A spatial mark–resight model augmented with telemetry data

RAHEL SOLLMANN,1,6 BETH GARDNER,1 ARIELLE W. PARSONS,2 JESSICA J. STOCKING,2 BRETT T. McCLINTOCK,3 THEODORE R. SIMONS,2 KENNETH H. POLLOCK,4 AND ALLAN F. O’CONNELL5

1North Carolina State University, Department of Forestry and Environmental Resources, Fisheries and Wildlife Program, 110 Brooks Avenue, Raleigh, North Carolina 27607 USA
2USGS, NC Cooperative Fish and Wildlife Research Unit, Department of Biology, North Carolina State University, Raleigh, North Carolina 27695 USA
3National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, Washington 98115 USA
4North Carolina State University, Department of Biology, 234 David Clark Labs, Raleigh, North Carolina 27695-7617 USA
5USGS Patuxent Wildlife Research Center, 10300 Baltimore Avenue, Beltsville, Maryland 20708 USA

Abstract. Abundance and population density are fundamental pieces of information for population ecology and species conservation, but they are difficult to estimate for rare and elusive species. Mark–resight models are popular for estimating population abundance because they are less invasive and expensive than traditional mark–recapture. However, density estimation using mark–resight is difficult because the area sampled must be explicitly defined, historically using ad hoc approaches. We developed a spatial mark–resight model for estimating population density that combines spatial resighting data and telemetry data. Incorporating telemetry data allows us to inform model parameters related to movement and individual location. Our model also allows 100% individual identification of marked individuals. We implemented the model in a Bayesian framework, using a custom-made Metropolis-within-Gibbs Markov chain Monte Carlo algorithm. As an example, we applied this model to a mark–resight study of raccoons (Procyon lotor) on South Core Banks, a barrier island in Cape Lookout National Seashore, North Carolina, USA. We estimated a population of 186.71 ± 14.81 individuals, which translated to a density of 8.29 ± 0.66 individuals/km2 (mean ± SD). The model presented here will have widespread utility in future applications, especially for species that are not naturally marked.

Key words: abundance; Bayesian statistics; camera trapping; density; mark–resight; Metropolis-within-Gibbs sampler; population estimation; Procyon lotor; raccoons; spatial capture–recapture; telemetry.

INTRODUCTION

Abundance and density are crucial pieces of information for both population ecology and species conservation. Capture–recapture models (e.g., Otis et al. 1978) are considered one of the most robust ways for estimating these population parameters, but these models require that all individuals sampled are identifiable. Individual identification using natural markings limits the number of species that can be sampled using noninvasive capture–recapture techniques, such as camera trapping. On the other hand, repeated physical captures are expensive, time consuming, and often risky for animals due to increased handling.

Mark–resight models (Arnason et al. 1991, Neal et al. 1993, White and Sheng 2001) present a less expensive and less invasive alternative, which is of particular interest when working with rare or endangered species. For mark–resight studies, a sample of individuals is captured and tagged (or otherwise marked) during a single marking event that takes place prior to the resighting surveys. The marked individuals constitute the individually identifiable portion of the study population; they do not have to be captured and handled again. Instead, a noninvasive technique such as camera-trapping or visual resighting can be used to collect “recapture” data on these individuals. In addition, sightings of unmarked individuals are recorded during resighting surveys, and both data types are used in the mark–resight model.

