Brazilian Journal of Probability and Statistics (2003), 17, pp. 17–40.
©Associação Brasileira de Estatística

A comparison of two models for estimating the size of a population containing uncatchable individuals

Cibele Queiroz da-Silva
Universidade Federal de Minas Gerais

Abstract: One of the assumptions for the majority of capture-recapture models is that all marked animals in the population have a unique mark. However, there are some populations of individuals to whom this assumption may be seriously violated. It may be the case for capture-recapture studies using photo-identification data where the individuals are identified through their natural marks. It may happen that the extent of marks does not allow for a unique identification or that there are many individuals without marks at all. Thus, such individuals are considered uncatchable and classical capture-recapture models are not adequate to estimate the size of the whole population. Working with bowhead whale *Balaena mysticetus* data, da Silva (1999) and da Silva *et al.* (2000) dealt with this problem and proposed models that account for the uncatchable part of the population. The proposed models incorporate some assumptions that may rise some criticism. So the aim of the present work is to evaluate in some extent how badly such assumptions may affect inferences about N. In this work we compare the simplest of the models suggested by da Silva *et al.* (2000) (multinomial model) with an alternative one which incorporates some more realistic assumptions. We concluded that the multinomial model is as good as the alternative model in the estimation of the population size N. Moreover, the multinomial model produces more precise estimates and it is much simpler and less computationally demanding than the alternative model.

Key words: Capture-recapture; EM algorithm; photo-identification.

1 Introduction

Capture-recapture methods for closed populations are useful for estimating the size N of an animal population. In its simplest form two subsequent samples of sizes n_1 and n_2 are taken from a population (single capture-release). In the first sample all the n_1 captured animals are uniquelly marked (using, for example, tags, tattoos, numbered rings, natural marks) and then they are returned to the population. After allowing some time for those animals to mix with the rest of the population, the second sample of n_2 animals is taken and the number m of marked individuals is recorded. If capture probability does not change from sample to sample, then m/n_2 estimates the proportion n_1/N of marked individuals in the population. Thus, by equating these two quantities the so called Petersen-Lincoln (see Petersen, 1896 and Lincoln, 1930) estimator of N, $\hat{N} = (n_1 n_2)/m$, is obtained. The more widely known model extensions from single to multiple capture-release were developed by Schnabel (1938), Chapman (1951) and Darroch

(1958). Multiple capture-release data consist of a sample of binary time series, the so called capture history of the individuals. For each individual i, $i = 1, \ldots, N$, captured at least once, a row in the data set is created and for the column corresponding to sampling occasion j, $j = 1, \ldots, t$, it is assigned value 1 for captures or recaptures or value 0 otherwise. However it is not necessary to have available the information about the individual capture histories, since the estimators may be based on a joint sufficient statistics such as the sample sizes and the number of distinct individuals observed over the experiment, etc. Thus, it is only necessary to know the observed values for those quantities in order to estimate population size. There are some capture-recapture softwares (like CAPTURE and MARK) that have a minimum sufficient statistic data entry option for this purpose.

One of the assumptions for the majority of capture-recapture models is that all marked animals in the population have a unique mark. By doing so an individual would not be counted more than once in the analysis. However, there are some populations of individuals to whom this assumption may be seriously violated. It may be the case for capture-recapture studies using photo-identification data where the individuals are identified through their natural marks. It may happen that the extent of marks does not allow for a unique identification or that there are many individuals without marks at all. Thus, such individuals are considered uncatchable and classical capture-recapture models are not adequate to estimate the size of the whole population. Williams *et al.* (1993) proposed a simple and intuitively appealing approach to estimate N under those circumstances, but it can be criticised since they used photographs of the same study to estimate quantities that were supposed to be estimated independently. That has an impact in the precision of the estimator of N, $\hat{N}$. Working with bowhead whale *Balaena mysticetus* photo-identification data, da Silva (1999) and da Silva *et al.* (2000) dealt with this problem and proposed models that account for the uncatchable part of the population. Confidence intervals for N were built using parametric bootstrap (Efron and Tibshirani, 1993). The proposed models incorporate some assumptions that may rise some criticism. The aim of the present work is to evaluate, in some extent, how badly such assumptions may affect inferences about N.

In this work we compare the simplest of the models suggested by da Silva *et al.* (2000), which we will call *model 1*, with an alternative model, *model 2*, which incorporates some more realistic assumptions. In Section 2 we present a brief description about real and simulated bowhead whale photo-id data. In Section 3 we describe *model 1*. In Section 4 we present *model 2*. An EM-algorithm procedure used in the estimation of parameters of interest is discussed in Section 5. In Section 6 we describe bootstrap procedures used to evaluate the precision of $\hat{N}$ and to construct confidence intervals for N. Finally, results and discussion are presented in Sections 7 and 8.

2 Actual and simulated bowhead data

The Bering-Chukchi-Beaufort Seas stock of bowhead whales spends summers in the Canadian Beaufort Sea, migrates in fall along the northern Alaskan Beaufort Sea coast into the Chukchi Sea near Siberia, then through the Bering Strait into the Bearing Sea. Braham *et al.* (1980) believe that the majority of the stock passes Point Barrow, Alaska, during the migration. Their narrow migration corridor and consistent timing of migration make Point Barrow an ideal area to sample the population (Rugh, 1990).

During the surveys an aircraft flies a systematic transect until whales are encountered. Medium-format (70 mm) cameras are used due to their good features for low-altitude aerial photographs. Multiple photographs are taken with an automatic film advance as the aircraft flies directly above the whales as to increase the chances of photographing marks.

Aerial photographs of bowhead whales suitable for identification of individuals using their natural markings have been collected in the Bering, Chukchi and Beaufort Seas since 1976. Most of the photographs have been collected by LGL Ltd. (LGL), the National Marine Mammal Laboratory (NMML) and Cascadia Research Collective (CRC). The collections are housed at LGL and NMML.

The use of capture-recapture methods to estimate whale population parameters from these data was not envisaged in the early years of the studies, so a single score that combined quality and identifiability was assigned to each photo. Inadequate quality screening of photographs can lead to violations of assumptions required for capture-recapture estimation, so Rugh *et al.* (1998) developed a revised scoring system for aerial photos of bowhead whales that addressed these problems. Since the largest available samples of photographs were collected in 1985 and 1986, Rugh (1990) and Whitcher *et al.* (1996) both suggested basing a population estimate on those years data sets. According to Whitcher *et al.* (1996), the only years in which 5% or more of the whales are estimated to have been photographed and identified are 1985 and 1986, and these are also the years with the largest number of photographs with appropriate quality scores. The 1985 and 1986 photographs, re-scored using the revised system, provided the data for the estimates of Section 7.5.

The use of simulated bowhead data was necessary to make possible the study of the statistical properties of the estimated parameters for the proposed models. Simulation of bowhead whale capture-recapture data was possible because a great deal is known about the bowhead whale population. In the simulation process were considered aspects such as age composition of the bowheads, survival rate and maturity, proportions of marked and unmarked whales in the population, samples sizes, population additions and deletions and number of photos of the whales. Values of the parameters used in the simulation were gathered from several sources, including papers and monographs describing the photographic surveys conducted by NMML and LGL. Important sources of information were summarized by Givens (1993) and Raftery *et al.* (1995), who present a Bayesian synthesis approach for making inferences about characteristics of interest given different sources of information which are linked by a deterministic population dynamics model. A number of those parameters used in the simulations came from the most likely population trajectory in a Bayesian synthesis analysis carrried out by Givens (pers. comm.).

