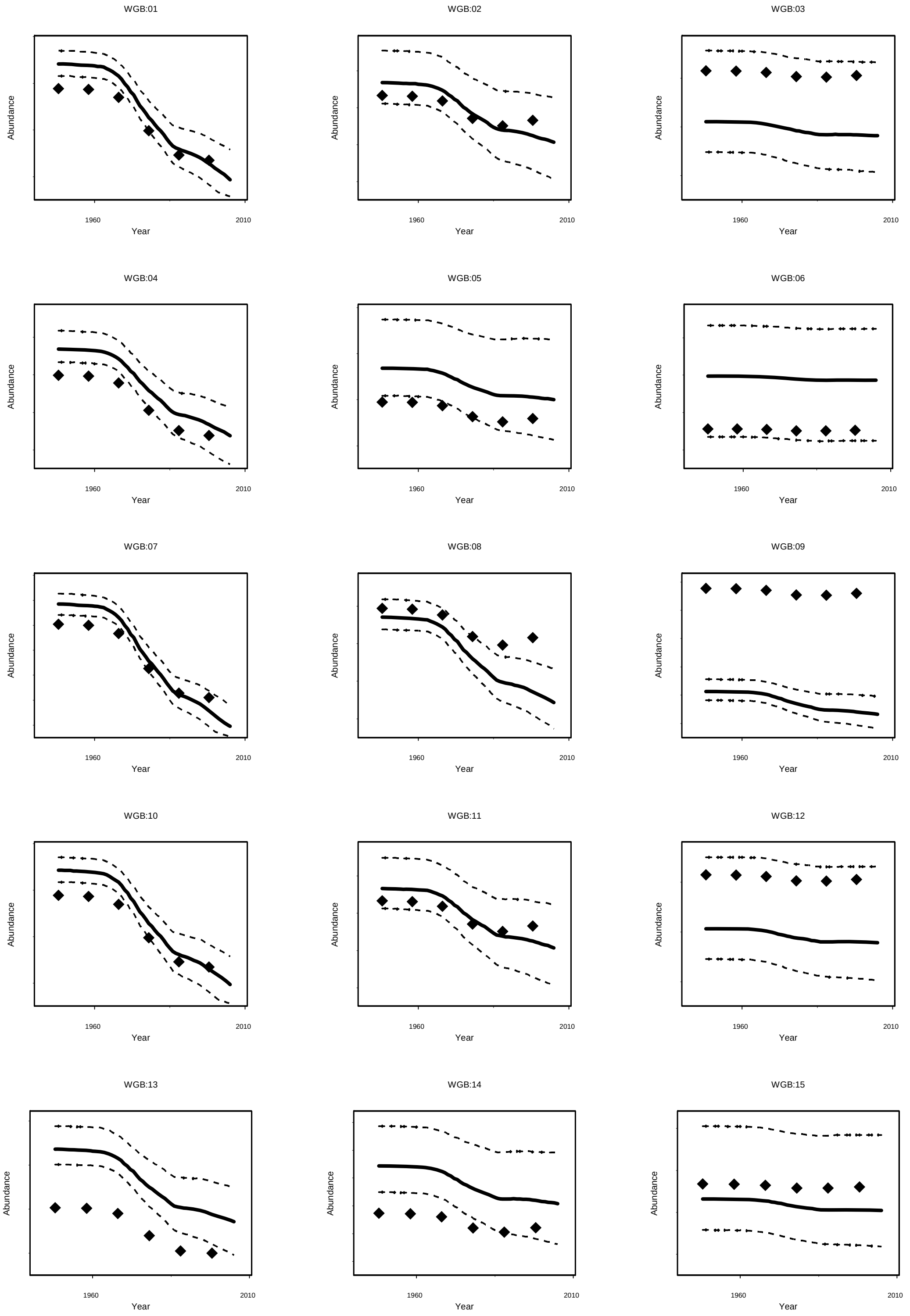


Figure 3: The projection of a single realisation of the **wgb** trials, with the true abundance given by the diamonds, and the median (solid curve) and the 90% credibility interval (dashed curves) for the corresponding Bayesian assessment.

