Bayesian estimate of Australian humpback whale calving interval under sparse resighting rates: 1987–2009

ROBERT WILLIAM RANKIN[1,2], DANIELA MALDINI[3] AND GREG KAUFMAN[1]

Contact e-mail: robertw.rankin@gmail.com

ABSTRACT

This study estimates a calving interval for humpback whales from a longterm photo-ID catalogue of 2,973 individuals resighted in Hervey Bay, East Australia. The study proposes a modification of two existing methods to handle partial identification of sex and age-classes of whales from visual surveys. One method truncates the data to just breeding females and discards all resighting events prior to the first observed breeding event. The second method utilises the multi-stage mark recapture (MSMR) framework and multi-event extension to include all resighted individuals and their entire encounter history. The performance of each method is assessed and the conditioning required to handle ambiguity of sex and age-classes is detailed, which is subtly different from most other mark-recapture methods. Both truncation and the multi-event methods led to similar estimates of calving intervals: 2.98 years (95% CI: 2.27–3.51) and 2.78 years (95% CI: 2.23–3.68) respectively. More importantly, estimates were more sensitive to the exact specification of resighting probabilities among age and sex classes than to the type of conditioning. However, the multi-event framework resulted in more precise estimates of other important life-history parameters such as apparent survival, and included a wider constituency of age and sex classes.

KEYWORDS: HUMPBACK; MODELLING; REPRODUCTION; BIRTH RATE; PHOTO-ID; AUSTRALASIA; MARK-RECAPTURE

INTRODUCTION

There is increasing evidence that estimates of life-history and demographic parameters may be as much a consequence of sampling and observational attributes as they are a reflection of the biological processes being measured (Clark and Gelfand, 2006; Cressie *et al.*, 2009; Halstead *et al.*, 2012; Pradel *et al.*, 2005) especially for sight-resight models of cetacean life-history parameters, such as calving intervals. Such studies are usually boat-based and employ photo-ID methods to collect encounter histories from often large, sparse populations. While crucial for the understanding of cetacean demographics, such studies suffer from a basic mismatch between 'terrestrial' observer and sub-surface subject, leading to predictable consequences for photo-ID data. For example, our dataset of the East Australia Group V Stock (EAGVS) of humpback whales (*Megaptera novaeangliae*) faces a number of sampling issues and analytical challenges, including:

(1) a large open population, leading to low resighting rates;

(2) different movement patterns among age and sex classes leading to skewed sex and age composition of sighted individuals (Craig and Herman, 2000; Craig and Herman, 1997; Forestell *et al.*, 2003); and

(3) infrequent sex-specific behaviours or visual cues of age (e.g. singing behaviour or genital photos) leading to low certainty about age and sex (Glockner-Ferrari and Ferrari, 1990; Glockner, 1983; Tyack, 1981; Winn and Winn, 1978).

Such ambiguity can result in biased estimates of life-history parameters (Kendall *et al.*, 2003) or highly variable estimates

depending on how non-identified individuals are treated versus other known classes (Hoffman *et al.*, 2010). The problem is further exasperated by decreasing resight rates in a growing population (Forestell *et al.*, 2011; Noad *et al.*, 2011), whereby more individuals results in fewer opportunities to resight the same individual. This is especially important to cetacean studies, which often depend on repeat observations to ascribe sex, e.g. getting positive confirmation of females' sex when observed with a calf, or declaring 'putative males' as true males based on never being seen with a calf.

For the EAGVS humpback whales, certainty of age and sex is common for only one type of observational event: females in mother-calf pairs, who are known to be in a breeding state. In contrast, only a fraction of males and females without calves can be confidently assigned to the adult class or the non-breeding portion of the population. Likewise, size-based classification of subadults (Clapham, 2000) is sensitive to distance from the boat, sea state and the presence of other animals for size comparison, leading to potential misclassification. Usually, this leads to a large portion of individuals being classed as 'unknowns'. This designation contains an unknown mixture of subadults, males and females observed only in the non-breeding state.

The goal of this study was to estimate an overall calving interval for the EAGVS humpback whales, under two different treatments of the aforementioned difficulties in sex and class designation. Calving intervals, and the inverse parameter, birth-rate, are critical to understanding the reproductive behaviour of long-living, uniparous mammals. They may also be important at the individual level to understand body condition (Wiley and Clapham, 1993) as well as important for population-level processes by

[1]*Pacific Whale Foundation, 300 Maalae Rd., Wailuku, HI 96793, USA.*
[2]*CSS-Dynamic Inc., 10301 Democracy Lane, Suite 300, Fairfax, VA 22030.*
[3]*Okeanis, P.O. Box 853, Moss Landing, CA 95039, USA.*

influencing the population rate-of-increase (Brandão *et al.*, 2000; Zerbini *et al.*, 2010).

Here two previously posited methods to estimate cetacean breeding propensity and calculate calving intervals are evaluated. The first method (hereafter referred to as the 'truncation method') was introduced by Barlow and Clapham (1997) to study humpback whales in the Gulf of Maine. The 'truncation' refers to conditioning the likelihood on a female's first observed birthing event, with subsequent elimination of: (1) females' encounter histories prior to their first observed breeding event; as well as (2) the encounter histories of all animals who have not been observed breeding, such as non-reproductive females, males, sub-adults and all other non-sexed/non-aged individuals. Such truncation is considered unbiased (Conn and Cooch, 2009), but under extremely low resight rates there is little assurance that the females observed with calves are a random sample of the population. Rather, heterogeneity of female breeding propensity (Rosenbaum *et al.*, 2002) may result in a truncated sample of higher-frequency breeders and thereby bias model estimates. Truncation also discards other potentially important life-history information, such as male and subadult mortality. The benefit of the truncation method is that it is relatively simple to implement, with few model parameters. In its original specification, Barlow and Clapham (1997) ignored the issue of resight probability, but the framework is easily extended to include imperfect detection, as is done in this study.

A second common method to model breeding propensity is the multi-state mark recapture (MSMR) framework (e.g. North Atlantic right whale *Eubalaena glacialis*; Fujiwara and Caswell, 2002). The method is advantageous as it can include individuals' entire encounter histories through a stage-structured transition matrix and thereby provide estimates of many important life-history parameters. Here, the conditional birth-interval probabilities (γ) of Barlow and Clapham (1997) are equivalent to the transition parameters for females moving between the breeding and non-breeding states. Crucially, the method is not conditioned on a female's first observed birthing event, and can therefore include all sex and ages classes, if all states are known without error. When there is misclassification and age/sex ambiguity, the 'multi-event' extension is necessary, which parses observational errors into two distinct unknowns (Nichols *et al.*, 2004; Pradel *et al.*, 2005): (1) uncertainty of detection, i.e. not knowing whether a cow has given birth or not because she has not been resighted; and (2) uncertainty of state from misclassification and partial identifiability of states, i.e. not knowing the sex or age-class of an individual who has been resighted (Conn and Cooch, 2009).

These two methods are investigated through simulations and analyses of a long-term humpback whale photo-ID dataset. The study begins with a matrix-based reformulation of the Barlow and Clapham (1997) truncation method, then assesses its performance under low or differential re-sighting probabilities between breeding and non-breeding females. Next the method is validated by re-examining the Gulf of Maine humpback whale dataset (Barlow and Clapham, 1997). To address the main objective of this study, a calving interval for EAGVS individuals sighted in Hervey Bay is estimated using model selection among 22 different models under both the truncation and the multi-event methods. Finally, select multi-event models are reformulated within a hierarchical framework to model heterogeneity of breeding probabilities, including simulations to evaluate the performance of both the truncation and multi-event framework under different levels of individual heterogeneity.

METHODS

Field methods

EAGVS humpback whales migrate annually along the eastern coast of Australia from their breeding grounds along the Great Barrier Reef to their austral summer feeding areas in the Southern Ocean (Chaloupka and Osmond, 1999). From 1987 to the present, we have conducted boat-based humpback whale surveys in Hervey Bay, Queensland (WGS1984 25°15.5'S, 152°51.7'E) between June and November. Hervey Bay is located in the southern portion of the EAGVS breeding grounds where migrating mother-calf pairs are typically observed in late-August to mid-September. The present study includes data collected since 1987 with the exclusion of 2001 and 2003 when no research was conducted in this area.

A detailed description of data collection and processing protocols are given in Kaufman *et al.* (1993), Chaloupka *et al.* (1999), Forestell *et al.* (2011) and Forestell *et al.* (2003). During whale encounters, adults (>12m; Clapham, 2000) were categorised as being male, female or unknown by visual observers or using photographic evidence. A breeding female was identified by its close proximity to an individual less than 6m in length (i.e. a calf). Males were identified by photographic evaluation of the genital slit or by the presence of singing behaviour (Tyack, 1981; Winn and Winn, 1978) or 'escorting' behaviour (Glockner-Ferrari and Ferrari, 1990; Glockner, 1983).

Fluke photograph quality and distinctiveness were graded by modified protocols for North Pacific humpback whales (Calambokidis *et al.*, 2008). Each fluke was given a score of 1 through 5 (good to bad) for five criteria: (1) proportion of fluke visible; (2) fluke angle; (3) photographer/lateral angle; (4) focus; and (5) exposure/contrast. A cumulative score was calculated by summing the scores of the five criteria. Flukes were not considered for analysis if they: (1) had a cumulative score exceeding 14; or (2) exceeded a score of 4 for focus or angle (criteria 4 and 5); or (3) did not have a picture of the fluke's central notch.

