Abundance of Bering-Chukchi-Beaufort bowhead whales (*Balaena mysticetus*) in 2004 estimated from photo-identification data

W.R. KOSKI*, J. ZEH+, J. MOCKLIN#, A.R. DAVIS*, D.J. RUGH#, J.C. GEORGE++ AND R. SUYDAM++

Contact e-mail: bkoski@lgl.com

ABSTRACT

Ice-based surveys near Point Barrow, Alaska, have been used to obtain most estimates of abundance for the Bering-Chukchi-Beaufort (B-C-B) stock of bowhead whales, but global warming has raised concerns that ice-based surveys may not be practical in the future. Aerial photographic surveys provide an alternative method for obtaining abundance estimates and may replace ice-based surveys. Aerial photographic surveys were conducted near Point Barrow during the spring migrations of bowhead whales in 2003 and 2004 and, in 2005, in the northern Bering Sea in spring and near Barrow in fall. The 2003 survey was the most complete photographic survey of the population conducted to date. These surveys provided photo-identification data for use in capture-recapture analyses. A screening procedure was used to define which whales captured in 2003, 2004 and/or 2005 were marked and could be reidentified if photographed on another occasion. An estimate of the number of marked whales was obtained using a closed population model for capture-recapture data. Several models were investigated, including models that accounted for heterogeneity in capture probabilities, but a simple model with no covariates produced the most precise estimate. To account for unmarked whales, the estimate of marked whales was divided by an estimate of the proportion of the bowhead population that was marked based on the 1989–2004 spring photographic surveys near Point Barrow. Abundance of the B-C-B bowhead population in 2004 (excluding calves) was estimated to be 12,631 with CV 0.2442, 95% bootstrap percentile confidence interval (7,900; 19,700) and 5% lower limit 8,400. These results were compared with results that used approximate variance expressions for the estimates of the number of marked whales, the proportion of the population that was marked and population abundance instead of using the bootstrap. The estimates of abundance in 2004 computed for comparison included one based on a modified Petersen estimate of the number of marked whales that omitted the 2005 data as well as the estimate of 12,631 described above. The comparison estimates also included estimates of abundance in 1985 computed from 1984–87 photographic survey data using the same methods. All the abundance estimates computed from photographic data were consistent with expectations based on independent abundance and trend estimates from the ice-based surveys conducted from 1978 to 2001.

KEYWORDS: ABUNDANCE ESTIMATE; MARK-RECAPTURE; SURVEY-AERIAL; PHOTO-ID; BOWHEAD WHALE; ARCTIC; BEAUFORT SEA

INTRODUCTION

Aerial photography projects conducted from 1981–2000 have provided much of the life history data that are available on the Bering-Chukchi-Beaufort (B-C-B) stock of the bowhead whale (Angliss *et al.*, 1995; da-Silva *et al.*, 2007; Koski *et al.*, 1992; 1993; 2006; Miller *et al.*, 1992; Nerini *et al.*, 1984; Rugh *et al.*, 1992b; Zeh *et al.*, 2002; 1993). The last major photographic effort during that period was conducted in 1992, although smaller scale photography projects were conducted during 1994 and 1998–2000.

The 1985 and 1986 photography projects also provided data that were used to make abundance estimates (da Silva *et al.*, 2000; da-Silva, 2003; da-Silva *et al.*, 2003; da-Silva and Tiburcio, 2010; Schweder, 2003) using closed population capture-recapture models. These estimates and their precision were similar to estimates from ice-based surveys in 1985 and 1986 (da Silva *et al.*, 2000). The capture-recapture estimates were based on photographic images of the midback zone of the whales scored as being of acceptable quality and identifiability (Rugh *et al.*, 1998). Zeh *et al.* (2000; 2002) developed a data screening method that allowed natural marks[1] in all four zones (rostrum, midback, lower back and flukes) to be used without risking failure to recognise recaptures because different zones of the whale were visible in images taken on different sampling occasions. This screening method provided larger sample sizes of naturally marked whales and increased precision of estimates based on their images. It was used to estimate annual survival probability of bowheads by Zeh *et al.* (2002) and da-Silva *et al.* (2007) using open population capture-recapture models; da-Silva *et al.* (2007) showed that accounting for heterogeneity in capture probabilities between moderately and highly marked whales improved precision of the survival estimate.

[1] 'Natural' marks include scars resulting from encounters with propellers, bullets and fishing gear as well as ice and killer whales. Since researchers do not capture, mark and release the whales, the term 'capture-recapture' rather than 'mark-recapture' is used in this paper. A naturally marked whale is 'captured' by obtaining a photograph of adequate quality to allow the whale to be categorised as marked during data screening and 'recaptured' when recognised in a subsequent photograph. An 'unmarked' whale is one with a photograph of adequate quality to determine that the screening method does not categorise it as 'marked'.

---

* *LGL Limited, environmental research associates, 22 Fisher St., P.O. Box 280, King City, Ontario, L7B 1A6, Canada.*
+ *University of Washington, Department of Statistics, Box 354322, Seattle, WA, 98195-4322, USA.*
# *National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 7600 Sand Point Way, NE, Seattle, WA, 98115-0070, USA.*
++ *Department of Wildlife Management, North Slope Borough, Box 69, Barrow, AK, 99723, USA.*

It was recognised that continuation of bowhead photography studies would provide information that would allow better definition of life history parameters of bowhead whales as has been done for other species of baleen whales such as right and humpback whales (Barlow and Clapham, 1997; Best *et al.*, 2001; Cooke *et al.*, 2001; Gabriele *et al.*, 2007; Payne *et al.*, 1990). In addition, there are concerns that global warming and unstable shore-fast ice might prevent successful completion of future ice-based surveys. This made it important to determine whether photographic data collected in two consecutive years and analysed using capture-recapture methods could provide adequately precise abundance estimates (i.e. sufficient for use in management as input data for the *Bowhead Strike Limit Algorithm* – e.g. IWC, 2003) to justify replacing ice-based with photographic surveys. The ice-based surveys are dependent on stable ice and weather conditions since they require observers to count whales from perches on the shore-fast ice that are close to leads through which the whales travel. In addition, it is important for the ice-based effort to include hydrophones to record whales that pass beyond viewing range. Snow, persistent fog and shifting ice can lead to failure of an ice-based survey. The aerial photography approach to estimating abundance is less sensitive to vagaries in ice cover but does require weather conditions suitable for conducting flights.

Long gaps between photographic surveys result in less precise estimates and difficulties in analysing data. Thus aerial photographic studies were conducted near Point Barrow, Alaska, during the spring bowhead migration in 2003 (12 April to 6 June) and 2004 (18 April to 7 June) to continue collection of photographs that could be used for better definition of life-history parameters and estimation of abundance. In addition, in connection with investigations of the structure of the B-C-B stock of bowheads, aerial photographic studies were conducted in 2005 in the northern Bering Sea (9 April to 2 May) and near Barrow prior to the main fall migration (6 to 9 September) (Koski *et al.*, 2007).

The 2003 data and data from the earlier spring photographic surveys near Point Barrow were used by Schweder *et al.* (2010) to estimate abundance, population growth rate and mortality. Their approach eliminated the need for data screening by modelling the probability of recognising a recapture as a function of degree of marking of the whale and quality of the images. However, they were not able to obtain capture-recapture estimates of 2003–05 abundance because the 2004 and 2005 data were not yet available and the 2003 data had not been checked for matches with other years.

Koski *et al.* (2008) computed modified Petersen estimates (Chapman, 1951) of the number of naturally marked bowheads for the only two pairs of years when photographic surveys provided adequate numbers of photographic captures and recaptures to support such estimates: 1985–86 and 2003–04. These were preliminary estimates because data from the 2005 surveys were not yet available and checking of data from the earlier surveys was ongoing. In addition, analytical methods were still under development. Koski *et al.* (2008) noted that their abundance estimates were completely independent from ice-based survey estimates used by the International Whaling Commission Scientific Committee (IWC SC) for giving management advice (IWC, 2003). The estimates from the two independent methods agreed well. In this paper, abundance estimates based on the modified Petersen estimate for 1985–86 and 2003–04 using updated data and methods are presented in order to facilitate comparisons with ice-based survey estimates and estimates based on three instead of two years of surveys.