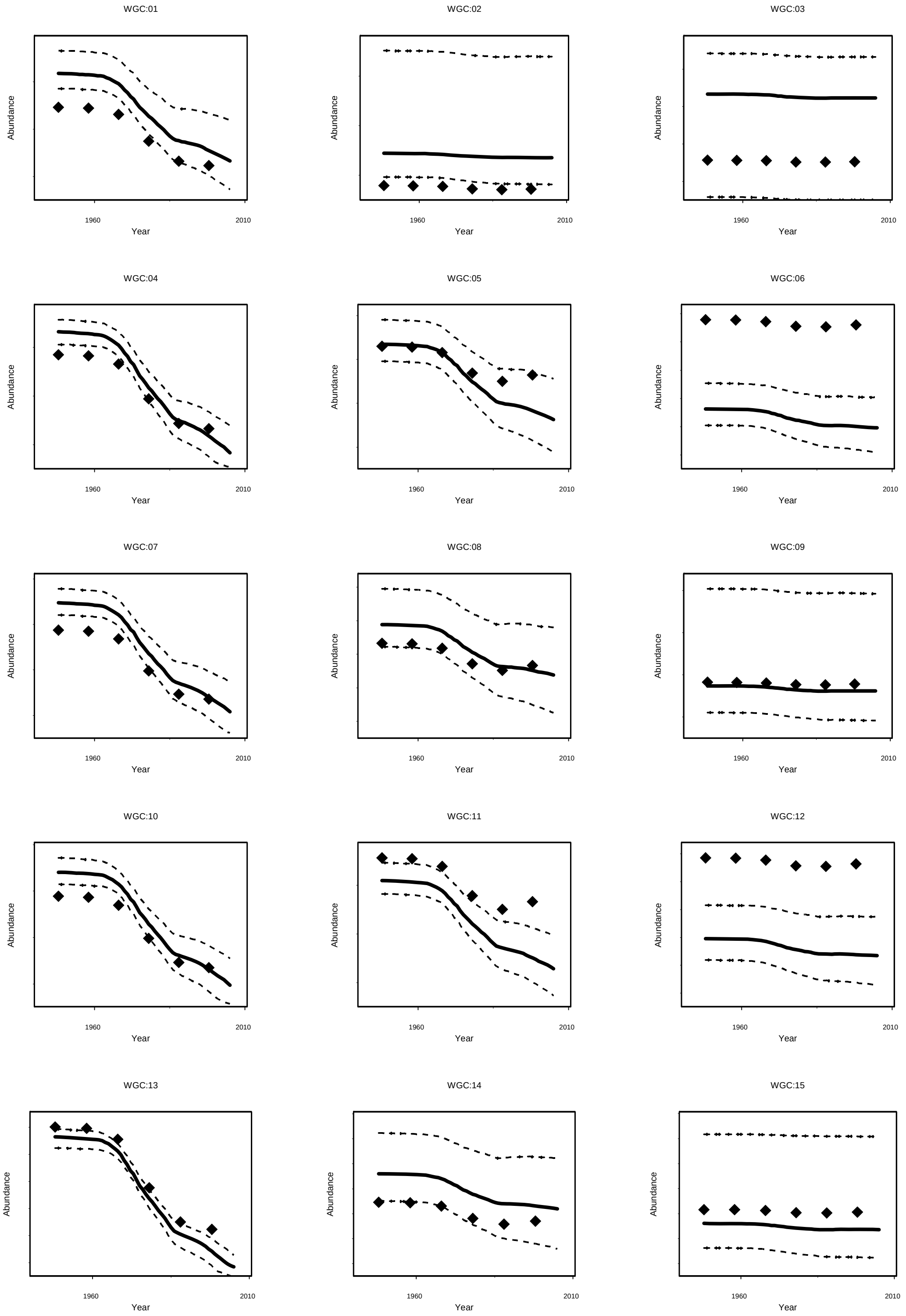
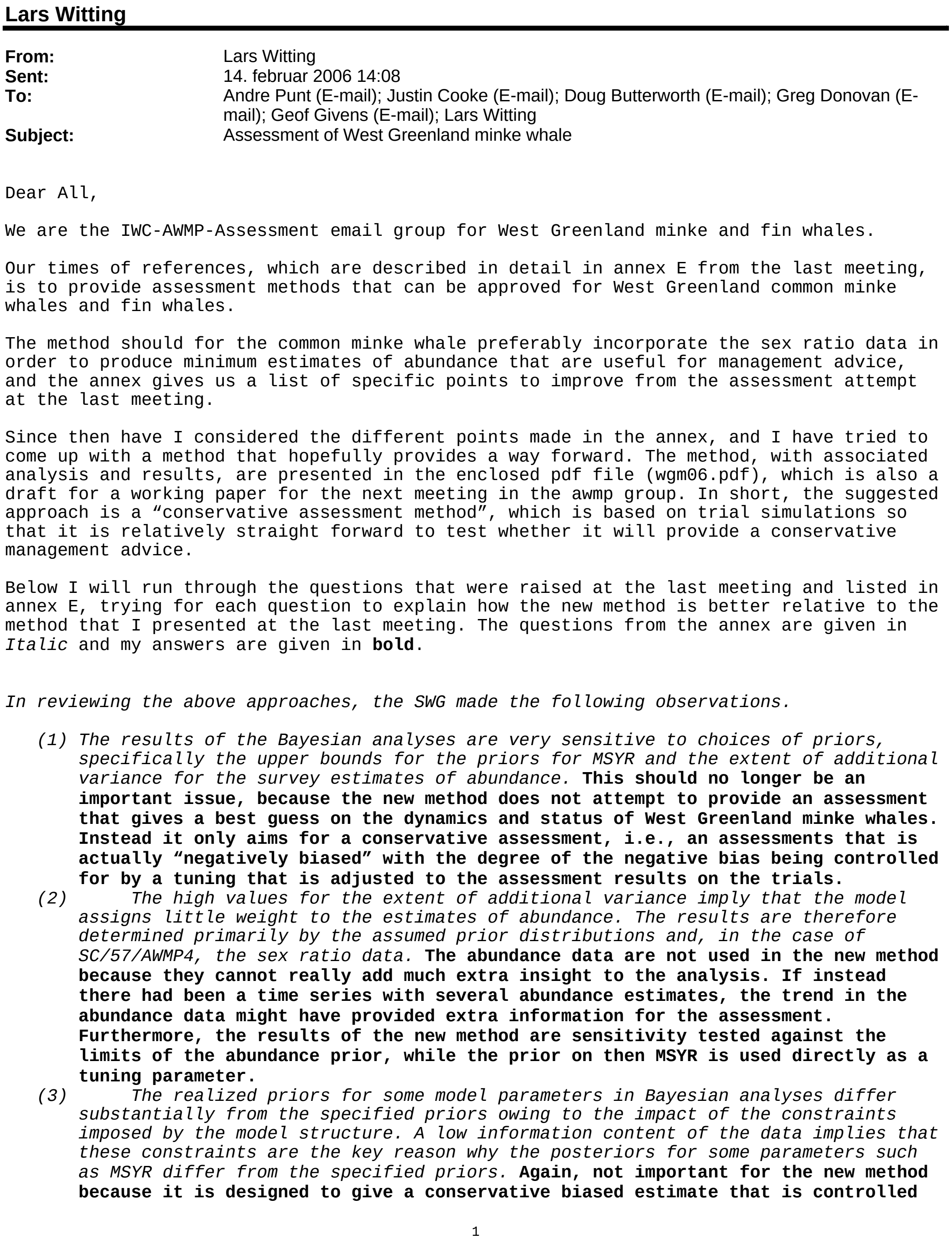


