Mark-Recapture with Multiple Non-invasive Marks

Simon Bonner and Jason Holmberg

August 27, 2012

Abstract

Non-invasive marks including pigmentation patterns, acquired scars, and genetic markers, are often used to identify individuals in mark-recapture experiments. Such data present new statistical challenges including that the marks may change over time, that the marks may be misidentified at non-negligible rates, and that only part of the population may possess a specific mark. Motivated by the analysis of data from the ECOCEAN online whale shark catalog, we consider issues that arise when individuals may be identified from multiple, non-invasive marks. Past approaches to this problem include modeling data from only one mark and combining inferences obtained from each mark separately under the assumption that marks are independent. We describe a Bayesian method that makes use of the data from multiple marks and properly accounts for the dependence between marks from the same individual. Results from the analysis of the ECOCEAN data and simulation studies are included and show that the new method provides more precise inference than analyzing the data from one mark while maintaining nominal coverage rates for credible intervals.

Keywords: Latent multinomial model; Mark-recapture; Multiple marks; Non-invasive marks; Photo-identification; Whale sharks
1 Introduction

Non-invasive marks (also called natural marks) include patterns in pigmentation, genetic markers, acquired scars, or other natural characteristics that allow researchers to identify individuals in a population without physical capture. Such marks have been used for a long time in mark-recapture studies of some classes of animals, particularly marine mammals, and are now being used more widely. The advantages of non-invasive marks over man-made marks are first that non-invasive marks can be observed passively from a distance or through the collection of secondary material (hair samples or scat) so that individuals do not have to be handled and second, depending on the type of mark, that every individual is marked from birth. However, non-invasive marks also present several modeling challenges including that marks may be misidentified at non-negligible rates, that individuals’ marks may change over time, and that certain types of marks may be restricted to a subset of the population. We consider the problem of modeling about the demographics of a population based on mark-recapture data when individuals may be identified from multiple non-invasive marks.

The specific application we consider concerns the aggregation of whale sharks in Ningaloo Marine Park (NMP), off the west coast of Australia, each year between April and July. During this time, individual whale sharks are located by tour companies and photographs are taken by tourists and tour operators who upload their images to the online ECOCEAN whale shark library. Whale sharks can be identified by the unique pattern of spots on their flanks, and computer assisted methods are used to match photographs in the library and generate capture histories which provide information about the timing of the sharks’ arrival and departure from NMP and their survival across years [Holmberg et al., 2009], Yoshizaki et al. (2009) and Yoshizaki et al.
(2011) list many more applications involving the use of non-invasive marks with wide range of species including photographs of large cats (cheetahs, snow leopards, and tigers), scar patterns on marine mammals (manatees and whales), skin patterns of reptiles and amphibians (snakes, crocodiles, and salamanders), and genetic marks (bears, wombats, and whales).

Though non-invasive marks provide advantages they also present modeling challenges that do not arise with man-made tags. Da-Silva et al. (2003) and Da-Silva (2006) consider difficulties that occur if only a portion of the population possesses the non-invasive mark; for example, if individuals are identified from body scars that they acquire over time. Some individuals are effectively uncatchable because they cannot be identified, and abundance estimates must be inflated to account for this portion of the population. A related challenge is that non-invasive marks may change over time, as individuals acquire new scars or their skin patterns change. Yoshizaki et al. (2011) calls these as “evolving marks” and develops a least squares based method that allows for estimation of population demographics under the assumption that it is impossible to match the mark before and after the change creating a “ghost capture history” (Yoshizaki et al., 2011, pg. 29).

Work has also been done on the problem of misidentification of non-invasive marks, particularly for genetic marks. Lukacs and Burnham (2005) develops an extension of the Jolly-Seber model to account for misidentification of genetic markers under the assumption that the same error cannot occur twice and that errors do not produces spurious matches with other individuals. This produces an apparent excess of individuals captured once from which the rate of misidentification is estimated. Yoshizaki et al. (2011) shows that the estimators of Lukacs and Burnham (2005) are biased, even in large samples, and presents an alternative least squares approach based on the same assumptions. Further, Wright et al. (2009) presents a method that relaxes
the assumptions that errors in identification when the probability that markers are misread can be estimated from multiple tests of each genetic sample.

The problem we address arises when a single individual can be identified from more than one non-invasive mark, but the marks are not always linked. This occurs in our example because a single whale shark may photographed from either the right or the left side, but the spot patterns on each side are distinct and cannot be linked without further information. This means that it may not be possible to connect observations of a shark photographed one from the left and later from the right and to know that the same individual was in fact encountered twice. Naively modeling the observed encounter histories will artificially inflate the number of individuals observed and create dependence between the encounter histories, violating a key assumption of most mark-recapture models. One solution is to construct encounter histories from a single mark, [Holmberg et al. (2009)](Holmberg:2009) considers only left-side photographs, but this may remove a large number of histories from the data. Alternatively, [Wilson et al. (1999)](Wilson:1999) combines inferences from each mark post hoc by averaging separates estimates of abundance and computing standard errors under the assumption that these estimates are independent. It is simple to show that the final point estimate has similar properties to the individual estimates – the overall estimate will be unbiased/consistent if each of the separate estimates is unbiased/consistent – but the assumption of independence is violated and the resulting standard errors will be biased low. [Madon et al. (2011)](Madon:2011) describes another method to estimate abundance by adjusting the sufficient statistics required to compute the Jolly-Seber estimator. While the approach seems plausible, we have concerns with the implementation; though the observed counts underestimate some of the statistics and overestimate others, [Madon et al. (2011)](Madon:2011) uses a single adjustment factor and constrains its value to be between 0 and 1. [Bonner (2012)](Bonner:2012) discusses these issues and shows that the correct adjustment factor is not constant.
and may be greater than 1 for some statistics. Moreover, the simulations [Madon et al. (2011)] indicates a clear problem in that the coverage of confidence intervals is much lower than their nominal value, even when the population is large and the capture probability is close to 1.

