

Methods in Statistical Ecology

S.T. Buckland  
E.A. Rexstad  
T.A. Marques  
C.S. Oedekoven

# Distance Sampling: Methods and Applications

 Springer

# Methods in Statistical Ecology

## **Series Editors**

Andrew P. Robinson, Melbourne, Victoria, Australia

Stephen T. Buckland, St. Andrews, United Kingdom

Peter Reich, St. Paul, USA

Michael McCarthy, Parkville, Australia

More information about this series at <http://www.springer.com/series/10235>



S.T. Buckland • E.A. Rexstad • T.A. Marques  
C.S. Oedekoven

# Distance Sampling: Methods and Applications



S.T. Buckland  
CREEM The Observatory  
University of St. Andrews  
St. Andrews, UK

E.A. Rexstad  
CREEM The Observatory  
University of St. Andrews  
St. Andrews, UK

T.A. Marques  
CREEM The Observatory  
University of St. Andrews  
St. Andrews, UK

C.S. Oedekoven  
CREEM The Observatory  
University of St. Andrews  
St. Andrews, UK

ISSN 2199-319X  
Methods in Statistical Ecology  
ISBN 978-3-319-19218-5  
DOI 10.1007/978-3-319-19219-2

ISSN 2199-3203 (electronic)  
ISBN 978-3-319-19219-2 (eBook)

Library of Congress Control Number: 2015941311

Springer Cham Heidelberg New York Dordrecht London  
© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

# Preface

This is the fourth book dedicated to distance sampling. The first (Buckland et al. 1993) appeared at around the same time as the first version of the *Distance* software (Laake et al. 1993). This ground-breaking book was largely replaced by *Introduction to Distance Sampling* (Buckland et al. 2001). A further book, *Advanced Distance Sampling* (Buckland et al. 2004), mostly covered advances since 1993. In parallel, many advances were introduced into the *Distance* software, summarized in Thomas et al. (2010). Since 2004, there have been many new developments, and it is becoming clearer which of these advances are of most use. In this book, we seek to cover the basic methods, and the various advances that we believe are of greatest value to practitioners. There is an increasing trend towards conducting distance sampling analyses using the statistical software R, rather than using the *Distance* software. Several of our examples reflect this.

In Chap. 1, we build on conventional plot sampling ideas to outline the basic methods of conventional distance sampling and the respective assumptions. Prior to Buckland et al. (1993), there was very little work on relaxing the four key distance sampling assumptions. Subsequently, methods for when animals on the line or at the point are not certain to be detected were developed, and Buckland et al. (2004) include a comprehensive treatment of double-observer methods. Since 2004, there has been work on incorporating measurement error, and allowing non-uniform animal distribution. Further, the importance of modelling animal availability in the context of estimating probability of detection on the line or at the point has been recognized in recent work. Incorporating animal movement models into distance sampling is a focus of current research.

In a marked departure from previous books on distance sampling, we bring design issues and field methods forward. In terms of the practical application of these methods, it is fundamental to get the design right and the field methods optimized. Data arising from poor survey design and/or field methods might or might not be salvageable using sophisticated analysis methods, but data from a good and carefully thought-out design with optimized field methods are generally simpler to analyse, and give more reliable estimates of abundance.

In Chap. 2, we address survey design issues for points and lines, covering aspects of stratification, survey effort guidelines and assessing designs by simulation. We also include a chapter on the design of distance sampling experiments (Chap. 3). Increasingly, distance sampling is being used not to estimate abundance in a study area, but to assess whether animal density varies by type of plot. For example, some plots may be subject to a treatment, while others act as a control, and we wish to assess whether the treatment has an effect on density. The treatment might relate to management of habitat, use of a pesticide, construction of a structure, *etc.*

Field methods and data checking are covered in Chap. 4. Trying to address data quality issues at the analysis stage is no substitute for gathering good quality data, with procedures for identifying and correcting recording errors.

At the heart of distance sampling methods is the concept of a detection function. The detection function represents the probability  $g(y)$  of detecting an animal, given that it is at a distance  $y$  from a line or point. Intuitively, this should be a non-increasing function of distance. In Chap. 5, we consider models for the detection function  $g(y)$ . For conventional distance sampling (CDS), we assume that this is a function of  $y$  alone, with  $g(0) = 1$  (*i.e.* an animal on the line or point is certain to be detected). For multiple-covariate distance sampling (MCDS), we allow the function to vary with other covariates in addition to distance. For mark-recapture distance sampling (MRDS), we also allow  $g(0) < 1$ . Modelling  $g(y)$  is crucial, because it allows us to estimate the proportion that we detect of animals that are located on the sampled strips or circles (the sampling units, which together comprise the covered region).

For CDS, we adopt a hybrid approach to modelling, so that model-based methods are used to estimate the detection function (that is, we fit a model to distances, to estimate for each animal detected its probability of detection), but design-based methods are used to give robust extrapolation from the sampling units to the wider survey region. In other words, we estimate abundance on the plots, and extrapolate to the whole study area, based on the random design, which ensures that the plots are representative of the area from which they were drawn. In Chap. 6, we address issues associated with design-based estimation of animal density and abundance.

An alternative to CDS methods is to consider distance sampling data as counts of detected animals within a set of sampling units. Viewed in this way, then it is natural to use count models such as Poisson or negative binomial to analyse the data. Chapters 7 and 8 present such methods. Distance sampling spatial models fit into this framework. Seminal work in this area stemmed from Hedley et al. (1999) and Hedley (2000). As the name implies, extrapolation beyond the sampling units to a wider study area is based on a model indexed in space.

In Chap. 7, we adopt such a model-based approach for counts. We consider an ‘offset’ within a count regression model to account for detectability. Thus, the approach has two stages: the detection function is estimated in the first, which provides an estimated offset to include in the count model in stage two. In designed

distance sampling experiments, this allows us to test whether treatment effects in the count model are significant, while for standard distance sampling studies, we can extrapolate beyond the samplers by modelling how counts vary through the survey region. We can also estimate an animal density surface, which allows density to be modelled as a function of spatial covariates such as habitat, and abundance estimates to be obtained for any sub-area by integrating the corresponding section of the estimated density surface.

While the two-stage approach is easier to implement, it is not conceptually elegant. Perhaps worse is that it raises issues relating to uncertainty propagation across the two stages. Therefore, in Chap. 8 we take these models further, exploring maximum likelihood and Bayesian methods to give flexible options for analysing distance sampling data adopting a single, integrated likelihood.

In Chap. 9, we describe variations on the theme of distance sampling, concentrating on extensions of line or point transect sampling that either are already in wide use or that have the potential to see wide use: cue counting, in which cues (for example, whale blows or calls, song bursts or calls for bird surveys, or primate calls) are recorded, from which animal density is estimated; trapping and lure point transects, in which a trap or lure is placed at each survey point; indirect surveys in which signs produced by the animal of interest (usually dung or nests) are surveyed; and three-dimensional surveys (such as sonar surveys of fish or radar surveys of migrating birds).

Distance sampling is used for a wide range of taxa and habitats. While most of the advice is general, some issues are taxon-specific. Chapter 10 discusses special issues that arise in applying distance sampling methods to the following taxa: songbirds, seabirds, cetaceans, primates, ungulates, butterflies and plants.

While we recommend to design surveys and collect field data to ensure that key assumptions are met, this is not always possible, and options are needed for when assumption failure cannot be ignored. Therefore, in Chap. 11, we return to assumptions, and consider ways to allow assumptions to be relaxed or removed. This will be done at the expense of requiring additional data and/or additional assumptions.

In Chap. 12, we finish by summarizing key issues, and providing a checklist for designing, conducting and analysing a distance sampling survey. We also discuss some of the new technologies that are being used to change the way surveys are implemented. Although this section is likely to become outdated very quickly, we believe that it is useful to practitioners to be aware of the technological advances that we feel will have most impact on the future implementation of distance sampling methods.

For key case studies, we provide data and files to allow the reader to implement the methods on the book website:

<http://www.creem.st-and.ac.uk/DS.M&A/>.

## Readers' Guide

We provide suggestions on how to use the book for several classes of readers. We hope that the book will be read in its entirety, but on first reading, we provide some suggested paths through the book.

- Those new to distance sampling should read Chaps. 1 and 2 for basic acquaintance with distance sampling principles. In addition, these readers should read Chap. 4 for thoughts on data collection procedures, Chaps. 5 and 6 for fitting detection functions and the estimation of density and abundance, and the checklist of Chap. 12.
- Those familiar with distance sampling should review Chaps. 1–6, noting in particular the discussion of experiments to assess effects of treatments upon abundance in Chap. 3. In addition, Chaps. 9–11 provide additional tools for dealing with special circumstances where traditional distance sampling methods may not perform satisfactorily.
- Readers interested in model-based methods to exploit the data from surveyed plots to draw inference on the study area should read Chaps. 7 and 8 in lieu of Chap. 6. The mathematical treatment of model-based methods in the unified likelihood framework of Chap. 8 is fairly sophisticated.
- We see distance sampling integrated with the assessment of experimental treatment effects (Chap. 3) as well as increased use of model based-inference (Chaps. 7 and 8) as areas of future methodological expansion. These chapters, along with the final chapter (Chap. 12), in which we speculate on the effects of technological advances upon distance sampling, should be read by those making contributions to methodological development of distance sampling.

## Acknowledgements

This book includes many illustrative examples based on real case studies. It is a pleasure to thank everyone who kindly provided data for our case studies.

The conservation buffer case studies with indigo buntings and northern bob-white quail coveys (Sects. 3.2.2.1, 8.5.1 and 8.5.2): The national CP33 monitoring program was coordinated and delivered by the Department of Wildlife, Fisheries, and Aquaculture and the Forest and Wildlife Research Center, Mississippi State University. The national CP33 monitoring program was funded by the Multistate Conservation Grant Program (Grants MS M-1-T, MS M-2-R), a program supported with funds from the Wildlife and Sport Fish Restoration Program and jointly managed by the Association of Fish and Wildlife Agencies, U.S. Fish and Wildlife Service, USDA-Farm Service Agency, and USDA-Natural Resources Conservation Service-Conservation Effects Assessment Project.

The cork oak woodland data (Sect. 5.2.2.4): Filipe S. Dias, Miguel N. Bugalho, Joana Marcelino (Centro de Ecologia Aplicada, Prof. Baeta Neves/Instituto Superior

de Agronomia, Universidade de Lisboa) and Jorge Orestes Cerdeira (Departamento de Matemática, Faculdade de Ciências da Universidade Nova de Lisboa).

The black-capped vireo data used in Sect. 5.4.2.3 were provided by Dr Bret Collier and collected under The U.S. Army Integrated Training Area Management Program, Office of the Secretary of Defense, Department of Defense, who provided funding support.

The long-tailed duck data used in Sect. 7.4.3 were collected by the National Environmental Research Institute of Denmark.

The crossbill lure point transect survey (Sect. 9.2.1): the Royal Society for the Protection of Birds and Scottish Natural Heritage co-commissioned this national survey.

The red deer dung data (Sect. 9.3.3): the Royal Society for the Protection of Birds conducted the surveys and funded the development of the survey design.

The chimpanzee nest data (Sect. 9.3.4): Hjalmar Kühl at the Max Planck Institute for Evolutionary Anthropology.

The DTAG data from Sect. 10.3.3: Peter Tyack for facilitating access to DTAG data. Tagging was performed under US National Marine Fisheries Service Research Permit Nos. 981-1578-02 and 981-1707-00 to PLT and with the approval of the Woods Hole Oceanographic Institution Animal Care and Use Committee. Jessica Ward and NUWC colleagues for processing data to obtain localizations of the animals.

The hare survey data (Sect. 11.1.2): David Tosh and Robbie Macdonald for access to the data; the Northern Ireland Environment Agency (formerly known as the Environment and Heritage Service) through their Natural Heritage Research Partnership with Queen's University Belfast.

The eastern grey kangaroo data survey (Sect. 11.1.3): Brett Howland for facilitating access to the data; ACT Government Ecologists Dr Don Fletcher and Claire Wimpenny for providing GPS collar data.

We thank Eric Howe, who reviewed a draft of the book, and Laura Marshall, who put together the DSSim case study. Suresh Kumar from the SPi TeX Support team was very helpful in sorting bibtex formatting issues.

STB thanks his wife Marlene for unwavering support, and his colleagues at CREEM, who provide a great environment to work on projects such as this book. He is especially grateful to those colleagues who have relieved him of various responsibilities, leaving him (relatively) free to do what he enjoys. He also thanks various collaborators from around the world, who have generated and/or helped solve many interesting problems in distance sampling.

ER acknowledges the support of his colleagues at CREEM, workshop participants and contributors to the distance sampling email list for presenting to him many facets of distance sampling; making him think more deeply about issues faced by practising field biologists. He is grateful to his wife Charisse for her support in this and other adventures.

TAM thanks his family and close friends for the continued support, in particular his grandmother Gabriela Silva for funding his first trip to St Andrews back in 2001. It has been a while since then, but nothing would have happened without it. Dinis

Pestana for pulling him to the dark side, Russell Alpizar-Jara for the first contact with distance sampling and the decisive push towards a PhD in St Andrews. Steve Buckland during the PhD and Len Thomas ever since, for believing in him and providing more support than he could have wished for. Anything I ever do in life will be partly yours. A big thank you to CREEMinals in general: it has been both a pleasure and a privilege to work in such a vibrant yet friendly environment. To Regina, with love. To Filipe, with a message: now both mom and dad have written a book, so it's your turn!

CSO thanks the CREEM team, friends and family for the continuous support during her PhD studies and as a postdoctoral research fellow. Furthermore, she is grateful to scientists outside CREEM: Kristine Evans and Wes Burger, both from Mississippi State University, Jeff Laake from the National Oceanic and Atmospheric Administration and Hans Skaug from the University of Bergen for collaborations over the past few years which are mentioned in this book. CSO was part-funded through the National Centre for Statistical Ecology by EPSRC/NERC grant EP/1000917/1.

St. Andrews, UK

Stephen T. Buckland  
Eric A. Rexstad  
Tiago A. Marques  
Cornelia S. Oedekoven

# Contents

**Part I Introduction, Survey Design and Field Methods**

<b>1</b>	<b>The Basic Methods</b>	3
1.1	Introduction	3
1.2	Line and Point Transect Sampling	5
1.3	The Montrave Case Study	6
1.4	Plot Sampling	8
1.4.1	Estimating Density and Abundance	8
1.4.2	Limitations of Plot Sampling	9
1.5	Line Transect Sampling	10
1.6	Point Transect Sampling	11
1.7	Assumptions	12
<b>2</b>	<b>Designing Surveys</b>	15
2.1	Point Transect Survey Designs	15
2.2	Line Transect Survey Designs	17
2.3	Stratification	19
2.3.1	Geographic Strata	19
2.3.2	Post-stratification	22
2.4	Survey Effort	23
2.4.1	Minimum Sample Size	23
2.4.2	Estimating Total Effort Required	23
2.5	Comparing Designs	25
2.5.1	Assessing Designs by Simulation	25
2.5.2	Example: Comparing the Efficiency of Parallel-Line and Zig-Zag Designs	26
<b>3</b>	<b>Designing Distance Sampling Experiments</b>	29
3.1	Introduction	29
3.2	Designed Experiments	30
3.2.1	Completely Randomized Designs	30
3.2.2	Randomized Block Designs	30



3.2.3	Factorial Designs .....	31
3.2.4	Repeated Measures Designs .....	31
3.2.5	BACI Designs .....	32
3.3	Impact Assessment Studies .....	33
<b>4</b>	<b>Field Methods and Data Issues .....</b>	<b>35</b>
4.1	Field Methods .....	35
4.1.1	Point Transect Sampling .....	36
4.1.2	Line Transect Sampling .....	37
4.2	Data Issues .....	42
4.2.1	Data Recording .....	42
4.2.2	Data Checking .....	43
4.2.3	Field Testing .....	48
 <b>Part II Design-Based and Model-Based Methods for Distance Sampling</b>		
<b>5</b>	<b>Modelling Detection Functions .....</b>	<b>53</b>
5.1	Model Selection and Goodness-of-Fit .....	53
5.1.1	Pooling Robustness .....	54
5.1.2	Information Criteria .....	56
5.1.3	Likelihood Ratio Test .....	58
5.1.4	Goodness-of-Fit Tests .....	58
5.2	Conventional Distance Sampling .....	61
5.2.1	Models for the Detection Function .....	61
5.2.2	Line Transect Sampling .....	64
5.2.3	Point Transect Sampling .....	74
5.2.4	Summary .....	82
5.3	Multiple-Covariate Distance Sampling .....	82
5.3.1	Adding Covariates to the Detection Function Model .....	82
5.3.2	MCDS Case Studies .....	84
5.4	Mark-Recapture Distance Sampling .....	93
5.4.1	Modelling the Detection Function .....	93
5.4.2	MRDS Case Studies .....	97
<b>6</b>	<b>Design-Based Estimation of Animal Density and Abundance .....</b>	<b>105</b>
6.1	Introduction .....	105
6.2	Plot Sampling .....	106
6.2.1	Quantifying Precision .....	106
6.2.2	The Montrave Case Study: Plot Sampling .....	107
6.3	Conventional Distance Sampling .....	109
6.3.1	Line Transect Sampling .....	109
6.3.2	Point Transect Sampling .....	113
6.3.3	Generalizing the Estimation Equations .....	115
6.4	Horvitz–Thompson-Like Estimators .....	116
6.4.1	The Horvitz–Thompson Estimator .....	116

6.4.2	The Horvitz–Thompson-Like Estimator for Conventional Distance Sampling .....	117
6.4.3	The Horvitz–Thompson-Like Estimator for Multiple-Covariate Distance Sampling .....	118
6.4.4	The Horvitz–Thompson-Like Estimator for Mark-Recapture Distance Sampling .....	123
<b>7</b>	<b>Model-Based Distance Sampling: Two-Stage Models .....</b>	<b>127</b>
7.1	Introduction .....	127
7.2	Plot Count Models .....	128
7.2.1	Stage One: Estimating Probability of Detection .....	128
7.2.2	Stage Two: Modelling the Counts .....	129
7.3	Plot Abundance Methods .....	132
7.4	Examples .....	132
7.4.1	The BACI Case Study .....	132
7.4.2	Density Surface Modelling: Spotted Dolphins in the Gulf of Mexico .....	135
7.4.3	The Nysted Windfarm Example .....	138
<b>8</b>	<b>Model-Based Distance Sampling: Full Likelihood Methods .....</b>	<b>141</b>
8.1	Introduction .....	141
8.2	Formulating the Likelihood .....	142
8.2.1	Conventional Distance Sampling .....	142
8.2.2	Multiple-Covariate Distance Sampling .....	144
8.2.3	Mark-Recapture Distance Sampling .....	146
8.2.4	Plot Count Models .....	146
8.2.5	Model Extensions .....	150
8.3	Model Fitting: Maximum Likelihood Methods .....	153
8.3.1	Maximum Likelihood Estimation .....	153
8.3.2	Estimating Precision .....	154
8.4	Model Fitting: Bayesian Methods .....	155
8.4.1	MCMC Algorithm .....	155
8.4.2	Model Averaging: Reversible-Jump MCMC .....	156
8.4.3	Bayesian Inference .....	157
8.5	Conservation Buffer Case Study .....	158
8.5.1	Indigo Buntings: A Full Likelihood Approach Using Maximum Likelihood Methods .....	158
8.5.2	Northern Bobwhite Surveys: A Bayesian Approach .....	160
 <b>Part III Distance Sampling Variations, Special Issues and Assumptions</b>		
<b>9</b>	<b>Variations on a Theme .....</b>	<b>167</b>
9.1	Trapping Webs and Trapping Line Transects .....	169
9.2	Trapping and Lure Point Transects .....	170
9.2.1	Lure Point Transect Case Study: Scottish Crossbills .....	171

