MARINE MAMMAL SCIENCE, 21(1):80–92 (January 2005)
© 2005 by the Society for Marine Mammalogy

MULTISITE MARK-RECAPTURE FOR CETACEANS: POPULATION ESTIMATES WITH BAYESIAN MODEL AVERAGING

J. W. Durban[1]
University of Aberdeen,
School of Biological Sciences, Lighthouse Field Station,
Cromarty, Ross-Shire IV11 8YJ, United Kingdom

D. A. Elston
Biomathematics and Statistics Scotland,
Environmental Modelling Unit,
The Macaulay Institute,
Aberdeen AB15 8QH, United Kingdom

D. K. Ellifrit
Center for Whale Research,
355 Smugglers Cove Road,
Friday Harbor, Washington 98250, U.S.A.

E. Dickson
Whale and Dolphin Conservation Society,
Moray Firth Wildlife Centre,
Spey Bay, Moray IV32 7PJ, United Kingdom

P. S. Hammond
NERC Sea Mammal Research Unit,
Gatty Marine Laboratory, School of Biology,
University of St Andrews,
St Andrews, Fife KY16 8LB, United Kingdom

P. M. Thompson
University of Aberdeen,
School of Biological Sciences, Lighthouse Field Station,
Cromarty, Ross-Shire IV11 8YJ, United Kingdom

Abstract

Mark-recapture techniques are widely used to estimate the size of wildlife populations. However, in cetacean photo-identification studies, it is often impractical to sample across the entire range of the population. Consequently, negatively biased population estimates can result when large portions of

[1] Current address: National Marine Mammal Laboratory, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, Washington 98115-6349, U.S.A.; e-mail: john.durban@noaa.gov.

a population are unavailable for photographic capture. To overcome this problem, we propose that individuals be sampled from a number of discrete sites located throughout the population's range. The recapture of individuals between sites can then be presented in a simple contingency table, where the cells refer to discrete categories formed by combinations of the study sites. We present a Bayesian framework for fitting a suite of log-linear models to these data, with each model representing a different hypothesis about dependence between sites. Modeling dependence facilitates the analysis of opportunistic photo-identification data from study sites located due to convenience rather than by design. Because inference about population size is sensitive to model choice, we use Bayesian Markov chain Monte Carlo approaches to estimate posterior model probabilities, and base inference on a model-averaged estimate of population size. We demonstrate this method in the analysis of photographic mark-recapture data for bottlenose dolphins from three coastal sites around NE Scotland.

Key words: mark-recapture, model selection, model averaging, Bayesian analysis, Markov Chain Monte Carlo, population size.

Mark-recapture methods are well established for estimating the abundance of wildlife populations. Conventionally, marking or tagging is used to uniquely identify individuals in successive capture samples, and mark-recapture models use information on the recapture rate to estimate population size (Seber 1982). The conventional approach of physical capture and marking has also been generalized to other types of individual detection, greatly increasing the range of species that are amenable to mark-recapture population analysis. For cetaceans, photographic identification using natural markings (*e.g.*, Wilson *et al.* 1999), or individual recognition from microsattelite genotypes (*e.g.*, Palsbøll *et al.* 1997), have increased the utility and application of the mark-recapture approach. However, despite these practical developments, considerable difficulties remain in the application of mark-recapture methods to reliably estimate cetacean abundance.

Population estimation with mark-recapture depends crucially on the model that is used. There is a wide array of mark-recapture models (Chao 2001), and the choice of which to use depends on matching characteristics of the sample data to the inherent assumptions made by each model. This is particularly important for populations that are sampled using non-conventional mark-recapture approaches, where randomized or predetermined sampling designs are often not feasible, and sampling is more opportunistic. One consequence of such non-standardized sampling is that animals are more likely to be captured in some locations and times than others, inducing heterogeneity in capture probabilities that violate the standard assumptions of conventional mark-recapture models (Seber 1982, Hammond 1986). In the case of wide-ranging cetacean populations, much of the heterogeneity may be introduced by the movement of individuals beyond the range of single study areas, and by individual differences in these ranging patterns (Hammond 1986). A more fundamental problem may also occur when it is impractical to sample throughout the population's entire range, namely that a substantial proportion of the population may remain unavailable for capture, leading to negatively biased population estimates (*e.g.*, Hammond 1986, Whitehead *et al.* 1986, Hammond *et al.* 1990).