Simulation and analysis

Imperfect detection and the truncation method

The R language (R Development Core Team, 2010) was used to generate 1,000 simulations of 1,000 female encounter histories over 15 years, using assumptions and parameters similar to the system studied by Barlow and Clapham (1997) i.e. zero mortality, birthing is only conditional on the time since the last breeding event and a maximum of five years are possible between birthing events. Calving histories were simulated based on similar unconditional birth-probabilities estimated in their study (0.046, 0.568, 0.318, 0.05, 0.017, for intervals spanning one to five years). Each five-year history represents a draw from a multinomial distribution of 32 possible five-year histories (year zero corresponding to a birthing event).

Imperfect detectability was simulated under nine different

scenarios representing 3×3 combinations of differential resight probabilities between females-with-calves (p_C) and without-calves (p_N): $p_C = 0.95, 0.5$ and 0.24, while p_N was a logit-scale function of the p_C, such that:

$$p_{N,t} = \frac{p_{C,t}e^{\delta}}{1 + p_{C,t}(e^{\delta} - 1)} \qquad (1)$$

where δ varied with values $\log(1)$ for equal detection, $\log(0.5)$ and $\log(0.25)$ for lower detection. All probabilities were equal across years ($p_{N,t} = p_{N,t+1}$), but later in the analyses, time-varying detection probabilities were allowed (hence the t subscripts in equation 1). The multinomial likelihood function was constructed in native R language, and was maximised with respect to the five conditional birth-interval probabilities using the 'optim' function. For each of 1,000 simulations, a calving interval was calculated based on the weighted-sum of intervals 1 through 5, whereby the weights were the unconditional birth-interval probabilities for each interval:

$$\sum_{t=1}^{5} t\gamma_t \left(\Pi_{y=1}^{t-1}(1 - \gamma_y) \right) \qquad (2)$$

where γ_t are the conditional birth-interval probabilities for interval t.

A matrix-based reformulation of the truncation method
Barlow and Clapham (1997) had the luxury of near-complete encounter histories for every animal in their study, and so could safely ignore imperfect detection and mortality. Doing so led to just 32 possible encounter histories. However, the burden of proof has since shifted to assume that detection probability is neither perfect nor homogeneous in ecological studies (MacKenzie and Kendall, 2002). The matrix formulation of stage-structured populations provides a natural means to incorporate every permissible state when a female has not been seen, thereby automating the likelihood calculation (Fujiwara and Caswell, 2002). Under this framework, the analysis can be performed in readily available open-source software such as OpenBUGS or E-SURGE (Choquet *et al.*, 2009). This is an important improvement on the Barlow and Clapham likelihood, which requires custom programming. Later in this study, the matrix-based method is validated by reanalysing the 1997 data from the Gulf of Maine.

Recast as matrices, the conditional birth-interval probabilities are transitions in a 7×7 row-stochastic matrix Γ, consisting of 1 breeding state, 5 non-breeding states and 1 final state for death, for a total of 7 states.

$$\Gamma = \begin{pmatrix} \gamma_1 & 1-\gamma_1 & 0 & 0 & 0 & 0 & 0 \\ \gamma_2 & 0 & 1-\gamma_2 & 0 & 0 & 0 & 0 \\ \gamma_3 & 0 & 0 & 1-\gamma_3 & 0 & 0 & 0 \\ \gamma_4 & 0 & 0 & 0 & 1-\gamma_4 & 0 & 0 \\ \gamma_5 & 0 & 0 & 0 & 0 & 1-\gamma_5 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

A distinction is made between the observed event (o) and the true state (z), the mapping of which is done with the matrix B. The observed events correspond to the three rows of B: row 1 is 'seen with calf', row 2 is 'seen without calf' and row 3 is 'unobserved'. The observed event 1 ('seen with calf') is unambiguously associated with the 'breeding female' state ($z = C$), but the observed event 2 ('seen without calf'), can be one of five possible non-breeding states (N_1–N_5; columns 2 through 6). Since only adult females are being considered in the truncation method, each N represents a non-breeding female at a different time-since-last-birthed. Each of these states has the option of moving into the next non-breeding interval or transitioning back into the breeding state with independent probabilities (γ_1–γ_5). Death is row 7, known as an absorbing state. Apparent survival is a 7×7 row-stochastic matrix Φ, and resight probabilities are parameterised as column-stochastic matrix P_t.

$$\Phi = \begin{pmatrix} \phi_F & 0 & 0 & 0 & 1-\phi_F \\ 0 & \phi_F & 0 & 0 & 1-\phi_F \\ & & \ddots & & \vdots \\ 0 & 0 & 0 & \phi_F & 1-\phi_F \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P_t = \begin{pmatrix} p_{C,t} & 0 & 0 & 0 & 0 \\ 0 & p_{N,t} & 0 & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & 0 & p_{N,t} & 0 \\ 1-p_{C,t} & 1-p_{N,t} & \cdots & 1-p_{N,t} & 0 \end{pmatrix}$$

It should be noted that the five-year maximum breeding interval was an empirical suggestion from Barlow and Clapham and was specific to their study (Barlow and Clapham, 1997). In the above matrix formulation, this maximum is explicitly coded by the value of 1 in element $\Gamma[1,6]$, i.e. females have a 100% probability of breeding again after five years. This maximum-interval is not fundamental to the truncation method and could be increased, decreased, or recast as a simple two-state Markovian system (see later), as is common in other stage-structured analyses of breeding states (Fujiwara and Caswell, 2002). An auxiliary interest of this study was to see how well the data could resolve estimates of γ_4 and γ_5, which are based on rare and difficult to observe events such as a female not breeding for five or more years. This system will be hereafter referred to as the 'memory' specification of Γ, to distinguish it from the two-state 'memoryless' Γ described later.

An individual's encounter history ($h|\gamma$) and all permissible unseen states can then be modelled with the use of an appropriate matrix multiplication algorithm, as suggested by Pradel *et al.* (2005):

$$L(h_i \mid o_{(e_i+1)}) = \left(\prod_{t=e_i+1}^{T} \Phi\Gamma D\big((P_t B)[o_t,\cdot]\big) \right) 1_7 . \qquad (3)$$

where e_i is the time at which individual i was first sighted with a calf; $P_t B$ is the matrix product of time-varying resight probability matrix P_t and B; $(P_t B)[o_t,\cdot]$ is the row vector corresponding to event o_t of the matrix $(P_t B)$; $D(\theta)$ is the

matrix with diagonal elements equal to an arbitrary vector θ and 1_7 is a column vector of 7 ones.

Using matrices, the calving interval can be derived in a more general way than equation 2 through the use of eigenvectors, and proceeds as follows: first, calculate the dominant positive eigenvector of $\Gamma[-7,-7]$ (i.e. the asymptotic distribution of states, minus the absorbing state 7); second, calculate the calving interval as the inverse of the ratio between element 1 (breeding) versus elements 2 through 6 (non-breeding) of the eigenvector.

Under this matrix formulation, the birth-interval probabilities are the same as estimated in Barlow and Clapham (1997) because both are conditioned on a female's first observed birthing event. To demonstrate this equivalency, the Barlow and Clapham Gulf of Maine humpback whale dataset (provided in their appendix) were reanalysed, plus inclusion of time-varying resight probabilities for both females-with-calves and females-without-calves, as well as including estimates of apparent survival. The likelihood was calculated using the RcppArmadillo matrix multiplication library in C++ and Rcpp (Eddelbuettel and Francois, 2011).

To use the matrix-based truncation method on the EAGVS dataset, the catalogue was reduced from 2,973 individuals to just 435 females sighted at least once with a calf. Six different models were analysed: equal detectability between females with and without calves ($p_{N,t} = p_{C,t}$), fully independent resight rates ($p_{*,t}$), or a logit-link between the two states ($p_{N,t} = f(p_{C,t})$) according to equation 1 where resight probabilities differed by the same amount each year on the logit-scale (δ). These three scenarios were also considered in the context of five non-breeding states versus six non-breeding states (i.e. expanding matrix Γ to 8×8 dimensions), representing rows 1–3 and 4–6 in Table 3, respectively. In all cases, birth-interval probabilities and apparent survival were time-invariant, while the resight probabilities were fully time-varying, as per Forestell *et al.* (2011). The latter was motivated by the idea that individuals should be more difficult to resight under a rapidly growing population (Noad *et al.*, 2011) and resight probabilities should therefore be different every year.

Markov chain Monte Carlo (MCMC) techniques were used under a Bayesian inferential paradigm to sample from the posterior distributions of birth-interval probabilities. The SCAM adaptive Metropolis-within-Gibbs algorithm was used to optimise proposal densities and ensure convergence of the chains (Haario *et al.*, 2005; Roberts and Rosenthal, 2009). Proposal densities were computed from univariate Normals on the logit-scale. This facilitated the adaptive MCMC algorithm and is not uncommon in other component-wise Bayesian samplers (Hall, 2012). Likewise, prior densities of resight and birthing probabilities were specified on the logit-scale with Normal densities *Norm(0, 1.682)*, such that the priors were flat to slightly concave on the probability scale (Gelman *et al.*, 2008) and helped to avoid chain exploration of extreme values on the logit scale. Apparent survival was specified with a slightly more informative prior of *Norm(logit(0.94), 1)*, based on the central tendency of 16 humpback whale studies reviewed by Zerbini *et al.* (2010). The large variance of the priors ensured that estimates were mostly driven by the likelihood.

Model parsimony and goodness-of-fit were compared by the Deviance Information Criterion (DIC; Spiegelhalter *et al.*, 2002) and posterior predictive checks (Gelman *et al.*, 1996; Gelman and Shalizi, 2012) in addition to assessing whether the apparent survival and calving intervals were biologically reasonable. For the posterior predictive check, we simulated each individuals' history conditioned on the year of first having been observed birthing, then compared simulated data versus observed data using the deviance function as the discrepancy statistic, as similarly used in other stage-structured marine mammal studies (Chilvers *et al.*, 2010).