9.2.2	Trapping Point Transect Case Study: Key Largo Woodrats .....	174
9.3	Indirect Surveys .....	179
9.3.1	Clearance Plot Method .....	180
9.3.2	The Standing Crop Method .....	181
9.3.3	Red Deer Dung Decay Rate Experiment .....	182
9.3.4	Chimpanzee Nest Survey .....	184
9.4	Cue Counting .....	187
9.4.1	Cue Counting for Whales .....	188
9.4.2	Cue Counting for Birds .....	189
9.5	Surveys in Three Dimensions .....	194
9.6	Radar, Sonar and Acoustic Surveys .....	195
9.6.1	Radar Surveys .....	195
9.6.2	Active Sonar Surveys .....	198
9.6.3	Passive Acoustic Surveys .....	199
<b>10</b>	<b>Taxon-Specific Issues .....</b>	<b>201</b>
10.1	Songbirds .....	201
10.1.1	Line Transect Versus Point Transect Surveys .....	201
10.1.2	Minimizing Bias from Bird Movement .....	202
10.1.3	Minimizing Bias from Failure to Detect All Birds on the Line or Point .....	205
10.1.4	Minimizing Bias from Inaccurate Distance Measurement .....	206
10.2	Seabirds .....	207
10.2.1	Shipboard Surveys .....	207
10.2.2	Aerial Surveys .....	208
10.3	Cetaceans .....	211
10.3.1	Shipboard Surveys .....	211
10.3.2	Aerial Surveys .....	215
10.3.3	Acoustic Surveys .....	216
10.4	Primates .....	220
10.4.1	Dealing with Primate Groups .....	221
10.4.2	Other Approaches .....	222
10.5	Ungulates .....	223
10.5.1	Ground-Based Surveys .....	223
10.5.2	Aerial Surveys .....	224
10.6	Butterflies .....	225
10.7	Plants .....	226
<b>11</b>	<b>Exchanging Assumptions for Data .....</b>	<b>231</b>
11.1	Non-random Transect Placement .....	232
11.1.1	Bias .....	232
11.1.2	Correcting for Bias .....	233
11.1.3	Example: Eastern Grey Kangaroos .....	235
11.1.4	Estimating Bias by Simulation .....	237
11.1.5	Other Approaches .....	239

11.2	Uncertain Detection on the Transect .....	240
11.3	Measurement Error Models .....	243
11.3.1	Bias Arising from Measurement Error .....	243
11.3.2	Accounting for Measurement Error .....	244
11.3.3	A Practical Example .....	246
11.3.4	Take-Home Messages .....	248
11.4	Addressing Bias from Animal Movement .....	249
11.4.1	Movement Independent of the Observer .....	249
11.4.2	Responsive Animal Movement .....	251
<b>12</b>	<b>Summary .....</b>	<b>253</b>
12.1	Check-List for a Good Survey .....	254
12.2	New Technology .....	257
12.2.1	Using Technology to Detect Animals .....	257
12.2.2	Data Recording .....	259
12.2.3	Distance Estimation .....	259
12.2.4	Aerial Surveys .....	260
12.2.5	Satellites and GPS .....	260
<b>Erratum</b>	.....	<b>E 1</b>
<b>References</b>	.....	<b>263</b>
<b>Index</b>	.....	<b>275</b>

**Part I**  
**Introduction, Survey Design**  
**and Field Methods**

# Chapter 1

## The Basic Methods

### 1.1 Introduction

One of the most common and pervasive questions in applied ecology relates to the size of a given population. How many animals are there? The question is intrinsically interesting, but perhaps even more important from an applied perspective, the actual answer has implications for most ecological processes affecting that population. The effective management of a population is not possible without knowing at least approximately how many individuals it includes. As an example, for a small population, a given mortality rate due to a newly introduced human disturbance might be important and a matter of concern, quickly leading the population to local extinction, but essentially not ecologically relevant for an abundant population. Therefore, knowledge about abundance is required to adequately interpret a wide variety of ecological processes affecting a given population.

Under all but the simplest of scenarios, total enumeration, known as a *census*, is not an option. Therefore, researchers depend on some kind of sampling method to answer this question, and will have to draw inferences on a population based on a sample. Conventional sampling ideas like randomness, replication and simple random sampling are important in this context. However, due to the special characteristics of wild populations, it is hard to implement conventional survey methods. One cannot simply divide the area of interest into a large number of sampling units and count all animals within a random selection of these. The difficulty will stem from three special characteristics of estimating abundance of wild populations. To begin with, we are interested in estimating the size of the population. Contrast that with most traditional sampling scenarios, where population size is assumed known, and we want to draw a random sample of individuals from the population to estimate a characteristic of the members of the population. For example, we might wish to estimate the mean height of students at a school based on selecting a small number of students at random. Secondly, even after selection of spatial sampling units, we cannot guarantee that all individuals present are detected. Thirdly, wild populations

exist in time and space but most are not static, and hence the traditional concept of sampling and sampling units gets blurred. Animals can move around and avoid observers, approach observers, *etc.* The relative speed (observer to animal speed) at which we observe the areas of interest will have an influence on the outcome of the sampling.

Because of these features, it is unsurprising that various methods have been developed explicitly to deal with them. A review of available methods is beyond the scope of this book, but the reader is referred to the excellent article by Schwarz and Seber (1999) for an overview. We note in passing that George A.F. Seber is one of the authors who contributed most to synthesizing the knowledge in this area with two earlier reviews (Seber 1986, 1992) and a book (Seber 1982) on the theme. The most commonly-used methods in this context are capture–recapture and distance sampling. While for the former we can think of sampling individuals, with information about detectability contained in the capture histories of different individuals, for the latter, we can think of sampling areas, with information on detectability being contained in the detection distances. The underlying idea to both approaches is that, if we can estimate the probability of detection for those objects of interest that we detect, we can also estimate how many were undetected. From that estimate of detectability, we can correct the observed counts for individuals missed on the sampled plots.

The term ‘distance sampling’ was introduced by Buckland et al. (1993) to include a suite of methods, including line and point transect sampling, in which animal density or abundance is estimated from a sample of distances to detected individuals. The first version of the *Distance* software (Laake et al. 1993) was released to allow users to implement the methods of the book.

*Introduction to Distance Sampling* (Buckland et al. 2001) updated the earlier book, and remains the standard reference for ‘conventional distance sampling’, corresponding with the *cds* engine of *Distance*. *Advanced Distance Sampling* (Buckland et al. 2004) addressed a range of advances, including multiple-covariate distance sampling (the *mcdds* engine of *Distance*), density surface modelling (the *dsm* engine of *Distance*), mark-recapture distance sampling for double-observer data (the *mrds* engine of *Distance*), and automated survey design (implemented in *Distance* with the aid of various Geographic Information System algorithms). The stand-alone version of *Distance* is summarized by Thomas et al. (2010).

As more advanced methods were developed, it became easier to implement them using the statistical software R. Initially, developments were implemented in *Distance*, with R called by *Distance*, so that the user was largely unaware of the role of R. However, to give users the full power of the *mrds* analysis engine with a simplified syntax and data structures, a version of *Distance* was released as an R package (Miller 2013).

Datasets and the means to implement methods (computer code or *Distance* projects) are provided for the book case studies at the book website:

<http://www.creem.st-and.ac.uk/DS.M&A/>.

Some of the case studies (Montrave, Sect. 1.3, and Hawaiian song birds, Sect. 5.3.2.1) are based upon the *Distance* software. The northern bobwhite

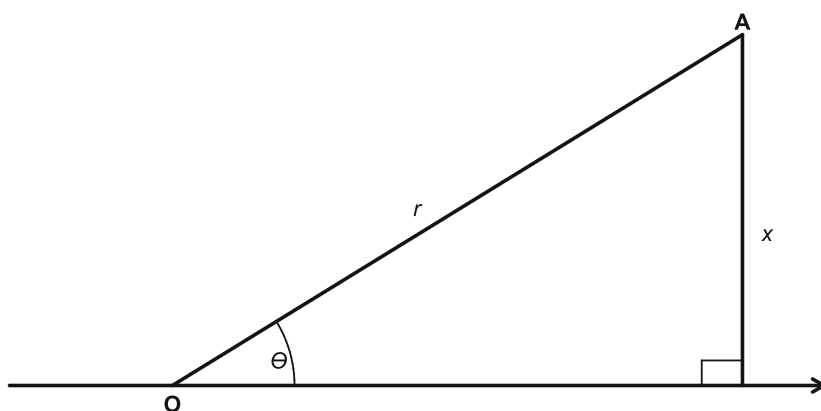
quail (Sect. 8.5.2) and Scottish crossbill (Sect. 9.2.1) case studies use specialized software and cannot be accomplished using `Distance`. Finally, the spotted dolphin (Sect. 7.4.2) and simulation (Sect. 2.5.2) case studies can be done either using `Distance` or using only R packages `dsm` and `DSsim`.

## 1.2 Line and Point Transect Sampling

There are two basic methods of distance sampling: line transect sampling and point transect sampling. All other methods may be considered as extensions of one or the other. In line transect sampling, a series of lines is distributed according to some design (usually a systematic grid of parallel lines, Chap. 2), and an observer travels along each line, searching for animals or animal *clusters*. We define a cluster to be a group of animals with a well-defined location for the group centre. For each animal or cluster detected, the observer measures or estimates the (perpendicular) distance  $x$  of the animal or cluster centre from the nearest part of the line. In many surveys, it is easier to measure or estimate the observer-to-animal (radial) distance  $r$  at the time the detection is made. If the sighting angle  $\theta$  is also measured, then the perpendicular distance may be found by simple trigonometry:  $x = r \sin \theta$  (Fig. 1.1).

For point transect sampling, the design comprises a set of points, often spaced on a systematic grid (Chap. 2). The observer travels to each point, and records any animal (or cluster) detected from the point, together with its distance  $r$  from the point.

Conceptually, distance sampling is a form of plot sampling in which not all animals on the plot are detected. In line transect sampling, the plots are rectangles of size  $2wl$  where  $l$  is the length of a given line (which may vary by line) and  $w$  is



**Fig. 1.1** An observer travels along the transect from left to right. When the observer is at point O, an animal is detected at A. The observer can record either the perpendicular distance  $x$  directly, or both the radial distance  $r$  and the sighting angle  $\theta$ , from which  $x = r \sin \theta$



the half-width of the strip; it is also the truncation distance for distances from the line, which might be pre-determined, so that detections further than  $w$  from the line are not recorded, or may be set by the analyst, so that distances greater than  $w$  are excluded from analysis. In point transect sampling, the plots are circles of size  $\pi w^2$ , where  $w$ , the plot radius, is again the truncation distance.

### 1.3 The Montrave Case Study

We will use the surveys of Buckland (2006) as a case study to illustrate standard distance sampling methods. We will refer to this study as the Montrave case study; it was conducted in an area of woodland and parkland at Montrave in Fife, Scotland, in the spring of 2004. Four species were recorded: common chaffinch (*Fringilla coelebs*), winter wren (*Troglodytes troglodytes*), great tit (*Parus major*) and European robin (*Erithacus rubecula*). Only territory-holding male birds were recorded, as some females may be undetectable even at close range. In addition to standard line and point transect sampling methods, a snapshot point transect method and a cue count survey were conducted (Buckland 2006). We use line transect and snapshot point transect data on the robin (Fig. 1.2) as a case study. We also use the robin data to illustrate both cue counting and plot sampling. For plot sampling, we truncate detections at 35 m, and assume that all male birds within that distance were detected.



**Fig. 1.2** The European robin is a strongly territorial species, occurring in woodland habitats throughout Europe. Photo: Steve Buckland

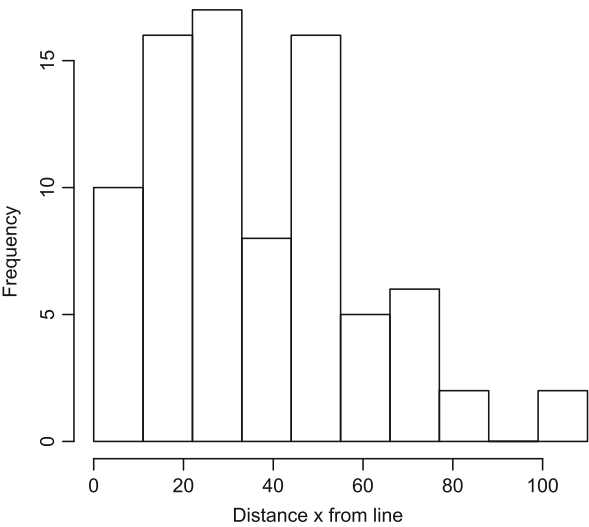
Survey design for this case study will be shown in Chap. 2. Estimation of the probability of detection will be addressed in Chap. 5, and estimating bird density will be covered in Chap. 6. For plot sampling only, density will be estimated in Sect. 6.2.2. A cue counting analysis is presented in Sect. 9.4.2.

We provide the data used to illustrate line transect sampling in Table 1.1. Note that no robins were detected from three of the lines. These lines must be included in the analysis; if they are not, mean density and hence abundance will be over-estimated. The distances are plotted in a histogram in Fig. 1.3.

The data for point transect sampling are shown in Table 1.2 and Fig. 1.4. Note the different shape of the histogram relative to Fig. 1.3. This arises because there is

**Table 1.1** The line transect data for European robins from the Montrave study

Line	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Length (m)	208	401	401	299	350	401	393	405	385	204	39	47	204	271	236	189	177	200	20
Distances	12	5	0	15	15	15	15	23	15	55		45	5	5	5		15	10	
	26	5	5	35	21	15	30	27	45			50	25	10	20		25	25	
	70	20	5	72	22	20	50	38	80				38	27	25		33	70	
		24	15	75	28	35	60	55					45	40	45		35		
		30	20	100	33	50	85	58					45	45					
		45	22		35	65		75						45					
		50	27		45	70		100											
		53	28		60														
				32															
				40															
				60															

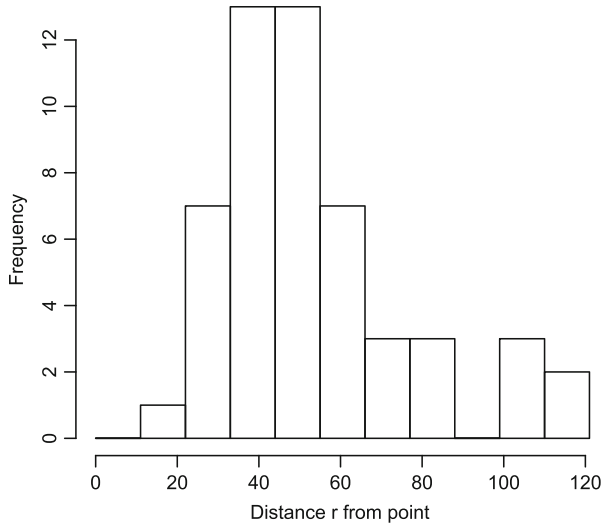


**Fig. 1.3** Distribution of robin detections by distance (line transect sampling)

**Table 1.2** The snapshot point transect data for European robins from the Montrave study

Point	1	2	3	6	7	8	9	11	13	14	15	16	19	20	21	22	24	25	26	27	28	30	31	32
Distances	100	35	40	40	115	60	40	65	42	45	35	45	20	55	27	30	25	45	25	35	55	100	25	25
			85		120	80			70	45	45	65	30	100	35	40	40	80	35	45				36
									70	60	55		65		45	70	50		40	60				
											55						60		50					

No robins were detected from points 4, 5, 10, 12, 17, 18, 23 and 29



**Fig. 1.4** Distribution of robin detections by distance (snapshot point transect sampling)

very little area within a small distance of a point, and so few birds are detected close to a point. Thus numbers of detections increase with distance, until the distance is sufficiently large that we fail to detect many of the birds.

## 1.4 Plot Sampling

### 1.4.1 Estimating Density and Abundance

Consider first the case that all animals on the sampled plots are detected. For line transects, we term this special case *strip transect sampling*, and for point transects, *circular plot sampling*. To estimate animal density  $D$  and hence abundance  $N = DA$ , where  $A$  is the size of the study region, we simply take the total number  $n$  of animals detected across the plots, and divide by the total size  $a$  of the plots:  $\hat{D} = n/a$  where the hat (  $\hat{\phantom{x}}$  ) indicates that this is an estimate based on our sample data.

The corresponding abundance estimate is  $\hat{N} = \frac{nA}{a}$ . For strip transect sampling with  $K$  lines, we have  $a = 2wL$  where  $L = \sum_{k=1}^K l_k$  is the total length of the  $K$  lines. Thus we have

$$\hat{D} = \frac{n}{2wL} \quad (1.1)$$

and

$$\hat{N} = \frac{nA}{2wL} . \quad (1.2)$$

For circular plot sampling with  $K$  points,  $a = K\pi w^2$  so that

$$\hat{D} = \frac{n}{K\pi w^2} \quad (1.3)$$

and

$$\hat{N} = \frac{nA}{K\pi w^2} . \quad (1.4)$$

### 1.4.2 Limitations of Plot Sampling

Because we assume that all animals on each plot are detected, either we must make the plots small, so they can be fully surveyed from the centre point or line, or effort must be expended searching the plot. The more time that is spent on the plot, the greater the bias arising from new animals entering the plot. Active searching of the plot may also disturb animals, which may move off the plot undetected; and it can be difficult to keep track of those animals already counted. Whether plots are small, and searched quickly, or large, and searched more slowly, many detected animals are likely to be beyond the plot. These detections are ignored, reducing precision. Further, it can be difficult to judge whether an animal close to a boundary, especially an animal that is moving around, is inside or outside the plot; observers have a tendency to include such animals in their counts, causing upward bias in abundance estimates. For standard line or point transect sampling, for which probability of detection at the plot boundary is typically low, such bias causes some upward bias in estimated probability of detection at the boundary, but there is a corresponding upward bias in the estimate of the effective size of the plot surveyed, and the two tend to cancel, so that there is little if any bias in resulting abundance estimates.

A further limitation of plot sampling is that it can be very difficult to judge whether the observer has successfully detected all animals on each plot. If distances of detected animals from the line or point are recorded, then we can explore whether numbers at larger distances are fewer than we would expect, if all animals are being detected. However, if we are recording these distances anyway, then we can make full use of the data by using line and point transect methods.

In some studies, no attempt is made to detect all animals on the plots, yet the plot counts are not corrected for detectability, on the grounds that only a measure of relative abundance is needed, given the study objectives. This rationale is valid provided detectability does not vary in any systematic way across plots or over time. Given that detectability is not estimated, this assumption typically cannot be tested. Further, studies show that detectability varies by many factors, including habitat, weather conditions, animal behaviour, time of day, time of year, observer experience, observer age, *etc.* It seems preferable therefore to adopt methods that allow detectability to be estimated and controlled for, if it is not practical to carry out complete counts on plots.

## 1.5 Line Transect Sampling

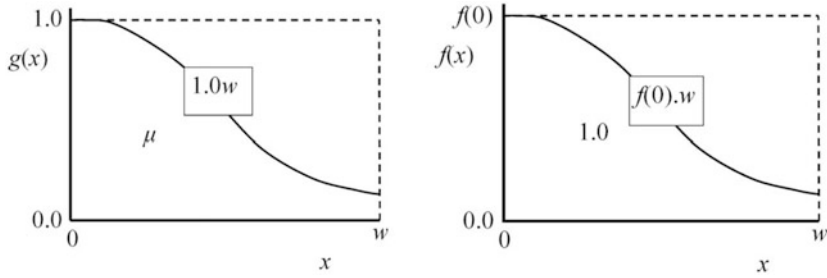
In conventional line transect sampling, we assume that all animals on the line are detected, but that probability of detection falls off smoothly with distance from the line. We use the recorded distances  $x$  of detected animals from the line to model the *detection function*,  $g(x)$ , which is defined to be the probability of detecting an animal that is at distance  $x$  ( $0 \leq x \leq w$ ) from the line. For the standard method, we assume that animals on the line are certain to be detected:  $g(0) = 1$ .

We may regard line transect sampling as plot sampling, in which the plots are long, narrow strips of half-width  $w$ , centred on the transect line, and for which not all animals on a plot are counted. For those animals on a plot, we denote the expected proportion counted by  $P_a$ . Thus for strip transect sampling,  $P_a = 1$ , but for line transect sampling,  $0 < P_a < 1$ . We show in Fig. 1.5 that  $P_a = \mu/w$  where  $\mu = \int_0^w g(x) dx$ . Hence, given an estimate  $\hat{g}(x)$  of the detection function  $g(x)$ , we can obtain an estimate  $\hat{P}_a$  of  $P_a$ . Equation (1.1) of strip transect sampling is replaced by:

$$\hat{D} = \frac{n}{2wL\hat{P}_a}. \quad (1.5)$$

Thus if we estimate that one half of the animals on the plot of half-width  $w$  are counted, we double the density estimate that would be obtained if we were to assume that all animals on the plot were counted. This is equivalent to estimating density under the assumption that our count of  $n$  animals was a complete count of animals on a strip of half-width  $\mu$ , which has led to  $\mu$  being termed the effective strip half-width.