3 Model 1

In this section we introduce some notation and restate the simplest of the models suggested by da Silva (1999) and da Silva *et al.* (2000). All the notation is written as to be used for bowhead whale photo-id data, but the model is rather general and it can be used for estimating the size of any closed population which presents characteristics analogous to the bowheads like the presence of animals in the population which are photographically uncatchable. Next we give further details about the bowheads surveys.

During the surveys, the aircraft flew a systematic transect until whales were encountered. Then multiple photographs were taken when directly above the whale or group of whales. In many cases, the plane circled above the whales until the photographers were confident that all had been photographed at least once. Therefore, the multinomial trials of assigning photos to whales are not independent because multiple photos are taken of the same whale under similar conditions. The violation of the independence assumption in the case of the whales may be fairly mild though, since for real bowhead whale data the average number of photos per whale is around 1.5, justifying the model used in (1).

We will denote as *good photos* the ones that were given a high score when submitted to a scoring system. They represent a category of high quality photos of marked and unmarked individuals. It is not possible to extract any information about the capture history of the individuals from bad quality photos. Thus, the models that will be presented will be conditioned on the good photos. However, we do not underestimate the usefulness of the bad photos to improve inferences about N. Actually, we are presently conducting some research on this issue.

In *model 1* capture probabilities vary over time, but not among individuals. The basic assumptions of this model are:

- The population remains constant throughout the experiment.
- In a given sample, all the animals have the same propensity of being caught, regardless of their previous capture history.
- The catch propensity may vary over time according to the effort spent to capture the individuals.
- All marked animals in good photos are correctly identified.

The capture history of the unmarked whales cannot be identified. Therefore, capture-recapture methods have been used to estimate only the identifiable part of the population. Such procedure is unsatisfactory for bowheads because the unmarked individuals constitute the majority of the population.

The primary aim of the model is to present a methodology for the estimation of the total population size of the bowhead whales when only part of the population is catchable. Contrary to most of the models for capture-recapture, we will make an extensive use of the information given by the photos of the unmarked individuals to help us estimate the parameters of interest.

3.1 Notation

- X_i^u is the number of good photos of unmarked whales taken at time i;
- X_i^m is the number of good photos of marked whales taken at time i;
- X_i is the total number of good photos taken at time i;
- n_i^m is the number of marked whales captured at time i;
- r is the total number of distinct marked whales captured over the study;
- $\{u_w\}$ is a set which includes the number of individuals with capture history w where w is a subset of the integers 1 to t designating the occasions of capture.
- U represents the event "selecting an unmarked individual in the population".
- M represents the event "selecting a marked individual in the population".
- G_j represents the event "taking a good photo of an encountered individual at time j".

The following relation is observed:

$$X_i = X_i^u + X_i^m.$$

The parameters in the model are:

- $N = N^m + N^u$ is the total number of individuals in the population;
- N^u is the total number of unmarked individuals in the population;
- N^m is the total number of marked individuals in the population;
- p_i is the probability that a given whale is photographed at sampling occasion i;
- $q_i = 1 - p_i$.

Let us assume that the probability of getting a good photo in occasion j is the same for marked and unmarked individuals, i.e., $P(G_j \mid M) = P(G_j \mid U)$. That is a reasonable assumption since it is in agreement with the analyst's procedure protocol to attach rank quality to a photo. The analyst gives a grade to each photo based only on photo quality, not on a whale's extent of marks. Thus the probability of taking a good photo in occasion j, $P(G_j)$, equals the probabilities above. Therefore, given that a good photo is taken at time j, the probability of it being from a marked individual is

$$P(M \mid G_j) = \frac{N^m}{N},\ j = 1, \ldots, t.$$

Thus the distribution of X_i^m is assumed to be binomial with parameters

$$\left(X_i, \ \ \frac{N^m}{N} \right),$$

and the distribution of $\{u_w\}$ is multinomial, given by Darroch's model (1958).

The reason why we used the binomial model above instead of a hypergeometric one with parameters N, N^m and X_i to represent the conditional distribution of X_i^m given X_i is that N and N^m are, in general, large compared to X_i, justifying a binomial approximation to hypergeometric distribution.

The joint distribution of X_i^m and $\{u_w\}$ is,

$$\begin{aligned}
P(\{u_w\}, \{X_i^m\}) &= P(\{X_i^m\} \mid \{u_w\}) P(\{u_w\}) \\
&= \left[\prod_{i=1}^{t} P\left(X_i^m \mid \sum_{w \supset i} u_w \right) \right] P(\{u_w\}) \\
&= \left[\prod_{i=1}^{t} P\left(X_i^m \mid n_i^m \right) \right] P(\{u_w\}).
\end{aligned}$$

The distribution of $(X_i^m \mid n_i^m)$ is truncated binomial because $X_i^m \geq n_i^m$.

3.2 The model

Now considering that the estimation of N^u and N^m is restricted by the relation

$$N = N^u + N^m,$$

it is natural to write N^u as being proportional to N^m (according to a proportionality constant γ, say) and write

$$N^u = \gamma N^m,$$

therefore,

$$N = N^m(1 + \gamma).$$

Further development of the model is simplified by the relation above. Now, the distribution of $(X_i^m \mid n_i^m)$ is expressed by:

$$P(X_i^m \mid n_i^m) = \frac{\binom{X_i}{X_i^m} \left[\frac{1}{1+\gamma}\right]^{X_i^m} \left[\frac{\gamma}{1+\gamma}\right]^{X_i^u} 1\{(X_i^u, X_i^m) \in B\}}{\sum_{j=n_i^m}^{X_i} \binom{X_i}{j} \left[\frac{1}{1+\gamma}\right]^{j} \left[\frac{\gamma}{1+\gamma}\right]^{X_i - j}} \tag{3.1}$$

where $B = \{(X_i^u, X_i^m) : n_i^m \le X_i^m \le X_i\}$ and $1\{(X_i^u, X_i^m) \in B\}$ is an indicator function for pairs (X_i^u, X_i^m) belonging to the set B.

3.3 The likelihood function

Conditionally on the good photos, the likelihood function is given by the conditioned joint distribution of X_i^m, $\{u_w\}$. Notice that $n_i = \sum_{w \supset i} u_w$, then it is not necessary to include n_i as a random variable in our model since $\{u_w\}$ is already there. Similar comments apply to equation (6) presented in Section 4.1.

$$
\begin{aligned}
L &= \{\prod_{i=1}^{t} P\left(X_i^m \mid X_i\right)\} P(\{u_w\}) \\
&= \prod_{i=1}^{t} \left[\frac{\binom{X_i}{X_i^m} \left[\frac{1}{1+\gamma}\right]^{X_i^m} \left[\frac{\gamma}{1+\gamma}\right]^{X_i^u} 1\{(X_i^u, X_i^m) \in B\}}{\sum_{j=n_i^m}^{X_i} \binom{X_i}{j} \left[\frac{1}{1+\gamma}\right]^{j} \left[\frac{\gamma}{1+\gamma}\right]^{X_i - j}} \right] \\
&\quad \times \frac{N^m!}{(N^m - r)! \prod_w u_w!} \prod_{i=1}^{t} p_i^{n_i^m} {q_i}^{N^m - n_i^m}.
\end{aligned}
$$

The maximum likelihood estimators of the capture probabilities are given by $\hat{p}_i = \frac{n_i^m}{\hat{N}^m}$, $i = 1, \cdots, t$, where the maximum likelihood estimator of N^m, $\hat{N}^m$, is a value which satisfies equation (2) (solved numerically).

$$
\prod_{i=1}^{t} (N^m - n_i^m) = \{N^m\}^{t-1} (N^m - r). \tag{3.2}
$$

The maximum likelihood estimator of γ is also obtained numerically. da Silva (1999) shows that when $n_i^m / X_i \cong 0$, $\hat{\gamma} = \sum_{i=1}^{t} X_i^u / \sum_{i=1}^{t} X_i^m$.