Our objectives are to construct a model that accounts for the dependence between multiple non-invasive marks and provides valid inference about the population's demographics while making use of all available data. We do so by constructing an explicit model of the observation process that allows for multiple marks to be observed on different occasions and then applying Bayesian methods of inference via Markov chain Monte Carlo (MCM) sampling. Our model can be cast as a special case of the latent multinomial model (LMM) of [Link et al. (2010)]. Section 2 describes the whale shark data and Section 3 gives details on the structure of our model and our methods of inference. We provide results from a simulation study and the application of our method to the whale shark data in Sections 4 and 5, and conclude with a discussion of the method and further extensions in Section 6.

2 Data

Data for our analysis was obtained from the ECOCEAN on-line whale shark library (available at [www.whaleshark.org](http://www.whaleshark.org)). This library contains photographs of whale sharks taken by recreational divers and tour operators worldwide and submitted online. Computer assisted algorithms are then used to identify the pattern of spots on the shark's flank and scan the database for matches with previous photographs which are confirmed visually. Quality control studies indicated that the probability of a mismatch is approximately .02 ([Holmberg et al., 2009](http://www.whaleshark.org)). The library has been operational since 2003, and over 41,000 photographs have been submitted by more
than 3300 contributors. Further details on the study site, the observation protocols, and the algorithms for matching photographs are provided in Holmberg et al. (2009).

We model only the data collected from the northern ecotourism zone of NMP during the 16 week period between April 1 and July 31, 2008, which we have divided into 8 capture periods of 2 weeks each. Sharks may be encountered multiple times during a single capture period and so five possible events may occur. In each 2 week period, a shark may:

1) not be encountered at all (event 0)

2) be photographed from the left only (event L),

3) be photographed from the right only (event R),

4) be photographed from both sides simultaneously on at least one encounter (event S), or

5) be photographed from both sides but never simultaneously (event B).

Problems with identification arise because it is only possible to match the skin patterns from the two sides of a shark if photographs of both sides are taken simultaneously during at least one capture period – i.e., there is at least one S in its encounter history. Otherwise, an individual that was photographed from both sides will contribute two encounter histories to the data set – one containing the observations of its right side and the other containing the observations of its left side.
3 Methods

3.1 Modeling

The challenge in modeling this data is that some encounter histories are unobservable so that the recorded histories may not reflect the true histories of the encountered individuals. Suppose that an individual’s true encounter history is 00L0B0R0. This history is not observable because the two sides of the individual were never photographed simultaneously and the marks on its right and left sides cannot be linked. Instead, the individual will contribute two observed histories to the data – 00L0L000 and 0000R0R0. Working in the opposite direction, if the observed data includes the histories 00L0L000 and 0000R0R0 then it is not possible to know if these histories come from one individual encountered on three occasions or from two individuals encountered twice each.

With $K = 8$ capture occasions there are $5^8 - 1 = 390,624$ possible true encounter histories (we ignore the zero history because our model conditions on an individual being captured once at least). Of these, 325,558 (83.3%) can occur in the observed data because they contain photographs from the left-side only, photographs from the right-side only, or at least one simultaneous encounter. The remaining 65,026 (16.7%) contain photographs taken from both the left and right sides, including via event B, but never simultaneously and are unobservable. In practice, we need not consider all 65,026 unobservable histories. Given a set of observed histories, the possible true histories can be constructed by combining each history containing only left-side photographs with each history containing only right-side photographs as follows: when event L appears in the first history and 0 appears in the second enter L in the combined history, when 0 appears in the first history and R in the second enter R in the combined history, and when L appears in the first and R in the second then B is
To account for the uncertainty in the true encounter histories when modeling the demographics of the population we adapt the LMM of [Link et al. (2010)]. This model applies when the latent counts of the true histories follow a multinomial distribution and determine the observed counts through a known, linear relationship. Suppose that $L$ unique histories were observed, let $n_j$ be the number of times that the $j^{th}$ history was observed, and set $n = (n_1, \ldots, n_L)^T$. Further, suppose that there are $L^*$ true histories that could have generated the observed histories, let $n^*_j$ be the number of individuals with the $j^{th}$ true history, and set $n^* = (n^*_1, \ldots, n^*_{L^*})^T$. The LMM supposes first that $n = An^*$ for some fixed matrix $A$ and second that $n^*$ has a multinomial distribution whose density, $f(n^*|n^*, \theta)$, depends on $n^*$, representing either the total population size or the number of unique individuals encountered, and the vector of parameters $\theta$. The likelihood for this model is proportional to the probability of $n$ given both $n^*$ and $\theta$ and is computed by summing $f(n^*|n^*, \theta)$ over all possible configurations of $n^*$ that satisfy equation (??) and produce the correct value of $n^*$ ($n^* = \sum n^*_j$). Evaluating this sum directly is generally not possible, and, instead, [Link et al. (2010)] applies Bayesian inference treating $n^*$ as a random variable and sampling from the joint posterior distribution

$$
\pi(n^*, n^*, \theta|n) \propto 1(n = An^*)f(n^*|n^*, \theta)\pi(n^*, \theta)
$$

(1)

where $\pi(n^*, \theta)$ is the chosen prior for $n^*$ and $\theta$.