Mark–resight models have been applied for decades to a range of wildlife survey techniques. Recently, a suite
of generalized likelihood-based mark–resight models has been developed. These models allow individual heterogeneity in detection, an unknown number of marked individuals, a lack of geographical closure, sampling with or without replacement, <100% individual identification of marked animals, and the combination of data from several surveys in a robust design framework (McClintock et al. 2009a, b, McClintock and White 2009). Although these mark–resight models are very flexible and provide robust estimates of abundance, they suffer from some of the same shortcomings as traditional capture–recapture models when it comes to estimating population density. Space has no explicit manifestation in these models. In order to estimate density, we need to define the area we sampled, and this generally relies on ad hoc approaches (e.g., Wilson and Anderson 1985, Parmenter et al. 2003). Further, heterogeneity in capture probability induced by different exposure of individuals to the survey locations cannot be modeled explicitly. Instead, it must be explained by individual covariates (if available) or approximated by classical models of latent heterogeneity such as the beta-binomial mixture model (Burnham 1972), the logit-normal mixture (Coul and Agresti 1999), or finite mixtures (Norris and Pollock 1996).

The development of spatial capture–recapture (SCR) models (Efford 2004, Royle and Young 2008, Gardner et al. 2009) has addressed both of these issues by incorporating individual location and movement relative to a trap array explicitly into the model, and by providing abundance estimates that are linked to a predefined area using a point process model. These models, however, are still fundamentally capture–recapture models, and thus require individual identification of all animals that are encountered.

Chandler and Royle (2013) generalize SCR models to sampling situations where not all individuals can be uniquely identified. When at least some of the individuals can be distinguished, this corresponds to a spatial mark–resight (SMR) situation. In SMR models, accuracy and precision of parameter estimates are directly related to the number of marked and thus identifiable individuals in the study (Chandler and Royle, 2013). Especially for species that are difficult to trap and mark, the marked portion of the population will usually be small, and marking will most likely involve radio or GPS tagging. Mark–resight studies traditionally relied on radio-telemetry to confirm which marked individuals were alive and available for resighting. This was because standard mark–resight models required knowing the exact number of marked individuals in the population (e.g., White and Shenk 2001, but see Arnason et al. 1991, McClintock et al. 2009a). Open-population mark–recapture and tag–return models have been augmented with telemetry data to obtain more precise estimates of survival and to separate mortality from emigration (e.g., Powell et al. 2000, Pollock et al. 2004). But in the context of SMR models, data obtained via telemetry can also provide additional information on individual location and movement that has not been previously utilized in mark–resight (or mark–recapture) models. Parameters related to individual location and movement are normally estimated from resighting data, and sparse spatial resighting information can render parameters inestimable (e.g., Marques et al. 2011, Sollmann et al. 2012). Sparse resighting data can arise from both a small number of marked individuals and from spacing resighting locations too widely relative to animal movement, so that individuals are unlikely to be recorded at more than one resighting location. Therefore, our motivation was to extend the SMR model by Chandler and Royle (2013) by utilizing additional telemetry data that is collected simultaneously to the resighting data with the objective to obtain useful population estimates even when sample sizes are relatively small or when a lack of knowledge of animal movement prevents adequate spatial study design. Further, we address the issue that in many sampling situations individual identification rate of marked individuals may be below 100%. We demonstrate the model on a study of raccoons (Procyon lotor) on South Core Banks, a barrier island in North Carolina, USA. This application highlights the power of combining multiple surveying and modeling techniques into a unified framework for estimating density. Improved protocols to estimate population parameters are particularly important for the conservation and management of those species that are rare, elusive, and/or difficult to sample.

Spatial Mark–Resight Model

SCR data are generally collected using an array of detectors or traps. SCR models assume that each individual i has an activity center, s, and these s are distributed following a homogeneous Poisson point process in the state-space S, an arbitrarily large area that includes the trap array. When individuals can be detected repeatedly at each trap during each of several trapping periods, we assume the number of detections of individual i at trap j and trapping period k, yijk, is a Poisson random variable with mean encounter rate \( \lambda_{ijk} \), which is assumed to be a decreasing function of the distance \( d_{ij} \) from trap j to the individual’s activity center s. Under a half-normal function,

\[
\lambda_{ijk} = \lambda_0 \times \exp \left( -\frac{d_{ij}^2}{2\sigma^2} \right)
\]

