Marc Kéry
J. Andrew Royle

APPLIED HIERARCHICAL MODELING IN ECOLOGY

Analysis of distribution, abundance and species richness in R and BUGS

Volume 1
Prelude and Static Models

Applied Hierarchical Modeling in Ecology

Analysis of distribution, abundance and species richness in R and BUGS

Applied Hierarchical Modeling in Ecology

Analysis of distribution, abundance and species richness in R and BUGS

Volume 1

Prelude and Static Models

Marc Kéry
Swiss Ornithological Institute,
Sempach, Switzerland

J. Andrew Royle
USGS Patuxent Wildlife Research Center,
Laurel MD, USA

ELSEVIER
AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2016 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-12-801378-6

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalogue record for this book is available from the Library of Congress

For information on all Academic Press publications visit our website at http://store.elsevier.com/

ELSEVIER
Book Aid International
Working together
to grow libraries in
developing countries
www.elsevier.com • www.bookaid.org

J. A. Royle was the principal author of chapters 2, 7, 8, 9. Use of product names does not constitute endorsement by the U.S. Government.

For Jim Nichols, who changed the way in which we think about Ecology

Contents

PART 1 PRELUDE

PART 2 MODELS FOR STATIC SYSTEMS

Additional Online Content: http://www.mbr-pwrc.usgs.gov/pubanalysis/keryroylebook/

CHAPTER

MODELING ABUNDANCE USING HIERARCHICAL DISTANCE SAMPLING

8

CHAPTER OUTLINE

Applied Hierarchical Modeling in Ecology. http://dx.doi.org/10.1016/B978-0-12-801378-6.00008-4

8.1 INTRODUCTION

Distance sampling (DS) is one of the most widely used statistical methods in ecology for estimating density or population size (Burnham et al., 1980; Buckland et al., 2001, 2004a; Williams et al., 2002). There is an enormous body of literature, which the two classics by Buckland et al. summarize succinctly. Clearly, our aim here is not to try and present an exhaustive overview of distance sampling but rather to summarize the salient features of this important methodology, in order to show how it fits into the larger picture of hierarchical modeling of spatially indexed abundance data.

Conventional distance sampling (CDS; Buckland et al., 2001) uses information on observed distances of animals from transect lines or observation points to characterize the detection probability of individuals. Under the eminently plausible hypothesis that detection probability is related to the distance between animals and the observer, one may obtain an estimate of absolute density. Often sampling is done from boats or planes in open environments and thus animals that are amenable to such sampling are widely studied using distance sampling methods, including ungulates, whales, and other marine mammals. More recently, CDS methods have become very popular in the sampling of birds using point counts (Buckland et al., 2001; Rosenstock et al., 2002). In point count surveys, distances are recorded from a point of observation (instead of along a transect), and this is usually referred to as point transect sampling. In this chapter, we use the terms ‘point count’ and ‘point transect’ synonymously. Distance sampling methods are attractive because they do not require that individuals be uniquely marked and recaptured (or resighted) through time. Furthermore, unlike most capture-recapture models (but not spatial capture-recapture models), distance sampling requires only a single sample of the population, making it “cheaper” in terms of logistics (because a single visit to a site is enough) and modeling (because no closure assumption is needed). Finally, distance sampling is one of the only methods, along with spatial capture-recapture (Royle et al., 2014), which accommodates the basic problem of unknown sample area (see also Section 6.10).

The main assumptions of the CDS method are: (1) animals are distributed uniformly in space, (2) detection probability is a function of distance and is equal to 1 at distance 0, (3) individuals are detected at their original location, i.e., there is no responsive movement, and (4) distances are measured without error. Usually assumption (1) is not stated explicitly but, instead, it is assumed that sample points or transects are distributed randomly (Buckland et al., 2001, p. 29). As far as the mechanics of distance sampling are concerned, these two are effectively equivalent assumptions about the system, the former being a model-based version of random sampling of individuals, the latter being more of a design-based argument.

Conventional distance sampling has been a popular sampling method for many decades. However, historically little attention has been paid to *modeling spatial variation* in abundance using distance sampling methods. While it has been standard practice to obtain distance sampling data at multiple sample units using essentially the type of meta-population design we have encountered in Chapters 1, 6 and 7, the data from such replicate samples have typically been pooled in order to estimate parameters of the detection probability model. Thus, information about, or explicit attention to, factors that influence abundance among sites has been neglected. This is unfortunate because spatial or spatiotemporal patterns in abundance are often the primary interest of ecological studies!

This is not to say that information from replicate sample units is not used at all in CDS—indeed, the spatial replicate sample units are used to estimate the *encounter rate variance*, i.e., the variance in n_s (number of encountered individuals) among replicate units $s = 1, 2, \ldots, S$. This provides a sort of nonparametric variance estimator while, on the other hand, in a formal hierarchical modeling framework, we use parametric models to describe variation in abundance among sample units. Instead of pooling data as in CDS, we think it makes sense, following the basic ideas of Chapters 6 and 7, to provide an explicit model for the variation of local population size N_s, the population size for spatial sampling unit s. We call this hierarchical distance sampling (HDS). By specifying a model for this latent variable, we can then build explicit models for distance sampling data that account for variation in population size (or local density) among sample units, thus facilitating inference about factors that influence spatial variation in abundance or the making of explicit spatial predictions of abundance.

While conventional distance sampling is very mature and established in ecology and wildlife science, HDS has only existed for a few years. There are two key conceptual papers that develop ideas of HDS. Hedley and Buckland (2004) adopt a two-stage estimation procedure where they use the usual distance sampling model for observed distances to estimate detection probability, pooling the data among sample units, and then in a second-stage procedure they fit a model (e.g., a Poisson GLM with an offset being a function of the probability of detection) to the observed count of individuals n_s for each of $s = 1, 2, \ldots, S$ sample units. Miller et al. (2013a) call this methodology *density surface modeling* and describe an R package named `dsm` and the implementation of this in the popular program Distance (Thomas et al., 2010). Royle et al. (2004) develop HDS as a formal hierarchical model in which the two components (detection and abundance) are simultaneously estimated in a single hierarchical model exactly analogous to the binomial or multinomial N-mixture model framework of Chapters 6 and 7. The HDS methodology is implemented in `unmarked` using the `distsamp` and `gdistsamp` functions and is the topic of this chapter and the next. We note that in this chapter we will temporarily deviate from our usual notation and for now index sites by s for $s = 1, \ldots, S$. The reason is that we also need to index individuals, for which we will use index i, with $i = 1, \ldots, M$, where M will usually be the number of individuals in an augmented data set; see below.

Since 2004, only a trickle of papers have appeared that develop HDS ideas or provide novel implementations. Chelgren et al. (2011b) and Moore and Barlow (2011) appear to be the first to do a Bayesian analysis of an HDS model in BUGS. Chelgren et al. (2011b) contains quite a few novel elements. They formulate an HDS model in continuous space (using the "ones trick" in BUGS) and provide a three-part hierarchical model with a binomial observation model for $n_s|N_s$, a Poisson model for N_s, and then the ordinary distance sampling model for the observed distances conditional on n_s. They also accommodate within-unit heterogeneity in density by zeroing out "nonhabitat" in the sampled region. Moore and Barlow (2011) have a temporal dimension and embed an exponential population model into their distance sampling observation model and also model group size (see also Pardo et al., 2015).

Shirk et al. (2014) adopt the Chelgren et al. formulation of the model and provide a nice application to sampling chameleons. Oedekoven et al. (2013) also use a variation of the three-part formulation of the HDS model but remove N from the model by summation to reduce this to a two-part model. Schmidt et al. (2012) and Schmidt and Rattenbury (2013) fit HDS models with variation in group size in BUGS using data augmentation (DA). They may be the first to use "S-fold data augmentation," i.e., doing DA for each (transect) population and then linking the different transects by modeling the data augmentation parameter ψ (see also Tenan et al., 2014b). Sillett et al. (2012) develop an application of likelihood-based HDS with covariates on detectability (e.g., the parameter σ of some detection function) and $E(N)$, and considered Poisson and negative binomial abundance models. This is the first paper using `unmarked`'s `gdistsamp` function. Chelgren et al. (2011b) and Shirk et al. (2014) included effects on σ in the context of Bayesian HDS models. Conn et al. (2012) develop a combined double-observer HDS model with group structure and a CAR formulation of spatial correlation, using a Bayesian analysis conducted with a custom MCMC algorithm implemented in R package `hierarchicalDS`. Amundson et al. (2014) develop an HDS model with time of removal and individual level effects (see Chapter 9). Finally, Niemi and Fernandez (2010) develop a spatial point process model for line transect data and Johnson et al. (2010), including their R package `Dspat`, develop a similar model (see also ver Hoef et al., 2014, and Pardo et al., 2015).

In this chapter we begin with a fairly detailed introduction to basic ideas of distance sampling models, absent the hierarchical structure of having multiple spatial sample units, i.e., conventional distance sampling, as it is covered in the classic textbook by Buckland et al. (2001) and implemented in the widely used Distance software (Thomas et al., 2010). We do this so that we can introduce the reader to the mechanics of formulating the distance sampling model, simulating data, and fitting the model for the two standard sampling contexts: (1) transects and (2) point counts ("point transects"). For both cases we consider both continuous and "binned" distance measurements. These are statistically equivalent models as the number of bins gets large, but the practical issues of their analysis, and especially their implementation in BUGS, are very different. Hence, it's useful to see and experiment with both formulations. In addition, there are technical distinctions having to do with whether we adopt a conditional or full likelihood formulation of the model, and also whether we analyze the data by classical likelihood or Bayesian analysis. We first cover all of these various manifestations of the conventional (nonhierarchical) distance sampling (CDS) model.

Only once these basic principles have been developed do we extend the ideas to hierarchical distance sampling, where we use hierarchical models to combine the data from sampling at multiple locations formally into a single joint model. The distance sampling protocol, combined with a model for abundance, produces what we'll call the hierarchical distance sampling (HDS) model. When continuous distance sampling measurements are binned into distance classes, a multinomial observation model is produced. Therefore, the multinomial mixture models of the previous chapter can be applied with only some minor technical modifications that we have to consider when computing the multinomial cell probabilities.

8.2 CONVENTIONAL DISTANCE SAMPLING

We first develop the basic concepts and technical details of "classic" distance sampling without thinking about spatial replication and hierarchical models. A simple way to motivate distance sampling is to think about our heuristic estimator of N derived by solving the relationship

$$E(n) = \overline{p}N.$$

Therefore, we can estimate N from a sample count n and an estimate of $\overline{p}$, the probability that an object (i.e., animal) appears in our sample of size n. The idea of distance sampling is to estimate $\overline{p}$ by modeling detection probability of objects as a function of distance x from the object to an observer recording data at a point or walking along a transect. This is done by specification of some function, the "detection (probability) function," $g(x; \theta)$, describing detection probability as a function of distance x and parameter(s) θ. That is, the detection function is a model for the probability of detection of an object conditional on its distance from the observer x, i.e., $g(x; \theta) = Pr(y = 1|x)$ in our usual notation of conditional probabilities, where y is a Bernoulli trial indicating detection ($y = 1$) or nondetection ($y = 0$). The traditional notation can be a little confusing because if we just write $g(x; \theta)$ then it looks like this could be a probability distribution for x, which it is not—rather it is the parameter of a Bernoulli probability mass function for a variable y, whether or not an object is detected conditional on x. This is why we will write this as $\Pr(y = 1|x)$ when we want to be clear that it is a probability of an event, that is, of being detected.

How is $\overline{p}$ related to this detection function $g(x;\ \theta)$? It is the marginal or *average* detection probability (therefore we write $\overline{p}$ instead of $\widehat{p}$), which is the probability that an individual in the population at large appears in the sample, and it is computed by averaging $g(x;\ \theta)$ over all possible values of x. Formally, the calculation is

$$\overline{p} \equiv \Pr(y = 1) = \int_x g(x;\theta)[x]dx \tag{8.1}$$

Note that the averaging is being done with respect to a probability density for x, denoted here by $[x]$ (using our established bracket notation), although we have yet to specify this quantity. Thus, the basic distance sampling model has two explicit and essential components:

1. The "observation model," which describes how individuals appear in the sample, characterized by the function $g(x;\ \theta)$.
2. The "process model," $[x]$, which describes how objects in the population are distributed with respect to the observer or the transect.

Conventional distance sampling adopts an explicit and intensive focus on inference about component (1), typically considering many and fairly complex models for the detection function and choosing among those by AIC (e.g., Miller and Thomas, 2015). Historically, very little attention has been focused on modeling the "process," i.e., the probability distribution $[x]$. Conversely, HDS adopts an explicit focus on modeling $[x]$ as we will see later in this chapter. It may seem like we're making a big deal about this because how can "distance from observer" be any kind of meaningful ecological process? Well, in and of itself it is not, but, in specific cases, the distribution of $[x]$ is precisely equivalent to the distribution of objects in space, and models for such things are usually called *point process models* (Illian et al., 2008; Wiegand and Moloney, 2014). In a sense then, HDS is all about merging an "observation model" that describes the detection of individuals conditional on where they are located during sampling with a "process model" that describes where the individuals are located.

Before elaborating on that concept in more detail we discuss how to obtain $\overline{p}$ from a sample of distance data obtained by surveying a transect and recording distances to each of n objects that are detected. Conceptually, we think it is extremely helpful to think about distance sampling as a logistic regression capture-recapture problem by introducing a population of N individuals each characterized by a pair of random variables $(y_i,\ x_i)$ where y_i is a binary indicator of whether we captured (or observed) that individual with $y_i = 1$ meaning "captured" and $y_i = 0$ "not captured," and x_i is the distance from the observer to the individual at the instantaneous time of sampling. Given the population of N (y, x) pairs, we only observe (y, x) for those n individuals having $y = 1$. So our "data" for a distance sampling study consist of the sample of distances $x_1,\ldots,x_n$ that appear in our sample conditional on the event that $y = 1$ (i.e., that the individual was detected).

To obtain the likelihood for a sample of distances we need to identify the probability distribution of the *observed* distances x, which is to say the probability distribution of x conditional on the event $y = 1$. This can be calculated from a simple application of Bayes' rule. The observed data are the values of x for which $y = 1$, and thus we seek to identify the probability distribution $[x|y = 1]$. Bayes' rule tells us that

$$[x|y = 1] = \frac{[y = 1|x][x]}{[y = 1]}$$

where $[y = 1|x]$ is the function that we choose to model detection probability as a function of distance—the "detection function." The other two components require some specific discussion and analysis: (1) $[x]$ is the *population distribution of distances*. Therefore we must specify this probability distribution in order to compute $[x|y = 1]$. (2) Once we specify the probability distribution $[x]$, then we can compute the denominator directly as $[y = 1] = \int_0^B [y = 1|x][x]dx$. This is the average probability of detection over the interval $[0, B]$, where B represents some maximal distance out to which individuals are counted. So, the probability density for distance observation x_i is the following:

$$[x_i|y_i = 1] = \frac{g(x_i;\theta)[x]}{\int_x g(x;\theta)[x]dx} \tag{8.2}$$

for whatever distribution for $[x]$ we choose.

What sorts of distributions make sense for x? To gain some intuition about this we note that the distribution for x is essentially a prior distribution on "distance from observer," and it can be derived equivalently from a prior distribution on the location of individuals in the surveyed region. To make life simple here we first assume that the survey is done along a linear transect of length L so that we can imagine the surveyed region is a long rectangle with a line running down the middle. We will assume that individuals are only counted up to some maximum distance, say B, and so the surveyed region is a rectangle of dimension $L \times 2B$. Lacking specific knowledge to the contrary, it is sensible to assume that individuals are distributed uniformly over the sampled rectangle. Let's define the individual locations by the coordinates $\mathbf{u}_i$ for $i = 1, 2, \ldots, N$. As it turns out, if individuals are uniformly distributed in space, then their *distances* to a transect (but not to a point, see below) also have a uniform distribution on the interval $[0, B]$, i.e., the density $[x] = 1/B$.

One of the important concepts of distance sampling is that the observed distances are biased with respect to the population distribution $[x]$. The conditional density in Eq. (8.2) makes it clear that the density of observed x should be proportional to the detection function. So if we simulate data under a half-normal detection model (with scale parameter σ, see Figure 8.1 left) the distribution of the observed distances, represented by blue histogram shown in Figure 8.1 (right), is clearly not uniform.

Under the assumption that individuals are uniformly distributed in space, so that distances are uniformly distributed on the interval $[0, B]$, $[x]$ cancels from the numerator and denominator of the conditional distribution given above and therefore does not further influence the likelihood contribution of each x_i (for point count data, things are slightly different in the sense that x doesn't cancel from the likelihood; see below). The likelihood for n observed distances is therefore

$$L(\mathbf{x};\theta) = \prod_{i=1}^{n} \frac{g(x_i;\theta)}{\int_x g(x;\theta)dx} \tag{8.3}$$

which we maximize to obtain $\widehat{\theta}$. It is worth pointing out that in order to evaluate the likelihood we have to do a numerical integration of the detection function over the support of x, the interval $[0, B]$. This is a key calculation because the integral in the denominator is also the average probability that an

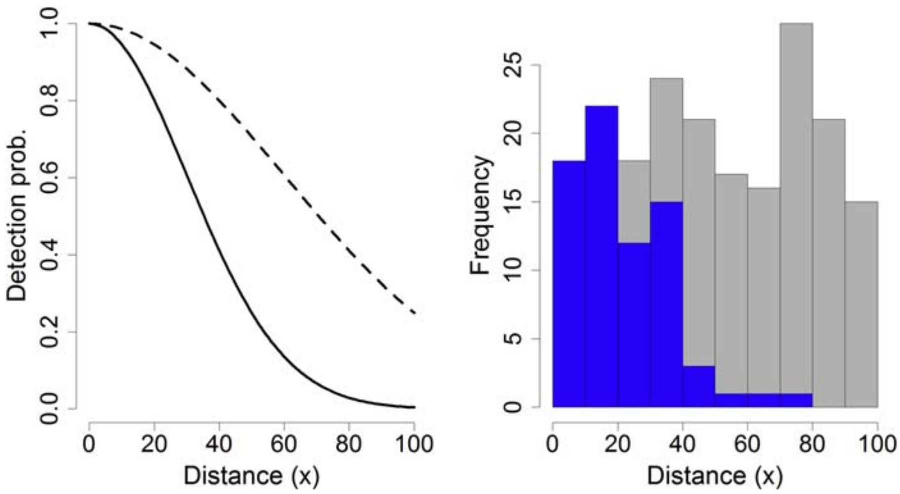

FIGURE 8.1

Half-normal detection function for two different values of σ (left: $\sigma = 30$ (solid) and $\sigma = 60$ (dashed)) and (right) histogram of a sample of true (gray) and observed distances (blue) for $\sigma = 30$; see Section 8.2.3 for R code (function `sim.ldata`).

individual in the population is encountered, i.e., $\overline{p}$ from Eq. (8.1), and thus it is instrumental to "converting" n to $\widehat{N}$. Once we obtain MLEs of the model parameters $\widehat{\theta}$, we evaluate the expression $\overline{p} = [y = 1] = \int_x [y = 1|x][x]dx$, which is the marginal (or average) probability of detection, and we can get an estimate of N directly by $\widehat{N} = n/\widehat{\overline{p}}$. It is customary in distance sampling not to estimate N but, rather, to estimate density D, which is related deterministically to N by dividing by the sampled area $\widehat{D} = \widehat{N}/(2 * L * B) = n/(2 * L * B * \widehat{\overline{p}})$. The denominator here is the *effective sample area* (more commonly $B * \widehat{\overline{p}}$ is called the effective strip half-width for transect sampling, see Buckland et al., 2001, p. 53).

We make four remarks here:

Remark 1: *Density and abundance*. Conventional distance sampling is talked about almost exclusively in terms of estimating density, $D = N/A$, where A is the area over which animals were counted. This area is not usually precisely defined; however, it is implicit in the estimation (i.e., buried under the hood) because formal bounds of integration for a distribution of distance must (usually) be specified, and this effectively implies an area A. We discuss this shortly. As a technical matter, estimation of N or D are statistically equivalent problems.

Remark 2: *Realized versus expected abundance*. In general, whenever an explicit model is placed on the unknown parameter N, this induces a distinction between realized and expected abundance and density (Efford and Fewster, 2013; Dorazio, 2013; Section 5.7.3 in Royle et al., 2014). Expected population size is $E(N)$ where the expectation is with respect to the distribution of N. Expected density is $E(N)/A$. The interpretation of these expected quantities is as the mean value of some hypothetical unit to which our model applies.

Remark 3: *Replicate transects*. Normally we have distances observed from multiple spatial units (transects, points), whereas we have so far only described the situation as if we had only a single spatial unit. Having multiple spatial sample units does not change the fundamental estimation problem nor the mechanics of how we achieve it in CDS. In this case, we simply pool all of the distances into one data set and do the analysis as just outlined. We show this with simulated data below.

Remark 4: *The uniformity assumption*. We stated an explicit assumption that distances from the transect are uniformly distributed. This is induced by the equivalent assumption that animal locations are distributed uniformly in two-dimensional space. However, sometimes formal distance sampling developments do *not* state such assumptions about distances or points or else they make general claims that they are not necessary. Instead, uniformity can be induced *by design* by randomly locating transects (Barry and Welsh, 2001). Our view is that the model assumption and sampling assumption yield equivalent statistical procedures, and so we're not too concerned with how you describe them.

8.2.1 THE FULL LIKELIHOOD

The previous section described estimation of N by first estimating the detection function parameter θ from the likelihood for the observed distances constructed for the n observations, a procedure that is naturally conditional on the event that $y = 1$. As a result, this is usually called the "conditional likelihood," and the estimator of N obtained by "adjusting" n is the *conditional estimator* of population size. However, it is also common in practice to use the so-called *full likelihood* (Borchers et al., 2002; p. 232 in Royle and Dorazio, 2008), which recognizes that n is also a stochastic outcome of the study and should be modeled. The distribution of n is

$$n \sim Binomial(N, \overline{p})$$

and to obtain the full likelihood we simply multiply the conditional likelihood by the binomial component for n. This yields (note: we leave the $[x]$ part in the conditional likelihood for a moment, instead of canceling it from both numerator and denominator):

$$L(\sigma, N) = \left\{ \prod_{i=1}^{n} \frac{g(x_i;\theta)[x]}{\int_x g(x;\theta)[x]dx} \right\} \frac{N!}{n!(N-n)!} \overline{p}^{n} (1-\overline{p})^{N-n}$$

where after some factorizing, canceling, and rearranging, we are left with:

$$L(\sigma, N) = \frac{N!}{n!(N-n)!} \left\{ \prod_{i=1}^{n} g(x_i;\theta)[x] \right\} (1-\overline{p})^{N-n}. \tag{8.4}$$

This resembles the usual full likelihood for every other capture-recapture type of model and, in particular, the individual covariate models (Borchers et al., 2002; Section 7.1 in Royle and Dorazio, 2008). Therefore, we can understand distance sampling as a special type of capture-recapture model where only a single ($J = 1$) sample is taken, and there is an individual covariate, distance x, measured on each observed individual. The full likelihood can be maximized to obtain the MLE of θ and N.

8.2.2 MODELS OF DETECTION PROBABILITY

So far we have just talked about the general concepts and mechanics of distance sampling and how to construct the likelihood of observed distances. However, much of practical distance sampling is focused on the encounter probability model $g(x;\theta)$ and, in particular, choosing among large classes of detection probability models to find models that fit the observed distance distribution (usually in the AIC sense). Here, we don't go into a catalog of types of models but simply mention that some common models include the following:

"half normal"	$g(x;\sigma) = \exp(-x^2/2\sigma^2)$
negative exponential	$g(x;\sigma) = \exp(-x/\sigma)$
hazard rate	$g(x;\sigma,b) = 1 - \exp(-h(x;\sigma,b))$ where $h(x;\sigma,b) = 1 - \exp(-(x/\sigma)^{-b})$

One key feature of these standard models is that they represent monotone decreasing functions of distance with one or two parameters. Another key feature is they all have a known intercept of 1, i.e., $g(0) = 1$, which is a requirement of conventional distance sampling models when we have no other ancillary data. We discuss generalizing this shortly (Alpizar-Jara and Pollock, 1996; Borchers et al., 1998). Why is it required that $g(0) = 1$? Think about the definition of $\bar{p}$:

$$\bar{p} = \Pr(y=1) = \int_x [y=1|x][x]dx$$

If our model for $[y=1|x]$ had some arbitrary intercept, say $[y=1|x] = \alpha * k(x;\sigma)$ where $k()$ was itself some function such that $k(0) = 1$, then the constant intercept α would be confounded with a level shift in density, i.e., a detection model with intercept α and density $[x] = 1/(2 * L * B)$ is equivalent to a detection model with intercept 1 and density $\alpha/(2 * L * B)$. The two are indistinguishable and, in fact, α just cancels from the conditional likelihood expression (Eq. (8.2)). In other words, an intercept in the detection function in CDS is not estimable using standard data for this design.

8.2.3 SIMULATING DISTANCE SAMPLING DATA

We demonstrate some of these basic distance sampling concepts by simulating an imaginary population of the extinct Chihuahuan musk oxen along a transect of length $L = 10$ km. We subject the individual musk oxen to detection by an observer traversing the transect and use a half-normal detection probability function (Figure 8.1). All of this goes according to the following:

```
strip.width <- 100    # one side of the transect, really half-width
sigma <- 30           # Scale parameter of half-normal detection function

# Define half-normal detection function
g <- function(x, sig) exp(-x^2/(2*sig^2)) # Function definition
g(30, sig=sigma)      # Detection probability at a distance of 30m

# Plot the detection function
par(mfrow=c(1,2))
curve(g(x, sig=30), 0, 100, xlab="Distance (x)", ylab="Detection prob.", lwd = 2, frame = F)
curve(g(x, sig=60), 0, 100, add=TRUE, lty = 2, lwd = 2)
```

```
# Define function to simulate non-hierarchical line transect data
sim.ldata <- function(N = 200, sigma = 30){
# Function to simulate line transect data under CDS.
# Function arguments:
#    N: number of individuals along transect with distance u(-100, 100)
#    sigma: scale parameter of half-normal detection function
# Function subjects N individuals to sampling, and then retains the value
# of x=distance only for individuals that are captured
par(mfrow = c(1,2))
# Plot the detection function
curve(exp(-x^2/(2*sigma^2)), 0, 100, xlab="Distance (x)", ylab="Detection prob.", lwd =
2, main = "Detection function", ylim = c(0,1)) # Plot detection function as function of sigma
text(80, 0.9, paste("sigma:", sigma))
xall <- runif(N, -100,100)    # Distances of all N individuals
hist(abs(xall), nclass=10, xlab = "Distance (x)", col = "grey", main = "True (grey) \nand
observed distances (blue)")  # Histogram of distances
g <- function(x, sig) exp(-x^2/(2*sig^2))
p <- g(xall, sig=sigma) # detection probability
y <- rbinom(N, 1, p)  # some inds. are detected and their distance measured
x <- xall[y==1]       # this has direction (right or left side of transect)
x <- abs(x)           # now it doesn't have direction
hist(x, col = "blue", add = TRUE)
return(list(N = N, sigma = sigma, xall = xall, x = x))
}

# Obtain a data set for analysis
set.seed(2015)                # If you want to get same results
tmp <- sim.ldata(sigma = 30) # Execute function and assign results to 'tmp'
attach(tmp)
```

We see that the blue histogram in Figure 8.1 vaguely resembles the half-normal detection probability function (you can increase the resemblance greatly by increasing *N*, e.g., to 10^6). Next, we will obtain the maximum likelihood estimates of the half-normal parameter σ (log-transformed to enforce a positive value) from the simulated data. To do this we define an R function that evaluates the conditional and full likelihoods and use `optim` to minimize the negative log-likelihood in each case.

```
# Conditional likelihood
Lcond <- function(lsigma){ # Define conditional nll
   sigma <- exp(lsigma)
   -1*sum(log(g(x,sig=sigma)/integrate(g, 0, 100, sig=sigma)$value/100))
}

# Call optim to maximize conditional likelihood
optim(log(30), Lcond, hessian=TRUE, method="Brent", lower=-5, upper=10)
$par
[1] 3.257716

$value
[1] 626.8964

[ ... output deleted ... ]
```

```
# Full likelihood
Lfull <- function(parm){   # Define full nll
   sigma <- exp(parm[1])
   n0 <- exp(parm[2])
   N <- length(x)+ n0
   pbar <- integrate(g, 0, 100, sig=sigma)$value/100
   -1*( lgamma(N+1) - lgamma(n0+1) + sum(log(g(x,sig=sigma)/100)) + n0*log(1-pbar) )
}

# Call optim to maximize full likelihood
optim(c(log(30), log(4)), Lfull, hessian=TRUE)
$par
[1] 3.259401 5.012220

$value
[1] 50.31808

[... output deleted ...]
```

In the first case, we get the MLE of log(σ), which we have to convert to the MLE of $\overline{p}$ and then compute $\widehat{N}_c = n/\widehat{\overline{p}}$. In the second case, we get the MLE of N directly by maximizing the full likelihood. Or rather, in this case, we estimate the logarithm of $n_0 = N - n$ and then have to back-transform and add back n to it for an estimate of N. To convert estimates of density we need simply divide the estimates of N by the area of the transect, which was 10 km long and 0.2 km wide (100 m on either side) = 2 km^2. This is all done as follows:

```
pbar <- integrate(g, 0, 100, sig=exp(3.26))$value/100
n <- length(tmp$x)

(Nhat.condl <- n/pbar)
[1] 223.6231
(Dhat.condl <- Nhat.condl/(10*.2))
[1] 111.8115

n0hat <- exp(5.01)
(Nhat.full <- n + n0hat)
[1] 222.9047
(Dhat.full <- Nhat.full/(10*.2))
[1] 111.4524
```

We find that the densities of Chihuahuan musk oxen are quite respectable this year, to say the least (C. Amundson, pers. comm.), being on the order of 111 per km^2. Perhaps we will open a harvest season.

8.2.4 BINNED DATA

It is common in applications of distance sampling to produce observations in distance bands. In that case, the observation model is multinomial and therefore the multinomial N-mixture models, which we developed in Chapter 7, are directly relevant. We develop the likelihood for this case here. Suppose observations are recorded into $h = 1,2,\ldots,H$ distance bands or strips on the intervals

$[b_0,b_1],(b_1,b_2],\ldots,(b_{H-1},b_H]$, where we define $b_0 = 0$ and $b_H = B$, the upper bound of recording distances. (Note that for pure notational convenience here we use H for the number of *observed* multinomial categories, while in Chapter 7 it was for *all* categories.) Let y_h be the frequency of encounters in distance interval h, and let $\mathbf{y} = (y_1,\ldots,y_H)$ denote the vector of frequencies with $n = \sum_h y_h$. We assert that the vector of observations $\mathbf{y}$ has a multinomial distribution:

$$\mathbf{y}|N \sim Multinomial(N; \{\pi_h\})$$

with parameters N and cell probabilities $\{\pi_h\}$. The number of individuals not detected will be denoted by $n_0 = N - n$, and the corresponding cell probability for these undetected individuals is $\pi_{H+1} = 1 - \sum_{h=1}^{H} \pi_h$.

It remains to define the cell probabilities π_h. These are, in words, "the probability that an individual occurs and is detected in distance class h" (see Buckland et al., 2001, p. 52), which is also "the probability that an individual is detected given that it occurs in class h *times* the probability that it occurs in class h." This is, using a formula,

$$\Pr(y = 1 \ and \ x \in h) = \Pr(y = 1|x \in h)\Pr(x \in h)$$

which we'll simplify by writing

$$\pi_h = \overline{p}_h \psi_h$$

where ψ_h is the probability that x is located in distance interval h, which is, for a line transect, $\psi_h = (b_{h+1} - b_h)/B$, i.e., just the interval width over the transect half-width, and is implied by the uniform distribution assumption. But what is $\Pr(y = 1|x \in h)$? It is the integral over the distance band of the detection function multiplied by the conditional probability of x, given that x is in distance band h:

$$\overline{p}_h = \int_{x \in h} \Pr(y = 1|x, x \in h)\Pr(x|x \in h)dx$$

Under the uniformity assumption x is also uniformly distributed in each interval, and so the conditional pdf of x is $[x|x \in h] = 1/(b_{h+1} - b_h)$. Putting this all together then, we just integrate the detection function over the interval with an adjustment for area

$$\pi_h = (1/B) * \int_{x \in h} \Pr(y = 1|x, x \in h).$$

In our on-going line transect example with the muskoxen suppose we use 10-m distance bands for distances between 0 and 100 m. Then the conditional probability density of x is 1/10 for each 10-m distance band, and we have to do the calculation

$$\overline{p}_h = \int_{b_{h-1}}^{b_h} \Pr(y = 1|x)/10 \, dx$$

and then the multinomial cell probabilities are

$$\pi_h = \psi_h \overline{p}_h = (1/10)\overline{p}_h$$

The last cell probability, π_0, is 1 minus the sum of the rest: $\pi_0 = 1 - \sum \pi_h$. Thus, binned distance data have a multinomial distribution with these cell probabilities and we can obtain the full likelihood directly:

$$L(\sigma, n_0; \mathbf{y}) = \frac{(n+n_0)!}{n!n_0!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_H^{y_H} \pi_0^{n_0} \tag{8.5}$$

where $n = \sum y_h$.

8.2.4.1 Conditional and Other Likelihoods

The conditional likelihood is easily derived by noting that the distribution of detections among the H distance classes, conditional on n, is also multinomial but with multinomial index n instead of N and conditional probabilities $\pi_h^c = \pi_h/(1-\pi_0)$. To uncondition on n, in order to obtain the full likelihood again, we note that $n \sim \textit{Binomial}(N, 1-\pi_0)$ and so the full likelihood is the product of the conditional likelihood and this binomial piece, which leads us back to Eq. (8.5). The point is, the conditional multinomial and the binomial for n together are exactly equivalent specifications to the multinomial likelihood in Eq. (8.5) (see Section 5.1.2 in Royle and Dorazio, 2008). In practice, some may analyze the full likelihood by retaining the two individual pieces, although there is no need to do this in most cases.