4 Model 2

Model 1, discussed in Section 3, has some good features like simplicity and little computational effort necessary to estimate the parameters of interest. However, we recognize that the sampling scheme used to collect the data has not been properly taken into account in the modelling. As mentioned in the second paragraph in the beginning of Section 3, the bowhead whale sampling scheme was such that multinomial trials of assigning photos to whales are not really independent because multiple photos are taken of the same whale under similar conditions and that violates the independence assumption.

A more realistic model may be described by assuming that the number of good photos per *captured* whale follows a truncated at zero Poisson distribution since an individual is *captured* only if it has at least one good photo. This distribution is appropriate for modelling the number of good photos of individual marked whales and, therefore, the joint distribution of the number of good photos is expressed as the product of truncated at zero i.i.d. Poisson distributions. However, in the case of the unmarked whales, the only information available is the total number of good photos, since we cannot tell how many individuals were encountered when the samples were taken because it is not possible to recognize an unmarked whale.

In this section we introduce a new model for estimating population size when part of the population is uncatchable. We will describe the model and then study the relative merits, in terms of bias and variance, of this model compared to model 1.

4.1 The model

We may think of the total number of good photos of the unmarked whales at time i, X_i^u, as being expressed by a random sum of i.i.d. random variables, i.e.,

$$X_i^u = x_{1i}^u + x_{2i}^u + \cdots + x_{n_i^u i}^u, \quad n_i^u = 1, \cdots, N^u$$

where

• x_{ji}^u describes the number of good photos taken of unmarked encountered individual j at time i.

• n_i^u is a random variable describing the number of unmarked individuals that were photographed at time i; its range varies from 1 to N^u.

x_{ji}^u has probability density function $f(\cdot)$. The n-fold convolution of the density $f(\cdot)$ is $f^{(n)}$ and it will help us defining the probability distribution of the random sum above as follows:

$$Pr(X_i^u = v) = \sum_{n=1}^{N^u} f^{(n)}(v) P_{N_e^{(n)}}. \tag{4.1}$$

The convolution $f^{(n)}$ accounts for the number of *encountered* unmarked whales in the sample. The probability distribution of the number of *encountered* unmarked whales is $P_{N_e^{(n)}}$.

We are modeling the number of good photos per unmarked whale using the same distribution assumed for such quantity for the marked individuals. Therefore,

$$f^{(k)}(v) = Pr\left(\sum_{n=1}^{k} x_n = v\right) = \frac{e^{-(k\lambda)}(k\lambda)^v}{v!}; \quad v = 0, 1, 2, \ldots; \quad k = 1, 2, \ldots.$$

The number of encountered unmarked whales at time i, N_e, will be modeled by a

Binomial distribution with parameters N^u and ${p_i}'$, where ${p_i}'$ is the *encountering* probability at time i. Capture probability at time i, p_i, represents the probability of taking a good photo of an encountered individual at time i and it is expressed by the *encountering* probability at time i, ${p_i}'$, times the probability of a whale having at least one good photo, i.e.,

$$p_i = {p_i}'(1 - e^{-\lambda_i}).$$

The Binomial model mentioned above is then expressed by

$$Pr(N_e = k) = \left(\begin{array}{c} N^u \\ k \end{array} \right) \left[\frac{p_i}{1 - e^{-\lambda_i}} \right]^k \left[\frac{1 - e^{-\lambda_i} - p_i}{1 - e^{-\lambda_i}} \right]^{N^u - k}.$$

In summary, we are expressing the distribution of X_i^u as the distribution of a random sum given as follows:

$$Pr(X_i^u = v) = \sum_{k=1}^{N^u} \frac{e^{-(k\lambda_i)}(k\lambda_i)^v}{v!} \left(\begin{array}{c} N^u \\ k \end{array} \right) \left[\frac{p_i}{1 - e^{-\lambda_i}} \right]^k \left[\frac{1 - e^{-\lambda_i} - p_i}{1 - e^{-\lambda_i}} \right]^{N^u - k}. \quad (4.2)$$

In the present approach we are considering that the number of good photos of unmarked whales that were encountered at time j does not depend on the number of good photos of the marked whales at that occasion. That follows because the total number of good photos is being treated as a random quantity given by the sum of the random number of good photos of marked and unmarked whales, as it happens in a cross-classification table whose cells are generated by independent Poisson processes.

Let X_{ij}^m represent the number of good photos of *encountered* marked whale j at time i, where $i = 1, \ldots, t$ and $j = 1, \ldots, N^m$ with values ranging from $0, 1, \ldots$. The likelihood function is defined by the joint distribution of the variables $\{ X_i^u \}$, $\{ X_{ij}^m \}$ and $\{ u_w \}$:

$$\begin{aligned}
L &= Pr(\{X_i^u\}) Pr(\{X_{ij}^m\} \mid \{u_w\}) Pr(\{u_w\}) \\
&= \left[\prod_{i=1}^{t} Pr(X_i^u = x_i^u) \right] \times \left[\prod_{i=1}^{t} \prod_{j=1}^{n_i^m} Pr(X_{ij}^m = x_{ij}^m \mid n_i^m) \right] \times Pr(\{u_w\}) \\
&= \prod_{i=1}^{t} \left\{ \sum_{k=1}^{N^u} \frac{e^{-(k\lambda_i)}(k\lambda_i)^{x_i^u}}{x_i^u!} \left(\begin{array}{c} N^u \\ k \end{array} \right) \left[\frac{p_i}{1 - e^{-\lambda_i}} \right]^k \left[\frac{1 - e^{-\lambda_i} - p_i}{1 - e^{-\lambda_i}} \right]^{N^u - k} \right\} \\
&\quad \times \prod_{i=1}^{t} \prod_{j=1}^{n_i^m} \left(\frac{1}{1 - e^{-\lambda_i}} \right) \left(\frac{e^{-\lambda_i} \lambda_i^{x_{ij}^m}}{x_{ij}^m!} \right) \\
&\quad \times \frac{N^m!}{(N^m - r)! \prod_w u_w!} \prod_i^t p_i^{n_i^m} {q_i}^{N^m - n_i^m}. \quad (4.3)
\end{aligned}$$

Estimation of the parameters of interest is facilitated by the use of the EM-algorithm. The missing characteristic here is the number of unmarked individuals which were encountered at each occasion, $\{n_i^u\}$. In the next section we briefly describe the EM-algorithm and apply it to the estimation of bowhead abundance.