We need a method for estimating  $g(x)$ . As we will see in Sect. 5.2.2.1, we can derive the *probability density function* corresponding to the distances  $x$  simply by rescaling  $g(x)$  so that it integrates to one — a property that all probability density functions share. Thus  $f(x) = g(x)/\mu$  for  $0 \leq x \leq w$ , and in particular,  $f(0) = 1/\mu$  because we assume that  $g(0) = 1$ . The advantage of this is that a probability density function represents the relative likelihood of different observations, and we have general methods for fitting such functions (Sect. 5.2.2.1). Having obtained the



**Fig. 1.5** The probability  $P_a$  that an animal within distance  $w$  of a line is detected may be represented by the proportion of the rectangle that is under the curve in these plots. If the curve is the detection function  $g(x)$  (left plot), this gives  $P_a = \mu/w$  where  $\mu = \int_0^w g(x) dx$  is the area under the curve, and the area of the rectangle is  $g(0).w = w$ , as  $g(0) = 1$ . For the probability density function (right plot), the area under the curve is 1.0 (a property of all probability density functions), and so  $P_a = 1/[w.f(0)]$ . Hence the alternative forms for  $\hat{D}$  in Eq. (1.6)

estimated function  $\hat{f}(x)$ , and evaluated it at  $x = 0$  to give  $\hat{f}(0)$ , we can write

$$\hat{D} = \frac{n}{2wL\hat{P}_a} = \frac{n}{2\hat{\mu}L} = \frac{n\hat{f}(0)}{2L}. \quad (1.6)$$

The conceptual relationship between  $P_a$ ,  $\mu$  and  $f(0)$  is shown in Fig. 1.5.

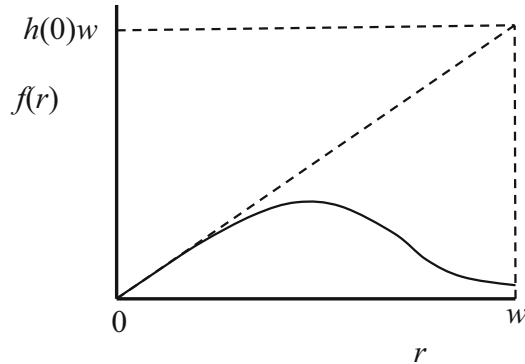
We address modelling of the detection function in greater detail in Sect. 5.2.2.1, and the estimation of density and abundance in Sect. 6.3.1.

## 1.6 Point Transect Sampling

As for line transect sampling, we start by estimating the detection function, which we now represent by  $g(r)$ , where  $r$  is distance from the point. Thus  $g(r)$  is the probability that an animal at distance  $r$  from a point is detected ( $0 \leq r \leq w$ ). We again assume  $g(0) = 1$ . Having fitted a model for the detection function, we can estimate the proportion  $P_a$  of animals available for detection (*i.e.* within distance  $w$  of a point) that are detected. Then Eq. (1.3) is extended as:

$$\hat{D} = \frac{n}{K\pi w^2 \hat{P}_a}. \quad (1.7)$$

As with line transect sampling, for the purpose of estimating  $P_a$ , it is convenient to work with the probability density function. However, unlike line transect sampling, this function does not now have the same shape as the detection function. We will see in Sect. 5.2.3 that the probability density function  $f(r)$  of detection distances  $r$  from the point is proportional to the detection function  $g(r)$  multiplied by distance



**Fig. 1.6** The probability  $P_a$  that an animal within distance  $w$  of a point is detected may be represented by the proportion of the triangle that is under the curve in this plot. The area under the curve is one, as the curve is a probability density function. The triangle has base  $w$  and height  $h(0)w$  where  $h(0)$  is the slope of the hypotenuse, so that its area is  $h(0)w^2/2$ , giving  $P_a = 2/[h(0)w^2]$ . Note that, if all animals out to distance  $w$  were detected, then the probability density function would be represented by the triangle, and we would have  $P_a = 1$

$r$ . As illustrated in Fig. 1.6, we have  $P_a = 2/[h(0)w^2]$  where  $h(0)$  is the slope of the probability density function evaluated at zero distance. Thus we can write Eq. (1.7) as:

$$\hat{D} = \frac{n}{K\pi w^2 \hat{P}_a} = \frac{n\hat{h}(0)}{2K\pi}. \quad (1.8)$$

We give details of estimating the detection function in Sect. 5.2.3, and we address density and abundance estimation in Sect. 6.3.2.

## 1.7 Assumptions

There is one key design assumption of distance sampling and three key model assumptions. These four assumptions are:

1. Animals are distributed independently of the lines or points.
2. Objects on the line or at the point are detected with certainty.
3. Distance measurements are exact.
4. Objects are detected at their initial location.

Most extensions of the above theory are to address failure of one or more of these assumptions. The first assumption is the key design assumption. Stated more fully, we assume that animals are distributed uniformly with respect to distance from the line (so that  $\pi(x) = 1/w$ ) for line transect sampling or that animals are

distributed according to a triangular distribution with respect to distance from the point (so that  $\pi(r) = 2r/w^2$ ) for point transect sampling. We call this a design assumption because a suitably randomized design will ensure that the assumption holds. (So-called plus-sampling into a bufferzone around the study region ensures that edge effects do not cause violation of this assumption (Strindberg et al. 2004), but for large study regions, edge effects tend to be negligible if this issue is ignored.) Thus lines or points should be positioned according to a random design. Usually, a systematic grid of lines or points is placed randomly over the study region.

Model assumptions (assumptions 2–4 above) are usually of greater concern, because we cannot ensure that they hold by adopting a suitable design. Instead we need to consider whether field methods can be adopted that will ensure low bias when they fail. Assumption 4 is related to animal movement, which may be in response to the observer or independent of the observer. If animals show responsive movement, then field methods should if possible ensure that animals are detected (and their locations recorded) before they respond. Movement that is independent of the observer should be slow relative to observer speed. If this is not possible (for example for point transect sampling, in which the observer does not move), then the standard methods may require modification. In Chap. 11, we will consider methods to address failures of each of these three model assumptions.

Other assumptions are usually of little consequence. We assume that detections are independent, but adopt estimation methods that are very robust to failures of this assumption, although assessing model fit is more problematic if there is strong violation. For clustered populations, we assume that the size of detected clusters is recorded without error. Using the default method in `Distance` for estimating mean cluster size, we can relax this assumption, so that for clusters on or close to the line or point, cluster size estimates are assumed to be unbiased. Another minor assumption is that, having detected an animal, the observer is able to avoid counting it a second time from the same line or point. Counting the same animal across multiple lines because it has moved between surveying one line and the next should not be a problem, provided that the animal movement is not in response to the observer. In the absence of directed movement such as migration, on average for every animal moving from line 1 to line 2, there is another animal that moves the other way, so that any effect tends to cancel out.



## Chapter 2

# Designing Surveys

Survey design is covered in greater detail by Buckland et al. (2001, pp. 230–248) and by Strindberg et al. (2004). We concentrate here on basic designs, and on issues that commonly need to be addressed. Some of those issues relate to analysis; in such cases, we provide forward references to where they are addressed. Although through most of the book, we describe line transect methods before point transect methods (for which analysis has greater complexity), for survey design, we consider point transect sampling first. This is because point transect surveys are simpler to design than are line transect surveys.

### 2.1 Point Transect Survey Designs

A standard point transect survey comprises points spanning the entire study area, and systematically spaced, so that they form a grid. The Montrave study is an example of this design: points are at the intersections of a grid comprising squares of side 100 m (Fig. 2.1).

In some studies, a simple random sample of units is preferred to a systematic sample. In this case, it is again convenient to define a grid of squares, so that the length of the side of each square is at least  $2w$ , where  $w$  is the distance from the point beyond which either detected animals are not recorded, or such detections are truncated at the analysis stage. (In the latter case, this value might be assessed from similar past studies.) A simple random sample of grid squares is then selected, and a point placed in the middle of each selected square. This ensures that the circular plots around selected points cannot overlap. See Fig. 7.2 of Strindberg et al. (2004).

For large study areas with wide separation between points, the above systematic design is inefficient, as the observer must travel a large distance from one point to the next, and then record just animals detectable from that point. In this case, cluster sampling is a better option: at each selected point, locate a small grid of points, with

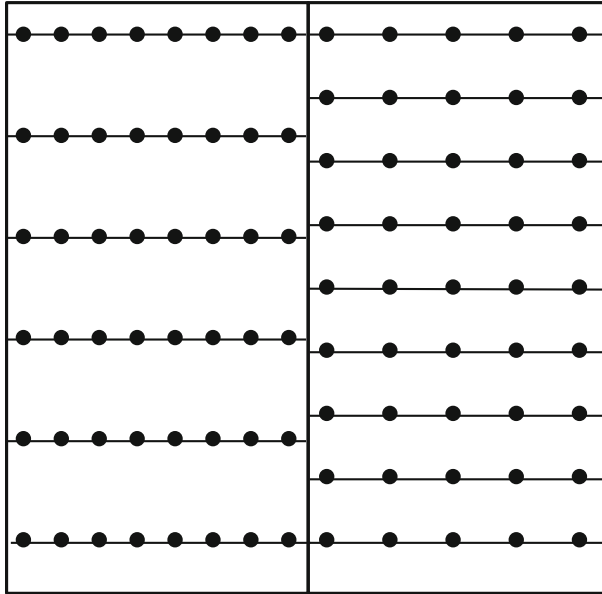


**Fig. 2.1** The survey design for the Montrave study. The surveyed woodland and parkland totalled about 33 ha in two blocks, surrounded by grazing and crops. For the point transect surveys, a systematic grid of 32 points (indicated by *circles*) was placed over the study area, with 100 m separation between neighbouring points. Each point was visited twice. For the line transect survey, there were 19 lines (11 in one block and eight in the other), with about 70 m separating successive lines. Each line was surveyed twice

separation distance between neighbouring points in the vicinity of  $2w$ . For analysis, each cluster (*i.e.* grid of points) will be treated as a single sampling unit.

Often, points are positioned along lines. If the separation of neighbouring points on the same line is less than the separation of neighbouring points on different lines, then this should be treated as cluster sampling, where each line represents one cluster of points. If the separation is the same, then the surveyed points form a regular grid through the study area, and each point may be treated as a sampling unit. See Fig. 2.2.

When points are placed at random within a study area, some will fall within distance  $w$  of the study area boundary. In such cases, the intersection of the circular plot with the study area is no longer a complete circle, which creates an ‘edge effect’ on animal availability. Such edge effects cause bias in abundance estimates.

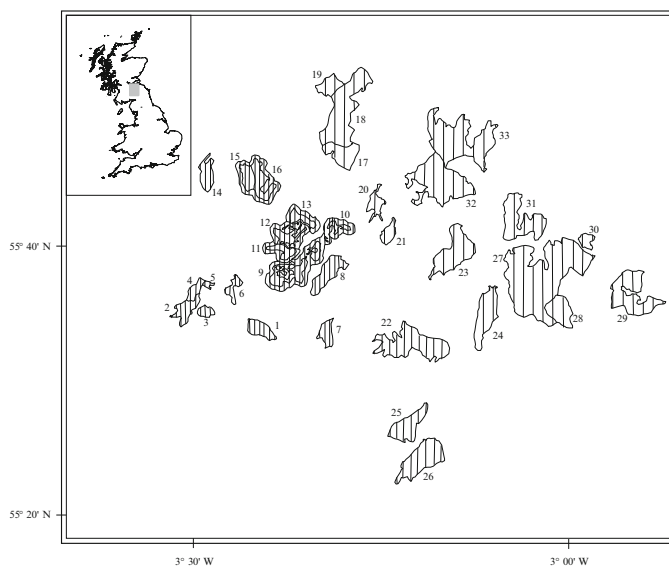


**Fig. 2.2** Point transect designs with points along lines. In the *left-hand plot*, points are spaced closely along lines, with large spacing between successive lines. For estimating precision, each line should be treated as a sampling unit, with points on a single line treated as a cluster sample. In the *right-hand plot*, if the lines are removed, it leaves a systematic sample of points, equally spaced through the area. For such a design, individual points are taken to be the sampling units

For large study areas, bias is generally negligible, and can be ignored. For very small study areas, or for study areas that are long and narrow (*e.g.* riparian habitat), the bias may be significant. An easy way to avoid bias is to include a bufferzone around the study area, extending a distance  $w$  beyond the study area boundary, and extend the sampling grid into this bufferzone. Points in the bufferzone are surveyed, but detected animals are only recorded if they are within the study area. Although this is an easy solution to implement, it can add significantly to the cost of the survey. Another solution is to survey only those points within the study area, and to choose a smaller value for the truncation distance  $w$  than would be normal. See Buckland (2006) for further information.

## 2.2 Line Transect Survey Designs

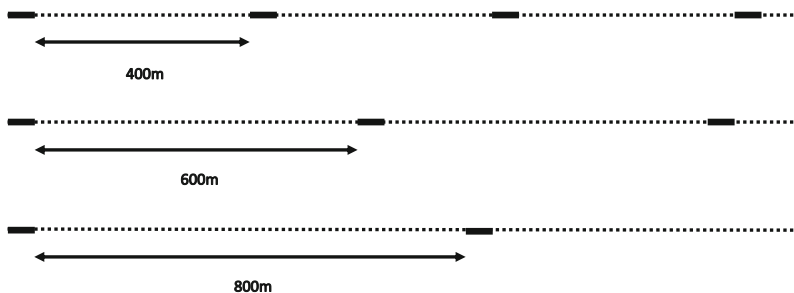
The ‘standard’ design for a line transect survey depends on whether it is practical or efficient to survey lines that cross the entire study area. The Montrave study is an example of a small study area, for which it is feasible to run each line from one boundary to the opposite boundary. In this case, the standard design is a set of parallel lines, systematically spaced (Fig. 2.1).



**Fig. 2.3** Map of Peeblesshire in southern Scotland, showing the 33 forest blocks surveyed to estimate abundance of sika deer. Also shown is the grid of parallel lines along which transects were placed. In high-density blocks, the parallel lines were spaced 400 m apart. In blocks of medium density, the separation was 600 m, and 800 m in low-density blocks. Reproduced from Marques et al. (2001). ©British Ecological Society

If it is more practical to have short lines relative to the dimensions of the study area, then a grid of short lines may be randomly superimposed over the study area. In Fig. 2.3, taken from Marques et al. (2001), we show the design used for estimating density of sika deer (*Cervus nippon*) pellet groups (from which deer density may be estimated, using the indirect survey methods of Sect. 9.3). For each survey block, it shows a grid of parallel lines, with each line spanning the width of the block. However, the full length of each line was not surveyed. Instead, 50 m lengths of each line were surveyed, with gaps between. The length of the gap was chosen to coincide with the separation between successive parallel lines (Fig. 2.4), so that the surveyed strips form an evenly-spaced grid across each block, allowing the individual 50 m sections of line to be used as sampling units, giving better estimates of precision.

In many shipboard and aerial surveys, there is significant cost in travelling from one transect to the next. In this circumstance, it is common to use a zigzag (or saw-tooth) design so that one line starts where the previous line ends. When width of the study area varies, it is problematic to generate zigzag designs that avoid bias. Strindberg and Buckland (2004) address this issue. We reproduce a figure of theirs that illustrates different methods in Fig. 2.5. This shows a sector of the Antarctic with strata defined for surveys of southern minke whales (*Balaenoptera bonaerensis*). The southern stratum has a complex boundary along the ice edge.



**Fig. 2.4** Diagram showing spacing between transects (*solid lines*) along lines (*dotted lines*). The *top line* shows the spacing used for high-density blocks in Peeblesshire, corresponding to the 400 m separation between successive lines, the *middle line* shows the spacing for medium-density blocks, and the *bottom line* for low-density blocks

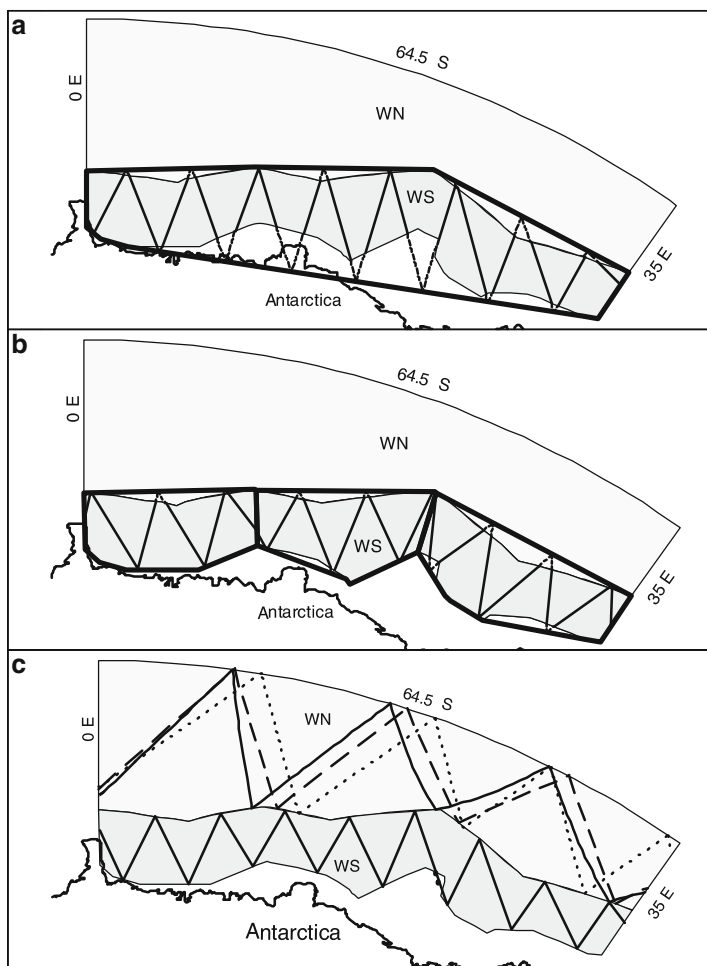
As for point transect surveys, edge effects are only of concern for small study areas. Bias from edge effects also tends to be smaller than for point transect surveys. A simple way to reduce or eliminate edge effects is to include a bufferzone around the study area, and project any lines that intersect the study area boundary into the bufferzone. The projected line is then surveyed out to a distance that it is no longer possible to detect animals within the study area. (Any animals that are located and detected in the bufferzone are not recorded.) See Fig. 7.5 of Strindberg et al. (2004). This method was implemented for the line transect surveys conducted as part of the Montrave study (Buckland 2006).

## 2.3 Stratification

Many of the uses of stratification in distance sampling have become less relevant since the advent of multiple-covariate distance sampling (Sect. 5.3). For example Buckland et al. (2001) considered stratifying by environmental conditions, by cluster size for clustered populations, by habitat, and by observer or platform. All of these are usually better dealt with using multiple-covariate distance sampling methods. Here we concentrate on geographic stratification, but also consider briefly post-stratification.

### 2.3.1 Geographic Strata

A stratified design is one that has been divided into units. In distance sampling, these units are most commonly geographical ‘blocks’ — strata that together comprise the study region. There are several reasons why strata might be used, and we discuss these in turn.