As a final point, we might think like a Bayesian here and assume that N is not a fixed number to estimate but, rather, is itself the realization of a random variable. If $N \sim \textit{Poisson}(\lambda)$, then this implies precisely that $n \sim \textit{Poisson}((1-\pi_0)\lambda)$, and we can estimate the parameter λ instead of N. This “Poisson integrated full likelihood” has the following form (see Royle et al., 2014, p. 192):

$$L(\boldsymbol{\theta}, \lambda) = \left\{ \prod_{i=1}^{n} g(x_i; \boldsymbol{\theta}) \right\} \lambda^n \exp(-\lambda(1-\pi_0))$$

(see also Borchers and Efford, 2008). Instead of a Poisson prior for N we can consider a $\textit{Binomial}(M, \psi)$ where M is prescribed, similar in spirit to the model for N in data augmentation (DA; Royle and Dorazio, 2012). For large M this approximates the Poisson prior but it yields a different likelihood, having the form (see Royle and Dorazio, 2008, p. 238):

$$\left(\prod_{i=1}^{n} \psi g(x_i; \boldsymbol{\theta}) \right) \frac{M!}{n!(M-n)!} (1 - \psi(1-\pi_0))^{M-n}.$$

This binomial integrated form of the full likelihood is equivalent to the model we would analyze using data augmentation (see below). These various considerations give us a number of essentially equivalent ways to analyze distance sampling models with binned data. It is worth knowing of these different formulations because one or another may have certain advantages in a given instance. For example, in Bayesian analysis of the distance sampling model we use a method of DA that is easily implemented in the BUGS language. The model implied by DA is the binomial integrated likelihood just shown.

8.2.4.2 Simulating Binned Distance Sampling Data

There are two ways to go about simulating binned distance sampling data, and we show both here. First, we can simulate continuous-space data exactly as we have done before with function `sim.ldata` and then aggregate into distance intervals. Second, we can simulate directly multinomial observations with cell probabilities π_h. Here is a script that does it both ways using a half-normal detection function. (To verify that the same cell probabilities are produced, you could execute it with very large population size N).

```
set.seed(2015)
# Design settings and truth (population size N and detection function g)
interval.width <- 10
strip.width <- 100    # half-width really (one side of transect)
nbins <- strip.width%/%interval.width
sigma <- 30           # Scale parameter of half-normal detection function
g <- function(x, sig) exp(-x^2/(2*sig^2)) # Half-normal detection function
N <- 200              # Population size

# Method 1: simulate continuous distances and put into intervals
x <- runif(N, -strip.width, strip.width) # Distance all animals
p <- g(x, sig=sigma)  # Detection probability
y <- rbinom(N, 1, p)  # only individuals with y=1 are detected
x <- x[y==1]          # this has direction (right or left side of transect)
x <- abs(x)           # now it doesn't have direction

# Compute the distance category of each observation
xbin <- x %/% interval.width + 1   # note integer division function %/%

# Multinomial frequencies, may have missing levels
y.obs <- table(xbin)

# Pad the frequencies to include those with 0 detections
y.padded <- rep(0,nbins)
names(y.padded) <- 1:nbins
y.padded[names(y.obs)] <- y.obs
y.obs <- y.padded
y.true <- c(y.obs, N-length(xbin)) # Last category is "Not detected"

# Relative frequencies by binning continuous data (pi). These should compare
#  with the cell probabilities computed below when N is very large
(y.rel <- y.true/N)    # Last category is pi(0) from above
(pi0.v1 <- y.rel[nbins+1])
0.635

# Compute detection probability in each distance interval
dist.breaks <- seq(0, strip.width, by=interval.width)
p <- rep(NA, length(dist.breaks)-1)
for(j in 1:length(p)){
  p[j] <- integrate(g, dist.breaks[j], dist.breaks[j+1],
     sig=sigma)$value / (dist.breaks[j+1]-dist.breaks[j])
}
```

```
round(p, 2)
[1] 0.98 0.88 0.71 0.51 0.33 0.19 0.10 0.05 0.02 0.01

# Compute the multinomial cell probabilities analytically. These are exact.
# psi = probability of occurring in each interval
interval.width <- diff(dist.breaks)
psi <- interval.width/strip.width
pi <- p * psi
sum(pi)                    # This is 1 - pi(0) from above
[1] 0.3756716
(pi0.exact <- 1-sum(pi))
[1] 0.6243284              # Compare with 0.635 above

# Method 2: Use rmultinom to simulate binned observations directly
# This includes 0 cells AND n0
pi[length(p)+1] <- 1 - sum(pi)
(y.obs2 <- as.vector(rmultinom(1, N, prob=pi)))
(y.obs2 <- y.obs2[1:nbins]) # Discard last cell for n0 (because not observed)
```

We see that, under this model, we expect to encounter about 38% of the individuals along the transect.

Now let's take our simulated data and obtain the MLEs of the model parameters. Keep in mind that we have only simulated a single multinomial sample, which we could think of as sampling one transect of a certain length or multiple transects but then pooling the resulting data. Shortly we will get on to the meta-population sampling context and consider having spatial replicates, but, for now we continue our focus on the basic analysis of distance sampling data. The likelihood is just a multinomial, so if we package up most of the previous simulation R code into a function that computes the likelihood, given the parameter values, a multinomial data vector, and the distance breaks, then we can use `optim` or `nlm` to obtain the MLEs. Note that the multinomial full likelihood must include the combinatorial term in N and, as before, we parameterize the model in terms of the number of uncaptured individual n_0 so that $N = n + n_0$. Note also that the observed multinomial frequencies may include zero counts in some distance bands, and we must be sure to pad the observed data vector when appropriate.

```
Lik.binned <- function(parm, data, dist.breaks){
# Note that the parameters are parm[1] = log(sigma), parm[2] = log(n0)

sigma <- exp(parm[1])
n0 <- exp(parm[2])
p <- rep(NA, length(dist.breaks)-1)
for(j in 1:length(p)) {
   p[j] <- integrate(g, dist.breaks[j], dist.breaks[j+1],
      sig=sigma)$value / (dist.breaks[j+1]-dist.breaks[j])
}
psi <- interval.width/strip.width
pi <- p * psi
pi0 <- 1-sum(pi)

N <- sum(data) + n0
-1*(lgamma(N+1)-lgamma(n0+1) + sum(c(data,n0)*log(c(pi,pi0))))
}
```

```
# Evaluate likelihood for some particular value of the parameters
Lik.binned(c(2,0), data=y.obs, dist.breaks=dist.breaks)
[1] 335.1482

# Obtain the MLEs for the simulated data
optim(c(2,0), Lik.binned, data=y.obs, dist.breaks=dist.breaks)
$par
[1] 3.263681 5.006211

$value
[1] -117.0331
```

The MLE of N is $\widehat{N} = \widehat{n}_0 + n = \exp(5.006) + n$ where $n = \sum y_h$ for the observed distance categories; this yields $\widehat{N} = \exp(5.006) + 73 = 222.3063$, not too far from the true data-generating value of $N = 200$ (and compare to the estimates of musk oxen abundance we obtained back in Section 8.2.3). (Note the object `y.obs2` simulated using the `rmultinom` function produces a different number of observed individuals due to randomness of random number generation.) The point of this was to build basic tools of simulating and fitting distance sampling data, for use later when we develop hierarchical distance sampling models, and so we don't do anything else with this for right now. At the same time, simulating data sets using R provides another and, to some, perhaps more intuitive, description of the basic distance sampling model than algebra.

8.2.5 POINT TRANSECT DATA

So-called "point transect data" are distance sampling data collected from circular point counts where an observer stands at a point and records distance to detected individuals within some radius B. Formulation of the likelihood for such data follows the same logic as for transect data, but this time the natural probability density for distance x is *not* uniform. Recall that, in the transect case, if we assume a uniform distribution of individuals, then distance is also uniform. But in the case of a circular sample unit, the uniform distribution of individuals implies a triangular distribution for distance. We can understand this by computing the cumulative probability distribution directly, noting that the probability that distance is less than any value x should be proportional to the area of a circle of radius x relative to a circle of radius B. That is, the cumulative distribution function F of distance x is (note that for clarity we use the uppercase X for the variable and x to represent a specific value):

$$F(x) = \Pr(X \leq x) = \frac{\pi x^2}{\pi B^2}$$

The probability density is then obtained by differentiating with respect to x, i.e., $f(x) = \partial F(x)/\partial x$, which produces:

$$f(x) = \frac{2x}{B^2}$$

This triangular distribution has increasing probability density with distance from the center of the circle, to account for the increasing area of successive annuli. Recall our general expression for the conditional likelihood:

$$[x_i | y_i = 1] = \frac{g(x_i; \theta)[x]}{\int_x [y = 1|x][x]dx}$$

In the case of point transects [x] is not constant and so it does *not* cancel out from the numerator and denominator and the likelihood has to retain the [x] $\equiv f(x)$ bit. Using the half-normal detection function model, the contribution of each observed distance x_i to the conditional likelihood looks like this (note that $1/B^2$ cancels from numerator and denominator):

$$L(\sigma|x_i) = \frac{\exp\left(x_i^2/2\sigma^2\right)2x_i}{\int\limits_x \exp(x^2/2\sigma^2)2xdx}$$

Shortly we will simulate some data and fit the model using the conditional likelihood expressed as an R function.

For binned point count data the vector of frequencies of encounters in each distance class, including the cell "not encountered," has a multinomial distribution with cell probabilities π_h for $h = 1,2,\ldots,H$ distance classes, and the last cell, $H+1$, corresponds to "not encountered." To compute these cell probabilities we have to do the integrations over successive annuli of the circle, and we have to make a smallish bit of math argumentation following our development in Section 8.2.4 above to do this. The multinomial cell probability π_h is, in words, "the probability that an individual is detected *and* in distance class h," which is the same as saying, "the probability that an individual is detected, given that it occurs in class h, times the probability that it occurs in class h." This is, using a formula,

$$\Pr(y = 1 \textit{ and } x \in h) = \Pr(y = 1|x \in h)\Pr(x \in h)$$

which we write simply as: $\pi_h = \bar{p}_h \psi_h$ where ψ_h is the probability that x is in distance interval h, which is $\psi_h = (A_{h+1} - A_h)/\pi B^2$, where $A_h = \pi r_h^2$ is the area of a circle having radius r_h. So, ψ_h here is just the area of the annulus over the whole area of the point count circle. But what is $\Pr(y = 1|x \in h)$? Our expression from before had us compute this integral:

$$\bar{p}_h = \int\limits_{x \in h} \Pr(y = 1|x, x \in h)\Pr(x|x \in h)dx.$$

Under the assumption that individuals are uniformly distributed in space, x has the triangular distribution on [0, B] as noted above. But for the interval it has a slightly different form; we have to compute $f(x|x \in h)$, which is $f(x|x \in h) = f(x)/\Pr(x \in h)$, and it works out that the pdf of x is $[x|x \in h] = 2x/(b_{h+1}^2 - b_h^2)$. Also, $\psi_h = (b_{h+1}^2 - b_h^2)/B^2$. Putting this all together, we just integrate the detection function over the interval with an adjustment for area:

$$\pi_h = \psi_h \int\limits_{b_h}^{b_{h+1}} \Pr(y = 1|x, x \in h) \times 2x\Big/\left(b_{h+1}^2 - b_h^2\right)$$

This might seem a bit conceptual, but let's see how this looks in the form of an R function, which we then apply to our point count situation:

```
# Define function to compute cell probs for binned distance sampling
cp.ri <-function(radius1, radius2, sigma){
   Pi <- 3.141593
   a <- Pi*radius2^2 - Pi*radius1^2
```

```
    integrate(function(x, s=sigma) exp(-x^2 / (2 * s^2)) * x, radius1,
      radius2)$value *(2*Pi/a)
  }

  # Define distance intervals and compute multinomial probabilities
  delta <- 0.5                  # Width of distance bins
  B <- 3                        # Max count distance
  dist.breaks <-seq(0, B, delta)  # Make the interval cut points
  nD <-length(dist.breaks)-1
  sigma <- 1
  p.x <-rep(NA,nD)              # Conditional detection probabilities
  for(i in 1:nD){
     p.x[i] <- cp.ri(dist.breaks[i], dist.breaks[i+1], sigma =1)
  }
  area <- 3.141593 * dist.breaks[-1]^2
  ring.area <- diff(c(0, area))
  # Pr(detection| in ring)*Pr(in ring)
  cp <- p.x* ring.area/sum(ring.area)
```

These cell probabilities are used below to simulate data using the `rmultinom` function or to construct the multinomial likelihood, which was given previously (for line transects) as,

$$L(\sigma, n_0; \mathbf{y}) = \frac{(n+n_0)!}{n!n_0!}\pi_1^{y_1}\pi_2^{y_2}\ldots\pi_H^{y_H}\pi_0^{n_0}.$$

Here, $n = \sum y_h$, but we have to go through the gyrations of computing the cell probabilities π_h for the case of a circle instead of a nice rectangular transect.

8.2.5.1 *Simulating Point Transect Data*

To simulate point transect data we can simulate individuals uniformly on a $2B \times 2B$ square and then toss out those individuals located $>B$ from the center point of the square. This produces continuous distance data, which we can then bin into distance classes. Alternatively, we could compute the multinomial cell probabilities and simulate multinomial (i.e., grouped) observations directly using `rmultinom`. We show both in this section. First, we define a function that will simulate a population of individuals on the square and return the required data objects and give summary plots (Figure 8.2):

```
  sim.pdata <- function(N=1000, sigma=1, B=3, keep.all=FALSE) {
  # Function simulates coordinates of individuals on a square
  # Square is [0,2*B] x[0,2*B], with a count location on the center
  # point (B,B)
  # Function arguments:
  #     N: total population size in the square
  #     sigma: scale of half-normal detection function
  #     B: circle radias
  #     keep.all: return the data for y = 0 individuals or not
```

```
# Plot the detection function
par(mfrow = c(1,2))
curve(exp(-x^2/(2*sigma^2)), 0, B, xlab="Distance (x)", ylab="Detection prob.", lwd = 2,
main = "Detection function", ylim = c(0,1))
text(0.8*B, 0.9, paste("sigma:", sigma))

# Simulate and plot simulated data
library(plotrix)
u1 <-runif(N, 0, 2*B)           # (u1,u2) coordinates of N individuals
u2 <- runif(N, 0, 2*B)
d <- sqrt((u1 - B)^2 + (u2 - B)^2) # distance to center point of square
plot(u1, u2, asp = 1, pch = 1, main = "Point transect")
N.real <- sum(d<= B)           # Population size inside of count circle

# Can only count indidividuals in the circle, so set to zero detection probability of
individuals in the corners (thereby truncating them):
p <- ifelse(d < B, 1, 0) * exp(-d*d/(2*(sigma^2)))
# Now we decide whether each individual is detected or not
y <- rbinom(N, 1, p)
points(u1[d <= B], u2[d <= B], pch = 16, col = "black")
points(u1[y==1], u2[y==1], pch = 16, col = "blue")
points(B, B, pch = "+", cex = 3, col = "red")
draw.circle(B, B, B)

# Put all of the data in a matrix:
#       (note we don't care about y, u, or v normally)

if(!keep.all){
   u1 <- u1[y==1]
   u2 <- u2[y==1]
   d <- d[y==1]
}
return(list(N=N, sigma=sigma, B=B, u1=u1, u2=u2, d=d, y=y, N.real=N.real))
}

# obtain a data set by distance sampling a population of N=1000 out to a distance of B=3
set.seed(1234)
tmp <-sim.pdata(N=1000, sigma=1, keep.all=FALSE, B=3) # produces Figure 8.2
attach(tmp)
```

Here we simulated a complete data set (Figure 8.2) but returned only the location coordinates of each individual (u_1, u_2) and the observed distances for captured ($y = 1$) individuals. We will use these locations later in the chapter. For now, we develop likelihood analyses of the distance data. We start by taking the data just simulated and "bin" them by using the integer division function `%/%`. To apply this to our simulated distance data, we do the following (and note we must *always* make sure that we have a vector of encounter frequencies that includes the zeros, i.e., the distance bins where nobody was detected!).

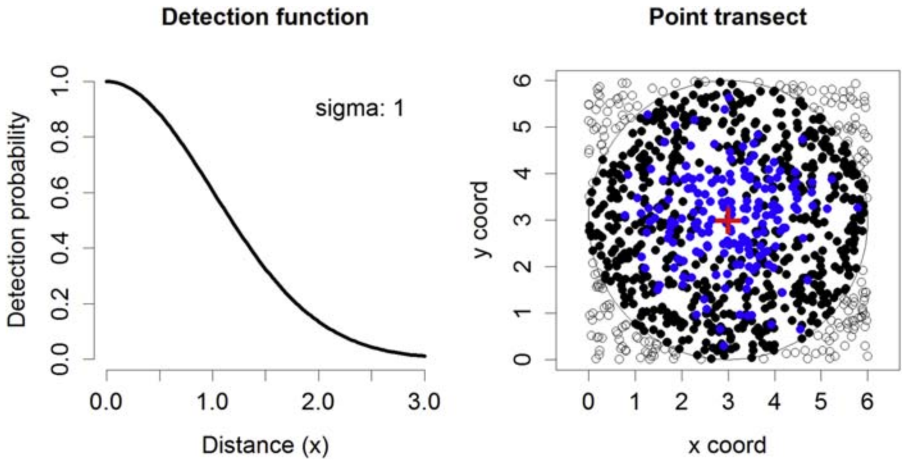

FIGURE 8.2

Plot of the simulation of point transect data (see R code `sim.pdata`). (left) Form of detection function for chosen value of sigma. (Right) Map of simulated point pattern. Locations of individuals inside of the maximum detection distance (circle) are black, individuals detected are blue, and the point count location is at the red cross.

```
# Bin the data and tabulate the bin frequencies. Be sure to pad the 0s!
delta <- 0.5                         # Width of distance bins
dist.breaks <-seq(0, B, delta)       # Make the interval cut points
dclass <- tmp$d %/% delta +1         # Convert distances to categorical distances
nD <-length(dist.breaks)-1           # How many intervals do we have ?
y.obs <- table(dclass)               # Next pad the frequency vector
y.padded <- rep(0, nD)
names(y.padded) <- 1:nD
y.padded[names(y.obs)] <- y.obs
y.obs <- y.padded
```

Next, we will simulate binned distance data with the `rmultinom` function which uses numerical construction of the multinomial cell probabilities based on the `cp.ri` function defined in the previous section:

```
cp <- c(cp, 1-sum(cp)) # Compute the last cell and add it to the vector

as.vector(rmultinom(n=1, size=1000, prob=cp))
[1] 25 59 74 53 25  9 755
```

We can check that these cell probabilities are in agreement with what we get when we simulate continuous distance data and then bin them as follows: we simulate an extremely large data set (e.g., $N = 10^5$) and compute the relative frequencies in each distance class. These should then be very close to the multinomial cell probabilities obtained by numerical integration.

8.2.5.2 Likelihood Analysis of Point Transect Data

To do likelihood analysis of the point transect data, we define an R function that evaluates the likelihood for a particular value of the parameter(s) and other arguments, such as the observed distance data and the upper distance bound of counting *B*. Here we provide three versions of the likelihood: (1) the multinomial full likelihood for binned data; (2) the full likelihood for continuous distance data; (3) the conditional likelihood for continuous data. (We omit the conditional likelihood based on the binned data.) In the following block of code, we define the three likelihood functions and then optimize each to obtain the estimated population size for the circular sample unit (note: you may want to re-create the data set from above in case you overwrote stuff in your R workspace).

```
# (1) Define multinomial likelihood for binned data
Lik.binned.point <- function(parm, data, dist.breaks){
   sigma <- exp(parm[1])
   n0 <- exp(parm[2])
   p.x <-rep(NA, nD)
   for(i in 1:nD){
      p.x[i] <- cp.ri(dist.breaks[i], dist.breaks[i+1], sigma =sigma)
   }
   area <- 3.141593 * dist.breaks[-1]^2
   ring.area <- diff(c(0, area))
   cp <- p.x* ring.area/sum(ring.area) # Pr(detection| in ring)*Pr(in ring)
   pi0 <- 1-sum(cp)
   N <- sum(data) + n0
   negLL <- -1*(lgamma(N+1)-lgamma(n0+1) + sum(c(data,n0)*log(c(cp,pi0))))
return(negLL)
}

# Fit model
mle1 <- optim(c(2,0), Lik.binned.point, data=y.obs, dist.breaks=dist.breaks)

# (2) Define full likelihood for continuous data
Lik.cont.point <- function(parm, data, B){
   sigma <- exp(parm[1])
   n0 <- exp(parm[2])
   n <- length(data)
   N <- n + n0
   p <- exp(-data*data/(2*sigma*sigma))
   f <- 2*data/(B^2)
   pbar <- integrate(function(r, s=sigma) exp(-r^2 / (2 * s^2)) * r, 0, B)$value*2/(B^2)
   negLL <- -1*sum( log(p*f/pbar )) -1*(lgamma(N+1) - lgamma(n0+1) +
      n * log(pbar) + n0*log(1-pbar))
   return(negLL)
}

# Fit model
mle2 <- optim(c(0, 5), Lik.cont.point, data=tmp$d, B=B, hessian=TRUE)

# Compare two solutions and with realized true value of N
(Nhat.binned <- length(tmp$d) + exp(mle1$par[2]))
[1] 792.9483
(Nhat.cont <- length(tmp$d) + exp(mle2$par[2]))
[1] 799.0507
tmp$N.real
[1] 797
```

These are different by about 1%! It is tempting to regard the estimate under the continuous distance model as being better, and, indeed, it is both closer to the truth and also the data were simulated in that way, so in this case, it is. However, in practice, we don't know the truth, and there is no theoretical reason to prefer the continuous distance estimator to the binned estimator. They are alternative models both being used as approximations to the truth, which we don't know.

Finally, we provide the conditional likelihood for the continuous distance data, which only has a single parameter σ (for the half-normal model) to be estimated:

```
# (3) Define conditional likelihood for continuous data
Lik.cond.point <- function(parm, data, B){
   sigma <- exp(parm)
   p <- exp(-data*data/(2*sigma*sigma))
   f <- 2*data/(B^2)
   pbar <- integrate(function(x, s=sigma) exp(-x^2 / (2 * s^2)) * x, 0, B)$value*2/(B^2)
   negLL <- -1*sum( log(p*f/pbar ))
   return(negLL)
}

# Fit the model
mle3 <- optim(c(0), Lik.cond.point, data=tmp$d, B=B, method="Brent", hessian=TRUE,
lower=-10, upper=10)

# Inspect the output
mle3
$par
[1] 0.01523024

$value
[1] 160.8509

[ . . . output truncated . . . ]

$hessian
          [,1]
[1,] 544.0608

# Estimated sigma
(sigma.hat <- exp(mle3$par))
[1] 1.015347
```

We see that $\widehat{\sigma} = \exp(0.0152) = 1.015$, which is pretty close to the true value of 1, and so we surmise that our likelihood implementation is likely correct. With the conditional estimator we don't obtain directly an estimate of density or population size. Instead we have to use the MLE of σ after the fact and compute $\overline{p}$, which we do as follows, finding that the conditional estimator of N $(= n/\overline{p})$ is in the same ballpark as the other two, and yet slightly different.

```
# Estimated average detection probability and conditional estimator of N
pbar <- integrate(
    function(x, s=sigma.hat) exp(-x^2 / (2 * s^2)) * x, 0, B)$value*2/(B^2)

 (Nhat.condl   <- length(d) / pbar)
[1] 800.2388
```

8.2.6 SENSITIVITY TO BIN WIDTH

When you write a paper that uses distance sampling with binned data, one criticism raised by a referee undoubtedly will be that it would be better to use a continuous data model instead of an "approximation." However, both the continuous distance model and the model for binned data are mere approximations to the actual data-generating process, which we don't know. Nevertheless, it is interesting to see how similar a discrete distance model is to a particular continuous data-generating model, which we can handily know if we happen to be simulating data.

To evaluate the effect of binning data that are generated from a truly continuous model, we show a small simulation study here that you can easily repeat for your own situation. We simulate data as above (all of the code is repeated here) using bin widths of $\delta = 0.5$, and we fit both the continuous and discrete distance models as in the previous section. We do 1000 Monte Carlo (simulation) replicates of each bin width scenario and, at the end, we compute the mean of $\widehat{N}$ and also the standard deviation. All of this goes as follows:

```
set.seed(1234)
simrep <- 1000                   # Number of sim reps
simout <- matrix(NA, nrow=simrep, ncol=3)
colnames(simout) <- c("N.real", "N.binned", "N.continuous")
delta <- 0.5                     # Set width of bins

# Begin simulation loop
for(sim in 1:simrep){
   tmp <- sim.pdata(N=1000, sigma=1, keep.all=FALSE, B=3)
   B <- tmp$B
   d <- tmp$d
   N.real <- tmp$N.real

   # Bin data, tabulate frequencies and pad 0s if necessary
   dist.breaks <- seq(0, B, delta)
   dclass <- d%/%delta + 1       # Convert distances to categorical distances
   nD <- length(dist.breaks) -1# How many intervals do we have ?
   y.obs <- table(dclass)        # Next pad the frequency vector
   y.padded <- rep(0, nD)
   names(y.padded) <- 1:nD
   y.padded[names(y.obs)] <- y.obs
   y.obs <- y.padded

   # Obtain the MLEs using both models
   binned.est <- optim(c(2,0), Lik.binned.point, data=y.obs,
     dist.breaks=dist.breaks)
   cont.est <- optim(c(1, 6), Lik.cont.point, data=d, B=B, hessian=TRUE)
   Nhat.binned <- length(d) + exp(binned.est$par[2])
   Nhat.cont <- length(d) + exp(cont.est$par[2])

   # Store results in a matrix
   simout[sim,] <- c(N.real, Nhat.binned, Nhat.cont)
}
```

```
# Now summarize the output
apply(simout, 2, mean)
      N.real     N.binned    N.continuous
    785.0580     782.5183        782.7314

sqrt(apply(simout, 2, var))
      N.real     N.binned     N.continuous
    12.84647     84.25812         82.70834
```

What we see here is essentially the same expected value of both estimators and an only very slightly increased standard deviation of the binned estimator. This is a general truth: binning has essentially no effect on bias and only negligibly decreases precision of the estimator *compared to fitting the correct continuous distance model*. Of course, in practice, we will not know the true model.

To gauge the sensitivity to bin width, we repeated the analyses for the same 1000 data sets using bin widths of 0.1, 0.2, 0.3, 0.5, 0.6, and 1.0 (all of these produce equal-width bins for $B = 3$). The results are tabulated as follows:

```
Width  Truth   Nhat.binned  SD.binned
0.1    785.06    782.82       82.80
0.2    785.06    782.90       82.82
0.3    785.06    782.45       83.49
0.5    785.06    782.52       84.26
0.6    785.06    782.19       84.38
1.0    785.06    783.51       89.27
```

What we see here is negligible bias in the estimated population size (much less than 1%), although we do see a systematic increase in the standard deviation of the estimator. For a bin width of 0.1 the SD is about the same as the continuous distance model, whereas for the bin width of 1.0 (one-third of the total count radius!) the SD increases by about 8%. What all of this means is that you'll suffer a small cost in terms of precision by using a bin width that is extremely coarse but no practical effect at all for bin widths that are roughly <10% of the count radius. Note that the mean MLE for all six cases is systematically less than the true average of 785.06. This is because the same 1000 data sets were used for each simulation and a different 1000 data sets will produce a different (higher or lower) discrepancy, on average. As always, we encourage you to play around with simulations to gain intuition and understanding of the effects of data collecting and analysis decisions.

8.2.7 SPATIAL SAMPLING

In practice we virtually always have more than a single sample unit. What do we do in this situation if we're doing conventional distance sampling ? Let's say we have S transects, which we imagine to be S multinomial samples with size (= local population size) N_s, then the conventional distance sampling approach is to just pool all of the distance data and fit a single conditional likelihood to it (remember our change of notation for site index; see Section 8.1). So the N_s are not involved in this at all. Spatial sampling is ignored. This is surprising, since probably most applications of CDS have an interest in assessing hypotheses about spatiotemporal variation in N_s, and yet when doing CDS we usually ignore the problem almost completely by using the conditional likelihood of the pooled data.

We say “almost” since spatial sampling is not entirely ignored. The conditional estimator of density is

$$\widehat{D} = n / \left(\widehat{p} * 2 * L * B \right)$$

To obtain the variance of this estimator we need to estimate the quantity *Var(n)*, which CDS estimates using the variance of n_s among sample units (Buckland et al., 2001, p. 79). The latter is usually called the “encounter rate variance” and can be thought of as a nonparametric estimator of the variance among spatial units. That is the only way in which spatial sampling is dealt with in CDS.

8.3 BAYESIAN CONVENTIONAL DISTANCE SAMPLING

While we have not yet analyzed a hierarchical distance sampling (HDS) model, we have learned the four basic operations: simulation and analysis of continuous and binned data for line transects and for point transects. For point transects, this is only slightly more complicated than for transects due to the different geometry. We have shown how to write out the likelihood and obtain MLEs. We now cover how to analyze these models using Bayesian methods, which will come in handy when we finally get to the analysis of hierarchical distance sampling models. Part of the reason for building up this material in such a leisurely way is that there is not “one way” to analyze HDS models, just as we saw with CDS models. The various ways of analyzing these models will all be useful in analyzing HDS models in different situations or using different BUGS engines. So, we will go through the various formulations of the models above (line/point transects, continuous/binned distances, conditional/full likelihood) using Bayesian methods implemented in BUGS. To implement continuous distance models in BUGS often takes a little bit of trickery because the probability distribution of the observed distances is not usually a standard form. It is easy to resolve this by using distance bins, in which case we can use a categorical or multinomial distribution, where we build the cell probabilities explicitly. And, as we saw previously, there is almost no statistical cost for using a discrete distance model, even when we happen to know the correct continuous distance model.

For analyzing the full likelihood in BUGS we use the idea of *parameter-expanded data augmentation* (PX-DA or DA for short; Royle et al., 2007a; Royle and Dorazio, 2012), which we also discussed briefly in Section 7.8.4 and will encounter again in Chapters 9, 11, and later. The idea of DA is that we take our data set of n observed encounters (and distances) and augment it with a large number of $M - n$ “not encountered” individuals, which necessarily have missing distance data. We further expand our model by introducing a set of binary latent variables (the data augmentation variables) z_i, which are indicators of whether an individual in the larger data set of size M is a “real” individual, so that the observation of 0 is a stochastic (sampling) zero, or whether it is a fixed zero, which is to say $y = 0$ with probability 1. We assume $z_i \sim$ *Bernoulli*(ψ), where ψ is the data augmentation parameter. (There is a sense in which DA transforms a capture-recapture type of model into an occupancy type of model, and our notation with z and ψ is intended to reflect this.) This formulation of the model is equivalent to putting a *Binomial*(M, ψ) prior distribution on population size N (see Section 8.2.4.1 above) and a *Uniform*(0,1) prior on ψ. Those two priors together imply that the marginal (induced) prior distribution for N is *Discrete Uniform*(0, M). Of course the binomial prior is roughly equivalent to a Poisson prior distribution when M is large, but even when it is not, it is not clearly a better or worse prior than the Poisson, just different. Next, we demonstrate the use of DA for line transect data with continuous and binned data measurements.