To apply the LMM to our problem, we need to identify the set of possible true histories that could have generated the observed data and the relationship between the corresponding counts. Suppose that the observed data contains $L_1$ unique histories including only left-side photographs, $L_2$ unique histories including only right-side
photographs, and \( L_3 \) unique histories with at least one encounter of both sides simultaneously so that \( L = L_1 + L_2 + L_3 \). Let \( W_1 \) be the \( L_1 \times K \) matrix whose rows contain those histories including only left-side photographs and \( n_1 \) the \( L_1 \)-vector of observed counts for each row of \( W_1 \). Define \( W_2, n_2 \) and \( W_3, n_3 \) similarly for those histories including only right side photographs and those including at least one simultaneous encounter so that the complete vector of observed counts is \( n = (n_1^T, n_2^T, n_3^T)^T \).

The additional entries in the set of possible true histories are then generated by combining each row of \( W_1 \) with each row of \( W_2 \), as described above. Let \( W_4 \) denote the \( L_4 \times K \) matrix of unobservable histories \( (L_4 = L_1 L_2) \) whose \( L_1 (i - 1) + j \) row is formed by combining row \( i \) of \( W_1 \) with row \( j \) of \( W_2 \). Let \( n_{4ij}^* \) denote the number of individuals encountered whose true encounter history given by the \( L_1 (i - 1) + j^{th} \) row of \( W_4 \) and \( n_4^* = (n_{411}^*, \ldots, n_{4L_11}^*, \ldots, n_{41L_2}^*, \ldots, n_{4L_1L_2}^*)^T \) so that the complete vector of true counts is \( n^* = (n_1^* T, n_2^* T, n_3^* T, n_4^* T)^T \).

The set of linear constraints mapping \( n^* \) to \( n \) is then obtained by considering how each true history contributes to each observed history. Specifically, the number of times that a history in \( W_1 \) or \( W_2 \), say \( w \), is observed must equal the sum of the counts for all of the true histories that could have generated \( w \). These constraints are determined by the set of \( L \) linear equations

\[
\begin{align*}
n_{1i} &= n_{1i}^* + \sum_{j=1}^{L_2} n_{4ij}^*, \quad i = 1, \ldots, L_1 \\
n_{2j} &= n_{2j}^* + \sum_{i=1}^{L_1} n_{4ij}^*, \quad j = 1, \ldots, L_2 \\
n_{3k} &= n_{3kj}^*, \quad k = 1, \ldots, L_3
\end{align*}
\]

one for each observed history. Note that there are \( L \) equations and \( L^* \) unknowns, implying that there are \( L^* - L = L_4^* \) free parameters. It is most convenient to choose
\( \mathbf{n}^*_4 \) as the set of free parameters, and the possible values can be bounded by noting that \( 0 \leq n^*_4_{ij} \leq \min(n_{1i}, n_{2j}) \). This limits the possible configurations of \( \mathbf{n}^*_4 \) and makes it easier to sample from the posterior distribution, as described in Section 3.3.

As an example, suppose that the observed data comprises the six histories listed in the top of Table 1. The possible true histories that generated this data include the six observed histories plus the four unobservable histories generated by combining the histories including only of left-side photographs (histories 1&2) with the histories including only right-side photographs (histories 3&4), as given in the bottom of the table. The relationship between \( \mathbf{n} \) and \( \mathbf{n}^* \) is then determined by the six linear equations

\[
\begin{align*}
n_1 &= n^*_1 + n^*_7 + n^*_9, \quad n_3 = n^*_3 + n^*_7 + n^*_8, \quad n_5 = n^*_5 \\
n_2 &= n^*_2 + n^*_8 + n^*_10, \quad n_4 = n^*_4 + n^*_9 + n^*_10, \quad n_6 = n^*_6.
\end{align*}
\]

The free parameters are \( n^*_7, \ldots, n^*_10 \) and these values are bounded so that \( 0 \leq n^*_7 \leq \min(n_1, n_3), 0 \leq n^*_8 \leq \min(n_2, n_3), 0 \leq n^*_9 \leq \min(n_1, n_4), \) and \( 0 \leq n^*_10 \leq \min(n_2, n_4) \).

[Table 1 about here.]

### 3.2 Model of Observed Encounter Histories

This framework for modeling data from multiple non-invasive marks is general and can incorporate any of the standard mark-recapture models for the true encounter histories. We are most interested in the timing of the sharks’ arrival and departure from NMP and have developed an extension of the Link-Barker-Jolly-Seber (LBJS) model [Link and Barker, 2005] which accounts for the presence of multiple marks. The LBJS model describes the demographics of a population through two sets of
parameters: 1) the recruitment rate, \( f_k, k = 1, \ldots, K - 1 \), defined as the number of individuals that enter the population between occasions \( k \) and \( k + 1 \) per individual present on occasion \( k \) and 2) the survival probability, \( \phi_k, k = 1, \ldots, K - 1 \), defined as the probability that an individual present on occasion \( k \) is also present on occasion \( k + 1 \). For the whale sharks, \( \phi_k \) corresponds to the probability that an individual remains at NMP rather than true survival. It is assumed that all emigration is permanent, that individuals behave independently, and that the apparent survival probability does not depend on how long an individual has been present (or any other factors).

The process of observing individuals is then modeled in terms of two further sets of parameters: 1) the probability that an individual is captured/encountered given that it is present in the study area and 2) the conditional probabilities of each event. Specifically, we define \( p_k \) to be the probability that an individual is photographed at least one time during occasion \( k \) given that it is present, \( k = 1, \ldots, K \), and \( \rho_j \) to be the probability that the individual is photographed from the left-side only \((j = 1)\), from the right-side only \((j = 2)\), from both sides simultaneously at least once \((j = 3)\), or from both sides but never simultaneously \((j = 4)\) during period \( k \). This last set of parameters is what distinguishes our model of the true histories from the original model of \cite{link2005}. Our model assumes that all individuals present on an occasion have the same probability of capture, that encounters are independent between individuals and across the occasions, and that the conditional event probabilities do not vary between individuals or over time.
3.3 Inference