We propose an approach to address these problems by using a set of spatially discrete study sites to more explicitly account for movement patterns and allow sampling to penetrate farther into the population. Lists can then be compiled

containing the identities of all individuals seen in each study site. The "overlap" of individuals between these study site lists can then be treated as recaptures, which are spatially rather than temporally ordered. However, with opportunistic sampling designs, study sites may be located due to logistical convenience rather than by design, and it may be necessary to model dependencies between samples due to geographic distance effects. For example, two sites close together are more likely to get dependent overlap of individuals than will sites far away. Such dependence among samples leads to bias of mark-recapture estimators derived under the assumption of independence (Chao 2001). In such cases, the fundamental relationship between the recapture rate and the number of undetected animals is likely to be unclear and non-linear. Unbiased estimates of population size therefore require a modeling approach that can account for dependencies of this kind.

Our approach is to use log-linear models (*e.g.*, Fienberg 1972, Cormack 1989) for modeling dependencies between capture samples, allowing different hypotheses about dependencies to be represented as alternative models (*e.g.*, King and Brooks 2001). Assessing the sensitivity of population estimates to model selection, and quantifying model selection uncertainty, therefore represent important components of inference (Buckland *et al.* 1997). This is especially important in non-conventional and opportunistic mark-recapture studies, where study design has not been tailored to suit a specific model. We therefore present a Bayesian approach (*e.g.*, Madigan and York 1997, King and Brooks 2001) for model selection based on posterior model probabilities, and show how inference about population size can be based on a model-averaged probability distribution. We demonstrate this method in the analysis of photographic mark-recapture data for bottlenose dolphins (*Tursiops truncatus*) from multiple sites in the coastal waters around NE Scotland.

Methods

We consider the situation where animals in a closed population are sampled at a number of S spatially discrete study sites, in such a way that the individuals sampled in each site can be uniquely identified. A list can then be compiled for each study site, containing the identities of all individuals identified in each site. The aim is to estimate the size of this closed population based on the overlap of these lists, representing reidentification of individuals across sites. We describe the development of this approach by means of a hypothetical design with $S = 3$ sampling sites. This is the minimum number of sites required to identify dependencies between lists, but this approach can be logically extended to include more study sites as required.

Overlap of identification lists between sites can be conveniently represented in a contingency table (*e.g.*, Dellaportes and Forster 1999), where the $C = 2^S$ cells of the table refer to discrete categories formed by combinations of the study areas, which serve as classifying factors. If an individual cell is denoted by $i = (1, \ldots, C)$, then the corresponding cell count, n_i, denotes the number of individuals that appear in each combination of study areas. For each study site, two levels are possible—either seen (1) or not seen (−1). With three sites there are therefore $C = 2^3 = 8$ cells describing the discrete classification of individuals in the population (*e.g.*, Table 1). The population estimate, N, can then be simply derived as the total sample size of the table:

Table 1. A contingency table for the multisite photographic identification data of bottlenose dolphins for three study sites in the coastal water of NE Scotland. The cell counts refer to the number of individuals with long-lasting natural markings that appear in each $i = 1, \ldots, 8$ distinct combination of study sites, specified through study site indicators, x_{i1}, x_{i2}, x_{i3}, for sites 1, 2, and 3, respectively. These indicators take values of 1 = identified, and −1 = not identified, in each of the three study sites.

		Study site indicators		
Cell, i	Cell count, n_i	x_{i1}	x_{i2}	x_{i3}
1	1	1	1	1
2	16	1	1	−1
3	3	1	−1	1
4	35	1	−1	−1
5	2	−1	1	1
6	2	−1	1	−1
7	16	−1	−1	1
8	??	−1	−1	−1

$$N = \sum n_i \tag{1}$$

The problem here is that we only observe the counts for seven of these cells ($i = 1, \ldots, 7$), with the missing cell count, n_8, denoting the number of missed animals that were not identified in any of the sites. The population estimate therefore requires a predicted value for the missing cell, n_8. Underlying this prediction are models of the count data and the parameters generating them, not just for the observed data, but the complete set of observed and unobserved counts $n_1, \ldots, n_8$. In particular, we adopt log-linear models to model the relationships between the cells in the table, and predict an estimate into the missing cell.

We assume that the cell counts n_i are independent Poisson random variables with mean μ_i, and we model the logarithm of the Poisson mean to be an additive regression function of study area effects:

$$\log(\mu_i) = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1} x_{i2} + \beta_5 x_{i1} x_{i3} + \beta_6 x_{i2} x_{i3} \tag{2}$$

where β is the vector of unknown parameters to be estimated and the x_i's are indicator variables for the study area classifying factors in the design of the contingency table (*e.g.*, Dellaportes and Forster 1999). Specifically, $x_{i0} = 1$ for all i, and β_0 therefore corresponds to an overall mean of the counts on the log-scale. The indicators x_{i1}, x_{i2}, x_{i3} take values of either 1 or −1 according to the attribute (1 = seen, −1 = not seen), for study areas 1, 2, and 3 respectively (*e.g.*, Table 1). The corresponding parameters, β_1, β_2, β_3, therefore represent the main effect of each study area on the overall mean β_0, describing the difference between the average of the μ_i's for cells relating to each study site, and the average of all μ_i's. The terms containing products of any two of these indicators define two-way interaction effects, with β_4, β_5, β_6 describing the strength and direction of effects from pairs of study sites 1:2, 1:3, and 2:3 respectively. To ensure that all parameters in this model are identifiable, a saturated log-linear model including a three-way interaction between all study sites cannot be included in the model (King and Brooks 2001).