5 The EM-algorithm

Garthwaite *et al.* (1995) and Tanner (1993) review the EM-algorithm, popularized by Dempster *et al.* (1977). The EM-algorithm is particularly useful when some of the data are missing. However, as noted by Dempster *et al.* (1977), other types of problems can be solved by the same general method. The EM-algorithm can be viewed as a data augmentation algorithm; rather than performing a complicated maximization or simulation, one augments the observed data with appropriate latent data which simplifies the calculation and subsequently performs a series of simple maximizations or simulations. According to Tanner (1993), the principle of data augmentation can be stated as follows: augment the observed data Y with latent data Z so that augmented likelihood function $p(\theta \mid Y, Z)$ is "simple". The maximization process is in general easier when one takes advantage of this simplicity. In fact, the EM-algorithm is an iterative method for locating the maxima of the likelihood function mentioned above. Using the jargon that has been adopted since the publication of the paper by Dempster *et al.* (1977), the available data is called *incomplete data*, while the augmented data is called *complete data*.

In the EM-algorithm each iteration consists of two steps, the E- or expectation step, and the M- or maximization step. To start the iterations it is necessary to input initial values for the vector of parameters, θ. Using such initial values, the E-step follows by estimating the missing data (latent data) using the conditional expectation of the missing data given the actual data. In the M-step, the maximum likelihood estimator of θ is computed from the "complete data", and using such estimates as initial values, the process continues until convergence. Next we present in detail the development of the EM algorithm for estimating population size of the bowhead whale.

5.1 Likelihood for the complete data

The likelihood function of the complete data is defined by the joint distribution of the variables $\{ X_i^u \}$, $\{ X_{ij}^m \}$, $\{ u_w \}$ and $\{ n_i^u \}$ and is given by

$$\begin{aligned}
L &= Pr(\{X_i^u\} \mid \{n_i^u\}) Pr(\{X_{ij}^m\} \mid \{u_w\}) Pr(\{u_w\}) Pr(\{n_i^u\}) \\
&= \left[\prod_{i=1}^{t} Pr(X_i^u = x_i^u \mid n_i^u) \right] \times \left[\prod_{i=1}^{t} \prod_{j=1}^{n_i^m} Pr(X_{ij}^m = x_{ij}^m \mid n_i^m) \right] \\
&\quad \times Pr(\{u_w\}) \times \prod_{i=1}^{t} Pr(n_i^u)
\end{aligned}$$

$$= \left[\prod_{i=1}^{t} \frac{e^{-\lambda_i n_i^u}(\lambda_i n_i^u)^{x_i^u}}{x_i^u!}\right] \times \prod_{i=1}^{t}\prod_{j=1}^{n_i^m}\left(\frac{1}{1-e^{-\lambda_i}}\right)\left(\frac{e^{-\lambda_i}\lambda_i^{x_{ij}^m}}{x_{ij}^m!}\right)$$

$$\times \frac{N^m!}{(N^m-r)!\prod_w u_w!}\prod_i^t p_i^{n_i^m} q_i^{N^m-n_i^m}$$

$$\times \prod_{i=1}^{t}\binom{N^u}{n_i^u}\left[\frac{p_i}{1-e^{-\lambda_i}}\right]^{n_i^u}\left[\frac{1-e^{-\lambda_i}-p_i}{1-e^{-\lambda_i}}\right]^{N^u-n_i^u}.$$

The log-likelihood of the complete data is maximized with respect to the λ_i and p_i when their respective score functions equal zero, i.e, $\dot{l}_{\lambda_i} = 0, \quad i = 1, \ldots, t$ and $\dot{l}_{p_i} = 0, \quad i = 1, \ldots, t$.
This log-likelihood function is maximized with respect to the N^u and N^m by solving the difference equations

$$\triangle \log L(N^u) = 0 \quad \text{and} \quad \triangle \log L(N^m) = 0,$$

i.e., the maximum likelihood estimate of N^u is such that

$$\log L(N^u) = \log L(N^u - 1)$$

and in such case that implies that $\hat{N}^u$ must solve the equation

$$t \log N^u - \sum_{i=1}^{t} \log(N^u - n_i^u) + \sum_{i=1}^{t} \log(1 - e^{-\lambda_i} - p_i) - \sum_{i=1}^{t} \log(1 - e^{-\lambda_i}) = 0. \quad (5.1)$$

Similarly, the maximum likelihood estimate of N^m is such that

$$\log\left(\frac{N^m}{N^m - r}\right) = -\log\left(\prod_{i=1}^{t}(1 - p_i)\right) \quad (5.2)$$

which implies that the maximum likelihood estimate of N^m is given by the expression

$$N^m = \frac{r}{1 - \prod_{i=1}^{t}(1 - p_i)}.$$

The number of unmarked individuals caught at time i, i.e., the "missing data", is estimated by the conditional expectation of the $\{n_i^u\}$ given $\{X_i^u\}$, i.e.,

$$E(n_i^u \mid X_i^u) = \int n_i^u f_{n_i^u \mid X_i^u} dn_i^u$$

$$= \frac{\int n_i^u f_{X_i^u|n_i^u} f_{n_i^u} dn_i^u}{f(X_i^u)}$$

$$= \frac{\displaystyle\sum_{k=1}^{N^u} k \frac{e^{-(k\lambda_i)}(k\lambda_i)^{x_i^u}}{x_i^u!} \left(\begin{array}{c} N^u \\ k \end{array} \right) \left[\frac{p_i}{1-e^{-\lambda_i}}\right]^k \left[\frac{1-e^{-\lambda_i}-p_i}{1-e^{-\lambda_i}}\right]^{N^u-k}}{\displaystyle\sum_{k=1}^{N^u} \frac{e^{-(k\lambda_i)}(k\lambda_i)^{x_i^u}}{x_i^u!} \left(\begin{array}{c} N^u \\ k \end{array} \right) \left[\frac{p_i}{1-e^{-\lambda_i}}\right]^k \left[\frac{1-e^{-\lambda_i}-p_i}{1-e^{-\lambda_i}}\right]^{N^u-k}}. \tag{5.3}$$

All equations above have to be solved using the following iterative procedure:

1. Take initial values
$$\Psi^{(0)} = \left(\{\lambda_i^{(0)}\}, \{p_i^{(0)}\}, N^{u(0)}, N^{m(0)}\right)$$
2. Using $\lambda_i^{(t)}$, $p_i^{(t)}$, and $N^{u(t)}$, calculate $\{\ n_i^{u(t)}\ \}$ using (5.3).
3. Maximize the log-likelihood of the complete data (5.1) as a function of $\{\ \lambda_i\ \}$and $\{\ p_i\ \}$, i.e., obtain $\{\ \hat{\lambda}_i^{(t)}\ \}$ and $\{\ \hat{p}_i^{(t)}\ \}$, and using those values solve equations (5.1) and (5.2) to obtain $\hat{N}^{u(t)}$ and $\hat{N}^{m(t)}$. In summary, obtain:
$$\Psi^{(t+1)} = \left(\{\lambda_i^{(t+1)}\}, \{p_i^{(t+1)}\}, N^{u(t+1)}, N^{m(t+1)}\right)$$
4. Using $\Psi^{(t+1)}$ repeat steps 2 and 3 until convergence for t=0,1,. . . .

An important property of the EM algorithm is that the likelihood of the observed data (incomplete data) always increases along an EM sequence, and convergence is attained when differences between two consecutive iterations of the log-likelihood of the *incomplete data*, equation 4.3, are arbitrarily small. In the present problem this property does not hold for the entire EM sequence. In the next section we discuss that problem and also the choice of initial points for the algorithm.