As noted by Link et al. (2010), the LMM lends itself to Bayesian inference via MCMC sampling with data augmentation, and the Bayesian specification of the model is completed by assigning prior distributions to all of the parameters. We have selected non-informative, hierarchical prior distributions where possible. Specifically, the survival and capture probabilities were assigned prior distributions

\[
\text{logit}(\phi_k) \sim N(\mu_\phi, \sigma_\phi^2), \quad k = 1, \ldots, K - 1
\]

\[
\text{logit}(p_k) \sim N(\mu_p, \sigma_p^2), \quad k = 1, \ldots, K
\]

with

\[
\mu_\phi, \mu_p \sim N(0, 2) \quad \text{and} \quad \sigma_\phi, \sigma_p \sim HT(3, .9)
\]

where \(HT\) denotes the half \(t\)-distribution with 3 degrees of freedom and scale parameter \(.9\). These hyperparameters were chosen because they provide distributions that are close to uniform on \((0, 1)\) for both the hierarchical means, \(\mu_\phi\) and \(\mu_p\), and the base demographic parameters, \(\phi_k\) and \(p_k\). The recruitment probabilities were assigned prior distributions of the form

\[
\log(f_k) \sim N(\mu_f, \sigma_f^2), \quad k = 1, \ldots, K - 1
\]

with

\[
\mu_f \sim N(0, .25) \quad \text{and} \quad \sigma_f \sim HT(3, .9),
\]

and the number of unique individuals encountered was assigned the discrete uniform prior on \(0, \ldots, M\) for some \(M\) guaranteed not to be smaller than \(n^*\). This is satisfied by the data dependent prior \(n^* \sim U\{0, \ldots, L_1 + L_2 + L_3\}\), because
\[ n^* = \sum_{j=1}^{L^*} n_j^* \leq \sum_{j=1}^{L} n_j, \text{ but the exact value does not affect our computations.} \]

Finally, the conditional event probabilities were assigned a Dirichlet prior with parameter \( \alpha = 1_4 \) implying that \( \rho_1, \rho_2, \rho_3, \) and \( \rho_4 \) were marginally distributed \textit{a priori} with identical beta distributions having mean .25 and probability .9 between .00 and .86. All prior distributions were assumed independent.

An MCMC algorithm for sampling from the distribution in (1) can then be implemented as follows. Letting \( n^*(k), n^*(k), \) and \( \theta^{(k)}, \) denote the values sampled on the \( k^{th} \) iteration, the values on the next iteration are generated in two steps:

1. Update \( n^*(k) \) and \( n^*(k) \) given \( \theta^{(k)} \) by:
   a) setting \( n^{*(\text{curr})} = n^*(k) \) and \( n^{*(\text{curr})} = n^*(k), \)
   b) updating \( n^*_{411} \) by:
      i) setting \( n^{*'} = n^{*(\text{curr})} \)
      ii) sampling a new value for \( n^*_{411} \) from \( \{0, \ldots, \min(n_{11}, n_{21})\} \) according to some proposal distribution, \( q(n^*_{411}|n^{*(\text{curr})}, \theta^{(k)}), \)
      iii) ensuring the constraints in equation (2) remain satisfied by setting
         \[
         n^*_{11} = n_1 - \sum_{j=1}^{L_2} n^*_{41j},
         \]
         \[
         n^*_{21} = n_2 - \sum_{i=1}^{L_1} n^*_{4i1}, \text{ and}
         \]
         \[
         n^* = 1^T n^{*'},
         \]
   iv) rejecting \( n^{*'} \) if \( n^*_{11} < 0 \) or \( n^*_{21} < 0 \) and otherwise accepting \( n^{*'} \) and setting
      \( n^{*(\text{curr})} = n^{*'} \) and \( n^{*(\text{curr})} = n^{*'} \) with probability:
      \[
      \alpha(n^*, n^{*'}) = \min \left\{ 1, \frac{f(n^{*'}|n^{*'}, \theta^{(k)}) q(n^*_{411}^{*(\text{curr})}|n^{*'}, \theta^{(k)})}{f(n^{*(\text{curr})}|n^{*(\text{curr})}, \theta^{(k)}) q(n^*_{411}^{*(\text{curr})}|n^{*(\text{curr})}, \theta^{(k)})} \right\}
      \]
c) repeating step b) for each \( n_{4ij}^* \), \( i = 1, \ldots, L_1 \), \( j = 1, \ldots, L_2 \), and finally
d) setting \( n^{*(k+1)} = n^{*(\text{curr})} \) and \( n^{*(k+1)} = n^{*(\text{curr})} \).

2. Update \( \theta^{(k)} \) given \( n^{*(k+1)} \) and \( n^{*(k+1)} \) through a series of Metropolis-Hastings (MH) steps appropriate for the selected model.

The steps for updating each element of \( n^* \) in this algorithm have an intuitive interpretation. In step i), a new count is proposed for one of the unobservable histories. In step ii), the counts of the corresponding observed histories in \( W_1 \) and \( W_2 \) are adjusted to maintain the correct observed counts. The new proposal is rejected immediately if this produces negative counts, and otherwise, a MH accept-reject step is conducted. One choice for \( q(n_{4ij}^*|n^*, \theta) \) is the discrete uniform on \( \{0, \ldots, \min(n_{1i}, n_{2j})\} \) in which case the proposal, \( n^{'*} \) is accepted whenever the likelihood is increased, \( f(n^*|n^{'*}, \theta) > f(n^{*(\text{curr})}|n^{*(\text{curr})}, \theta) \).