**Fig. 2.5** Two strata (WS and WN) used for surveying southern minke whales in the Antarctic. (a) A single realization of a zigzag sampler is shown in the convex hull of WS. The sampler is discontinuous (because sections outside the survey region are not covered) and thus inefficient. (b) WS has been divided into three almost convex sub-regions, and a zigzag sampler placed within the convex hull of each, reducing the sampler discontinuity. (c) A zigzag sampler is generated in the nonconvex WS giving the least sampler discontinuity in this instance. Within WN three types of zigzag sampler are illustrated: equal-angle (*dotted line* — this method has a constant angle with a defined main axis, and places too much effort in narrow sections and too little in wide sections when width of the survey area varies), equal-spaced (*dashed line* — based on equally-spaced waypoints along the boundaries, this method has much lower bias) and adjusted-angle (*solid line* — this method is designed to be unbiased, but angle varies continuously, so that transects are not straight). Reproduced from Strindberg and Buckland (2004) with kind permission of Springer Science+Business Media

*Density or abundance estimates may be needed for individual strata, as well as for the entire study area.* By making any sub-area of interest into a stratum in the design, design-based estimates of density and abundance may be obtained for each stratum. Commonly, there are too few data in at least some strata to fit a separate detection function in each stratum. One solution to this is to assume a common detection function across strata. However, if the assumption that all strata have the same detection function is false, stratum-specific estimates of abundance will be biased. The pooling robustness property (Sect. 5.1.1) ensures that total abundance (summed across strata) is approximately unbiased, provided the sampling intensity is the same in each stratum.

A better solution to small sample size per stratum is to assume a common detection function model across strata, but to allow the scale parameter of the model to differ among strata. This can be achieved by using multiple-covariate distance sampling with stratum as a factor-type covariate (Sect. 5.3).

Akaike's information criterion (Sect. 5.1.2) may be used to select between the three options: a separate detection function for each stratum; a single detection function for all strata; or the intermediate option of a single detection function model across strata, in which stratum is a factor-type covariate.

We note that if a model-based approach is used (Chaps. 7 and 8), we can obtain density or abundance estimates in any section of the study area. For the model-based approach, a density surface is estimated which models how density varies through the study region. Integrating under sections of that surface allows us to obtain density or abundance estimates for the respective section. Adopting this approach, stratification of the design is unnecessary, although the back-up and simpler option of a design-based approach, in case the model-based approach fails, may then be more problematic; splitting the effort and sightings retrospectively into strata (termed post-stratification, covered in Sect. 2.3.2) might result in too little effort in a stratum of interest.

The sika deer example of Fig. 2.3 is one in which separate reliable estimates of density are needed for every stratum (forest block), so that an assessment of the number of deer to cull in each block can be made, if the spread of this introduced species is to be halted.

*Density may vary appreciably over the study region.* If we know in advance where densities are high or low, we may wish to avoid expending substantial search effort in low-density areas, where few animals will be detected. In the survey of sika deer pellet groups (Fig. 2.3), forest blocks were assigned to three categories (high, medium and low density) on the basis of prior knowledge, and greater survey effort assigned in areas of higher density. In Antarctic minke whale surveys, relatively greater effort is assigned to the smaller ice-edge strata, where whale density is high (e.g. Fig. 2.5).

One downside of placing greater effort in areas of high density is that less effort is placed in areas where detections may in any case be few. Thus it may not be possible to quantify reliably the abundance in low-density strata. If low-density strata account only for a small percentage of total abundance, and stratum-specific estimates are

not needed, this is unlikely to be a problem. However, sometimes low-density strata are large compared with high-density strata (*e.g.* Fig. 2.5), and in that circumstance, a substantial proportion of the total population may be in those strata. In those cases, care should be taken to ensure sufficient detections in the low-density strata. As noted above, multiple-covariate distance sampling in which strata are taken as a factor-type covariate can reduce these problems.

Another factor to consider is whether there is a single objective of estimating total abundance within the study area, or whether there are other objectives, for example separate abundance estimates by stratum, modelling the distribution of animals through the study area, relating animal density to habitat, *etc.* In this case, we may prefer to allocate effort in proportion to stratum size, rather than assign more effort where densities are high.

*Different parts of the study region might have very different characteristics.* By defining these parts to be different strata, we have greater flexibility to ensure that the design and field methods are appropriate for the stratum. For example detectability may be lower in woodland habitats than in open habitats, so that greater sampling intensity may be needed in woodland, or field methods to ensure that the key assumptions are met may change between woodland and open environments.

### 2.3.2 *Post-stratification*

Sometimes, the design is unstratified, but it is convenient to ‘post-stratify’ the data before analysis. For example, we may need separate estimates of the male and female components of a population, and different detection functions might be appropriate for the two sexes. Post-stratification allows us to obtain the estimates we need, even though we did not stratify the design. Another example is if we wish to estimate animal density in each main habitat type, but we do not have the habitat types mapped to allow geographic strata to be defined at the design stage. In this case, habitat type may be recorded during the survey, and transects split into sections corresponding to the different habitat types, thus allowing post-stratification of the data. (In this example, we can readily estimate density for each habitat type, but we cannot obtain corresponding abundance estimates, as we do not know the area of each habitat type in the study region. Given enough lines, we could estimate the proportion of the total area that is of each type from the proportion of total line length that passes through each type, and hence estimate abundance corresponding to each habitat type. However, such estimates may not be useful if animals frequently move between habitat types, especially if one habitat type is preferred during daylight when surveys are carried out, and another habitat type at night.)



## 2.4 Survey Effort

In determining how much survey effort is needed, there are two principal considerations: what is the minimum sample size for reliable estimation, and what precision is required to meet the survey objectives? We consider each of these issues below.

### 2.4.1 *Minimum Sample Size*

There are two components to sample size: the number of line or point transects in the design, and the number of animals (or animal clusters for clustered populations) detected. Buckland et al. (2001, p. 232) recommend a minimum of 10–20 replicate lines or points to allow reliable estimation of variance of encounter rate. If the population is very patchily distributed, then more lines or points are needed, to ensure that the variability through the study area is adequately represented and estimated. Buckland et al. (2001, pp. 240–241) also suggest that at least 60–80 animals (or animal clusters) should be detected for reliable estimation of the detection function for line transect sampling, and at least 75–100 for point transect sampling. Analyses are possible with smaller numbers of detections, but it can then be a matter of luck whether the detection function is well-modelled. For example, when sample size is small, just two or three additional sightings on or very close to the line or point can affect estimation appreciably.

The advent of multiple-covariate distance sampling (Sect. 5.3) has resulted in the relaxation of sample size demands in some circumstances. For example, if the design is stratified, and detectability varies by stratum, the above numbers of detections apply to each stratum if a separate detection function is fitted to the data in each stratum. If however we use a single detection function model across strata, with stratum as a factor-type covariate, then we just need 60–80 detections in total for line transect sampling, and 75–100 for point transect sampling. Further, for a rare species for which numbers of detections are small, it may be possible to pool data with those for a more common but similar species, including species as a factor-type covariate in the model (Alldredge et al. 2007; Marques et al. 2007). This approach is illustrated in Sect. 5.3.2.2.

### 2.4.2 *Estimating Total Effort Required*

In the following, we assume that a target coefficient of variation has been identified. The coefficient of variation of an estimate  $\hat{D}$  of animal density  $D$  is defined as  $cv(\hat{D}) = se(\hat{D})/\hat{D}$ , where  $se(\hat{D})$  is the standard error of  $\hat{D}$  (see Sect. 6.3.1.2 for line transect sampling and Sect. 6.3.2.2 for point transect sampling). We do not give separate equations for populations in which animals occur in clusters because the

additional variation in estimating mean cluster size is typically small, and so has little effect on estimates of effort needed. See Buckland et al. (2001, pp. 240–248) for a more complete treatment, including formulae for clustered populations and for stratified designs.

#### 2.4.2.1 Point Transect Sampling

Suppose that we conduct a pilot survey, in which  $K_0$  points are surveyed. If this yields sufficient data to estimate density  $D$  and its coefficient of variation by  $\hat{D}$  and  $cv(\hat{D})$  respectively, then a simple method to estimate  $K$ , the number of points required to achieve the target coefficient of variation, denoted by  $cv_t(\hat{D})$ , is to take

$$K = \frac{K_0 \{cv(\hat{D})\}^2}{\{cv_t(\hat{D})\}^2} . \quad (2.1)$$

Pilot surveys are seldom sufficiently large to allow this approach. Suppose we detected  $n_0$  animals from the  $K_0$  points in the pilot survey, where  $n_0$  is too small to allow estimation of  $D$ . We can now take

$$K = \frac{K_0}{n_0} \frac{b}{\{cv_t(\hat{D})\}^2} . \quad (2.2)$$

Buckland et al. (2001, p. 245) recommend setting  $b = 3$ .

#### 2.4.2.2 Line Transect Sampling

The formulae for line transect sampling are very similar. First suppose that a pilot survey, conducted along lines of total length  $L_0$ , provides sufficient data to estimate density  $D$  and its coefficient of variation by  $\hat{D}$  and  $cv(\hat{D})$  respectively, and the target coefficient of variation is again  $cv_t(\hat{D})$ . Then we estimate the total length of line for the main survey as

$$L = \frac{L_0 \{cv(\hat{D})\}^2}{\{cv_t(\hat{D})\}^2} . \quad (2.3)$$

If the pilot data are insufficient for this method, and  $n_0$  animals are detected, we instead take

$$L = \frac{L_0}{n_0} \frac{b}{\{cv_t(\hat{D})\}^2} , \quad (2.4)$$

where again we can set  $b = 3$  (Buckland et al. 2001, p. 242).

## 2.5 Comparing Designs

### 2.5.1 Assessing Designs by Simulation

Different designs will have different costs associated with them, and will vary in efficiency. If appropriate randomization is not implemented, the design can also lead to substantial bias in abundance estimates. When designing a survey, it is useful to carry out simulations to assess the efficiency and bias associated with different design options. The R package `DSsim` (Marshall 2014) and version 7 of `Distance` allow such simulations to be carried out.

To assess designs by simulation, the study area must first be defined. A description of the population is also required, including the spatial distribution from which animal locations will be generated within the study area. As the true spatial distribution of animals is unknown, typically locations are generated from several different models, selected to span the plausible range for the true distribution. Population size  $N$  can be fixed for any given set of simulations, or it can be given a Poisson distribution, based on a specified spatial distribution of animal densities. Potential survey designs are then specified, from which transects can be generated.

Having generated the survey transects, the detection process is simulated. A model for the detection function is defined by the user, and for each animal within distance  $w$  of a line or point, its probability of being detected is determined by the detection function (see Chap. 5). Thus if an animal is at distance  $y \leq w$  from the nearest line or point, then its probability of detection is set by evaluating the detection function at distance  $y$ . The resulting distance data (*i.e.* the distances of detected animals from the line or point) are analysed, to estimate abundance  $N$  in the study area (and any other parameters of interest). The user may select multiple models to be fitted to each set of simulated distance data, and the model with the lowest value for AIC (default), AICc or BIC (Sect. 5.1.2) will be selected.

The process is repeated, so that for each model of the spatial distribution of animals, and for each design, we have a number say  $B$  of realizations of the animal distribution and of the design, and the corresponding estimates  $\hat{N}_b$  of  $N$ ,  $b = 1, \dots, B$ . Typically, we might choose a value of  $B$  between 100 and 1000. If the design is not randomized but selected subjectively, exactly the same design is used in all  $B$  simulations.

We can summarize the results to assess the performance of each design for each model of the population. Denote the mean of the simulated estimates by

$\hat{N} = \sum_{b=1}^B \hat{N}_b / B$ , and the standard deviation by  $\text{sd}(\hat{N}) = \sqrt{\frac{\sum_{b=1}^B (\hat{N}_b - \hat{N})^2}{B-1}}$ . Then

more efficient designs will give smaller values of  $\text{sd}(\hat{N})$ . Further, if both the design and the method of analysis are good, we should find that  $\text{sd}(\hat{N})$  is close to the mean of the  $B$  standard errors of  $\hat{N}$  calculated when analysing the  $B$  datasets.

A design that generates a small value of  $\text{sd}(\hat{N})$  has high efficiency, but this is of limited use if it also gives biased estimates of abundance. The mean  $\hat{N}$  should be close to  $N$ . We can estimate the percent bias by  $\frac{\hat{N}-N}{N} \times 100\%$ . Further, if

$$|\hat{N} - N| > 1.96\text{se}(\hat{N}), \quad (2.5)$$

where  $\text{se}(\hat{N}) = \text{sd}(\hat{N})/\sqrt{B}$ , then approximately we have evidence for bias at the 5 % level; conversely, if  $|\hat{N} - N| \leq 1.96\text{se}(\hat{N})$ , our design does not generate significant bias at the 5 % level. (The vertical bars in these expressions indicate that we take the positive value of the difference.)

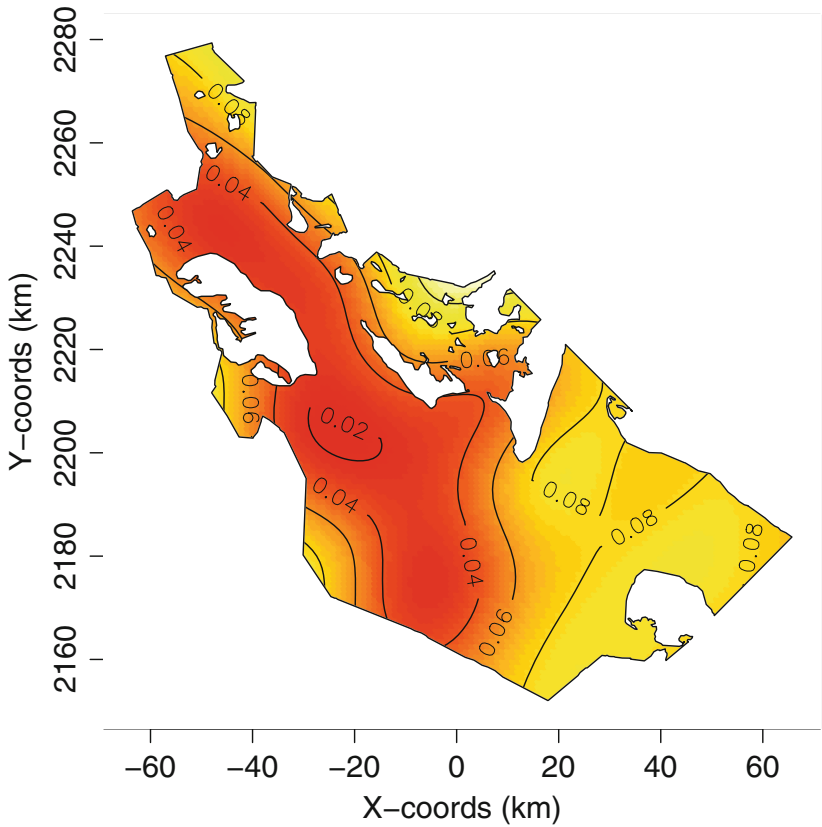
### 2.5.2 Example: Comparing the Efficiency of Parallel-Line and Zig-Zag Designs

We used `DSsim` to compare the efficiencies of three designs for surveying a simulated population. The study area is shown in Fig. 2.6. Also shown on this map is the density surface used to generate a population of fixed size  $N = 1500$  animals.

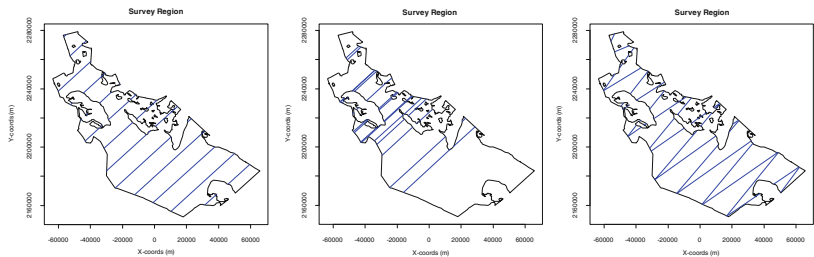
The three designs considered were: a systematic sample of parallel lines with a random start; a simple random sample of parallel lines; and a zigzag sample, based on equally-spaced waypoints (see Fig. 2.5). One hundred sets of transects were generated for each design, and the survey process was simulated for each replicate assuming a half-normal detection function with scale parameter  $\sigma = 500$  m, and truncation distance  $w = 1000$  m. A single realization of each design is shown in Fig. 2.7.

We summarize the performance of each design in Table 2.1. Although the total distance covered by each method is approximately the same, we see that the distance covered while ‘on effort’ (*i.e.* while surveying) is much greater for the zigzag design, because there is no off-effort travel from the end of one line to the start of the next. Consequently, the zigzag design gives higher precision.

If this higher precision came at the expense of bias, we might still prefer the parallel line designs. However, applying Eq. (2.5), we find no evidence of bias for any of the three methods. For the systematic parallel lines design,  $|\hat{N} - N| = |1478 - 1500| = 22$ , while  $1.96\text{se}(\hat{N}) = 1.96 \times 208/\sqrt{100} = 41$ ; for the random parallel lines design,  $|\hat{N} - N| = |1468 - 1500| = 32$ , while  $1.96\text{se}(\hat{N}) = 1.96 \times 219/\sqrt{100} = 43$ ; and for the zigzag design,  $|\hat{N} - N| = |1504 - 1500| = 4$ , while  $1.96\text{se}(\hat{N}) = 1.96 \times 115/\sqrt{100} = 23$ . Thus none of the means differ significantly from the true abundance of  $N = 1500$  at the 5 % level.



**Fig. 2.6** The study area used for the simulation study, together with the density surface from which our simulated population was generated. Red indicates low density, through to white high density. Contours show densities as numbers  $\text{km}^{-2}$



**Fig. 2.7** A single realization of each design: systematic sample of parallel lines (*left*); simple random sample of parallel lines (*middle*); zigzag sample with equally-spaced waypoints (*right*)

**Table 2.1** Summary of simulation results for each design

	Systematic parallel lines	Random parallel lines	Zigzag design
Mean distance travelled on effort (km)	474	473	701
Mean sample size	146	146	222
Mean abundance estimate	1478	1468	1504
Estimated percent bias	−1.4	−2.1	0.2
Mean std error of abundance estimate	210	205	172
Std dev of abundance estimates	208	219	115

All estimates are based on 100 simulations. True abundance was  $N = 1500$ . Total distance travelled (whether on effort or not) was approximately the same for each design

In cases where on-effort surveying is more costly or slower per unit distance than off-effort travel, the gain, if any, of zigzag sampling will be less. Zigzag surveys tend to be especially useful for shipboard surveys conducted in large survey regions, and also for aerial surveys when separation between successive lines in a parallel lines design would be large.

Another point to note from Table 2.1 is that for the zigzag design, the standard deviation of the abundance estimates from the simulations (115) is appreciably smaller than the mean of the standard error estimates (172). If our standard error estimates were unbiased, we would expect these two values to be about the same. Thus the method gives greater precision, at least for this example, than the standard errors would suggest.

## Chapter 3

# Designing Distance Sampling Experiments

### 3.1 Introduction

For most of this book, we assume that the main objective is to estimate abundance of the population of interest. Sometimes, we might instead wish to assess the effects (if any) of a management regime or a development on population densities. For example, does spraying a forest block with pesticide to eradicate an insect pest have a detrimental effect on songbird densities? Or does a proposed conservation measure in farmland have the desired effect on species of conservation concern?

In the above examples, it should be possible to establish replicate plots, in which case designed experiments are feasible. Typically in a designed experiment using distance sampling, there is only one treatment, and it is essential to compare that treatment with a control. There might be no time element (we simply compare densities on treatment plots with those on control plots at the time of the surveys), or we may have repeat visits to plots, and are interested in change over time. We cannot assume that repeat counts at the same plot are independent.

Much of the theory for designed experiments was developed with agricultural trials in mind, for example to compare yields for different crop varieties or growth weights of livestock on different diets. Designed experiments in distance sampling tend to differ from designed experiments in agriculture in several key respects. First, in agriculture, the data tend to be yields or weight gains, and are assumed independent. In distance sampling, the basic data are counts by plot, so that methods for the analysis of counts should generally be used. If we assume that detectability is the same everywhere, we can simply analyse the counts, assuming that they are independent. However, if we need to allow for variability in probability of detection (and especially if detectability varies by treatment), we need to correct the counts for detectability. A consequence is that corrected counts are not independent, as it is unlikely that we have sufficient data on a plot to estimate detectability separately for each plot.

Another key difference is that for agricultural trials, the experiment is likely to be conducted on a number of adjacent plots all within the same field, or in a number of adjacent livestock pens. By contrast, for a designed distance sampling experiment, it is likely that potential plots are spread widely through a large study area.