where \( \lambda_0 \) is the baseline trap encounter rate, or the expected number of times an individual i would encounter a (hypothetical) trap located precisely at \( s_i \); \( \sigma \) is the scale parameter of the half-normal function and can be related to animal movement (Reppucci et al. 2011). When only part of the population can be individually identified, as is the case in mark–resight studies, the individual encounter histories \( y_{ijk} \) are
partially latent: Only \( y_{ijk} \) for the \( m \) individually identifiable animals in our study population are fully observed. For the unmarked individuals, we only observe the accumulated counts \( n_{jk} \); unobserved encounter histories are essentially missing data. Bayesian analysis is a natural paradigm for inference in models with latent variables and missing data, because general algorithms exist for simulating from the posterior conditional on the latent variables and missing data. Using Metropolis-within-Gibbs (MwG) Markov chain Monte Carlo (MCMC) sampling (Gelman et al. 2004:283–310), we can update those \( y_{ijk} \) for animals that cannot be individually identified using their full conditional distribution. Under a sampling scheme of individuals being resighted across an array of \( J \) locations and with a Poisson encounter process, this full conditional is multinomial with sample size \( n_{jk} \) as follows:

\[
y_{ijk} \sim \text{Multinomial} \left( n_{jk}, \frac{\lambda_{ijk}}{\sum_{u=m+1}^{M} \lambda_{ijk}} \right)
\]

where \( u = \{m + 1, \ldots, N\} \) is an index vector of the \( N - m = U \) unmarked individuals. Model parameters are then updated conditional on the full set of encounter histories.

Here, we assume that \( m \) is known, which is reasonable under certain conditions, for example, when telemetry devices can be used to confirm the availability of tagged individuals for resighting (McClintock and White 2007). But \( N \), the total population size, remains an unknown quantity. To estimate \( N \), we employ data augmentation (Tanner and Wong 1987, Royle et al. 2007, Royle and Dorazio 2012). Formally, data augmentation is equivalent to assuming a Discrete-Uniform(0, \( M \)) prior for the parameter \( N \), for \( M \) sufficiently large, which is a natural noninformative prior for \( N \). In standard SCR models this formulation is equivalent to augmenting the data set with “all-zero” encounter histories, i.e., \( y_{ijk} = 0 \) for all \( j \) and \( k \), where \( i = m + 1, \ldots, M \) correspond to hypothetical individuals that were never observed. \( N \) is then estimated as the sum of an individual auxiliary variable:

\[
z_i \sim \text{Bernoulli}(\Psi)
\]

where \( i = 1, 2, 3, \ldots, M \) and \( z_i = 1 \) if the animal is part of the population and 0 otherwise. Introduction of \( z \) changes the observation model to

\[
y_{ijk} | z_i \sim \text{Poisson}(\lambda_{ijk} \times z_i).
\]

The prior probability of \( \Psi \) is Uniform (0, 1), which is equivalent to assuming a Discrete-Uniform prior over \( \{0, M\} \) on \( N \). In spatial mark–resight models where \( m \) is known, we can actually observe the “all-zero” encounter histories of marked animals. Thus, estimating \( N \) reduces to estimating the number of unmarked individuals \( U \), both photographed and never observed. In this situation, \( z_i = 1 \) for marked individuals, \( i = 1, \ldots, m \) (whether resighted or not), and \( M - m = \) size of hypothetical unmarked population of \( S \). By updating the latent encounter histories we assign records of unmarked individuals to some of these hypothetical individuals, so that their encounter histories are no longer “all-zero.” Density \( D \) can be derived by dividing \( N \) by the area of \( S \) and is invariant to the size of \( S \) as long as \( S \) is chosen large enough to include all individuals that could have been exposed to the resighting array.

In addition to the parametric model assumptions, the SMR model assumes: demographic closure of the population, no loss of marks, correct identification of animals as marked or unmarked, and no misidentification when individual identity can be established. Further, as in all mark–resight studies, marked individuals are assumed to be a representative sample from the overall population.