Koski *et al.* (2008) suggested that a more precise estimate of 2004 abundance might be obtained without additional surveys by accounting for heterogeneity in capture probabilities as a function of predictors such as whether whales were highly or only moderately marked (da-Silva *et al.*, 2007; Schweder *et al.*, 2010). They also observed that an estimate based on 2005 as well as 2003–04 data would be more precise. These ideas are pursued in this paper. Methodological improvements have also been made. Most important are refinement of the method for estimating the proportion of whales that are marked and development of a bootstrap approach for assessing precision in addition to the approach based on the delta method. Estimates based on 1989–2005 data are compared with estimates based on 1984–87 data computed using the same methods.

METHODS

Collecting and processing of images

Field and laboratory methods for the pre-2003 surveys (1984–94) have been documented (Angliss *et al.*, 1995; Koski *et al.*, 1992; Rugh *et al.*, 1992a; 1998) and described (Koski *et al.*, 2006). The 2003–05 aerial photographic studies were conducted jointly by LGL Limited (LGL), the North Slope Borough Department of Wildlife Management (NSB-DWM) and the Alaska Fisheries Science Center's National Marine Mammal Laboratory (NMML) with support from the Minerals Management Service (MMS). Field and laboratory methods were similar to those of the earlier studies.

Following each field season, the film was developed, labelled, duplicated and stored in acid-free archive sheets for future analyses. The data documenting each image were entered into an *Excel* spreadsheet for future integration into the 'Bowhead Whale Photography Database' described in Koski *et al.* (2006). Images obtained in 2003–05 were digitised at 4,000 dots per inch; most of the digitised images were cropped and printed to nearly fill 12.7cm × 17.8cm (5in × 7in) colour prints, which are suitable for comparing images to identify matches i.e., recaptures (Rugh *et al.*, 1992a). Printed images were checked against the original film transparencies and the data files to ensure that all images were scanned and printed.

Researchers at LGL and NMML have shared all tasks. NMML researchers have taken the lead on scoring images for photo quality and identifiability (as per Rugh *et al.*, 1998). LGL researchers have taken the lead on within-year matching for the 2003–05 studies, assembling the database, and measuring whales. NMML researchers did within-year matching of images from 2004 for verification of the same effort at LGL. Researchers at both NMML and LGL provided final determination of within-year matches. LGL and NMML researchers independently identified between-year matches. After both groups completed their matching efforts, match results were compared and discussed, and final match determinations were made.

Images were screened using the method of Zeh *et al.*

(2000; 2002) to determine whether they were of acceptable quality for use in capture-recapture analyses and, if so, whether they were of marked or unmarked whales. Quality is scored as 1+ (best), 1–, 2+, 2– or 3 (worst) in each of four zones on the whale's body: rostrum, midback, lower back and flukes. If a zone is scored as 3, it is not acceptable for use in defining the whale as marked for capture-recapture analyses except in the rare cases in which identifiability is scored as H+, H– or M+. Identifiability in each zone is scored as H+ (highly marked), H–, M+ or M– (moderately marked); U+, U– or U (unmarked); or X meaning the zone is not depicted clearly enough in the photo to determine mark status. Scores of X almost always correspond to quality 3. It is assumed that if a zone scored as quality 3 receives an identifiability score of M+ or better, it can be used in defining a whale as marked because that whale would be recognised in a subsequent image of the zone.

In defining the whale (as opposed to the zone) as marked, whales marked in the midback zone are first defined as marked. Then whales with a midback image quality of at least 2+ that were never scored as marked in the midback zone are defined as marked if they are marked on the rostrum. Whales are added to the list of marked whales similarly if they are adequately marked on the flukes or lower back and unmarked in the zones already considered. The end product of the screening process is a list of marked whales that is used in the capture-recapture analysis. This screening method, as well as the natural differences in how well marked individual whales are, leads to heterogeneity in capture probabilities that should be accounted for in analyses (da Silva *et al.*, 2000; da-Silva *et al.*, 2007; Schweder *et al.*, 2010). Covariates created during the screening process can be used to account for heterogeneity under the model used for estimating abundance of the marked population.

Estimating abundance of the 1+ population

An estimate N of bowhead abundance can be computed from photo-identification data using a closed population capture-recapture model to obtain an estimate N^m of the number of naturally marked whales and accounting for unmarked whales by dividing by an estimate p^* of the proportion of the bowhead population that is naturally marked. This abundance estimate is:

$$N = N^m / p^* \qquad (1)$$

See p.72 of Seber (1982) or equation (1) of da Silva *et al.* (2000).

A rough estimate of the variance of N can be derived using the delta method under the assumption that N^m and p^* are statistically independent (Seber, 1982). It can be written as

$$V(N) = V(N^m) / (p^*)^2 + (N^m)^2[V(p^*) / (p^*)^4] \qquad (2)$$

The square root of the right-hand side of equation (2) provides an estimate of the standard error (SE) of N. Calves are not included in computing either N^m or p^*, so N is an estimate of the size of the 1+ (non-calf) population.

Precision of N can also be assessed using a bootstrap procedure. This is not simple given equation (1) because a bootstrap for N^m is based on sampling capture histories of individual marked whales (Buckland and Garthwaite, 1991) while p^* is computed from images of marked and unmarked whales and the effort expended to collect those images. There can be several photographic images of an individual marked whale in a given year and no images in another year included in its capture history. Some images contributing to the capture history of a marked whale are not included in computing p^*, and no images of unmarked whales contribute to the capture histories. The natural sampling unit for a bootstrap on p^* is a survey flight because hours of effort are recorded for each survey flight. While some survey flights produce no images, most produce images of both marked and unmarked whales.

In a given bootstrap replicate, some marked whales with images from a given flight may be represented in the bootstrap sample of capture histories and others may not. If a marked whale with images from a given flight is indicated by the bootstrap sample of capture histories to have been seen in the year of the flight, the flight must be part of the sample of flights. This is because in most cases whales were seen in only one flight in a given year. Although there may be multiple photographs of an individual whale from a single flight, most animals are migrating, so there is a low probability that they will still be in the area during a subsequent flight. If a marked whale photographed on an included flight is not represented in the capture history sample, images of that whale (and a proportional number of unmarked whale images from the flight) must be omitted because the whale cannot both be and not be in the bootstrap replicate. In the next three sections, we describe in more detail how this is accomplished.

Once the bootstrap samples have been defined N^m, p^* and N can be computed for each bootstrap replicate. The standard deviations of the bootstrap values provide standard errors; e.g. if there are *nboot* replicates, the SE of N is given by the standard deviation (SD) of the *nboot* values computed for N.

Estimating abundance of the marked population

The estimate N^m can be obtained using the closed capture model of Huggins (1989; 1991) as implemented in Program MARK (White and Burnham, 1999). The 2003–05 data on marked whales can be treated as representing three sampling occasions (if spring and fall 2005 samples are combined) or four occasions. Recapture probabilities $c(t)$ can be treated as equal to or different from initial capture probabilities $p(t)$, where t denotes the sampling occasion. In initial analyses of the capture-recapture data, spring and fall 2005 were treated as separate sampling occasions, Sp2005 and Fa2005. All 5 recaptures and 49 of the initial captures in 2005 occurred in Sp2005; only 12 initial captures were in Fa2005. When the same models for $p(t)$ and $c(t)$ were fit to the three-occasion and four-occasion data, the estimates N^m of the number of marked whales were generally similar, but N^m from the three-occasion model was somewhat more precise. This is to be expected since the three-occasion model has one less capture probability parameter to estimate than the four-occasion model. Therefore, four-occasion models were not considered further.

Linear or logit models for $p(t)$ and/or $c(t)$ can include covariates that differ among the whales and are expected to influence these capture and/or recapture probabilities, e.g. the identifiability scores that indicate how well marked the whales are. Except in the case of the simplest model discussed below, the parameters that determine $p(t)$ and/or $c(t)$ are estimated via maximum conditional likelihood while N^m is obtained using a method of moments (Huggins, 1989). Either AIC (Akaike, 1974) or the Bayesian Information Criterion (BIC – Schwarz, 1978) can be used in selecting the best model. In this paper, BIC is used because it chooses more parsimonious models. Models with lower BIC explain the data better than those with higher BIC. However, BIC is a function of the maximised likelihood, which involves only $p(t)$, $c(t)$ and any covariates, along with the model assumed for them. Since N^m, the parameter of primary interest, is a derived parameter, a measure of how well the model permits it to be estimated is also needed. Its CV is used for this purpose. An over-parameterised model may produce $p(t)$ and $c(t)$ that fit the data well, but if CV(N^m) is too large, it is not a useful model for our purposes.