5.2 Starting points for the algorithm and stopping rule

In the EM-algorithm theory, certain regularity conditions must be observed to assure that the likelihood of the incomplete data converges monotonically to some value L^*. According to McLachlan and Krishnan (1997), p. 85, if the likelihood of the incomplete data is unimodal in its parameter space and a certain differentiability condition is satisfied, then any EM sequence converges to the unique MLE, irrespective of its starting point $\Psi^{(0)}$. In the present problem the log-likelihood function is not differentiable at all parameters, since N^u and N^m are not continuous functions. This is the reason why our results depend on the choice of the starting points. However, as we shall see later, such dependency will not compromise a comparision of the performance of models 1 and 2.

According to McLachlan and Krishnan (1997), in problems where the likelihood of the incomplete data has several stationary points, convergence of the EM sequence to either type (local or global maximizers, saddle points) depends on the choice of starting point $\Psi^{(0)}$. In the present problem, even though the log-likelihood function of the incomplete data is not concave since it is not continuous over all the parameters (here we are using the result from a theorem by Madden (1986) which says that if a function $f : R^n \rightarrow R$ is concave, then f is continuous) graphical analysis indicates a function pretty well behaved with respect to the $p_i's$ and $\lambda_i's$, with no saddle points or local maxima (when N^u and N^m are maintained fixed). However, we observe that different values of N^u describe different surfaces for the log-likelihood that may intercept one another. That implies that each surface defined by a different value of N^u may be maximized with respect to λ_i and p_i, $i = 1, \cdots, t$ for distinct values of these parameters. In regions we believed to be close to the global maximum, the log-likelihood function for each iteration of the EM-algorithm may switch around surfaces causing non-monotonic increase of the log-likelihood.

In order to define a stopping rule for the procedure, we defined a criterion based on the analysis of changes in the log-likelihood of the incomplete data from subsequent iterations, Δl. Once the maximum is reached, Δl should vary around zero. By assuming that Δl has mean μ and variance σ^2, then the interval $\mu \pm 2\sigma$ may give us an indication of stabilization of the process.

We analyzed non-interceptable windows of size 10, i.e., at each ten EM-iterations we computed the interval above. If the standard deviation of Δl was smaller or equal to a predetermined level of tolerance δ and if the interval contained zero, we stopped the iterations. The idea of this procedure is based on the work of Chan and Ledolter (1995), who work with Monte Carlo EM algorithm using a Markov chain sampling technique (Gibbs sampling) in the calculation of the expectation in the E-step of the EM algorithm. However, in the authors' framework the analyzed intervals are of the form $\mu \pm L\sigma$, where L is defined according to some quantities derived from Monte Carlo and Gibbs sampling procedures.

Figure 1 shows the changes in log-likelihood, Δl, for a given bowhead whale simulated data set among the 100 ones we used in the analysis that will be presented in Section 7. After 60 iterations the changes in log-likelihood seem to represent a noise process. Using the criterion described in Section 5.2, the conditions for the stopping rule were met for the 10 length window between iterations 30 and 39. The vertical line in the lower right panel of Figure 1 represents the point where the stopping rule terminated the iterative process.

As we mentioned before, the EM-algorithm in our case depends on the initial values. The best initial values we could come up with were the estimated values of the parameters given by the multinomial model that was discussed in Section 3. In the next section we present, for each of the models proposed, a bootstrap approach used to estimate variance of $\hat{N}$ and confidence intervals for N.

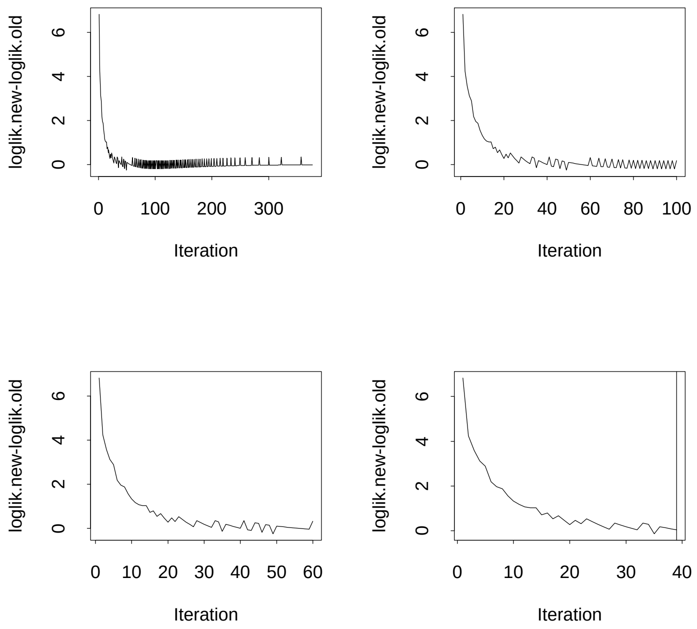

Figure 1 *Stopping rule for the EM-algorithm for the alternative model. Each plot shows a smaller range of iterations. The vertical line in the plot in the lower right hand corner shows where stopping rule terminated iterations.*

6 Evaluating the variance of the estimators via bootstrap

We follow Buckland (1980) and others in using the parametric bootstrap to estimate standard error. In the parametric bootstrap setting we draw B samples of size n from the distribution $\hat{F}_{par}$, an estimate of F derived from a parametric model for the data. Where parameters are needed to specify the distribution, estimates of these parameters computed from the original data are used.

The choice between nonparametric and parametric bootstrap in capture-recapture is addressed by Buckland and Garthwaite (1991). They note that even though the nonparametric bootstrap is more widely used and more familiar than the parametric bootstrap, the latter allows us to choose which underlying distribution model to assume for the data. Mark-recapture provides an example where the nonparametric bootstrap makes specific parametric assumptions that are not immediately apparent. That may lead a user to bootstrap on the wrong sampling unit, or to conclude erroneously that the results are more robust than those from a parametric approach.

For **model 1**, presented in Section 3, a parametric bootstrap approach for estimating variance of $\hat{N}$ and confidence intervals for N involves the following steps:

1. Obtain the "original data" by running the data simulation program once or by using the real bowhead data.

2. Estimate the parameters $\hat{N}^m$, $\hat{p}_1, \cdots, \hat{p}_4$, $\hat{\gamma}$, and $\hat{N}$ from the data obtained in step 1.

3. Using the estimated capture probabilities $(\hat{p}_1, \cdots, \hat{p}_4)$ and population size of the marked whales $(\hat{N}^m)$, simulate the number of individuals with a given capture history w, u_w^*, by using Darroch's multinomial model. With that procedure we obtain the sample sizes $(n_1^{m*}, \cdots, n_4^{m*})$ and r^* to be used in calculating the estimate N^{m*} (and p_i^*'s) from the bootstrap sample.

4. Simulate truncated binomial distributions. The total number of photos X_i obtained in the "data" at each occasion i was kept fixed at its value in the original data and was divided among marked and unmarked whales as follows:

   (i) A truncated binomial distribution with parameters
   $$\left(X_i, \quad \frac{1}{1+\hat{\gamma}}\right)$$
   was simulated to obtain the number of good photos of marked whales;

   (ii) The number of good photos of unmarked whales was obtained by subtraction.