This MCMC sampling algorithm was implemented in JAGS using the rjags interface to the R software package \( \text{[Plummer, 2003, 2011, Team, 2012]} \). Three chains were run in parallel in order that convergence could be monitored with the Gelman-Rubin-Brooks (GRB) diagnostic \( \text{[Brooks and Gelman, 1998]} \) as implemented in the R package CODA \( \text{[Plummer et al., 2006]} \). One challenge we faced was obtaining initial values for different sets of parameters were dispersed but consistent with one other. To do this, we generated initial values for each chain in three steps. First, we simulated random values of \( \phi_k \) and \( f_k \), \( k = 1, \ldots, K - 1 \), under one of three schemes –

1) \( \text{logit}(\phi_k) \sim N(\text{logit}(.9), .1) \) and \( \text{log}(f) \sim N(\text{log}(.1), .1) \), 2) \( \text{logit}(\phi_k) \sim N(., 1) \) and \( \text{log}(f) \sim N(\text{log}(.5), .9) \), 3) \( \text{logit}(\phi_k) \sim N(\text{log}(1.1), .1) \) and \( \text{log}(f) \sim N(\text{log}(.9), .1) \).

We then obtained values of \( p_k \), \( k = 1, \ldots, K \), by fitting the LBJS model to the data obtained from the left-side photographs only but treating the sampled values of \( \phi_k \) and \( f_k \) as fixed. Finally, we obtained values of \( n^* \) by fitting the our model...
to the full data set but treating all of $\phi_k$, $f_k$, and $p_k$ as fixed. An R package which implements these methods and provides functions to format the data and to call JAGS through the rjags interface is available from the website of the first author at www.simon.bonners.ca/MultiMark.

4 Simulation Study

To assess the performance of the model presented in the previous section we conducted simulation studies under a variety of scenarios. Here we present the results from three simulation scenarios which illustrate our main results.

In the first scenario, we compared the performance of the new model (henceforth the two-sided model) with two alternative models. First, we considered models using reduced data sets containing only data from the left or right side photographs (the one-sided models). Capture histories were constructed by combining all events that include a left-side photograph, namely L, S, and B, and all right-side photographs were ignored, or vice versa. The models we fit to this data were equivalent to the LBJS model with prior distributions as given in Section 3.3. Second, we developed a Bayesian equivalent of the method of combining inferences from the two sides under the assumption of independence as in (Wilson et al., 1999) (combined inference). To do this, we fit separate models to the data from the left and right side photographs and then averaged the values sampled on each iteration of the separate MCMC samplers prior to computing summary statistics. For example, let $\phi_t^{(k,L)}$ and $\phi_t^{(k,R)}$ represent the values of $\phi_t$ drawn on the $k^{th}$ iterations of the MCMC samplers run separately for models of the the left and right-side data and $\text{Var}^{(L)}(\phi_t)$ and $\text{Var}^{(R)}(\phi_t)$ be the separate posterior variances. Combined inference for $\phi_t$ was obtained by first computing the
inverse variance weighted average of the samples drawn on the $k^{th}$ iteration

$$
\phi_t^{(k)} = \frac{\text{Var}^{(R)}(\phi_t)\phi_t^{(k,L)} + \text{Var}^{(L)}(\phi_t)\phi_t^{(k,R)}}{\text{Var}^{(L)}(\phi_t) + \text{Var}^{(R)}(\phi_t)}
$$

and then computing summary statistics from the new chain $\phi_t^{(1)}, \phi_t^{(2)}, \ldots$. The advantage of this method is that posterior standard deviations and credible intervals can be computed directly from the new chain without distributional assumptions. Note that the mean of the values in the new chain is exactly equal to the inverse variance weighted average of means from the separate chains. We expected that, in general, posterior standard deviations and credible intervals from the one-sided models would be larger/wider than the corresponding standard deviations and intervals from the two-sided model, and that posterior standard deviations and credible intervals produced by combined inference would be smaller/narrower than those from the two-sided model but would not achieve the nominal coverage probability.

One hundred data sets for this scenario were generated under the assumption that each event was equally likely conditional on capture ($\rho_1 = \rho_2 = \rho_3 = \rho_4 = .25$). Simulated capture histories included 10 capture occasions, and data were generated by simulating true capture histories for individuals sequentially until 200 observed capture histories were produced (each true history contributing either 0, 1, or 2 histories to the observed data). Survival and capture probabilities for each occasion were generated from independent logit-normal distributions with means $\text{logit}(.8)$ and $\text{logit}(.5)$, and standard deviation $.30$. Recruitment rates were generated from independent log-normal distributions with mean $\text{log}(.25)$ and standard deviation $.30$. The median number of true histories simulated before 200 observed histories were obtained was 164 (min=150,max=180), the median number of unique individuals observed at least one time was 138 (min=127,max=148), and the median number of captures was 2.
Table 2 presents statistics comparing the mean-squared error (MSE) of the posterior means and the mean width and estimated coverage probability of the 95% credible intervals obtained from the alternative models. The MSE of the two-sided model and the combined-inference were similar for all parameters and smaller than those of the one-sided model by between 10% and 25%. Credible intervals for both the one-sided and two-sided model achieved the nominal coverage rate for all parameters, but the credible intervals for the one-sided model were wider by approximately 10%. In comparison, the credible intervals from the combined inference were narrower than those of the two-sided model by 20% or more but failed to achieve the nominal coverage rate.

In the second scenario, we simulated data under the same model except that the capture probabilities were generated from a logit-normal distribution with mean $\text{logit}(0.5)$. The result is that capture probabilities were lowered and the number of observations per individual decreased. From the 100 data sets generated, the median number of true histories simulated before 200 observed histories were obtained was 196 ($\text{min}=170, \text{max}=235$), the median number of unique individuals observed at least one time was 131 ($\text{min}=119, \text{max}=143$), and the median number of captures per observed individual was 1 ($\text{min}=1, \text{max}=9$). In this scenario, the MSEs of the posterior mean survival probabilities was similar between the two-sided model and the combined inference, but the two-sided model produced higher MSEs for both the recruitment and population growth rates. Credible intervals from the two-sided model were again narrower than those of the one-sided model by approximately 10%, but coverage probabilities from both models were above the nominal rate. Finally, although the credible intervals from the combined inference were narrower than those of the two-sided model by approximately 20%, the coverage of these intervals was also increased.
and was near or above the nominal value. This result was due to the combination
of non-informative priors and the relatively small amount of information in the data
when capture probabilities are low; we discuss this point in Section 6.