The simplest defensible model for the bowhead data, since different numbers of hours of survey effort and different survey conditions characterised the three years, is a model with different values for $p(2003)$, $p(2004)$ and $p(2005)$ with $c(2004) = p(2004)$ and $c(2005) = p(2005)$. This is the model discussed in Chapter 4 of Seber (1982) as the generalised hypergeometric model (Chapman, 1952; Darroch, 1958). White *et al.* (1982) and Buckland and Garthwaite (1991) refer to this model as Model M_t. It will be referred to as Model M_t in this paper. Under this model, Program MARK computes the maximum likelihood estimate N^m as the largest root of the quadratic equation

$$(N^m)^2\,(m_2 + m_3) - N^m\,(n_1n_2 + n_1n_3 + n_2n_3) + n_1n_2n_3 = 0 \quad (3)$$

where n_1 is the number of naturally marked whales photographed in 2003, n_2 the number photographed in 2004, n_3 the number photographed in 2005 and m_2 and m_3 the number of recaptures in 2004 and 2005 respectively. The estimated variance $V(N^m)$ of N^m is computed as in Seber (1982) using an asymptotic variance derived by Darroch (1958):

$$V(N^m) = 1 \,/\, [1/(N^m - r) + 2 \,/\, N^m - 1 \,/\, (N^m - n_1) - 1/(N^m - n_2) - 1 \,/\, (N^m - n_3)] \quad (4)$$

where r is the number of different whales caught during the three sampling occasions. As noted in the previous section, this variance can also be estimated as the variance of *nboot* bootstrap values N^m. However, the variance of N^m is of less interest when the bootstrap is used because the bootstrap provides a direct estimate of the variance of N in place of the function of N^m, p^* and their variances given by equation (2).

Seber (1982) gives an expression for the bias b of N^m from an asymptotic result of Darroch (1958) which for our case of three sampling occasions reduces to

$$b = \{[2 \,/\, N^m - 1 \,/\, (N^m - n_1) - 1 \,/\, (N^m - n_2) -1 \,/\, (N^m - n_3)]^2 + [2 \,/\, (N^m)^2 - 1 \,/\, (N^m - n_1)^2 - 1 \,/\, (N^m - n_2)^2 - 1 \,/\, (N^m - n_3)^2]\} \,/\, \{2[1 \,/\, (N^m - r) + 2 \,/\, N^m - 1 \,/\, (N^m - n_1) - 1 \,/\, (N^m - n_2) - 1 \,/\, (N^m - n_3)]^2\} \quad (5)$$

As already noted, equations (3) and (4) assume that the population of naturally marked bowheads is closed, i.e. the effects of emigration, immigration, mortality and recruitment on the size of the marked population are negligible so that this size can be assumed to be constant over the period during which the data are collected, e.g. 2003–05. This bowhead population has a high survival rate (Zeh *et al.*, 2002), a modest annual rate of increase (George *et al.*, 2004; Zeh and Punt, 2005), a consistent migration pattern that brings it past Point Barrow and into the Beaufort Sea each spring which makes it easy to photograph (Braham *et al.*, 1984; Moore and Reeves, 1993) and stable natural markings that permit the whales to be identified over periods of many years (Koski *et al.*, 1992; Rugh *et al.*, 1992a; 2008; 1992b). It has been shown via simulations based on bowhead photo-identification and natural history data by da Silva *et al.* (2000) that the closed population assumption does not lead to biased estimates over a two-year sampling period in the bowhead case.

Thus the closed population assumption over a three-year sampling period seems reasonable. Comparing abundance estimates based on two-year capture histories, where N^m is the modified Petersen estimate (Chapman, 1951), with those based on three-year capture histories, where N^m is obtained by subtracting the bias given by equation (5) from N^m given by equation (3), provides a check on this assumption. It is important to correct the Model M_t estimate for bias to make the comparison valid because the modified Petersen estimate can be assumed to have negligible bias unless there are fewer than seven recaptures (Robson and Regier, 1964).

If the population continued to increase in 2003–05 as in 1978–2001 (George *et al.*, 2004), the assumed constant abundance would be most representative of 2004. Therefore, the abundance estimate was assigned to that year. Using the same reasoning, the abundance estimates based on 1985–86 and 1984–86 capture histories were assigned to 1985; 1984 was chosen as the additional year for the latter estimate because the number of marked whales successfully photographed in 1984 was similar to the number in 2005.

Bootstrap on capture histories to obtain bootstrap values

Buckland and Garthwaite (1991, pp.257–9), describe how to carry out either a parametric or a nonparametric bootstrap on the capture histories under Model M_t. Capture probability in sample t is estimated by $n_t \,/\, N^m$ where n_t is the number of marked whales actually captured in sample t and N^m is the estimate of the number of marked whales, in our case the estimate given by equation (3) corrected for the bias estimated from equation (5). The probability of each possible capture history, including the capture history of marked whales that were never captured in a photograph, is estimated from the $n_t \,/\, N^m$ values under a multinomial model. Seber (1982) notes that this multinomial model and Model M_t lead to the same maximum likelihood estimates N^m of abundance.

Buckland and Garthwaite (1991) favour a parametric bootstrap carried out by drawing N^m capture histories from the assumed multinomial distribution. However, we needed to draw from the observed capture histories of the marked whales in order to determine which whales were, and which were not, included in each bootstrap sample. An entry 000 for the never captured whales was added to the observed capture histories and sampled with probability $(1 - n_1 \,/\, N^m)(1 - n_2 \,/\, N^m)(1 - n_3 \,/\, N^m)$.

The single marked whale captured in 2003, 2004 and 2005 (capture history 111) was sampled with probability $(n_1 \,/\, N^m)(n_2 \,/\, N^m)(n_3 \,/\, N^m)$. Each other observed capture history was represented by more than one marked whale, so the corresponding multinomial model probability was divided among the whales. For example, eight whales were captured in 2003 and 2004 but not 2005, so each of those whales was sampled with probability $[(n_1 / N^m)\ (n_2 / N^m)\ (1 - n_3 / N^m)] / 8$ to represent capture history 110.

Estimating the proportion of the population that is marked

Koski *et al.* (2008) used data from spring, summer and autumn photographic surveys from 1981 to 2004 to estimate proportion marked. In this paper, data from 1984–87 spring photographic surveys near Point Barrow were used to estimate the proportion of the bowhead population that is naturally marked for 1985 abundance estimates. Data from 1989–92 and 1994 spring surveys as well as the 2003 and 2004 surveys were used for 2004 abundance estimates. These are the most appropriate surveys to use for this purpose because they were designed to sample the entire migrating population. Data from the 2005 surveys cannot be used for estimating proportion marked because those surveys were not conducted during the spring migration near Point Barrow. The Sp2005 survey was designed to sample naturally marked whales, and the Fa2005 survey covered only a few days prior to the main fall migration. Although summer/autumn surveys in earlier years attempted to sample the whole population, they did not always succeed due to age segregation on the summering grounds.

As in Koski *et al.* (2008), separate surveys were used for the 1985 and 2004 abundance estimates so that those estimates would be statistically independent. In this paper, more years were assigned to the 2004 estimate to increase its precision.

The estimate p^* is based on all images with midback quality better than 3 and midback identifiability better than X. The data screening procedure of Zeh *et al.* (2000; 2002) that was used results in the majority of marked whales being marked on their midbacks, and to qualify for the list of marked whales on the basis of marks in another zone, they must be unmarked on their midbacks. Therefore images of the midback zone scored X do not contribute to defining whales as marked or unmarked. The restriction to quality better than 3, not imposed by Koski *et al.* (2008), is to avoid positive bias in p^* due to well marked whales recognisable as marked even in some images of lowest quality.