   This provides the data needed to calculate the estimate γ^*.

5. Calculate the parameter estimates, including N^*, from the bootstrap sample.

6. Repeat steps 3-5 B times.

In the above steps, * denotes data or an estimate from the bootstrap sample.

For **model 2**, presented in Section 4, the parametric bootstrap approach for estimating variance of $\hat{N}$ and confidence intervals for N is similar to the last one and consists of the following steps:

1. Obtain the "original data" by running the data simulation program once or by using the real bowhead data.

2. Estimate the parameters $\hat{N}^m$, $\hat{N}^u$, $\hat{p}_1, \cdots, \hat{p}_4$, and $\hat{\lambda}_1, \cdots, \hat{\lambda}_4$ from the data obtained in step 1.

3. Using the estimated capture probabilities $(\hat{p}_1, \cdots, \hat{p}_4)$ and population size of the marked whales $(\hat{N}^m)$, simulate the number of individuals with a given capture history w, u^*_w, by using Darroch's multinomial model. With that procedure we obtain the sample sizes $(n_1^{m*}, \cdots, n_4^{m*})$ and r^*.

4. Simulate the number of good photos of the n_i^{m*} marked whales that were captured at sampling occasion i by using a truncated at zero Poisson distribution with parameter $\hat{\lambda}_i$.

5. Simulate the number of good photos of the encountered unmarked whales at each sampling occasion using the following two steps:

   (i) Simulate the number k_i of unmarked encountered whales at time i through a Binomial distribution with parameters $\hat{N}^u$ and $\frac{\hat{p}_i}{1-e^{-\hat{\lambda}_i}}$;

   (ii) Simulate the number of good photos of those k_i whales through a Poisson distribution with parameter $k_i\hat{\lambda}_i$.

6. Using the simulated data estimate the parameters using model 2.

7. Repeat steps 3-6 B times.

The standard deviation of N^* over the B bootstrap samples, $s.e.^*$, estimates the standard error of $\hat{N}$. The difference between the mean of the B values N^* and $\hat{N}$, $bias^*$, estimates the bias of $\hat{N}$. Confidence intervals may be found using the percentile method (Efron, 1981; Buckland and Garthwaite, 1991) as follows. Order the N^* from smallest to largest, and denote the ordered list by $\hat{N}_{(j)}$. Approximate $100(1-2\alpha)\%$ confidence limits are then given by $\hat{N}_{(k)}$ and $\hat{N}_{(k')}$, where $k = (B+1)\alpha$ and $k' = (B+1)(1-\alpha)$, both rounded to the nearest integer value.

The determination of the number B of bootstrap replications depends on the application. Efron (1981) suggests that bootstrap estimates of standard error usually have relatively little bias, and very seldom are more than $B = 200$ replications needed for estimating a standard error. Many more replications are needed to obtain a good estimate of bias or for construction of confidence intervals. The percentile method depends on the tail of the distribution where fewer samples occur. Buckland and Garthwaite (1991) advocate that for a 95% confidence interval $B = 1000$ should be satisfactory, whereas $B = 200$ is likely to be inadequate. In Section 7.3 we give some more details about those calculations. In the next section we provide some more description about the simulated data in order to enhance comprehension about the results obtained from the comparison of the proposed models to be discussed in Sections 7.2-7.4.

7 Results

7.1 Some simulation details

We applied the procedure described in Section 6.0 to 100 bowhead whale simulated data sets. As we mentioned in Section 2, we did not allow for differences between the population sizes for the years of 1985 and 1986 since we intended to simulate data from a closed population. For the simulated data we considered a population size of $N = 6,734$ individuals. This value represents the average between the estimated population sizes for the years of 1985 and 1986 (6,649 and 6,820, respectively) obtained from the most likely population trajectory in a Bayesian synthesis analysis carried out by Givens (pers. comm.).

The capture probabilities used in the simulations were

$$p_1 = 0.095, \quad p_2 = 0.158, \quad p_3 = 0.060, \text{ and } p_4 = 0.065.$$

In order to simulate the number of good-quality photos taken from each whale photographed, we supposed that quantity to follow a zero truncated Poisson distribution. Using some data sets available with data restricted to the mid-back portion of the whale (since it is considered to be the most likely region to be photographed and recognized) we estimated the average number of photos/whale using only the good quality photos. That number was 1.541 photos/whale. For the parameter of a Poisson distribution, we wanted a value λ that after truncation gives an average of 1.541 photos/whale. A value of $\lambda = 0.938$ satisfies this criterion. This parameter is a general one for the population without considering differences among groups of whales, i.e., regardless of the degree of marking. That is so because according to the sampling protocol the whales were photographed regardless of whether or not they were marked.

7.2 Model comparisons

According to Figure 2 we noticed that the population size estimate depended on the initial values of the parameters. The upper left panel of Figure 2 makes pointwise comparisons of the 100 estimated values of N obtained by the alternative model when fixed initial values for the parameters were used and those values obtained by the multinomial model. The upper right panel of Figure 2 makes the same kind of comparison when initial values for the alternative model were given according to estimated values obtained by the multinomial model. We observe a very good agreement of the estimated values in the last case.

The lower left panel of Figure 2 compares the estimated values of N when fixed initial values for the parameters for the alternative model were used, when those initial values were given according to the respective estimated values given by the multinomial model, and the estimated values of N obtained by the multinomial model. We observe that box-plot "1" (estimated values using fixed initial points) presents much smaller variability when compared to the other two, which are very similar, though most of the estimated values are below the true value of the population size in the simulations which is given by the horizontal line at value 6,734.

The scatterplot of the estimated values using multinomial and alternative models describes a straight line passing through zero indicating very good agreement of results. The evaluation of the precision of $\hat{N}$ under each of the models is discussed in the next section.

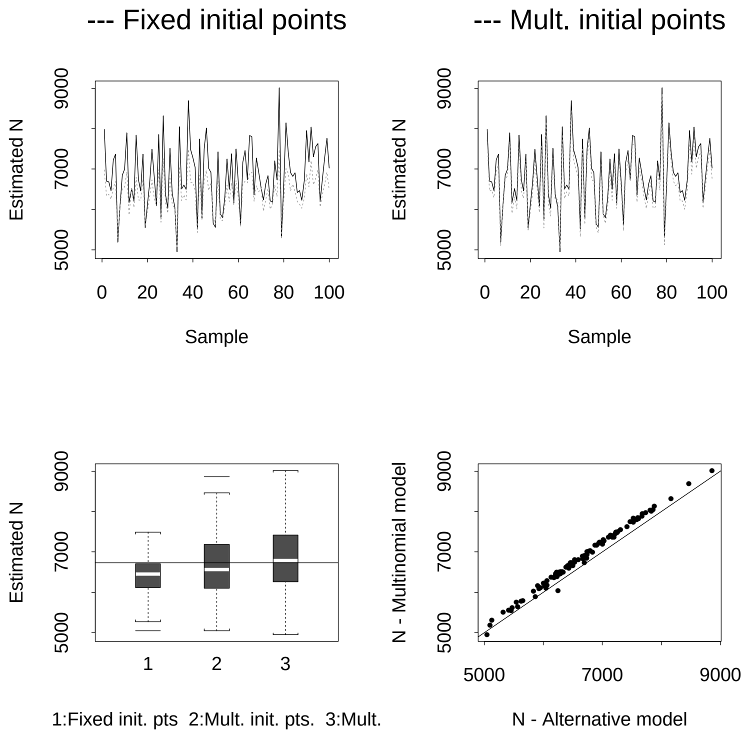

Figure 2 *Comparing the multinomial to the alternative model*

7.3 Variance comparision using simulated data

As we discussed in Section 6, it is necessary to choose an adequate value for the number B of bootstrap samples to be drawn. For the 100 simulated data sets we estimated variance using 300 bootstrap samples, drawn from each model, using the procedures described in Section 6.