8.3.1 BAYESIAN ANALYSIS OF LINE TRANSECT DATA

We illustrate a Bayesian analysis of distance sampling data from a transect using the famous impala data set from Burnham et al. (1980; analysis modified from Royle and Dorazio 2008, p. 235). In this study, distance data were collected along a 60 km transect. If we use a transect width of 1000 m, the total area is 60 km^2, which we'll use to convert estimated *N* to estimated density, *D*. The line transect situation is especially easy to deal with in BUGS because we can specify the uniform distribution for distance explicitly and then, conditional on the distances, the observation model is specified as a simple Bernoulli trial, like in a logistic regression. The Bayesian formulation of the distance sampling model therefore makes clear the elegant hierarchical structure of distance sampling as involving a process model (the distribution of individuals) and an observation model (the detection or nondetection of individuals; for this we use the half-normal model throughout). Next, we input the data directly into the R workspace, package things up, and run BUGS as follows:

```
# Get data and do data-augmentation
# Observed distances (meters) in the impala data set
x <- c(71.93, 26.05, 58.47, 92.35, 163.83, 84.52, 163.83, 157.33,
22.27, 72.11, 86.99, 50.8, 0, 73.14, 0, 128.56, 163.83, 71.85,
30.47, 71.07, 150.96, 68.83, 90, 64.98, 165.69, 38.01, 378.21,
78.15, 42.13, 0, 400, 175.39, 30.47, 35.07, 86.04, 31.69, 200,
271.89, 26.05, 76.6, 41.04, 200, 86.04, 0, 93.97, 55.13, 10.46,
84.52, 0, 77.65, 0, 96.42, 0, 64.28, 187.94, 0, 160.7, 150.45,
63.6, 193.19, 106.07, 114.91, 143.39, 128.56, 245.75, 123.13,
123.13, 153.21, 143.39, 34.2, 96.42, 259.81, 8.72)

B <- 500 # Strip half-width. Larger than max observed distance
nind <- length(x)

# Analysis of continuous data using data augmentation (DA)
nz <- 200 # Augment observed data with nz = 200 zeroes
y <- c(rep(1, nind), rep(0, nz)) # Augmented inds. have y=0 by definition
x <- c(x, rep(NA, nz)) # Value of distance are missing for the augmented

# Bundle and summarize data set
str( win.data <- list(nind=nind, nz=nz, x=x, y=y, B=B) )

# Save text file with BUGS model
cat("
model {

# Priors
sigma ~ dunif(0,1000)  # Half-normal scale
psi ~ dunif(0,1)       # DA parameter

# Likelihood
for(i in 1:(nind+nz)){
   # Process model
   z[i] ~ dbern(psi)  # DA variables
   x[i] ~ dunif(0, B) # Distribution of distances
```

```
    # Observation model
    logp[i] <- -((x[i]*x[i])/(2*sigma*sigma)) # Half-normal detection fct.
     p[i] <- exp(logp[i])
     mu[i] <- z[i] * p[i]
    y[i] ~ dbern(mu[i])    # Simple Bernoulli measurement error process
}

# Derived quantities
N <- sum(z[1:(nind + nz)]) # Population size
D <- N / 60                # Density, with A = 60 km^2 when B = 500
}
",fill=TRUE,file="model1.txt")

# Inits
zst <- y
inits <- function(){ list (psi=runif(1), z=zst, sigma=runif(1,40,200)) }

# Params to save
params <- c("N", "sigma", "D")

# Experience the raw power of BUGS and summarize marginal posteriors
library(R2WinBUGS)
bd <- "c:/Program Files/WinBUGS14/"   # May have to adapt for your computer
out1 <- bugs(win.data, inits, params, "model1.txt", n.thin=2,n.chains=3,
   n.burnin=1000, n.iter=11000, debug=TRUE, DIC=FALSE, bugs.dir=bd)
print(out1, 3)
         mean     sd  2.5%   25%   50%   75%  97.5%  Rhat n.eff
N     221.468 24.746 174.0 204.0 222.0 240.0 267.00 1.004   600
sigma 131.466 10.798 112.8 123.8 130.6 138.2 155.20 1.003  1100
D       3.691  0.412   2.9   3.4   3.7   4.0   4.45 1.004   600
```

Next, we provide an analysis of the impala data but using binned data to demonstrate the BUGS implementation using data augmentation. We first need to convert the distance data into distance bins, which we define here to be 50 m bins. Then we specify the model in BUGS using the `dcat` distribution for individual distance class observations. In BUGS we have to define detection probability for each interval, which we do by evaluating the half-normal detection probability function at the midpoint of each interval (input as data), which will look like this:

```
log(p[g]) <- -midpt[g] * midpt[g] / (2 * sigma * sigma)
```

We also have to compute the probability mass for each distance interval:

```
pi[g] <- delta / B  # probability of x in each interval

# Analysis of binned data using data augmentation
delta <- 50              # Width of distance bins
xg <- seq(0, B, delta)   # Make the interval cut points
dclass <- x %/% delta + 1 # Convert distances to distance category
nD <- length(xg) -1      # N intervals = length(xg) if max(x) = B
```

```
# Bundle data
# Note data changed to include dclass, nD, bin-width delta and midpt
midpt <- xg[-1] - delta/2 # Interval mid-points
str( win.data <- list (nind=nind, nz=nz, dclass=dclass, y=y, B=B,
   delta=delta, nD=nD, midpt=midpt) )  # Bundle and summarize

# BUGS model specification
cat("
model{
# Priors
psi ~ dunif(0, 1)
sigma ~ dunif(0, 1000)

# Likelihood
# Construct conditional detection probability and Pr(x) for each bin
for(g in 1:nD){           # midpt = mid point of each cell
   log(p[g]) <- -midpt[g] * midpt[g] / (2 * sigma * sigma)     # half-normal model
   pi[g] <- delta / B     # probability of x in each interval
}

for(i in 1:(nind+nz)){
   z[i] ~ dbern(psi)              # model for individual covariates
   dclass[i] ~ dcat(pi[])         # population distribution of distance class
   mu[i] <- z[i] * p[dclass[i]]  # p depends on distance class
   y[i] ~ dbern(mu[i])
}
# Derived quantities: Population size and density
N <- sum(z[])
D <- N / 60
}
",fill=TRUE, file = "model2.txt")

# Inits function
zst <- y # DA variables start at observed value of y
inits <- function(){ list (psi=runif(1), z=zst, sigma=runif(1,40,200)) }

# Parameters to save
params <- c("N", "sigma", "D")

# Unleash WinBUGS and summarize posteriors
out2 <- bugs(win.data, inits, params, "model2.txt", n.thin=2, n.chains=3,
   n.burnin=1000, n.iter=11000, debug=TRUE, DIC=FALSE, bugs.dir = bd)
print(out2, 2)
        mean    sd   2.5%    25%    50%    75%  97.5% Rhat n.eff
N     218.48 25.27 170.00 200.00 218.00 236.00 267.00 1.01   320
sigma 134.24 11.16 114.90 126.30 133.30 141.20 158.50 1.00   840
D       3.64  0.42   2.83   3.33   3.63   3.93   4.45 1.01   320
```

These are similar to those obtained previously using the continuous distance model, and they should become more similar as we decrease the bin width, and perhaps also by increasing the number of MCMC iterations so as to reduce Monte Carlo error.

8.3.2 OTHER FORMULATIONS OF THE DISTANCE SAMPLING MODEL

We have shown transect models with binned and continuous distances analyzed in BUGS using data augmentation. But there are many other exciting formulations of distance sampling models. For example, in Chapter 9 we will provide a formulation of the model not in terms of distance but in terms of location of encounter. This is one of our favorites. Another type of model that might be useful to develop is that in which we have binned data but parameterize the model in terms of latent continuous observations. The observed bin data are a "cut" of the continuous data that have the standard DS model for continuous data. This is an interesting idea because it allows us to "downscale" the observations, or make predictions, to a finer scale than the available observations. We think that this can be done directly in JAGS using its function for interval censoring. Finally, it is possible to formulate conditional likelihood models in BUGS. This is somewhat more complicated because the conditional distribution of the distances cannot be specified directly. However, we can formulate the model for binned data and then compute the conditional cell probabilities explicitly in the BUGS code. Then we can use either a multinomial distribution for distance bin frequencies or the categorical distribution for individual distance bin observations. We will show this in the next section in the context of point transect data, but we leave it to you as an exercise for the transect case.

8.3.3 A TREATISE ON THE INTEGRATION OF MATHEMATICAL FUNCTIONS IN ONE DIMENSION

Before we proceed with a development of Bayesian analysis of point transect data in BUGS, we first discuss the basic concept of *integration using the "rectangular rule"* whereby, to compute the integral of some function $f(x)$, we approximate the function by a bunch of rectangles centered at points x_i and then sum up the area of those rectangles. This is precisely the approach we used in the analysis of binned transect data in the previous section. The virtue of being able to do this is that it gives us a way of parameterizing *any* distribution in BUGS if we just know a formula for its pdf. Instead of the "ones trick" or the "zeros trick" (Section 5.8, and p. 204–206 in Lunn et al., 2013), we just compute the area under chunks of the curve and use `dcat` as a model for a binned version of the variable. The error in making a discrete approximation to any continuous distribution is usually negligible compared to the MC error in our MCMC analysis, provided we use enough bins.

To demonstrate this we show, in Figure 8.3, the right side of a normal kernel, and we ask, "what is the area under this curve?" Of course, we can use the `integrate` function to compute that directly (which we do below). Or, we can line up a bunch of rectangles as shown in Figure 8.3 and sum up the area of those rectangles. All of the code for doing this together is as follows:

```
sigma <- 2                          # normal scale (standard deviation)
curve(exp(-x^2 / (2*sigma^2)), 0, 10, frame = F)

delta <- 1                          # bin width
mid <- seq(0.5, 9.5, delta)         # 10 rectangles
f.mid <- exp(-mid^2 / (2*sigma^2))
barplot(f.mid, add=T, space=0, col="grey", width=delta)
curve(exp(-x^2 / (2*sigma^2)), 0, 10, add = TRUE, col = "blue", lwd = 3)
```

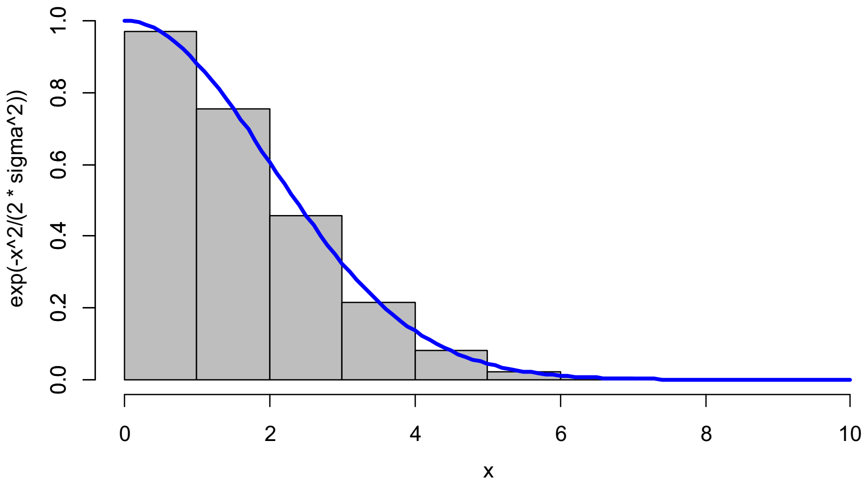

FIGURE 8.3

All you need to know about integration in one figure. The area under the blue curve is approximated by the area of the gray rectangles.

```
# Integral done using the integrate function
integrate( function(x){ exp(-x^2/(2*sigma^2)) }, lower=0, upper=100)
2.506628 with absolute error < 8.1e-07

# Summing up the 10 rectangular areas:
areas <- f.mid * delta
sum(areas)
[1] 2.506627
```

At the end of the day we see no practical difference between these two results, and therefore it stands to reason that if we model continuous distributions in BUGS using even a moderate number of rectangles, we would not expect to be badly led astray. Moreover, as we've said a few times already in this chapter and will continue to say some more, discrete distributions are perfectly reasonable models of random variables, without even having to think about them as approximations to any continuous thing.

8.3.4 BAYESIAN ANALYSIS OF POINT TRANSECT DATA

As with the formulation of the conditional likelihood for line transect data, the problem with analyzing point transect data in BUGS is that there is no built-in distribution for the distances, which we noted, in Section 8.2.5, has a triangular distribution. This can be dealt with in several ways by using the "zeros trick" or the "ones trick" (see Chelgren et al., 2011b, for a neat distance sampling application) or, alternatively, we can analyze the model for binned data on the circle and use the `dcat` or multinomial model. We show that here. This is sufficient in practice because of course we can always use a huge number of distance intervals to obtain what essentially is a continuous distance model if that was necessary. But also, as we discussed in Section 8.2.6, the continuous distance model is not any more correct than an a similar step function model. The key thing is to identify, for each individual, which distance category it belongs in and then compute the probabilities for that categorical random variable. We can use these to specify a model based either on the conditional likelihood or we can use specify an 'unconditional' model based on data augmentation. We show both of these.

The mathematical argumentation to define the distance class probabilities goes like this: The probability density of detections is the product of the detection function (here, a half-normal) and the density of x:

$$\Pr(\text{detection in } x) = \Pr(\text{detection}|x)\Pr(x) = \exp\left(-\frac{1}{2\sigma^2}x^2\right)f(x),$$

where $f(x)$ = pdf of radial distance from a point (for a transect $f(x)$ is constant). The probability distribution of radial distance x on a circle of radius B is:

$$f(x) = \frac{2x}{B^2}$$

(there is more mass in a distance band as you move far away from the point). As we showed back in Section 8.2.5, we need to integrate $\exp\left(-\frac{1}{2\sigma^2}x^2\right)f(x)$ over distance bands to get multinomial cell probabilities (this is what `unmarked` does, see Section 8.4.3). But from our treatise on the integration of one-dimensional functions, we know that, approximately, the multinomial cell probabilities should be "width times height" of a rectangle centered at x_h and therefore:

$$\pi(x_h) = \Pr(x_h - \delta/2 \le x \le x_h + \delta/2) \approx \exp\left(-\frac{1}{2\sigma^2}x_h^2\right)f(x_h)\delta$$

(this is the rectangular approximation to an integral). So we can choose x_h to be the mid-points of our intervals or we can use many very narrow intervals and then add them up into coarser bins. To implement a conditional model we need to compute conditional distance class probabilities:

$$\pi^c(x_h) = \frac{\pi(x_h)}{1 - \pi_0},$$

where the denominator: $1 - \pi_0 = \Pr(capture) = \sum_h \pi(x_h)$

Note that the conditional distance class probabilities are used when we analyze the conditional likelihood version of the model in BUGS, in which case N is a derived parameter. We can also analyze the full likelihood version of the model by data augmentation, which we also show below. The full likelihood DA version uses the probabilities $f(x_h)$ as the distribution for the population of true distances and then also models detection/nondetection of each individual, y_i.

We can simulate binned distance sampling data either directly by simulating categorical random variables, or we can simulate continuous distance data and bin the data as we did previously. We will simulate continuous data here, bin the data into classes, and then use the categorical distribution in BUGS to fit the point transect model.

```
### Version 1: Point count data in BUGS (conditional likelihood)
# Simulate a data set and harvest the output
set.seed(1234)
tmp <- sim.pdata(N=200, sigma=1, keep.all=FALSE, B=3)
attach(tmp)
```

```
# Chop the data into bins
delta <- 0.1              # width of distance bins for approximation
xg <- seq(0, B, delta)   # Make the mid points and chop up the data
midpt <- xg[-1] - delta/2

# Convert distances to categorical distances (which bin?)
dclass <- d %/% delta + 1
nD <- length(midpt) # how many intervals
nind <- length(dclass)

# Bundle and summarize data set
str( win.data <- list(midpt=midpt, delta=delta, B=B, nind=nind, nD=nD, dclass=dclass) )

# BUGS model specification, conditional version
cat("
model{

# Prior for single parameter
sigma ~ dunif(0, 10)

# Construct cell probabilities for nD cells (rectangle approximation)
for(g in 1:nD){   # midpt[g] = midpoint of each distance band
   log(p[g]) <- -midpt[g] * midpt[g] / (2*sigma*sigma)
   pi[g] <- (( 2 * midpt[g] ) / (B*B)) * delta
   f[g] <- p[g] * pi[g]
   fc[g] <- f[g] / pcap
}
pcap <- sum(f[]) # capture prob. is the sum of all rectangular areas

# Categorical observation model
for(i in 1:nind){
   dclass[i] ~ dcat(fc[])
}

# Derived quantities: population size and density
N <- nind / pcap
D <- N/(3.141*B*B)
}
",fill=TRUE, file="model3.txt")

# Inits function
inits <- function(){list (sigma=runif(1, 1, 10)) }

# Params to save
params <- c("sigma", "N","D")

# MCMC settings
ni <- 62000  ;  nb <- 2000  ;  nt  <-  2  ;  nc <- 3

# Run BUGS and summarize posteriors
bd <- "c:/Program Files/WinBUGS14/"    # May have to adapt this to your computer
```

```
out3 <- bugs(win.data, inits, params, "model3.txt", n.thin=nt,
n.chains=nc, n.burnin=nb, n.iter=ni, debug=FALSE, bugs.dir = bd)


## Version 2: point count data (full likelihood with data augmentation)
# Do data augmentation (for same simulated data set)
M <- 400
nz <- M - nind
y <- c(rep(1, nind), rep(0, nz))
dclass <- c(dclass, rep(NA, nz))

# Bundle and summarize data set
str( win.data <- list(midpt=midpt, delta=delta, B=B, nind=nind, nD=nD, dclass=dclass,
y=y, nz=nz) )

# BUGS model
cat("
model{

# Priors
sigma ~ dunif(0, 10)
psi ~ dunif(0, 1)

# Construct cell probabilities for nD cells (rectangle approximation)
for(g in 1:nD){            # midpt[g] = midpoint of each distance band
   log(p[g]) <- -midpt[g] * midpt[g] / (2*sigma*sigma)
   pi[g] <- ((2 * midpt[g]) / (B * B)) * delta
   pi.probs[g] <- pi[g] / norm
   f[g] <- p[g] * pi[g]
   fc[g] <- f[g] / pcap     # conditional probabilities
}
pcap <- sum(f[])# capture prob. is the sum of all rectangular areas
norm <- sum(pi[])

# Categorical observation model
for(i in 1:(nind+nz)){
    z[i] ~ dbern(psi)
    dclass[i] ~ dcat(pi.probs[])
    mu[i] <- p[dclass[i]] * z[i]
    y[i] ~ dbern(mu[i])
}

# Derived quantities: population size and density
N <- sum(z[])
D <- N/(3.141*B*B)

}
",fill=TRUE,file="model4.txt")

# Inits
inits <- function(){list (sigma=runif(1,1,10), psi=runif(1) ) }
```

```
# Parameters to save
params <- c("sigma", "N","D","psi")

# MCMC settings
ni <- 62000  ;  nb <- 2000  ;  nt  <-  2  ;  nc <- 3

# Run BUGS and summarize posteriors
out4 <- bugs(win.data, inits, params, "model4.txt", n.thin=nt,
   n.chains=nc, n.burnin=nb, n.iter=ni, debug=FALSE, bugs.dir = bd)

# Compare posterior summaries
print(out3,2)    # Conditional likelihood
Inference for Bugs model at "model3.txt", fit using WinBUGS,
 3 chains, each with 62000 iterations (first 2000 discarded), n.thin = 2
 n.sims = 90000 iterations saved
            mean    sd   2.5%    25%    50%    75%  97.5%  Rhat n.eff
sigma       1.11  0.13   0.91   1.02   1.10   1.18   1.41     1 90000
N         142.96 27.20  93.71 123.80 141.60 160.60 200.00     1 90000
D           5.06  0.96   3.31   4.38   5.01   5.68   7.08     1 90000
deviance  246.92  1.59 245.80 245.90 246.30 247.30 251.40     1 90000

print(out4,2)    # Full likelihood
Inference for Bugs model at "model4.txt", fit using WinBUGS,
 3 chains, each with 62000 iterations (first 2000 discarded), n.thin = 2
 n.sims = 90000 iterations saved
            mean    sd   2.5%    25%    50%    75%   97.5%  Rhat n.eff
sigma       1.09  0.12   0.90   1.01   1.08   1.16    1.37     1 28000
N         150.16 35.14  91.00 125.00 147.00 172.00  229.00     1 51000
D           5.31  1.24   3.22   4.42   5.20   6.08    8.10     1 51000
psi         0.38  0.09   0.22   0.31   0.37   0.43    0.58     1 59000
deviance  396.00 18.01 361.80 383.70 395.70 407.90  432.40     1 44000
```

We see a slight inconsistency between the two analyses, both producing posterior means of N slightly less than the true value of $N = 152$ (`=sum(tmp$N.real)`). However, relative to the uncertainty of these estimates (quantified by the posterior standard deviation), the discrepancy between the two estimates is fairly small.

8.4 HIERARCHICAL DISTANCE SAMPLING (HDS)

Now we transition from the basic elements of conventional distance sampling models to situations where we have distance sampling data collected at S spatial locations, usually either transects or point count locations, but we could also have a mixture of both or even strange shapes or irregular transects. (Remember our change of notation for sites, which now have index s, which runs from 1 to S.) As we've noted a few times before, the traditional way to deal with this in distance sampling is to pool the distance data from all S spatial locations and estimate the parameter(s) of the detection function, e.g., σ for a half-normal. This is used to obtain an estimate of density, and then the variance is based on the encounter rate variance, which does use some information from among the sample units. However, conventional distance sampling does not directly address problems of spatial inference either in the form of modeling variation in N_s or local density across sample units s or making explicit predictions at other transects or point locations. We would argue that modeling variation in N among sample units is critically important and, indeed, often the primary interest in studies that use distance sampling. Therefore, HDS models should be in every ecologist's toolbox.

The models we deal with here assume that N_s is the population size of spatial sample unit s, and they don't make any explicit assumptions about "*within sample unit*" variation in density. Rather, they assume that the average covariate value defined for the sample unit is meaningful for explaining among sample unit variation. Thus, when we assume that $N_s \sim Poisson(\lambda_s)$ the parameter λ_s is constant for the sample location s and represents the mean for the sample unit. This is *not* to say that HDS models assume that density is constant within a sample unit, just that the aggregate density is adequately modeled by the covariates defined for the sample unit. We discuss this more in Chapter 9.

8.4.1 HDS DATA STRUCTURE AND MODEL

To develop distance sampling in an explicit meta-population setting, we suppose that S distinct spatial units are sampled using the distance sampling protocol. These might be transects or point counts for birds, distributed in some region (e.g., a park or forest). Distance x is naturally viewed as a continuous measurement, but for now we jump right into the discrete distance class formulation (we discuss continuous measurements shortly). Thus, we consider binned data here, wherein distances are recorded in discrete intervals from the central point of observation for each site. Let $h = 1, 2, \ldots, H$ index the distance classes, with end points, or *distance breaks* $(c_1, c_2), (c_2, c_3), \ldots, (c_H, c_{H+1})$. Here, c_{H+1} is the maximum distance at which birds were counted, or the radius of the point count (which we called B previously). Let y_{sh} be the observed count of individuals in distance class h for site $s = 1, 2, \ldots, S$. The data structure is summarized in Table 8.1.

We follow the basic ideas of the binomial and multinomial mixture models of Chapters 6 and 7, and assume that sample unit s has local abundance N_s, which is a random variable having a suitable distribution. For now we assume:

$$N_s \sim Poisson(\lambda_s)$$

Table 8.1 Typical hierarchical distance sampling data structure. For each of S transects we have encounter frequencies in each of a number of distance classes (three illustrated here). In addition, we may have one or more site-level covariates (v).

Transect	dclass 1 (0–50 m)	dclass 2 (50–100 m)	dclass 3 (100–200 m)	Covariate1	Covariate2
Transect 1	2	0	1	v_{11}	v_{12}
Transect 2	3	0	0	v_{21}	v_{22}
Transect 3	2	1	1	v_{31}	v_{32}
⋮	⋮	⋮	⋮	⋮	
Transect S	4	2	0	v_{S1}	v_{S2}

where one or more covariates (v) may influence the expected abundance, λ_s, on a suitable scale:

$$\log(\lambda_s) = \beta_0 + \beta_1 v_s.$$

In addition, we assume the detection frequencies in each of the H distance classes have, conditional on the population size N_s, a multinomial distribution:

$$(y_{s1}, \ldots, y_{sH}) \sim Multinomial(N_s, \boldsymbol{\pi}_s)$$

where π_{sh} is the multinomial cell probability for distance class h and sample unit s—these depend on detection-function parameter(s) σ. These are computed exactly as we've described previously for either line or point transects. If there are no site covariates then there are no additional considerations. If, on the other hand, we also have covariates that influence detection probability and vary across sites, then we have to compute the multinomial cell probabilities separately for each site. It would be natural to model such covariates on the parameter σ, allowing this parameter to vary as a function of covariates that may be site specific (Marques et al., 2007). For example:

$$\log(\sigma_s) = \alpha_0 + \alpha_1 v_s$$

The scale parameter (σ) is a continuous, nonnegative number, hence, it is natural to apply a linear model of covariates on a transformed scale, typically the log, as for the expected count (λ) in a Poisson GLM.

8.4.2 HDS IN `unmarked`

The `unmarked` package has two specific functions for fitting HDS models. The older function `distsamp` assumes a basic closed population model (i.e., for one sample occasion) and allows only for a Poisson abundance model: $N_s \sim Poisson(\lambda_s)$. The more general (and newer) function `gdistamp` allows for a type of simple open population structure (see Chapter 9) and also for a negative binomial abundance distribution. Both of these functions accept only binned distance data, i.e., multinomial distance class frequencies.

In this section we mainly consider the `distsamp` function, which works about the same way as `multinomPois` (Chapter 7), where the abundance parameters N_s are marginalized out of the multinomial likelihood according to

$$[\mathbf{y}_s|\alpha, \beta] = \sum_{N_s=0}^{\infty} [\mathbf{y}_s|N_s, \alpha][N_s|\beta]$$

In practice, we truncate the upper bound of summation (called K in `unmarked`). As do other N-mixture model fitting functions in `unmarked`, `distsamp` uses as a default for K the maximum observed count at a site plus 100. For the Poisson abundance model this likelihood reduces to the product of independent Poisson components (as in Section 7.4), which is very efficient to compute with. As with other `unmarked` functions, `distsamp` and `gdistsamp` have helper functions, called `unmarkedFrameDS` and `unmarkedFrameGDS`, for packaging up the data for use by either fitting function. These functions take the basic data and some metadata and set it all up in an `unmarkedFrame` for analysis by either fitting function and certain summary functions. We demonstrate their use shortly. The `distsamp` function itself is used roughly like the multinomial *N*-mixture functions of Chapter 7, and it has a few critical arguments as follows (not all arguments shown):

```
distsamp(formula, data, keyfun=c("halfnorm", "exp", "hazard", "uniform"),
   output=c("density", "abund"), unitsOut=c("ha", "kmsq"), starts, ...)
```

`formula`: Double right-hand formula describing detection covariates followed by abundance covariates. ~1 ~1 would be an intercepts-only model.

`data`: object of class unmarkedFrameDS, containing response matrix, covariates, distance interval cut points, survey type ("line" or "point"), transect lengths (for survey = "line"), and units ("m" or "km") for cut points and transect lengths. See example for set up.

`keyfun`: One of the following detection functions: "halfnorm", "hazard", "exp", or " uniform." See details.

`output`: Model either "density" or "abund"

`unitsOut`: Units of density. Either "ha" or "kmsq" for hectares and square kilometers, respectively.

`starts`: Vector of starting values for parameters.

A more versatile function that allows the fitting of negative binomial abundance models is the `gdistsamp` function, which has the following structure:

```
gdistsamp(lambdaformula, phiformula, pformula, data,
       keyfun =c("halfnorm", "exp", "hazard", "uniform"),
       output = c("abund","density"), unitsOut = c("ha", "kmsq"),
       mixture = c("P", "NB"), K, starts, method = "BFGS",
       se = TRUE, rel.tol=1e-4, ...)
```

This function will also handle a type of open population structure, allowing for random availability to sampling with a parameter ϕ (corresponding to a temporary emigration probability $1-\phi$). We cover such models in Chapter 9. The temporal structure is accommodated via a third formula argument, "`phiformula`," in addition to formulas for the expected abundance ($E(N)$, λ) and p. Note that the three formulas are separated in a `gdistsamp` call by a comma while there is no comma between the two components of the hierarchical model in `distsamp`. **And, importantly, the order of the formulas is "lambda, phi, p," whereas the order in the double formula in `distsamp` is "p, lambda."** Because there are two "state" parameters (λ and ϕ), certain summary functions such as `predict`, which previously required `type=state` or `type=det`, require that you now specify which state parameter to predict (`type=lambda`, `type=phi` or `type=det`).

The abundance distribution (Poisson or negative binomial) is specified by the "`mixture`" argument. As with the `pcount` (Chapter 6) and the `gmultmix` function (see Section 7.5.3), the negative binomial parameterization used in `gdistsamp` contains the mean, λ, and logarithm of the negative binomial "size" parameter, say $log(\tau)$, with a variance of $\lambda + \lambda^2/\tau$. Therefore, as $1/\tau \rightarrow 0$ or $\tau \rightarrow \infty$, the negative binomial tends to the Poisson (i.e., no overdispersion is indicated). The `gdistsamp` function relies on the same basic technology as `gmultmix`. In general it computes the marginal likelihood by summing over possible values of N from $N = 0$ up to some finite value $N = K$. Thus K has to be specified either by the user or it defaults to 100 plus the maximum count at a site. Sensitivity of the estimates to the choice of K, beyond some large number such as this default, may indicate problems with parameter estimability (Couturier et al., 2013; Dennis et al., 2015a). The main arguments to the `gdistsamp` function are defined as follows:

```
lambdaformula: A right-hand side formula describing the abundance
                 covariates.
   phiformula: A right-hand side formula describing the availability
                 covariates.
     pformula: A right-hand side formula describing the detection function
                 covariates.
         data: An object of class 'unmarkedFrameGDS"
       keyfun: One of the following detection functions: "halfnorm",
                 "hazard", "exp", or "uniform." See details.
       output: Model either "density" or "abund"
     unitsOut: Units of density. Either "ha" or "kmsq" for hectares and
                 square kilometers, respectively.
      mixture: Either "P" or "NB" for the Poisson and negative binomial
                 models of abundance.
            K: An integer value specifying the upper bound used in the
                 integration.
```

In the following section, we apply `distsamp` and `gdistsamp` to the analysis of a distance sampling data set on the island scrub-jay.

8.4.3 EXAMPLE: ESTIMATING THE GLOBAL POPULATION SIZE OF THE ISLAND SCRUB-JAY (ISSJ)

The island scrub-jay (*Aphelocoma insularis*; Figure 8.4) is a species that is endemic to Santa Cruz Island, California (Figures 8.5 and 8.6), and of some conservation interest to the National Park Service (NPS) and other organizations due to the extremely local distribution of the species and previous reports of low and declining population sizes. Our esteemed colleague T. S. Sillett and others initiated an island-wide survey in 2008 to obtain a statistical estimate of population size. The study was reported in Sillett et al. (2012), and we reproduce some of the analyses here. The island scrub-jay data are available in `unmarked` by typing `data(issj)`.

The data are distance sampling point count data from 307 point count locations (Figure 8.5) with counts made out to 300 m. For analysis, the raw distance data were binned into three 100-m distance classes because nearby birds were responding to the observer (by moving closer, representing responsive movement), so it was believed that the large distance classes should mitigate that affect. The objectives were to (1) estimate the global population size; (2) produce a map of the distribution of the population (i.e., $E(N)$), as a function of local habitat conditions; and (3) make predictions of $E(N)$ under alternative/historical landscapes. Until recently, the island had been heavily grazed by livestock, and an intense removal effort successfully eradicated the livestock causing vegetation to return to historical conditions. But we have the vegetation map for the state of the island under heavy grazing, and so we want make a hypothetical statement about how many jays there may have been.

FIGURE 8.4

A proud island scrub-jay (*Aphelocoma insularis*). *(Photo credit: Melanie Klein.)*

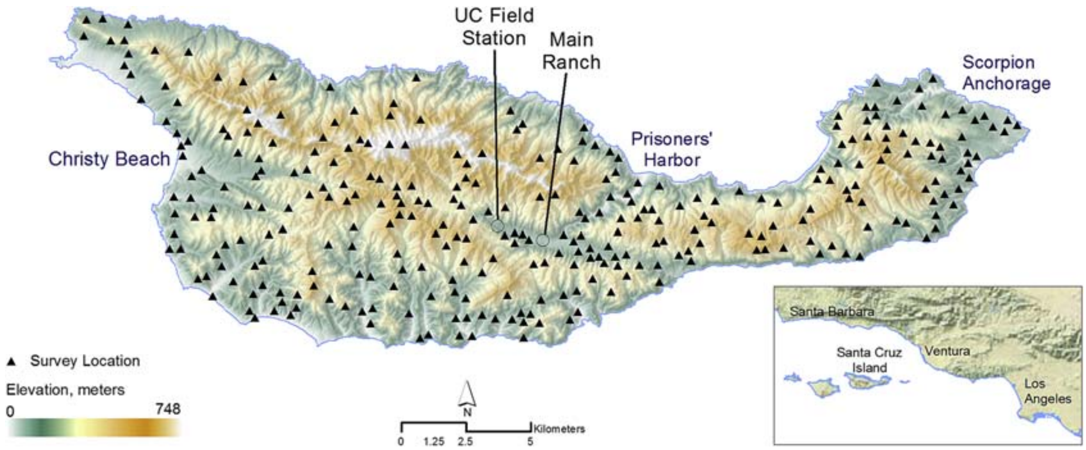

FIGURE 8.5

Santa Cruz Island, Channel Islands, California. The 307 distance sampling point count locations are shown as solid triangles.

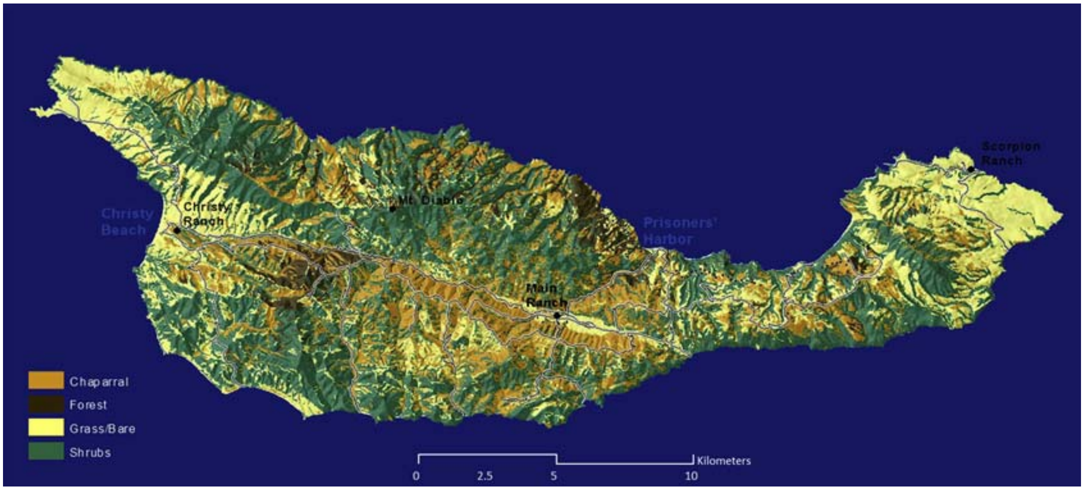

FIGURE 8.6

Current habitat and topography of Santa Cruz island.