Figure 4: The projection of a single realisation of the **wgc** trials, with the true abundance given by the diamonds, and the median (solid curve) and the 90% credibility interval (dashed curves) for the corresponding Bayesian assessment.

Appendix



Lars Witting

From: Lars Witting
Sent: 14. februar 2006 14:08
To: Andre Punt (E-mail); Justin Cooke (E-mail); Doug Butterworth (E-mail); Greg Donovan (E-mail); Geof Givens (E-mail); Lars Witting
Subject: Assessment of West Greenland minke whale

Dear All,

We are the IWC-AWMP-Assessment email group for West Greenland minke and fin whales.

Our times of references, which are described in detail in annex E from the last meeting, is to provide assessment methods that can be approved for West Greenland common minke whales and fin whales.

The method should for the common minke whale preferably incorporate the sex ratio data in order to produce minimum estimates of abundance that are useful for management advice, and the annex gives us a list of specific points to improve from the assessment attempt at the last meeting.

Since then have I considered the different points made in the annex, and I have tried to come up with a method that hopefully provides a way forward. The method, with associated analysis and results, are presented in the enclosed pdf file (wgm06.pdf), which is also a draft for a working paper for the next meeting in the awmp group. In short, the suggested approach is a “conservative assessment method”, which is based on trial simulations so that it is relatively straight forward to test whether it will provide a conservative management advice.

Below I will run through the questions that were raised at the last meeting and listed in annex E, trying for each question to explain how the new method is better relative to the method that I presented at the last meeting. The questions from the annex are given in *Italic* and my answers are given in **bold**.

In reviewing the above approaches, the SWG made the following observations.