The MSMR multi-event framework

The above reformulation of the Barlow and Clapham method (adding resighting probabilities, apparent survival and variable number of non-birthing states) does not change its fundamental conditioning on the first observed birthing event. This is distinctive from other MSMR frameworks which are conditioned on individuals' first capture event, regardless of state and therefore include more observations and more individuals in the model.

MSMR also assumes certainty of state: lacking such certainty requires the multi-event extension. The multi-event allows the inclusion of full encounter histories of females (conditioned on first capture) as well as all other individuals in the study (Avril *et al.*, 2012; Conn and Cooch, 2009; Pradel *et al.*, 2005). This is made possible by incorporating two additional observational processes into the traditional MSMR: (1) the process of assigning individuals seen in events ($o_{i,t}$) to different states z (the mapping of which is handled by matrix B); and (2) the initial probability of encountering an individual in different states z (matrix Π_t). For the EAGVS dataset, six events were considered, which correspond to the rows of matrix B: row 1 = 'identified as a subadult'; 2 = 'adult seen with a calf'; 3 = 'identified as a non-breeding adult female'; 4 = 'identified as an adult male'; 5 = 'unidentified adult' or 'unknown adult'; and 6 = 'unobserved'. Five true states corresponding to the columns of all matrices: column 1 = 'subadult' (J), 2 = 'breeding female' (C), 3 = 'non-breeding adult female' (N), 4 = 'adult male'(M) and 5 = 'dead'. Individuals in state z are seen in event o with probabilities: β_J = successfully seen as a subadult; β_A = successfully seen as an adult (regardless of sex); β_F = successfully seen as a non-breeding female (i.e. a genital photograph); and β_M = male (observed singing or confirmed with a genital photograph). Here, the probability of identifying a mother-calf pair as a breeding female is 1 (unlike in Kendall *et al.*, 2003) because calves maintain a close physical proximity to their mothers.

$$B = \begin{pmatrix} \beta_J & 0 & (1-\beta_A) & (1-\beta_A) & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \beta_A\beta_F & 0 & 0 \\ 0 & 0 & 0 & \beta_A\beta_M & 0 \\ 1-\beta_J & 0 & \beta_A(1-\beta_F) & \beta_A(1-\beta_M) & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Gamma = \begin{pmatrix} 1-\gamma_J & 0.5\gamma_J & 0 & 0.5\gamma_J & 0 \\ 0 & \gamma_1 & 1-\gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 1-\gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Phi = \begin{pmatrix} \phi_J & 0 & 0 & 0 & 1-\phi_J \\ 0 & \phi_F & 0 & 0 & 1-\phi_F \\ 0 & 0 & \phi_F & 0 & 1-\phi_F \\ 0 & 0 & 0 & \phi_M & 1-\phi_M \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P_t = \begin{pmatrix} p_{J,t} & 0 & 0 & 0 & 0 \\ 0 & p_{C,t} & 0 & 0 & 0 \\ 0 & 0 & p_{N,t} & 0 & 0 \\ 0 & 0 & 0 & p_{M,t} & 0 \\ 1-p_{J,t} & 1-p_{C,t} & 1-p_{N,t} & 1-p_{M,t} & 0 \end{pmatrix}$$

Apparent survival is parameterised as matrix Φ, with different values for subadults, females and males. State-dependent and time-varying resight probabilities are denoted by P_t. The transition matrix Γ is considered to have a nuisance 'maturation' parameter γ_J (a subadult transitions to the breeding population, with 50% becoming breeding females and 50% becoming males) and two female states: non-breeding and breeding. This specification of Γ is slightly different from that of the truncation method and will hereafter be referred to as the 'memoryless' matrix specification of Γ. It is 'memoryless' because females may transition to a breeding or non-breeding state based only on their previous state and not upon their states up to five years previously. The memoryless specification is not an essential feature of the multi-event or MSMR framework and Γ could have been expanded to include as many non-breeding states as in the original Barlow and Clapham parameterisation. The memoryless model is more common in MSMR and was used here due to counter-intuitive point estimates from the memory specification (discussed later).

The second important feature of the multi-event extension is the parameterisation of Π, defined by Pradel (2005) as 'the probability of being in state $[z_e]$ when first encountered [at time e].'

$$\Pi_t = \begin{pmatrix} \pi_{J,t} & 1 & (1-\pi_{J,t})\pi_F & 1-\pi_{J,t}-\pi_F & 0 \end{pmatrix}$$

Like the Γ and Φ matrices, it has a column for each state z, but unlike them, its row does not sum to one. Rather, our simulations (below) demonstrated that the values of π_C must be set to one (i.e. a female initially encountered with a calf is known 100% to be in the breeding state), while π_J, π_N and π_M are stochastic and together sum to one. Death is by definition equal to a zero encounter probability. Under this formulation, one can interpret Π to weight the resight history ($t > e_i$) by the probability that individual i was originally in one of the three possible states (akin to the weightings suggested in Fujiwara and Caswell, 2002).

The likelihood is similarly calculated as in equation 3, but now the initial sighting event o_i at time $t = e_i$ is explicitly modelled:

$$L(h_i \mid o_{(e_i)}) = \Pi_{t=e_i} D\big(B[o_{e_i};]\big)\Big(\prod_{t=e_i+1}^{T} \Phi\Gamma D\big((P_t B)[o_t;]\big)\Big)\mathbf{1}_5 \quad (4)$$

As before, plausible models were constructed (Table 4) and comparisons made with the DIC and posterior predictive checks. Models varied based on the specification of P_t and Π. The elements of P_t were allowed to be fully time varying, but were linked among different states by three options: equality (e.g. $p_{J,t} = p_{N,t}$ for each t), a logit-link to another state (equation 1), or fully independent. Rather than compute every combination of linkages among states (81 combinations for P_t), attention was focused on the following generalities: p_C should vary from the other states, because females with calves are suspected to migrate along coastal routes more so than the other states (Craig and Herman, 2000; Forestell *et al.*, 2003), while non-breeding females, males and subadults are more or less similar to each other. π_J also varied across models as either a time-invariant or time-varying parameter. This was motivated by the strong spike in estimated EAGVS abundances in the late 2000s (Forestell *et al.*, 2011) which may imply that there is a non-stable age distribution between subadults and adults. The ratio of males to females was assumed to be time-invariant (e.g. $\pi_{M,*}$, conditioned on $\pi_{J,t}$).

The 14 models had the same priors on the logit-scale, with γ_Z, $\pi_{Z,t}$, $\beta_Z \sim Norm(0, 1.682)$. The priors on $p_{Z,t}$ and φ_F were specified as previously as in the truncation models. Assuming that males may have slightly higher mortality than females (Ramp *et al.*, 2010) φ_M was set, $\sim Norm(logit(0.92),1)$. φ_J was set $\sim Norm(logit(0.8),1)$ based on post-yearling estimates of previous EAGVS studies (Hoffman *et al.*, 2010) and humpback calves' survival in the review by Zerbini *et al.* (2010).

As detailed above, the component-wise SCAM Metropolis-within-Gibbs (Haario *et al.*, 2005) sampler was used to sample from the posterior distributions of model parameters, within chains between 40,000 to 100,000 (depending on model complexity). All chains were inspected visually to ensure convergence and efficient mixing.

Individual heterogeneity: simulations
Increasingly, mark-recapture studies utilise hierarchical models to incorporate individual heterogeneity in life-history parameters (Lebreton, 1995; Link and Barker, 2005; Schofield and Barker, 2011) and observation errors (Fletcher *et al.*, 2012). Not only can random effect models provide a more realistic portrait of variation (i.e. overdispersion), but they can re-balance an estimate away from individual sighting events to the population mean (i.e. 'shrinkage', Halstead *et al.*, 2012). In this way, females who breed more and are resighted more will not influence a parameter's estimate drastically more than an individual who breeds less and is resighted less. However, less is known about how hierarchical models perform when a life-history parameter is estimated from a sample that is contaminated with individuals for whom the parameter does not apply (e.g. male subadults should have no influence on breeding propensity).

A further simulation was performed to test the performance of the truncation method and the multi-event

method in the face of individual heterogeneity in birth-interval probabilities. Individual effects (ε_i) were simulated as belonging to a zero-centred Normal distribution with variance σ_γ^2 and linked to the parameters γ_1 and γ_2 via a logit-link:

$$\gamma_{y,i} = \frac{\gamma_y e^{\varepsilon_i}}{1 + \gamma_y(e^{\varepsilon_i} - 1)} \varepsilon_i \sim Norm(0, \sigma_y^2) \qquad (5)$$

Six simulations were run which differed according to two values for the variance ($\sigma_\gamma^2 = 0.5$ and 1) and three values for the resight probabilities of females-without-calves ($p_N, \cdot = 0.5$, 0.25, 0.125). These were different from $p_C, \cdot$ which was set to 0.5 for all simulations, while $p_J, \cdot$ and $p_M, \cdot$ were equal to $p_N, \cdot$. Other life-history parameters were set at $\gamma_1 = 0.1$, $\gamma_2 = 0.55$ (for a true calving interval of 2.64 years), $\gamma_J = 0.15$, $\varphi_J = 0.87$, $\varphi_F = 0.96$, $\varphi_M = 0.91$. The process also included a probabilistic mis-identification of subadults ($1-\beta_J = 0.7$), males ($1-\beta_M = 0.9$) and females without calves ($1-\beta_N = 0.9$). Populations were grown from an initial 400 individuals and sampled after 5 years for 15 years to approximate a stable-age-distribution. Each scenario was repeated 1,000 times and parameters were estimated according to either method using the R *optim* function.