After the restriction to the images just described, each image was given a weight. That weight was 1.0 for the vast majority of the images. However, following Koski *et al.* (2006), images of cows accompanied by calves were given less weight because of increased effort to photograph cow-calf pairs and the greater amount of time spent at the surface by calves. Cows and yearlings travelling together were given intermediate weight because, like cows with calves, increased effort is made to photograph them, but their surface times are similar to other non-calves. Summing these weights is equivalent to counting the images with each weight, multiplying by the weight and summing the weighted counts. Koski *et al*. (2008) computed p^* as

$$p^* = (\textit{sum of weights for images of marked whales}) / (\textit{sum of weights for all images}) \qquad (6)$$

They used images collected before 1988 to compute p^* for their 1985–86 abundance estimate and those collected after 1988 for the 2003–04 estimate. They used the same cow-calf and cow-yearling weights for both estimates.

In this paper, following Koski *et al.* (2004), time at the surface was estimated from data on surfacing, respiration and diving (SRD) behaviour during the spring migrations of 1989–91 and 1994. Durations of surfacings and dives were recorded for 248 calf SRD cycles and 302 SRD cycles of other whales. Calves were found to spend 1.71 times as long at the surface as other whales, with SE = 0.14 based on 2000 bootstrap replicates. To account for uncertainty in this factor, values were drawn from the bootstrap values used to obtain the SE just cited when a bootstrap analysis was conducted to obtain standard errors for N^m, p^* and N.

To allow for the possibility of changes over time in the extra effort expended to photograph cow-calf pairs, this factor was computed separately from the 1984–87 spring surveys and the 1989–2004 spring surveys. The ratio of images per whale for cows with calves to images per whale for whales not accompanied by a calf or yearling during the part of the migration when calves were seen defines the factor. It was 1.56 in 1984–87 and 1.46 (SE = 0.09) in 1989–2004. The SE of the latter value was estimated via the bootstrap by calculating the ratio from just the eligible images included in each bootstrap sample. A bootstrap analysis was not conducted for estimates obtained from the 1984–87 data because there were many complications to be dealt with, including shifts in migration timing in 1985 (Koski *et al.*, 2006) and 1987. Thus the weights for the 1984–87 calculations were 0.641 = 1/1.56 for cows and yearlings seen together and 0.375 = 1/1.56 × 1/1.71 for cows seen with calves. The corresponding 1989–2004 values were 0.685 = 1/1.46 and 0.401 = 1/1.46 × 1/1.71.

Seventeen whales not accompanied by calves or yearlings that lingered near Point Barrow for three days or more between 19 May and 6 June 2004, a behaviour almost never observed in other years, were omitted from the 1989–2004 calculations described in the previous paragraph. These whales were photographed as many as 17 times on as many as 6 different days, on average 4.99 times as often as other whales not part of a pair. Images of these whales were given weight 0.200 = 1/4.99 in computing p^*.

As in Koski *et al.* (2006), the migration was divided into 'weeks' and the weeks' proportions of marked whales combined to obtain the overall proportion. This approach avoids positive or negative bias in p^* that could result if a week with unusually large numbers of marked whales was oversampled or undersampled, respectively, and (6) was used to compute p^*. The weeks for 1989–2004 are the seven weeks in Koski *et al.* (2006). The more limited sample for 1984–87 required reducing the number of weeks to five by merging the first week with the second and the penultimate with the last. Dates for 1985 were shifted by 9 days as in Koski *et al.* (2006) to account for the late migration that year. Koski *et al.* (2006) did not examine 1987 data because usable lengths were not obtained in 1987. We found that the 1987 migration appeared to be late by about 6 days and shifted its dates accordingly.

Data are available on the number of hours of photographic survey effort for each of the spring surveys near Point

Barrow. The hours of effort for each week were summed over the relevant survey years to obtain $effort_w$ = total hours of effort for week w. It was assumed that if each week had the same amount of effort, the number of images per week would be related to the fraction of the bowhead population migrating past Point Barrow during that week. Under this assumption, p^* can be computed as follows:

$$p^* = [\Sigma(M_w / effort_w)] / [\Sigma(A_w / effort_w)] \qquad (7)$$

where M_w = sum of weights of week w images of marked whales, A_w = sum of weights of all week w images and Σ represents summation over weeks.

A rough estimate $V(p^*)$ of the variance of p^* can be computed under the assumption that M_w follows a binomial distribution with parameters A_w, p_w as

$$V(p^*) = \Sigma W_w^{\,2} p_w (1 - p_w) / A_w \qquad (8)$$

where $p_w = M_w / A_w$ and $W_w = (A_w / effort_w) / \Sigma(A_w / effort_w)$.

The variance of p^* was also estimated as the square of the SD of p^* values computed from the images and effort in bootstrap samples for the 1989–2004 data. For each bootstrap replicate, the sample of capture histories was drawn first as described in the previous section. For each week, flights which obtained images of whales included in the capture history sample were included in the sample of survey flights for the week. Flights which obtained images of marked whales, none of which were included in the capture history sample, were excluded from the sample of survey flights, along with all images obtained on those flights. Among the included flights, if some marked whales photographed were and others were not included in the capture history sample, images of those that were not were excluded. A proportional number of images of unmarked whales chosen at random from the same flight were also excluded in order to keep the proportion marked for images from the flight unchanged. The remaining flights to make up the correct total number of flights for the week were sampled, with replacement, from the flights not already included or excluded, i.e. flights in 2003 and 2004 during which no usable photographs of marked whales were obtained and all flights near Point Barrow during the 1989–92 and 1994 surveys.

Once the sample of flights for each bootstrap replicate was defined, the factor representing the extra effort expended to photograph cow-calf and cow-yearling pairs was computed from the images obtained on the sampled flights as described above and used instead of 1.46. A bootstrap value for calves' extra time at the surface was used instead of 1.71. Then p^* was computed from equation (7) using these bootstrap values and the images of adequate quality and hours of effort from the sampled flights.

Using covariates to account for heterogeneity in capture probabilities

The covariates considered to model differences in capture probabilities among whales were

- *ib* best identifiability score in any of the four zones (b midback, r rostrum, f fluke, l lower back);
- *brfl* zone that defined the whale as marked;
- *zib* best identifiability score in the zone that defined the whale as marked;
- *zqb* best quality score in the zone that defined the whale as marked;
- *nz* number of zones with marks;
- *photos* maximum number of acceptable quality photos of the whale per sampling occasion.

These covariates were considered singly, and all possible pairs were considered. Various codings for each covariate were considered. Covariates were coded to values between 0 and 1 to avoid the need for standardisation within Program MARK. After initial exploratory analyses, we considered only three-occasion models for the 2003–05 data, with recapture probability the same as initial capture probability on each sampling occasion after the first, i.e. $c(t) = p(t)$, $t = 2, 3$. Zeh *et al.* (2002) noted that the assumption $c(t) = p(t)$ is appropriate for photo-identification data because the animals are not physically captured and there is no reason to suppose that the act of photographing a whale should make it more or less likely to be photographed on another occasion as might happen with captured animals. Both linear and logit models were considered. When it was possible to hypothesise which covariate values were expected to lead to higher capture probabilities, the covariate was coded so that its coefficient would likely be positive if the hypothesis was correct. E.g. both *ib* and *zib* were coded with M– = 0, M+ = 0.1, H– = 0.2 and H+ = 0.3 or with M (– or +) = 0 and H (– or +) = 1.

Initial exploratory analyses included attempts to model capture probability as a function of hours of photographic survey effort instead of allowing a different model intercept for each sampling occasion t. These were unsuccessful because of differences in the surveys not reflected in hours of effort. The 2003 survey covered the early part of the migration, when young unmarked whales predominate, more thoroughly than the 2004 survey. Consequently fewer marked whales per hour of effort were captured in 2003 than in 2004. The Sp2005 survey, which accounted for over 80% of the 2005 captures, targeted marked whales. However, high winds reduced the quality of the photos (Koski *et al.*, 2007). Since there were relatively few usable photos, the number of marked whales captured per hour of effort in 2005 was low. Thus all models discussed below include a different intercept for each t.