To do the analysis in `unmarked` we load the data and do a few other bookkeeping things such as computing the area of the point count circle to use as an offset so that density in ha is reported, and we build the `unmarkedFrameDS`.

```
# Load, view and format the ISSJ data
library(unmarked)
data(issj)
```

```
round(head(issj), 2)
  issj[0-100] issj(100-200] issj(200-300]    x         y  elevation forest  chaparral
1        0           0             2  234870.1  3767154      51.39   0.02       0.24
2        0           0             0  237083.0  3766804     156.88   0.01       0.47
3        0           0             0  235732.0  3766717     144.81   0.02       0.77
4        0           0             0  237605.0  3766719     184.27   0.26       0.21
5        0           0             0  234239.1  3766570     111.35   0.00       0.00
6        0           0             0  235005.1  3766420     204.13   0.16       0.34

# Package things up into an unmarkedFrame
covs <- issj[,c("elevation", "forest", "chaparral")]
area <- pi*300^2 / 100^2              # Area in ha
jayumf <- unmarkedFrameDS(y=as.matrix(issj[,1:3]),
   siteCovs=data.frame( covs, area),
   dist.breaks=c(0, 100, 200, 300),
   unitsIn="m", survey="point")
```

We note that the island scrub-jay (ISSJ) data comes with site covariates, which are elevation of the point, and cover types `forest` and `chaparral`. We input these along with the point count area (constant for all points) using the `siteCovs` argument to the `unmarkedFrameDS`. We also specify the distance breaks and the units of distance intervals, and declaring `survey="point"`, that the sample unit is a point count circle (clearly, distances and distance breaks must be in the same units). Now we're ready to fit a few models, which we do like this, first with chaparral as a covariate on both the detection scale σ and on expected abundance λ and also elevation as a covariate on λ, and then, the second model has a constant σ:

```
# Fit model 1
(fm1 <- distsamp(~chaparral ~chaparral + elevation + offset(log(area)),
    jayumf, keyfun="halfnorm", output="abund"))

Call:
distsamp(formula = ~chaparral ~ chaparral + elevation + offset(log(area)),
    data = jayumf, keyfun = "halfnorm", output = "abund")

Abundance:
             Estimate      SE       z    P(>|z|)
(Intercept) -3.50982  0.31261  -11.23  2.99e-29
chaparral    4.11503  0.62458    6.59  4.44e-11
elevation   -0.00216  0.00073   -2.96  3.11e-03

Detection:
                    Estimate    SE      z     P(>|z|)
sigma(Intercept)      5.02  0.161  31.15  5.65e-213
sigmachaparral       -1.07  0.319  -3.36   7.73e-04

AIC: 964.7203

# Fit model 2

(fm2 <- distsamp(~1 ~chaparral + elevation + offset(log(area)),
    jayumf, keyfun="halfnorm", output="abund"))
```

```
Call:
distsamp(formula = ~1 ~ chaparral + elevation + offset(log(area)),
    data = jayumf, keyfun = "halfnorm", output = "abund")

Abundance:
             Estimate       SE      z   P(>|z|)
(Intercept) -2.71972  0.200946 -13.53  9.77e-42
chaparral    2.12760  0.309172   6.88  5.92e-12
elevation   -0.00212  0.000728  -2.91  3.59e-03

Detection:
 Estimate      SE     z  P(>|z|)
     4.58  0.0488  93.9        0

AIC: 976.2306
```

We see the model with chaparral on both σ and λ is favored by a wide margin according to AIC. We check the goodness-of-fit of this model by bootstrapping the `fitstats` function first introduced in Section 7.5.4 (see also Section 6.8).

```
 (pb <- parboot(fm1, fitstats, nsim=1000, report=5))
 (c.hat <- pb@t0[2] / mean(pb@t.star[,2])) # c-hat as ratio of observed
                                 # and mean of expected value of Chi2 (under H0)
                                 # (see, e.g., Johnson et al., Biometrics, 2010)
    Chisq
2.590553

residuals(fm1)                   # Can inspect residuals
plot(pb)                         # Not shown
print(pb)

Call: parboot(object = fm1, statistic = fitstats, nsim = 1000, report = 5)

Parametric Bootstrap Statistics:
                t0  mean(t0 - t_B)  StdDev(t0 - t_B)  Pr(t_B > t0)
SSE            421           262.7              16.5             0
Chisq         2357          1447.1              66.1             0
freemanTukey   210            42.9              10.1             0

t_B quantiles:
              0%  2.5%  25%  50%  75%  97.5%  100%
SSE          110   130  147  158  169    193   213
Chisq        739   806  864  903  950   1056  1364
freemanTukey 131   147  161  167  174    187   202

t0 = Original statistic compuated from data
t_B = Vector of bootstrap samples
```

The bootstrap analysis shows that the model does not fit at all, with not a single bootstrap sample falling to the right of the observed value, for any of the three fit statistics. That is, under the Null hypothesis of a fitting model, we don't expect to see any more extreme values of the fit statistics than their value for the observed data set. The “c-hat” statistic (Johnson et al., 2010) indicates a fairly high degree of overdispersion (2.59). This suggests that there is more unexplained variation in the data than allowed for by the distributional assumptions of the model. In the worst case this could mean that the main inference, e.g., regarding covariate effects, is wrong (i.e., that the model is structurally wrong), while in the much less dramatic case it could simply mean that we have unstructured noise, which would make the SEs too small and CIs too narrow. To mitigate that, we go through a more detailed process of model fitting, evaluation, and prediction by expanding the covariate structure of the model. First, however, we will standardize the covariates in the `unmarkedFrame` because this generally causes the fitting and analysis functions to perform more smoothly (i.e., often it avoids various types of numerical errors or errors due to bad starting values). You can repeat the analysis below without standardizing the covariates to see what happens.

```
# Standardize the covariates
sc <- siteCovs(jayumf)
sc.s <- scale(sc)
sc.s[,"area"] <- pi*300^2 / 10000  # Don't standardize area
siteCovs(jayumf) <- sc.s
summary(jayumf)

unmarkedFrameDS Object

point-transect survey design
Distance class cutpoints (m):  0 100 200 300

307 sites
Maximum number of distance classes per site: 3
Mean number of distance classes per site: 3
Sites with at least one detection: 76

Tabulation of y observations:
   0   1   2   3   4   5   6   9 <NA>
 833  53  19   9   1   2   3   1    0

Site-level covariates:
   elevation           forest              chaparral             area
 Min.   :-1.4884   Min.   :-0.49215   Min.    :-1.1562   Min.   :28.27
 1st Qu.:-0.7974   1st Qu.:-0.49215   1st Qu. :-0.8721   1st Qu.:28.27
 Median :-0.1687   Median :-0.44295   Median  :-0.2014   Median :28.27
 Mean   : 0.0000   Mean   : 0.00000   Mean    : 0.0000   Mean   :28.27
 3rd Qu.: 0.6650   3rd Qu.:-0.06982   3rd Qu. : 0.6872   3rd Qu.:28.27
 Max.   : 3.5731   Max.   : 5.42362   Max.    : 2.8809   Max.   :28.27

# Fit a bunch of models and produce a model selection table.
fall <- list()  # make a list to store the models

# With the offset output=abund is the same as output = density
fall$Null <- distsamp(~1 ~offset(log(area)), jayumf, output="abund")
fall$Chap. <- distsamp(~1 ~chaparral + offset(log(area)), jayumf,
     output="abund")
```

```
fall$Chap2. <- distsamp(~1 ~chaparral+I(chaparral^2)+offset(log(area)),
    jayumf, output="abund")
fall$Elev. <- distsamp(~1 ~ elevation+offset(log(area)), jayumf,
    output="abund")
fall$Elev2. <- distsamp(~1 ~ elevation+I(elevation^2)+offset(log(area)),
    jayumf, output="abund")
fall$Forest. <- distsamp(~1 ~forest+offset(log(area)), jayumf,
    output="abund")
fall$Forest2. <- distsamp(~1 ~forest+I(forest^2)+offset(log(area)),
    jayumf, output="abund")
fall$.Forest <- distsamp(~forest ~offset(log(area)), jayumf,
    output="abund")
fall$.Chap <- distsamp(~chaparral ~offset(log(area)), jayumf,
    output="abund")
fall$C2E. <- distsamp(~1 ~ chaparral + I(chaparral^2) + elevation +
    offset(log(area)),jayumf, output="abund")
fall$C2F2. <- distsamp(~1 ~chaparral + I(chaparral^2) + forest +
    I(forest^2)+offset(log(area)), jayumf, output="abund")
fall$C2E.F <- distsamp(~forest ~chaparral+I(chaparral^2)+elevation+
    offset(log(area)), jayumf, output="abund")
fall$C2E.C <- distsamp(~chaparral ~chaparral + I(chaparral^2) + elevation +
    offset(log(area)), jayumf, output="abund")

# Create a fitList and a model selection table
 (msFall <- modSel(fitList(fits=fall)))

         nPars     AIC   delta    AICwt   cumltvWt
C2E.C        6  951.35    0.00  9.9e-01       0.99
C2E.         5  961.01    9.66  7.9e-03       1.00
C2E.F        6  962.95   11.60  3.0e-03       1.00
Chap2.       4  965.95   14.60  6.7e-04       1.00
C2F2.        6  968.13   16.78  2.2e-04       1.00
Chap.        3  981.39   30.04  3.0e-07       1.00
.Chap        3 1007.02   55.67  8.1e-13       1.00
Forest2.     4 1015.07   63.72  1.4e-14       1.00
Elev2.       4 1017.33   65.98  4.7e-15       1.00
Elev.        3 1018.10   66.75  3.2e-15       1.00
Null_D       2 1018.12   66.77  3.1e-15       1.00
Null         2 1018.12   66.77  3.1e-15       1.00
Forest.      3 1019.65   68.30  1.5e-15       1.00
.Forest      3 1020.08   68.73  1.2e-15       1.00

# Check out the best model
fall$C2E.C

Call:
distsamp(formula = ~chaparral ~ chaparral + I(chaparral^2) +
    elevation + offset(log(area)), data = jayumf, output = "abund")
```

```
Abundance:
                 Estimate     SE      z    P(>|z|)
(Intercept)        -2.562 0.1589 -16.12  1.75e-58
chaparral           1.230 0.1602   7.68  1.64e-14
I(chaparral^2)     -0.282 0.0775  -3.64  2.68e-04
elevation          -0.238 0.0926  -2.57  1.02e-02

Detection:
                   Estimate      SE      z   P(>|z|)
sigma(Intercept)     4.686  0.0682  68.75 0.000000
sigmachaparral     - 0.208  0.0626  -3.32 0.000892

AIC: 951.3504

# Check out the goodness-of-fit of this model
(pb.try2 <- parboot(fall$C2E.C, fitstats, nsim=1000, report=5))
Call: parboot(object = fall$C2E.C, statistic = fitstats, nsim = 1000, report = 5)

Parametric Bootstrap Statistics:
               t0  mean(t0 - t_B)  StdDev(t0 - t_B) Pr(t_B > t0)
SSE           425           267.7             16.13            0
Chisq        2197          1285.9             70.54            0
freemanTukey  207            43.4              9.76            0

t_B quantiles:
               0% 2.5% 25% 50% 75% 97.5% 100%
SSE           112  128 147 156 168   190  233
Chisq         752  794 864 904 946  1057 1419
freemanTukey  134  144 157 164 170   183  195

# Express the magnitude of lack of fit by an overdispersion factor
(c.hat <- pb.try2@t0[2] / mean(pb.try2@t.star[,2])) #  Chisq
2.411948
```

Once again we see the fit is pretty bad, even considering the more complex covariate structures, and the overdispersion ratio is only negligibly smaller. We could think about trying to improve on this by considering more complex covariate models. However, there may be excess Poisson variation that simply cannot be explained by the available covariates. For example, jays are not uniformly distributed, and there is some amount of aggregation that might be explained by overdispersion (i.e., unstructured additional noise to the Poisson variation). So next, we try fitting a negative binomial model using the `gdistamp` function. To do that we have to create a new `unmarkedFrame` using the `unmarkedFrameGDS` constructor function, which takes at a minimum one new argument called `numPrimary`, which is the number of sampling occasions within which it is reasonable to assume a closed population was sampled. In a normal distance sampling survey, we view the sampling as instantaneous and so we specify `numPrimary=1`. However, if we did a survey of the same points separated in time by days, weeks, or even years, then `numPrimary` would be the number of such temporal surveys. We don't discuss modeling temporal structure here (see Sections 9.5–9.7 in this and Chapter 14 in volume 2).

```
covs <- issj[,c("elevation", "forest", "chaparral")]
area <- pi*300^2 / 100^2              # Area in ha
```

```
jayumf <- unmarkedFrameGDS(y=as.matrix(issj[,1:3]),
   siteCovs=data.frame(covs, area), numPrimary=1,
   dist.breaks=c(0, 100, 200, 300),
   unitsIn="m", survey="point")
sc <- siteCovs(jayumf)
sc.s <- scale(sc)
sc.s[,"area"] <- pi*300^2 / 10000 # Don't standardize area
siteCovs(jayumf) <- sc.s
summary(jayumf)

# Fit the model using gdistsamp and look at the fit summary
(nb.C2E.C <- gdistsamp( ~chaparral + I(chaparral^2) + elevation +
   offset(log(area)), ~1, ~chaparral, data =jayumf, output="abund",
   mixture="NB", K = 150))

gdistsamp(lambdaformula = ~chaparral + I(chaparral^2) + elevation +
     offset(log(area)), phiformula = ~1, pformula = ~chaparral,
     data = jayumf, output = "abund", mixture = "NB", K = 150)

Abundance:
                Estimate    SE      z   P(>|z|)
(Intercept)       -2.516 0.198 -12.73  4.17e-37
chaparral          1.432 0.229   6.25  4.01e-10
I(chaparral^2)    -0.376 0.114  -3.28  1.04e-03
elevation         -0.227 0.146  -1.55  1.20e-01

Detection:
             Estimate     SE      z   P(>|z|)
(Intercept)     4.679  0.0658  71.14 0.000000
chaparral      -0.199  0.0600  -3.32 0.000905

Dispersion:                         # Note the NB dispersion parameter
 Estimate    SE     z    P(>|z|)    #  scale is log(tau)
    -1.02 0.215 -4.73  2.23e-06

AIC: 695.4445
```

This produces a long list of warnings of this sort:

```
42: In log(cp[J + 1]) : NaNs produced
```

These are related to having near 0 probability in the very last cell (individuals > 300 m away) and in general are not a problem. The size parameter is $\exp(-1.02) = 0.36$, which, as we noted above, is the τ parameter in the negative binomial distribution. The AIC of `gdistsamp` is not comparable to that of `distsamp`. If you run the same Poisson model using both functions, you get a different AIC! This is because the likelihood construction is completely different. However, we will use this model here to carry out some further analysis. We check the model fit using our parametric bootstrap procedure:

```
(pb.try3 <- parboot(nb.C2E.C, fitstats, nsim=1000, report=5))

Call: parboot(object = nb.C2E.C, statistic = fitstats, nsim = 1000, report = 5)
```

```
Parametric Bootstrap Statistics:
               t0  mean(t0 - t_B) StdDev(t0 - t_B) Pr(t_B > t0)
SSE           430           97.4             99.7     0.143856
Chisq        2200          868.1            161.2     0.000999
freemanTukey  211           20.8             20.9     0.159840

t_B quantiles:
               0% 2.5%  25%  50%  75% 97.5% 100%
SSE           117  188  265  315  379   576 1133
Chisq         921 1073 1225 1313 1425  1706 2556
freemanTukey 113  150  177  190  205   234  260

There were 50 or more warnings (use warnings() to see the first 50)

(c.hat <- pb.try3@t0[2] / mean(pb.try3@t.star[,2])) #
   Chisq
1.65186
```

This also produces many warnings of the previously mentioned variety. But, on the brighter side, this model does fit in a slightly more satisfactory way according to two out of three of our fit statistics. And, the overdispersion ratio is reduced by nearly 50% and so things appear to be more tolerable for this model. We could also produce predictions that are corrected for overdispersion as we did for the multinomial mixture models (see Section 7.9.4).

Next we use the results to produce an estimate of population size. We first define a function `getN`, which computes the sum of the predicted values for a given model object, and then we can apply it to any model we wish, and we can also use it as an input to the `parboot` function to produce uncertainty measures (SEs, CIs). For comparison, we also compute the predictions the old-fashioned way by constructing the model matrix and doing the linear algebra “by hand.” In addition, we compute the Best Unbiased Predictor (BUP) of local abundance.

```
# *Expected* population size for the sample points
getN <- function(fm, newdata=NULL)
   sum(predict(fm, type="lambda", newdata=newdata)[,1])

getN(nb.C2E.C)
[1] 889.6142

# This does the same thing as the following commands
X <- model.matrix(~chaparral+I(chaparral^2)+elevation+log(offset(area)),
       siteCovs(jayumf))
head(X) # The design matrix
  (Intercept)  chaparral I(chaparral^2)   elevation  log(offset(area))
1           1 -0.1218243     0.01484117 -1.20607849          3.341954
2           1  0.8384709     0.70303345 -0.36132054          3.341954
3           1  2.1319298     4.54512461 -0.45797193          3.341954
4           1 -0.2737078     0.07491594 -0.14196112          3.341954
5           1 -1.1562240     1.33685389 -0.72588125          3.341954
6           1  0.2989173     0.08935153  0.01704522          3.341954
```

```
# Prediction of total expected population size at the sample points
sum(exp(X %*% c(coef(nb.C2E.C, type="lambda"), 1)))
[1] 889.6142

# Empirical Bayes estimates of posterior distribution:
# Pr(N=x | y, lambda, sigma) for x=0,1,...,K
re.jay <- ranef(nb.C2E.C, K = 150)

# *Realized* population size
sum(bup(re.jay, "mean"))
[1] 827.4331
```

So there are about 889 ISSJs on the total area sampled by the 307 point counts, based on the fitted mean of the Poisson model. On the other hand, if we use the best unbiased predictor we have only about 827 ISSJs on the 307 point counts. In general, the two predictions should not be the same because the BUP "adjusts" the predictions toward the data (the observed counts) and so uses some additional information. The BUP is conditional on the particular sample at hand.

Next, we do two further summary analyses of the ISSJ models. First, we produce a graphical display of the effect of chaparral on expected local population size, and then we show a predictive map of expected density over the whole island (Figure 8.7).

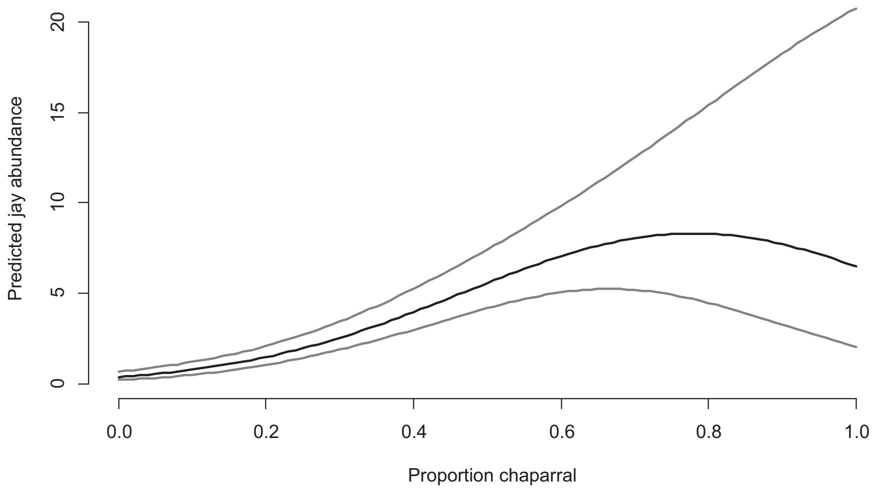

FIGURE 8.7

Response curve of the expected abundance, E(N), of island scrub-jays per 28 ha pixel to the covariate chaparral (with 95% CI limits).

```
summary(jayumf) # Note the range of chaparral which we need to know

[...output shortened...]

Site-level covariates:
   elevation           forest           chaparral          area
 Min.   :-1.4884  Min.   :-0.49215  Min.   :-1.1562  Min.   :28.27
 1st Qu.:-0.7974  1st Qu.:-0.49215  1st Qu.:-0.8721  1st Qu.:28.27
 Median :-0.1687  Median :-0.44295  Median :-0.2014  Median :28.27
 Mean   : 0.0000  Mean   : 0.00000  Mean   : 0.0000  Mean   :28.27
 3rd Qu.: 0.6650  3rd Qu.:-0.06982  3rd Qu.: 0.6872  3rd Qu.:28.27
 Max.   : 3.5731  Max.   : 5.42362  Max.   : 2.8809  Max.   :28.27

# Create a new data frame with area 28.27 ha, the area of a 300 m circle
chap.orig <- seq(0, 1, 0.01)   # Values from 0 to 1 prop. chaparral
chap.pred <- (chap.orig - mean(issj$chaparral)) / sd(issj$chaparral)
newdat <- data.frame(chaparral = chap.pred, elevation = 0, area=28.27)

# Expected values of N for covariate values in "newdat"
E.N <- predict(fall$C2E.C, type="state", newdata=newdat, appendData=TRUE)
head(E.N)
  Predicted        SE     lower     upper  chaparral elevation  area
1 0.3606945 0.1108475 0.1974927 0.6587612 -1.1562240         0 28.27
2 0.3907380 0.1163551 0.2179779 0.7004205 -1.1134578         0 28.27
3 0.4228468 0.1220244 0.2401844 0.7444256 -1.0706916         0 28.27
4 0.4571217 0.1278586 0.2642091 0.7908898 -1.0279254         0 28.27
5 0.4936646 0.1338616 0.2901490 0.8399297 -0.9851591         0 28.27
6 0.5325784 0.1400385 0.3181008 0.8916661 -0.9423929         0 28.27

# Make a plot of the response curve for the grid of chaparral values
plot(chap.orig, E.N[,"Predicted"], xlab="Proportion chaparral", ylab="Predicted jay
abundance", type="l", ylim = c(0, 20), frame = F, lwd = 2)
matlines(chap.orig, E.N[,3:4], lty = 1, col = "grey", lwd = 1)
```

Finally, now we take the habitat map for the whole island (Figure 8.6) and we predict the expected abundance, *E(N)*, on every pixel of the map. These pixels are 9 ha pixels instead of 28 ha sample units, and so we have to account for that area change. In addition, because models were fitted with standardized covariates, we need to appropriately standardize the landscape variables by exactly the same mean and SD used for the data in the analysis. To do that we first look at the attributes of the scaled site covariates, for which we computed several pages of R code previously:

```
attributes(sc.s) # means are "scaled:center". SDs are "scaled:scale"
$dim
[1] 307  4

$dimnames
$dimnames[[1]]
NULL
```

```
$dimnames[[2]]
[1] "elevation" "forest"  "chaparral" "area"

$'scaled:center'
  elevation     forest  chaparral       area
202.0023616  0.0673357  0.2703592 28.2743339

$'scaled:scale'
  elevation     forest  chaparral       area
124.8818069  0.1368199  0.2338295  0.0000000
```

And now we can apply these values of the mean and SD to the grid variables and then predict for each pixel of the Santa Cruz landscape.

```
cruz.s <- cruz    # Created a new data set for the scaled variables
cruz.s$elevation <- (cruz$elevation*0.3048-202)/125
cruz.s$chaparral <- (cruz$chaparral-0.270)/0.234
cruz.s$area <- (300*300)/10000 # The grid cells are 300x300m=9ha
EN <- predict(nb.C2E.C, type="lambda", newdata=cruz.s)

# Total population size (by summing predictions for all pixels)
getN(nb.C2E.C, newdata=cruz.s)
[1] 2282.039

# Parametric bootstrap for CI
# A much faster function could be written to doing the sum
set.seed(2015)
(EN.B <- parboot(nb.C2E.C, stat=getN, nsim=1000, report=5))

Call: parboot(object = nb.C2E.C, statistic = getN, nsim = 1000, report = 5)

Parametric Bootstrap Statistics:
   t0  mean(t0 - t_B) StdDev(t0 - t_B) Pr(t_B > t0)
1 890          -13.1              160        0.535

t_B quantiles:
     0% 2.5%  25%  50%   75% 97.5%  100%
t*1  481   607  788  903 1004  1237 1425

t0 = Original statistic compuated from data
t_B = Vector of bootstrap samples
```

So we have a population size estimate and a 95% confidence interval, and now let's make a map of the predictions (i.e., create a species distribution map in terms of the expected abundance). To do this we use the `raster` package to create a raster stack using the land cover variables (standardized) that are provided when the ISSJ data are loaded. Then we use the `predict` function with this raster stack.

```
library(raster)
cruz.raster <- stack(rasterFromXYZ(cruz.s[,c("x","y","elevation")]),
   rasterFromXYZ(cruz.s[,c("x","y","chaparral")]),
   rasterFromXYZ(cruz.s[,c("x","y","area")]))
```

```
names(cruz.raster) # These should match the names in the formula
[1] "elevation" "chaparral" "area"
plot(cruz.raster)           # not shown
# Elevation map on the original scale (not shown)
plot(cruz.raster[["elevation"]]*125 + 202, col=topo.colors(20),
main="Elevation (in feet) and Survey Locations", asp = 1)
points(issj[,c("x","y")], cex=0.8, pch = 16)
```

The `predict` function will use a raster stack having the appropriate covariates and produce raster output of predictions, SEs, and lower and upper confidence limits (Figure 8.8). Where else can you get such goodness from one function call? Here we show the Holy Grail of population ecology: a spatial

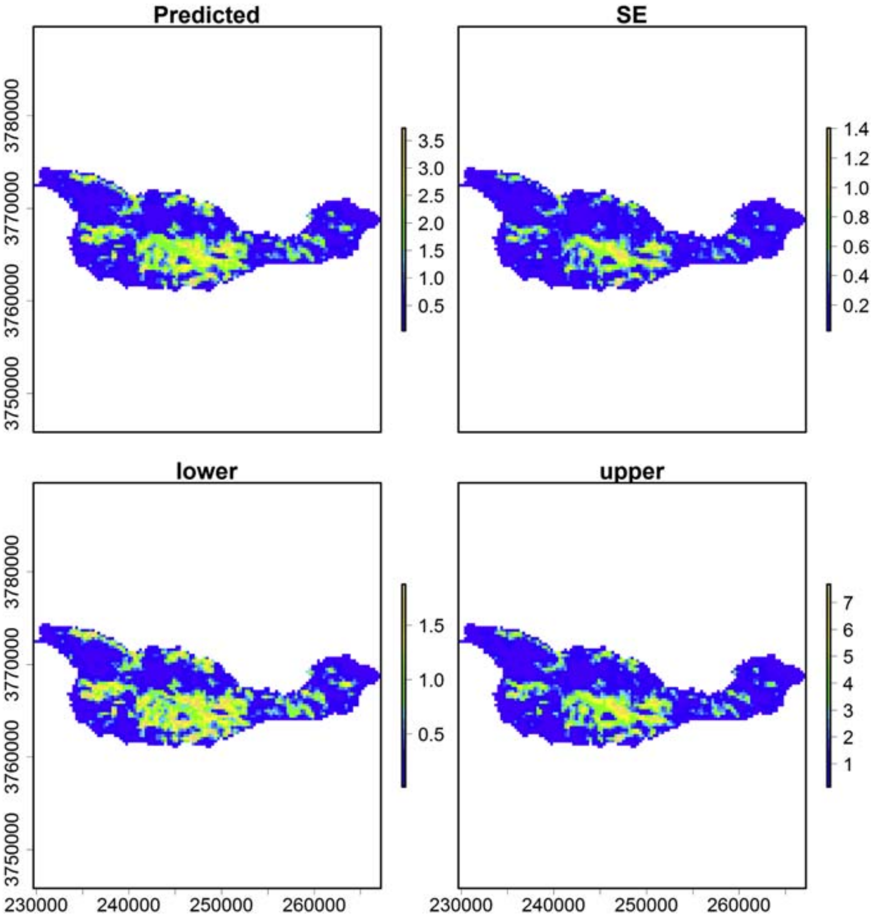

FIGURE 8.8

Global species distribution map of the island scrub-jay (*Aphelocoma insularis*), an endemic on the island of Santa Cruz, California, based on the best model in the model set. Predictions show the expected abundance (λ_s) under the AIC-best negative binomial model for every pixel. Three maps are shown to depict uncertainty in these predictions: the prediction SE and the lower and upper limit of a 95% prediction interval.

map depicting global population size predictions from the distance sampling model applied at the landscape scale:

```
EN.raster <- predict(nb.C2E.C, type="lambda", newdata=cruz.raster)
  doing row 1000 of 5625
  doing row 2000 of 5625
  doing row 3000 of 5625
  doing row 4000 of 5625
  doing row 5000 of 5625
plot(EN.raster, col = topo.colors(20), asp = 1)  # See Figure 8.8
```

While we conclude our analysis here merely by showing an estimate of the global population distribution of this important species, we note that the ultimate objective of this analysis was to use the model to make predictions of population size and distribution using the 1985 land cover of the island (pre-sheep cull; see Sillett et al., 2012, for more detail).

8.5 BAYESIAN HDS

Bayesian hierarchical distance sampling can be implemented in a number of different ways (following our developments of Section 8.3). There are two basic formulations that we demonstrate here: (1) The conditional (three-part) formulation of the model using either continuous or discrete data, which is similar to that which we outlined for the multinomial mixture model in Section 8.3. This three-part formulation of the model is similar to Chelgren et al. (2011b) and Shirk et al. (2014) and also similar to Hedley and Buckland (2004; although, they didn't do a joint estimation of the parameters from the different model components); and (2) formulation of the model for either discrete or continuous data using data augmentation. We should note before getting into the details that for some problems it might be perfectly reasonable to just pool all of the data and analyze one big data set having a single parameter *N*, the population size among all sampled populations. This may be reasonable to do if estimating overall abundance or mean density was the primary objective and the investigation of patterns in the variation among sample units was not important.

8.5.1 SIMULATING HDS DATA

We start by developing some familiarity with the data structure and processing by defining a function for simulating HDS data and fitting models to it. The function `simHDS` (with its default arguments shown) is called as follows:

```
simHDS(type="line", nsites = 100, mean.lambda = 2,
   beta.lam = 1, mean.sigma = 1, beta.sig = -0.5, B = 3, discard0=TRUE)
```

The function arguments mean the following:

- `type` lets you choose between either a line (`type = "line"`) or a point (`type = "point"`) transect protocol.
- `nsites` is the number of sites
- `alpha.lam` (= `log(mean.lambda)`) and `beta.lam` are the intercept and the slope of a log-linear regression of expected abundance per site on a habitat covariate
- `alpha.sig` (= `log(mean.sigma)`) and `beta.sig` are the intercept and the slope of a log-linear regression of scale parameter σ of the half-normal detection function on wind speed

- `B` is the strip half width
- `discard0=TRUE` subsets to sites at which >0 individuals were captured. You may or may not want to do this depending on how the model is formulated, so be careful.

Calling the function produces a visualization of the generated data set (see Figure 8.9 for type = “line” and Figure 8.10 for type = “point”).

By default we simulate line transect data for 100 sites, with abundance N_s for transect s having a Poisson distribution with a mean that depends on some simulated site covariate, “habitat.” We also incorporate an effect of another site-specific covariate, wind speed, which we assume affects the observation model via the detection function (specifically parameter σ; Marques et al., 2007). We now execute the function to obtain a point or a line transect data set (with default arguments).

```
set.seed(1234)
tmp1 <- simHDS("point")     # Point transect
tmp2 <- simHDS()            # Line transect (this is the default)
str(tmp1)                   # Look at function output
List of 14
 $ type       : chr "point"
 $ nsites     : num 100
 $ mean.lambda: num 2
 $ beta.lam   : num 1
 $ mean.sigma : num 1
 $ beta.sig   : num -0.5
 $ B          : num 3
 $ data    : num [1:76, 1:5] 2 3 6 13 21 22 24 29 31 31 ...
  ..- attr(*, "dimnames")=List of 2
  .. ..$ : NULL
  .. ..$ : chr [1:5] "" "y" "u" "v" ...
 $ B        : num 3
 $ nsites   : num 100
 $ habitat  : num [1:100] -1.207  0.277 1.084 -2.346 0.429 ...
 $ wind     : num [1:100] 0.643  0.113 -0.73 1.071  0.105 ...
 $ N        : int [1:100] 0  6  6  0  6  3 0  4  1 1 ...
 $ N.true   : int [1:100] 0  5  5  0  3  2 0  3  1 1 ...
```

Note that `N.true` is the number of animals with distance ≤ B, so for a line transect, `N = N.true`, while for a point transect, `N.true ≤ N` because we are simulating on a square.

Now we have a nice set of distance sampling data collected at 100 sites and with two site-specific covariates called `habitat` (affecting local abundance) and `wind` (affecting the distance out to which individuals are detected).

As always, it is extremely useful to play around with data simulation functions with changed arguments to train your intuition about a certain modeled process and also about the statistical model we use to make an inference about the parameters in this process; see Exercise 7.

8.5.2 BAYESIAN HDS USING DATA AUGMENTATION

We have discussed data augmentation (Royle et al., 2007a) several times in previous chapters, and we analyzed the distance sampling model using the Impala data with DA in Section 8.3 for both binned and continuous distance measurements. Here we apply these ideas to HDS models. In general, every capture-recapture model can be analyzed using data augmentation. And, distance sampling can be regarded as

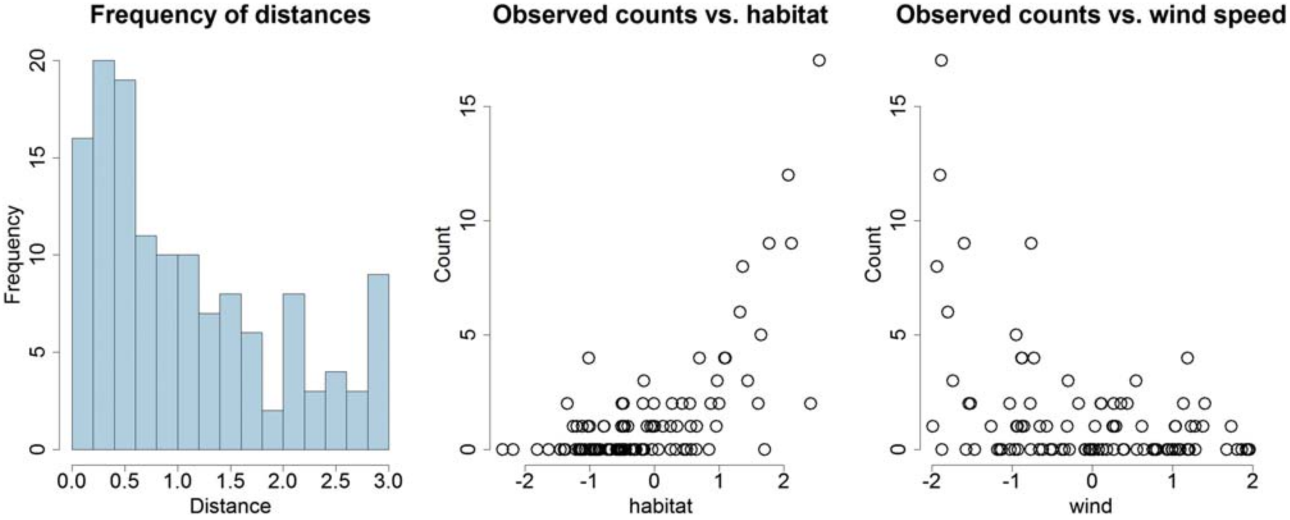

FIGURE 8.9

Visualization produced when the function `simHDS` is run for line transects. Histogram shows *observed* distances.