**Incorporating telemetry data**

By using a half-normal detection function, we can relate the parameters \( \sigma \) and \( s \) directly to those from a bivariate normal movement model, with mean = \( s_i \), and variance-covariance matrix \( \Sigma \), where the variance in both dimensions is \( \sigma^2 \), and covariance is 0. Ordinarily, these parameters are estimated directly from the spatial distribution of individual recaptures/resightings. Telemetry data, however, provide more detailed information on individual location and movement, since the resolution and extent of the data are not limited by the trapping grid and potentially more locations can be accumulated through telemetry than resighting. By assuming that the \( R_i \) locations of individual \( i \), \( l_i \), are a bivariate normal random variable

\[
l_i \sim \text{Normal}_2(s_i, \Sigma).
\]

We can estimate \( \sigma \) as well as \( s_i \) for the collared individuals directly from telemetry location data using their full conditional distributions within the MwG sampler as follows:

\[
[\sigma | l, s] \propto \left[ \prod_{i=1}^{m} \prod_{r=1}^{R} \frac{1}{2\pi\sigma^2} \exp \left( -\frac{1}{2} \left( \frac{(lx - sx_i)^2}{\sigma^2} + \frac{(ly - sy_i)^2}{\sigma^2} \right) \right) \right] \times [\sigma]
\]

\[
[s_i | l, \sigma] \propto \left[ \prod_{r=1}^{R} \frac{1}{2\pi\sigma^2} \exp \left( -\frac{1}{2} \left( \frac{(lx - sx_i)^2}{\sigma^2} + \frac{(ly - sy_i)^2}{\sigma^2} \right) \right) \right] \times [s_i].
\]

Here \( [\sigma] \) and \( [s_i] \) denote the density functions of the prior distributions for these parameters, which we chose to be uniform(0, \( +\infty \)) (\( \sigma \)) and uniform over \( S \) (\( s \)). Under the standard mark–resight assumption that marked individ-
uals are a representative sample of the population, the estimate of $r$ can be applied to the entire population. For the unmarked individuals, $s_i$ is estimated conditional on the latent encounter histories (see Appendix for a complete description of the full conditional distributions and the algorithms used to sample from these distributions).

**Incomplete individual identification**

Often during resighting, it may be possible to see that an individual is tagged, but impossible to determine its individual identity. In this case, the individual encounter histories of marked animals are essentially incomplete, and if we used these incomplete data to inform the detection parameter of the model, we would underestimate detection/trap encounter rate and overestimate abundance. McClintock et al. (2009a, b) suggest an intuitive means of correcting for this bias in a nonspatial model framework when dealing with a Poisson encounter model by adding the number of unidentified records per marked individual to the average encounter rate of marked individuals in the model likelihood. This procedure assumes that the inability to identify a marked individual occurs at random throughout the population. We can translate this concept to SMR models, where we are interested in the individual and trap specific encounter rate, $\lambda_{ijk}$. Assuming that the sum of all correctly identified records of marked individuals, $\Sigma y_{ik}$, is a Binomial random variable with the number of trials being the total number of records of marked individuals, $\Sigma y_{im}$, we can estimate the probability of identifying the record of a marked individual, say, $c$. For the marked individuals we can then multiply $\lambda_{ik}$ with $c$ to account for the fact that we observe incomplete individual encounter histories. Since this identification issue does not apply to unmarked individuals, their baseline trap encounter rate remains simply $\lambda_{0i}$. In addition to assuming that failure to identify tagged individuals occurs at random throughout the population, this approach also assumes that it occurs at random throughout space.