Inclusion of a three-way interaction would add an extra parameter, with a total of eight parameters in the vector β exceeding the number of observed cells.

This general model can be further simplified by omitting one or more of the parameters corresponding to interactions between study sites. However, interaction terms may express real features such as geographic distance effects and, therefore, the inclusion of some, or all, of these interactions may well be warranted. We therefore consider a suite of possible alternative models, differing in the inclusion of different combinations of two-way interaction effects. All models included the level of the counts (β_0), and the main effects for each study site ($\beta_1 + \beta_2 + \beta_3$), to allow for differential observation effort and sightings frequencies at the different sites. We follow the approach of Kuo and Mallick (1998) for expanding the regression equation to incorporate all possible subsets of interaction effects (β_4, β_5, β_6) by adding indicator variables as parameters. This involves the introduction of a vector δ_n of $n = 3$ binary indicator variables, to switch each of the three possible two-way interaction variables either in or out of the composite model depending on their relevance to the observed data:

$$\log(\mu_i) = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \delta_1 \beta_4 x_{i1} x_{i2} + \delta_2 \beta_5 x_{i1} x_{i3} + \delta_3 \beta_6 x_{i2} x_{i3} \tag{3}$$

Consequently, there are $2^3 = 8$ models of interest corresponding to the inclusion or exclusion of the three interaction terms in different combinations:

(1) $\delta_1 = 0,\ \delta_2 = 0,\ \delta_3 = 0$ (no interactions)
(2) $\boldsymbol{\delta_1 = 1},\ \delta_2 = 0,\ \delta_3 = 0$ (1:2 interaction only)
(3) $\delta_1 = 0,\ \boldsymbol{\delta_2 = 1},\ \delta_3 = 0$ (1:3 interaction only)
(4) $\delta_1 = 0,\ \delta_2 = 0,\ \boldsymbol{\delta_3 = 1}$ (2:3 interaction only)
(5) $\boldsymbol{\delta_1 = 1},\ \boldsymbol{\delta_2 = 1},\ \delta_3 = 0$ (1:2 + 1:3)
(6) $\boldsymbol{\delta_1 = 1},\ \delta_2 = 0,\ \boldsymbol{\delta_3 = 1}$ (1:2 + 2:3)
(7) $\delta_1 = 0,\ \boldsymbol{\delta_2 = 1},\ \boldsymbol{\delta_3 = 1}$ (1:3 + 2:3)
(8) $\boldsymbol{\delta_1 = 1},\ \boldsymbol{\delta_2 = 1},\ \boldsymbol{\delta_3 = 1}$ (1:2 + 1:3 + 2:3)

There are now two inferences of interest in this variable selection problem. First we are interested in identifying suitable log-linear models from this suite of candidates, and second, we are interested in producing estimates of population size and associated uncertainty under these models. We adopt a Bayesian approach to fully quantify both parameter and model uncertainty. Bayesian inference is probabilistic, and is based on the conditional "posterior" probability distribution of unknown parameters given the data (Gelman *et al.* 1995). However, to calculate these conditional probabilities, a joint probability distribution must first be specified to describe the relationships between all unknowns and the data. Therefore, having specified the Poisson probability model (the likelihood) that links the unknown parameters to the data, we must specify a prior probability distribution for each of the unknown parameters, and model forms.

Because the counts are modeled on the log scale, the overall mean level of the counts, β_0, can be assigned a Normal prior distribution, with mean zero and high variance (*e.g.*, $\beta_0 \sim N(0, 100)$). With a large variance this prior will be vague in the sense that it is essentially flat in the region of likely values for this analysis, and is therefore expected to have minimal effect on the analysis. In the absence of relevant prior information, the main and interaction effects can also be specified as Normally

distributed about zero with a common but unknown variance, σ^2 (*e.g.*, Dellaportes and Forster 1999, King and Brooks 2001). Prior specification therefore only requires a prior to be set on the variance parameter σ^2. In order to allow either positive or negative non-zero effects to emerge, we let prior knowledge about this variance be vague by additionally specifying a hyper-prior distribution for prior σ^2:

$$\beta_{1,\ldots,6} \sim \mathrm{N}(0, \sigma^2) \tag{4}$$

$$\sigma^2 \sim G^{-1}(c, d) \tag{5}$$

where $G(c, d)$ denotes a gamma distribution with mean c/d and variance c/d^2. To examine the sensitivity of inference to the priors assigned to these effect terms, we adopt three different hyper-prior distributions for σ^2: $G^{-1}(1,1)$, $G^{-1}(0.1,0.1)$, $G^{-1}(0.01,0.01)$.