In the third scenario, we simulated data from the same model used in the first
scenario except that both marks were seen with probability one each time an indi-
vidual was captured ($\rho_3 = 1$). This represents the extreme situation in which there is
complete dependence between the two marks and no uncertainty in the true capture
histories. Modeling the data from the left-side photographs, right-side photographs,
or from both-sides using our model produces exactly the same results, and so we only
compare the model of the left-side photographs with combined inference. The me-
dian number of histories simulated in the 100 data sets before 200 observed histories
were obtained was 215 (min=204,max=227) and the median number of captures per
observed individual was 2 (min=1,max=10). In this scenario, the point estimates
produced by the two models were almost exactly equal and the MSE of the two mod-
els was indistinguishable. However, there were clear differences in the performance of
the 95% credible intervals. While the intervals produced by combined-inference were,
on average, 30% narrower the coverage of these intervals was well below the nominal
value.

[Table 2 about here.]

5 Results

The data provided in the ECOCEAN whale shark library contained a total of 96
observed encounter histories for the 2008 study period. Of these, 27 histories (28%) were constructed from left-side photographs alone, 24 (25%) were constructed from
right-side photographs alone, and 45 (47%) contained at least one encounter with
photographs taken from both sides simultaneously. Along with the model presented in Section 3, we computed inferences for $p$, $f$, and $\phi$ from the three alternative models described in Section 4.

Table 3 provides posterior summary statistics for the LBJS parameters model obtained from the two-sided model. Inferences about all parameters are relatively imprecise because of the relatively small number of individuals captured and the low capture probabilities, but the posterior means follow the expected patterns. Point estimates for the survival probability (the probability that a whale shark remains NMP between occasions) are at or above .90 in the first 2 periods, below .70 in the last two periods, and about .80 in between. The posterior mean recruitment rate is very high in week 2, suggesting that most of the sharks entered during this period, and lower thereafter. This table also provides summary statistics for the population growth rate, $\lambda_k = \phi_k + f_k$, $k = 1, \ldots, K - 1$, computed as a derived parameter. Although the 95% credible intervals for $\lambda_k$ cover 1.00 for all $k$, the point estimates are greater than 1.00 for the first two periods, close to 1.00 in the next three periods, and less than .75 in the last two periods. This suggests that the aggregation of whale sharks grew during the first two periods, remained almost steady during the next 3 periods, and declined during the last two periods. This supports the hypothesis that whale sharks aggregate at NMP to feed after the major coral spawn which occurred between April 9 and 12 in 2008 (Chalmers, 2008, pg. 33).

Table 4 provides posterior summary statistics for the conditional event probabilities. These results show that when sharks were encountered photographs were most often taken from both sides simultaneously ($\hat{\rho}_3 = .45(.36, .54)$) and that the probabilities that an individual was photographed from either the left- or right-side only

[Table 3 about here.]
were similar ($\hat{\rho}_1 = .29(.20, .38)$ versus $\hat{\rho}_2 = .21(.13, .29)$).

The posterior mean of $n^*$, the number of unique sharks encountered during the 2008 season, was 88 with 95% credible interval (82,93). The full posterior distribution of $n^*$ is shown in Figure[1] and compared with the prior distribution of $n^*$ generated by simulating data sets from the prior predictive distribution conditional on there being 96 observed capture histories and at least 72 true histories (the minimum number given that 24 of 96 observed histories included right-side photographs alone). Whereas the prior distribution of $n^*$ is close to uniform the posterior distribution is strongly peaked and concentrates 95% of its mass between 82 and 93. We conclude that between 3 (3.1%) and 14 (14.6%) of the sharks encountered during the 2008 season were photographed from both the left- and right-sides on separate occasions without ever being matched.

[Table 4 about here.]

[Figure 1 about here.]

Comparisons of the three chains starting from different initial values provided no evidence of convergence problems. Traceplots all indicated that the three chains converged within the burn-in period, GRB diagnostic values were all less than 1.02, and the estimated MCMC error was less than 2.6% of the posterior standard deviation for each parameter. Based on these results, we are confident that the chains converged and were run for sufficiently long to obtain reliable summary statistics.

The plots in Figure[2] compare inferences for the survival, recruitment, and growth rates from the four alternative models. Posterior means from the four models are all very similar and the 95% credible intervals for all parameters overlap considerably. Comparison of the widths of the 95% credible intervals from the left- and right-side
data alone showed that the two-sided model provided improved inference in most, but not all, parameters. On average, the 95% credible intervals for the recruitment rates produced by the two-sided model were 93% and 69% as wide as those produced from the left- and right-side data alone. The 95% credible intervals for the survival probabilities produced by the two-sided model were 78% as wide as those from the right-side data, on average, but 103% as wide as those from the left-side data. This last result seems to be caused by issues with the upper bound on the survival probabilities as the 95% credible intervals for the logit transformed survival probabilities produced from the two-sided model were, on average, 90% and 89% as wide as those obtained from the left- and right-side data alone. Credible intervals produced via combined inference were on average 12% smaller than those obtained from the two-sided model; however, based on the results in the previous section, we believe that these intervals would not achieve the nominal coverage rate and do not reflect the variability of the parameters correctly.