Individual heterogeneity and the multi-event model
The two most supported multi-event models from the EAGVS analysis were selected for reanalysis within a hierarchical context including individual variation in female birth-interval probabilities. The computational cost of incorporating individual-level random effects is significant and was thus only applied to birth-interval probabilities, the main parameter of interest in this study.

Priors, proposal densities and MCMC techniques were applied as previously in the non-hierarchical models. As in equation 5, individual effects were considered from a Normal distribution and linked to population parameters γ_1 and γ_2 via a logit-link. For the group variance parameter, σ_γ^2, a Gamma prior was applied with shape = 1 and rate = 1.5, with the intention that the distribution of individual heterogeneity should at least tend towards being flat on the probability scale, or at least not concave. The posterior distributions of parameter estimates were sampled from the series-products of the Multi-event likelihood (equation 4), the probability density of the Normal distribution for random effects and each parameters' prior density $\pi(\theta)$:

$$P(\theta \mid h) \sim \left(\prod_{i=1}^{n} L_{Multievent}(h_i \mid \varepsilon_i \theta) \right) \left(\prod_{i=1}^{n} L_{Norm}(\varepsilon_i \mid \theta) \right) \pi(\theta)$$

RESULTS
Summaries
The EAGVS Hervey Bay catalogue consisted of 2,973 individuals seen over 22 years. Of these individuals, 1,263 (42.5%) were observed as subadults at least once during their encounter history, 435 (14.6%) were females seen at least once with a calf, 49 (1.6%) were confirmed males and 2,446 (82.2%) could only be identified as an 'unknown adult', i.e. their sex was not confirmed.

The annual counts of observed events varied heavily among years. There was a mean of 64.4 sightings of subadults per year (range of 3–282), 31.5 females-with-calves per year (1–95), 3 non-breeding females per year (0–8), 2.4 males per year (0–6) and 96.4 unknown adults per year (7–289).

Consecutive sightings of the same female were low. There were 214 resightings of the same female consecutively for two years, 28 resightings of the same female consecutively for three years, and 3 for four years consecutively. There were 17 resightings of the same female with a calf for two years consecutively and just one instance of the same female being resighted with a calf for three years consecutively. Conversely, there were 142 instances of the same female seen consecutively for two years without a calf, 16 instances of the same female seen consecutively for three years without a calf and 2 instances of the same female being seen without a calf for four years consecutively.

Simulation of imperfect detection and the truncation method
The truncation method provided unbiased estimates of the true calving interval (2.44 years) when the females-with-calves and the females-without-calves had the same detection probability (Table 1), even when detection probabilities were low. The bias was larger and significant with increasing difference between females with- and without-calves, as well as at lower overall detection probabilities.

A matrix-based reformulation of the truncation method
For the Gulf of Maine humpback whale data (from the appendix of Barlow and Clapham, 1997), all three matrix-based models provided similar estimates of the birth-interval probabilities and calving intervals (Table 2). There were no significant differences among parameter estimates or compared to the point estimates from the original study (row 1 in Table 2). However, the uncertainty in the four and five year intervals (γ_4, γ_5) was very high, with the latter spanning most of the probability scale (i.e. the 95% CI was ~0.07–0.88

Table 1

Simulation of the effects of differential resighting probabilities between females-with-calves and females-without-calves on calving interval estimates.

		Proportional difference in resight probabilities (females-with-calves versus females-without-calves) on the logit-scale		
		log(1)	log(0.5)	log(0.25)
Resight probabilities of females-with-calves	0.95	2.43 (2.42–2.44)[1]	2.41 (2.4–2.42)	2.38 (2.37–2.39)
	0.5	2.42 (2.4–2.44)	2.24 (2.22–2.25)	1.9 (1.88–1.92)
	0.25	2.43 (2.37–2.47)	2.14 (2.1–2.18)	1.68 (1.64–1.71)

[1]95% distribution of simulation Maximum Likelihood Estimator.

Table 2

Comparison of the Gulf of Maine humpback whale calving interval data under a matrix-based reparameterisation including resight probabilities and apparent survival.

Model	g.o.f[3]	Calving interval (years)[4]	Birth interval probabilities				
			γ_1	γ_2	γ_3	γ_4	γ_5
$p = 1, \varphi = 1$[1]		2.38 (s.e 0.10)	0.04	0.59	0.803	0.61	0.458
$p(f(z),t)\ \varphi(\cdot,\cdot)$[2]	0.21	2.48 (2.34–2.66)	0.04 (0.02–0.08)	0.59 (0.5–0.67)	0.77 (0.6–0.88)	0.65 (0.27–0.9)	0.43 (0.07–0.87)
$p(\cdot,t)\ \varphi(\cdot,\cdot)$	0.19	2.49 (2.33–2.64)	0.04 (0.02–0.09)	0.59 (0.51–0.68)	0.77 (0.64–0.89)	0.64 (0.29–0.9)	0.4 (0.06–0.88)
$p(Z,t)\ \varphi(\cdot,\cdot)$	0.19	2.51 (2.37–2.7)	0.04 (0.02–0.08)	0.58 (0.5–0.66)	0.76 (0.61–0.88)	0.68 (0.26–0.89)	0.44 (0.08–0.89)

[1]Probability of resight p and apparent survival φ fixed as 1; from Barlow and Clapham (1997), shown for comparison. [2]The first argument of p and φ denote whether parameters vary by breeding state (Z), or are equal among states ($\cdot$), or vary according to a logit-link $f(Z)$ among states Z (equation 1). The second argument denotes time varying estimates (t) or an estimate equal among all years ($\cdot$). [3]Posterior predictive check. [4]Posterior median and 95% credibility intervals in parentheses.

for all three models). Resight probabilities for both females with- and without-calves were high in almost all years (>0.8 for both classes) and similar across model specifications.

For the EAGVS dataset, the truncation method provided calving interval estimates which varied heavily with the specification of the resight probabilities: equal resight probabilities between females with- and without-calves (models 1 and 4 in Table 3) resulted in much lower calving interval estimates than independent or logit-linked specifications (~1.44 years versus between 2.43–3.01 years respectively). In contrast, changing the maximum birth interval from 5 years to 6 had less effect on the calving interval estimates and the differences were not significant among sister models with the same resight-probability specifications (i.e. models 1, 2 and 3 versus 4, 5 and 6 respectively in Table 3). The models with the most support by DIC were 2 and 5, both with a logit-link specification of p. The better model 2 included 5 birth-interval probabilities and estimated a calving interval of 2.98 years (95% CI: 2.27–3.51 years) while 5 estimated 2.69 (2.07–3.41).

In most model specifications, the uncertainty of the 3rd, 4th, 5th (and 6th) birth-interval probabilities were very large, with γ_4, γ_5 and γ_6 having 95% CI's which spanned most of the probability interval (e.g. for the best fitting model, the interval for γ_4 was 0.06–0.8). Furthermore, point estimates of γ_3, γ_4 and γ_5 were lower than the estimate of γ_2, unlike the Barlow and Clapham results.

All models had similar estimates of apparent survival, with means varying between 0.92–0.95. The best model (2) had a point estimate of 0.92 with a long left-tailed distribution (95% CI: 0.25–0.999).

All models had adequate goodness-of-fit statistics (>>0.05 and <<0.95), but models with a logit-link (including the model with the lowest DIC) had poorer fit statistics than other models (e.g. 0.83 for the selected model 2).

The multi-event framework
All multi-event models had better goodness-of-fit statistics than the truncation models, with most multi-event models having a value of ~0.5 (range of 0.47 to 0.63). Unfortunately,

Table 3

Model selection and parameter estimates for East Australia humpback whales, using the truncated dataset.

Model	No. of non-breeding states (N)	Model[1] description p_C	Model[1] description p_N	No. of parameters	ΔDIC	GOF[2]	Calving Interval (years)[3]	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	φ_F
1	5	t	p_C	26	52.8	0.53	1.44 (1.33–1.59)	0.61 (0.46–0.71)	0.94 (0.85–0.98)	0.67 (0.08–0.98)	0.58 (0.04–0.97)	0.58 (0.05–0.97)	–	0.92 (0.29–1)
2	5	t	$f(p_C)$	27	0	0.83	2.98 (2.27–3.51)	0.08 (0.05–0.15)	0.47 (0.32–0.69)	0.37 (0.18–0.7)	0.34 (0.06–0.8)	0.38 (0.06–0.83)	–	0.92 (0.25–1)
3	5	t	t	46	8.7	0.64	3.01 (2.06–3.53)	0.09 (0.05–0.19)	0.5 (0.3–0.76)	0.37 (0.18–0.83)	0.36 (0.05–0.82)	0.4 (0.08–0.9)	–	0.94 (0.25–1)
4	6	t	p_C	27	52.5	0.53	1.43 (1.33–1.58)	0.6 (0.47–0.71)	0.95 (0.84–0.98)	0.68 (0.09–0.98)	0.57 (0.05–0.98)	0.54 (0.06–0.97)	0.52 (0.05–0.97)	0.95 (0.23–1)
5	6	t	$f(p_C)$	28	0.6	0.8	2.69 (2.07–3.41)	0.1 (0.06–0.17)	0.51 (0.33–0.75)	0.42 (0.19–0.84)	0.46 (0.07–0.93)	0.44 (0.05–0.94)	0.85 (0.13–0.99)	0.93 (0.25–1)
6	6	t	t	47	8.2	0.63	2.43 (1.92–3.48)	0.11 (0.06–0.22)	0.56 (0.32–0.84)	0.43 (0.18–0.93)	0.41 (0.07–0.95)	0.47 (0.06–0.96)	0.83 (0.1–0.99)	0.93 (0.2–1)

[1]Columns only indicates those parameters which vary between models. All models include a time invariant parameter for apparent female survival, φ_F. t = independent and fully time stochastic; $f(p_z)$ = logit-link function to variable p_z; and p_z = same as variable p_z. [2]Posterior predictive check. [3]Posterior median and 95% credibility intervals. [4]γ_y = conditional birth interval probability for interval y; φ_F = apparent female survival.