Model uncertainty is incorporated into inference by specifying a prior distribution across the candidate models, through the specification of priors on the three indicator variables. These indicators are assigned prior distributions such that the prior probability of including any two-way interaction in the model was 0.5:

$$\delta_1, \delta_2, \delta_3 \sim \mathrm{Bernoulli}(0.5) \tag{6}$$

A uniform prior across models is thus specified in terms of the eight discrete patterns of inclusion or exclusion of each of the three interaction terms through the indicator variables.

Having completed the specification of the joint probability distribution for this model, we adopt Markov chain Monte Carlo methods (MCMC; *e.g.*, Brooks 1998) to sample the marginal posterior distributions of interest. In particular, we use the Gibbs sampling MCMC method (Casella and George 1992), which has been shown to be a simple yet versatile method for sampling from multivariate distributions. Gibbs sampling from this model can be implemented using the WinBUGS software (Lunn *et al.* 2000, Ntzoufras 2002) to produce a sequence of sampled values from the posterior distribution of log-linear parameters and derived population size for each of the eight model formulations. For the full model with indicator variables, the proportion of iterations in which a particular model formulation is selected can be monitored and used to provide an estimate of the relative probability of each model. When sampling simultaneously from this joint model space, estimates of population size are composed of MCMC samples from different models, with the proportion of samples originating from each model corresponding to the relative model probabilities (Godsill 2001). Population estimates will therefore be model-averaged in the sense that they are weighted by these model probabilities.

An Example: Multisite Photo-identification of Bottlenose Dolphins Around NE Scotland

To demonstrate this method, we analyzed photographic identification data for bottlenose dolphins collected from three study sites in the coastal waters around NE Scotland (Fig. 1). These sites, denoted 1, 2, and 3, each covered between 25 and 100 km^2, and were separated by minimum swimming distances of approximately 60 km between sites 1 and 2, 240 km between 2 and 3, and 300 km between 1 and 3. These sites were not randomly or equally spaced, but took advantage of existing

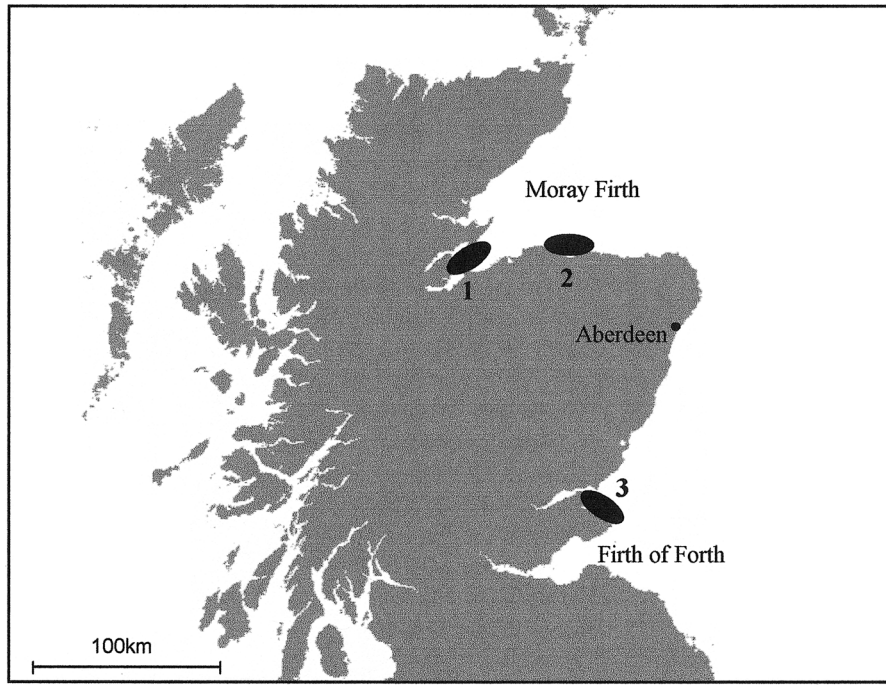

Figure 1. A map of Scotland illustrating the locations of the three coastal study areas 1, 2, and 3, marked as shaded ellipses, where bottlenose dolphin photo-identification surveys were conducted.