Environmental impact studies can also be conducted using designed experiments, when adequate replication is possible. However, if we wish to assess the impact of say a large offshore windfarm development, we are unable to have replicate treatment plots; there is just one treatment plot, corresponding to the footprint of the windfarm. (If that footprint was divided into plots, we could not assume that those plots are independent; such replication is termed pseudoreplication, a concept popularized by Hurlbert (1984).) We could have replicate control plots, but as we are unable to replicate the treatment plot, this is of limited value. In this case, a different strategy is required, which we outline in Sect. 3.3.

## 3.2 Designed Experiments

### 3.2.1 *Completely Randomized Designs*

Suppose we have  $p$  plots, and wish to test  $t$  treatments,  $T_1, \dots, T_t$ . Typically, we would choose  $p$  to be a multiple of  $t$ , so that we can replicate each treatment the same number  $p/t$  of times. However, if one of the treatments is a control, and comparisons of each treatment with the control is of greater interest than comparisons of treatments with each other, we might choose to have greater replication for the control. In a completely randomized design, treatments are randomly allocated to plots.

In the simplest case, there would be a single line or point on each plot, although we might choose to have say a grid of lines or points on each plot, and pool the data over the grid for analysis.

### 3.2.2 *Randomized Block Designs*

Suppose that we again have  $p$  plots and  $t$  treatments, but now suppose that the plots are grouped into  $b$  blocks. For a balanced design with no replication in each cell (where a cell corresponds to one treatment and block combination), we would choose  $p = bt$ , so that each of the  $bt$  combinations is used on exactly one plot.

A special case of the randomized block design is the matched pairs design, in which  $t = 2$  (usually treatment and control), and each pair corresponds to a block, so that  $b = p/2$ .



### 3.2.2.1 The Conservation Buffer Case Study

A matched-pairs design was used to compare bird densities in crop fields with and without a conservation buffer (Evans et al. 2013). Incentives were offered to establish linear patches of native herbaceous vegetation (*i.e.* field buffers) along field margins to provide temporary habitat for upland bird species. For setting up the monitoring scheme, a minimum of 40 landowner contracts were randomly selected from each state participating in the programme. The objective was to estimate the effect (if any) of field buffers on bird densities in row-crop production systems across nine ecological regions (13 US states).

For each treatment plot, a nearby control plot was identified, so that the design was a matched-pairs design, with one point on each plot of each pair. Each control plot was chosen within 1–3 km of the corresponding treatment plot, in a field without a buffer but with a similar cropping system and landscape.

We refer to this study as the conservation buffer study. We will return to it in Sect. 8.5.1, where we use maximum likelihood methods to assess the effect of conservation buffers on densities of indigo buntings (*Passerina cyanea*), and again in Sect. 8.5.2, where a Bayesian analysis of the full likelihood is carried out to quantify the effect of conservation buffers on densities of coveys of the northern bobwhite quail (*Colinus virginianus*).

### 3.2.3 Factorial Designs

In a full factorial design, we have two or more treatment factors, and each combination of levels of the factors occurs once. For example suppose that we have a study of songbirds in managed forest, and we wish to assess whether there is an effect on bird densities of using a pesticide, and of thinning. We might use a  $2 \times 2$  factorial design, for which there are four plots, one with pesticide applied and thinning, one with pesticide applied but no thinning, one with thinning but no pesticide applied, and one with neither thinning nor pesticide applied. As precision would be very poor if we restricted our design to four plots, and because there would be no true replication from which to estimate precision, we would add replicates of the design, both to allow reliable estimation of precision and to increase sample size, and hence power for detecting effects.

### 3.2.4 Repeated Measures Designs

Suppose we wish to compare the effect on animal density of two different management regimes for woodland. Any difference between the two regimes may be slow to emerge, and may change over time. In this circumstance, we would want to survey the same set of woodland plots repeatedly over time, rather than make

single visits to a large number of plots. Variation in the data from one visit to another for a given plot will tend to be less than if different plots are visited, so that treatment effects can be detected with greater power. A variant of a repeated measures design is the cross-over design where if we have for example a treatment and a control, the control plots from the first time point are the treatment plots in the second time point and *vice versa*.

### 3.2.5 BACI Designs

Before/After Control/Impact (BACI) designs are used to assess the effect of an impact of some development or management action on animal density. Plots are randomly assigned to treatment (control or impact), and each plot is visited (possibly multiple times) both before and after impact. For the design to be effective, it must be possible to arrange for the impact to occur on several replicate plots. Thus the method is not suitable for example to assess the impact of constructing a single offshore windfarm on seabird densities, because we do not have replicate windfarms (see Sect. 3.3).

#### 3.2.5.1 BACI Case Study

Our BACI case study was designed to assess whether prescribed burning of ponderosa pine (*Pinus ponderosa*) plots (Fig. 3.1) affects bird densities (Dickson et al. 2009; Russell et al. 2009). The case study is based on Buckland et al. (2009), who analysed the data on two warbler species: Grace’s warbler (*Dendroica graciae*) and yellow-rumped warbler (*Dendroica coronata*).

Four National Forests in the southwestern United States were selected as study sites: the Kaibab, Coconino, and Apache-Sitgreaves National Forests in Arizona,



**Fig. 3.1** The effect of prescribed burning on ponderosa pine woodland. Shown here is a single location pre-burn (*left*) and 1 year post-burn (*right*). Photos: Victoria Saab

and the Gila National Forest in New Mexico. On each site, a prescribed fire treatment plot was paired with one (Coconino and Gila) or two (Apache-Sitgreaves and Kaibab) control plots of similar extent and vegetation composition. In our analyses, for sites with two control plots, we combine the data on the control plots as if they had been a single contiguous plot. The four sites were surveyed for birds during each summer from 2003 to 2005. Prescribed fire treatments were applied in the fall of 2003 (Apache-Sitgreaves and Coconino) and spring of 2004 (Gila and Kaibab), so that surveys were conducted for 1 year before impact, and for 2 years post-impact. For each pair of plots, one was burned and the other was not. We are interested in the treatment-by-year interaction, as under the alternative hypothesis that burning affects bird densities, we expect any change from pre-burning to post-burning to differ between burn and control plots.

Point count sampling was conducted at 25–50 points per plot, where points were located at least 250 m apart and 150 m from the edge of the study plot. Point counts began just after the dawn chorus and were completed within 5 h. Each point was surveyed either three or four times during the breeding season between late May and early July. At each point, observers recorded all birds detected for 5 min and used a laser rangefinder to estimate the distance between the observer and the detected bird as 0–10 m, 10–25 m, 25–50 m, 50–75 m, 75–100 m and >100 m.

Several covariates that might affect detectability were recorded: site, plot, cluster size (number of birds in a detected group), visit number, observer, species and year. Additionally, data were recorded on the vegetation surrounding each point, including large (>23 cm dbh) tree and snag densities, and fire effects data to compute a composite burn severity index.

We present analysis of the resulting data in Sect. 7.4.1.

### 3.3 Impact Assessment Studies

In impact assessment studies, it is not possible to establish replicate plots (needed if we are to use an experimental design approach to assessing impact) unless the impact can be replicated at different locations. It can also be difficult to define appropriate control plots in heterogeneous environments. In the absence of replication and/or control plots, we need another basis for drawing inference. A practical approach is to quantify the spatial distribution of animals in a wider area spanning the impact zone, both before and after impact, and then to test which areas show significant changes in density. An example of this approach is given by Petersen et al. (2011), who assessed the impact of constructing the Nysted offshore windfarm on the distribution of wintering long-tailed ducks (*Clangula hyemalis*). As only one windfarm was constructed in the area of interest, replication of the impact was not possible. (Multiple plots could be established within the footprint of the windfarm, but they could not be assumed independent.)

The Nysted windfarm comprises 72 wind turbines and is located 10 km southwest of Gedser, south-eastern Denmark. Most turbines were erected in summer or

autumn of 2002. Every year from 2000 to 2007, apart from 2006, aerial surveys were conducted of a large area spanning the windfarm footprint and surrounding waters, corresponding to a bio-geographical unit for wintering long-tailed ducks. The survey design was 26 north–south transects in a systematic design with 2 km separation between successive lines.

We summarize the analysis of the resulting data in Sect. [7.4.3](#).

## Chapter 4

# Field Methods and Data Issues

As methods developers and data analysts, we often see data long after they have been collected. In some instances, no amount of clever analysis can retrieve a study that has fallen foul of poor field methods. In this chapter, we define what is meant by ‘poor field methods’, and by contrast, what constitutes ‘good’ field methods; and we describe means by which ‘poor’ field methods can be diagnosed and remedied before data collection is completed.

There are two aspects to fieldwork: preparation before entering the field, and data collection. Allocation of sampling effort that fails to respect the tenets of design-based inference (*i.e.* locations sampled are representative of locations over which inference is to be drawn) will produce flawed results.

Difficulties will also arise if the design assumption that animals are distributed independently of the line or point (see Sect. 1.7) is not met. For these reasons, sampling for example near roads or easily accessible areas is likely to result in biased estimates of abundance. If a suitably randomized scheme is not possible, investigators should assess the potential for bias arising from failure of this assumption prior to commencing data collection.

Assumptions exist to make data analysis easier. The need to meet key assumptions necessitates careful attention during data collection. As we will see in Chap. 11, when assumptions cannot be met because of challenges in data collection, assumptions can be exchanged for the collection of additional data.

### 4.1 Field Methods

Field methods employed during data collection should be developed with the three key model assumptions of Sect. 1.7 in mind: that animals on the line or point should be detected with certainty, that animals should be detected at their initial location

before any movement, and that distances should be measured without error. It is important that observers are trained to understand these assumptions, and how to ensure that they are met.

To ensure that all animals at or very close to the line or point are detected, the observer might move slowly and quietly (line transects), or remain longer at the point (point transects). Another option is to have more than one observer, to increase probability of detection. Technology such as thermal imagers or acoustic sensors can help to ensure that this assumption is met. In marine surveys for example where observers have a tendency to search too far out at the expense of missing animals close to the line, training needs to stress the importance of not missing such animals. The principal aim of the observer is *not* to maximize the number of detections or to compete with other observers, but to ensure that animals at the line or point are not missed.

Animal movement, whether responsive or not, may be problematic. Field methods should seek to allow observers to detect animals before responsive movement occurs, or if the animal is detected because it flushes, then the location recorded should be where it flushed from. Non-responsive movement should be slow on average relative to the speed of the observer. This may conflict with the requirement that probability of detection for animals on the line or point is certain; the need for certain detection may require that observers move slowly, while animal movement may dictate that observers move quickly. If these conflicting needs cannot be reconciled, then more complex methods may be required, for example double-observer methods (Sect. 4.1.2.4) so that we do not have to assume that animals on the line or point are certain to be detected.

Technology has reduced the difficulties in estimating distances accurately. We will return to this issue in the appropriate sections below.

### **4.1.1 Point Transect Sampling**

Errors in estimating distances generate appreciably greater bias for point transect sampling than for line transect sampling (Sect. 11.3). It is important therefore to give observers effective training, and laser rangefinders should always be provided. Even for surveys in which most detections are aural, a laser rangefinder is invaluable in measuring the distance to an estimated location.

Point transect sampling is mostly used for songbird surveys. On arriving at a point, typically observers are trained to allow some time for birds to resume normal behaviour after the initial disturbance, and then to record for a fixed time, remaining at the point. Observer training is straightforward for this approach, but it has several difficulties. First, distance sampling is a snapshot method; conceptually, animals are assumed to be frozen at their initial location while the survey takes place. For line transect sampling, non-responsive movement (*i.e.* movement independent of the observer) generates little bias provided that movement is slow relative to the speed of the observer (Sect. 11.4). However, for point transect sampling, any

movement during a count is problematic. Second, in closed habitats, most detections of songbirds tend to be aural, and for an observer who remains at the point, estimating the distance to such detections is difficult (Alldredge et al. 2008). Third, it may prove difficult to keep track of detected birds for the duration of the count, so that double-counting may occur, generating upward bias in density estimates. Fourth, the length of the count is a compromise in the hope that downward bias arising from failure to detect all birds at the point cancels with the upward bias from bird movement and double-counting. In multi-species surveys, the appropriate length of count will vary by species, and will be unknown.

Buckland (2006) proposed a snapshot field method. In it, the observer arrives at a point, and takes several minutes ahead of a snapshot moment to assess what is on the plot. The snapshot moment might be defined by setting a timer on arrival at the point. The observer then records the positions of birds at the snapshot moment. He or she may take as long as is necessary after the snapshot moment to confirm locations, and if a location cannot be confirmed, the bird is not recorded. After the snapshot moment, the observer may move up to say 15 m from the point, to help confirm locations, and to improve estimates of distances of birds from the point. This enables the observer to use triangulation to identify locations of birds that are only heard, or allows the observer to change his or her viewpoint, which may help in locating detected birds.

This snapshot method requires more training, so that observers fully understand what they need to achieve, and how best to use the greater flexibility. It assumes that a bird at or very near the point at the snapshot moment is certain to be located and recorded. If this is problematic, cue counting (Sect. 9.4) or mark-recapture distance sampling (Sects. 5.4.1 and 6.4.4) should be considered.

## ***4.1.2 Line Transect Sampling***

### **4.1.2.1 Terrestrial Surveys**

The key design assumption of distance sampling is that lines are placed independently of animal locations. It is often difficult to achieve this in terrestrial surveys. A vehicle may be preferred to increase observer speed (thus reducing bias from non-responsive movement of animals) and sample size. However, this may restrict transect lines to roads and tracks (Sect. 11.1). Densities along roads and tracks may be unrepresentative, for example because of greater disturbance, or because they avoid less accessible terrain and habitats, or because the roads and tracks themselves create open space and edge habitat that may attract animals. It may prove necessary to do additional off-track transects by foot, to allow calibration of density estimates from track transects. In this case, overall precision might be greater by conducting the whole survey on foot, at the expense of less coverage and smaller samples, due to the imprecision in estimating the calibration factor for correcting track density estimates.

Access issues will also arise for surveys conducted on foot, either due to topography or because access is restricted. If it is impractical to conduct surveys along random transects, point transect sampling should be considered as an alternative: it is easier to reach a random point by the easiest or most accessible route than to follow the route of a random line, so that the ideal design is less compromised.

When detection distances are very small (*e.g.* 1–2 m as occurs in dung (Sect. 9.3) and plant (Sect. 10.7) surveys), it is important to have a marked line to measure to, as otherwise, there is a tendency to record short distances as zero. If a high proportion of detection distances is recorded as zero, reliable analysis of the data is not possible. A common strategy is for one person to pull a rope, or lay down a biodegradable thread, while another searches for objects and measures accurate distances.

When detection distances are larger, it is less important to have a marked line, but there should still be a means of locating the line reasonably accurately. This might be achieved for example by one fieldworker identifying the bearing, while another walks ahead, searching for animals. This helps to avoid bias for example when an observer deviates from the ideal transect to skirt around less accessible habitat; distances of detected animals should then be from the ideal line, not from the route taken by the observer.

Often with foot surveys, target animals hear the observer approaching, and slip away undetected. In these circumstances, field methods should be developed to avoid this as much as possible. For example in closed habitats, transects might be cut in advance, to allow quieter and faster passage of the observers. In such cases, the cut should be minimal, so as not to create an obvious path and hence greater disturbance. Further, observers cannot then see along the transect to a great distance, which would create greater bias from non-responsive animal movement, as animals crossing the transect well ahead of the observer might be spotted and recorded as on the line.

Laser rangefinders should be considered an essential item of equipment for most terrestrial surveys.

#### 4.1.2.2 Aerial Surveys

Special issues arise with aircraft surveys because of the speed of travel. On the positive side, non-responsive animal movement is slow relative to the speed of the observer, so that movement bias is not usually a problem. However, responsive movement may be a problem, if animals have time to move away from the path of the aircraft before being detected and recorded by observers who may have a restricted sideways-only view.

Because of the speed of the platform, it can be difficult to estimate or measure distances of detected animals from the line, and the task may become impossible in high-density areas. If the terrain is relatively flat, this can be addressed by placing markers on the wing struts, together with markers on the window (see Fig. 7.9 of Buckland *et al.* (2001)). When the observer aligns these markers, they define strips of known width and known distance from the transect. Observers then simply count



the number of animals detected in each strip, and analysis of grouped distance data is carried out (Sect. 5.2.2.2). Optical clinometers can also be used to measure vertical sighting angle. Vertical sighting angle along with aircraft altitude can be trigonometrically converted to perpendicular distance from the transect.

Visibility below the aircraft is usually poor. Some aircraft are designed with bubble and/or belly windows for good downward visibility, as illustrated in Fig. 7.14 of Buckland et al. (2001). Without such aircraft, the transect is usually considered to be offset to a distance where visibility is good, for which it is reasonable to assume that probability of detection is certain. If search is conducted from both sides of the plane, there is thus a gap down the centre of the surveyed strip (Fig. 7.8 of Buckland et al. (2001)). Helicopters offer some advantages for aerial surveys because they give less obstructed views forward and below the aircraft, thereby reducing the need for an offset, as well as reducing the difficulties of responsive movement if the observer is able to detect animals before they respond to the aircraft.

If some animals are missed even if they are on the (possibly offset) transect line, double-observer methods (Sect. 4.1.2.4) may be used, which allows estimation of the detection function without having to assume certain detection at any distance (Sect. 5.4.1). Becker and Quang (2009) take this a step further. They do not offset the line, but model the fall-off in detectability under the aircraft.

Increasingly, visual observers are being replaced by high-resolution imagery (Sect. 10.2.2.2), together with strip transect methods (Sect. 6.2.2.1) for analysis. This allows the aircraft to fly higher, causing less disturbance, and fewer problems in areas with complex topography. Further, the data can be validated in a way that visual counts cannot. Technology is likely to result in substantial changes to how aerial surveys are conducted in the future (Sect. 12.2). It seems likely that many studies will soon be conducted using high-resolution imagery from drones ('unmanned aerial vehicles').

### 4.1.2.3 Shipboard Surveys

Shipboard surveys are typically costly, and so ship time should be used to full effect. This generally means having a number of observers, perhaps split into two teams, so that continuous search effort can be conducted during daylight hours. There should be sufficient observers on duty at any one time to ensure, if possible, that all animals close to the line are detected. If this is not feasible, then double-observer methods should be used (Sect. 4.1.2.4).

Animals often respond to a vessel, sometimes avoiding it and sometimes being attracted to it. Field methods should seek to ensure that such animals are detected before they respond. This might mean carrying out a behavioural study as part of a pilot study, to assess over what distance animals respond. This may reveal that observers should search with hand-held binoculars rather than by naked eye, or even that large tripod-mounted binoculars (see Fig. 7.5 of Buckland et al. (2001)) should be used. Such binoculars are especially effective if animals are more-or-less continuously visible, such as occurs with larger schools of dolphins, as they

allow a wider strip to be searched, thus increasing sample size. For animals only intermittently available for detection, such as whales, porpoise or small groups of dolphin, or diving seabirds, the narrow field of view can result in animals close to the trackline being missed. Further, given the narrow field of view, it may be necessary to have multiple observers. They should be trained to ensure that their combined search pattern is effective, which may require some degree of overlap in search area of the observers.

Sighting conditions (*e.g.* sea state, glare) should be routinely recorded. It may prove necessary to discard data recorded in poor conditions. Porpoise for example can easily pass undetected even when very close, and the encounter rate drops dramatically in high sea states. If double-observer methods are used, such covariates can be included in the models to reduce bias arising from heterogeneity in detection probabilities. Other covariates that might be recorded for possible inclusion in the detection function model include observer identity, animal behaviour, and group size for clustered populations.

Laser rangefinders cannot measure distances on water, unless there is an object to hit, but they are still very useful in marine surveys for carrying out distance estimation experiments, or to give observers feedback on their ability to estimate distances, by measuring distances to floating objects. Reticles in binoculars are often the most effective way of estimating distances to detected animals at sea. These measure the distance down from the horizon, which is converted to distance from the observer (Buckland et al. 2001, pp. 256–258). These observer-to-animal distances  $r$  can be converted to distances  $x$  of detected animals from the transect provided that the sighting angle  $\theta$  is recorded (Fig. 1.1). If angles are estimated by eye, considerable rounding can be expected. Especially problematic is that angles close to zero tend to be rounded to zero, resulting in a perpendicular distance of zero. Such systematic rounding to zero makes it very difficult to identify and fit a good model for the detection function. If high-powered tripod-mounted binoculars are used, angle rings on the tripod are a simple and accurate way to estimate angles — though many observers still have a tendency to round to the nearest  $5^\circ$ , unless trained not to! Otherwise, angle boards (see Fig. 7.10 of Buckland et al. (2001)) may be used: an arrow on the board is pointed at the location of a detected animal, and the corresponding angle read. Observers should be trained to record the angle accurately, not to round; even though there is imprecision in aligning the arrow, discouraging rounding of the resulting angle prevents systematic rounding of small angles to zero.