A covariate coding or a pair of covariates was rejected if its use resulted in a failure by Program MARK to fit the model successfully. In some cases, MARK provided error messages indicating that the model could not be fit. These included 'no numerical convergence', 'numerical convergence suspect', 'beta number x is a singular value' and 'error number x from VA09AD optimisation routine' with x an integer indicating the offending parameter or error. For example, when *nz* was coded 1 = 0.1, 2 = 0.2, 3 = 0.3 and 4 = 0.4 there was no numerical convergence. In other cases, MARK provided no message indicating problems in fitting the model, but output values provided a clear indication of failure. Such output values included SE = 0 for most or all of the estimated parameters, BIC values more than 300 times as large as those from successful fits and estimates N^m smaller than the number of marked whales with capture histories in the data. For example, when *photos* was coded 1 = 0.1, 2 = 0.2, 3 = 0.3, 4 = 0.4 and >4 = 0.5 MARK provided no error message and reported BIC = 955.8, but all parameter estimates had SE = 0.

If none of these obvious failures occurred, BIC and CV(N^m) were used to evaluate covariate codings and models. BIC = $-2 \log(L) + n_{par} \log(n)$, where L is the likelihood of the model for capture probability, n_{par} the number of parameters and n the sample size. MARK computes n as (number of sampling occasions) × (number of marked whales providing capture histories). When no covariates were used, $n_{par} = 3$ since the model included only the intercept parameter for each t. Each coding for each covariate required only a single parameter for the covariate, so $n_{par} = 4$ if a single covariate was used and $n_{par} = 5$ if a pair of covariates was used. CV(N^m) evaluates the additional parameter N^m. For each covariate, it was possible to find at least one coding producing BIC ≤ 973.1 and CV(N^m) ≤ 0.2707. These 'best' codings all came from linear models, so logit models were not considered further. In cases where one coding was better in terms of BIC and another better in terms of CV(N^m), results are reported for both if BIC ≤ 973.1 and CV(N^m) ≤ 0.2707.

Computing confidence intervals

Buckland (1992) was followed in using the method of Burnham *et al*. (1987) to compute confidence intervals (CI). For example, a 95% CI for N is

$$(N / C,\ N \times C),\ \text{where } C = \exp\left[1.96 \sqrt{\log_e(1 + V(N) / N^2)}\right] \tag{9}$$

The percentiles of sorted bootstrap values also provide confidence limits (Buckland and Garthwaite, 1991).

RESULTS

The number of images in the Bowhead Whale Photography Database, the number suitable for use in estimating proportion marked and the number of marked whales from the photographic studies in 1989–92, 1994 and 2003–05 are shown in Table 1 by year. For each of the years 2003–05 used in the capture-recapture analyses in this paper, the number of marked whales identified for the first time in each year and the recaptures are also shown. The analogous 1984–87 data are shown in Table 2. Table 2 is slightly more complicated than Table 1. In Table 1, 2003 is the first year for both the modified Petersen estimate and the Model M_t estimate. In Table 2, 1985 is the first year for the modified Petersen estimate while 1984 is the first year for the Model M_t estimate, so for the latter estimate both 1985 and 1986 have recaptures of whales captured in 1984.

For 2003–05, as shown in Table 1, the method of data screening used in Zeh *et al*. (2000; 2002) produced a sample with $n_1 = 150$ marked whales captured in 2003, $n_2 = 210$ in 2004 and $n_3 = 66$ in 2005, representing 412 different marked whales with capture histories for 2003–05. Among these histories, $m_2 = 9$ whales were recaptured in 2004 and $m_3 = 5$ whales in 2005. The 14 recaptures were of 13 different whales; only one whale was recaptured in both 2004 and 2005.

Different models for capture probabilities and the resulting estimates of the number of marked whales N^m were explored only using the 2003–05 data. Table 3 summarises N^m and its precision obtained from different models, both with and without a covariate characterising individual whales. The estimate N^m from the model without a covariate is given by equation (3). No bias corrections were used for N^m and its CV in Table 3. Equation (5) provided an estimate of bias for N^m given by equation (3), but estimates of bias for the models in Table 3 with a covariate were not available.

The models for capture probabilities in Table 3 were ranked by BIC, so the first model in the table explains capture probabilities best and the last model in the table is the worst based on that criterion. The best model involving each covariate singly was included. All were linear rather than logit models. All had BIC ≤ 973.1 and CV(N^m) ≤ 0.2707. The rank of CV(N^m) is also shown in Table 3, with 1 the best (lowest CV) and 10 the worst of the models shown. For all models in which a pair of covariates was considered, both BIC and CV(N^m) were the same or larger than for the best model involving one of the covariates singly. Consequently, no models with two covariates were included in Table 3.

Recall that covariates were coded such that a positive coefficient would represent the expected result. Coefficients were considered statistically significant (indicated by a Yes under Sig?) if they were significantly different from zero at the 5% level. The *ib* and *zib* coefficients in Table 3 indicate that more highly marked whales are more likely to be captured, though only the *zib* coefficient obtained when *zib* = 0 for moderately marked whales and *zib* = 1 for highly marked whales is statistically significant.

The model best in terms of BIC is worst in terms of CV(N^m), so it is clearly not the best model for the purpose of abundance estimation. That model and the third best model in terms of BIC involve the *photos* covariate. One would expect that having more than one photograph of a whale on a sampling occasion would make it easier to determine if that

Table 1

Numbers of images and marked whales by year and in total used in the 2004 abundance estimates. The 1989–2004 data are from surveys near Point Barrow during the spring migration. The 2005 data are from a spring survey in the Bering Sea and flights in early September near Point Barrow. In 2003–05 initial captures and recaptures that provided the data for estimating the number of marked whales in 2004 are also shown. The modified Petersen estimate used only 2003 and 2004 captures and recaptures while the Model M_t estimate used 2003–05 captures and recaptures. Initial captures and recaptures are not shown for years not used in the capture-recapture analyses to emphasise that matching to determine which whales captured in 2003–05 were first captured before 2003 has not yet been done.

	1989	1990	1991	1992	1994	2003	2004	2005	Total
Number of images	705	677	615	670	283	1,455	1,766	1,081	7,252
Number of images for computing p^*	419	409	402	384	156	967	1,295	0	4,032
Number of marked whales photographed	88	60	69	61	16	150	210	66	720
Initial captures	–	–	–	–	–	150	201	61	412
Recaptures	–	–	–	–	–	0	9	5	14

Table 2

Number of images and marked whales by year and in total used in the 1985 abundance estimates. The images for computing p^* are from surveys near Point Barrow during the spring migration. The numbers of initial captures and recaptures in each of the years used in estimating the number of marked whales in 1985 are also shown. The captures in the first year used in each capture-recapture analysis are treated as initial captures even though some of the whales were captured prior to that year. No captures before the first year used in the analysis are considered in defining whales as recaptured, e.g. eight whales captured in both 1984 and 1985 and four captured in both 1984 and 1986 are treated as initial captures for the modified Petersen estimate, which is based only on 1985 and 1986 capture-recapture data.

	1984	1985	1986	1987	Total
Number of images	1,156	2,788	1,450	403	5,797
Number of images for computing p^*	12	774	508	226	1,520
Number of marked whales photographed	63	254	162	24	503
Initial captures, modified Petersen estimate	–	254	143	–	397
Recaptures, modified Petersen estimate	–	0	19	–	19
Initial captures, Model M_t estimate	63	246	139	–	448
Recaptures, Model M_t estimate	0	8	23	–	31

whale was marked and hence to capture or recapture it. Both codings for *photos* shown in Table 3 were based on that expectation. Nevertheless, *photos* had a significant negative coefficient in both cases. This may be because this covariate represents a property of the sampling occasion rather than the whale. Among the whales captured on only one sampling occasion, 76% of the marked whales captured in Sp2005 had only one photo, compared to 41% to 47% on the other occasions. The negative coefficient apparently allowed for a better model for capture probabilities but resulted in estimates N^m with relatively poor precision.

One would also expect that having more zones marked would increase the probability of capture. However, the relatively large BIC and insignificant negative coefficient of *nz* in Table 3 suggest that it is not a useful covariate. It is less clear how *brfl* should be coded. Its position at the bottom of Table 3 suggests it is not a useful covariate. Its insignificant negative coefficient may reflect the relative ease of obtaining images of the midback compared to the lower back. Similarly, *zqb* does not appear to be a useful covariate; whale identifiability was a better predictor of capture probability than photo quality.