photo-identification studies (1 and 3), and dolphin watching activities (2). Dolphin identification photographs were collected during boat-based surveys that were conducted over a five-month period between May and September 2001. The population was assumed to be closed to births, deaths, and migration over this relatively short period, an assumption that is reasonable for this isolated population of relatively long-lived individuals (Wilson *et al.* 1999, Parsons *et al.* 2002). Logistics and budget prevented these surveys from being conducted in a coordinated strategy across sites, but rather surveys were conducted opportunistically in all three sites over the same simultaneous five-month period. The survey methodology in these sites was targeted to cover areas known to be used by dolphins, in order to maximize the number of dolphin encounters. Canon digital SLR cameras, and 35-mm SLR cameras with color film were used to photograph as many individuals as possible during each encounter. Photographic data from each of the three sites were analyzed together at a central location (Lighthouse Field Station, University of Aberdeen), where a long-term photographic catalog of individual dolphins is maintained. The analysis involved grading all images separately for both photographic quality and individual distinctiveness. Individual identifications were then assigned to only high-quality images of distinctive individuals. Full details of the photographic analysis can be found in Wilson *et al.* (1999).

In total, 1,770 high quality identification photographs were taken, resulting in the identification of 75 different individual dolphins with "long-lasting" marks

Table 2. Estimates of log-linear model parameters for the full model incorporating each possible main effect and interaction term. Estimates are presented as the mean (standard deviation) of the marginal posterior distribution for each parameter, under three different prior specifications for the variance σ^2 of the effect parameters $\beta_{1,\ldots,6}$. Prior 1 uses $\sigma^2 \sim G^{-1}(1, 1)$; prior 2 is $\sigma^2 \sim G^{-1}(0.1, 0.1)$; and prior 3 is $\sigma^2 \sim G^{-1}(0.01, 0.01)$.

Term	Parameter	Prior 1	Prior 2	Prior 3
Overall mean	β_0	1.75 (0.28)	1.80 (0.26)	1.80 (0.25)
Main effect site 1	β_1	−0.04 (0.22)	−0.02 (0.21)	−0.01 (0.21)
Main effect site 2	β_2	−0.69 (0.19)	−0.65 (0.18)	−0.64 (0.18)
Main effect site 3	β_3	−0.74 (0.26)	−0.69 (0.24)	−0.67 (0.24)
1:2 interaction	β_4	0.27 (0.29)	0.25 (0.26)	0.24 (0.25)
1:3 interaction	β_5	−0.68 (0.29)	−0.65 (0.26)	−0.65 (0.26)
2:3 interaction	β_6	−0.03 (0.27)	−0.01 (0.24)	−0.01 (0.24)

(*e.g.*, Wilson *et al.* 1999). The three study sites were sampled to different extents. Most (1,585) of the high-quality photographs were taken in site 1, with 66 in site 2, and 119 in site 3, resulting in the identification of 54, 22, and 21 different individuals in each of these sites, respectively. The overlap of individuals between sites is represented as a contingency table (Table 1). This overlap represents recaptures of individuals in a mark-recapture context, with the study site lists representing capture "occasions" defined by space instead of time.

For each of the three different prior specifications for the model parameters, the Gibbs sampling procedure was run for 100,000 iterations following a burn-in period of the same length. Inference about log-linear parameters was relatively insensitive to changes in the prior for the variance (Table 2). For all three priors used, the estimated main effect for each of the study sites ($\beta_{1,\ldots,3}$) reflected the number of individuals identified in each site, with largest estimated effect attaching to study site 1 and the smallest estimated effect for study site 3. Estimates of the interaction effects reflected the geographical proximity of the study sites, with an estimated negative interaction (β_5) between the most geographically separated sites (1 and 3), and a positive interaction (β_4) between the two closest sites (1 and 2). In contrast, there is little evidence for a strong interaction between sites 2 and 3, with the posterior distribution for β_6 centered close to zero.

Inference about model selection was also insensitive to changes in the prior for the variance of the model parameters, with the same qualitative order of model probabilities and very similar quantitative values (Table 3). The strong negative interaction between the geographically distant study sites (1:3 interaction) had a high probability of remaining in the model, with a probability of 0.94, 0.95, and 0.95 for models including this interaction term for prior specifications 1, 2, and 3 respectively. However, there was uncertainty about the best model for inference. Relatively high posterior model probabilities were estimated for both the model with only the 1:3 interaction and the model in which the negative 1:3 interaction was combined with the positive 1:2 interaction. Moderate posterior probability was assigned to the model incorporating all three of the interaction effects, but only low or negligible probability was estimated for models without the 1:3 interaction.