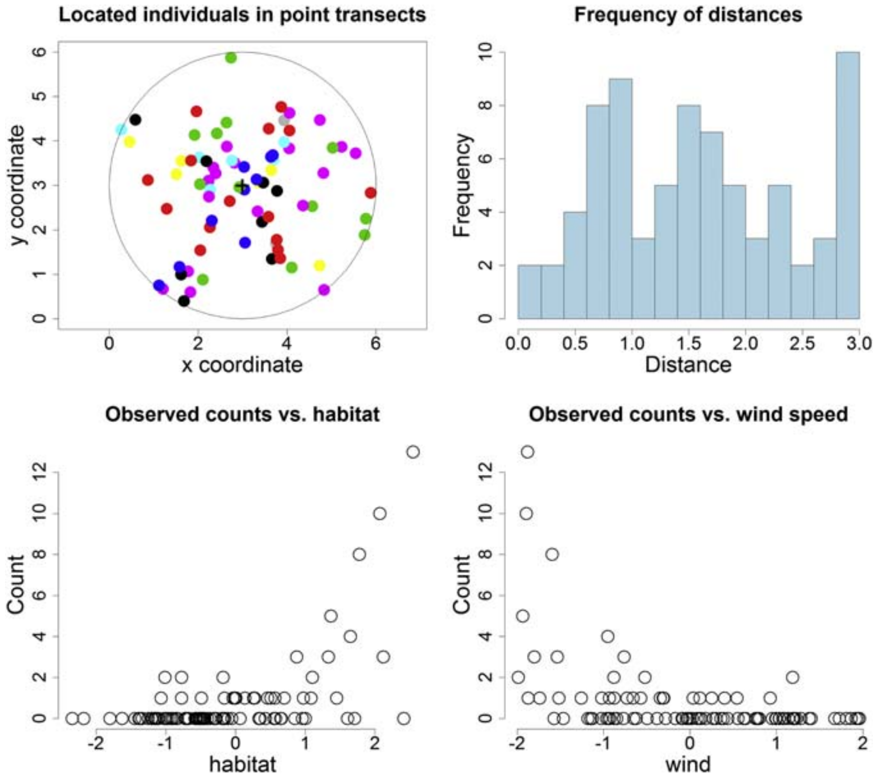

FIGURE 8.10

Visualization produced when the function `simHDS` is run for point transects. Colors in the top-left panel denote different sites. Histogram shows *observed* distances.

just another capture-recapture model with an individual covariate “distance,” which affects p (also without the “recapture”). To apply DA to distance sampling, we augment a data set of captured individuals ($y = 1$) with a large number of uncaptured individuals (i.e., $y = 0$), and we recognize that the resulting “augmented data set” is a zero-inflated version of the known-N data set (similar to the relationship between an occupancy model and a logistic regression of detections at sites that are known to be occupied). That is, some of the added zeros are sampling zeros and some of them are structural zeros (fixed zeros that are not missed individuals). We can express this zero-inflated Bernoulli model directly in BUGS. It essentially recasts the capture-recapture (here distance sampling) model as a site occupancy model (see Chapter 10). In distance sampling, the only nuance is that p_i depends on distance and the distance “data” must be input as missing values for the augmented individuals, i.e., we have a site-occupancy model with a partially missing, site-level covariate, which are estimated as part of the model.

Now for site-structured data, i.e., HDS, where we have distance sampling data from a number S of sites, we have to consider how to get the multisite structure integrated into this DA formulation of the model. Such a framework for accomplishing this was described in Converse and Royle (2012), Royle et al. (2012), and Royle and Converse (2014). The main idea is simply to pool the data into one big data set having rows $i = 1,2,\ldots,M$, which includes the n observed individuals (from among all sites) and a large number, $M - n$, of unobserved individuals. In addition, we add an additional individual covariate, which is the site membership of each individual, say $site_i$. The observed site membership of individuals is treated as a *categorical individual covariate*, which, for the augmented individuals, is missing data that can be estimated. In this formulation of the model we have both sites and individuals (as in Section 7.8.4) and so we need to be careful with our indexing, which will be slightly different from other places in the book. As in Chapter 7, we will use the index i for individual and $s = 1,2,\ldots,S$ for sites. We require a prior distribution for the categorical individual covariate, which we specify as follows in the BUGS language:

```
site[i] ~ dcat( site.probs[]),
```

where `site.probs` is a vector of length S defined as

`site.probs[s]=` $\frac{\lambda_s}{\sum\limits_s \lambda_s}$.

Here, λ_s is the expected abundance at site s, which may of course depend on site-specific covariates. This model is *implied* by the assumption that

$$N_s \sim Poisson(\lambda_s)$$

with

$$\log(\lambda_s) = \beta_0 + \beta_1 x_s.$$

The data augmentation part of the model includes a set of latent variables z_i which are Bernoulli trials taking on the value $z_i = 1$ when an *individual* observation corresponds to a real individual and $z_i = 0$ when it corresponds to a structural zero. As before, we assume $z_i \sim Bernoulli(\psi)$, where ψ is the data augmentation parameter. **One extremely important caveat** is that in this use of DA for such site-structured or stratified models, the intercept parameter β_0 of the abundance model is confounded with the DA parameter ψ (Royle and Converse 2014). They are equivalent parameters, and the model must be fitted by imposing a constraint, either by setting $\beta_0 = 0$ or setting $\psi = \sum_{s=1}^{S} \lambda_s / M$, where M is the total number of individuals in the augmented data set (see Royle et al., 2014, p. 314).

To summarize, we can do hierarchical distance sampling by: (1) including individual "site membership" as a categorical covariate; (2) specifying the site membership covariate as a categorical random variable with cell probabilities proportional to the Poisson mean parameter λ_s; and (3) specifying covariate effects directly on λ_s as in our usual binomial and multinomial N-mixture models.

For illustration, we also show how covariates can be introduced into the detection model, which we typically do by specifying a linear model for $log(\sigma)$. We assume that we had measured a covariate `wind` speed, which affects detection probability via its effect on the *detection function*. Having measured wind speed at the time of our survey at every site, it is a site covariate (if we had temporal replicate measurements at a site, it would be an *observational covariate*). Thus, the observation part of our HDS model for every individual i in the augmented data set is this (note that sites are now indexed s and the scale parameter σ varies by site according to the log-linear model on wind speed):

$$y_i \sim Bernoulli(p_i)$$

$$p_i = \exp\left(- d_{\mathrm{i}} * d_{\mathrm{i}} \Big/ \left(2 * \sigma^2_{s(i)}\right)\right)$$

$$\log(\sigma_s) = \alpha_0 + \alpha_1 * wind_s$$

Note that the first line specifies a relationship for individuals, the second one between individuals and an individual-specific covariate (d_i) and a site-specific parameter σ_s, and the third line a relationship purely at the site level. In line two, we emphasize the relationship between individual and site by use of the double subscript in $\sigma_{s(i)}$, which specifies membership of individual i to site s.

We now demonstrate using our simulated data from 100 line transect surveys.

```
# Recreate line transect data set
set.seed(1234)
tmp <- simHDS()          # Line transect (default)
attach(tmp)

# Data augmentation: add a bunch of "pseudo-individuals"
nz <- 500                          # Augment by 500
nind <- nrow(data)
y <- c(data[,2], rep(0, nz))       # Augmented detection indicator y
site <- c(data[,1], rep(NA, nz))   # Augmented site indicator,
                                   #    unknown (i.e., NA) for augmented inds.
d <- c(data[,5], rep(NA,nz))       # Augmented distance data (with NAs)

# Bundle and summarize data set
str( win.data <- list(nsites=nsites, habitat=habitat, wind=wind, B=B, nind=nind, nz=nz,
y=y, d=d, site=site) )
win.data$site                      # Unknown site cov. for augmented inds.

# BUGS model for line transect HDS (NOT point transects!)
cat("
model{
# Prior distributions
beta0 ~ dunif(-10,10)    # Intercept of lambda-habitat regression
beta1 ~ dunif(-10,10)    # Slope of log(lambda) on habitat
alpha0 ~ dunif(-10,10)   # Intercept of log(sigma) (half-normal scale)
alpha1 ~ dunif(-10,10)   # Slope of log(sigma) on wind

# psi is a derived parameter under DA for stratified populations
psi <- sum(lambda[]) / (nind+nz)

# 'Likelihood' (sort of...)
for(i in 1:(nind+nz)){                        # i is index for individuals
   z[i] ~ dbern(psi)                          # Data augmentation variables
   d[i] ~ dunif(0, B)                         # Distance uniformly distributed
   p[i] <- exp(-d[i]*d[i]/(2*sigma[site[i]]*sigma[site[i]])) # Det. function
   mu[i] <- z[i]* p[i]                        # 'straw man' for WinBUGS
   y[i] ~ dbern(mu[i])                        #  Bernoulli random variable
   site[i] ~ dcat(site.probs[1:nsites]) # Population distribution among sites
}

# Linear models for abundance and for detection
for(s in 1:nsites){                              # s is index for sites
  # Model for abundance
  # next line not necessary, but allows to make predictions
  N[s] ~ dpois(lambda[s])                        # Realized abundance at site s
  log(lambda[s]) <- beta0 + beta1*habitat[s] # Linear model abundance
  site.probs[s] <- lambda[s] / sum(lambda[])

 # Linear model for detection
 log(sigma[s]) <- alpha0 + alpha1*wind[s]
}
```

```
# Derived parameters: total population size and average density across all sites
Ntotal <- sum(z[])
area <- nsites*1*2*B                            # Unit length == 1, half-width = B
D <- Ntotal/area
}
",fill=TRUE , file = "model1.txt")

# Inits
zst <- c(rep(1, sum(y)), rep(0, nz)) # ... and for DA variables
inits <- function(){list(beta0=0, beta1=0, alpha0=0, alpha1=0, z=zst)}

# Parameters to save
params <- c("alpha0", "alpha1", "beta0", "beta1", "psi", "Ntotal", "D")

# MCMC settings
ni <- 12000  ;  nb <- 2000  ;  nt <- 2  ;  nc <- 3

# Call BUGS (ART 33 min) ...
bd <- "c:/Program Files/WinBUGS14/" # Never forget this for WinBUGS
out1 <- bugs(win.data, inits, params, "model1.txt", n.thin=nt,
  n.chains=nc, n.burnin=nb, n.iter=ni, debug=TRUE, bugs.dir = bd)

# ... or try JAGS for a change (ART 6 min)
library(jagsUI)    # never forget to load jagsUI
out1 <- jags(win.data, inits, params, "model1.txt", n.thin=nt,
   n.chains=nc, n.burnin=nb, n.iter=ni)
```

We note that JAGS completes the analysis about five times faster than WinBUGS (and would be faster still with argument `parallel = TRUE`). This is a good example of why testing things out in both BUGS engines can be helpful and efficient (you might also want to try out OpenBUGS).

```
# Summarize posterior output
print(out1, 2)
              mean    sd    2.5%     50%   97.5% overlap0    f  Rhat n.eff
alpha0        0.02  0.11   -0.17    0.02    0.28     TRUE 0.56     1 14232
alpha1       -0.66  0.10   -0.88   -0.65   -0.48    FALSE 1.00     1  4213
beta0         0.58  0.15    0.29    0.58    0.85    FALSE 1.00     1  1818
beta1         0.94  0.08    0.79    0.94    1.10    FALSE 1.00     1  1077
psi           0.41  0.04    0.33    0.40    0.50    FALSE 1.00     1  9092
Ntotal      259.11 24.40  215.00  258.00  310.00    FALSE 1.00     1 15000
D             0.43  0.04    0.36    0.43    0.52    FALSE 1.00     1 15000

# Truth in data simulation (note alpha0 and beta0 are log transformed)
 $ mean.lambda: num 2                 # exp(beta0) above
 $ beta.lam   : num 1                 # beta1
 $ mean.sigma : num 1                 # exp(alpha0)
 $ beta.sig   : num -0.5              # alpha1
```

The posterior means of all parameters are not too far from their data-generating values. In addition, the posterior mean of `Ntotal` is 259.11 (CRI 215–310), which agrees quite well with the true total population size obtained by summing N_s over all S sites (note: only 136 individuals were detected).

```
sum(tmp$N.true)
[1] 305
```

Next, we consider the same analysis, but as if we had binned distance data. For this model we use the categorical distribution in BUGS. While there is nothing technically novel here, note that when we have the covariate on σ (`wind`) we must define a large matrix of site and distance-class specific detection probabilities, and this will slow things down substantially. This takes a few extra lines of code to make use of categorical data. We use the same simulated data set from earlier in this Section (which you may have to re-create in case you wrote over things in your R workspace). We first convert the distance data into categorical distance classes as we've done before, and then write out the BUGS model into a file and set things up for a BUGS run.

```
# Prepare data
delta <- 0.1                           # width of distance bins for approx.
midpt <- seq(delta/2, B, delta)        # make mid-points and chop up data
dclass <- d %/% delta + 1              # convert distances to cat. distances
nD <- length(midpt)                    # Number of distance intervals

# Bundle and summarize data set
str( win.data <- list (y=y, dclass=dclass, site=site, midpt=midpt, delta=delta, B=B,
nind=nind, nz=nz, nsites=nsites, nD=nD, habitat=habitat, wind=wind) )

# BUGS model specification for line-transect HDS (NOT point transects!)
cat("
model{
# Prior distributions
alpha0 ~ dunif(-10,10)
alpha1 ~ dunif(-10,10)
beta0  ~ dunif(-10,10)
beta1  ~ dunif(-10,10)

psi <- sum(lambda[])/(nind+nz)         # psi is a derived parameter

for(i in 1:(nind+nz)){                 # Loop over individuals
  z[i] ~ dbern(psi)                    # DA variables
  dclass[i] ~ dcat(pi[site[i],])       # Population distribution of dist class
  mu[i] <- z[i] * p[site[i],dclass[i]]  # p depends on site AND dist class
  y[i] ~ dbern(mu[i])                  # Basic Bernoulli response in DS model
  site[i] ~ dcat(site.probs[1:nsites]) # Site membership of inds
}

for(s in 1:nsites){                    # Loop over sites
# Construct cell probabilities for nD cells
for(g in 1:nD){                        # midpt = mid point of each cell
  log(p[s,g]) <- -midpt[g]*midpt[g]/(2*sigma[s]*sigma[s])
  pi[s,g] <- delta/B                   # Probability of x per interval
  f[s,g] <- p[s,g]*pi[s,g]             # pdf of observed distances
}

  # not necessary  N[s]~dpois(lambda[s]) except for prediction
  N[s] ~ dpois(lambda[s])              # Predict abundance at each site
```

```
    log(lambda[s]) <- beta0 + beta1 * habitat[s] # Linear model for N
    site.probs[s] <- lambda[s]/sum(lambda[])
    log(sigma[s]) <- alpha0 + alpha1*wind[s] # Linear model for sigma
  }

  # Derived parameters: total abundance and mean density across all sites
  Ntotal <- sum(z[])       # Also sum(N[]) which is size of a new population
  area <- nsites*1*2*B     # Unit length == 1, half-width = B
  D <- Ntotal/area
  }
  ",fill=TRUE, file = "model2.txt")

  # Inits
  zst <- c(rep(1, sum(y)), rep(0, nz))
  inits <- function(){list (alpha0=0, alpha1=0, beta0=0, beta1=0, z=zst) }

  # Params to save
  params <- c("alpha0", "alpha1", "beta0", "beta1", "psi", "Ntotal","D")

  # MCMC settings
  ni <- 12000  ;  nb <- 2000  ;  nt <- 2  ;  nc <- 3

  # Run JAGS with parallel processing (ART 1 min)
  library(jagsUI)
  out2 <- jags(win.data, inits, params, "model2.txt", n.thin=nt,
     n.chains=nc, n.burnin=nb, n.iter=ni, parallel = FALSE)
```

We summarize the posterior samples for each parameter (output truncated), the means of which, as before, are similar to the data-generating values.

```
print(out2,2)
          mean    sd    2.5%     50%   97.5% overlap0    f Rhat n.eff
alpha0    0.02  0.11   -0.17    0.01    0.26     TRUE 0.55 1.00   838
alpha1   -0.66  0.10   -0.88   -0.65   -0.48    FALSE 1.00 1.00  6469
beta0     0.57  0.14    0.29    0.58    0.85    FALSE 1.00 1.01   266
beta1     0.94  0.08    0.79    0.94    1.11    FALSE 1.00 1.01   224
psi       0.41  0.04    0.33    0.40    0.49    FALSE 1.00 1.00   959
Ntotal  258.89 23.45  217.00  258.00  309.00    FALSE 1.00 1.00   834
D         0.43  0.04    0.36    0.43    0.52    FALSE 1.00 1.00   834
```

The results are, not surprisingly, numerically very similar to the analysis based on data augmentation of the previous section.

8.5.3 BAYESIAN HDS USING THE THREE-PART CONDITIONAL MULTINOMIAL MODEL

It is possible to specify HDS models in BUGS without using data augmentation. In Section 7.6 we discussed a specific formulation of multinomial models for implementation in BUGS to get around our inability to specify a random variable as a multinomial index. We will use that same formulation here. The basic idea is that we deconstruct the multinomial observation model by first conditioning on n_s for each site s, so that instead of having a multinomial/Poisson mixture model, we have the three-part multinomial/binomial/Poisson mixture model as follows:

$$y_s | n_s \sim Multinomial(n_s, \boldsymbol{\pi}_s^c) \tag{1}$$

where $\pi_k^c = \pi_k/(1 - \pi_0)$, the index k here representing the kth element of the vector $\boldsymbol{\pi}_s^c$,

$$n_s | N_s \sim Binomial(N_s, 1 - \pi_0) \tag{2}$$

$$N_s \sim Poisson(\lambda_s) \tag{3}$$

The first component is the model for distance class of the observed n_s individuals, the second describes imperfect detection of the N_s individuals leading to count n_s, and the third is our usual model for spatial variation in local abundance N_s. The key thing is that the multinomial of the first component has index n_s, which is observed, and so we're conditioning on observed data, not on the latent variable N_s. This three-part hierarchical model is easily implemented in BUGS. Chelgren et al. (2011b) may have been the first to do this, but they also formulated the model for continuous distance point count data using the "zeros trick" in BUGS. We prefer to use a fine binning if a nearly-continuous model is desired. In practice, we also usually specify the first stage in BUGS using a categorical observation model for individual observations, instead of the multinomial model for the distance class frequencies (as described above). This is a much more versatile formulation, which allows considerable flexibility to expand the model (see Chapter 9).

We simulate another data set using the same function as before, but this time *do not discard* the data from the sites where no animals were detected (`discard0=F`) so that those sites are carried as observed zeros in the data set and we can analyze the zero-filled data.

```
# Simulate line transect data set
set.seed(1234)
tmp <- simHDS(type="line", discard0=FALSE)
attach(tmp)

# Get number of individuals detected per site
# ncap = 1 plus number of detected individuals per site
ncap <- table(data[,1])                # ncap = 1 if no individuals captured
sites0 <- data[is.na(data[,2]),][,1]   # sites where nothing detected
ncap[as.character(sites0)] <- 0        # Fill in 0 for sites with no detections
ncap <- as.vector(ncap)

# Prepare other data
site <- data[!is.na(data[,2]),1]       # site ID of each observation
delta <- 0.1                           # distance bin width for rect. approx.
midpt <- seq(delta/2, B, delta)        # make mid-points and chop up data
dclass <- data[,5] %/% delta + 1       # convert distances to cat. distances
nD <- length(midpt)                    # Number of distance intervals
dclass <- dclass[!is.na(data[,2])]     # Observed categorical observations
nind <- length(dclass)                 # Total number of individuals detected

# Bundle and summarize data set
str( win.data <- list(nsites=nsites, nind=nind, B=B, nD=nD, midpt=midpt, delta=delta,
ncap=ncap, habitat=habitat, wind=wind, dclass=dclass, site=site) )
```

```
# BUGS model specification for line-transect HDS (NOT point transects!)
cat("
model{
# Priors
alpha0 ~ dunif(-10,10)
alpha1 ~ dunif(-10,10)
beta0 ~ dunif(-10,10)
beta1 ~ dunif(-10,10)

for(i in 1:nind){
   dclass[i] ~ dcat(fc[site[i],]) # Part 1 of HM
}

for(s in 1:nsites){
# Construct cell probabilities for nD multinomial cells
  for(g in 1:nD){                  # midpt = mid-point of each cell
    log(p[s,g]) <- -midpt[g] * midpt[g] / (2*sigma[s]*sigma[s])
    pi[s,g] <- delta / B           # Probability per interval
    f[s,g] <- p[s,g] * pi[s,g]
    fc[s,g] <- f[s,g] / pcap[s]
  }
 pcap[s] <- sum(f[s,])             # Pr(capture): sum of rectangular areas
 ncap[s] ~ dbin(pcap[s], N[s])     # Part 2 of HM
 N[s] ~ dpois(lambda[s])           # Part 3 of HM
 log(lambda[s]) <- beta0 + beta1 * habitat[s] # Linear model abundance
 log(sigma[s])<- alpha0 + alpha1*wind[s]      # Linear model detection
}
# Derived parameters
Ntotal <- sum(N[])
area <- nsites*1*2*B  # Unit length == 1, half-width = B
D <- Ntotal/area
}
",fill=TRUE, file = "model3.txt")

# Inits
Nst <- ncap + 1
inits <- function(){list(alpha0=0, alpha1=0, beta0=0, beta1=0, N=Nst)}

# Params to save
params <- c("alpha0", "alpha1", "beta0", "beta1", "Ntotal","D")

# MCMC settings
ni <- 12000  ;  nb <- 2000  ;  nt <- 1  ;  nc <- 3

# Run JAGS (ART 1 min) and summarize posteriors
library(jagsUI)
out3 <- jags(win.data, inits, params, "model3.txt", n.thin=nt,
   n.chains=nc, n.burnin=nb, n.iter=ni)
print(out3, 2)
```

```
          mean    sd   2.5%    50%  97.5% overlap0    f Rhat n.eff
alpha0    0.00  0.11  -0.19  -0.01   0.23     TRUE 0.54    1 11116
alpha1   -0.66  0.10  -0.87  -0.66  -0.49    FALSE 1.00    1  4359
beta0     0.64  0.14   0.35   0.64   0.90    FALSE 1.00    1  4330
beta1     0.91  0.08   0.76   0.91   1.07    FALSE 1.00    1 10802
Ntotal  266.98 24.52 223.00 266.00 318.00    FALSE 1.00    1  3085
D         0.44  0.04   0.37   0.44   0.53    FALSE 1.00    1  3085
```

This is not too bad (true $N = 305$) and the chains seem to be mixing well, so we're satisfied that this formulation of the model is viable. We could work up the three-part model for continuous distance data, but, as in Chelgren et al. (2011b) and Shirk et al. (2014), we would have to use the ones or zeros trick to implement this (but see Section 9.8.1 for an alternative formulation that avoids having to use the ones or zeros trick). However, there would be no practical difference between doing that and just using a large number of relatively narrow distance intervals (remember our treatise on integration in Section 8.3.2).

8.5.4 POINT TRANSECT HDS USING THE CONDITIONAL MULTINOMIAL FORMULATION

We now show how to apply the three-part conditional multinomial model to point transect data. As before, we simulate a data set but now keep the nondetection sites in the data set so that when we process the data we have a record of sites with $n = 0$ (where no individual was detected), and be sure to specify `type = "point"`.

```
# Simulate a point count data set using our simHDS function
set.seed(1234)
tmp <- simHDS(type="point", discard0=FALSE)
attach(tmp)

# Prepare data
# Number of individuals detected per site
ncap <- table(data[,1])                  # ncap = 1 if no individuals captured
sites0 <- data[is.na(data[,2]),][,1]     # Sites where nothing was seen
ncap[as.character(sites0)] <- 0          # Fill in 0 for sites with no detections
ncap <- as.vector(ncap)                  # Number of individuals detected per site

# Other data
site <- data[!is.na(data[,2]),1]         # Site ID of each observation
delta <- 0.1                             # Distance bin width for rect. approx.
midpt <- seq(delta/2, B, delta)          # Make mid-points and chop up data
dclass <- data[,5] %/% delta + 1         # Convert distance to distance category
nD <- length(midpt)                      # Number of distance intervals
dclass <- dclass[!is.na(data[,2])]       # Observed categorical observations
nind <- length(dclass)                   # Total number of individuals detected

# Bundle and summarize data set
str( win.data <- list(nsites=nsites, nind=nind, B=B, nD=nD, midpt=midpt,
delta=delta, ncap=ncap, habitat=habitat, wind=wind, dclass=dclass,site=site) )
```

```
# BUGS model specification for point transect data
cat("
model{
# Priors
alpha0 ~ dunif(-10,10)
alpha1 ~ dunif(-10,10)
beta0 ~ dunif(-10,10)
beta1 ~ dunif(-10,10)

for(i in 1:nind){
  dclass[i] ~ dcat(fc[site[i],]) # Part 1 of HM
}
for(s in 1:nsites){
   # Construct cell probabilities for nD distance bands
  for(g in 1:nD){                   # midpt = mid-point of each band
    log(p[s,g]) <- -midpt[g] * midpt[g] / (2 * sigma[s] * sigma[s])
    pi[s,g] <- ((2 * midpt[g] ) / (B * B)) * delta # prob. per interval
    f[s,g] <- p[s,g] * pi[s,g]
    fc[s,g] <- f[s,g] / pcap[s]
 }
 pcap[s] <- sum(f[s,])              # Pr(capture): sum of rectangular areas

 ncap[s] ~ dbin(pcap[s], N[s])      # Part 2 of HM
 N[s] ~ dpois(lambda[s])            # Part 3 of HM
 log(lambda[s]) <- beta0 + beta1  * habitat[s] # Linear model abundance
 log(sigma[s]) <- alpha0 + alpha1*wind[s]      # Linear model detection
}

# Derived parameters
Ntotal <- sum(N[])
area <- nsites*3.141*B*B
D <- Ntotal/area
}
",fill=TRUE, file="model4.txt")

# Inits
Nst <- ncap + 1
inits <- function(){list(alpha0=0, alpha1=0, beta0=0, beta1=0, N=Nst)}

# Params to save
params <- c("alpha0", "alpha1", "beta0", "beta1", "Ntotal","D")

# MCMC settings
ni <- 12000  ;  nb <- 2000  ;  nt <- 1  ;  nc <- 3

# Run BUGS (ART 2.3 min) and summarize posteriors
out4 <- bugs(win.data, inits, params, "model4.txt", n.thin=nt,
   n.chains=nc, n.burnin=nb, n.iter=ni, debug=TRUE, bugs.dir = bd)

print(out4, 2)
Inference for Bugs model at "model4.txt", fit using WinBUGS,
 3 chains, each with 12000 iterations (first 2000 discarded)
  n.sims = 30000 iterations saved
```

```
           mean    sd   2.5%    25%    50%    75%  97.5% Rhat n.eff
alpha0    -0.06  0.11  -0.26  -0.13  -0.06   0.01   0.18    1  7500
alpha1    -0.56  0.11  -0.80  -0.62  -0.55  -0.49  -0.38    1 30000
beta0      0.41  0.24  -0.09   0.24   0.41   0.57   0.87    1  3100
beta1      0.97  0.11   0.76   0.90   0.97   1.05   1.19    1  4800
Ntotal  227.92 40.85 159.00 198.00 224.00 253.00 317.00    1  4900
D          0.08  0.01   0.06   0.07   0.08   0.09   0.11    1  4900
```

Once again we obtain results very similar to the data generating values – we have only one individual difference between the posterior mean and the actual realized population size (and note that only 76 individuals were detected, so this estimation problem was not trivial). Of course density is estimated much lower in this case because the sample units are circles of radius $B = 3$ and not rectangles as before, and therefore the total area is much larger.

```
sum(tmp$N.true)                # True realized population size
[1] 227
sum(!is.na(tmp$data[,"y"]))    # Observed index for population size (Johnson, 2008)
[1] 76
```

8.5.5 BAYESIAN HDS ANALYSIS OF THE ISSJ DATA

Finally, we illustrate an application of HDS using the ISSJ data. We could use either the data augmentation or the three-part model formulations of HDS for the ISSJ data. We'll use the three-part model formulation here and leave as an exercise for you to figure out how to implement the DA version of the model. From Section 8.4.3 we found that a negative binomial abundance model with quadratic effect of chaparral and linear effect of elevation on mean abundance, and a linear effect of chaparral on the distance function parameter σ was our preferred model, and provided a reasonable fit to the data. So we fit a model that is similar here, but minus the chaparral effect on σ (we have you do this as an exercise). And, instead of a negative binomial abundance model, we will illustrate the fitting of a Poisson lognormal model of overdispersion, where we add a site effect with standard deviation σ_{site} to the linear predictor of the expected abundance. First, we have to convert the vector of frequencies for each site to individual distance class observations.

```
# Load the ISSJ data
library(unmarked)
data(issj)

# Prepare some data
nD <- 3                      # Number of intervals
delta <- 100                 # Interval width
B  <- 300                    # Upper bound (max. distance)
midpt <- c(50, 150, 250)     # mid points

# Convert vector frequencies to individual distance class
H <- as.matrix(issj[,1:3])
nsites <- nrow(H)
ncap <- apply(H, 1, sum)     # Number of individuals detected per site
dclass <- rep(col(H), H)     # Distance class of each individual
```

```
nind <- length(dclass)                              # Number of individuals detected
elevation <- as.vector(scale(issj[,c("elevation")])) # Prepare covariates
forest <- as.vector(scale(issj[,"forest"]))
chaparral <- as.vector(scale(issj[,"chaparral"]))

# Bundle and summarize data set
str( win.data <- list(nsites=nsites, nind=nind, B=B, nD=nD, midpt=midpt,delta=delta,
ncap=ncap, chaparral=chaparral, elevation=elevation, dclass=dclass) )

# BUGS model specification
cat("
model{
# Priors
sigma ~ dunif(0,1000)
beta0 ~ dunif(-10,10)
beta1 ~ dunif(-10,10)
beta2 ~ dunif(-10,10)
beta3 ~ dunif(-10,10)
sigma.site ~ dunif(0,10)
tau <- 1/(sigma.site*sigma.site)
# Specify hierarchical model
for(i in 1:nind){
    dclass[i] ~ dcat(fc[])                  # Part 1 of HM
}

# Construct cell probabilities for nD cells
for(g in 1:nD){                             # midpt = mid-point of each cell
  log(p[g]) <- -midpt[g] * midpt[g] / (2 * sigma * sigma)
  pi[g] <- ((2 * midpt[g]) / (B * B)) * delta # prob. per interval
  f[g] <- p[g] * pi[g]
  fc[g] <- f[g] / pcap
}
pcap <- sum(f[])                            # Pr(capture): sum of rectangular areas
for(s in 1:nsites){
  ncap[s] ~ dbin(pcap, N[s])                # Part 2 of HM
  N[s] ~ dpois(lambda[s])                   # Part 3 of HM
   log(lambda[s]) <- beta0 + beta1*elevation[s] + beta2*chaparral[s] +
   beta3*chaparral[s]*chaparral[s] + site.eff[s]
                                            # Linear model for abundance
 site.eff[s] ~ dnorm(0, tau)                # Site log normal 'residuals'
}
# Derived params
Ntotal <- sum(N[])
area <- nsites*3.141*300*300/10000          # Total area sampled, ha
D <- Ntotal/area
}
",fill=TRUE, file="model5.txt")

# Inits
Nst <- ncap + 1
inits <- function(){list (sigma = runif(1, 30, 100), beta0 = 0, beta1 = 0, beta2 = 0,
beta3 = 0, N = Nst, sigma.site = 0.2)}
```

```
# Params to save
params <- c("sigma", "beta0", "beta1", "beta2", "beta3", "sigma.site", "Ntotal","D")

# MCMC settings
ni <- 52000  ;  nb <- 2000  ;  nt <- 2  ;  nc <- 3
```

When we run WinBUGS, we may sometimes get an undefined real result error, which appears to be related to a bad choice of initial values for `sigma`. Simply try again until the algorithm works (also remember to define the object `bd`, which gives the WinBUGS Windows address)....

```
# Run BUGS (ART 0.9 min) and summarize posteriors
 out5 <- bugs(win.data, inits, params, "model5.txt", n.thin=nt,
    n.chains=nc, n.burnin=nb, n.iter=ni, debug=TRUE, bugs.dir = bd)
```

… or else you can run JAGS, which we never observed to crash for numerical over/underflow. However, you may get the `'Observed node inconsistent with unobserved parent'` error if you use a prior for `sigma` that is too far away from the bulk of the posterior mass.

```
out5 <- jags(win.data, inits, params, "model5.txt", n.thin=nt,
   n.chains=nc, n.burnin=nb, n.iter=ni)

# Run JAGS (ART 0.5 min) and summarize posteriors
print(out5, 3)
             mean    sd   2.5%    50%  97.5% overlap0    f  Rhat  n.eff
sigma      102.31  4.69  93.75 102.08 112.13    FALSE 1.00     1   9031
beta0       -0.10  0.25  -0.62  -0.09   0.36     TRUE 0.65     1   6307
beta1       -0.24  0.16  -0.56  -0.24   0.07     TRUE 0.94     1   1405
beta2        1.20  0.23   0.77   1.20   1.67    FALSE 1.00     1   3031
beta3       -0.52  0.15  -0.83  -0.52  -0.24    FALSE 1.00     1   1224
sigma.site   1.53  0.18   1.20   1.52   1.92    FALSE 1.00     1   2859
Ntotal     664.84 71.48 535.00 661.00 814.00    FALSE 1.00     1   4885
D            0.08  0.01   0.06   0.08   0.09    FALSE 1.00     1   4885
```

This is only one model, not exactly the best model from our analysis with `unmarked`, but the estimated effects are reasonably consistent. Under our Poisson lognormal (PLN) model for abundance we obtain an estimated `Ntotal` which is a bit less than the BUP (of about 827) from `unmarked`. Since we fitted the negative binomial model in the latter, which is similar but not identical to the PLN, some difference might be expected there.