(1) *The results of the Bayesian analyses are very sensitive to choices of priors, specifically the upper bounds for the priors for MSYR and the extent of additional variance for the survey estimates of abundance.* **This should no longer be an important issue, because the new method does not attempt to provide an assessment that gives a best guess on the dynamics and status of West Greenland minke whales. Instead it only aims for a conservative assessment, i.e., an assessments that is actually “negatively biased” with the degree of the negative bias being controlled for by a tuning that is adjusted to the assessment results on the trials.**

(2) *The high values for the extent of additional variance imply that the model assigns little weight to the estimates of abundance. The results are therefore determined primarily by the assumed prior distributions and, in the case of SC/57/AWMP4, the sex ratio data.* **The abundance data are not used in the new method because they cannot really add much extra insight to the analysis. If instead there had been a time series with several abundance estimates, the trend in the abundance data might have provided extra information for the assessment. Furthermore, the results of the new method are sensitivity tested against the limits of the abundance prior, while the prior on then MSYR is used directly as a tuning parameter.**

(3) *The realized priors for some model parameters in Bayesian analyses differ substantially from the specified priors owing to the impact of the constraints imposed by the model structure. A low information content of the data implies that these constraints are the key reason why the posteriors for some parameters such as MSYR differ from the specified priors.* **Again, not important for the new method because it is designed to give a conservative biased estimate that is controlled**

for by running assessments on trials, with the prior on the MSYR being used as a tuning parameter.

(4) *The approach used in SC/57/AWMP4 to make use of the data on the sex ratio of the catch has the potential to determine a lower bound for the abundance of the total stock (rather than just that component that feeds off West Greenland). However, at present, the fits to the data on sex ratio are poor.* **Fig.~1 in the enclosed paper illustrates the CI on the sex ratio, showing that the large between year variation of the point estimates is likely a reflection of uncertainty in the historical sex ratio estimates, and as such it is only the overall trend in the data that is important for the model to explain. There is no overall trend in the reported sex ratio with time, and this absence can be explained by the underlying model as shown last year. The absence of a trend, however, should not necessarily be present in our preferred models because we do not aim for models that give the best estimate of the true dynamics of West Greenland minke whales, but instead for models that give conservative estimates. In result some of the conservative models show an increase in the male fraction with time because they are conservative, and this divergence is not problematic as it is a natural part of the process of generating a conservative estimate.**

(5) *The penalty imposed on equilibrium abundance in SC/57/AWMP4 is highly influential, including on the lower bound of equilibrium abundance and MSYR, but the tuning levels are essentially arbitrary.* **There is no longer such a tuning method in the current approach. The tuning of the conservatism in the assessment is instead performed by comparing assessments results to trial information, and thus it is no longer arbitrary but possible to control do a desired level of conservatism.**

(6) *The production model assessments assume that the estimates of abundance pertain to absolute population size although this assumption is likely to be invalid to some (possibly substantial) extent.* **No longer relevant as no abundance estimates are used.**

(7) *In the case of the fin whale assessment, the posterior median time trajectory of 1+ abundance did not correspond well to the observed estimates of abundance.* **Not relevant for the minke whale that is considered here.**

Recognising that the consistently skewed sex ratio in the West Greenland common minke whale catches (see Item 2.1 and Appendix 7) is a conspicuous feature of the fishery, the SWG agreed that the sex-ratio data should be incorporated into future attempts at assessments because they can in principle provide information about the lower bound for the total abundance of the stock. However, any assessment based on these data must examine the sensitivity of the results to assumptions associated with their inclusion, including sensitivity to: (1) the assumption that the catch is taken uniformly from all age-classes greater than age 1; **I have also included trials that are based on an increasing and decreasing harvest selectivity over minke whale age.** *(2) The assumption that there have been no changes in sex selectivity over time;* **I have now include trials that have an increasing trend in the female bias of a sex selective hunt and/or a sex specific dispersal.** *And (3) the form of the likelihood function for the sex ratio data.* **I had some discussion with Andre on this in December to understand the point. His main point was that the likelihood function might better be described by a beta distribution on the female fraction, than by a log normal distribution on the ratio between the male fraction and the female fraction. As you will se in Fig.~1 in the paper, I tested the assumption of log normality on the ratio of the fractions and found that it was very close to log normal. In result I maintained the log normal distribution for the likelihood function.** *The SWG agreed that it might be valuable to base future preliminary assessments for common minke whales off West Greenland on maximum likelihood methods because they are not affected by the choice of priors.* **As explained above I chose a different approach to avoid the problem with influential priors.**

I am looking forward for your input to guide the way for an acceptable conservative assessment for West Greenland common minke whales.

Best wishes, Lars