Inference about the size of the population with long-lasting marks was highly dependent on the chosen model, but remained insensitive to changes in the prior for the variance. Models with the strong negative 1:3 interaction generally produced lower estimates of population size than models not including this effect,

Table 3. Posterior model probabilities, *P* (to two decimal places), and summary statistics for population size of dolphins possessing long-lasting markings, *N*. Estimates are presented for eight log-linear models corresponding to the inclusion of different sets of interaction terms between study sites 1, 2, and 3, along with an estimate of population size averaged across all models. Data are presented for the posterior median (M) and the 95% probability interval (PI) of values encompassing 95% of the posterior density, under the three different prior specifications. Prior 1 uses $\sigma^2 \sim G^{-1}(1, 1)$; prior 2 is $\sigma^2 \sim G^{-1}(0.1, 0.1)$; and prior 3 is $\sigma^2 \sim G^{-1}(0.01, 0.01)$.

	Prior 1			Prior 2			Prior 3		
		N			*N*			*N*	
Model	*P*	M	95% PI	*P*	M	95% PI	*P*	M	95% PI
No interaction	0.00	113	92–151	0.00	112	92–149	0.00	111	91–148
1:2	0.04	211	126–473	0.03	206	123–467	0.03	205	123–467
1:3	0.34	83	76–99	0.33	83	76–99	0.34	83	76–99
2:3	0.00	111	91–149	0.00	110	90–147	0.00	109	90–146
1:2, 1:3	0.32	93	77–179	0.33	91	77–165	0.32	91	77–162
1:2, 2:3	0.02	279	135–1,087	0.01	263	130–1,013	0.01	262	129–999
1:3, 2:3	0.17	81	75–96	0.17	81	76–96	0.17	81	76–95
1:2, 1:3, 2:3	0.11	90	76–270	0.12	90	76–221	0.12	90	76–207
Average	—	85	76–263	—	85	76–221	—	85	76–224

demonstrating the consequences of ignoring this negative interaction. Conversely, models that included the positive 1:2 interaction generally produced higher estimates of population size than the corresponding models with this term absent. These between-model differences, along with the uncertainty about the choice of model, were reflected in the posterior estimates of population size produced by averaging across all candidate models (Table 3). The model-averaged estimate of population size of 85 (95% probability interval = 76–263) can be viewed as a compromise estimate, which shrinks the estimates from the individual models towards a common value. This compromise is weighted, with the magnitude and directionality of this shrinkage relating to the relative model probabilities and their associated population estimates. The model-averaged 95% probability interval was generally wider than the interval for individual models, as this estimate was obtained by sampling across all candidate models and therefore accounted for model uncertainty. However, this model-averaged 95% probability interval did not cover the widest interval of all models combined, because little posterior probability was assigned to the models with very high upper bounds, and therefore the posterior distribution for population size was sampled from these models on only a small proportion of iterations.

Discussion

This example demonstrates the utility of this multisite approach for estimating the size of wide-ranging cetacean populations. By using multiple study areas from throughout the population's range, and by estimating geographical dependencies between study sites, we can directly address the problems caused by wide-ranging and heterogeneous movement patterns, when animals are more likely to be encountered in some areas than others. This sampling method therefore provides

an alternative to the logistically difficult approach of surveying across the full range of the population. Our model-averaged estimate for population size in this example of around 85 (95% probability interval = 76–263) is very similar to the maximum likelihood estimate of 80 (95% confidence interval = 66–113) for the number of animals with long-lasting marks in this population produced by Wilson *et al.* (1999), who employed a rigorous and standardized survey methodology and a complex mark-recapture model that allowed for individual heterogeneity. The multisite approach that we have introduced allows us to produce similar estimates when survey effort is more opportunistic, and limited to a number of geographically discrete sites. By estimating geographical dependencies between sites, this method enables useful population estimates to be obtained from data collected from surveys located for practical convenience rather than by random design.