6 Conclusion

The simulation results presented in Section 4 illustrate the main advantages of our model over the previous approaches to analyzing mark-recapture data with multiple non-invasive marks In general, estimates from our model will be more precise than estimates based on only one mark in that posterior standard deviations will be smaller and credible intervals will be narrower but still achieve the nominal coverage rate. In contrast, the apparent gain from combining estimators computed separately for each mark under the assumption of independence is artificial and credible/confidence intervals computed by these methods will not achieve the nominal coverage rate –
particularly when the probability that both marks are seen simultaneously is high and the separate estimators are highly dependent. The differences between the methods are smaller when the capture probabilities are low (as in Simulation 2) but this is not surprising given that we have selected non-informative prior distributions. The posterior distribution is influenced more by the prior as the amount of information in the data decreases, and this would not have occurred if more informative priors had been selected.

The disadvantage of the two-sided model is that the computations are more complex and more time consuming. The major challenge is that the number of possible true capture histories grows relative to the square of the number of observed capture histories. Possible solutions include deriving more efficient methods of computation or approximating the posterior distribution in some way. Further research is needed in this area before applying our model to large data sets.

In the analysis of the whale shark data, we have focused on understanding the dynamics of the population (i.e., survival, recruitment, and population growth rates) but the method is easily adapted to estimate abundance. Following Link et al. (2010), one can include the null encounter history (vector of 0s) to the set of possible true histories and extend the vector of counts for the true histories to \( \mathbf{n}^* = (n_1^*, \ldots, n_4^*, n_0^*) \) where \( n_0^* \) represents the number of individuals in the population that were never captured. The value \( n^* = \mathbf{1}^T \mathbf{n}^* \) would then denote the total population size. Note that the prior on \( n^* \) must also be modified by increasing \( M \) to allow \( n^* \) to be larger than the number of observed histories.

Our model can also be extended to combine data from any number of marks by expanding the set of constraints, and we expect that including more marks will strengthen differences between the alternative methods seen in the simulation study. The ratio between the median width of the credible intervals from the one-sided model
and the combined inference is close to 1.4 for all parameters in all three simulation scenarios – very near the $\sqrt{2}$ reduction in standard deviation that occurs when two identically distributed and perfectly correlated random variables are averaged. Generally, posterior standard deviations computed under the assumption of independence should decrease relative to the square root of the number of marks used, but the gain is artificial and the coverage of credible/confidence intervals will decrease further below the nominal value as more marks are used. In comparison, our model will account for the dependence between marks and will provide more precise inference without sacrificing coverage.

One unusual aspect of the whale shark study is that individuals may be encountered multiple times during each capture period. This would not occur if the capture periods were truly instantaneous, and in such cases both marks could only be observed in the same period if they were seen simultaneously. Event B could never occur. Our model could be fit to such data simply by specifying, \textit{a priori}, that $P(\rho_4 = 0) = 1$. and other researchers have begun to develop similar methods specific to such data (Conn and McClintock, pers. comm.).

**Acknowledgments**

We thank Laura Cowen for providing comments on a previous draft of the manuscript and Matt Schofield for valuable discussions during the development of our model. Support for this work was provided in part by the NSF-Kentucky EPSCOR Grant (NSF Grant No. 0814194).
References

Bonner, S. J. (2012). Comment on: A new method for estimating animal abundance with two sources of data in capture-recapture studies. *Methods in Ecology and Evolution* Submitted.

Brooks, S. P. and Gelman, A. E. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7, 434–455.

Chalmers, A. (2008). *Temporal and spatial variability in coral condition at Sandy Bay, Ningaloo*. PhD thesis, The University of Western Australia.

Da-Silva, C. Q. (2006). Asymptotics for a population size estimator of a partially uncatchable population. *Annals of the Institute of Statistical Mathematics* 59, 603–615.

Da-Silva, C. Q., Rodrigues, J., Leite, J. G., and Milan, L. A. (2003). Bayesian estimation of the size of a closed population using photo-id data with part of the population uncatchable. *Communications in Statistics - Simulation and Computation* 32, 677–696.

Holmberg, J., Norman, B., and Arzoumanian, Z. (2009). Estimating population size, structure, and residency time for whale sharks *rhincodon typus* through collaborative photo-identification. *Endangered Species Research* 7, 39–53.

Link, W. A. and Barker, R. J. (2005). Modeling association among demographic parameters in analysis of open population capture-recapture data. *Biometrics* 61, 46–54.
Link, W. A., Yoshizaki, J., Bailey, L. L., and Pollock, K. H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics* **66**, 178–185.

Lukacs, P. M. and Burnham, K. P. (2005). Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. *Journal of Wildlife Management* **69**, 396–403.

Madon, B., Gimenez, O., McArdle, B., Scott Baker, C., and Garrigue, C. (2011). A new method for estimating animal abundance with two sources of data in capture-recapture studies. *Methods in Ecology and Evolution* **2**, 390–400.

Plummer, M. (2003). Jags : A program for analysis of bayesian graphical models using gibbs sampling jags. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, pages 1–10, Vienna, Austria.

Plummer, M. (2011). *rjags: Bayesian graphical models using MCMC*. R package version 3-5.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6**, 7–11.

Team, R. D. C. (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Wilson, B., Hammond, P. S., and Thompson, P. M. (1999). Estimating size and assessing trends in a coastal bottlenose dolphin population. *Ecological Applications* **9**, 288–300.

Wright, J. A., Barker, R. J., Schofield, M. R., Frantz, A. C., Byrom, A. E., and
Gleeson, D. M. (2009). Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. *Biometrics* **65**, 833–40.

Yoshizaki, J., Brownie, C., Pollock, K. H., and Link, W. A. (2011). Modeling misidentification errors that result from use of genetic tags in capture-recapture studies. *Environmental and Ecological Statistics* **18**, 27–55.