8.6 SUMMARY

Distance sampling is an extremely important methodology in wildlife ecology and management. There is a very well worked-out theory and a truly huge body of literature (e.g., Buckland et al., 2001, 2004a), along with comprehensive software (Distance), R packages, user groups etc.; see, e.g., Thomas et al. (2010) and `distancesampling.org`. In this chapter, we did of course not aim to summarize all of distance sampling, but rather, we gave an introductory view of how distance sampling fits into the concept of hierarchical models, as we present them in this book. We gave a broad overview of various formulations of the distance sampling model, including conditional and full likelihood, point and line transects, and likelihood and Bayesian analysis. There are many formulations of distance sampling! We showed this vast scope of distance sampling implementations in BUGS because each one may be advantageous in a given situation.

Almost all studies that employ distance sampling collect data on multiple spatial sample units, either line transects or point counts (sometimes called point transects in the distance sampling literature). We call models that formally account for the variation in abundance *N* or in density among sample units hierarchical distance sampling (HDS) models, and we discussed analysis of HDS models in `unmarked` and in the BUGS language.

The R package `unmarked` implements models for data that are recorded or summarized into discrete distance intervals so that the resulting models are a variation of the multinomial *N*-mixture models of Chapter 7. We have argued in this chapter that there is no practical reason to favor continuous distance data and models over distance interval (i.e., binned) data models in principle, since a discrete distance model is no more or less an approximation to truth than is a continuous distance model. Both are just alternative approximations to the true but unknown detection function. And, they may or may not be close approximations to each other depending on the continuous model being approximated and the discreteness of the distance bins. If you really must have a continuous model, you can achieve this without any appreciable loss of precision simply by using many narrow distance bins.

We have introduced two general formulations of HDS in BUGS that can be useful: the conditional multinomial or “three-part” hierarchical formulation of the model, and the formulation based on data augmentation. The use of DA gives us an individual formulation of the model, and so this formulation should be advantageous for situations where we need to model individual covariates such as sex or other characteristics or when the observation unit is a group or cluster of individuals. We discuss some more advanced HDS models in the next chapter and then in chapters 14 and 24 in volume 2. One challenge to implementing point transect HDS models in BUGS is that nonstandard distributions are involved. In this case, we use binned data and approximate the cell probabilities using rectangular approximations to the area under the curve. An alternative would be to use the zeros or ones trick in BUGS (Chelgren et al., 2011b; Shirk et al., 2014).

Hierarchical distance sampling is a relatively recent advance that shows great promise to addressing fundamental problems related to the modeling of spatial variation in abundance or density. The HDS framework is not only flexible but, because it is so easy to implement in the BUGS language, ecologists can easily extend the ideas we have provided here to solve their own problems. Moreover, `unmarked` contains novel HDS modeling capabilities in a user-friendly and standardized analysis framework and will likely contain additional capabilities in the future. Given the accessibility of HDS models and the importance of modeling spatial variation in ecology, we think HDS models will become the de facto standard for the analysis of distance sampling data in the near future. HDS models are so important that we address several additional extensions in the next chapter (and more in volume 2!).

EXERCISES

1. The estimator of density as a function of the conditional MLE of $\overline{p}$ is: $\widehat{D} = n/(\widehat{\overline{p}} * L * 2 * B)$. Using basic statistical arguments, what is the variance of this estimator?
2. “Prove” by simulation that if individuals are randomly distributed on a rectangle then the distance to the center line has a uniform distribution. “Prove” by simulation that if individuals are randomly distributed about a point in space then the distance to the point has a triangular distribution. We put quotes around the verb ‘to prove’ because of course this does not represent a mathematical proof which you may also be able to derive.

3. Play around with the data simulation function (i.e., vary *N* and *sigma*) for line transects (this is function `sim.ldata` in Section 8.2.3) to train your intuition about line transect distance sampling.
4. Play around with the data simulation function for point transects (function `sim.pdata` in Section 8.2.5) to train your intuition about point transect distance sampling.
5. In Section 8.3.4, run a simulation study to "prove" that the two estimators (conditional likelihood and full likelihood) are about unbiased in the frequentist sense of the word.
6. In Section 8.4.3 (the ISSJ analysis using `unmarked`) see by how much the density estimate would be biased if the habitat effect of chaparral on p had NOT been taken account of. And by how much would we have erred in our global population estimate (assuming that the model in the section is the correct one of course) by assuming that chaparral does not affect the measurement error of density?
7. Play around with the data simulation function for hierarchical line and point transect sampling (function `simHDS` in Section 8.5.1) to train your intuition about HDS. Vary everything you can, especially, type, number of sites, average abundance, average half-normal scale parameter, and also strip half width (B).
8. Implement the HDS model using data augmentation for the ISSJ data (we used the three-part hierarchical model in Section 8.5.5).
9. For the ISSJ data in BUGS using the three-part model (Section 8.5.5), figure out how to model an effect of chaparral on the detection scale parameter σ.
10. For the analysis developed in question (9), modify the model to have a negative binomial abundance model instead of a Poisson-lognormal abundance model, and compare the inferences under the two alternatives of "overdispersed Poisson" models.

References

Aarts, G., Fieberg, J., Matthiopoulos, J., 2012. Comparative interpretation of count, presence-absence and point methods for species distribution models. Methods Ecol. Evol. 3, 177−187.

Abadi, F., Gimenez, O., Arlettaz, R., Schaub, M., 2010a. An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. Ecology 91, 7−14.

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., Schaub, M., 2010b. Estimation of immigration rate using integrated population modeling. J. Appl. Ecol. 47, 393−400.

Adams, M.J., Chelgren, N.D., Reinitz, D., Cole, R.A., Rachowicz, L.J., Galvan, S., McCreary, B., Pearl, C.A., Bailey, L.L., Bettaso, J., Bull, E.L., Leu, M., 2010. Using occupancy models to understand the distribution of an amphibian pathogen, *Batrachochytrium dendrobatidis*. Ecol. Appl. 20, 289−302.

Aing, C., Halls, S., Oken, K., Dobrow, R., Fieberg, J., 2011. A Bayesian hierarchical occupancy model for track surveys conducted in a series of linear, spatially correlated, sites. J. Appl. Ecol. 48, 1508−1517.

Aitkin, M., 1991. Posterior bayes factors. J. R. Stat. Soc. Series B (Methodol.) 53, 111−142.

Alldredge, M.W., Pollock, K.H., Simons, T.R., Collazo, J.A., Shriner, S.A., 2007. Time-of-detection method for estimating abundance from point-count surveys. Auk 124, 653−664.

Alpizar-Jara, R., Pollock, K.H., 1996. A combination line transect and capture-recapture sampling model for multiple observers in aerial surveys. Environ. Ecol. Stat. 3 (4), 311−327.

Amstrup, S.C., McDonald, T.L., Manly, B.F.J. (Eds.), 2005. Handbook of Capture-recapture Analysis. Princeton University Press, Princeton NJ.

Amundson, C.L., Royle, J.A., Handel, C.M., 2014. A hierarchical model combining distance sampling and time removal to estimate detection probability during avian point counts. Auk 131, 476−494.

Andrewartha, H.G., Birch, L.C., 1954. The Distribution and Abundance of Animals. University of Chicago Press, Chicago.

Arnason, A.N., 1972. Parameter estimates from mark-recapture experiments on two populations subject to migration and death. Res. Pop. Ecol. 13, 97−113.

Arnason, A.N., Schwarz, C.J., 1999. Using POPAN-5 to analyse banding data. Bird Study 46, 157−168.

Augustin, N.H., Mugglestone, M.A., Buckland, S.T., 1996. An autologistic model for the spatial distribution of wildlife. J. Appl. Ecol. 33, 339−347.

Baddeley, A., Berman, M., Fisher, N.I., Hardegen, A., Milne, R.K., Schuhmacher, D., Shah, R., Turner, R., 2010. Spatial logistic regression and change-of-support in Poisson point processes. Electron. J. Stat. 4, 1151−1201.

Baddeley, A., Turner, R., 2005. Spatstat: an R package for analyzing spatial point patterns. J. Stat. Softw. 12, 1−42.

Baillie, S.R., 1991. Integrated population monitoring of breeding birds in Britain and Irland. Ibis 132, 151−166.

Bailey, L.L., Hines, J.E., Nichols, J.D., MacKenzie, D.I., 2007. Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. Ecol. Appl. 17, 281−290.

Bailey, L.L., MacKenzie, D.I., Nichols, J.D., 2014. Advances and applications of occupancy models. Methods Ecol. Evol. 5, 1269−1279.

Bailey, L.L., Simons, T.R., Pollock, K.H., 2004. Estimating site occupancy and species detection probability parameters for terrestrial salamanders. Ecol. Appl. 14, 692−702.

Balmford, A., Green, R.E., Jenkins, M., 2003. Measuring the changing state of nature. Trends Ecol. Evol. 18, 326−330.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall/CRC Press.

Banks-Leite, C., Pardini, R., Boscolo, D., Cassano, C.R., Püttker, T., Barros, C.S., Barlow, J., 2014. Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. J. Appl. Ecol. 51, 849–859.

Bardos, D.C., Guillera-Arroita, G., Wintle, B.A., 2015. Valid auto-models for spatially autocorrelated occupancy and abundance data. Methods Ecol. Evol. http://dx.doi.org/10.1111/2041-210X.12402.

Barker, R.J., 1997. Joint modeling of live-recapture, tag-resight, and tag-recovery data. Biometrics 53, 666–677.

Barnagaud, J.Y., Barbaro, L., Papaix, J., Deconchat, M., Brockerhoff, E., 2014. Habitat filtering by landscape and local forest composition in native and exotic New Zealand birds. Ecology 95, 78–87.

Barry, S.C., Welsh, A.H., 2001. Distance sampling methodology. J. R. Stat. Soc. Series B 63, 31–53.

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2014. Fitting Linear Mixed-Effects Models Using lme4. http://arxiv.org/abs/1406.5823.

Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances. Phil. Trans. R. Soc. A 53, 370–418.

Beesley, L.S., Gwinn, D.C., Price, A., King, A.J., Gawne, B., Koehn, J.D., Nielsen, D.L., 2014. Juvenile fish response to wetland inundation: how antecedent conditions can inform environmental flow policies for native fish. J. Appl. Ecol. 51, 1613–1621.

Begon, M., Harper, J.L., Townsend, C.R., 1986. Ecology: Individuals, Populations and Communities. Blackwell, Oxford.

Beissinger, S.R., 2002. Population viability analysis: past, present, future. In: Beissinger, S.R. (Ed.), Population Viability Analysis. The University of Chicago Press, Chicago, pp. 5–17.

Bellier, E., M. Kéry, M. Schaub. Dynamic N-mixture models with density–dependence in vital rates. In review.

Berliner, L.M., 1996. Hierarchical Bayesian time series models. In: Hanson, K., Silver, R. (Eds.), Maximum Entropy and Bayesian Methods. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 15–22.

Besbeas, P., Freeman, S.N., Morgan, B.J.T., Catchpole, E.A., 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. Biometrics 58, 540–547.

Bibby, C.J., Burgess, N.D., Hill, D.A., Mustoe, S., 2000. Bird Census Techniques, second ed. Academic Press.

Bled, F., Royle, J.A., Cam, E., 2011a. Assessing hypotheses about nesting site occupancy dynamics. Ecology 92, 938–951.

Bled, F., Royle, J.A., Cam, E., 2011b. Hierarchical modeling of an invasive spread: case of the Eurasian collared dove *Streptopelia decaocto* in the USA. Ecol. Appl. 21, 290–302.

Bolker, B.M., 2008. Ecological Models and Data in R. Princeton University Press, Princeton, New Jersey.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24, 127–135.

Bonner, S.J., Schwarz, C.J., 2006. An extension of the Cormack-Jolly-Seber model for continuous covariates with application to *Microtus pennsylvanicus*. Biometrics 62, 142–149.

Borchers, D.L., Buckland, S.T., Zucchini, W., 2002. Estimating Animal Abundance. Springer, London.

Borchers, D.L., Efford, M.G., 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. Biometrics 64, 377–385.

Borchers, D.L., Stevenson, B.C., Kidney, D., Thomas, L., Marques, T.A., 2015. A unifying model for capture–recapture and distance sampling surveys of wildlife populations. J. Am. Stat. Assoc. 110, 195–204.

Borchers, D.L., Zucchini, W., Fewster, R.M., 1998. Mark-recapture models for line transect surveys. Biometrics 1207–1220.

Bornand, C.N., Kéry, M., Bueche, L., Fischer, M., 2014. Hide-and-seek in vegetation: time-to-detection is an efficient design for estimating detectability and occurrence. Methods Ecol. Evol. 5, 433–442.

Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E., Pollock, K.H., 1998. Estimating species richness: the importance of heterogeneity in species detectability. Ecology 79, 1018–1028.

Boyce, M.S., 2010. Presence-only data, pseudo-absences, and other lies about habitat selection. Ideas Ecol. Evol. 3, 26−27.

Broms, K.M., Johnson, D.S., Altwegg, R., Conquest, L.L., 2014. Spatial occupancy models applied to atlas data show Southern Ground Hornbills strongly depend on protected areas. Ecol. Appl. 24, 363−374.

Brooks, S.P., Catchpole, E.A., Morgan, B.J.T., 2000. Bayesian animal survival estimation. Stat. Sci. 15, 357−376.

Brooks, S.P., Gelman, A., 1998. Alternative methods for monitoring convergence of iterative simulations. J. Comput. Gr. Stat. 7, 434−455.

Brooks, S.P., King, R., Morgan, B.J.T., 2004. A Bayesian approach to combining animal abundance and demographic data. Anim. Biodivers. Conserv. 27, 515−529.

Brown, J.H., Maurer, B.A., 1989. Macroecology: the division of food and space among species on continents. Science 243, 1145−1150.

Brownie, C., Anderson, D.R., Burnham, K.P., Robson, D.S., 1985. Statistical Inference from Band Recovery Data - a Handbook. In: US Fish and Wildlife Service, vol. 156. Resource Publication, Washington DC.

Brownie, C., Hines, J.E., Nichols, J.D., Pollock, K.H., Hestbeck, J.B., 1993. Capture-recapture studies for multiple strata including non-Markovian transitions. Biometrics 49, 1173−1187.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L., 2001. Introduction to Distance Sampling. Oxford University Press, Oxford.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L. (Eds.), 2004a. Advanced Distance Sampling. Oxford University Press, Oxford, 414 pp.

Buckland, S.T., Newman, K.B., Fernandez, C., Thomas, L., Harwood, J., 2007. Embedding population dynamics models in inference. Stat. Sci. 22, 44−58.

Buckland, S.T., Newman, K.B., Thomas, L., Koesters, N.B., 2004b. State-space models for the dynamics of wild animal populations. Ecol. Model. 171, 157−175.

Burnham, K.P., 1993. A theory for combined analysis of ring recovery and recapture data. In: Lebreton, J.D. (Ed.), Marked Individuals in the Study of Bird Populations. Birkhäuser, Basel, pp. 199−213.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information Theoretic Approach. Springer, New York.

Burnham, K.P., Anderson, D.R., Laake, J.L., 1980. Estimation of density from line transect sampling of biological populations. Wildl. Monogr. 72, 3−202.

Burton, A., Sam, M., Balangtaa, C., Brashares, J., 2012. Hierarchical multi-species modeling of carnivore responses to hunting, habitat and prey in a West African protected area. PLoS ONE 7 (5), e38007.

Cam, E., Link, W.A., Cooch, E.G., Monnat, J.Y., Danchin, E., 2002a. Individual covariation in life-history traits: seeing the trees despite the forest. Am. Nat. 159, 96−105.

Cam, E., Nichols, J.D., Hines, J.E., Sauer, J.R., Alpizar-Jara, R., Flather, C.H., 2002b. Disentangling sampling and ecological explanations underlying species-area relationships. Ecology 83, 1118−1130.

Cam, E., Nichols, J.D., Sauer, J.R., Hines, J.E., 2002c. On the estimation of species richness based on the accumulation of previously unrecorded species. Ecography 25, 102−108.

Carlin, B.P., Chib, S., 1995. Bayesian model choice via Markov chain Monte Carlo methods. J. R. Stat. Soc. Series B (Methodol.) 57, 473−484.

Carlin, B.P., Louis, T.A., 2009. Bayesian Methods for Data Analysis. CRC Press/Taylor & Francis Group, Boca Raton.

Carrillo-Rubio, E., Kéry, M., Morreale, S.J., Sullivan, P.J., Gardner, B., Cooch, E.G., Lassoie, J.P., 2014. Use of multispecies occupancy models to evaluate the response of bird communities to forest degradation associated with logging. Conserv. Biol. 28, 1034−1044.

Casella, G., Berger, R.L., 2002. Statistical inference (Vol. 2). Duxbury, Pacific Grove, CA.

Caswell, H., 2001. Matrix Population Models. Construction, Analysis, and Interpretation. Sinauer Associates, Sunderland, Massachusetts.

Catchpole, E.A., Kgosi, P.M., Morgan, B.J.T., 2001. On the near-singularity of models for animal recovery data. Biometrics 57, 720–726.

Catchpole, E.A., Morgan, B.J.T., 1997. Detecting parameter redundancy. Biometrika 84, 187–196.

Catchpole, E.A., Morgan, B.J.T., Viallefont, A., 2002. Solving problems in parameter redundancy using computer algebra. J. Appl. Stat. 29, 625–636.

Caughley, G., 1974. Bias in aerial survey. J. Wildl. Manage. 38, 921–933.

Chambert, T., Miller, D.A.W., Nichols, J.D., 2015. Modeling false positive detections in species occurrence data under different study designs. Ecology 96, 332–339.

Chandler, R.B., 2015. Modeling Variation in Abundance Using Capture-Recapture Data. Available at: cran.r-project.org/web/packages/unmarked/vignettes/cap-recap.pdf.

Chandler, R.B., Clark, J.D., 2014. Spatially explicit integrated population models. Methods Ecol. Evol. 5, 1351–1360.

Chandler, R.B., King, D.I., 2011. Habitat quality and habitat selection of golden-winged warbler in Costa Rica: application of hierarchical models for open populations. J. Appl. Ecol. 48, 1038–1047.

Chandler, R.B., King, D.I., Chandler, C.C., 2009a. Effects of management regime on the abundance and nest survival of shrubland birds in wildlife openings in northern New England, USA. For. Ecol. Manage. 258, 1669–1676.

Chandler, R.B., King, D.I., DeStefano, S., 2009b. Scrub-shrub bird habitat associations at multiple spatial scales in beaver meadows in Massachusetts. Auk 126, 186–197.

Chandler, R.B., King, D.I., Raudales, R., Trubey, R., Chandler, C., Chavez, V.J.A., 2013. A small-scale land-sparing approach to conserving biological diversity in tropical agricultural landscapes. Conserv. Biol. 27, 785–795.

Chandler, R.B., Royle, J.A., 2013. Spatially explicit models for inference about density in unmarked or partially marked populations. Ann. Appl. Stat. 7, 936–954.

Chandler, R.B., Royle, J.A., King, D.I., 2011. Inference about density and temporary emigration in unmarked populations. Ecology 92, 1429–1435.

Chelgren, N.D., Adams, M.J., Bailey, L.L., Bury, R.B., 2011a. Using multilevel spatial models to understand salamander site occupancy patterns after wildfire. Ecology 92, 408–421.

Chelgren, N.D., Samora, B., Adams, M.J., McCreary, B., 2011b. Using spatiotemporal models and distance sampling to map the space use and abundance of newly metamorphosed Western toads (*Anaxyrus boreas*). Herpetol. Conserv. Biol. 6, 175–190.

Chen, G., Kéry, M., Plattner, M., Ma, K., Gardner, B., 2013. Imperfect detection is the rule rather than the exception in plant distribution studies. J. Ecol. 101, 183–191.

Chen, G., Kéry, M., Zhang, J., Ma, K., 2009. Factors affecting detection probability in plant distribution studies. J. Ecol. 97, 1383–1389.

Choquet, R., Cole, D., 2012. A hybrid symbolic-numerical method for determining model structure. Math. Biosci. 236, 117–125.

Choquet, R., Lebreton, J.D., Gimenez, O., Reboulet, A.M., Pradel, R., 2009a. U-CARE: utilities for performing goodness of fit tests and manipulating CApture-REcapture data. Ecography 32, 1071–1074.

Choquet, R., Rouan, L., Pradel, R., 2009b. Program E-SURGE: a software application for fitting multievent models. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations. Springer, New York, pp. 845–865.

Clark, J.S., 2007. Ecological Data Models with R. Princeton University Press, Princeton, NJ.

Clark, J.S., Gelfand, A.E. (Eds.), 2006. Hierarchical Modelling for the Environmental Sciences. Oxford University Press, Oxford, England.

Clark, J.S., Gelfand, A.E., Woodall, C.W., Zhu, K., 2014. More than the sum of the parts: forest climate response from joint species distribution models. Ecol. Appl. 24, 990–999.

Cole, D.J., 2012. Determining parameter redundancy of multi-state mark-recapture models for sea birds. J. Ornithol. 152 (Suppl. 2), 305−315.

Collier, B.A., Groce, J.E., Morrison, M.L., Newnam, J.C., Campomizzi, A.J., Farrell, S.L., Mathewson, H.A., Snelgrove, R.T., Carroll, R.J., Wilkins, R.N., 2012. Predicting patch occupancy in fragmented landscapes at the rangewide scale for an endangered species: an example of an American warbler. Divers. Distrib. 18, 158−167.

Conn, P.B., Laake, J.L., Johnson, D.S., 2012. A hierarchical modeling framework for multiple observer transect surveys. PLoS ONE 7, e42294.

Conroy, M.J., Runge, J.P., Barker, R.J., Schofield, M.R., Fonnesbeck, C.J., 2008. Efficient estimation of abundance for patchily distributed populations via two-phase, adaptive sampling. Ecology 89, 3362−3370.

Converse, S.J., Royle, J.A., 2012. Dealing with Incomplete and Variable Detectability in Multi-year, Multi-site Monitoring of Ecological Populations. Design and Analysis of Long-term Ecological Monitoring Studies. Cambridge University Press, Cambridge, England, United Kingdom, pp. 426−442.

Cooch, E., White, G., 2014. Program MARK: A Gentle Introduction. Available in pdf format for free download at http://www.phidot.org/software/mark/docs/book.

Cook, R.D., Jacobson, J.O., 1979. A design for estimating visibility bias in aerial surveys. Biometrics 735−742.

Cormack, R.M., 1964. Estimates of survival from the sighting of marked animals. Biometrika 51, 429−438.

Coull, B.A., Agresti, A., 1999. The use of mixed logit models to reflect heterogeneity in capture-recapture studies. Biometrics 55, 294−301.

Couturier, T., Cheylan, M., Bertolero, A., Astruc, G., Besnard, A., 2013. Estimating abundance and population trends when detection is low and highly variable: a comparison of three methods for the Hermann’s tortoise. J. Wildl. Manage. 77, 454−462.

Crainiceanu, C.M., Ruppert, D., Wand, M.P., 2005. Bayesian analysis for penalized spline regression using WinBUGS. J. Stat. Softw. 14, 1−25.

Crewe, T.L., Taylor, P.D., Lepage, D., 2015. Modeling systematic change in stopover duration does not improve bias in trends estimated from migration counts. PLoS ONE 10 (6), e0130137. http://dx.doi.org/10.1371/journal.pone.0130137.

Cressie, N., Calder, C.A., Clark, J.S., Ver Hoef, J.M., Wikle, C.K., 2009. Accounting for uncertaity in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecol. Appl. 19, 553−570.

Cressie, N., Wikle, C.K., 2011. Statistics for Spatio-Temporal Data. Wiley.

Cribari-Neto, F., Zeileis, A., 2010. Beta regression in R. J. Stat. Softw. 34, 1−24. http://www.jstatsoft.org/v34/i02/.

Dail, D., Madsen, L., 2011. Models for estimating abundance from repeated counts of an open population. Biometrics 67, 577−587.

Dellaportas, P., Forster, J.J., Ntzoufras, I., 2002. On Bayesian model and variable selection using MCMC. Stat. Comput. 12, 27−36.

Dénes, F.V., Silveira, L.F., Beissinger, S.R., 2015. Estimating abundance of unmarked animal populations accounting for imperfect detection and other sources of zero inflation. Methods Ecol. Evol. 6, 543−556.

Dennis, B., 1996. Discussion: should ecologists become Bayesians? Ecol. Appl. 6, 1095−1103.

Dennis, E.B., Morgan, B.J.T., Freeman, S.N., Brereton, T., Roy, D.B., 2015c. A Generalised Abundance Index for Seasonal Invertebrates. Technical report UKC/SMSAS/14/006. University of Kent.

Dennis, E.B., Morgan, B.J.T., Freeman, S.N., Roy, D.B., Brereton, T., 2015b. Dynamic models for longitudinal butterfly data. J. Agric Biol. Environ. Stat. http://dx.doi.org/10.1007/s13253-015-0216-3.

Dennis, E.B., Morgan, B.J.T., Ridout, M.S., 2015a. Computational aspects of N-mixture models. Biometrics 71, 237−246.

de Valpine, P., 2003. Better inferences from population-dynamics experiments using Monte Carlo state-space likelihood methods. Ecology 84, 3064−3077.

de Valpine, P., 2009. Shared challenges and common ground for Bayesian and classical analysis of hierarchical statistical models. Ecol. Appl. 19, 584−588.

de Valpine, P., Turek, D., Paciorek, C.J., Anderson-Bergman, C., Temple Lang, D., Bodik, R. Programming with models: writing statistical algorithms for general model structures with NIMBLE. J. Comput. Gr. Stat. (in review).

DeWan, A.A., Zipkin, E.F., 2010. An integrated sampling and analysis approach for improved biodiversity monitoring. Environ. Manage. 45, 1223−1230.

Denwood, M.J., 2015. runjags: An R package providing interface utilities, parallel computing methods and additional distributions for MCMC models in JAGS. J. Stat. Softw. http://cran.r-project.org/web/packages/runjags/.

Dice, L., 1938. Some census methods for mammals. J. Wildl. Manage. 2, 119−130.

Diefenbach, D.R., Marshall, M.R., Mattice, J.A., Brauning, D.W., Johnson, D.H., 2007. Incorporating availability for detection in estimates of bird abundance. Auk 124, 96−106.

Dixon, P.M., 2006. Bootstrap resampling. Encycl. Environmetrics (Online, Wiley). http://dx.doi.org/10.1002/9780470057339.vab028.

Dobson, A., Barnett, A., 2008. An Introduction to Generalized Linear Models. CRC/Chapmann & Hall, Boca Raton.

Dodd, C.K., Dorazio, R.M., 2004. Using counts to simultaneously estimate abundance and detection probabilities in salamander surveys. Herpetologica 60, 468−478.

Doherty Jr., P.F., Sorci, G., Royle, J.A., Hines, J.E., Nichols, J.D., Boulinier, T., 2003. Sexual selection affects local extinction and turnover in bird communities. Proc. Nat. Acad. Sci. 100, 5858−5862.

Dorazio, R.M., 2007. On the choice of statistical models for estimating occurrence and extinction from animal surveys. Ecology 88, 2773−2782.

Dorazio, R.M., 2012. Predicting the geographic distribution of a species from presence-only data subject to detection errors. Biometrics 68, 1303−1312.

Dorazio, R.M., 2013. Bayes and empirical Bayes estimators of abundance and density from spatial capture-recapture data. PLoS ONE 8, e84017.

Dorazio, R.M., 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. Glob. Ecol. Biogeogr. 23, 1472−1484.

Dorazio, R.M., Connor, E.F., 2014. Estimating abundances of interacting species using morphological traits, foraging guilds, and habitat. PLoS ONE 9, e94323.

Dorazio, R.M., Connor, E.F., Askins, R.A., 2015. Estimating the effects of habitat and biological interactions in an avian community. PLoS ONE 10, e0135987. http://dx.doi.org/10.1371/journal.pone.0135987.

Dorazio, R.M., Gotelli, N.J., Ellison, A.M., 2011. Modern methods of estimating biodiversity from presence-absence surveys. In: Grillo, O., Venora, G. (Eds.), Biodiversity Loss in a Changing Planet. InTech, ISBN 978-953-307-707-9.

Dorazio, R.M., Jelks, H.L., Jordan, F., 2005. Improving removal-based estimates of abundance by sampling a population of spatially distinct subpopulations. Biometrics 61, 1093−1101.

Dorazio, R.M., Kéry, M., Royle, J.A., Plattner, M., 2010. Models for inference in dynamic metacommunity systems. Ecology 91, 2466−2475.

Dorazio, R.M., Martin, J., Edwards, H.H., 2013. Estimating abundance while accounting for rarity, correlated behavior, and other sources of variation in counts. Ecology 94, 1472−1478.

Dorazio, R.M., Mukherjee, B., Zhang, L., Ghosh, M., Jelks, H.L., Jordan, F., 2008. Modeling unobserved sources of heterogeneity in animal abundance using a Dirichlet process prior. Biometrics 64, 635−644.

Dorazio, R.M., Rodriguez, D.T., 2012. A Gibbs sampler for Bayesian analysis of site-occupancy data. Methods Ecol. Evol. 3, 1093−1098.

Dorazio, R.M., Royle, J.A., 2003. Mixture models for estimating the size of a closed population when capture rates vary among individuals. Biometrics 59, 351−364.

Dorazio, R.M., Royle, J.A., 2005. Estimating size and composition of biological communities by modeling the occurrence of species. J. Am. Stat. Assoc. 100, 389–398.

Dorazio, R.M., Royle, J.A., Söderström, B., Glimskär, A., 2006. Estimating species richness and accumulation by modeling species occurrence and detectability. Ecology 87, 842–854.

Dupuis, J.A., 1995. Bayesian estimation of movement and survival probabilities from capture-recapture data. Biometrika 82, 761–772.

Dupuis, J.A., Bled, F., Joachim, J., 2011. Estimating the occupancy rate of spatially rare or hard to detect species: a conditional approach. Biometrics 67, 290–298.

Edwards, A.W., 1992. Likelihood, expanded ed. Johns Hopkins, Baltimore.

Efford, M., 2004. Density estimation in live-trapping studies. Oikos 106, 598–610.

Efford, M.G., Borchers, D.L., Mowat, G., 2013. Varying effort in capture-recapture studies. Methods Ecol. Evol. 4, 629–636.

Efford, M.G., Dawson, D.K., 2009. Effect of distance-related heterogeneity on population size estimates from point counts. Auk 126, 100–111.

Efford, M.G., Dawson, D.K., 2012. Occupancy in continuous habitat. EcoSphere 3, 1–12. Article 12.

Efford, M.G., Fewster, R.M., 2013. Estimating population size by spatially explicit capture-recapture. Oikos 122, 918–928.

Efron, B., 1982. The jackknife, the bootstrap and other resampling plans, Vol. 38. Society for industrial and applied mathematics, Philadelphia.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall/CRC Press.

Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697.

Ellis, M.M., Ivan, J.S., Tucker, J.M., Schwartz, M.K., 2015. rSPACE: spatially based power analysis for conservation and ecology. Methods Ecol. Evol. 6, 621–625.

Etterson, M.A., Niemi, G.J., Danz, N.P., 2009. Estimating the effects of detection heterogeneity and overdispersion on trends estimated from avian point counts. Ecol. Appl. 19, 2049–2066.

Farnsworth, G.L., Nichols, J.D., Sauer, J.R., Fancy, S.G., Pollock, K.H., Shriner, S.A., Simons, T.R., 2005. Statistical Approaches to the Analysis of Point Count Data: A Little Extra Information Can Go a Long Way. USDA Forest Service. Gen. Tech. Rep. PSW-GTR-191.

Farnsworth, G.L., Pollock, K.H., Nichols, J.D., Simons, T.R., Hines, J.E., Sauer, J.R., 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119, 414–425.

Ferraz, G., Nichols, J.D., Hines, J.E., Stouffer, P.C., Bierregaard Jr., R.O., Lovejoy, T.E., 2007. A large-scale deforestation experiment: effects of patch area and isolation on Amazon birds. Science 315, 238–241.

Fewster, R.M., Buckland, S.T., Siriwardena, G.M., Baillie, S.R., Wilson, J.D., 2000. Analysis of population trends for farmland birds using generalized additive models. Ecology 81, 1970–1984.

Fiske, I., Chandler, R., 2011. unmarked: an R package for fitting hierarchical models of wildlife occurrence and abundance. J. Stat. Softw. 43, 1–23.

Fithian, W., Elith, J., Hastie, T., Keith, D.A., 2014. Bias correction in species distribution models: pooling survey and collection data for multiple species. Methods Ecol. Evol. 6, 424–438.

Fithian, W., Hastie, T., 2013. Finite-sample equivalence in statistical models for presence-only data. Ann. Appl. Stat. 7, 1917–1939.

Fujisaki, I., Mazzotti, F.J., Dorazio, R.M., Rice, K.G., Cherkiss, M., Jeffery, B., 2011. Estimating trend in alligator populations from nightlight survey data. Wetlands 31, 147–155.

Fukaya, K., Royle, J.A., 2013. Markov models for community dynamics allowing for observation error. Ecology 94, 2670–2677.

Garrard, G.E., Bekessy, S.A., McCarthy, M.A., Wintle, B.A., 2008. When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. Austral Ecol. 33, 986–998.

Garrard, G.E., Bekessy, S.A., McCarthy, M.A., Wintle, B.A., 2015. Incorporating detectability into environmental impact assessment for threatened species. Conserv. Biol. 29, 216–225.

Garrard, G.E., McCarthy, M.A., Williams, N.S.G., Bekessy, S.A., Wintle, B.A., 2013. A general model of detectability using species traits. Methods Ecol. Evol. 4, 45–52.

Gelfand, A.E., Smith, A.F., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 85, 398–409.

Gelfand, A.E., Hills, S.E., Racine-Poon, A., Smith, A.F., 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. Journal of the American Statistical Association 85 (412), 972–985.

Gelfand, A.E., Schmidt, A.E., Wu, S., Silander Jr., J.A., Latimer, A., Rebelo, A.G., 2005. Modelling species diversity through species level hierarchical modelling. Appl. Stat. 54, 1–20.

Gelfand, A.E., Silander Jr., J.A., Wu, S., Latimer, A., Lewis, P.O., Rebelo, A.G., Holder, M., 2006. Explaining species distribution patterns through hierarchical modeling. Bayesian Anal. 1, 41–92.