Table 4
Model selection and parameter estimates using the multi-event framework.

Model	Model framework	π_J	p_J	p_C	p_N	p_M	No. of parameters	ΔDIC	GOF[2]	Calving interval (years)[3]	Demographic parameter estimates[4]					
											γ_J	γ_1	γ_2	φ_J	φ_F	φ_M
7	multievent	•	p_C	t	p_C	p_C	32	451	0.63	1.5 (1.39–1.62)	0.17 (0.12–0.24)	0.8 (0.73–0.86)	0.39 (0.29–0.52)	0.74 (0.67–0.79)	0.99 (0.98–1)	0.9 (0.87–0.93)
8	*multievent*	•	$f(p_C)$	t	$f(p_C)$	$f(p_C)$	35	440	0.63	2.83 (2.31–3.81)	0.17 (0.1–0.27)	0.13 (0.07–0.21)	0.45 (0.31–0.61)	0.72 (0.61–0.79)	0.99 (0.97–1)	0.93 (0.9–0.96)
9	*multievent*	•	p_N	t	t	p_N	52	385	0.54	2.73 (2.23–3.37)	0.11 (0.07–0.15)	0.13 (0.07–0.22)	0.5 (0.36–0.63)	0.8 (0.74–0.84)	0.97 (0.96–0.98)	0.94 (0.91–0.96)
10	*multievent*	•	$f(p_N)$	t	t	p_N	53	370	0.55	2.75 (2.33–3.37)	0.17 (0.1–0.27)	0.13 (0.07–0.2)	0.49 (0.37–0.62)	0.7 (0.6–0.78)	0.98 (0.96–0.99)	0.94 (0.92–0.97)
11	*multievent*	•	$f(p_C)$	t	t	p_N	53	369	0.54	2.7 (2.3–3.2)	0.21 (0.14–0.33)	0.14 (0.08–0.23)	0.51 (0.4–0.63)	0.68 (0.58–0.76)	0.98 (0.97–0.99)	0.94 (0.91–0.97)
12	*multievent*	•	$f(p_N)$	t	t	$f(p_C)$	54	426	0.55	2.55 (2.11–3.14)	0.15 (0.09–0.25)	0.15 (0.09–0.27)	0.55 (0.41–0.68)	0.74 (0.64–0.82)	0.98 (0.97–0.99)	0.92 (0.89–0.95)
13	*multievent*	•	t	t	t	$f(p_N)$	73	386	0.54	2.58 (2.17–3.33)	0.38 (0.21–0.51)	0.15 (0.07–0.25)	0.52 (0.36–0.65)	0.53 (0.44–0.64)	0.98 (0.97–0.99)	0.95 (0.92–0.97)
14	*multievent*	t	p_C	t	p_C	p_C	51	68.9	0.56	1.5 (1.39–1.62)	0.17 (0.12–0.24)	0.82 (0.74–0.87)	0.37 (0.27–0.47)	0.73 (0.67–0.8)	0.99 (0.97–1)	0.87 (0.85–0.89)
15	*multievent*	t	$f(p_C)$	t	$f(p_C)$	$f(p_C)$	54	50.8	0.59	3.64 (2.69–5.07)	0.13 (0.08–0.22)	0.1 (0.06–0.18)	0.34 (0.21–0.49)	0.71 (0.62–0.79)	0.99 (0.97–1)	0.91 (0.89–0.93)
16	*multievent*	t	p_N	t	t	p_N	71	5.6	0.49	2.96 (2.36–4.26)	0.08 (0.05–0.13)	0.11 (0.06–0.2)	0.45 (0.27–0.59)	0.82 (0.76–0.87)	0.97 (0.95–0.98)	0.91 (0.88–0.92)
17	***multievent***	t	$f(p_N)$	t	t	p_N	**72**	**1.2**	**0.48**	**3.27 (2.48–4.66)**	**0.11 (0.06–0.19)**	**0.1 (0.06–0.18)**	**0.4 (0.24–0.56)**	**0.72 (0.64–0.8)**	**0.97 (0.96–0.99)**	**0.92 (0.9–0.94)**
18	***multievent***	t	$f(p_C)$	t	t	p_N	**72**	**1.1**	**0.5**	**2.89 (2.36–3.63)**	**0.16 (0.1–0.26)**	**0.12 (0.07–0.2)**	**0.46 (0.33–0.58)**	**0.68 (0.59–0.77)**	**0.98 (0.97–0.99)**	**0.92 (0.89–0.93)**
19	*multievent*	t	$f(p_N)$	t	t	$f(p_C)$	73	19.7	0.5	1.15 (1.1–1.23)	0.2 (0.12–0.31)	0.9 (0.85–0.92)	0.7 (0.59–0.8)	0.7 (0.6–0.78)	0.99 (0.97–1)	0.9 (0.87–0.92)
20	***multievent***	t	t	t	t	$f(p_N)$	**92**	**0**	**0.49**	**2.78 (2.23–3.68)**	**0.18 (0.11–0.31)**	**0.13 (0.07–0.22)**	**0.48 (0.32–0.62)**	**0.64 (0.53–0.72)**	**0.98 (0.96–0.99)**	**0.92 (0.9–0.94)**
21	*multievent, random effects*	t	$f(p_N)$	t	t	p_N	73[5]	1287	0.49	3.46 (2.53–4.99)	0.11 (0.04–0.02)	0.09 (0.06–0.02)	0.36 (0.05–0.18)	0.73 (0.22–0.17)	0.97 (0.64–0.53)	0.92 (0.96–0.8)
22	*multievent. random effects*	t	t	t	t	$f(p_N)$	93[5]	1111	0.46	2.94 (2.23–3.86)	0.18 (0.03–0.02)	0.12 (0.1–0.02)	0.45 (0.06–0.3)	0.63 (0.29–0.21)	0.98 (0.53–0.6)	0.92 (0.96–0.72)

[1]Columns only indicate those parameters which vary between models. All models include initial encounter probabilities (π_Z), classification errors (β_Z), a subadult maturation rate (γ_J), and apparent survival (φ_Z). t = independent and fully time stochastic; • = time-invariant; $f(p_Z)$ = logit-link function to variable p_Z; and p_Z = same as variable p_Z. [2]Posterior predictive check. [3]Posterior median and 95% credibility intervals. [4]γ_J = maturation parameter from subadult to adult; γ_y = conditional birth interval probability for interval y; φ_Z = apparent survival. [5]Minimum estimate of the number of parameters for random effects models, counting only the variance parameter as an extra parameter.

the DIC's could not be compared between the truncation and multi-event models because they use different data (the former being a subset of the latter).

The model with the most support by DIC was model 20 (Table 4). The largest differences among DIC values occurred as a result of differing specifications of time-invariant and time-varying estimates of π_J. The latter models yielded estimates of $\pi_{J,t}$ which varied heavily by year, being low for the time between 1987 to 1993 (<0.1), and much higher for years after 1996 (>0.48).

Otherwise, poorer fits and larger DIC values were common among models which had more constraints on resighting probabilities (models 7, 8, 14 and 15) as compared to models which had some combination of independent and logit-linked resight probabilities by sex and age. The models with the most support (ΔDIC <3), had similar or equal resight probabilities among males (p_M) and females without calves (p_N), while the probability of resighting a subadult (p_J) differed from all other adult classes (much greater than p_M and p_N). In general, all resighting rates seemed to decline over the years (see Fig. 1, based on model 20), consistent with a growing population.
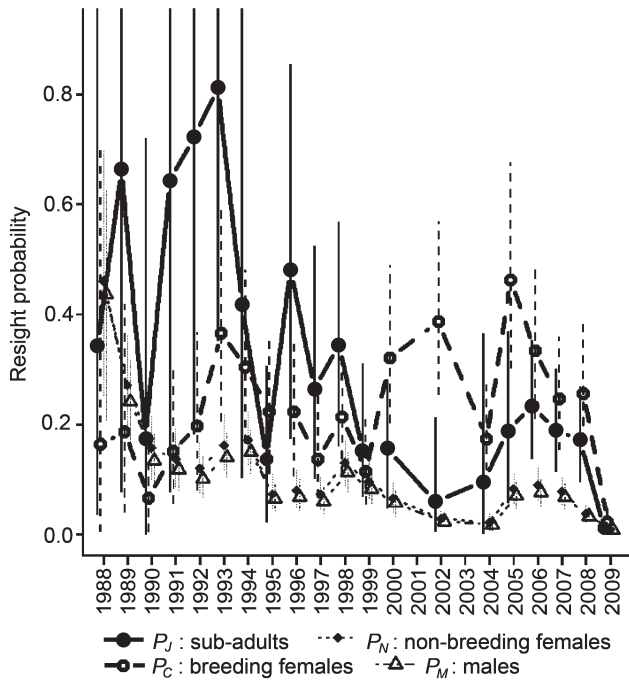
Fig. 1. Estimated humpback whale resight probabilities for age-sex classes.

Fig. 2. Simulation results from six scenarios under differential resighting probabilities between females-with-calves (p_C = 0.5) and females-without-calves (p_N = 0.125, 0.25 and 0.5; y-axis), and degree of variation in individual heterogeneity (σ_γ^2 = 0.5 and 1) in birth-interval probabilities. Open circles are means of ML Estimators across simulations according to the truncation method (dark circles and lines) and the multi-event method (line circles and lines).