Model M_t with no covariates and N^m obtained from equation (3) had the lowest CV(N^m) and was second best in terms of BIC. Although Table 3 suggests that some of the covariates considered might contribute to a better model when matching of the 2003–05 data with the 1981–2000 data is complete and the full dataset is available, N^m from equation (3) provides the best estimate among those in Table 3. As shown in Table 3, $N^m = 3{,}909$ and SE(N^m) = 993. The bias of N^m from equation (5) is $b = 250$, so the bias-corrected estimate of the abundance of the marked population is 3,659. We correct for the estimated bias so that N based on the Model M_t estimate will be comparable to N based on the unbiased modified Petersen estimate.

The modified Petersen estimate based on the 2003 and 2004 captures and recaptures was 3,185 (SE = 906). It was estimated to be unbiased using an approximation due to Robson and Regier (1964) given by Seber (1982), p. 60. The estimated proportion of the bowhead population that is naturally marked to be used in equation (1) to obtain 2004 abundance estimates is $p^* = 0.28968$. This estimate was computed from equation (7) using data from the 1989–92, 1994, 2003 and 2004 surveys conducted near Point Barrow during spring migration. Using $V(p^*)$ given by equation (8), it was estimated that SE(p^*) = 0.00707. The corresponding values obtained from the 1984–87 data for 1985 abundance estimates are $p^* = 0.33937$ and SE(p^*) = 0.01225. Thus total 1+ abundance in 2004 is estimated as $N = 3{,}185/0.28968 = 10{,}995$ using the modified Petersen estimate and as $N = 3{,}659/0.28968 = 12{,}631$ using the Model M_t estimate. These values of N–with delta method CVs and confidence limits based on equations (2), (8) and (9) and the estimate v^* (Seber, 1982, p. 60) for the modified Petersen estimate or equation (4) for the Model M_t estimate–are shown in Table 4. The corresponding results for 1985 are also shown in

Table 3

Estimates N^m of the number of marked whales for various models for capture probabilities $p(t)$ and recapture probabilities $c(t)$ as a function of sampling occasion and whale-specific covariates (Huggins, 1989; 1991). For each sampling occasion t after the first, it is assumed that $c(t) = p(t)$.

Covariate	N^m	SE(N^m)	CV(N^m)	CV rank	BIC	ΔBIC	Coefficient + or – Sig?	Deviance
Photos 1, 2, >2 coded as 0.1, 0.2, 0.3	4,466	1,209	0.2707	10	955.6	0.0	– Yes	927.2
None	3,909	993	0.2540	1	966.9	11.3	N/A	945.5
Photos >1 versus 1 coded as 1 versus 0	4,128	1,066	0.2582	8	967.9	12.3	– Yes	939.4
zib H versus M coded as 1=H– or H+, 0=M– or M+	4,111	1,063	0.2586	9	968.2	12.6	+ Yes	939.7
ib H versus M coded as 1=H– or H+, 0=M– or M+	4,031	1,032	0.2560	7	969.9	14.3	+ No	941.4
zib M–, M+, H–, H+ coded as 0.0, 0.1, 0.2, 0.3	3,991	1,019	0.2553	4	971.0	15.4	+ No	942.5
nz 1, 2, >2 coded as 0.0, 0.1, 0.2	4,007	1,025	0.2558	6	971.5	15.9	– No	943.0
ib M–, M+, H–, H+ coded as 0.0, 0.1, 0.2, 0.3	3,969	1,012	0.2550	3	971.9	16.3	+ No	943.4
zqb 3, 2–, 2+, 1–, 1+ coded as 0.0, 0.1, 0.2, 0.3, 0.4	3,973	1,015	0.2555	5	972.5	16.9	– No	944.1
brfl b, r, f, 1 coded as 0.1, 0.2, 0.3, 0.4	3,950	1,006	0.2547	2	973.1	17.5	– No	944.6

Table 4

Estimates N of B-C-B bowhead 1+ abundance in 1985 and 2004 with CVs and confidence limits based on the delta method or bootstrap. Estimates that include data from two years are based on the modified Petersen estimate of the number of marked whales. Estimates that include data from three years are based on the Model M_t estimate of the number of marked whales. Confidence limits based on the delta method CV are computed using equation (9). Bootstrap percentile confidence limits are shown when the CV is based on the bootstrap SE. Lower and Upper denote the ends of a 95% confidence interval in either case.

			Confidence limits		
Estimate	N	CV	Lower	Upper	Lower 5th percentile
1985–86 estimate, delta method	6,120	0.1997	4,150	9,020	4,420
1984–86 estimate, delta method	6,129	0.1695	4,410	8,520	4,650
2003–04 estimate, delta method	10,995	0.2855	6,400	19,000	6,900
2003–05 estimate, delta method	12,631	0.2727	7,500	21,300	8,100
2003–05 estimate, bootstrap	12,631	0.2442	7,900	19,700	8,400

Table 5

Means and standard deviations (SD) over 2,000 bootstrap replicates for key parameter estimates used in computing the Table 4 estimate of B-C-B bowhead 1+ abundance in 2004 based on Model M_t with CV and confidence limits estimated via the bootstrap. The bootstrap CV of N is SD(N)/N with N the estimate and SD(N) from this table. The confidence limits in Table 4 are from percentiles of the sorted bootstrap replicate values N.

Parameter	N	N^m	b	p^*
Estimate	12,631	3,909	250	0.28968
Mean over bootstrap replicate values	12,307	3,880	271	0.29345
SD over bootstrap replicate values	3,084	1,081	189	0.00715

Table 4. For both 1985 and 2004, the addition of a third year of data improved precision as measured by the delta method CVs.

Because the delta method estimate of the variance of N given by equation (2) and $V(N^m)$ and $V(p^*)$ given by equations (4) and (8) are only rough estimates based on assumptions that may not hold, variances for N, N^m and p^* were also obtained directly via the bootstrap procedure described above for the estimate of abundance in 2004 based on Model M_t. In order to obtain reliable percentile bootstrap confidence limits, *nboot* = 2,000 bootstrap replicates were used (Buckland and Garthwaite, 1991; da Silva *et al.*, 2000). The means and SDs over the bootstrap replicate values for N, N^m, bias b from equation (5) and p^* from equation (7) are shown in Table 5. Recall that in the absence of bias, the mean over the bootstrap replicate values should be close to the corresponding estimate and the SD gives the SE of the estimate. The resulting bootstrap CV for N and percentile bootstrap confidence limits are given in Table 4.

DISCUSSION

The estimates of total 1+ abundance and measures of their precision in Table 4 are consistent with our expectations concerning bowhead abundance based on completely independent ice-based survey data (George *et al.*, 2004; Zeh and Punt, 2005). George *et al.* (2004) estimated 2001 abundance as 10,470 with SE = 1,351 (CV = 0.129) and 95% CI 8,100 to 13,500. They estimated the annual rate of increase for the population from 1978 to 2001 as 3.4% (95% CI 1.7% to 5%). The estimate of Zeh and Punt (2005) was 10,545 (CV = 0.128) for 2001. If the trend line fit by George *et al.* (2004) is projected forward, the expected abundance in 2004 is 11,811; the point on the trend line for 1985 is 6,295. The 2004 abundance estimates in Table 4 are about 800 whales away from 11,811 and their average is 11,813. The 1985 estimates in Table 4 are within 175 whales of the trend line value. In other words, all the estimates N in Table 4 are considerably closer to the values expected based on the ice-based surveys than the CVs in Table 4 indicate they might be.

The CVs in Table 4 are higher than CVs for the 1988, 1993 and 2001 abundance estimates from ice-based surveys in Table 4 of Zeh and Punt (2005). Those ice-based surveys had more comprehensive acoustic monitoring of whales that passed too far offshore to be seen than the earlier surveys, leading to improved precision. However, our Table 4 CVs are comparable to or lower than their Table 4 CVs for the ice-based survey estimates obtained before 1988. E.g. the 1985 estimate in Table 4 of Zeh and Punt (2005) has a CV of 0.253, and the remaining pre-1988 CVs range from 0.215 to 0.345. The 1985 estimate based on Model M_t in our Table 4 has a CV of 0.1695.