Gelman, A., 2005. Analysis of variance: why is it more important than ever (with discussion). Ann. Stat. 33, 1–53.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 1, 515–534.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2014. Bayesian Data Analysis, third ed. CRC/Chapman & Hall, Boca Raton.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, second ed. Chapman and Hall, London.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/hierarchical Models. Cambridge University Press, Cambridge.

Gelman, A., Meng, X.-L., Stern, H.S., 1996. Posterior predictive assessment of model fitness via realized discrepancies (with discussion). Stat. Sinica 6, 733–807.

Gelman, A., Pardoe, I., 2006. Bayesian measures of explained variance and pooling in multilevel (hierarchical) models. Technometrics 48, 241–251.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7, 457–511.

George, E.I., McCulloch, R.E., 1993. Variable selection via Gibbs sampling. J. Am. Stat. Assoc. 88, 881–889.

Ghosh, S., Gelfand, A.E., Zhu, K., Clark, J.S., 2012. The k-ZIG: flexible modeling for zero-inflated counts. Biometrics 68, 878–885.

Gilks, W.R., Thomas, A., Spiegelhalter, D.J., 1994. A language and program for complex Bayesian modelling. Statistician 43, 169–177.

Gilroy, J.J., Edwards, F.A., Uribe, C.A.M., Haugaasen, T., Edwards, D.P., 2014b. Surrounding habitats mediate the trade-off between land-sharing and land-sparing agriculture in the tropics. J. Appl. Ecol. 51, 1337–1346.

Gilroy, J.J., Prescott, G.W., Cardenas, J.S., Castaneda, P.G.D., Sanchez, A., Rojas-Murcia, L.E., Uribe, C.A.M., Haugaasen, T., Edwards, D.P., 2015. Minimizing the biodiversity impact of Neotropical oil palm development. Glob. Change Biol. 21, 1531–1540.

Gilroy, J.J., Woodcock, P., Edwards, F.A., Wheeler, C., Uribe, C.A.M., Haugaasen, T., Edwards, D.P., 2014a. Optimizing carbon storage and biodiversity protection in tropical agricultural landscapes. Glob. Change Biol. 20, 2162–2172.

Gimenez, O., Choquet, R., Lebreton, J.D., 2003. Parameter redundancy in multistate capture-recapture models. Biomet. J. 45, 704–722.

Gimenez, O., Covas, R., Brown, C.R., Anderson, M.D., Bomberger Brown, M., Lenormand, T., 2006a. Nonparametric estimation of natural selection on a quantitative trait using mark-recapture data. Evolution 60, 460–466.

Gimenez, O., Crainiceanu, C., Barbraud, C., Jenouvrier, S., Morgan, B.J.T., 2006b. Semiparametric regression in capture-recapture modeling. Biometrics 62, 691–698.

Gimenez, O., Viallefont, A., Catchpole, A.E., Choquet, R., Morgan, B.J.T., 2004. Methods for investigating parameter redundancy. Anim. Biodivers. Conserv. 27, 561–572.

Gimenez, O., Viallefont, A., Charmantier, A., Pradel, R., Cam, E., Brown, C.R., Anderson, M.D., Brown Bomberger, M., Covas, R., Gaillard, J.M., 2008. The risk of flawed inference in evolutionary studies when detectability is less than one. Am. Nat. 172, 441−448.

Gimenez, O., Blanc, L., Besnard, A., Pradel, R., Doherty, P.F., Marboutin, E., Choquet, R., 2014. Fitting occupancy models with E-SURGE: hidden Markov modelling of presence−absence data. Methods Ecol. Evol. 5, 592−597.

Giovanini, J., Kroll, A.J., Jones, J.E., Altman, B., Arnett, E.B., 2013. Effects of management intervention on post-disturbance community composition: an experimental analysis using Bayesian hierarchical models. PLoS ONE 8, e59900. http://dx.doi.org/10.1371/journal.pone.0059900.

Gopalaswamy, A.M., Royle, J.A., Delampady, M., Nichols, J.D., Karanth, K.U., Macdonald, D.W., 2012. Density estimation in tiger populations: combining information for strong inference. Ecology 93, 1741−1751.

Gotelli, N.J., McGill, B.J., 2006. Null versus neutral models: what's the difference? Ecography 29, 793−800.

Govindan, B.N., Swihart, R.K., 2015. Community structure of acorn weevils (Curculio): inferences from multispecies occupancy models. Can. J. Ecol. 93, 31−39.

Grimm, V., 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? Ecol. Model. 115, 129−148.

Green, A.W., Hooten, M.B., Grant, E.H.C., Bailey, L.L., 2013. Evaluating breeding and metamorph occupancy and vernal pool management effects for wood frogs using a hierarchical model. J. Appl. Ecol. 50, 1116−1123.

Grosbois, V., Gimenez, O., Gaillard, J.M., Pradel, R., Barbraud, C., Clobert, J., Möller, A.P., Weimerskirch, H., 2008. Assessing the impact of climate variation on survival in vertebrate populations. Biol. Rev. 83, 357−399.

Gu, W., Swihart, R.K., 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. Biol. Conserv. 116, 195−203.

Guélat, J., Kéry, M. Effects of spatial autocorrelation and imperfect detection on large-scale maps of abundance (in review).

Guillera-Arroita, G., 2011. Impact of sampling with replacement in occupancy studies with spatial replication. Methods Ecol. Evol. 2, 401−406.

Guillera-Arroita, G., Lahoz-Monfort, J.J., 2012. Designing studies to detect changes in species occupancy: power analysis under imperfect detection. Methods Ecol. Evol. 3, 860−869.

Guillera-Arroita, G., Lahoz-Monfort, J.J. Species occupancy estimation and imperfect detection: shall surveys continue after the first detection? Adv. Stat. Anal. (in press).

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, K., Lentini, P.E., McCarthy, M.A., Tingley, R., Wintle, B.A., 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Glob. Ecol. Biogeogr. 24, 276−292.

Guillera-Arroita, G., Lahoz-Monfort, J.J., MacKenzie, D.I., Wintle, B.A., McCarthy, M.A., 2014a. Ignoring imperfect detection in biological surveys is dangerous: a response to 'Fitting and Interpreting Occupancy Models'. PLoS ONE 9, e99571.

Guillera-Arroita, G., Morgan, B.J.T., Ridout, M.S., Linkie, M., 2011. Species occupancy modelling for detection data collected along a transect. J. Agric. Biol. Environ. Stat. 3, 301−317.

Guillera-Arroita, G., Morgan, B.J.T., Ridout, M.S., Linkie, M., 2012. Models for species detection data collected along transects in presence of abundance-induced heterogeneity and clustering in the detection process. Methods Ecol. Evol. 3, 358−367.

Guillera-Arroita, G., Ridout, M.S., Morgan, B.J.T., 2010. Design of occupancy studies with imperfect detection. Methods Ecol. Evol. 1, 131−139.

Guillera-Arroita, G., Ridout, M.S., Morgan, B.J.T., 2014b. Two-stage Bayesian study design for species occupancy estimation. J. Agric. Biol. Environ. Stat. 19, 278−291.

Guzy, J.C., Price, S.J., Dorcas, M.E., 2013. The spatial configuration of greenspace affects semi-aquatic turtle occupancy and species richness in a suburban landscape. Landscape Urban Plan. 117, 46−56.

Hammond, P.S., Berggren, P., Benke, H., Borchers, D.L., Collet, A., Heide-Jørgensen, M.P., Heimlich, S., Hiby, A.R., Leopold, M.F., Øien, N., 2002. Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters. J. Appl. Ecol. 39, 361–376.

Hanski, I., 1998. Metapopulation dynamics. Nature 396, 41–49.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, Boca Raton.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.

Hayne, D.W., 1949. Calculation of size of home range. J. Mammal. 30, 1–18.

Hayes, D.B., Monfils, M.J., 2015. Occupancy modeling of bird point counts: Implications of mobile animals. J. Wildl. Manage. http://dx.doi.org/10.1002/jwmg.943.

He, F., Gaston, K.J., 2000. Estimating species abundance from occurrence. Am. Nat. 156, 553–559.

Hector, A., Bell, T., Hautier, Y., Isbell, F., Kéry, M., Reich, P.B., van Ruijven, J., Schmid, B., 2011. Bugs in the analysis of biodiversity experiments: species richness and composition are of similar importance for grassland productivity. PLoS ONE 6, e17434.

Hedley, S.L., Buckland, S.T., 2004. Spatial models for line transect sampling. J. Agric. Biol. Environ. Stat. 9, 181–199.

Heikkinen, J., Högmander, H., 1994. Fully Bayesian approach to image restoration with an application in biogeography. Appl. Stat. 43, 569–582.

Henden, J.A., Yoccoz, N.G., Ims, R.A., Langeland, K., 2013. How spatial variation in areal extent and configuration of labile vegetation states affect the riparian bird community in Arctic tundra. PLoS ONE 8, e63312. http://dx.doi.org/10.1371/journal.pone.0063312.

Hestbeck, J.B., Nichols, J.D., Malecki, R.A., 1991. Estimates of movement and site fidelity using mark-resight data of wintering Canada Geese. Ecology 72, 523–533.

Higa, M., Yamaura, Y., Koizumi, I., Yabuhara, Y., Senzaki, M., Ono, S., 2015. Mapping large-scale bird distributions using occupancy models and citizen data with spatially biased sampling effort. Divers. Distrib. 21, 46–54.

Hines, J.E., 2006. PRESENCE – Software to Estimate Patch Occupancy and Related Parameters. USGS-PWRC. www.mbr-pwrc.usgs.gov/software/presence.shtml.

Hines, J.E., Nichols, J.D., Collazo, J.A., 2014. Multiseason occupancy models for correlated replicate surveys. Methods Ecol. Evol. 5, 583–591.

Hines, J.E., Nichols, J.D., Royle, J.A., MacKenzie, D.I., Gopalaswamy, A.M., Samba Kumar, N., Karanth, K.U., 2010. Tigers on trails: occupancy modeling for cluster sampling. Ecol. Appl. 20, 1456–1466.

Hostetler, J.A., Chandler, R.B., 2015. Improved state-space models for inference about spatial and temporal variation in abundance from count data. Ecology 96, 1713–1723.

Hobbs, N.T., Hooten, M.B., 2015. Bayesian Models: A Statistical Primer for Ecologists. Princeton University Press, Princeton, NJ.

Holt, B.G., Rioja-Nieto, R., MacNeil, M.A., Lupton, J., Rahbek, C., 2013. Comparing diversity data collected using a protocol designed for volunteers with results from a professional alternative. Methods Ecol. Evol. 4, 383–392.

Holtrop, A.M., Cao, Y., Dolan, C.R., 2010. Estimating sampling effort required for characterizing species richness and site-to-site similarity in fish assemblage surveys of wadeable Illinois streams. Trans. Am. Fish. Soc. 139, 1421–1435.

Holyoak, M., Leibold, M.A., Holt, R.D. (Eds.), 2005. Metacommunities: Spatial Dynamics and Ecological Communities. University of Chicago Press, Chicago, IL.

Homyack, J.A., O'Bryan, C.J., Thornton, J.E., Baldwin, R.F., 2014. Anuran assemblages associated with roadside ditches in a managed pine landscape. For. Ecol. Manage. 334, 217–231.

Hooten, M.B., Hobbs, N.T., 2015. A guide to Bayesian model selection for ecologists. Ecol. Monogr. 85, 3–28.

Hunt, S.D., Guzy, J.C., Price, S.J., Halstead, B.J., Eskew, E.A., Dorcas, M.E., 2013. Responses of riparian reptile communities to damming and urbanization. Biol. Conserv. 157, 277–284.

Hurlbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54, 187–211.

Hutchinson, R.A., Valente, J.V., Emerson, S.C., Betts, M.G., Dietterich, T.G., 2015. Penalized likelihood methods improve parameter estimates in occupancy models. Methods Ecol. Evol. http://dx.doi.org/10.1111/2041-210X.12368. Article first published online: March 31, 2015.

Illian, J.B., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. Wiley, Chichester.

Illian, J.B., Sørbye, S.H., Rue, H., Hendrichsen, D.K., August 2012. Using INLA to fit a complex point process model with temporally varying effects – a case study. J. Environ. Stat. 3 (7). http://www.jenvstat.org/v03/i07.

Iknayan, K.J., Tingley, M.W., Furnas, B.J., Beissinger, S.R., 2014. Detecting diversity: emerging methods to estimate species diversity. Trends Ecol. Evol. 29, 97–106.

Ives, A.R., Zhu, J., 2006. Statistics for correlated data: phylogenies, space, and time. Ecol. Appl. 16, 20–32.

Jackman, S., 2009. Bayesian Analysis for the Social Sciences. John Wiley & Sons.

Jackman, S., 2012. pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University. Department of Political Science, Stanford University, Stanford, California. R package version 1.04.4. http://pscl.stanford.edu.

Jenni, L., Kéry, M., 2003. Timing of autumn bird migration under climate change: advances in long-distance migrants, delays in short-distance migrants. Proc. R. Soc. Lond. Series B 270, 1467–1471.

Jolly, G.M., 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52, 225–247.

Johnson, D.H., 2008. In defense of indices: the case of bird surveys. J. Wildl. Manage. 72, 857–868.

Johnson, D.S., Conn, P.B., Hooten, M.B., Ray, J.C., Pond, B.A., 2013. Spatial occupancy models for large data sets. Ecology 94, 801–808.

Johnson, D.S., Laake, J.L., Ver Hoef, J.M., 2010. A model-based approach for making ecological inference from distance sampling data. Biometrics 66, 310–318.

Johnson, F.A., Dorazio, R.M., Castellón, T.D., Martin, J., Garcia, J.O., Nichols, J.D., 2014. Tailoring point counts for inference about avian density: dealing with nondetection and availability. Nat. Res. Model. 27, 163–177.

Jones, J.E., Kroll, A.J., Giovanini, J., Duke, S.D., Ellis, T.M., Betts, M.G., 2012. Avian species richness in relation to intensive forest management practices in early seral tree plantations. PLoS ONE 7, e43290. http://dx.doi.org/10.1371/journal.pone.0043290.

Jonsen, I.D., Mills Flemming, J., Myers, R.A., 2005. Robust state-space modeling of animal movement data. Ecology 86, 2874–2880.

Joseph, L.N., Elkin, C., Martin, T.G., Possingham, H., 2009. Modeling abundance using N-mixture models: the importance of considering ecological mechanisms. Ecol. Appl. 19, 631–642.

Kadane, J.B., Lazar, N.A., 2004. Methods and criteria for model selection. J. Am. Stat. Assoc. 99, 279–290.

Kalinowski, S.T., Taper, M.L., Creel, S., 2006. Using DNA from non-invasive samples to identify individuals and census populations: an evidential approach tolerant of genotyping errors. Conserv. Genet. 7, 319–329.

Karanth, K.U., Nichols, J.D., 1998. Estimation of tiger densities in India using photographic captures and recaptures. Ecology 79, 2852–2862.

Karanth, K.U., Nichols, J.D., Kumar, N.S., Hines, J.E., 2006. Assessing tiger population dynamics using photographic capture-recapture sampling. Ecology 87, 2925–2937.

Karanth, K.K., Nichols, J.D., Sauer, J.R., Hines, J.E., Yackulic, C.B., 2014. Latitudinal gradients in North American avian species richness, turnover rates and extinction probabilities. Ecography 37, 626–636.

Karr, J.R., 1990. Biological integrity and the goal of environmental legislation: lessons for conservation biology. Conserv. Biol. 4, 244–250.

Kellner, K., 2015. jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses. R package version 1.3.7. http://CRAN.R-project.org/package=jagsUI.

Kellner, K.F., Swihart, R.K., 2014. Accounting for Imperfect Detection in Ecology: A Quantitative Review. PLoS ONE 9, e111436.

Kendall, W.L., Hines, J.E., Nichols, J.D., 2003. Adjusting multistate capture-recapture models for misclassification bias: manatee breeding proportions. Ecology 84, 1058–1066.

Kendall, W.L., Hines, J.E., Nichols, J.D., Grant, E.H.C., 2013. Relaxing the closure assumption in occupancy models: staggered arrival and departure times. Ecology 94, 610–617.

Kendall, W.L., Nichols, J.D., Hines, J.E., 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. Ecology 78, 563–578.

Kéry, M., 2002. Inferring the absence of a species - a case study of snakes. J. Wildl. Manage. 66, 330–338.

Kéry, M., 2004. Extinction rate estimates for plant populations in revisitation studies: importance of detectability. Conserv. Biol. 18, 570–574.

Kéry, M., 2008. Estimating abundance from bird counts: binomial mixture models uncover complex covariate relationships. Auk 125, 336–345.

Kéry, M., 2010. Introduction to WinBUGS for Ecologists - a Bayesian Approach to Regression, ANOVA, Mixed Models and Related Analyses. Academic Press, Burlington.

Kéry, M., 2011a. Species richness and community dynamics — a conceptual framework. In: O'Connell, A.F., Nichols, J.D., Karanth, K.U. (Eds.), Camera Traps in Animal Ecology: Methods and Analyses. Springer, Tokyo, pp. 207–231.

Kéry, M., 2011b. Towards the modeling of true species distributions. J. Biogeogr. 38, 617–618.

Kéry, M., Dorazio, R.M., Soldaat, L., van Strien, A., Zuiderwijk, A., Royle, J.A., 2009. Trend estimation in populations with imperfect detection. J. Appl. Ecol. 46, 1163–1172.

Kéry, M., Gardner, B., Monnerat, C., 2010a. Predicting species distributions from checklist data using site-occupancy models. J. Biogeogr. 37, 1851–1862.

Kéry, M., Gardner, B., Stoeckle, T., Weber, D., Royle, J.A., 2011. Use of spatial capture-recapture modeling and DNA data to estimate densities of elusive animals. Conserv. Biol. 25, 356–364.

Kéry, M., Gregg, K.B., 2003. Effects of life-state on detectablity in a demographic study of the terrestrial orchid *Cleistes bifaria*. J. Ecol. 91, 265–273.

Kéry, M., Gregg, K.B., 2004. Demographic analysis of dormancy and survival in the terrestrial orchid *Cypripedium reginae*. J. Ecol. 92, 686–695.

Kéry, M., Gregg, K.B., Schaub, M., 2005a. Demographic estimation methods for plants with unobservable life-states. Oikos 108, 307–320.

Kéry, M., Guillera-Arroita, G., Lahoz-Monfort, J.J., 2013. Analysing and mapping species range dynamics using dynamic occupancy models. J. Biogeogr. 40, 1463–1474.

Kéry, M., Madsen, J., Lebreton, J.D., 2006. Survival of Svalbard pink-footed geese *Anser brachyrhynchus* in relation to winter climate, density and land-use. J. Anim. Ecol. 75, 1172–1181.

Kéry, M., Royle, J.A., 2008. Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. J. Appl. Ecol. 45, 589–598.

Kéry, M., Royle, J.A., 2009. Inference about species richness and community structure using species-specific occupancy models in the national Swiss breeding bird survey MHB. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations, pp. 639–656.

Kéry, M., Royle, J.A., 2010. Hierarchical modeling and estimation of abundance in metapopulation designs. J. Anim. Ecol. 79, 453–461.

Kéry, M., Royle, J.A., Schmid, H., 2005b. Modeling avian abundance from replicated counts using binomial mixture models. Ecol. Appl. 15, 1450−1461.

Kéry, M., Royle, J.A., Schmid, H., 2008. Importance of sampling design and analysis in animal population studies: a comment on Sergio et al. J. Appl. Ecol. 45, 981−986.

Kéry, M., Royle, J.A., Schmid, H., Schaub, M., Volet, B., Häfliger, G., Zbinden, N., 2010b. Site-ocupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. Conserv. Biol. 24, 1388−1397.

Kéry, M., Schaub, M., 2012. Bayesian Population Analysis Using WinBUGS: A Hierarchical Perspective. Academic Press.

Kéry, M., Schmidt, B.R., 2008. Imperfect detection and its consequences for monitoring for conservation. Community Ecol. 9, 207−216.

Kéry, M., Spillmann, J.H., Truong, C., Holderegger, R., 2006. How biased are estimates of extinction probability in revisitation studies. J. Ecol. 94, 980−986.

King, R., Morgan, B.J.T., Gimenez, O., Brooks, S.P., 2009. Bayesian Analysis for Population Ecology. Chapmann & Hall, Boca Raton.

Knape, J., Korner-Nievergelt, F., 2015. Estimates from non-replicated population surveys rely on critical assumptions. Methods Ecol. Evol. 6, 298−306.

Koenen, K.K., DeStefano, S., Krausman, P.R., 2002. Using distance sampling to estimate seasonal densities of desert mule deer in a semidesert grassland. Wildl. Soc. Bull. 30, 53−63.

Koneff, M.D., Royle, J.A., Otto, M.C., Wortham, J.S., Bidwell, J.K., 2008. A double-observer method to estimate detection rate during aerial waterfowl surveys. J. Wildl. Manage. 72, 1641−1649.

Korner-Nievergelt, F., von Felten, S., Roth, T., Almasi, B., Guélat, J., Korner-Nievergelt, P., 2015. Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and Stan, first ed. Academic Press.

Krebs, C.J., 2009. Ecology: The Experimental Analysis of Distribution and Abundance, sixth ed. Benjamin Cummings, San Francisco.

Kroll, A.J., Ren, Y., Jones, J.E., Giovanini, J., Perry, R.W., Thill, R.E., White, D., Wigley, T.B., 2014. Avian community composition associated with interactions between local and landscape habitat attributes. For. Ecol. Manage. 326, 46−57.

Kuo, L., Mallick, B., 1998. Variable selection for regression models. Sankhya 60B, 65−81.

Laake, J.L., Collier, B.A., Morrison, M.L., Wilkins, R.N., 2011. Point-based mark-recapture distance sampling. J. Agric. Biol. Environ. Stat. 16, 389−408.

Lachish, S., Gopalaswamy, A.M., Knowles, S.C.L., Sheldon, B.C., 2012. Site-occupancy modelling as a novel framework for assessing test sensitivity and estimating wildlife disease prevalence from imperfect diagnostic tests. Methods Ecol. Evol. 3, 339−348.

Lahoz-Monfort, J.J., Guillera-Arroita, G., Wintle, B.A., 2014. Imperfect detection impacts the performance of species distribution models. Glob. Ecol. Biogeogr. 23, 504−515.

Langtimm, C.A., Dorazio, R.M., Stith, B.M., Doyle, T.J., 2011. New aerial survey and hierarchical model to estimate manatee abundance. J. Wildl. Manage. 75, 399−412.

Latimer, A.M., Wu, S., Gelfand, A.E., Silander Jr., J.A., 2006. Building statistical models to analyse species distributions. Ecol. Appl. 16, 33−50.

Le Cam, L., 1990. Maximum likelihood — an introduction. ISI Review 58, 153−171.

Lee, Y., Nelder, J.A., 2000. Two ways of modeling overdispersion in non-normal data. App. Stat. 49, 591−598.

Lee, Y., Nelder, J.A., Pawitan, Y., 2006. Generalized Linear Models with Random Effects. Unified Analysis via H-likelihood. Chapman and Hall/CRC, Boca Raton, FL.

Lele, S.R., 2010. Model complexity and information in the data: could it be a house built on sand? Ecology 91, 3493−3496.

Lele, S.R., 2015. Is Non-informative Bayesian Analysis Appropriate for Wildlife Management: Survival of San Joaquin Kit Fox and Declines in Amphibian Populations. arXiv:1502.00483 [q-bio.QM].

Lele, S.R., Dennis, B., 2009. Bayesian methods for hierarchical models: are ecologists making a Faustian bargain? Ecol. Appl. 19, 581−584.

Lele, S.R., Keim, J.L., 2006. Weighted distributions and estimation of resource selection probability functions. Ecology 87, 3021−3028.

Lele, S.R., Moreno, M., Bayne, E., 2012. Dealing with detection error in site occupancy surveys: what can we do with a single survey? J. Plant Ecol. 5, 22−31.

Lewis, T.L., Lindberg, M.S., Schmutz, J.A., Bertram, M.R., Dubour, A.J., 2015. Species richness and distributions of boreal waterbird broods in relation to nesting and brood-rearing habitats. J. Wildl. Manage. 79, 296−310.

Linden, D.W., Roloff, G.J., 2013. Retained structures and bird communities in clearcut forests of the Pacific Northwest, USA. For. Ecol. Manage. 310, 1045−1056.

Link, W.A., 1999. Modeling pattern in collections of parameters. J. Wildl. Manage. 63, 1017−1027.

Link, W.A., 2013. A cautionary note on the discrete uniform prior for the binomial N. Ecology 94, 2173−2179.

Link, W.A., Barker, R.J., 2006. Model weights and the foundations of multimodel inference. Ecology 87, 2626−2635.

Link, W.A., Barker, R.J., 2010. Bayesian Inference with Ecological Applications. Academic Press, London.

Link, W.A., Eaton, M.J., 2012. On thinning of chains in MCMC. Methods Ecol. Evol. 3, 112−115.

Link, W.A., Royle, J.A., Hatfield, J.S., 2003. Demographic analysis from summaries of an age-structured population. Biometrics 59, 778−785.

Link, W.A., Sauer, J.R., 2002. A hierarchical analysis of population change with application to Cerulean warblers. Ecology 83, 2832−2840.

Link, W.A., Yoshizaki, J., Bailey, L.L., Pollock, K.H., 2010. Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. Biometrics 66, 178−185.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O., 2008. SAS for Mixed Models, second ed. SAS Institute, Cary, NC.

Little, R.J.A., 2006. Calibrated Bayes: a bayes/frequentist roadmap. Am. Stat. 60, 213−223.

Little, R.J.A., Rubin, D.B., 2002. Statistical Analysis with Missing Data, second ed. Wiley, New York.

Liu, J.S., Wu, Y.N., 1999. Parameter expansion for data augmentation. J. Am. Stat. Assoc. 94, 1264−1274.

Lukacs, P.M., Burnham, K.P., 2005. Estimating population size from DNA-based closed capture−recapture data incorporating genotyping error. J. Wildl. Manage. 69, 396−403.

Lunn, D., Jackson, C., Best, N., Thomas, A., Spiegelhalter, D., 2013. The BUGS Book: A Practical Introduction to Bayesian Analysis. Chapman and Hall/CRC.

Lunn, D.J., Spiegelhalter, D., Thomas, A., Best, N., 2009. The BUGS project: evaluation, critique and future directions. Stat. Med. 28, 3049−3067.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10, 325−337.

Lynch, H.J., Thorson, J.T., Shelton, A.O., 2014. Dealing with under- and overdispersed count data in life history, spatial, and community ecology. Ecology 95, 3173−3180.

Lyons, J.E., Royle, J.A., Thomas, S.M., Elliott-Smith, E., Evenson, J.R., Kelly, E.G., Milner, R.L., Nysewander, D.R., Andres, B.A., 2012. Large-scale monitoring of shorebird populations using count data and N-mixture models: Black Oystercatcher (*Haematopus bachmani*) surveys by land and sea. Auk 129, 645−652.

MacEachern, S.N., Berliner, L.M., 1994. Subsampling the Gibbs sampler. Am. Stat. 48, 188−190.

MacKenzie, D.I., 2005. What are the issues with presence-absence data for wildlife managers? J. Wildl. Manage. 69, 849−860.

MacKenzie, D.I., Bailey, L.L., 2004. Assessing the fit of site-occupancy models. J. Agric. Biol. Environ. Stat. 9, 300−318.

MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G., Franklin, A.B., 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84, 2200−2207.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83, 2248−2255.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Hines, J.E., Bailey, L.L., 2006. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence. Elsevier, San Diego.

MacKenzie, D.I., Nichols, J.D., Seamans, M.E., Gutierrez, R.J., 2009. Modeling species occurrence dynamics with multiple states and imperfect detection. Ecology 90, 823−835.

MacKenzie, D.I., Nichols, J.D., Sutton, N., Kawanishi, K., Bailey, L.L., 2005. Improving inferences in population studies of rare species that are detected imperfectly. Ecology 86, 1101−1113.

MacKenzie, D.I., Royle, J.A., 2005. Designing occupancy studies: general advice and allocating survey effort. J. Appl. Ecol. 42, 1105−1114.

Magnusson, W.E., Caughley, G.J., Grigg, G.C., 1978. A double-survey estimate of population size from incomplete counts. J. Wildl. Manage. 42, 174−176.

Marques, F.F.C., Buckland, S.T., 2003. Incorporating covariates into standard line transect analyses. Biometrics 59, 924−935.

Marques, T.A., Buckland, S.T., Bispo, R., Howland, B., 2013. Accounting for animal density gradients using independent information in distance sampling surveys. Stat. Methods Appl. 22, 67−80.

Marques, T.A., Buckland, S.T., Borchers, D.L., Tosh, D., McDonald, R.A., 2010. Point transect sampling along linear features. Biometrics 66, 1247−1255.

Marques, F.F., Buckland, S.T., Goffin, D., Dixon, C.E., Borchers, D.L., Mayle, B.A., Peace, A.J., 2001. Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. J. Appl. Ecol. 38, 349−363.

Marques, T.A., Thomas, L., Fancy, S.G., Buckland, S.T., 2007. Improving estimates of bird density using multiple-covariate distance sampling. Auk 124, 1229−1243.

Marsh, H., Sinclair, D.F., 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. J. Wildl. Manage. 53, 1017−1024.

Martin, T.G., Kuhnert, P.M., Mergersen, K., Possingham, H.P., 2005. The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. Ecol. Appl. 15, 266−280.

Martin, J., Nichols, J.D., McIntyre, C.L., Ferraz, G., Hines, J.E., 2009. Perturbation analysis for patch occupancy dynamics. Ecology 90, 10−16.

Martin, J., Royle, J.A., Mackenzie, D.I., Edwards, H.H., Kéry, M., Gardner, B., 2011. Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. Methods Ecol. Evol. 2, 595−601.

Mata, L., Goula, M., Hahs, A.K., 2014. Conserving insect assemblages in urban landscapes: accounting for species-specific responses and imperfect detection. J. Insect Conserv. 18, 885−894.

Matechou, E., Dennis, E.B., Freeman, S.N., Brereton, T., 2014. Monitoring abundance and phenology in (multivoltine) butterfly species: a novel mixture model. J. Appl. Ecol. 51, 766−775.

Mattsson, B.J., Zipkin, E.F., Gardner, B., Blank, P.J., Sauer, J.R., Royle, J.A., 2013. Explaining local-scale species distributions: relative contributions of spatial autocorrelation and landscape heterogeneity for an avian assemblage. PLoS ONE 8, e55097.

Mazerolle, M.J., 2015. AICcmodavg: Model Selection and Multimodel Inference Based on (Q)AIC(c). R package version 2.0-3. http://CRAN.R-project.org/package=AICcmodavg.

McCarthy, M.A., 2007. Bayesian Methods for Ecology. Cambridge University Press, Cambridge.

McCarthy, M.A., Masters, P., 2005. Profiting from prior information in Bayesian analyses of ecological data. J. Appl. Ecol. 42, 1012−1019.

McCarthy, M.A., Moore, J.L., Morris, W.K., Parris, K.M., Garrard, G.E., Vesk, P.A., Rumpff, L., Giljohann, K.M., Camac, J.S., Bau, S.S., Friend, T., Harrison, B., Yue, B., 2013. The influence of abundance on detectabiliy. Oikos 122, 717–726.

McClintock, B.T., Bailey, L.L., Pollock, K.H., Simons, T.R., 2010a. Unmodeled observation error induces bias when inferring patterns and dynamics of species occurrence via aural detections. Ecology 91, 2446–2454.

McClintock, B.T., Nichols, J.D., Bailey, L.L., MacKenzie, D.I., Kendall, W.L., Franklin, A.B., 2010b. Seeking a second opinion: uncertainty in disease ecology. Ecol. Lett. 13, 659–674.

McCoy, E.D., Heck Jr., K.L., 1987. Some observations on the use of taxonomic similarity in large-scale biogeography. J. Biogeogr. 14, 79–87.

McCrea, R.S., Morgan, B.J.T., 2014. Analysis of Capture-recapture Data. Chapman & Hall/CRC Press, Boca Raton, FL, USA.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman & Hall, London.

McCulloch, C.E., Searle, S.R., 2001. Generalized, Linear, and Mixed Models. Wiley, New York.

McIntyre, A.P., Jones, J.E., Lund, E.M., Waterstrat, F.T., Giovanini, J.N., Duke, S.D., Hayes, M.P., Quinn, T., Kroll, A.J., 2012. Empirical and simulation evaluations of an abundance estimator using unmarked individuals of cryptic forest-dwelling taxa. For. Ecol. Manage. 286, 129–136.

McKann, P.C., Gray, B.R., Thogmartin, W.E., 2013. Small sample bias in dynamic occupancy models. J. Wildl. Manage. 77, 172–180.

McKenny, H.C., Keeton, W.S., Donovan, T.M., 2006. Effects of structural complexity enhancement on eastern red-backed salamander (*Plethodon cinereus*) populations in northern hardwood forests. For. Ecol. Manage. 230, 186–196.

McManamay, R.A., Orth, D.J., Jager, H.I., 2014. Accounting for variation in species detection in fish community monitoring. Fish. Manage. Ecol. 21, 96–112.

McNew, L.B., Handel, C.M., 2015. Evaluating species richness: biased ecological inference results from spatial heterogeneity in detection probabilities. Ecol. Appl. 25, 1669–1680.

Mead, R., 1988. The Design of Experiments: Statistical Principles for Practical Applications. Cambridge University Press, Cambridge UK.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092.

Mihaljevic, J.R., Joseph, M.B., Johnson, P.T.J., 2015. Using multispecies occupancy models to improve the characterization and understanding of metacommunity structure. Ecology 96, 1783–1792.

Millar, R.B., 2009. Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' factors. Biometrics 65, 962–969.

Miller, A., 2002. Subset Selection in Regression. Chapman & Hall/CRC.

Miller, D.A., Bailey, L.L., Grant, E.H.C., McClintock, B.T., Weir, L., Simons, T.R., 2015. Performance of species occurrence estimators when basic assumptions are not met: a test using field data where true occupancy status is known. Methods Ecol. Evol. 6, 557–565.