**APPLICATION**

We applied the spatial mark–resight model supplemented with telemetry data to a camera-trap and radio-telemetry data set from the raccoon population on South Core Banks, a barrier island within Cape Lookout National Seashore, North Carolina. In many coastal areas high raccoon densities correlate with high levels of nest failure in shorebirds and sea turtles (e.g., Stancyk et al. 1980, Ratnaswamy et al. 1997, Erwin et al. 2001, Schulte 2012). Among the objectives of this study was the collection of baseline population and demographic data of raccoons. Between May and September 2007, 131 individuals were marked with large individually numbered cattle tags attached to dog collars; 44 of these tagged individuals were equipped with radio collars (ATS, Asanti, Minnesota, USA). Collared individuals were located using a VHF receiver and antenna, and their locations were estimated approximately weekly. Twenty infrared 35-mm camera traps (TrailMaster; Goodson and Associates, Lenexa, Kansas, USA) were set up in May 2007 along the length of South Core Banks, spaced ~1.77 km apart (Fig. 1). Each camera trap was armed from 21:00 hours to 05:00 hours every day from 18 May to 21 October 2007. From 17 November 2007 until 25 January 2008 cameras were only armed for one week per month. Since tagging occurred until 31 September, we only considered camera-trapping data from 1 October 2007 to 22 January 2008 in our analysis, which is also likely to approximate a demographically closed population. During this period, 104 marked individuals, 38 radio-collared, were alive and available for resighting with camera traps.

We defined the state-space $S$ as the entire area of South Core Banks (Fig. 1). A change in the number of photo-captures over the course of the study suggested a variation of detection rate with time. Since date recording in some cameras malfunctioned, photographic records could only be assigned to the time interval between subsequent trap checks, and we refer to these intervals between checks as sampling periods. These periods ranged from 2 to 43 days, and we accounted for
the differences in sampling effort by standardizing $\lambda_0$ to 7-d intervals (e.g., a 2-d period corresponds to two-sevenths of a standardized period). We allowed the standardized $\lambda_0$ and $c$ to change with sampling period.

We implemented the model using a custom-made MwG algorithm (see Supplement 1 for R code) in the software R 2.13.0 (R Development Core Team 2011). We ran three MCMC chains of the algorithm with 200,000 iterations each, discarding the first 10,000 iterations as burn-in. To check for chain convergence, we calculated the Gelman-Rubin statistic R-hat (Gelman et al. 2004) using the R package coda (Plummer et al. 2006). Values below 1.1 indicate convergence; for all parameters in our model, R-hat was $<1.1$.

For comparison we also implemented the spatial mark–resight model without telemetry data and the model described by Chandler and Royle (2013) without telemetry data or individual identity (i.e., treating all individual encounter histories as latent), using the same MCMC settings.

Camera-traps recorded 117 pictures of unmarked raccoons, 33 pictures of 18 marked and identifiable raccoons, and 49 pictures of marked, but not individually identifiable animals. Eighty-six of the 104 marked raccoons, and 49 pictures of marked, but not individually identifiable animals. Eighty-six of the 104 marked individuals were never photographed. The number of photographs of the remaining marked individuals ranged from one to seven. An average of 16.32 photographs of the remaining marked individuals.