Yoshizaki, J., Pollock, K. H., Brownie, C., and Webster, A. (2009). Modeling misidentification errors in capture-recapture studies using photographic identification of evolving marks. *Ecology* **90**, 3–9.
Figure 1: Comparison of the prior and posterior distribution of $n^*$. The prior distribution of $n^*$ conditional on there being 170 observed capture histories is shown by the histogram with white bars. The posterior distribution of $n^*$ is shown by the histogram with grey bars.
Figure 2: Comparison between the two-sided model and the three alternative models. The plots on the left-side of the figure compare the posterior means (points) and 95% credible intervals (vertical lines) of the survival probability (top), recruitment rate (middle), and population growth rate (bottom) obtained from the four models. The plots on the right side of the figure display the posterior standard deviations from the three alternative models relative to the posterior standard deviation from the two-sided model. Results from the two-sided model are represented by the circles, from the left-side photographs only by the upward pointing triangle, from the right-side photographs only by the downward pointing triangles, and from combined inference by the diamonds.
Table 1: Example of possible observed and true capture histories. Suppose that the data comprises the six observed histories given in the top of the table. The possible true histories that may have generated this data include these six plus the four additional histories in the bottom of the table.

| Occasion | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------|---|---|---|---|---|---|---|---|
| 1        | 0 | 0 | L | 0 | L | 0 | 0 | 0 |
| 2        | 0 | 0 | 0 | 0 | L | 0 | 0 | 0 |
| 3        | 0 | 0 | R | 0 | 0 | 0 | 0 | 0 |
| 4        | 0 | 0 | 0 | R | R | 0 | 0 | 0 |
| 5        | 0 | 0 | S | B | R | 0 | 0 | 0 |
| 6        | S | 0 | S | 0 | 0 | 0 | 0 | 0 |
| 7        | 0 | 0 | B | 0 | L | 0 | 0 | 0 |
| 8        | 0 | 0 | B | 0 | L | 0 | 0 | 0 |
| 9        | 0 | 0 | L | R | B | 0 | 0 | 0 |
| 10       | 0 | 0 | 0 | R | B | 0 | 0 | 0 |
Table 2: Performance of the estimates from the three simulation studies. Each column of the table presents the MSE of the posterior mean relative to the MSE of the posterior mean of the one-sided model, and the median width and estimated coverage probability of the 95% credible intervals for the survival probability ($\phi$), recruitment rate ($f$), and growth rate ($\lambda$) for one of the three models – one-sided (OS), two-sided (TS), or combined-inference (CI). Descriptions of the models are provided in the text in Section 4.

|        | Simulation 1 |         | Simulation 2 |         | Simulation 3 |         |
|--------|--------------|---------|--------------|---------|--------------|---------|
|        | OS           | TS      | CI           | OS      | TS           | CI      |
| $\phi$ | MSE          | .10     | .89          | .87     | .81          | .81     | .10     | .81          | .81     | .10     | .81          | .81     | .10     | .81          | .81     | .10     |
|        | Width        | .23     | .20          | .16     | .30          | .28     | .21     | .17     | .12     |
|        | Cover        | .97     | .96          | .90     | .98          | .99     | .93     | .95     | .84     |
| $f$    | MSE          | .10     | .88          | .81     | .86          | .72     | .10     | .81          | .81     | .10     | .81          | .81     | .10     |
|        | Width        | .35     | .31          | .24     | .48          | .42     | .32     | .26     | .18     |
|        | Cover        | .97     | .95          | .90     | .99          | .98     | .96     | .95     | .84     |
| $\lambda$ | MSE          | .10     | .88          | .82     | .86          | .73     | .10     | .81          | .81     | .10     | .81          | .81     | .10     |
|        | Width        | .41     | .36          | .29     | .56          | .50     | .38     | .31     | .22     |
|        | Cover        | .98     | .97          | .95     | .99          | .99     | .97     | .97     | .87     |

30
Table 3: Posterior summary statistics for the demographic parameters $\phi_k$, $f_k$, $\lambda_k$, and $p_k$ obtained from the two-sided model. The columns of the table provide posterior means followed with equal-tailed 95% credible intervals.

| Occ ($k$) | Survival ($\phi_k$) | Recruitment ($f_k$) | Growth ($\lambda_k$) | Capture ($p_k$) |
|-----------|----------------------|---------------------|----------------------|-----------------|
| 1         | 0.90(0.67,1.00)      | 0.36(0.00,1.93)     | 1.26(0.76,2.83)      | 0.23(0.08,0.43) |
| 2         | 0.92(0.73,1.00)      | 2.40(0.08,6.41)     | 3.31(1.00,7.33)      | 0.19(0.05,0.33) |
| 3         | 0.82(0.54,1.00)      | 0.17(0.00,0.72)     | 0.99(0.64,1.56)      | 0.26(0.15,0.43) |
| 4         | 0.77(0.45,0.99)      | 0.09(0.00,0.36)     | 0.85(0.51,1.20)      | 0.22(0.13,0.34) |
| 5         | 0.82(0.49,1.00)      | 0.23(0.00,0.79)     | 1.05(0.63,1.65)      | 0.22(0.12,0.36) |
| 6         | 0.48(0.14,0.96)      | 0.06(0.00,0.29)     | 0.54(0.17,1.12)      | 0.25(0.14,0.42) |
| 7         | 0.66(0.16,0.99)      | 0.09(0.00,0.42)     | 0.75(0.20,1.28)      | 0.20(0.06,0.37) |
| 8         | –                    | –                   | –                    | 0.18(0.03,0.34) |
Table 4: Posterior summary statistics for the conditional event probabilities.

| Event (j) | Cond. Prob. ($\rho_j$) |
|-----------|------------------------|
| 1         | 0.29(0.20,0.38)        |
| 2         | 0.21(0.13,0.29)        |
| 3         | 0.45(0.36,0.54)        |
| 4         | 0.06(0.01,0.13)        |