Surveys conducted from small boats are considerably less costly than shipboard surveys, but also have greater restraints on number of observers, on where they can operate, and on how far from the line observers can search.

#### 4.1.2.4 Double-Observer Surveys

Double-observer methods are useful for when detection of animals on the line is uncertain, as is the case for example with cetaceans. Double-observer data can be

analysed using mark-recapture distance sampling methods (Sects. 5.4.1 and 6.4.4). The two observers (or in some cases, the two teams of observers) correspond to the two sampling occasions of mark-recapture. In two-sample mark-recapture, the sampling occasions are ordered: animals are caught and marked in the first sample, and any recaptures occur in the second sample. By contrast, each observer in double-observer methods can fulfil either role. Whether we wish to exploit this symmetry depends on circumstances.

One strategy is to have observer 1 search an area first, setting up trials, and then to have observer 2 search that area, and to record which animals detected by observer 1 are 'recaptured'. This might be achieved by having observer 1 search with powerful tripod-mounted binoculars from a ship, while observer 2 searches with hand-held binoculars or naked eye, so that observer 1 has finished searching an area before observer 2 searches it. Another strategy is to have two aircraft flying in tandem, with one observer in each. By separating the two areas of search, any dependence across the two observers (caused for example by an animal behaving in a way that makes it more easily detectable to both observers) is weakened or even removed (Sect. 6.4.4.4). Because such heterogeneity in the probability of detection of different animals generates strong bias in mark-recapture estimates if it is not modelled, this is potentially an important advantage.

If we have the two observers searching the same area simultaneously, we generate more data (because either observer can 'recapture' an animal that was 'marked' by the other), and it is easier to identify animals that are detected by both observers than when they search the same area at different times. The cost however is the need to model the heterogeneity in the probabilities of detection, or to adopt more complex analysis methods to reduce bias from such heterogeneity.

Whichever approach is adopted, strategies are needed to allow reliable identification of duplicate detections (those detected by both observers). For shipboard surveys in which observer 1 searches further ahead of the ship than does observer 2, observer 1 stops searching after detecting an animal, and attempts to track it in. Although observer 2 should be unaware of any detections made by observer 1, observer 1 can be informed of detections made by observer 2, to allow a judgement of whether observer 2 detects the tracked animal. This is so-called one-way independence.

In the case of tandem aircraft, it is not possible to track a detection, so a model allowing for uncertainty in identifying duplicate detections can be adopted (Hiby and Lovell 1998).

Secondary information should be recorded on detected animals to aid the identification of duplicate detections. This might include recording exact times of cues and animal behaviour, in addition to the usual distance sampling data: distance from the line and size of group in the case of clustered populations.

For surveys of whales, a survey of cues (whale blows) is sometimes preferred. If some whales are unavailable because they do not surface while in detection range, then whale-based methods are biased unless availability is estimated or modelled, whereas cue-based methods do not assume that all whales are available. Also, for double-observer surveys, exact times of blows can be recorded, making

the task of identifying duplicate cues much easier than the equivalent task of identifying animals detected by both observers in whale-based surveys. However, the disadvantage is that cue rate must be estimated, and this can be problematic (Sect. 9.4).

## 4.2 Data Issues

### 4.2.1 Data Recording

In every survey, there is a compromise to be found between the bare minimum and information overload. The bare minimum comprises a record of effort, together with, for each animal detected, an estimate of its distance from the line or point. If an estimate of abundance in the study area is required, rather than just an estimate of mean density, then the size of the study area is required. This has the virtue of simplicity, but limits analysis options.

Effort is line length multiplied by number of visits to a line for line transect sampling, and number of visits to a point for point transect sampling. If only one side of a line is surveyed, the effort is halved. Similarly, if only a sector of a circle around a point is surveyed, then effort is multiplied by  $\theta/(2\pi)$  where  $\theta$  is the sector angle in radians.

Very often, there is interest in modelling animal density as a function of location, habitat or other variables. When initiating new surveys, the objectives may be more basic than this, but objectives often change over time. Hence it is worth considering whether it is practical to collect data to allow a spatial modelling approach later. Thus plot locations should be recorded (*e.g.* grid reference of a point, or of the two endpoints of a line), together with animal locations. Thus in addition to recording the distance of a detected animal from a line or point, we might record in the case of line transect sampling which side of the line the animal was located, and the animal's position along the line. (Distances along the line can also be useful for exploring density gradients in the study area.) For point transect sampling, we might record the bearing of the animal from the point together with its distance.

Whether for spatial modelling or modelling the detection function, it is worth considering what covariates are worth recording. Covariates may be static (*i.e.* do not vary over time, such as altitude) or dynamic (vary over time, such as sea state). Further, covariates might be associated with individual detections (such as animal behaviour, cluster size, gender, species) or with effort (such as observer, sea state, habitat), or they might be spatial covariates available throughout the study area (such as altitude or sea surface temperature available from satellite data).

Careful consideration should be given to how data are recorded. Paper recording forms might be used, or data might be entered onto a computer or data entry device. If it is important that observers are not distracted from searching, but there is no dedicated data recorder, voice-sensitive recording devices might be used, to allow

observers to record data without breaking off their search. This is especially useful for aerial surveys. If data are stored electronically, frequent back-ups should be made.

However data are recorded, time should be allocated at the end of each day to checking data, and perhaps to entering data onto a computer, if datasheets or voice recorders are used. In this way, if there are missing data, or errors in the data, corrections can be made while memories are fresh. More importantly, the data can be reviewed, to assess whether there are any problems (see Sect. 4.2.2). If there are, changes to field procedures might be implemented to resolve problems early, before too many data are compromised.

## 4.2.2 Data Checking

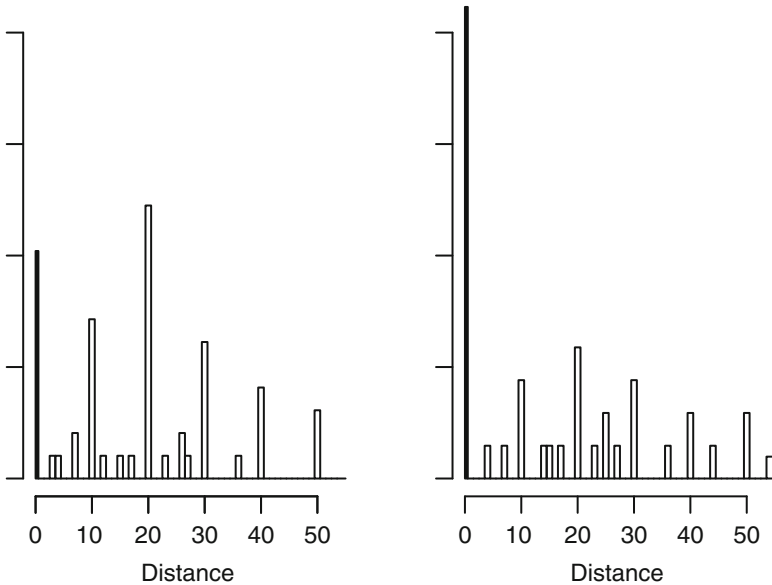
### 4.2.2.1 Heaping

Heaping in the data (rounding of distances, and possibly angles, to favoured values) is generally evident if distances are plotted in a histogram with a large number of intervals. If in line transect sampling, sighting distances and angles are recorded, from which perpendicular distances are calculated (see Fig. 1.1), rounding is more evident if sighting distances and angles are plotted in a histogram, rather than perpendicular distances. A common problem in shipboard surveys is rounding of angles to zero, which results in estimated perpendicular distances of zero.

The effects of heaping can be reduced by grouping distances into intervals, and analysing using the methods of Sect. 5.2.2.2 (line transect sampling) or 5.2.3.2 (point transect sampling). The intervals should be chosen such that as far as possible, all observations end up in the correct distance interval. This is achieved by ensuring that each favoured distance for rounding is roughly in the middle of a distance band, so that distances that belong to a different band are unlikely to be rounded to that distance.

Two examples of heaping in line transect data are shown in Fig. 4.1. In the left-hand plot, although there is clear rounding to the nearest 10 m, this has relatively little impact on estimation. If we analyse the data as if the distances are exact (Sect. 5.2.2.1; for point transect sampling, it would be Sect. 5.2.3.1), the effective strip half-width assuming a half-normal model is estimated to be  $\hat{\mu} = 35.6$  m, while if we group data into intervals of 0–5 m, 5–15 m, 15–25 m, 25–35 m, 35–55 m, we obtain an estimate of  $\hat{\mu} = 34.1$  m. If we instead assume a hazard-rate model, we obtain  $\hat{\mu} = 37.6$  m for exact data and  $\hat{\mu} = 38.3$  m for grouped data.

The right-hand plot of Fig. 4.1 shows what can occur for example in shipboard surveys, where observers either have been poorly trained to record angles or have not been provided with aids to allow reliable estimation of angles (see Sect. 4.1.2.3). Now if we analyse the data as before, analyses under the half-normal model are still similar:  $\hat{\mu} = 36.0$  m for exact data and  $\hat{\mu} = 34.0$  m for grouped data. However, the fit of the model is poor. If we fit the more flexible hazard-rate model, it



**Fig. 4.1** Two datasets showing strong heaping. The *left-hand plot* shows  $n = 49$  distances from the line, and rounding to the nearest 10 m is evident. The *right-hand plot* shows  $n = 34$  detection distances, of which eight were recorded as zero. The data are ‘spiked’, as a consequence of rounding to zero

attempts to fit the spike in the data at zero distance, and as a consequence, we obtain very different estimates from those obtained under the half-normal model, and also very different estimates depending on whether we group the data or not:  $\hat{\mu} = 2.8\text{m}$  from exact data and  $\hat{\mu} = 21.5\text{m}$  from grouped data. If we tried different choices of group interval, we would find sensitivity to this choice. Reliable estimation from these data is not possible unless we somehow obtain information on the rounding process to allow us to correct for it.

#### 4.2.2.2 Responsive Movement

Responsive movement is sometimes detectable from the data. For example if animals close to the line or point move away from the observer before detection, but remain in detection range, then there is a tendency to see excess detections at mid-distances, and too few detections at short distances. A good example of this for point transect sampling is shown in Fig. 10.2. However, if animals that respond to the observer flee beyond detection range, then there is unlikely to be any indication in the data of the problem. The result is an abundance estimate that is biased low.

Some animals are attracted towards the observer, for example dolphins approaching a ship to ride the bow wave, or a songbird investigating an intruder on its territory. An excess of detections at short distances relative to mid-distances may

be indicative of this. If the fitted detection function falls more quickly with distance from the line or point than might be expected, attraction to the observer should be considered as a possible cause. For example, field experience might suggest that almost all animals at 20 m from the line or point should be detected, but the fitted detection function might give an estimated probability of detection of just 0.5 at 20 m.

#### 4.2.2.3 Animals Missed on the Line or Point

In cases such as marine mammals or diving seabirds, or songbirds in dense canopy, there is no guarantee that all animals on the line or point will be detected. Standard distance sampling data give no indication of such a problem, and consequently, standard analysis will give estimates of abundance that are biased low. If it is thought that this occurs, double-observer methods might be adopted. If one observer detects an animal on the line or point that the other does not, this tells us that probability of detection is less than one. Further, we can use mark-recapture distance sampling methods to allow estimation without assuming that animals on the line or point are certain to be detected. Another solution to the problem is to adopt cue-based methods, which assume that the cue (*e.g.* a whale blow or a songburst for a bird), rather than the animal, is certain to be detected if it is very close to the observer. This assumption may also be relaxed by combining cue counting with double-observer methods. Double-observer field methods are discussed in Sect. 4.1.2.4, and analysis methods in Sect. 5.4. Further discussion on missing animals at the line or point appears in Sect. 11.2, and cue-based methods are addressed in Sect. 9.4.

#### 4.2.2.4 Overdispersion

When distance data are plotted by distance interval in a histogram, overdispersion in the plotted frequencies may be apparent. This arises when individual detections are not independent. The effect is seldom strong enough to be detectable. An exception to this is when cue counting is conducted from points (Sect. 9.4.2). In this case, the animal may give several cues from the same location; as each cue is separately recorded, this results in multiple records at exactly the same distance from the point. An example of this is shown in Fig. 9.13. Even when we plot the data with just a few, wide distance intervals (Fig. 9.14), the shape of the histogram is still not as smooth as we would wish.

Another circumstance that gives rise to overdispersion is when animals occur in clusters, but are recorded individually. This may be the preferred option if clusters are spread out so that the location of the cluster centre is difficult to identify, or more distant animals in the cluster may be undetected. See Sect. 10.4.1.

Analysis methods are very robust to failures of the independence assumption. However, when overdispersion is as extreme as in Fig. 9.13, goodness-of-fit tests

cannot be used to assess the adequacy of the model fit, and AIC can be an unreliable guide (Sect. 9.4.2).

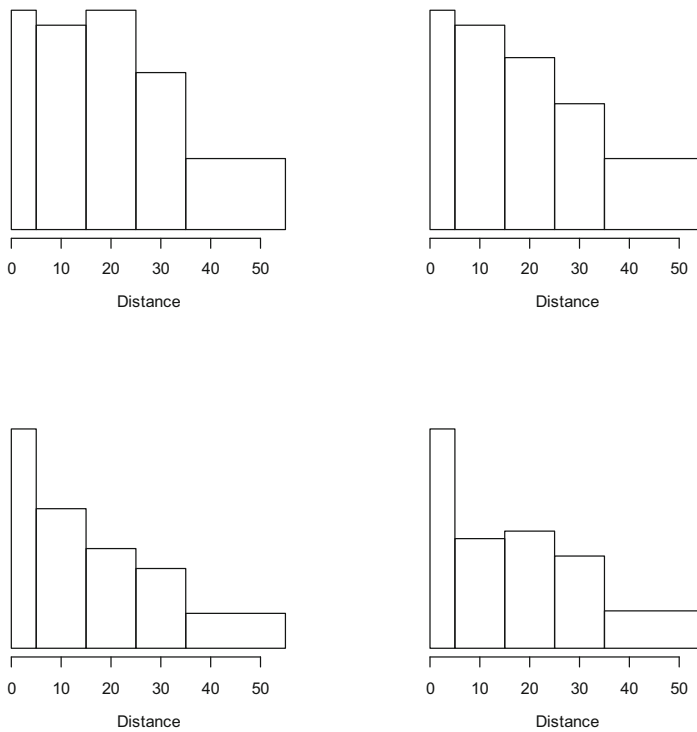
#### 4.2.2.5 Biased Estimation of Cluster Size

For populations that occur in clusters, we assume that the size of detected clusters is recorded without error. We can plot recorded cluster size against distance from the observer to assess whether there is a relationship. However, a relationship may arise because larger clusters are easier to detect (an example of size-biased sampling) and so are over-represented at larger distances (in which case there will be a positive trend in size with distance), or because the size of more distant clusters is underestimated (which would lead to a negative trend in recorded size with distance). If we are lucky, these two sources of bias will approximately cancel. For conventional distance sampling, the default method in software *Distance* is to conduct a regression of log cluster size on estimated probability of detection; we can then estimate mean cluster size when probability of detection is one. This corresponds to clusters on the line or at the point, where we expect no size bias. If we can also assume that cluster size is accurately estimated for such clusters, then this method also corrects for underestimation of the size of more distant clusters. See Sect. 6.3.1.3.

#### 4.2.2.6 Poor Search Pattern

Training of field staff should include how to search for animals to ensure that resulting data are easy to model. We show plots indicating different shapes for the detection function for line transect sampling in Fig. 4.2. We would prefer a histogram of the type shown in the top left plot. The data here suggest that the detection function has a wide shoulder, which means that the probability of detection stays close to one for some distance from the line. We would expect different models that fit the data well to give very similar estimates of animal density in this case. The top right plot by contrast suggests that the detection function has a narrow shoulder, with the detection probability dropping appreciably below one at relatively small distances from the line. We might expect to see more variability in density estimates using different models for the detection function, but estimation should nevertheless be reasonably reliable. The bottom left plot shows ‘spiked’ data, with apparently a very steep fall-off in probability of detection with distance from the line. Different models may yield very different estimates of animal density for these data. If the spike is real, indicating that some animals are unlikely to be detected unless very close to the line, then we need to attempt to fit the spike. This will result in very poor precision. If the spike is an artefact, arising from rounding many distances to zero, or from animals being attracted towards the observer, then a model





**Fig. 4.2** Plots of number of detections by distance interval. In the *top left plot*, the data suggest that the detection function has a wide shoulder. The *top right plot* suggests a narrow shoulder for the detection function. The *bottom left plot* shows ‘spiked’ data, while the *bottom right plot* indicates a good shape except that there are extra detections at very small distances

that cannot fit the spike, such as the half-normal without adjustments, is likely to give lower bias. The bottom right plot is similar, except that there appears to be a wide shoulder, but with extra observations very close to the line. This can occur due to the strategy of ‘guarding the trackline’ adopted in some shipboard surveys. Suppose the main observers search with binoculars, and some animals on or near the line avoid detection. An additional observer might search by naked eye to detect such animals. However, they search a much smaller area than do the observers using binoculars, and when the two sets of data are combined, we can obtain data like those shown in the bottom right plot of Fig. 4.2. It is not possible to model such data reliably, when the effect is this large. A double-observer approach (Sect. 4.1.2.4) might be the better option in this case.

Similar considerations apply to point transect data, except that there are relatively fewer detections close to the point, so that patterns at small distances tend to be less clear. Field methods that produce a wide shoulder for the detection function are more critical for point transect sampling than for line transect sampling, because

the relative lack of data close to the point results in greater variability in density estimates arising from different models for the detection function.

#### 4.2.2.7 Diagnostics of Poor Data

The previous subsections, dealing with challenges to analysis presented by awkward data, suggest a routine of plotting that should be adopted as a matter of course. A set of plots of collected data should be prepared on a daily basis to look for telltale signs of heaping, responsive movement, cluster size bias or problematic search patterns. Plots of data are not guaranteed to detect all of these problems. Nevertheless, just as diagnostic plots are an accepted practice in standard regression analysis, similar rigour in assessing distance sampling data *while data collection is on-going* is key to producing robust data to produce defensible estimates of animal density or abundance.

### 4.2.3 Field Testing

Because there are subtleties associated with field craft in both the sighting of animals and the recording of data, often with the aide of field gear, it is imperative that the entire enterprise be tested prior to data collection in earnest. It is axiomatic that if a field study lacks a pilot study, then the first field season will equate to a pilot study.

Field personnel should both be trained to collect and record high quality data, and assessed to determine if their training has been effective. Testing should assess the ability of field crews to

- properly identify the species of interest,
- accurately measure group size (if animals occur in clusters),
- accurately measure distances of detected animals from the point or line, and
- detect all animals on the transect line.

Trials for each of these skills can be set up under field conditions and crew members can be periodically provided with refresher training. The costs of training as well as periodic reassessment should be incorporated into the project budget.

Not only should there be training and assessment of the data collection process, but also of other aspects of the field enterprise. Checks should be made of the equipment used to make distance measurements, *e.g.* rangefinders, clinometers, altimeters, angle boards, binocular reticles, tape measures. Equipment used to measure effort (maps, stopwatches, GPS units) should be checked for accuracy. Anything used for data entry and recording, *e.g.* voice-activated microphones, handheld computers and paper forms, all need to be tested by the crew members collecting the data.

The complete field protocol, from laying out transects, through collection of data, to daily diagnostics of gathered data, to daily backup of collected data should be

practised by all members of the field crew to build in redundancy in crew expertise. Most importantly, apprising all members of the field crew of the steps from data collection through to data analysis enables the field crews to make enlightened decisions in the field when unforeseen circumstances arise.