The CVs in our Table 4 that are lower than the pre-1988 ice-based survey CVs are those of the 1985 estimates N. This is because, as can be seen by comparing Table 2 with Table 1, there were more initial captures and recaptures for both the Petersen estimate and the Model M_t estimate N^m for 1985 than for 2004. There were no summer surveys in 2003–05 and only one brief September survey. This contrasts with 1984–86, when summer and autumn photographic surveys in the Beaufort Sea in addition to spring surveys were conducted. The 1985 surveys were particularly comprehensive. Images from these summer and autumn surveys provided many of the initial captures and recaptures of marked whales for the 1985 estimates. There were many fewer images available from spring surveys near Point Barrow for computing p^* for the 1985 estimates, but this had relatively little impact on their precision.

Since the 1985 estimates in Table 4 are the most precise, they are the most useful for comparing the estimates based on the modified Petersen estimate computed from two years of capture-recapture data with those from Model M_t based on three years of capture-recapture data. This comparison is of interest as a check on whether the closed population assumption on which both estimates are based is acceptable over three years for bowheads. The very close agreement between the two values of N suggests that it is.

Recall that the bootstrap analysis was based on the assumption that the true number of marked whales in 2004 was 3,659. This is N^m = 3,909 from Model M_t, corrected for bias b = 250. The corresponding means over the bootstrap replicate values in Table 5 are N^m = 3,880 and b = 271, both quite close to the expected values. SD(b) in Table 5 indicates high variability of the estimated bias over the bootstrap replicates; b is always estimated to be positive as expected, ranging from 51 to 2,279. The mean of p^* over the bootstrap replicate values is 0.29345, close to the estimate p^* = 0.28968. The mean over the bootstrap replicate values of N in Table 5 is 12,307, very close to (3,880 – 271) / 0.29345 = 12,299 and reasonably close to the estimate N = 12,631 in Table 4. None of the Table 5 means suggest problems with the bootstrap analysis. It is interesting that the median of the bootstrap replicate values of N is 11,767. This is very close to the value of 11,811 in 2004 predicted from the trend in the ice-based survey estimates. From Table 5 the estimate SE(N) = 3,084 is obtained. It is similar to the delta method SE of 3,444 computed from equation (2). Taking the above discussion into account, we recommend using the bootstrap CV and percentile method confidence limits in Table 4 for the estimate N = 12,631 of total 1+ abundance in 2004.

Using a model for heterogeneity in capture probabilities under which whales highly marked on the midback were more likely to be captured than those only moderately marked, da Silva *et al.* (2000) showed via simulation that a capture-recapture estimate of abundance that does not account for heterogeneity in capture probabilities when present can be slightly negatively biased. An estimate that accounts for the heterogeneity may be slightly positively biased. However, the biases were small compared to the SE of the estimates. The estimate that accounted for heterogeneity was slightly more precise. The simulated data from which these results were obtained had many more captures and recaptures than even the 1984–86 data shown in Table 2.

The estimates N^m in Table 3 based on assuming and attempting to account for heterogeneity in capture probabilities are somewhat larger than the estimate that assumes homogeneity. This is consistent with the bias results of da Silva *et al.* (2000). However, the estimates that account for heterogeneity are slightly less precise than the one that does not. Recall that da-Silva *et al.* (2007) showed that accounting for heterogeneity in capture probabilities between moderately and highly marked whales improved precision of the estimate of annual survival probability using a much larger dataset. More captures and recaptures than are available in the 2003–05 dataset are needed to assess covariate effects on capture probabilities and abundance estimates. It may be necessary to use an open population model (da-Silva *et al.*, 2007; Schweder *et al.*, 2010) and many more than three years of data to obtain an adequate number of captures and recaptures.

ACKNOWLEDGEMENTS

Many people have assisted in the acquisition of photographs and data that are present in the current bowhead whale photographic database. People who have contributed significantly to several of the studies include Robyn Angliss, Lisa Baraff, Howard Braham, Jim Cubbage, Rolph Davis, Bob Evans, Gary Miller, Mary Nerini, John Richardson, Kim Shelden, Dave Withrow and Bernd Würsig. John Brandon of NMML, Steve McLean of LGL and Devin Bates of the NSB-DWM were primary observers and data recorders during the 2003–04 aerial surveys. Darren Ireland and Meike Holst, of LGL, Rita Acker, Paula Earp, Crystal Pike and several other members of the NSB-DWM assisted as observers. Steve McLean was the photographer for 3 days (2–4 May 2003) when WRK was not available. Charles Pike, Marco Colella and Ralph Aiken of Commander Northwest piloted the aircraft. The staff at the NSB-DWM, including Directors Taqulik Hepa and Charles Brower, were particularly helpful and supportive of the studies.

We especially thank the Whaling Captains Association at Barrow for allowing us to conduct these studies near their village during the whaling seasons. The majority of the funding was provided through the NSB-DWM via Department funding; NOAA Coastal Impact Assessment Program and US National Marine Fisheries Service (NMFS) through a special NOAA appropriation for bowhead whale stock structure analyses. MMS and BP Exploration, Alaska, Inc. provided financial support for analysis of the 2003–04 data.

We thank Cibele da-Silva, Rolph Davis, Jeff Breiwick, Devin Johnson, Doug Butterworth and Andre Punt for reviewing and providing comments that improved the paper.

This paper is dedicated to two people who were major contributors to past photography studies. Gary Miller died in 2004 after a brief struggle with cancer. Gary contributed to all aspects of our photography studies and conducted the vast majority of matching of images collected from 1981 to 2000. Ralph Aiken was lost to an aircraft accident in Alaska in 2006. Ralph made a tremendous contribution to bowhead whale and arctic marine mammal science as he safely flew us thousands of kilometres over the ocean for various research programs. We will miss them.

Photographic surveys were conducted under Scientific Research Permits 580, 791, 782–1438 and 782–1791 (NMML) and 670 (LGL) issued by NMFS under the provisions of the US Marine Mammal Protection Act and US Endangered Species Act.

REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6): 716–23.

Angliss, R.P., Rugh, D.J., Withrow, D.E. and Hobbs, R.C. 1995. Evaluations of aerial photogrammetric length measurements of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*). *Rep. int. Whal. Commn* 45: 313–24.

Barlow, J. and Clapham, P.J. 1997. A new birth-interval approach to estimating demographic parameters of humpback whales. *Ecology* 78(2): 535–46.

Best, P.B., Brandão, A. and Butterworth, D.S. 2001. Demographic parameters of southern right whales off South Africa. *J. Cetacean Res. Manage.* (special issue) 2: 161–69.

Braham, H.W., Krogman, B.D. and Carroll, G.M. 1984. Bowhead and white whale migration, distribution and abundance in the Bering, Chukchi and Beaufort Seas, 1975–78. *NOAA Technical Report NMFS* SSRF-778: 39pp. NTIS PB84-157908. [Available from: *www.nmfs.gov*].

Buckland, S.T. 1992. Report of the Scientific Committee, Annex H. Proposal for standard presentation of abundance estimates. *Rep. int. Whal. Commn* 42: 235.

Buckland, S.T. and Garthwaite, P.H. 1991. Quantifying precision of mark-recapture estimates using the bootstrap and related methods. *Biometrics* 47: 255–68.

Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C. and Pollock, K.H. 1987. *Design and Analysis Methods for Fish Survival Experiments Based on Release-Recapture*. *American Fisheries Society Monographs* No. 5. American Fisheries Society, Bethesda, Maryland. x+437pp.

Chapman, D.G. 1951. Some properties of the hypergeometric distribution with applications to zoological censuses. *Univ. Calif. Public. Stat.* 1: 131–60.

Chapman, D.G. 1952. Inverse, multiple and sequential sample censuses. *Biometrics* 8: 286–306.

Cooke, J.G., Rowntree, V.J. and Payne, R.S. 2001. Estimates of demographic parameters for southern right whales (*Eubalaena australis*) observed off Península Valdés, Argentina. *J. Cetacean Res. Manage.* (special issue) 2: 125–32.