Miller, D.A., Nichols, J.D., McClintock, B.T., Grant, E.H.C., Bailey, L.L., Weir, L., 2011. Improving occupancy estimation when two types of observational errors occur: non-detection and species misidentification. Ecology 92, 1422–1428.

Miller, D.A.W., Nichols, J.D., Gude, J.A., Rich, L.N., Podruzny, K.M., Hines, J.E., Mitchell, M.S., 2013b. Determining occurrence dynamics when false positives occur: estimating the range dynamics of wolves from public survey data. PLoS ONE 8, e65808.

Miller, D.L., Burt, M.L., Rexstad, E.A., Thomas, L., 2013a. Spatial models for distance sampling data: recent developments and future directions. Methods Ecol. Evol. 4, 1001–1010.

Miller, D.L., Thomas, L., 2015. Mixture models for distance sampling detection functions. PLoS ONE 10, e0118726. http://dx.doi.org/10.1371/journal.pone.0118726.

Moore, J.E., Barlow, J., 2011. Bayesian state-space model of fin whale abundance trends from a 1991−2008 time series of line-transect surveys in the California Current. J. Appl. Ecol. 48, 1195−1205.

Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E., Fryxell, J.M., 2004. Extracting more out of relocation data: Building movement models as mixtures of random walks. Ecology 85, 2436−2445.

Mordecai, R.S., Mattsson, B.J., Tzilkowski, C.J., Cooper, R.J., 2011. Addressing challenges when studying mobile or episodic species: hierarchical Bayes estimation of occupancy and use. J. Appl. Ecol. 48, 56−66.

Murtaugh, P.A., 2007. Simplicity and complexity in ecological data analysis. Ecology 88, 56−62.

Nelder, J.A., 1965a. The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. Proc. R. Soc. Series A 283, 147−162.

Nelder, J.A., 1965b. The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. Proc. R. Soc. Series A 283, 163−178.

Newman, K.B., Buckland, S.T., Lindley, S.T., Thomas, L., Fernandez, C., 2006. Hidden process models for animal population dynamics. Ecol. Appl. 16, 74−86.

Newman, K., Buckland, S.T., Morgan, B., King, R., Borchers, D.L., Cole, D., Besbeas, P., Gimenez, O., Thomas, L., 2014. Modelling Population Dynamics, Model Formulation, Fitting and Assessment Using State-space Methods. Springer.

Nichols, J.D., Bailey, L.L., O'Connell, A.F., Talancy, N.W., Grant, E.H.C., Gilbert, A.T., Annand, E.M., Husband, T.P., Hines, J.E., 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. J. Appl. Ecol. 45, 1321−1329.

Nichols, J.D., Boulinier, T., Hines, J.E., Pollock, K.H., Sauer, J.R., 1998a. Estimating rates of local species extinction, colonization, and turnover in animal communities. Ecol. Appl. 8, 1213−1225.

Nichols, J.D., Boulinier, T., Hines, J.E., Pollock, K.H., Sauer, J.R., 1998b. Inference methods for spatial variation in species richness and community composition when not all species are detected. Conserv. Biol. 12, 1390−1398.

Nichols, J.D., Hines, J.E., MacKenzie, D.I., Seamans, M.E., Gutierrez, R.J., 2007. Occupancy estimation and modeling with multiple states and state uncertainty. Ecology 88, 1395−1400.

Nichols, J.D., Hines, J.E., Sauer, J.R., Fallon, F.W., Fallon, J.E., Heglund, P.J., 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117, 393−408.

Nichols, J.D., Thomas, L., Conn, P.B., 2009. Inferences about landbird abundance from count data: recent advances and future directions. In: Thomson, D.L., Cooch, E.G., Conroy, M.J. (Eds.), Modeling Demographic Processes in Marked Populations. Springer, New York, pp. 201−235.

Nichols, J.D., McIntyre, C.L., Ferraz, G., Hines, J.E., 2009. Perturbation analysis for patch occupancy dynamics. Ecology 90, 10−16.

Niemi, A., Fernandez, C., 2010. Bayesian spatial point processmodeling of line transect data. J. Agric. Biol. Environ. Stat. 15, 327−345.

NIMBLE Development Team, 2015. NIMBLE: An R Package for Programming with BUGS Models, Version 0.4. http://r-nimble.org.

Norris III, J.L., Pollock, K.H., 1996. Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics 639−649.

Ntzoufras, I., 2009. Bayesian Modeling Using WinBUGS. Wiley, Hoboken, New Jersey.

O'Brien, T.G., Baillie, J.E.M., Krueger, L., Cuke, M., 2010. The Wildlife Picture Index: monitoring top trophic levels. Anim. Conserv. 13, 335−343.

O'Hara, R.B., Sillanpää, M.J., 2009. A review of Bayesian variable selection methods: what, how and which. Bayesian Anal. 4, 85−118.

Oedekoven, C.S., Buckland, S.T., Mackenzie, M.L., Evans, K.O., Burger Jr., L.W., 2013. Improving distance sampling: accounting for covariates and non-independency between sampled sites. J. Appl. Ecol. 50, 786−793.

Oedekoven, C.S., Buckland, S.T., MacKenzie, M.L., King, R., Evans, K., Burger, W., 2014. Bayesian methods for hierarchical distance sampling models. J. Agric. Biol. Environ. Stat. 19, 219−239.

Olea, P.P., Mateo-Tomas, P., 2011. Spatially explicit estimation of occupancy, detection probability and survey effort needed to inform conservation planning. Divers. Distrib. 17, 714–724.

Otis, D.L., Burnham, K.P., White, G.C., Anderson, D.R., 1978. Statistical inference from capture data on closed animal populations. Wildl. Monogr. 62, 1–135.

Ovaskainen, O., Hottola, J., Siitonen, J., 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. Ecology 91, 2514–2521.

Ovaskainen, O., Soininen, J., 2011. Making more out of sparse data: hierarchical modeling of species communities. Ecology 92, 289–295.

Pacifici, K., Zipkin, E.F., Collazo, J.A., Irizarry, J.I., DeWan, A., 2014. Guidelines for a priori grouping of species in hierarchical community models. Ecol. Evol. 4, 877–888.

Papaïx, J., Cubaynes, S., Buoro, M., Charmantier, A., Perret, P., Gimenez, O., 2010. Combining capture-recapture data and pedigree information to assess heritability of demographic parameters in the wild. J Evol Biol 23, 2176–2184.

Pardo, M.A., Gerrodette, T., Beier, E., Gendron, D., Forney, K.A., et al., 2015. Inferring cetacean population densities from the absolute dynamic topography of the ocean in a hierarchical Bayesian framework. PLoS ONE 10, e0120727.

Pavlacky, D.C., Blakesley, J.A., White, G.C., Hanni, D.J., Lukacs, P.M., 2012. Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. J. Wildl. Manage. 76, 154–162.

Pearce, J.L., Boyce, M.S., 2006. Modelling distribution and abundance with presence-only data. J. Appl. Ecol. 43, 405–412.

Penteriani, V., 2003. Breeding density affects the honesty of bird vocal displays as possible indicators of male/territory quality. Ibis 145, E127–E135 (on-line).

Phillips, S.J., Dudik, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175.

Phillips, S.J., Elith, J., 2013. On estimating probability of presence from use-availability or presence-background data. Ecology 94, 1409–1419.

Pinheiro, J.C., Bates, D.M., 2000. Mixed-Effects Models in S and S-Plus. Springer, New York.

Plummer, M., 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Hornik, K., Leisch, F., Zeileis, A. (Eds.), Proceedings of the 3rd International Workshop in Distributed Statistical Computing (DSC 2003), March 20–22. Technische Universität, Vienna, Austria, pp. 1–10.

Plummer, M., 2015. rjags: Bayesian Graphical Models Using MCMC. R package version 3-15. http://CRAN.R-project.org/package=rjags.

Pollock, J.F., 2006. Detecting population declines over large areas with presence-absence, time-to-encounter, and count survey methods. Conserv. Biol. 20, 882–892.

Pollock, K.H., 1982. A capture-recapture design robust to unequal probability of capture. J. Wildl. Manage. 46, 752–757.

Pollock, K.H., Nichols, J.D., Brownie, C., Hines, J.E., 1990. Statistical inference for capture-recapture experiments. Wildl. Monogr. 107, 3–97.

Pollock, K.H., Nichols, J.D., Simons, T.R., Farnsworth, G.L., Bailey, L.L., Sauer, J.R., 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics 13, 105–119.

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A., McCarthy, M.A., 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). Methods Ecol. Evol. 5, 397–406.

Post van der Burg, M., Bly, B., Vercauteren, T., Tyre, A.J., 2011. Making better use of monitoring data from low density species using a spatially explicit modeling approach. J. Appl. Ecol. 48, 47–55.

Potts, J.M., Elith, J., 2006. Comparing species abundance models. Ecol. Model. 199, 153–163.

Pradel, R., 2005. Multievent: an extension of multistate capture-recapture models to uncertain states. Biometrics 61, 442−447.

Pradel, R., Hines, J.E., Lebreton, J.D., Nichols, J.D., 1997. Capture-recapture survival models taking account of transients. Biometrics 53, 60−72.

Purvis, A., Hector, A., 2000. Getting the measure of biodiversity. Nature 405, 212−219.

Qian, S.S., Shen, Z., 2007. Ecological applications of multilevel analysis of variance. Ecology 88, 2489−2495.

Railsback, S.F., Grimm, V., 2012. Agent-Based and Individual-Based Modeling. Princeton University Press, Princeton, USA.

Ramsey, D.S.L., Caley, P.A., Robley, A., 2015. Estimating population density from presence-absence data using a spatially explicit model. J. Wildl. Manage. 79, 491−499.

Renner, I.W., Elith, J., Baddeley, A., Fithian, W., Hastie, T., Phillips, S.J., Popovic, G., Warton, D.I., 2015. Point process models for presence-only analysis. Methods Ecol. Evol. 6, 366−379.

Renner, I.W., Warton, D.I., 2013. Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. Biometrics 69, 274−281.

Riddle, J.D., Mordecai, R.S., Pollock, K.H., Simons, T.R., 2010. Effects of prior detections on estimates of detection probability, abundance and occupancy. Auk 127, 94−99.

Ridout, M.S., Besbeas, P., 2004. An empirical model for underdispersed count data. Stat. Model. 4, 77−89.

Robert, C., Casella, G., 2010. Introducing Monte Carlo Methods with R. Springer Science & Business Media.

Rosenstock, S.S., Anderson, D.R., Giesen, K.M., Leukering, T., Carter, M.F., 2002. Landbird counting techniques: current practices and an alternative. Auk 119, 46−53.

Rota, C.T., Fletcher Jr., R.J., Dorazio, R.M., Betts, M.G., 2009. Occupancy estimation and the closure assumption. J. Appl. Ecol. 46, 1173−1181.

Roth, T., Amrhein, V., 2009. Estimating individual survival using territory occupancy data on unmarked animals. J. Appl. Ecol. 47, 386−392.

Rout, T.M., Moore, J.L., McCarthy, M.A., 2014. Prevent, search or destroy? A partially observable model for invasive species management. J. Appl. Ecol. 51, 804−813.

Royle, J.A., 2004a. Generalized estimators of avian abundance from count survey data. Anim. Biodivers. Conserv. 27, 375−386.

Royle, J.A., 2004b. N-mixture models for estimating population size from spatially replicated counts. Biometrics 60, 108−115.

Royle, J.A., 2006. Site occupancy model with heterogeneous detection probabilities. Biometrics 62, 97−102.

Royle, J.A., 2008. Modeling individual effects in the Cormack-Jolly-Seber model: a state-space formulation. Biometrics 64, 364−370.

Royle, J.A., 2009. Analysis of capture-recapture models with individual covariates using data augmentation. Biometrics 65, 267−274.

Royle, J.A., Chandler, R.B., Sollmann, R., Gardner, B., 2014. Spatial Capture-Recapture. Academic Press.

Royle, J.A., Chandler, R.B., Sun, C.C., Fuller, A.K., 2013. Integrating resource selection information with spatial capture-recapture. Methods Ecol. Evol. 4, 520−530.

Royle, J.A., Chandler, R.B., Yackulic, C., Nichols, J.D., 2012. Likelihood analysis of species occurrence probability from presence-only data for modeling species distributions. Methods Ecol. Evol. 3, 545−554.

Royle, J.A., Converse, S.J., 2014. Hierarchical spatial capture-recapture models: modelling population density in stratified populations. Methods Ecol. Evol. 5, 37−43.

Royle, J.A., Converse, S.J., Link, W.A., 2012. Data augmentation for hierarchical capture-recapture models arXiv preprint arXiv:1211.5706.

Royle, J.A., Dawson, D.K., Bates, S., 2004. Modeling abundance effects in distance sampling. Ecology 85, 1591−1597.

Royle, J.A., Dorazio, R.M., 2006. Hierarchical models of animal abundance and occurrence. J. Agric. Biol. Environ. Stat. 11, 249–263.

Royle, J.A., Dorazio, R.M., 2008. Hierarchical Modeling and Inference in Ecology. The Analysis of Data from Populations, Metapopulations and Communities. Academic Press, New York.

Royle, J.A., Dorazio, R.M., 2012. Parameter-expanded data augmentation for Bayesian analysis of capture-recapture models. J. Ornithol. 152, 521–537.

Royle, J.A., Dorazio, R.M., Link, W.A., 2007a. Analysis of multinomial models with unknown index using data augmentation. J. Comput. Gr. Stat. 16, 67–85.

Royle, J.A., Kéry, M., 2007. A Bayesian state-space formulation of dynamics occupancy models. Ecology 88, 1813–1823.

Royle, J.A., Kéry, M., Gauthier, R., Schmid, H., 2007b. Hierarchical spatial models of abundance and occurrence from imperfect survey data. Ecol. Monogr. 77, 465–481.

Royle, J.A., Kéry, M., Guélat, J., 2011. Spatial capture-recapture models for search-encounter data. Methods Ecol. Evol. 2, 602–611.

Royle, J.A., Link, W.A., 2005. A general class of multinomial mixture models for anuran calling survey data. Ecology 86, 2505–2512.

Royle, J.A., Link, W.A., 2006. Generalized site occupancy models allowing for false positive and false negative errors. Ecology 87, 835–841.

Royle, J.A., Nichols, J.D., 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84, 777–790.

Royle, J.A., Nichols, J.D., Kéry, M., 2005. Modelling occurrence and abundance of species when detection is imperfect. Oikos 110, 353–359.

Royle, J.A., Young, K.G., 2008. A hierarchical model for spatial capture-recapture data. Ecology 89, 2281–2289.

Rubin, D.B., 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. Ann. Stat. 12, 1151–1172.

Ruiz-Gutierrez, V., Zipkin, E.F., 2011. Detection biases yield misleading patterns of species persistence and colonization in fragmented landscapes. Ecosphere 2, article 61.

Ruiz-Gutierrez, V., Zipkin, E.F., Dhondt, A.A., 2010. Occupancy dynamics in a tropical bird community: unexpectedly high forest use by birds classified as non-forest species. J. Appl. Ecol. 47, 621–630.

Ruppert, D., 2002. Selecting the number of knots for penalized splines. J. Comput. Gr. Stat. 11, 735–757.

Ruppert, D., Wand, M., Carroll, R., 2003. Semiparametric Regression. Cambridge University Press, Cambridge, UK.

Russell, J.C., Stjernman, M., Lindstrom, A., Smith, H.G., 2015. Community occupancy before-after-control-impact (CO-BACI) analysis of Hurricane Gudrun on Swedish forest birds. Ecol. Appl. 25, 685–694.

Russell, R.E., Royle, J.A., Saab, V.A., Lehmkuhl, J.F., Block, W.M., Sauer, J.R., 2009. Modeling the effects of environmental disturbance on wildlife communities: avian responses to prescribed fire. Ecol. Appl. 19, 1253–1263.

Sadoti, G., Zuckerberg, B., Jarzyna, M.A., Porter, W.F., 2013. Applying occupancy estimation and modeling to the analysis of atlas data. Divers. Distrib. 19, 804–814.

Sanderlin, J.S., Block, W.M., Ganey, J.L., 2014. Optimizing study design for multi-species avian monitoring programmes. J. Appl. Ecol. 51, 860–870.

Sauer, J.R., Link, W.A., 2002. Hierarchical modeling of population stability and species group attributes from survey data. Ecology 86, 1743–1751.

Sauer, J.R., Link, W.A., 2011. Analysis of the North American breeding bird survey using hierarchical models. Auk 128, 87–98.

Sauer, J.R., Blank, P.J., Zipkin, E.F., Fallon, J.E., Fallon, F.W., 2013. Using multi-species occupancy models in structured decision making on managed lands. J. Wildl. Manage. 77, 117–127.

Schaub, M., Abadi, F., 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. J. Ornithol. 152 (Suppl. 1), 227−237.

Schaub, M., Gimenez, O., Sierro, A., Arlettaz, R., 2007. Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. Conserv. Biol. 21, 945−955.

Schaub, M., Jakober, H., Stauber, W., 2013. Strong contribution of immigration to local population regulation: evidence from a migratory passerine. Ecology 94, 1828−1838.

Schaub, M., Kéry, M., 2012. Combining information in hierarchical models improves inferences in population ecology and demographic population analyses. Anim. Conserv. 15, 125−126.

Schaub, M., Reichlin, T.S., Abadi, F., Kéry, M., Jenni, L., Arlettaz, R., 2012. The demographic drivers of local population dynamics in two rare migratory birds. Oecologia 168, 97−108.

Schmid, H., Zbinden, N., Keller, V., 2004. Überwachung der Bestandsentwicklung häufiger Brutvögel in der Schweiz. Schweizerische Vogelwarte, Sempach.

Schmidt, B.R., Kéry, M., Ursenbacher, S., Hyman, O.J., Collins, J.P., 2013. Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen. Methods Ecol. Evol. 4, 646−653.

Schmidt, J.H., Rattenbury, K.L., 2013. Reducing effort while improving inference: estimating Dall's sheep abundance and composition in small areas. J. Wildl. Manage. 77, 1048−1058.

Schmidt, J.H., Rattenbury, K.L., Lawler, J.P., MacCluskie, M.C., 2012. Using distance sampling and hierarchical models to improve estimates of Dall's sheep abundance. J. Wildl. Manage. 76, 317−327.

Schofield, M.R., Barker, R.J., 2008. A unified capture-recapture framework. J. Agric. Biol. Environ. Stat. 13, 458−477.

Schofield, M.R., Barker, R.J., MacKenzie, D.I., 2009. Flexible hierarchical mark-recapture modeling for open populations using WinBUGS. Envir. Ecol. Stat. 16, 369−387.

Seber, G.A.F., 1965. A note on the multiple recapture census. Biometrika 52, 249−259.

Seber, G.A.F., 1982. The Estimation of Animal Abundance and Related Parameters. Charles Griffin & Company Ltd, London.

Shirk, P.L., Linden, D.W., Patrick, D.A., Howell, K.M., Harper, E.B., Vonesh, J.R., 2014. Impact of habitat alteration on endemic Afromontane chameleons: evidence for historical population declines using hierarchical spatial modelling. Divers. Distrib. 20, 1186−1199.

Sillett, T.S., Chandler, R.B., Royle, J.A., Kéry, M., Morrison, S.A., 2012. Hierarchical distance-sampling models to estimate population size and habitat-specific abundance of an island endemic. Ecol. Appl. 22, 1997−2006.

Sollmann, R., Gardner, B., Chandler, R.B., Royle, J.A., Sillett, T.S., 2015. An open population hierarchical distance sampling model. Ecology 96, 325−331.

Sollmann, R., Gardner, B., Williams, K., Gilbert, A., Veit, R., 2016. A hierarchical distance sampling model to estimate abundance and covariate associations of species and communities. Methods Ecol. Evol. in press.

Solymos, P., 2010. dclone: data Cloning in R. R J. 2, 29−37. http://journal.r-project.org/.

Solymos, P., Lele, S.R., Bayne, E., 2012. Conditional likelihood approach for analyzing single visit abundance survey data in the presence of zero inflation and detection error. Environmetrics 23, 197−205.

Solymos, P., Matsuoka, S.M., Bayne, E.M., Lele, S.R., Fontaine, P., Cumming, S.G., Stralberg, D., Schmiegelow, F.K.A., Song, S.J., 2013. Calibrating indices of avian density from non-standardized survey data: making the most of a messy situation. Methods Ecol. Evol. 4, 1047−1058.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measure of model complexity and fit. J. R. Stat. Soc. Series B 64, 583−639.

Stanley, T.R., Royle, J.A., 2005. Estimating site occupancy and abundance using indirect detection indices. J. Wildl. Manage. 69, 874−883.

Stearns, S.C., 1992. The Evolution of Life Histories. Oxford University Press, Oxford.

Stefanski, L.A., 2000. Measurement error models. J. Am. Stat. Assoc. 95, 1353−1358.

Strebel, N., Kéry, M., Schaub, M., Schmid, H., 2014. Study of phenology by flexible estimation and modeling of seasonal detectability peaks. Methods Ecol. Evol. 5, 483–490.

Stone, M., 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. Journal of the Royal Statistical Society. Series B (Methodological) 44–47.

Sturtz, S., Ligges, U., Gelman, A., 2005. R2WinBUGS: a package for running WinBUGS from R. J. Stat. Softw. 12, 1–16.

Su, Y.-S., Yajima, M., 2014. R2jags: A Package for Running Jags from R. R package version 0.04-01. http://CRAN.R-project.org/package=R2jags.

Sutherland, C., Brambilla, M., Pedrini, P., Tenan, S. A multi-region community model for inference about geographic variation in species richness. Methods Ecol. Evol. (in review).

Sutherland, C., Elston, D.A., Lambin, X., 2012. Multi-scale processes in metapopulations: contributions of stage structure, rescue effect, and correlated extinctions. Ecology 93, 2465–2473.

Sutherland, C., Elston, D.A., Lambin, X., 2013. Accounting for false positive detection error induced by transient individuals. Wildl. Res. 40, 490–498.

Sutherland, C., Elston, D.A., Lambin, X., 2014. A demographic, spatially explicit patch occupancy model of metapopulation dynamics and persistence. Ecology 95, 3149–3160.

Tanadini, L., 2010. Heterogeneity Effects in the N-Mixture Model (Unpublished Masters thesis). University of Neuchâtel, Neuchâtel, Switzerland.

Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. J. Am. Stat. Assoc. 82, 528–540.

Tavecchia, G., Besbeas, P., Coulson, T., Morgan, B.J.T., Clutton-Brock, T.H., 2009. Estimating population size and hidden demographic parameters with state-space modeling. Am. Nat. 173, 722–733.

Tenan, S., O'Hara, R.B., Hendriks, I., Tavecchia, G., 2014a. Bayesian model selection: the steepest mountain to climb. Ecol. Model. 283, 62–69.

Tenan, S., Pradel, R., Tavecchia, G., Igual, J.M., Sanz-Aguilar, A., Genovart, M., Oro, D., 2014b. Hierarchical modelling of population growth rate from individual capture–recapture data. Methods Ecol. Evol. 5, 606–614.

Thomas, A., O'Hara, B., Ligges, U., Sturtz, S., 2006. Making BUGS open. R News 6, 12–17.

Thomas, L., Buckland, S.T., Newman, K.B., Harwood, J., 2005. A unified framework for modelling wildlife population dynamics. Aust. N.Z. J. Stat. 47, 19–34.

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A., Burnham, K.P., 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. J. Appl. Ecol. 47, 5–14.

Tingley, M.W., Beissinger, S.R., 2013. Cryptic loss of montane avian richness and high community turnover over 100 years. Ecology 94, 598–609.

Tobler, M.W., Hartley, A.Z., Carrillo-Percastegui, S.E., Powell, G.V.N., 2015. Spatiotemporal hierarchical modelling of species richness and occupancy using camera trap data. J. Appl. Ecol. 52, 413–421.

Trolle, M., Kéry, M., 2003. Estimation of ocelot density in the Pantanal using capture-recapture analysis of camera-trapping data. J. Mammal. 84, 607–614.

Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K., Possingham, H.P., 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecol. Appl. 13, 1790–1801.

ver Hoef, J.M., Boveng, P.L., 2007. Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data. Ecology 88, 2766–2772.

ver Hoef, J.M., Cameron, M.F., Boveng, P.L., London, J.M., Moreland, E.E., 2014. A spatial hierarchical model for abundance of three ice-associated seal species in the eastern Bering Sea. Stat. Methodol. 17, 46–66.

ver Hoef, J.M., Frost, K.J., 2003. A Bayesian hierarchical model for monitoring harbor seal changes in Prince William Sound, Alaska. Environ. Ecol. Stat. 10, 201–219.

ver Hoef, J.M., Jansen, J.K., 2007. Space-time zero-inflated count models of harbour seals. Environmetrics 18, 697−712.

ver Hoef, J.M., Boveng, P.L., 2015. Iterating on a single model is a viable alternative to multimodel inference. J. Wildlife Manag. 79, 719−729.

Waldmann, P., 2009. Easy and Flexible Bayesian Inference of Quantitative Genetic Parameters. Evolution 63, 1640−1643.

Warren, C.C., Veech, J.A., Weckerly, F.W., O'Donnell, L., Ott, J.R., 2013. Detection heterogeneity and abundance estimation in populations of Golden-cheeked warblers (*Setophaga chrysoparia*). Auk 130, 677−688.

Warton, D.I., Shepherd, L.C., 2010. Poisson point process models solve the "pseudo-absence problem" for presence-only data in ecology. Ann. Appl. Stat. 4, 1383−1402.

Webster, R.A., Pollock, K.H., Simons, T.R., 2008. Bayesian spatial modeling of data from avian point surveys. J. Agric. Biol. Environ. Stat. 13, 121−139.

Wells, K., Bohm, S.M., Boch, S., Fischer, M., Kalko, E.K.V., 2011. Local and landscape-scale forest attributes differ in their impact on bird assemblages across years in forest production landscapes. Basic Appl. Ecol. 12, 97−106.

Welsh, A.H., Lindenmayer, D.B., Donnelly, C.F., 2013. Fitting and interpreting occupancy models. PLoS ONE 8, e52015.

Wenger, S.J., Freeman, M.C., 2008. Estimating species occurrence, abundance, and detection probability using zero-inflated distributions. Ecology 89, 2953−2959.

White, G.C., 2005. Correcting wildlife counts using detection probabilities. Wildl. Res. 32, 211−216.

White, G.C., Burnham, K.P., 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46, 120−139.

White, A.M., Zipkin, E.F., Manley, P.N., Schlesinger, M.D., 2013a. Conservation of avian diversity in the Sierra Nevada: moving beyond a single-species management focus. PLoS ONE 8, e63088. http://dx.doi.org/10.1371/journal.pone.0063088.

White, A.M., Zipkin, E.F., Manley, P.N., Schlesinger, M.D., 2013b. Simulating avian species and foraging group responses to fuel reduction treatments in coniferous forests. For. Ecol. Manage. 304, 261−274.

Whittaker, R.J., Willis, K.J., Field, R., 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. J. Biogeogr. 28, 453−470.

Wiegand, T., Moloney, K.A., 2014. A Handbook of Spatial Point Pattern Analysis in Ecology. Chapman and Hall/CRC Press, Boca Raton.

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and Management of Animal Populations. Academic Press, San Diego.

Wintle, B.A., Bardos, D.C., 2006. Modeling species−habitat relationships with spatially autocorrelated observation data. Ecol. Appl. 16, 1945−1958.

Wood, S.N., 2006. Generalized Additive Models. An Introduction with R. Chapman & Hall/CRC, Boca Raton, USA.

Woodworth, G.G., 2004. Biostatistics: A Bayesian Introduction. Wiley.

Woodward, P., 2011. Bayesian Analysis Made Simple: An Excel GUI for WinBUGS. Chapman & Hall/CRC.

Wright, J.A., Barker, R.J., Schofield, M.R., Frantz, A.C., Byrom, A.E., Gleeson, D.M., 2009. Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. Biometrics 65, 833−840.

Wu, G., Holan, S.H., Wikle, C.K., 2013. Hierarchical bayesian spatio-temporal Conway−Maxwell Poisson models with dynamic dispersion. J. Agric. Biol. Environ. Stat. 18, 335−356.

Wu, G., Holan, S.H., Nilon, C.H., Wikle, C.K., 2015. Bayesian binomial mixture models for estimating abundance in ecological monitoring studies. Ann. Appl. Stat. 9, 1−26.

Yackulic, C.B., Reid, J., Davis, R., Hines, J.E., Nichols, J.D., Forsman, E., 2012. Neighborhood and habitat effects on vital rates: expansion of the Barred Owl in the Oregon coast ranges. Ecology 93, 1953−1966.

Yamaura, Y., 2013. Confronting imperfect detection: behavior of binomial mixture models under varying circumstances of visits, sampling sites, detectability, and abundance, in small-sample situations. Ornithol. Sci. 12, 73–88.

Yamaura, Y., Kéry, M., Royle, J.A. Study of biological communities subject to imperfect detection: bias and precision of multispecies N-mixture abundance models in small-sample situations. Ecol. Res., in press.

Yamaura, Y., Royle, J.A., Kubio, K., Tada, T., Ikeno, S., Makino, S., 2011. Modelling community dynamics based on species-level abundance models from detection/nondetection data. J. Appl. Ecol. 48, 67–75.

Yamaura, Y., Royle, J.A., Shimada, N., Asanuma, S., Sato, T., Taki, H., Makino, S., 2012. Biodiversity of man-made open habitats in an underused country: a class of multispecies abundance models for count data. Biodivers. Conserv. 21, 1365–1380.

Yang, H.C., Chao, A., 2005. Modeling animals' behavioral response by Markov chain models for capture–recapture experiments. Biometrics 61, 1010–1017.

Yoccoz, N.G., Nichols, J.D., Boulinier, T., 2001. Monitoring biological diversity in space and time. Trends Ecol. Evol. 16, 446–453.

Yoshizaki, J., Pollock, K.H., Brownie, C., Webster, R.A., 2009. Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. Ecology 90, 3–9.

Zellweger-Fischer, J., Kéry, M., Pasinelli, G., 2011. Population trends of brown hares in Switzerland: the role of land-use and ecological compensation areas. Biol. Conserv. 144, 1364–1373.

Zipkin, E.F., DeWan, A., Royle, J.A., 2009. Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling. J. Appl. Ecol. 46, 815–822.

Zipkin, E.F., Royle, J.A., Dawson, D.K., Bates, S., 2010. Multi-species occurrence models to evaluate the effects of conservation and management actions. Biol. Conserv. 143, 479–484.

Zipkin, E.F., Grant, E.H.C., Fagan, W.F., 2012. Evaluating the predictive abilities of community occupancy models using AUC while accounting for imperfect detection. Ecol. Appl. 22, 1962–1972.

Zipkin, E.F., Sillett, T.S., Grant, E.H.C., Chandler, R.B., Royle, J.A., 2014b. Inferences about population dynamics from count data using multistate models: a comparison to capture-recapture approaches. Ecol. Evol. 4, 417–426.

Zipkin, E.F., Thorson, J.T., See, K., Lynch, H.J., Grant, E.H.C., Kanno, Y., Chandler, R.B., Letcher, B.H., Royle, J.A., 2014a. Modeling structured population dynamics using data from unmarked individuals. Ecology 95, 22–29.

Zippin, C., 1956. An evaluation of the removal method of estimating animal populations. Biometrics 12, 163–189.

Zuur, A.F., Saveliev, A.A., Ieno, E.N., 2012. Zero-inflated Models and Generalized Linear Mixed Models with R. Highlands Statistics.

Marc Kéry • J. Andrew Royle

APPLIED HIERARCHICAL MODELING IN ECOLOGY

Analysis of distribution, abundance and species richness in R and BUGS

Volume 1: Prelude and Static Models

Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance, and species richness in R and BUGS offers a new synthesis of hierarchical modeling in studies of plant and animal distribution, abundance, and community characteristics such as species richness using data collected in meta-population designs. This type of site-stratified data is extremely widespread in ecology and its applications in such areas as biodiversity monitoring, conservation biology, and fisheries, and wildlife management.

This first volume begins by explaining the basic concepts of distribution and abundance, hierarchical models and their Bayesian and likelihood analysis, linear and generalized linear models, data simulation, and use of R and BUGS software for fitting traditional linear, generalized linear, and simple mixed models. The second part of the volume presents six "monographs" on the hierarchical modeling of distribution and abundance for single and multiple species, using classical and Bayesian inference and based on such common data as counts, capture/recapture, distance sampling, and detection/nondetection ("presence/absence"). *Applied Hierarchical Modeling in Ecology, Volume 1* serves as an indispensable manual for practicing field biologists, as a graduate-level text for students in ecology, conservation biology, fisheries/wildlife management, and related fields, and as an overview of this field for statisticians.

This book:

- Provides a synthesis of important classes of hierarchical models about distribution, abundance, and species richness while explicitly accommodating measurement error (imperfect detection)
- Emphasizes data simulation as a way of studying and understanding stochastic processes and statistical models
- Takes a dual inference paradigm approach by the seamless integration of both likelihood-based and Bayesian inference
- Uses the dominant software for likelihood and Bayesian analyses in ecology: R and BUGS
- Is written in a very accessible style with a large number of fully worked examples that will kick-start your analyses of your own data
- Includes access to a companion website containing data sets, code, errata, solutions to exercises, and further information, and is supported by two e-mail lists maintained by the authors
- Lays the foundation for Volume 2 of *Applied Hierarchical Modeling in Ecology,* wherein dynamic and more advanced versions of these models will be presented

Marc Kéry is a population ecologist with the Swiss Ornithological Institute. He is the author of about 80 peer-reviewed journal articles and two books on a wide range of topics, including the analysis of large-scale monitoring programs, demographic population analyses, experimental design for animal and plant surveys, and the population ecology of rare species.

J. Andrew Royle is a research statistician at the US Geological Survey's Patuxent Wildlife Research Center. His research is focused on the application of probability and statistics to ecological problems, especially those related to animal sampling and demographic modeling, about which he has authored over 100 peer-reviewed journal articles and three books.

Cover photograph: Greater Papuan Emperor – *Anax maclachlani* (Drawing by A.G. Orr)

ACADEMIC PRESS
An imprint of Elsevier
store.elsevier.com

ISBN 978-0-12-801378-6