**Part II**  
**Design-Based and Model-Based**  
**Methods for Distance Sampling**

## Chapter 5

# Modelling Detection Functions

Distance sampling provides a rigorous framework for estimating detectability, allowing us to correct counts of detected animals in covered areas for those that were missed. The fundamental concept involved in estimating detectability in the distance sampling context is the *detection function*, which represents the probability of detecting an object of interest as a function of its distance from the line or point. Thus a key step in any distance sampling analysis is to choose a plausible and parsimonious model for the detection function.

In this chapter, we introduce tools to help us select a suitable model for the detection function, and to assess how well it fits our data (Sect. 5.1). We then consider conventional distance sampling (Sect. 5.2), in which probability of detection is assumed to be a function of distance from the line or point alone, and for which animals at the line or point are certain to be detected. A set of plausible functions is introduced to model the detection function. Next, we allow the detection function to depend on additional covariates, such as habitat type, observer, animal behaviour or weather conditions, using multiple-covariate distance sampling (MCDS, Sect. 5.3). Finally, we relax the assumption that all animals on the line or point are detected, using mark-recapture distance sampling (MRDS, Sect. 5.4): instead of assuming  $g(0) = 1$ , we use information on number of animals detected by either one or both of two observers to account for failure to detect all animals on the line or at the point.

### 5.1 Model Selection and Goodness-of-Fit

As noted above, a central concept in distance sampling is that of the detection function. This function, usually referred to as  $g(y)$ , where  $y$  represents either the distance  $x$  from a line or the distance  $r$  from a point, describes how probability of

detection falls off as a function of distance from the line or point. A good model for the detection function should have the following properties.

*Shoulder.* The model should possess a shoulder. That is, the probability of detection should remain at or close to one as distance from the line or point increases from zero, before falling away at larger distances (Sect. 5.2.1). This is the so-called *shape criterion* (Buckland et al. 2001, pp. 68–69).

*Non-increasing.* The model should be a non-increasing function of distance from the line or point (Sect. 5.2.1). That is, the probability of detection at a given distance cannot be greater than the probability of detection at any smaller distance.

*Model robust.* As we never know the true shape of the detection function, we need flexible models that can fit a variety of shapes. The purpose of adjustment terms (Sect. 5.2.1) is to provide this flexibility.

*Pooling robust.* Although in conventional distance sampling, we model probability of detection as a function of distance alone, in reality it is a function of many factors. We can attempt to include these factors in the model (Sect. 5.3). Otherwise, we need to rely on the pooling robustness property, which states that inference from our model should be largely unaffected if we fail to include the various factors that influence detectability.

*Estimator efficiency.* Other things being equal, we would prefer a model that gives high precision (*i.e.* small standard errors). However, we should not select a model that gives high precision unless we are confident that it also satisfies the above properties, in which case it should have low bias.

While the first two properties stem from intuitive considerations about the search process and the resulting detectability pattern, the last three relate to the desirable statistical properties of the model for the detection function.

Below, we first consider the issue of pooling robustness in greater detail, and then we briefly describe tools used to select a suitable model.

### 5.1.1 Pooling Robustness

Pooling robustness is a key concept in distance sampling. For those more familiar with estimating abundance using mark-recapture, it seems implausible that estimates from distance sampling should be largely unaffected if there is unmodelled heterogeneity in probability of detection, yet for conventional distance sampling, this is the case. A mathematical proof of the pooling robustness property of distance sampling estimators is given in Burnham et al. (2004, pp. 389–392). Here we focus on when it applies and when it does not, and we illustrate both cases with examples.

When pooling robustness applies, we can model probability of detection as a function of distance from the line or point, while ignoring other factors that may affect detectability. Thus there is no need to record covariates on individuals (such as gender, whether calling, whether in a cluster), environment (such as habitat, thickness of vegetation, visibility) or observer (such as identity, experience, number of observers). This is a significant advantage relative to mark-recapture, for which such heterogeneity is very problematic (Link 2003). Hence it is important to understand when the property applies.

Pooling robustness does not apply for mark-recapture distance sampling, which is unsurprising given that it does not apply to mark-recapture in general. Thus if probability of detection at the line or point is not certain ( $g(0) < 1$ ), and double-observer methods are used to address this, every attempt should be made both to reduce heterogeneity through use of standardized field methods and to model remaining heterogeneity.

Provided we can reasonably assume that  $g(0) = 1$ , and provided heterogeneity is not extreme, pooling robustness applies to an overall abundance estimate. An example of extreme heterogeneity might be surveys of a songbird in dense forest in the breeding season, for which the male may be detectable at great range due to its song, while the female might be undetectable unless very close. In such an extreme case, it may be better to estimate the more detectable component of the population, separately from the less detectable component. For the songbird example, precision on the female component may be very poor; a better option might be to estimate the number of singing males, allowing estimation of the density of territories. Whether it is then reasonable to assume that, on average, there is one female for every male depends on the species.

If analyses are stratified, but a common detection function is assumed for the separate strata, then stratum-specific abundance estimates are not pooling robust, although an estimate of overall abundance obtained by summing the stratum-specific estimates is, provided the proportion of the stratum sampled is equal across strata. Again using the songbird example, if females were sufficiently detectable to allow their inclusion in the analysis, but still less detectable than males, then if we assume a common detection function, we will overestimate number of males in the population and underestimate number of females, although total population size will be estimated with little bias. The stratum-specific bias can be avoided by fitting separate detection functions for males and females, or by including gender as a factor in multiple-covariate distance sampling.

If a survey region is stratified into two habitats, and detectability is lower in one habitat than the other, the stratum-specific abundance estimates will again be biased, if we assume that the same detection function applies to both habitats. Total abundance across habitats will only have the pooling robustness property if effort is in proportion to stratum area. For example if one stratum is twice the size of the other, it should have twice the survey effort.

We illustrate pooling robustness using a dataset on trees in Sect. 5.2.2.4.

### 5.1.2 Information Criteria

Information criteria can be used for selecting between competing models, or quantifying the degree of support for each model. In conventional distance sampling, we can use information criteria to select between say a half-normal and a hazard-rate detection function. The most widely used information criterion is Akaike's information criterion (AIC), and this is the one that we will use throughout this book. Conceptually, most applied statisticians prefer to use information criteria rather than hypothesis testing, but mathematically, it is easy to show that using AIC to compare two nested models, for which one model has a single additional term, is equivalent to using a likelihood ratio test (Sect. 5.1.3) with size 0.157 (15.7 %).

For a given model with maximized likelihood  $\hat{\mathcal{L}}$ , information criteria are of the form  $-2 \log_e \hat{\mathcal{L}}$  with a penalty added, where the penalty is a function of the number of parameters in the model. For AIC, we have

$$\text{AIC} = -2 \log_e \hat{\mathcal{L}} + 2q \quad (5.1)$$

where  $q$  is the number of parameters in the model. (Thus  $q$  is the number of parameters in the key function plus the number of adjustment terms.) AIC is evaluated for each model, and the model with the smallest AIC value is deemed best.

For ease of comparison,  $\Delta\text{AIC}$  values are often used. These are formed by subtracting the AIC value corresponding to the best model (*i.e.* the one with the smallest AIC) from the AIC value of each model in turn. Thus the best model has  $\Delta\text{AIC} = 0$ . Burnham and Anderson (2002) suggested that models with a  $\Delta\text{AIC}$  value of around two or less should be deemed to be well supported by the data, with decreasing support the larger  $\Delta\text{AIC}$  becomes; models with values of  $\Delta\text{AIC}$  above ten should be considered to be very implausible.

AIC is the default model selection method in software *Distance*. Two other methods are provided. The first of these, AICc, includes a small-sample correction for when the data are normally distributed. As the method is not necessarily better for non-normally distributed observations, it is not the default method in *Distance*. It is defined as

$$\text{AICc} = -2 \log_e \hat{\mathcal{L}} + 2q + \frac{2q(q+1)}{n-q-1} . \quad (5.2)$$

When sample size  $n$  is large and  $q$  is small (as for most distance sampling datasets), AICc and AIC are approximately the same.

Also provided in *Distance* is the Bayes information criterion (BIC):

$$\text{BIC} = -2 \log_e \hat{\mathcal{L}} + q \log_e n . \quad (5.3)$$



The penalty  $q \log_e n$  exceeds  $2q$  for  $n \geq 8$ , so that for realistic sample sizes, BIC never selects a larger model than does AIC, and often selects a smaller one. If it is believed that truth is low-dimensional, and that one of the competing models represents truth, then BIC has better properties than AIC. If however, truth is considered to be high-dimensional, and if we are seeking to find a ‘best approximating model’ based on the data, AIC has the better properties. In our examples, we use AIC, which reflects better our philosophy.

A forwards stepping procedure may be adopted to select adjustment terms; this is computationally more efficient than evaluating AIC for every possible combination of adjustment terms, and in most cases, leads to the same choice of model. Thus we start with the key function on its own for model 1, and the key function with just the first adjustment term added as model 2. If model 1 has a smaller value of AIC, then use it. Otherwise, select model 2, then define model 3 to include a second adjustment term. If model 2 has the smaller AIC, we select it; otherwise we compare model 3 with model 4 (which has three adjustment terms). This process continues until the simpler of the two models is not rejected, or until we reach the maximum number of adjustments allowed by the software.

Values of information criteria may be transformed into weights, for model averaging or for quantifying relative support for the models (Buckland et al. 1997; Burnham and Anderson 2002). Model averaging is especially useful for when different models have comparable information criterion values, yet give rather different estimates of animal density; the approach then allows uncertainty over which model is appropriate to be incorporated into estimation. The AIC weight for model  $m$  is defined by

$$w_m = \frac{\exp(-AIC_m/2)}{\sum_m \exp(-AIC_m/2)} \quad (5.4)$$

where  $AIC_m$  is AIC for model  $m$ . (The weights are unaffected if calculated using  $\Delta AIC$  values in place of AIC values.)

In line transect sampling, if two models give quite different estimates, rather than using model averaging, it is usually worth considering what aspect of the data gives rise to the difference. This may allow one model fit to be rejected, for example because an assumption failure has distorted the fit. This can occur for example if small distances from the line tend to be rounded to zero, perhaps causing the hazard-rate model to fit a spurious spike; or animals avoiding the observer may result in the hazard-rate model fitting an implausibly wide shoulder. By contrast, in point transect sampling, there is relatively little information to judge fit at small distances, and this can result in different plausible model fits yielding quite different estimates of abundance. In this circumstance, model averaging may prove useful, allowing uncertainty to be better quantified.

### 5.1.3 Likelihood Ratio Test

The likelihood ratio test is only relevant for testing between nested models, which limits their usefulness relative to information criteria, which can be used to compare non-nested models too. Thus the likelihood ratio test may be used to test for inclusion of adjustment terms, having selected a key function, but it cannot be used to select between different key functions.

Suppose we have model 1, with  $m_1$  adjustment terms, and we denote the maximum value of its likelihood by  $\hat{\mathcal{L}}_1$ , and model 2, with  $m_1 + m_2$  adjustment terms and maximized likelihood  $\hat{\mathcal{L}}_2$ . Then the test statistic is

$$\chi^2 = -2 \log_e \left( \hat{\mathcal{L}}_1 / \hat{\mathcal{L}}_2 \right) = -2 \left[ \log_e \hat{\mathcal{L}}_1 - \log_e \hat{\mathcal{L}}_2 \right] \quad (5.5)$$

If model 1 is the true model, then the test statistic has a  $\chi^2$  distribution with  $m_2$  degrees of freedom.

A forwards stepping procedure is again most convenient for testing for inclusion of adjustment terms. Thus we start with the key function on its own for model 1 ( $m_1 = 0$ ), and the key function with just the first adjustment term added as model 2 ( $m_2 = 1$ ). If model 1 is not rejected (*i.e.* the  $p$ -value exceeds the size of the test, typically taken to be 0.05 or 5%), then use it. Otherwise, test the model with one adjustment against the one with two adjustments, and so on.

### 5.1.4 Goodness-of-Fit Tests

Information criteria do not provide an absolute measure of how well a model fits the data — instead, they allow the models to be ranked according to which fit best, and they quantify the relative support that each of the competing models has, given the data. However, even the best model as judged by AIC might provide a poor fit to the data.

Goodness-of-fit tests allow formal testing of whether a detection function model provides an adequate fit to data. The  $\chi^2$  goodness-of-fit test cannot be used on continuous data, so that it is of limited use for testing MCDS or MRDS models. It is useful for testing models using conventional distance sampling. However, if distances are not grouped, they must first be categorized into groups to allow the test to be conducted. Thus there is a subjective aspect to the test, and different analysts, using different group cutpoints, may reach different conclusions about the adequacy of the model. By contrast, the Kolmogorov–Smirnov and Cramér–von Mises tests can be applied to continuous data. All three tests can be carried out using the *Distance* software.

Goodness-of-fit tests, as their name implies, assess how well the model fits the data. It is important to realise that, in the event of a poor fit, there are two possible explanations. First is that the model is unable to approximate adequately

the true detection function. However, it may be that the model provides an excellent approximation, but that there is a problem with the data. The most common problem is rounding of distances to favoured values, such as multiples of 10 m. Non-independent detections can also generate significant test statistics. An extreme example of this is cue counting from points, where a single bird may give many songbursts from the same location during the recording period. Hence goodness-of-fit tests cannot be used in that context to assess model fit (Sect. 9.4.2).

In surveys with large numbers of detections, goodness-of-fit tests tend to indicate poor fit even when a perfectly good model has been used. This may be simply because with many observations, very small departures from the assumed model can be detected. More usually, it reflects an assumption failure, such as rounding of distances, non-independence of detections, or responsive movement. If the effect is relatively small, then any resulting bias is likely to be of little concern, despite the apparently poor fit. Indications of poor fit when sample size is small are of greater concern.

#### 5.1.4.1 The $\chi^2$ Goodness-of-Fit Test

Suppose we have  $n$  animals detected, with counts of  $n_1, n_2, \dots, n_u$  in  $u$  distance intervals (see Sect. 5.2.2.2), either by recording grouped distances in the field, or by subsequently defining interval cutpoints and counting how many recorded distances fall within each interval. Denote the interval cutpoints by  $c_0, c_1, \dots, c_u$ , where  $c_0 = 0$  unless the data are left-truncated, and  $c_u = w$ . The  $n_j$  ( $j = 1, 2, \dots, u$ ) are the observed counts by interval, while for a given model for the detection function, the corresponding expected counts are given by  $n$  multiplied by the proportion  $\hat{\pi}_j$  of the fitted probability density function,  $\hat{f}(y)$ , that lies within each interval:

$$\hat{\pi}_j = \int_{c_{j-1}}^{c_j} \hat{f}(y) dy \quad (5.6)$$

The  $\chi^2$  statistic is then defined by

$$\chi^2 = \sum_{j=1}^u \frac{(n_j - n\hat{\pi}_j)^2}{n\hat{\pi}_j}. \quad (5.7)$$

This statistic has an approximate  $\chi^2$  distribution with  $u - q - 1$  degrees of freedom under the null hypothesis that our model is the true detection function, where  $q$  is the number of parameters in the model that we have estimated. This approximation can be poor if there are small expected counts; a rough guide is that expected counts of less than 5 should be avoided. For continuous data, cutpoints can be chosen accordingly. If data are collected in bins, and an expected count is small for a given bin, the data from that bin can be pooled with the data from the neighbouring bin, thus reducing the number of bins by one.

The effects of rounding of continuous data on the  $\chi^2$  goodness-of-fit statistic can be ameliorated by suitable choice of cutpoints. For example if many distances are rounded to the nearest 10 m, cutpoints can be chosen at 5 m, 15 m, 25 m, 35 m, . . . . In other words, try to select cutpoints that are mid-way between favoured rounding distances, so that few distances are recorded in the wrong distance interval. This is one circumstance where the subjective choice of intervals in the  $\chi^2$  test can be used to advantage.

### 5.1.4.2 The Kolmogorov–Smirnov Test

The Kolmogorov–Smirnov test can be applied to continuous or discrete data. It is thus preferable to the  $\chi^2$  test for MCDS and MRDS methods. Suppose we have a model  $g(y, \mathbf{z})$  for the detection function, where  $\mathbf{z}$  is a vector of covariates. We give two such models in Sect. 5.3: the half-normal model (Eq. (5.40)) and the hazard-rate model (Eq. (5.41)), where the dependence on  $\mathbf{z}$  is modelled through the scale parameter (Eq. (5.39)).

The probability density function of the distance  $y_i$  of the  $i^{\text{th}}$  detected animal, given covariates  $\mathbf{z}_i$ , is denoted by  $f_{y|\mathbf{z}}(y_i|\mathbf{z}_i)$ . Note that in general, this function is different for each observation  $y_i$  (although if there are no covariates, then it is the same for all observations). This creates a problem, as the Kolmogorov–Smirnov test assumes that all observations have the same distribution. However, the test is carried out on values of the *cumulative distribution function*, defined as  $F_i = F_{y|\mathbf{z}}(y_i|\mathbf{z}_i) = \int_0^{y_i} f_{y|\mathbf{z}}(y|\mathbf{z}_i) dy$ . The values of this function are all independently and identically distributed as  $\text{uniform}(0,1)$ , so we can think of using the cumulative distribution function as a transformation of the observations  $y_i$ , which do not all have the same distribution, to  $F_i$ , that do. We can then conduct the test on the  $F_i$ . The remaining complication is that we must estimate these values.

Having maximized the conditional likelihood of Eq. (5.42), we estimate  $f_{y|\mathbf{z}}(y_i|\mathbf{z}_i)$  by  $\hat{f}_{y|\mathbf{z}}(y_i|\mathbf{z}_i)$  and hence  $F_{y|\mathbf{z}}(y_i|\mathbf{z}_i)$  by  $\hat{F}_{y|\mathbf{z}}(y_i|\mathbf{z}_i)$ , which we denote  $\hat{F}_i$ . Having evaluated  $\hat{F}_i$  for every observation, we rank these values from smallest to largest. The Kolmogorov–Smirnov test statistic is then defined as

$$D_{KS} = \max_i \left\{ \left| \frac{i}{n} - \hat{F}_{(i)} \right|, \left| \frac{i-1}{n} - \hat{F}_{(i)} \right| \right\} \quad (5.8)$$

for  $i = 1, \dots, n$ , where  $\hat{F}_{(i)}$  indicates the ordered values. Thus the test statistic is the biggest difference (ignoring sign) between the estimated cumulative distribution and the *empirical distribution function*, which is defined to be zero for  $F < F_{(1)}$ , one for  $F \geq F_{(n)}$ , and  $i/n$  for  $F_{(i)} \leq F < F_{(i+1)}$ ,  $i = 1, \dots, n-1$ .

### 5.1.4.3 The Cramér–von Mises Test

The Cramér–von Mises test is very similar to the Kolmogorov–Smirnov test, except that the test statistic is a function of the sum of the squared differences between the cumulative distribution function and the empirical distribution function. As a consequence, it uses more information, and should have higher power, although in practice, the power is very similar. As with the Kolmogorov–Smirnov test, for MCDS and MRDS, it must be carried out on the  $\hat{F}_i$  rather than the observations  $y_i$ .

A weighted version of the test is possible. Software `Distance` carries out both the unweighted test and a test in which differences close to zero distance from the line or point are given greater weight than other differences, as it is the fit close to zero distance that is most important for reliable estimation. See Burnham et al. (2004, pp. 388–389) for more details.

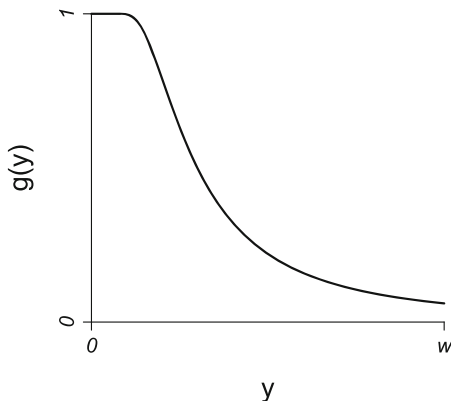
## 5.2 Conventional Distance Sampling

Conventional distance sampling was described in Chap. 1. It refers to the case when all assumptions of Sect. 1.7 hold, and when the detection function is modelled as a function of distance from the line or point only, relying on the pooling robustness property (Sect. 5.1.1) to yield estimates of density with low bias.