In order to compare variances calculated under models 1 and 2, for each one of the 100 simulated data sets we calculated bootstrap variances of $\hat{N}$. For model 1, the average of the bootstrap standard errors of $\hat{N}$ was $\overline{s.e.^*_1} = 827$, while for model 2 it was $\overline{s.e.^*_2} = 948$. Figure 3 shows a scatterplot of the 100 standard errors estimated under each model. As we can observe, the bootstrap standard errors obtained under model 2 tend to be larger than those obtained under moder 1. Actually, that happened in 96% of the time. That may be due to the larger number of parameters to estimate in model 2 compared to model 1. The initial values for N used in the EM procedure also seem to play a rule in the precision of $\hat{N}$ under model 2. As we observe from the lower left panel of Figure 2, the box-plots of the estimated values of N calculated using different approaches to stablish the starting points of N display non-negligible differences in the range of estimated values of N. The use of a fixed and unique starting point for N produces less variable estimated values of N than the use of distinct initial points for each one of the different simulated samples.

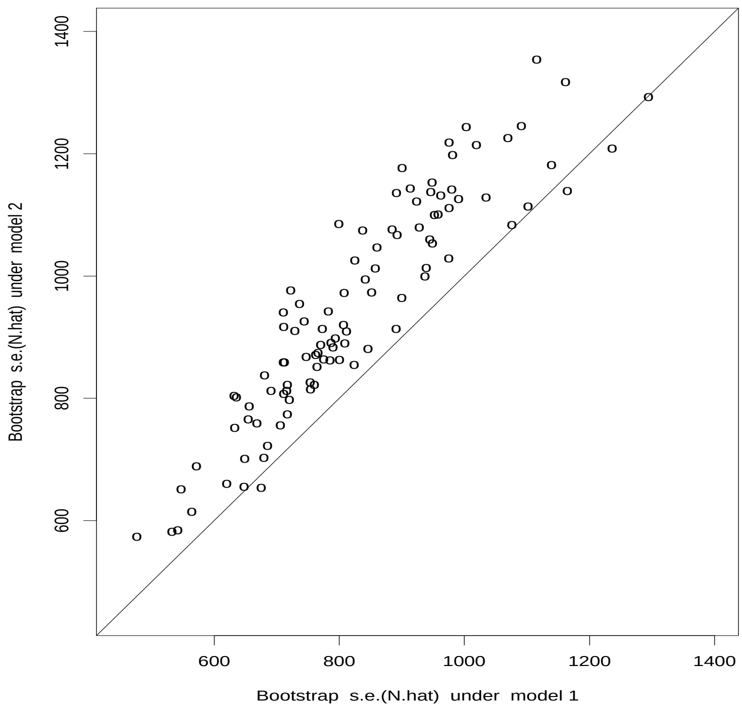

Figure 3 *Comparing the bootstrap s.e.*($\hat{N}$) *calculated under models 1 and 2.*

Regarding bias evaluation, we found that models 1 and 2 tend to produce biased estimates of N of about the same magnitude but in different directions. Model 1 tends to produce positively biased estimates ($\overline{bias_1^*} = 210.5$) while model 2 tends to produce negatively biased ones ($\overline{bias_2^*} = -206.5$). These last two values were based on only 100 bootstrap samples due to processing time constraints. Better bias estimates would require a much larger computing effort. Actually, working with model 1, 3,000 bootstrap samples and the same 100 simulated data, da Silva *et al.* (2000) found that the average bootstrap bias under model 1 was 78, a negligible one. Considering all the discussed aspects in the comparison of models 1 and 2, we recommend the use of model 1 despite some of its assumptions are not fully realistic. As it was presented, model 1 tend to produce more precise estimates of N than model 2. Besides that, model 1 is far easier to implement. In the next section we apply the two models to real bowhead data.

7.4 An application to real bowhead data

In this section we present an application of models 1 and 2 to real bowhead data. In the actual data set four sampling occasions (spring 1985, summer 1985, spring 1986, and summer 1986) are considered. The variables in the data base were used to create a data set containing records with the following information: Four columns indicate the capture histories of the bowheads, with 1 indicating that the whale was captured, and 0 that it was not captured in the sampling occasion represented by the column. Four columns indicating the number of good photos obtained for each of the captured individuals by sampling occasion. There were 1,677 records in the data set, 229 belonging to marked individuals, with 16 of the 229 captured on more than one occasion. The real data set only considers the adult bowhead population, i.e., calves are excluded.

Table 1 summarizes the results of the analysis. Variance and confidence intervals were obtained according to the bootstrap procedures described in Section 6. Following da Silva et al. (2000), we used $B = 3,000$ bootstrap samples for each model. As observed in the simulated data, the estimated value of N obtained from model 1 is larger than the one obtained using model 2. This tendency is also reflected in the confidence intervals for N. When model 2 was used, the standard error of $\hat{N}$ was slightly larger than that obtained using model 1. Generally speaking, both population estimates and their standard errors are comparable to the estimates obtained from the combined visual and acoustic census efforts conducted near Point Barrow in 1985 and 1986. Raftery and Zeh (1998) applied generalised removal method to another bowhead whale data set that is totally independent from the bowhead photo-id data. The other data set is a combination of visual and acoustic bowhead data collected over the years. Using such dataset Raftery and Zeh (1998) obtained estimates of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (including calves) of 6,039 (SE=1,915) and 7,734 (SE=1,450) for 1985 and 1986 respectively. The estimates in Table 1 also compare well with the 1985 and 1986 estimates of 6,649 and 6,849 (excluding calves) from the Bayesian synthesis analysis of Givens (pers. comm.). The estimates of Givens incorporate the estimates of Raftery and Zeh (1998) and additional data on bowhead whale population dynamics.

Table 1 *Real bowhead data - summary of the results.*

model	$\hat{N}$	$s.e.(\hat{N})$	Intervals (95%)
1	7,022	2,017	(4,701;12,561)
2	6,580	2,049	(3,705;11,476)

8 Discussion

Model 2 was built aiming to make better use of the photo-id data available and to incorporate more realistic assumptions than model 1. For the marked whales we have information about the number X_{ij}^m of good photos taken from *encountered* marked whale j at time i. This kind of individual bowhead data was considered when building model 2 while in model 1, for each sampling occasion i, we only accounted for the sum of those quantities, X_j^m, $j = 1, \cdots, t$, defined in Section 3.1. In model 1 it was assumed that X_j^m follows a Binomial distribution. However, since multiple photos of an individual may occur, the independence assumption of the Binomial model is violated. Though, as we said before, this kind of violation might be a mild one since it is known from the data that the average number of good photos per whale is about 1.5. Nevertheless, in order to avoid this kind of violation and taking full advantage of the data available, in model 2 it was assumed that X_{ij}^m follows a Poisson distribution.