Apparent survival estimates were consistent across models, with lower values for subadults (φ_J, range of means: 0.53–0.82) than males (φ_M, 0.87–0.95) and both being lower than females' survival (φ_F, 0.97–0.99). The selected model (20) estimated an apparent survival of 0.64 (95% CI: 0.53–0.72) for subadults, 0.92 (0.90–0.94) for males and 0.98 (0.96–0.99) for females. In general, the multi-event models produced more precise estimates of female apparent survival than the truncation models (e.g. 95% CI of 0.96–0.99 for model 20 versus 0.25–1 for the truncation model 2. Maturation rates (γ_J) varied from (0.08–0.38) across models, with the best model estimating 0.18 (95% CI: 0.11–0.31).

Point estimates of the calving interval varied widely by model specifications (range 1.15–3.64 years) and had a ΔDIC-weighted model average value of 2.97 years. Model 20 had an estimate of 2.78 years (95% CI: 2.23–3.68 years). Uncertainty estimates were slightly larger than those of the truncation models.

Individual heterogeneity: simulations
Both the truncation method and the multi-event method produced parameter estimates which were close to the simulation specifications and there were no significant differences in the resulting calving interval estimates among the different scenarios (Fig. 2). There was a slight negative bias apparent in both methods, which increased with greater individual variation (σ_γ^2 = 1) and greater difference between resight probabilities of breeding versus non-breeding females (especially $p_{N,*}$ = 0.125). The multi-event framework was consistent in producing smaller variance estimates for the calving interval than the truncation method in all cases and seemed to have slightly less bias.

Individual heterogeneity and the multi-event model
For the EAGVS, both multi-event random effect models estimated larger calving intervals and greater uncertainty than their corresponding non-hierarchical model, e.g. 2.78
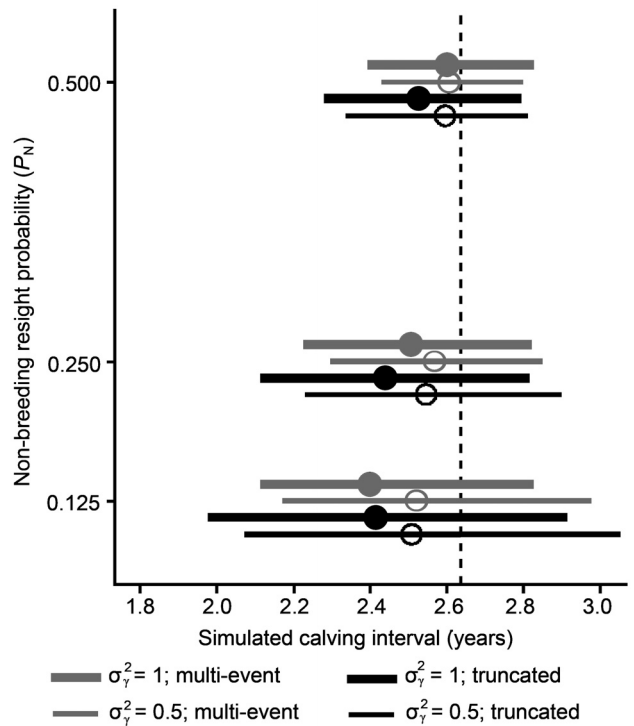
years (95% CI: 2.23–3.68) for model 20 versus 2.94 years (2.23–3.86 years) for model 22. The estimated group variances of the random effect models were low, e.g. 0.035 (95% CI: 0.018–0.045) for the best model 22. Other demographic parameters, such as class-specific survival and maturity, were similar between hierarchical and non-hierarchical models, with differences of less than 2 percentage points and small differences in uncertainty.

Both hierarchical models had adequate goodness-of-fit statistics (0.49 and 0.45 for models 21 and 22 respectively), while DIC values were much larger than the non-hierarchical models (with ΔDIC values of 1,287.3 and 1,111 respectively for models 21 and 22). However, the DIC has less theoretical and empirical support in hierarchical models (Celeux *et al.*, 2006; Jordan, 2011).

DISCUSSION

This study provides an updated formulation of two methods to estimate birthing probabilities and derive calving interval estimates (Barlow and Clapham, 1997; Fujiwara and Caswell, 2002; Pradel *et al.*, 2005) when considering low resighting rates and partial-identification of age and sex classes. We demonstrate the matrix construction and careful conditioning of likelihoods to model either truncated data or multi-state data, which are subtly different from most other sight-resight studies. These updated formulations provide a more accessible means to conduct these analyses, whereas the original likelihood from Barlow and Clapham required custom programming. For example, both methods can be analysed according to either a frequentist approach, e.g. E-SURGE (Choquet *et al.*, 2009), or through a latent-state

Bayesian framework, e.g. the BUGS language (Kéry and Schaub, 2012). Importantly, the matrix formulation facilitates the incorporation of other life-history parameters, such as survival, which may be more important than reproduction for the population dynamics of long-lived marine species (Benton and Grant, 1999; Brault and Caswell, 1993; Crowder *et al.*, 1994).

Modeling implications

Simulations suggested that imperfect detection has little practical implication when the overall encounter rates are high and similar between breeding and non-breeding states (i.e. p_z in the range of 0.7 of 0.95). This seemed to be the case for the Gulf of Maine dataset, where ignoring resight rates had no effect on the estimates of birth-interval probabilities. This may be a great relief to some researchers, as ignoring resight probabilities greatly simplifies analysis. However, the early studies of the Gulf of Maine population may be more of an exception than a representative example. A more typical scenario may be the EAGVS dataset, where resight rates were crucial to the final estimates of calving intervals. Furthermore, the benchmark of ecological analysis is moving towards more explicit modelling of observational errors versus biological processes (Clark and Gelfand, 2006; Cressie *et al.*, 2009). De facto inclusion of observational errors is more conservative, may improve the precision of estimates (Barker and Kavalieris, 2001) and may provide useful information on resighting and survival.

Considering the analysis of the EAGVS dataset, the truncation method and the multi-event framework both provided similar calving interval estimates with overlapping 95% credibility intervals. Simulations likewise demonstrated that the two methods had similar performance under strong individual heterogeneity. Truncating the data resulted in poorer goodness-of-fit statistics as compared to the multi-event models, but this is not unexpected considering the large difference in the number of parameters (e.g. 27 versus 92). Calving interval estimates varied much more within either framework based on the particular specification of resight probabilities. This suggests that it is important to control variation in detection and encounter probabilities at the design stage and focus attention on minimising artefacts of sampling and effort. The one domain where the multi-event framework showed a distinct benefit was by including class-specific survival estimates, which are otherwise discarded by the truncation method (e.g. males). The method also improved the precision of female apparent survival estimates.

This study provides some initial insights into the ongoing issue of individual and/or age-specific heterogeneity in life-history parameters (Robbins, 2007; Rosenbaum *et al.*, 2002). Our simulations suggest that a combination of individual heterogeneity, low-resighting rates and partial-identifiability of non-breeders can lead to slight biases in both the truncation and multi-event estimates, whereby calving interval estimates are somewhat lower than the true population value in either framework. Hierarchical models with individual random effects may be a promising way to 'shrink' the estimate away from the mean of the encounters towards the mean of the population (Halstead *et al.*, 2012). As anticipated, the two random effect multi-event models increased the calving interval estimates from 2.78 to 2.94

years and from 3.27 to 3.46 years (however, both differences were non-significant). This is consistent with the idea that photo-ID catalogues are biased to females who breed more and are more likely to be observed, which subsequently biases estimates from non-hierarchical models.

We assumed a Normal distribution for individual effects, as is similarly done in the popular E-SURGE software (Choquet *et al.*, 2009) and most other ecological studies. Despite its popularity, the Normal distribution may be an inadequate distribution when samples are thought to be skewed (Lachos *et al.*, 2009). In particular, it is unclear how unclassified individuals (of which only some are actually breeding females) may influence the random effects distribution for a female-specific parameter, such as birth-interval probabilities. In this case, the distribution may be better understood as a Contaminated Normal distribution (Tukey, 1960). Conversely, if the act of truncation biases data towards higher-frequency breeders, than a Skewed Normal distribution may be more appropriate to rebalance the posterior distribution to lower values (Bandyopadhyay *et al.*, 2012; Lachos *et al.*, 2009). In both cases, further research and simulations are required to investigate the robustness of the Normal distribution and whether attempts to correct one bias will merely introduce bias in the opposite direction. Latent state simulations and clever Gibbs samplers may be another promising avenue (Clark *et al.*, 2005) by drawing hyperparameters exclusively from the cohort of (sex-stochastic) females within a MCMC iteration.

Biological implications

The final models for the truncation method (2) and the multi-event model (22) estimated calving intervals of 2.98 years (95% CI: 2.27–3.51) and 2.78 years (95% CI: 2.23–3.68), respectively. These correspond to birth-rates of 0.33 (0.28–0.44) and 0.36 (0.27–0.44) and are somewhat lower than the point estimates of six studies reviewed by Zerbini *et al.* (2010) which range between 0.37–0.44. However, some of these estimates are likely too low, given the difficulty of observing rare events like a female delaying reproduction for four years or more. Nonetheless, if these other studies' estimates are accepted, then there is a seeming disconnect between the EAGVS slow birth-rate and its strong population rate-of-increase (Forestell *et al.*, 2011; Noad *et al.*, 2011). However, such population-level measures generally have a weak coupling to the birth-rate of long-living marine species (Benton and Grant, 1999; Brault and Caswell, 1993; Crowder *et al.*, 1994). Instead, population parameters arise from interactions among many life-history parameters, such as age-of-first parturition and survival, the latter being particularly important in elasticity studies. In order to make inferences on growth and recruitment, the multi-event extension seems like a promising approach to model recruitment-like parameters from backwards-time encounter histories (Link and Barker, 2005; Pradel, 1996), given its origins in MSMR. To our knowledge, this has not been done under partial and mis-identification of states and will require further development and simulations. Until further study, it is not possible to compare the birth-rate to the population rate of increase.