Alternative mark-recapture models do exist for the situation when captured individuals can be categorized into one of multiple states, where state could refer to a geographical site. Darroch (1961) proposed a "stratified Petersen" estimator where marking and recovery can be stratified by both space and time (*e.g.*, Schwarz and Taylor 1998). However, application of this estimator to geographical strata requires coordinated capture and recovery periods in multiple sites, which may not be possible when making use of opportunistic cetacean photo-identification effort that cannot easily be co-ordinated across sites. Additionally, this approach is a stratified version of the simple Lincoln-Petersen estimator (Seber 1982), which does not account for correlations between capture and recapture probabilities that may exist through geographical dependencies (Schwarz and Taylor 1998). More recently developed "multistrata" mark-recapture models (*e.g.*, Brownie *et al.* 1993, Schwarz *et al.* 1993) allow movement to be modeled directly through the estimation of transition probabilities between strata, which can be geographical sites. However, estimation of transition probabilities using this approach also requires coordinated sampling efforts at all sites over a series of survey periods. Additionally, such multistrata mark-recapture approaches are primarily intended to estimate transition probabilities between states and capture probabilities of animals within each state, but not population size. Using such methods, it may be possible to use state-specific capture probabilities to estimate abundance of animals in each study area strata at a particular time (*e.g.*, Whitehead 2001), but this falls short of an estimate of population size. This is particularly true when the study areas cover only a very small part of the total population's range, and therefore individuals can also transit to unobserved areas. In contrast to these methods, the multisite mark-recapture approach that we have introduced is directly focused on producing estimates of population size from opportunistic data collected simultaneously in multiple sites. Rather than requiring a complicated sampling design, our approach is based on established methods for the analysis of simple contingency tables (Fienberg 1972), enabling log-linear models to account for dependencies between geographical sites.

The practicality of this multisite approach is enhanced by the adoption of modern Bayesian statistical approaches. Bayesian methods have been repeatedly advocated as well suited for the analysis and communication of uncertainty in ecological data analysis (Ellison 1996, Wade 2000, Link *et al.* 2002). Here we have demonstrated how this utility extends to model determination, which can be based directly on the probability of competing models, estimated using MCMC methods (*e.g.*, Kuo and Mallick 1998). However, because inference about population size can be very

sensitive to the choice of log-linear model, emphasis should be placed on accounting for model-selection uncertainty (Buckland *et al.* 1997), especially where study design has not been tailored to suit a specific model. The Bayesian MCMC approach also provides a straightforward framework for incorporating model selection uncertainty directly into inference through model averaging. The model-averaged estimate of population size communicates the full extent of the uncertainty from both sparse data and model selection choices, which likely explains the considerably larger 95% probability intervals from our Bayesian estimate compared to the 95% confidence intervals from the single model employed by Wilson *et al.* (1999). All the analyses presented here can be implemented using the freely available WinBUGS software (Lunn *et al.* 2000, Link *et al.* 2002, Ntzoufras 2002), and program codes to perform these analyses can be obtained from the primary author.

There are certain practical limitations to the utility of this multisite mark recapture approach. The production of unbiased population estimates using this approach requires validation of the usual set of closed population mark-recapture assumptions (Hammond 1986). Notably, although the multisite method will account for much of the heterogeneity in capture probabilities due to movement, individuals will likely posses inherently different movement and capture probabilities that are not completely captured by the model. Several other mark-recapture approaches have been proposed for modeling such individual heterogeneity (Chao 2001), however these approaches generally model heterogeneity while ignoring sample dependencies. The incorporation of both dependencies and individual heterogeneity remains one of the most pressing problems in mark-recapture population analysis. The multisite approach also relies on the assumption that all animals in the population have some chance of occurring in a sample from at least one of the sites. That this assumption has been met may often be difficult to determine, but we suggest they it can be best achieved by careful location of the sampling sites to cover at least the known ranges of the population. If a population shifts its range (*e.g.*, Wilson *et al.* 2004), or as knowledge about a population's range changes, study sites can be moved or added to maintain effective penetration of the sampling into the target population. King and Brooks (2001) have presented a method to assess the utility of additional data sources for population estimation from multiple sources. This type of procedure could also prove useful for guiding survey design to facilitate the monitoring of wide-ranging cetacean populations using multisite mark-recapture.

Acknowledgments

JWD was supported by a postgraduate scholarship awarded by the Carnegie Trust for the Universities of Scotland; DAE was funded by the Scottish Executive Environment and Rural Affairs Department. The dolphin case study was supported by the Whale and Dolphin Conservation Society; the Moray Firth Wildlife Centre, ChevronTexaco and Talisman Energy (UK) Ltd, and assisted by Tim Barton, Alice Mackay, Kate Grellier, Karl Nielson, and Tony Archer. We thank two anonymous reviewers for helpful comments.

Literature Cited

Brooks, S. P. 1998. Markov chain Monte Carlo method and its application. The Statistician 47:69–100.

BROWNIE, C., J. E. HINES, J. D. NICHOLS, K. H. POLLOCK AND J. B. HESTBECK. 1993. Capture-recapture studies for multiple strata including non-Markovian transitions. Biometrics 49:1173–1187.

BUCKLAND, S. T., K. P. BURNHAM AND N. H. AUGUSTIN. 1997. Model selection: An integral part of inference. Biometrics 53:603–618.

CASELLA, G., AND E. I. GEORGE. 1992. Explaining the Gibbs sampler. The American Statistician 46:167–174.

CHAO, A. 2001. An overview of closed capture-recapture models. Journal of Agricultural, Biological and Environmental Statistics 6:158–175.