Precision in the estimation of N was evaluated using parametric bootstrap. Deduction of analitic formulation for the variance of $\hat{N}$ is a difficult problem to tackle. The likelihood functions related to each one of the models do not belong to a regular family of distributions since the range of u_w is not independent of N^m - actually, $0 \leq u_w \leq N^m$. Thus, the large-sample theory of maximum likelihood estimators is not valid. A possible approach to persue is to try to derive a covariance matrix using the theory developed by Sanathanan (1972).

Simulated bowhead data revealed that models 1 and 2 tend to produce very comparable estimates of N. However, model 1 tended to provide more precise estimates of N than model 2. It seems that not only the higher number of parameters to estimate in model 2 but also the initial points used in the EM algorithm play some rule in the variance of $\hat{N}$ under model 2. Therefore, we recommend the use of model 1 even though some assumptions in this model are not fully realistic.

Dependency of the solutions on initial points is caused by the discrete nature of N^m and N^u. Thus, the problem of maximizing the likelihood as a function of those discrete parameters seems to be an $\mathcal{NP}$-hard optimization problem (see Garey and Johnson, 1979) to whom methods producing pseudo-independent solutions from initial points would require a very large number of trials making the methodology of no practical interest. Our simulations indicate that the heuristics used in model 2 results in solutions that are very close to the ones we obtained for model 1. We understand that besides the acknowledged dependency on starting points, the results obtained are useful for biologists working with estimation of animal abundance based on photo-identification data.

The estimated values of N in Table 1 agree with the results from the simulations discussed in the previous sections: the estimated value of N obtained under model 1 was larger than that obtained under model 2. Besides that, $s.e.(\hat{N})$ obtained under model 1 was slightly smaller than that obtained under model 2.

Models 1 and 2 applied to real bowhead data both produced estimates (see Table 1) that compared well with the estimates obtained by Raftery and Zeh (1998) (6,039 (SE=1,915) and 7,734 (SE=1,450) for 1985 and 1986 respectively - including calves) and those obtained by Givens (pers. comm.), 6,649 and 6,820 for 1985 and 1986 respectively (excluding calves). Both confidence intervals of Table 1 cover the 1985 and 1986 estimates of 6,649 and 6,849 (excluding calves) from the Bayesian synthesis analysis of Givens (pers. comm.), which reflect the best information available on the size of the Bering-Chukchi-Beaufort Seas stock of bowhead whales in those years. Since the population size data and estimation methods on which the Givens estimates are based are completly different from the data and methods used here, our results provide independent confirmation for the population estimates currently accepted by the IWC Scientific Committee.

Acknowledgements

The author is grateful to Dr. Frederico Cruz (UFMG/ICEx-Departamento de Estatística) for helpful discussions and to the anonymous referees for a careful reading of the manuscript and for constructive suggestions which considerably improved the paper. Data for this analysis were provided through photographic aerial surveys conducted by LGL Limited (King City, Ontario, Canada) and the National Marine Mammal Laboratory (NMFS, NOAA, Seattle, Washington,U.S.A.). We particularly thank Bill Koski and Gary Miller of LGL and Dave Rugh of the National Marine Mammal Laboratory for their work on the surveys, the scoring and matching of photographs, and creation of the data base.

(Received October, 2001. Accepted January, 2003.)

References

Braham, H.W., Fraker, M. and Krogman, B. (1980). Spring migration of the Western Artic population of bowhead whales. *Mar. Fis. Rev.*, **42**, 36-46.

Buckland, S.T. (1980). A modified analysis of the Jolly-Seber capture-recapture model. *Biometrics*, **36**, 419-35.

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

da Silva, C. Q. (1999). Capture-recapture estimation of bowhead whale population size using photo-identification data. Doctoral Thesis, University of Washington, Seattle, WA, i-x + 190pp.

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 Research Management*, **2**, 45-61.

Chan, K.S. and Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving counts. *Journal of the American Statistical Association*, **90**, 242-252.

Chapman, D.G. (1951). Some properties of the hypergeometric distribution with applications to zoological sample sensuses. *Univ. Cal. Publ. Stat.*, **1**, 131-160.

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

Darroch, J.N. (1959). The multiple-recapture census. II: Estimation when there is immigration or death. *Biometrika*, **46**, 336-51.

Dempster, A.P., Laird, N.M., Rubin, D.R. (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society B*, **39**, 1-38.

Efron, B. (1981). Nonparametric standard errors and confidence intervals (with discussion). *Can. J. Stat.,* **9**, 139-72.

Efron, B. and Tibshirani, R.J., and Jones, B. (1993). *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Garey, M.R. and Johnson, D.S. (1979). *Computers and Intractability.* San Francisco: Freeman.

Garthwaite, P.H., Jolliffe, I.T., and Jones, B. (1995). *Statistical Inference.* 1st ed. UK: Prentice Hall International Limited.

Givens, G.H. (1993). A Bayesian framework and importance sampling methods for synthesising multiple sources of evidence and uncertainty linked by a complex mechanistic model. Doctoral Dissertation Thesis, University of Washington, Seattle, WA 98195. i-xii + 157pp.

Lincoln, F.C. (1930). Calculating waterfowl abundance on the basis of banding returns. *U.S. Department of Agriculture Circular* (No. 118).

McLachlan, G.J. and Krishnan, T. (1997). *The EM algorithm and extensions.* New York: John Wiley and Sons, Inc.

Madden, P. (1986). *Concavity & Optimization in Microeconomics.* New York: Basil Blackwell Ltd.

Petersen, C.G.J. (1896). The yearly immigration of young plaice into the Limfjord from the German Sea. *Rep. Danish. Biol. Sta.*, **6**, 1-48.

Raftery, A.E., and Zeh, J.E. (1995). Inference from a deterministic population dynamics model for bowhead whales. *Journal of the American Statistical Association*, **90**, 402-30.

Raftery, A.E., and Zeh, J.E. (1998). Estimating bowhead whale population size and rate of increase from the 1993 census. *Journal of the American Statistical Association*, **93**, 451-463.

Rugh, D.J. (1980). Bowhead whales reidentified through aerial photography near Point Barrow, Alaska. *Rep. Int. Whal. Commn*, **12**, 289-94. Special issue.

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 reidentifiability in aerial photographs of bowhead whales. *Rep. Int. Whal. Commn*, **48**, 501-12. Special issue.

Sanathanan, L. (1972). Estimating the size of a multinomial population. *The Annals of Mathematical Statistics*, **43**, 142-152.

Schnabel, Z.E. (1938). The estimation of the total fish population of a lake. *Amer. Math. Month.*, **45**, 348-352.

Tanner, M.A. (1993). *Tools for statistical inference.* 2nd ed. New York: Springer-Verlag.

Whitcher, B.J., Zeh, J.E., Rugh, D.J., Koski, W.R. and Miller, G.W. (1996). Estimationg of adult bowhead whale survival rates. Paper SC/48/AS12 presented to the IWC SC, June 1996.

Williams. J.A., Dawson, S.M. and Slooten, E. (1993). Distribution of bottlenosed dolphins (*Tursiops truncatus*) in Doubtful Sound, New Zealand. *Can. J. Zool.*, **71**, 2080-2088.

Cibele Queiroz da-Silva
Departamento de Estatística
Universidade Federal de Minas Gerais
31270-901, Belo Horizonte, MG, Brazil
cibeleqs@est.ufmg.br