The multi-event method provided more sensible estimates of survival and facilitated the estimation of male apparent

survival, which has only a few examples in the scientific literature on humpback whales. The truncation method provided female a survival estimate of 0.92 (95% CI: 0.25–1), which is lower than earlier estimates of 0.945 from the same data (Forestell *et al.*, 2011), while the multi-event's estimate was 0.98 (95% CI: 0.96–0.99) and is more in line with Zerbini and colleagues' review of 16 populations in 8 studies (with an average across studies of ~0.95). Male apparent survival was lower than the estimates for females, a ranking which is seen in some studies (Ramp *et al.*, 2010) but is opposite in others (Robbins, 2007). Subadult apparent survival was lower than many calf survival estimates (Zerbini *et al.*, 2010) and other subadult survival estimates. However, literature values vary a lot, from as low as 0.702 (Rosenbaum *et al.*, 2002) to greater than 0.9 (Robbins, 2007).

The 'maturation rate' in this study is a nuisance parameter without a clear biological interpretation. Its inverse (plus one year for the time spent as a calf) could be viewed as an average lower-bound to the age-of-first parturition (6.56 years; 95% CI: 4.23–13.50 years), which is similar to the lowest values reported by Zerbini *et al.* (2010) which ranged between 5.9–11 years. This nuisance specification is not an essential feature of the multi-event framework, but was necessary in our case because subadults and adults were clearly mis-classified as one another. Other datasets can and should model the true age-of-first parturition if encounter histories have more observations of calves seen later as adults.

Number of non-breeding states
An auxiliary but useful insight comes from the differing number of female breeding states considered among the models, i.e. whether a female who has not bred in two years should be considered to be in a different health and reproductive state than a female who just bred the previous year (Wiley and Clapham, 1993). This distinction is codified in the 'memory' versus 'memoryless' specifications of Γ. Barlow and Clapham (1997) conceived of serial transitions along a finite number of non-breeding states, whereas most MSMR studies consider just two memoryless states, such that females can remain in either state indefinitely. Neither specification is intrinsic to the truncated or multi-event likelihood, yet both could be used within either framework.

We suggest that the benefit of either specification is contingent upon having to deal with high or low resight probabilities. The Barlow and Clapham (1997) study had high resighting rates with confirmed sightings of females delaying breeding attempts for 5 or 6 years consecutively. This facilitated the distinction between γ_3 versus γ_2 and has an intuitive appeal: it suggests that females who have not calved in three years are more likely to breed than females who have not calved in two years. With lower resighting rates, the data has less power to resolve rarer, longer-interval calving events, such that γ_3, γ_4, and γ_5 have very wide CI's which span most of the probability scale. In the case of the EAGVS, this led to counter-intuitive point estimates of γ_3, γ_4 and γ_5 which were lower than γ_2. This is less a commentary about the breeding decisions of individual females and is more likely a consequence of the sparse data. In this case,

the memoryless system seems like a reasonable compromise, whereby females are assumed to decide to breed based only on their previous state. An alternative approach could be to apply stronger priors on later intervals (γ_3, γ_4, γ_5) to reflect the belief that females should be more likely to breed, not less likely, as time-since-breeding increases.

The memoryless system also circumvents the question about what is the appropriate maximum number of intervals to consider for the memory specification of Γ. Barlow and Clapham (1997) answered this question based on their ability to calculate an unconditional probability 0.984 of breeding within 5 years, leaving only a small residual probability of 0.016 that females would breed after longer intervals. For the EAGVS, the point estimate of this residual probability was 0.123 after five years (model 2) and 0.012 after six (model 4). By the Barlow and Clapham criteria, this may suggest that six intervals are adequate for the EAGVS. Alternatively, this could also be an artefact of the spread of uncertainty across γ_y due to low resighting rates. We expected that adding an extra year would increase the overall calving interval estimate. Instead, our results suggested that the six-year models did not result in significantly longer calving intervals than the five-year models, and in some cases the calving interval estimates were actually slightly lower. In either case, the functional specification of resight probabilities appears to be much more important than the specific specification of breeding versus non-breeding states. Again, this reaffirms the need for robust study design and control over encounter probabilities to strengthen inferences on the reproductive biology of cetaceans. It also places a high bar on sight-resight surveys when the model outputs may be used to inform other biological hypotheses, such as health condition and reproductive status (Miller *et al.*, 2011; Wiley and Clapham, 1993) or evolutionary processes via structural equation modelling (Cubaynes *et al.*, 2011).

CONCLUSIONS
This study reviews the issue of using large amounts of unclassified individuals in sight-resight analyses for inferences on life-history parameters. In particular, care is necessary when conditioning the likelihood, either by conditioning on just the known breeding females, or with explicit inclusion of an entry process to include the entire sample of individuals. Both methods led to similar conclusions on calving intervals (~3 years for the EAGVS), but modelling individuals' entire encounter history, under a multi-event framework, expanded the range of parameters and improved the precision of apparent survival. Both methods are more sensitive to the exact functional specification of resight probabilities than to different specifications of the breeding and non-breeding states (e.g. 5 versus 6 maximum intervals, or a 2 stage Markovian transition matrix). Individual heterogeneity in breeding propensity, especially under low resighting rates, may result in a slight bias in photo-ID catalogues towards more frequent observations of higher-frequency breeders. Through simulation and analysis, this study provides weak evidence that such may bias population estimates. Further simulations and developments are required to link the calving interval methods to population level processes, such as recruitment.

REFERENCES

Avril, A., Letty, J., Léonard, Y., Santin-Janin, H. and Pontier, D. 2012. A multi-event model to study stage-dependent dispersal in radio-collared hares: when hunting promotes costly transience. *Ecology* 93(6): 1,305–16.

Bandyopadhyay, D., Lachos, V.H., Castro, L.M. and Dey, D.K. 2012. Skew-Normal/Independent linear mixed models for censored responses with applications to HIV viral loads. *Biomet. J.* 54(3): 405–25.

Barker, R.J. and Kavalieris, L. 2001. Efficiency gain from auxiliary data requiring additional nuisance parameters. *Biometrics* 57(2): 563–66.

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

Benton, T.G. and Grant, A. 1999. Elasticity analysis as an important tool in evolutionary and population ecology. *Trends in Ecological Evolution* 14(12): 467–71.

Brandão, A., Butterworth, D.S. and Brown, M.R. 2000. Maximum possible humpback whale increase rates as a function of biological parameter values. *J. Cetacean Res. Manage. (Suppl.)* 2: 192–93.

Brault, S. and Caswell, H. 1993. Pod-specific demography of killer whales (*Orcinus orca*). *Ecology* 74(5): 1,444–54.

Calambokidis, J., Falcone, E.A., Quinn, T.J., Burdin, A.M., Clapham, P.J., Ford, J.K.B., Gabriele, C.M., LeDuc, R., Mattila, D., Rojas-Bracho, L., Straley, J.M., Taylor, B.L., Urban R, J., Weller, D., Witteveen, B.H., Yamaguchi, M., Bendlin, A., Camacho, D., Flynn, K., Havron, A., Huggins, J. and Maloney, N. 2008. SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific. Final report for Contract AB133F-03-RP-00078, US Department of Commerce Western Administrative Center, Seattle, Washington. [Available at *http://www.cascadiaresearch.org/SPLASH/SPLASH-contract-report-May08.pdf*].

Celeux, G., Forbes, F., Robert, C.P. and Titterington, D. 2006. Deviance information criteria for missing data models. *Bayesian Anal.* 1(4): 651–74.

Chaloupka, M. and Osmond, M. 1999. Spatial and seasonal distribution of humpback whales in the Great Barrier Reef. pp.89–106. *In*: Musick, J.A. (eds). *Life in the Slow Lane: Ecology and Conservation of Long-lived Marine Mammals*. American Fisheries Society Symposium, Berthesda, MD. 265pp.

Chaloupka, M., Osmond, M. and Kaufman, G. 1999. Estimating seasonal abundance trends and survival probabilities of humpback whales in Hervey Bay (east coast Australia). *Mar. Ecol. Prog. Ser.* 184: 291–301.

Chilvers, B.L., Wilkinson, I.S. and Mackenzie, D. 2010. Predicting life-history traits for female New Zealand sea lions (*Phocarctos hookeri*): integrating short-term mark recapture data and population modeling. *J. Agric. Bio. Env. Stat.* 15(2): 259–78.

Choquet, R., Rouan, L. and Pradal, R. 2009. Program E-SURGE: a software application for fitting multi-event models. pp.845–65. *In*: Thomson, D.L., Cooch, E., Conroy, M.J., Patil, G., Gregoire, T., Lawson, A.B. and Nussbaum, B.D. (eds). *Modeling Demographic Processes in Marked Populations*. Springer Science and Business Media, New York. 1131pp. [Available at: *http://www.springerlink.com/content/k24557h05h033551/abstract/* Accessed June 2012].

Clapham, P.J. 2000. The humpback whale: seasonal feeding and breeding in a baleen whale. pp.173–96. *In*: Mann, J., Connor, R.C., Tyack, P.L. and Whitehead, H. (eds). *Cetacean Societies. Field Studies of Dolphins and Whales*. The University of Chicago Press, Chicago. 433pp.

Clark, J.S., Ferraz, G., Oguge, N., Hays, H. and DiConstanzo, J. 2005. Hierarchical Bayes for structured, variable populations: from recapture data to life-history prediction. *Ecol.* 86(8): 2232–44.