CORMACK, R. M. 1989. Log-linear models for capture-recapture. Biometrics 45:395–413.

DARROCH, J. N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. Biometrika 48:241–260.

DELLAPORTES, P., AND J. J. FORSTER. 1999. Markov chain Monte Carlo model determination for hierarchical and graphical log-linear models. Biometrika 86:615–633.

ELLISON, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications 6:1036–1046.

FIENBERG, S. E. 1972. The multiple recapture census for closed populations and incomplete 2^k contingency tables. Biometika 59:591–603.

GELMAN, A., J. B. CARLIN, H. S. STERN AND D. B. RUBIN. 1995. Bayesian data analysis. Chapman and Hall, London, UK.

GODSILL, S. J. 2001. On the relationship between Markov chain Monte Carlo methods for model uncertainty. Journal of Computational and Graphical Statistics 10(2):1–19.

HAMMOND, P. S. 1986. Estimating the size of naturally marked whale populations using capture-recapture techniques. Report of the International Whaling Commission (Special Issue 8):253–282.

HAMMOND, P. S., R. SEARS AND M. BERUBE. 1990. A note on problems in estimating the number of blue whales in the Gulf of St. Lawrence from photo-identification data. Report of the International Whaling Commission (Special Issue 12):135–141.

KING, R., AND S. P. BROOKS. 2001. On the Bayesian analysis of population size. Biometrika 88:317–336.

KUO, L., AND B. MALLICK. 1998. Variable selection for regression models. Sankhya B 60:65–81.

LINK, W. A., E. CAM, J. E. NICHOLS AND E. G. COOCH. 2002. Of BUGS and birds: Markov chain Monte Carlo for hierarchical modelling in wildlife research. Journal of Wildlife Management 66:277–291.

LUNN D. J., A. THOMAS, N. BEST AND D. SPIEGELHALTER. 2000. WinBUGS—a Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10:325–337.

MADIGAN, D., AND J. C. YORK. 1997. Bayesian methods for estimation of the size of a closed population. Biometrika 84:19–31.

NTZOUFRAS, I. 2002. Gibbs variable selection using BUGS. Journal of Statistical Software 7(7). Available at www.jstatsoft.org.

PALSBØLL, P. J., J. ALLEN, M. BÉRUBE, P. J. CLAPHAM, T. P. FEDDERSEN, P. S. HAMMOND, R. R., HUDSON, H. JØRGENSEN, S. KATONA, A. H. LARSEN, F. LARSEN, J. LIEN, D. K. MATTLIA, J. SIGURJONSSON, R. SEARS, T. SMITH, R. SPONER, P. STEVICK AND N. OIEN. 1997. Genetic tagging of humpback whales. Nature 388:767–769.

PARSONS, K., L. NOBLE, B. REID AND P. THOMPSON. 2002. Mitochondrial genetic diversity and population structuring of UK bottlenose dolphins (*Tursiops truncatus*): Is the NE Scotland population demographically and geographically isolated? Biological Conservation. 108:175–182.

SCHWARZ, C. J., J. F. SCHWEIGERT AND A. N. ARNASON. 1993. Estimating migration rates using tag-recovery data. Biometrics 49:177–193.

SCHWARZ, C. J., AND C. G. TAYLOR. 1998. Use of the stratified-Petersen estimator in fisheries

management: Estimating the number of pink salmon (*Oncorhnchus gorbuscha*) spawners in the Fraser River. Canadian Journal of Fisheries and Aquatic Science 55:281–296.

SEBER, G. A. F. 1982. The estimation of animal abundance and related parameters, Second edition. Charles Griffin and Company Ltd., London, UK.

WADE, P. R. 2000. Bayesian methods in conservation biology. Conservation Biology 14:1308–1316.

WHITEHEAD, H. 2001. Analysis of animal movement using opportunistic individual identifications: Application to sperm whales. Ecology 82:1417–1432.

WHITEHEAD, H., R. PAYNE AND M. PAYNE. 1986. Population estimate for the right whales off Peninsula Valdes, Argentina, 1971–1976. Report of the International Whaling Commission (Special Issue 10):169–171.

WILSON, B., P. S. HAMMOND AND P. M. THOMPSON. 1999. Estimating size and assessing trends in a coastal bottlenose dolphin population. Ecological Applications 9:388–300.

WILSON, B., R. J. REID, K. GRELLIER, P. M. THOMPSON AND P. S. HAMMOND. 2004. Considering the temporal when managing the spatial: A population range expansion impacts protected areas based management for bottlenose dolphins. Animal Conservation. 7:331–338.

Received: 10 February 2004
Accepted: 10 July 2004