Darroch, J.N. 1958. The multiple-recapture census. I: Estimation of a closed population. *Biometrika* 45: 343–59.

da Silva, C.Q., Zeh, J., Madigan, D., Laake, J., Rugh, D., Baraff, L., Koski, W. and Miller, G. 2000. Capture-recapture estimation of bowhead whale population size using photo-identification data. *J. Cetacean Res. Manage.* 2(1): 45–61.

da-Silva, C.Q. 2003. A comparison of two models for estimating the size of a population containing uncatchable individuals. *Brazil. J. Probability Stat.* 17: 17–40.

da-Silva, C.Q., Gomes, P.V.S. and Stradioto, M.A. 2007. Bayesian estimation of survival and capture probabilities using logit link and photoidentification data. *Computational Stat. Data Anal.* 51: 6521–34.

da-Silva, C.Q., Rodrigues, J., Leite, J.G. and Milan, L.A. 2003. Bayesian estimation of the size of a closed population using photo-id data with part of the population uncatchable. *Commun. Stat. Simulation Computation* 32(3): 677–96.

da-Silva, C.Q. and Tiburcio, J.D. 2010. Empirical Bayes estimation of the size of a closed population using photo-identification data. *J. Cetacean Res. Manage.* 11(1): 23–30.

Gabriele, C.M., Straley, J.M. and Neilson, J.L. 2007. Age at first calving of female humpback whales in southeastern Alaska. *Mar. Mamm. Sci.* 23(1): 226–39.

George, J.C., Zeh, J., Suydam, R. and Clark, C. 2004. Abundance and population trend (1978–2001) of western Arctic bowhead whales surveyed near Barrow, Alaska. *Mar. Mamm. Sci.* 20(4): 755–73.

Huggins, R.M. 1989. On the statistical analysis of capture experiments. *Biometrika* 76: 133–40.

Huggins, R.M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* 47: 725–32.

International Whaling Commission. 2003. Report of the Scientific Committee. Annex E. Report of the Standing Working Group on the Development of an Aboriginal Subsistence Whaling Management Procedure (AWMP). *J. Cetacean Res. Manage.* (Suppl.) 5:154–255.

Koski, W.R., Davis, R.A., Miller, G.W. and Withrow, D.E. 1992. Growth rates of bowhead whales as determined from low-level aerial photogrammetry. *Rep. int. Whal. Commn* 42: 491–99.

Koski, W.R., Davis, R.A., Miller, G.W. and Withrow, D.E. 1993. Reproduction. pp.239–74. *In*: Burns, J.J., Montague, J.J. and Cowles, C.J. (eds). *The Bowhead Whale.* Special Publication No. 2. The Society for Marine Mammalogy, Lawrence, Kansas. 787pp.

Koski, W.R., Miller, G.W., Richardson, W.J. and Würsig, B. 2004. Bowhead whale (*Balaena mysticetus*) mothers and calves during spring migration in the Alaskan Beaufort Sea: Movements, behaviour and life history data. Paper SC/56/BRG27 presented to the IWC Scientific Committee, July 2004, Sorrento, Italy (unpublished). 31pp. [Paper available from the Office of this Journal].

Koski, W.R., Mocklin, J., Davis, A.R., Zeh, J., Rugh, D.J., George, J.C. and Suydam, R. 2008. Preliminary estimates of 2003–2004 Bering-Chukchi-Beaufort bowhead whale (*Balaena mysticetus*) abundance from photo-identification data. Paper SC/60/BRG18 presented to the IWC Scientific Committee, June 2008, Santiago, Chile (unpublished). 7pp. [Paper available from the Office of this Journal].

Koski, W.R., Rugh, D.J., Punt, A.E. and Zeh, J. 2006. An approach to minimise bias in estimation of the length-frequency distribution of bowhead whales (*Balaena mysticetus*) from aerial photogrammetric data. *J. Cetacean Res. Manage.* 8(1): 45–54.

Koski, W.R., Rugh, D.J., Zeh, J.E., George, J.C., Suydam, R.S., Davis, A.R., Mocklin, J. and Trask, K. 2007. A review of bowhead aerial photographic studies in 2003–06. Paper SC/59/BRG6 presented to the IWC Scientific Committee, May 2007, Anchorage, USA (unpublished). 6pp. [Paper available from the Office of this Journal].

Miller, G.W., Davis, R.A., Koski, W.R., Crone, M.J., Rugh, D.J., Withrow, D.E. and Fraker, M.A. 1992. Calving intervals of bowhead whales – an analysis of photographic data. *Rep. int. Whal. Commn* 42: 501–06.

Moore, S.E. and Reeves, R.R. 1993. Distribution and movement. pp.313–86. *In*: Burns, J.J., Montague, J.J. and Cowles, C.J. (eds). *The Bowhead Whale*. Special Publication No. 2. Society for Marine Mammalogy, Lawrence, KS. 787pp.

Nerini, M.K., Braham, H.W., Marquette, W.M. and Rugh, D.J. 1984. Life history of the bowhead whale, *Balaena mysticetus,* (Mammalia, Cetacea). *J. Zool. (Lond.)* 204: 443–68.

Payne, R., Rowntree, V., Perkins, J.S., Cooke, J.G. and Lankester, K. 1990. Population size, trends and reproductive parameters of right whales (*Eubalaena australis*) off Peninsula Valdes, Argentina. *Rep. int. Whal. Commn* (special issue) 12: 271–78.

Robson, D.S. and Regier, H.A. 1964. Sample size in Petersen mark-recapture experiments. *Trans. Amer. Fish. Soc.* 93: 215–26.

Rugh, D.J., Braham, H.W. and Miller, G.W. 1992a. Methods for photographic identification of bowhead whales, *Balaena mysticetus*. *Can. J. Zool.* 70: 617–24.

Rugh, D.J., Koski, W.R., George, J.C. and Zeh, J. 2008. Interyear re-identification of bowhead whales during their spring migration past Point Barrow, Alaska, 1984–1994. *J. Cetacean Res. Manage.* 10(3): 195–200.

Rugh, D.J., Miller, G.W., Withrow, D.E. and Koski, W.R. 1992b. Calving intervals of bowhead whales established through photographic identifications. *J. Mammal.* 73(3): 487–90.

Rugh, D.J., Zeh, J.E., Koski, W.R., Baraff, L.S., Miller, G.W. and Shelden, K.E.W. 1998. An improved system for scoring photo quality and whale identifiability in aerial photographs of bowhead whales. *Rep. int. Whal. Commn* 48: 501–12.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–4.

Schweder, T. 2003. Abundance estimation from multiple photo surveys: confidence distributions and reduced likelihoods for bowhead whales off Alaska. *Biometrics* 59: 974–83.

Schweder, T., Sadykova, D., Rugh, D. and Koski, W. 2010. Population estimates from aerial photographic surveys of naturally and variably marked bowhead whales. *JABES* 15(1): 1–19.

Seber, G.A.F. 1982. *The Estimation of Animal Abundance and Related Parameters*. 2nd ed. Charles Griffin and Company Ltd., London. i–xvii+654pp.

White, G.C., Anderson, D.R., Burnham, K.P. and Otis, D.L. 1982. *Capture-Recapture and Removal Methods for Sampling Closed Populations*. Los Alamos National Laboratory, Los Alamos. xvi+235pp. [Available from NTIS, 5285 Port Royal Road, Springfield, VA 22161].

White, G.C. and Burnham, K.P. 1999. Program MARK: Survival estimation from populations of marked animals. *Bird Study* 46(Supplement): 120–38.

Zeh, J., Poole, D., Miller, G., Koski, W., Baraff, L. and Rugh, D. 2000. Survival of bowhead whales, *Balaena mysticetus*, estimated from 1981–98 photo-identification data. Paper SC/52/AS19 presented to the IWC Scientific Committee, June 2000, in Adelaide, Australia (unpublished). 10pp. [Paper available from the Office of this Journal].

Zeh, J., Poole, D., Miller, G., Koski, W., Baraff, L. and Rugh, D. 2002. Survival of bowhead whales, *Balaena mysticetus*, estimated from 1981–98 photo-identification data. *Biometrics* 58: 832–40.

Zeh, J.E., Clark, C.W., George, J.C., Withrow, D., Carroll, G.M. and Koski, W.R. 1993. Current population size and dynamics. pp.409–89. *In*: Burns, J.J., Montague, J.J. and Cowles, C.J. (eds). *The Bowhead Whale*. Special Publication No. 2. The Society for Marine Mammalogy, Lawrence, Kansas. 787pp.

Zeh, J.E. and Punt, A.E. 2005. Updated 1978–2001 abundance estimates and their correlations for the Bering-Chuckchi-Beaufort Seas stock of bowhead whales. *J. Cetacean Res. Manage* 7(2): 169–75.

Date received: December 2009.
Date accepted: January 2010.