Clark, J.S. and Gelfand, A.E. 2006. *Hierarchical Modelling for the Environmental Sciences: Statistical Methods and Applications*, Oxford University Press, Oxford. 205pp.

Conn, P.B. and Cooch, E. 2009. Multistate capture-recapture analysis under imperfect state observation: an application to disease models. *J. Appl. Ecol.* 46(2): 486–92.

Craig, A.S. and Herman, L. 2000. Habitat preferences of female humpback whales (*Megaptera novaengliae*) in the Hawaiian Islands are associated with reproductive state. *Mar. Ecol. Prog. Ser.* 193: 209–16.

Craig, A.S. and Herman, L.M. 1997. Sex differences in site fidelity amd migration of humpback whales (*Megaptera novaengliae*) to the Hawaiian Islands. *Can. J. Zool.* 75(11): 1,923–33.

Cressie, N.A., Calder, C., Clark, J., Hoef, J. and Wikle, C. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierachical modeling. *Ecol. Appl.* 19(3): 553–70.

Crowder, L., Crouse, D., Heppell, S.S. and Martin, T. 1994. Predicting the impact of turtle excluder devices on loggerhead sea turtle populations. *Ecol. Appl.* 4: 437–45.

Cubaynes, S., Doutrelant, C., Gregoire, A., Perret, P., Faivre, B. and Gimenez, O. 2011. Testing hypothesis in evolutionary ecology with imperfect detection: capture-recapture structural equation modeling. *Ecol.* 93(2): 248–55.

Eddelbuettel, D. and Francois, R.E. 2011. REPP: seamless R and C++ integration. *Journal of Statistical Software* 40(8): 1–18.

Fletcher, D., Lebreton, J.D., Marescot, L., Schaub, M., Gimenez, O., Dawson, S. and Slooten, E. 2012. Bias in estimation of adult survival and asymptotic population growth rate caused by undetected capture heterogeneity. *Meth. Ecol. Evol.* 3(1): 206–16.

Forestell, P., Kaufman, G.D. and Chaloupka, M. 2011. Long term trends in abundance of humpback whales in Hervey Bay, Australia. *J. Cetacean Res. Manage. (special issue 3)*: 235–42.

Forestell, P.H., Kaufman, G. and Chaloupka, M. 2003. Migratory characteristics of humpback whales (*Megaptera novaengliae*) in Hervey Bay and the Whitsunday Islands, Queensland, Australia: 1993–1999. Final Report to the Environmental Protection Agency, Queensland Parks and Wildlife Service Brisbane, Australia, October 2003.

Fujiwara, M. and Caswell, H. 2002. Estimating population projection matrices from multi-stage mark-recapture data. *Ecol.* 83(12): 3,257–65.

Gelman, A., Jakulin, A., Pittau, M.G. and Su, Y. 2008. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* 2(4): 1,360–83.

Gelman, A., Meng, X. and Stern, H. 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sinica* 6: 733–59.

Gelman, A. and Shalizi, C. 2012. Philosophy and the practice of Bayesian statistics. *Brit. J. Math. Stat. Psych.* 66(1): 8–38.

Glockner-Ferrari, D.A. and Ferrari, M.J. 1990. Reproduction in the humpback whale (*Megaptera novaengliae*) in Hawaiian waters, 1975–1988: the life history, reproductive rates, and behaviour of known individuals identified through surface and underwater photography. *Rep. int. Whal. Commn* 12: 161–69.

Glockner, D.A. 1983. Determining the sex of humpback whales in their natural environment. pp.447–64. *In*: Payne, R. (eds). *Communication and Behavior of Whales*. AAAS Selected Symposium 76. Westview Press, Boulder, Colorado. 643pp.

Haario, H., Saksman, E. and Tamminen, J. 2005. Componentwise adaptation for high dimensional MCMC. *Comput. Stat.* 20(2): 265–73.

Hall, B. 2012. LaplacesDemon: Software for Bayesian Interference. CRAN. R package version 12.04.02. [Available at: *http://cran.r-project.org/web/packages/LaplacesDemon/index.html*].

Halstead, B.J., Wylie, G.D., Coates, P.S., Valcarcel, P. and Casazza, M.L. 2012. 'Exciting statistics': the rapid development and promising future of hierachical models for population ecology. *Anim. Conserv.* 15(2): 133–35.

Hoffman, W.S., Kaufman, G. and Jule, K. 2010. Estimation of survival. Recruitment and realized growth rates of the East Australia humpback population (BS-1) using temporal symmetry models. Paper SC/62/SH14 presented to the IWC Scientific Committee, June 2010, Agadir, Morocco (unpublished). 17pp. [Paper available from the Office of this Journal].

Jordan, M.B. 2011. What are the open problems in Bayesian statistics? *Int. Soc. Bayesian Anal. Bull.* 18(1): 568.

Kaufman, G.D., Lagerquist, B.A., Forestell, P.H. and Osmond, M.G. 1993. *Humpback Whales of Australia: a Catalogue of Individual Whales Identified by Fluke Photographs*. 1st ed. Queensland Department of Environment and Heritage, Brisbane, Queensland, Australia. 207pp.

Kendall, W.L., Hines, J.E. and Nichols, J.D. 2003. Adjusting multistate capture-recapture models for misclarification bias: Manatee breeding proportions. *Ecol.* 84(4): 1,058–66.

Kéry, M. and Schaub, M. 2012. Estimation of survival and movement from capture–recapture data using multistate models pp.263–313. *In*: Elsevier, I. (eds). *Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective*, Waltham, MA. 537pp.

Lachos, V.H., Dey, D.K. and Cancho, V.G. 2009. Robust linear mixed models with skew-normal independent distributions from a Bayesian perspective. *J. Stat. Plan. Infer* 139(12): 4,098–110.

Lebreton, J.D. 1995. The future of population dynamic studies using marked individuals: A statistician's perspective. *J. App. Stat.* 22(5–6): 1,009–30.

Link, W.A. and Barker, R.J. 2005. Modeling association among demographic parameters in analysis of open population capture–recapture data. *Biometrics* 61(1): 46–54.

MacKenzie, D.I. and Kendall, W.L. 2002. How should detection probability be incorporated into estimates of relative abundance? *Ecol.* 83(9): 2,387–93.

Miller, C.A., Reeb, D., Best, P.B., Knowlton, A.R., Brown, M.W. and Moore, M.J. 2011. Blubber thickness and its relationship with reproduction, life history status and possibly prey abundance in right whales *Eubalaena glacialis* and *Eubalaena australis*. *Mar. Ecol. Prog. Ser.* 438: 267–283.

Nichols, J.D., Kendall, W.L., Hines, J.E. and Spendelow, J.A. 2004. Estimation of sex-specific survival from capture–recapture data when sex is not always known. *Ecol.* 85(12): 3,192–201.

Noad, M.J., Cato, D.H. and Paton, D. 2011. Absolute and relative abundance estimates of Australian east coast humpback whales (*Megaptera novaeangliae*). *J. Cetacean Res. Manage. (Special Issue 3)*: 243–52.

Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* 52: 703–09.

Pradel, R., Gimenez, O. and Lebreton, J.D. 2005. Principles and interest of GOF tests for multistate capture-recapture models. *Animal Biodiversity Conserv.* 28: 189–204.

R Development Core Team, F. 2010. R: A Language and Environment for Statistical Computing. (Austria, Ed.). R Foundation for Statistical Computing. ISBN 3–900051–07–0. [Available at: *http://www.r-project.org*].

Ramp, C., Berube, M., Palsboll, P., Hagen, W. and Sears, R. 2010. Sex-specific survival in the humpback whale (*Megaptera novaeangliae*) in the Gulf of St. Lawrence, Canada. *Mar. Ecol. Prog. Ser.* 400: 267–76.

Robbins, J. 2007. Structure and dynamics of the Gulf of Maine humpback whale population. Doctoral Thesis, University of St Andrews, Scotland. [Available at: *http://hdl.handle.net/10023/328*].

Roberts, G.O. and Rosenthal, J.S. 2009. Examples of adaptive MCMC. *J. Comp. Graph. Stat.* 18(2): 349–67.

Rosenbaum, H.C., Weinrich, M.T., Stoleson, S.A., Gibbs, J.P., Baker, C.S. and DeSalle, R. 2002. The effect of differential reproductive success on population genetic structure: Correlations of life history with matrilines in humpback whales of the Gulf of Maine. *J. Hered.* 93(6): 389–99.

Schofield, M.R. and Barker, R.J. 2011. Full open population capture–recapture models with individual covariates. *J. Agric. Biol. Environ. Stat.* 16(2): 253–68.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van Der Linde, A. 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc.* (Series B Stat. Methodol.) 64(4): 583–639.

Tukey, J.W. 1960. A survey of sampling from contaminated distributions. pp.448–85. *In*: Olkin, I., S. G. Ghurye, S.G., Hoeffding, W., Madow, W.G. and Mann, H.B. (eds). *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Standford University Press, Stanford, CA. 517pp.

Tyack, P. 1981. Interactions between singing Hawaiian humpback whales and conspecifics nearby. *Behav. Ecol. Sociobiol.* 8: 105–16.

Wiley, D.N. and Clapham, P.J. 1993. Does maternal condition affect the sex ratio of offspring in humpback whales. *Anim. Behav.* 46(2): 321–24.

Winn, H.E. and Winn, L.K. 1978. The song of the humpback whale *Megaptera novaeangliae* in the West Indies. *Mar. Biol.* 47(2): 97–114.

Zerbini, A.N., Clapham, P.J. and Wade, P.R. 2010. Assessing plausible rates of population growth in humpback whales from life-history data. *Mar. Biol.* 157(6): 1225–36.