In our raccoon case study, although a large number of individuals were tagged, photographic data of tagged raccoons was surprisingly sparse, with only 33 identifiable pictures over the entire four-month course of the study. Analysis of the photographic data set without the telemetry data or without individual identity did not yield usable estimates of abundance, since models did not converge. Chandler and Royle (2013) report that SCR models without individual identity do not perform well when trap spacing is wide relative to individual movement, because under these circumstances point counts approximate independent and identically distributed (iid) Poisson counts and the spatial autocorrelation structure needed to identify model parameters is lost. Similarly, Sollmann et al. (2012) report that in regular SCR models (i.e., with full individual identification), wide trap spacing relative to individual movement resulted in few spatially spread-out recaptures and rendered model parameters non-estimable. In this study, camera traps were spaced on average 1.77 km apart, based on prior knowledge of raccoon movements. Nevertheless, this spacing turned out to be about 3.5 times our estimate of $\sigma$. Consequently, we observed very few individual marked raccoons captured at more than one trap. Under these circumstances, the telemetry data provide the necessary spatial information to estimate the movement and activity centers of individual animals and make the model identifiable. Thus, telemetry data can provide an “insurance” against sparse resighting data, not only when there is no prior information to base trap spacing on, but also when animals behave in an unexpected manner.
Telemetry tags (or any other kind of mark for that matter) are designed to be as small and unobtrusive as possible so that they do not impair the natural behavior and movement of animals. Therefore, visual identification may not always be possible, and often researchers may only be able to ascertain that an individual is tagged. When individual identification of marked animals is impossible, nonspatial mark–resight models can be applied using the observed counts of marked and unmarked individuals (McClintock and White 2010), but these models cannot account for heterogeneity in detection among individual animals. In a spatial framework, we could ignore marked and unmarked status completely and apply the model by Chandler and Royle (2013) without individual identity, but parameter estimates, when estimable, are inherently less precise when no individual information is available. Therefore, being able to retain the individual identity of records that can be identified while accounting for an identification rate <100% is extremely useful. In nonspatial mark–resight models, moderate to high levels of individual heterogeneity in detection in combination with moderate proportions of unidentified marks (10–20%) can lead to poor confidence interval coverage of N (White and Shenk 2001, Magle et al. 2007). The effects of different rates of individual identifiability on estimates of SMR models remain to be explored, but researchers should strive for high rates of individual identification.

The assumption that failure to identify marked individuals occurs at random throughout the population should hold as long as all individuals are identified from the same type of tags. It is, however, possible that identifying marked individuals could be nonrandom with respect to space, for example when using different types of camera traps, observers with different resighting skills, or when spatially varying habitat conditions influence tag identification probabilities. These potential biases can be explicitly stated in the model; for example, different types of camera traps or observers could be allowed to have different identification rates, or the rates could be modeled as a function of habitat covariates. Alternatively, model development could explore assigning records of marked but unidentified individuals to hypothetical individuals in a fashion similar to how unmarked records are assigned to hypothetical individuals in this model.

Spatial or individual sources of variation in λ₀ and σ can easily be incorporated in a fashion analogous to regular SCR models (for SCR examples, see Gardner et al. 2009, Sollmann et al. 2011). Conceptually, SMR models can also easily be extended to temporally or spatially replicated surveys to borrow information across time or space. This is especially advantageous when data are sparse, as is often the case in surveys of rare or cryptic species.

In summary, the spatial mark–resight model framework is flexible enough to account for a variety of factors that may influence individual movement and detection, as well as sampling-related parameters such as the identification rate of tagged individuals. As such, the approach is applicable to a wide range of population estimation problems, and it is readily incorporated into existing telemetry studies to yield robust density estimates. The model enables researchers to make optimal use of all existing data and estimate population parameters without the need for additional invasive methods. This feature is especially useful for difficult-to-study, cryptic species where often only a small sample of the population is marked.

Acknowledgments
Funding support was provided by the Natural Resource Preservation Program of the U.S. Geological Survey and National Park Service. We are grateful to the resource management staff at Cape Lookout NS for their logistical support. We thank J. Andrew Royle, Richard B. Chandler, Brian Reich and two anonymous reviewers for comments that greatly improved the manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not constitute endorsement by the U.S. government.

Literature Cited
Burnham, K. P. 1972. Estimation of population size in multiple capture-recapture studies when capture probabilities vary among animals. Dissertation, Oregon State University, Corvallis, Oregon, USA.

SUPPLEMENTAL MATERIAL

Appendix

Detailed description of the Metropolis-within-Gibbs MCMC sampler for a spatial mark–resight model augmented with telemetry location data and with individual identification rate of marks <100% (Ecological Archives E094-046-A1).

Supplement

R code and data to run spatial mark–resight model for raccoon camera trap and telemetry data (Ecological Archives E094-046-S1).