### 5.2.1 Models for the Detection Function

Models for the detection function  $g(y)$  (where  $y = x$ , the perpendicular or shortest distance of a detected animal from the line for line transect sampling, and  $y = r$ , the distance of a detected animal from the point for point transect sampling) should have a ‘shoulder’ (Fig. 5.1). In mathematical terms, we say that the slope of the detection function at zero distance, denoted by  $g'(0)$ , is zero. In practical terms, we assume that not only  $g(0) = 1$  (certain detection at the line or point — the first assumption of Sect. 1.7), but also that  $g(y)$  stays at or close to one for animals at small distances from the line or point. Theoretical considerations indicate that a shoulder should normally exist, provided that  $g(0) = 1$  (Buckland et al. 2004, p. 338). More pragmatically, if the true detection function has a wide shoulder, then different models for the function will tend to give similar estimates of density, while if the detection function has no, or only a narrow, shoulder, different models can give rise to very different estimates of density, even if they fit the data equally well. Thus field methods should be adopted that ensure that probability of detection stays close to one for some distance from the line or point, as illustrated in Fig. 5.1.

**Fig. 5.1** A good model for the detection function should have a shoulder, with probability of detection staying at or close to one at small distances from the line or point. At larger distances, it should fall away smoothly. The truncation distance  $w$  corresponds to the half-width of a strip (line transect sampling) or the radius of the circular plot (point transect sampling)



Models for  $g(y)$  should fall away smoothly at middle distances from the line or point. The fall should not be too rapid, and the function should be a non-increasing function of  $y$ . At large distances, the function should level off at or close to zero (Fig. 5.1).

The *Distance* software provides four models for the detection function, but also allows series adjustment terms to be added to the model, for when the simple model (termed a ‘key function’) on its own does not provide an adequate fit. The first key function is the uniform model:

$$g(y) = 1, \quad 0 \leq y \leq w. \quad (5.9)$$

This model should be used to carry out a strip transect analysis, assuming that all animals in the strip of half-width  $w$  are detected. If series adjustment terms are added, then the model is also useful for analysing line or point transect data.

The second key function is the half-normal model:

$$g(y) = \exp \left[ \frac{-y^2}{2\sigma^2} \right], \quad 0 \leq y \leq w. \quad (5.10)$$

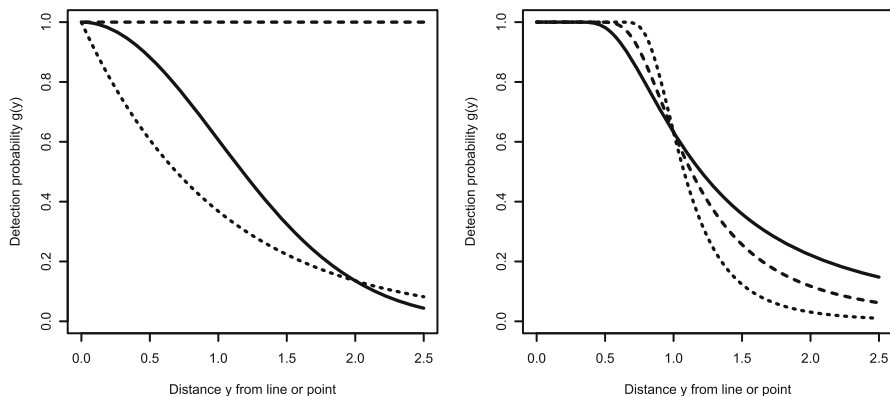
Note that  $g(0) = 1$  as required. The parameter  $\sigma$  is a scale parameter; varying it does not change the shape of the detection function, but does affect how quickly probability of detection falls with distance from the line or point.

The third key function is the hazard-rate model:

$$g(y) = 1 - \exp \left[ (-y/\sigma)^{-b} \right], \quad 0 \leq y \leq w. \quad (5.11)$$

It too has a scale parameter  $\sigma$ , but it also has a shape parameter  $b$ , giving it greater flexibility than the other models.

The fourth key function is the negative exponential model. All four models are shown in Fig. 5.2. The negative exponential model is in *Distance* for historical



**Fig. 5.2** Plots of the key functions available in `Distance`. *Left plot*: half-normal model  $g(y) = \exp[-y^2/(2\sigma^2)]$  with  $\sigma = 1$  (solid line); uniform model  $g(y) = 1/w$  (dashed line); negative exponential model  $g(y) = \exp(-y/\sigma)$  with  $\sigma = 1$  (dotted line). *Right plot*: hazard-rate model  $g(y) = 1 - \exp[(-y/\sigma)^{-b}]$  with  $\sigma = 1$  and  $b = 2$  (solid line),  $b = 3$  (dashed line) and  $b = 5$  (dotted line). In each case, truncation distance  $w = 2.5$

reasons, having been the first rigorously-developed line transect model (Gates et al. 1968), but we do not recommend its use, as it has no ‘shoulder’ (Sect. 5.1). Often, assumption failure leads to spiked data, for which there are many detections close to the line or point, with a sharp fall-off with distance. For example, in line transect sampling with poor estimation of distance, many detected animals may be recorded as on the line (zero distance). In such cases, a model selection tool such as AIC might select the negative exponential model because of its spiked shape. However, we should consider whether the negative exponential is a plausible model *a priori* for a detection function. If all animals on the line are certain to be detected, it is implausible that many animals just off the line will be missed. This is confirmed by hazard-rate modelling of the detection process (Hayes and Buckland 1983); even a sharply-spiked hazard of detection leads to a detection function with a shoulder. Thus there are good reasons not to fit a spiked detection function, even if the distance data appear spiked. Instead, we should look critically at field methods, to understand why the data are spiked, and how we might avoid such data in the future. Apart from measurement error (rounding to zero distance), spiked data can occur due to animal movement (e.g. dolphins approaching a ship to ride the bow wave), or because probability of detection on the line is less than one (in which case probability of detection might fall rapidly with distance).

The software `Distance` allows three types of adjustment to be made to a key function. These adjustments are not needed unless the fit of the key function to the distance data is poor. By default, `Distance` uses Akaike’s Information Criterion and forward stepping to decide whether to add adjustment terms (Sect. 5.1.2). These default options can be changed by the user.

The three types of adjustment are as follows. The first is a cosine series. If used in conjunction with the uniform key function, this gives the Fourier

series model. The second is a Hermite polynomial series. These have orthogonal properties with respect to the half-normal key function, and together, these give the Hermite polynomial model. The third is a simple polynomial series. In practice, any adjustment type can be used with any key function, and *maximum likelihood* methods are used to fit the models. Full details are given by Buckland et al. (2001, pp. 58–68).

### 5.2.2 Line Transect Sampling

A key assumption of conventional line transect sampling is that lines are placed independently of animal locations. This is achieved by placing the lines according to a randomized design — usually, a systematic grid of equally-spaced lines, randomly placed over the survey region. This ensures that animals available for detection are on average uniformly distributed with respect to distance from the line. (By having sufficient lines in the design, any chance non-uniformity averages out across the lines.) Hence any decline in numbers of detections with distance from the line reflects a fall in the probability of detection with distance.

#### 5.2.2.1 Exact Distance Data

In statistical terms, we can model the relative frequencies of observed detection distances by fitting a *probability density function*  $f(x)$  to the distances  $x$ ,  $0 \leq x \leq w$ . Because animals on the surveyed strip and so available for detection are distributed uniformly with respect to distance from the line, we can assume that  $f(x)$  has exactly the same shape as  $g(x)$ . More formally,  $f(x)$  is proportional to  $\pi(x)g(x)$  where  $\pi(x)$  represents the distribution of animals (whether detected or not) with distance from the line, and is uniform here, so that  $\pi(x) = 1/w$ ,  $0 \leq x \leq w$ . Thus  $\pi(x)$  is in fact independent of  $x$ . A valid probability density function must integrate to one, and so we have

$$f(x) = \frac{\pi(x)g(x)}{\int_0^w \pi(x)g(x) dx} \quad (5.12)$$

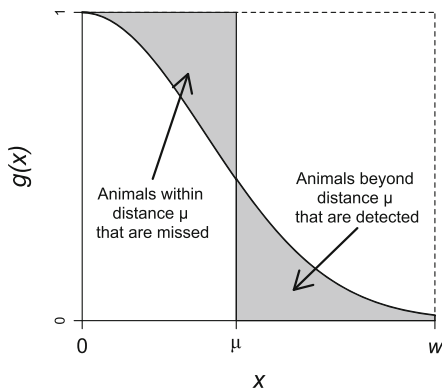
which due to the uniform distribution simplifies to

$$f(x) = \frac{g(x)}{\mu} \quad (5.13)$$

where

$$\mu = \int_0^w g(x) dx. \quad (5.14)$$





**Fig. 5.3** The effective strip half-width  $\mu$  is the distance for which as many animals are detected at distance greater than  $\mu$  (but less than  $w$ ) as are missed closer to the line than  $\mu$ . Thus the two shaded areas have the same size, and  $\mu$  is the distance for which, if you were able to do a complete count of the strip extending a distance  $\mu$  either side of the line, you would expect to detect the same number of animals as were detected within a distance  $w$  of the line

The quantity  $\mu$  is the area under the detection function (Fig. 1.5), and is termed the *effective strip half-width*, because it is the distance from the line at which the expected number of animals detected beyond distance  $\mu$  (but within  $w$ ) equals the expected number of animals missed within a distance  $\mu$  of the line (Fig. 5.3).

Thus for line transect sampling, the uniform model  $g(x) = 1$  for  $0 \leq x \leq w$  (i.e. strip transect sampling, Sect. 6.2.2.1) gives  $\mu = w$  and

$$f(x) = 1/w, \quad 0 \leq x \leq w. \quad (5.15)$$

In general, the integration of Eq. (5.14) does not have a closed analytic form, and is solved numerically. See Buckland et al. (2001, pp. 61–68) for details. Below, to illustrate how to fit a model using maximum likelihood methods, we take a model for which the integral does exist in closed form: the half-normal model without truncation.

If we denote the distances from the line of the  $n$  detected animals by  $x_1, x_2, \dots, x_n$ , then the *likelihood function*, conditional on  $n$ , is given by their joint probability density function. If we assume that the distances  $x_i$  are independent, this likelihood is given by:

$$\mathcal{L}_x = \prod_{i=1}^n f(x_i) = \frac{\prod_{i=1}^n g(x_i)}{\mu^n}. \quad (5.16)$$

We propose a model for  $f(x)$ ; in our example, this is the half-normal model with  $w = \infty$ :

$$g(x) = \exp \left[ \frac{-x^2}{2\sigma^2} \right] \quad (5.17)$$

(note that  $g(0) = 1$ ) so that

$$f(x) = \frac{\exp\left[\frac{-x^2}{2\sigma^2}\right]}{\mu} \quad (5.18)$$

where

$$\mu = \int_0^\infty g(x) dx = \sqrt{\frac{\pi\sigma^2}{2}}. \quad (5.19)$$

Hence

$$\mathcal{L}_x = \mu^{-n} \exp\left[\frac{-\sum_{i=1}^n x_i^2}{2\sigma^2}\right] = \left\{\frac{2}{\pi\sigma^2}\right\}^{n/2} \exp\left[\frac{-\sum_{i=1}^n x_i^2}{2\sigma^2}\right]. \quad (5.20)$$

We now find the value of  $\sigma^2$  that maximizes  $\mathcal{L}_x$ . This is most easily done by taking the logarithm of  $\mathcal{L}_x$ , differentiating the resulting function with respect to  $\sigma^2$ , and setting to zero, giving  $\hat{\sigma}^2 = \frac{\sum_{i=1}^n x_i^2}{n}$ . This in turn allows us to estimate  $\mu$ :

$$\hat{\mu} = \sqrt{\frac{\pi \sum_{i=1}^n x_i^2}{2n}}. \quad (5.21)$$

More generally, the proportion  $P_a$  of animals available for detection (*i.e.* within distance  $w$  of a line) that are actually detected may be expressed as the effective area searched divided by the total area of the sample plots:

$$P_a = \frac{2\mu L}{2wL} = \frac{\mu}{w}. \quad (5.22)$$

Further, it follows from Eq. (5.13) that  $\mu = \frac{g(x)}{f(x)}$  for any  $x$  in  $(0, w)$ . In particular, because  $g(0) = 1$  by assumption,  $\mu = 1/f(0)$ .

Thus the three parameters  $P_a$ ,  $\mu$  and  $f(0)$  are related as:

$$P_a = \frac{\mu}{w} = \frac{1}{wf(0)}. \quad (5.23)$$

Given a maximum likelihood estimate of any one of these parameters, the invariance property of maximum likelihood estimators allows us to calculate the maximum likelihood estimate of the other two. For example, given  $\hat{\mu}$ , we have:

$$\hat{P}_a = \frac{\hat{\mu}}{w} \quad (5.24)$$

and

$$\hat{f}(0) = \frac{1}{\hat{\mu}}. \quad (5.25)$$

Because we are using maximum likelihood methods, we can estimate the standard error of  $\hat{f}(0)$  by first estimating the Fisher information matrix (Buckland et al. 2001, pp. 61–68). We can then estimate the standard error of  $\hat{\mu}$  and of  $\hat{P}_a$  using the approximate result  $\text{cv}[\hat{f}(0)] = \text{cv}(\hat{\mu}) = \text{cv}(\hat{P}_a)$ , so that  $\text{se}(\hat{\mu}) = \hat{\mu}\text{se}[\hat{f}(0)]/\hat{f}(0)$  and  $\text{se}(\hat{P}_a) = \hat{P}_a\text{se}[\hat{f}(0)]/\hat{f}(0)$ .

### 5.2.2.2 Grouped Distance Data

We have assumed that exact distances of detected animals from the line are recorded. Often, distances are grouped into bins. For example, in aerial surveys, the speed may be too great for the observer to record accurate distances to each detected animal, especially in areas of high density. Instead, it is easier simply to count the number of animals by distance bin, where cutpoints between bins are delineated for example by aligned markers on wing struts and windows. In this case, the model for  $f(x)$  is fitted to the grouped data, using a multinomial likelihood:

$$\mathcal{L}_m = \frac{n!}{m_1! \dots m_u!} \prod_{j=1}^u f_j^{m_j} \quad (5.26)$$

where  $m_j$  is the number of animals counted in distance interval  $j$ ,  $j = 1, \dots, u$ ,  $n = \sum_{j=1}^u m_j$ , and  $f_j = \int_{c_{j-1}}^{c_j} f(x) dx$ , where  $c_0, c_1, \dots, c_u$  are the cutpoints between bins, with  $c_u = w$ , and  $c_0 = 0$  (unless data are left-truncated; Buckland et al. 2001, pp. 153–154).

Maximum likelihood methods are again used to fit the model, as described by Buckland et al. (2001, pp. 62–64).

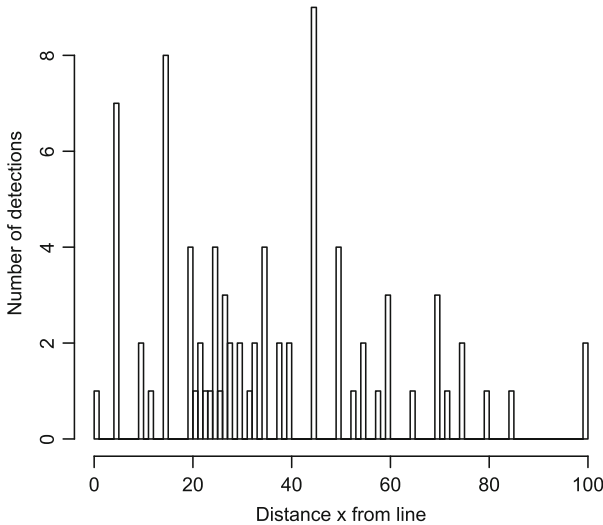
Often, distances are not binned in the field, but substantial rounding error may occur. In this circumstance, especially if there is rounding to zero distances, estimation may be more robust if distances are binned before analysis. In this case, the interval cutpoints that define the bins should be chosen to avoid favoured rounding distances (Buckland et al. 2001, p. 155). For example, if there is rounding to the nearest 10 m, cutpoints might be chosen at 15 m, 25 m, 35 m,  $\dots$ .

Little precision is lost by binning data. However, information is lost for assessing whether assumptions hold, and whether models for the detection function provide an adequate fit. If it is feasible to estimate exact distances, then such data should be recorded using aids (*e.g.* laser rangefinders, gps, clinometers, reticles) to reduce estimation error.

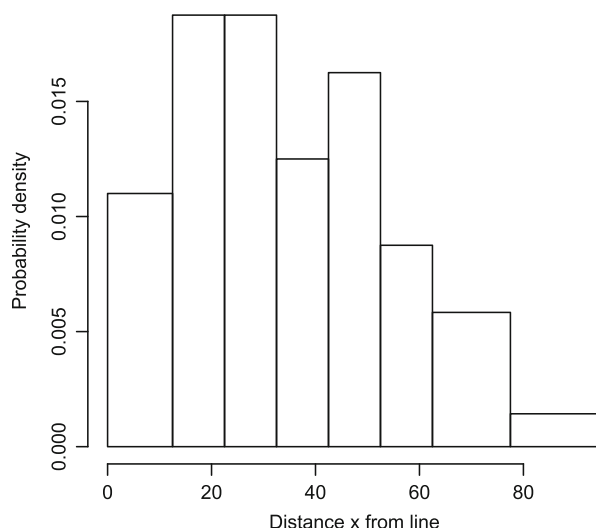
If data are analysed as exact, distance intervals still need to be defined for plotting distance data in histograms, and for conducting  $\chi^2$  goodness-of-fit tests. The Montrave case study on the book website explains how to specify interval cutpoints in *Distance* for analyses of exact distance data and of grouped distance data.

### 5.2.2.3 The Montrave Case Study: Line Transect Sampling

We use the robin data to illustrate line transect sampling. In total, 82 robins were detected, and the corresponding distance estimates ranged from 0 to 100 m. We see from Fig. 5.4 that there is rounding in the data. Seven distances were each recorded four times or more, and all are divisible by 5: 5, 15, 20, 25, 35, 45, 50 m. In principle, distances could be measured to the nearest metre, as a laser rangefinder was used. However, as is typical of songbird surveys in woodland habitat, many more birds were heard than seen. While the laser rangefinder was useful for measuring to visible objects, and aided distance estimation to singing birds that were not visible, nevertheless distance estimation to the nearest metre simply was not possible. To assess model fit, as distinct from revealing rounding in the data, we would like to select cutpoints both for our histogram and for the  $\chi^2$  goodness-of-fit test to avoid these favoured distances. Here, we use the following cutpoints: 0, 12.5, 22.5, 32.5, 42.5, 52.5, 62.5, 77.5, 95 m. Note the use of wider intervals at larger distances, where smaller numbers of birds are detected. For goodness-of-fit testing, a rough guide is that we should avoid expected numbers of observations of less than 5 for any interval (Sect. 5.1.4.1). Note too that we have truncated our distances at  $w = 95$  m. We see from Fig. 5.4 that this truncates two of the 82 detections. Generally, some truncation increases the robustness of the analysis, and Buckland et al. (2001, p. 16) suggest truncation where probability of detection is estimated to be around 0.15. As we will see, this suggests truncating at somewhere close to 80 m. In reality for these data, it makes very little difference. However, for studies where probability of detection shows much more heterogeneity among individuals, resulting in a longer upper ‘tail’



**Fig. 5.4** Estimated distances of the 82 robin detections from the line for line transect sampling, plotted by 1 m interval



**Fig. 5.5** Distribution of robin detections by distance (line transect sampling). Note that probability density is plotted on the y-axis rather than counts, because interval width varies, so that untransformed counts are not a valid guide to the shape of the detection function

to the distribution of distances, more truncation is typically needed, and choice of truncation distance is then important. In such cases, exploratory analyses should use different choices of truncation distance, before selecting a suitable value. Indicators that more truncation may be needed include poor model fit, the need for two or more adjustment terms to improve the fit, or observations at distances much greater than (*e.g.* more than three times) the mean detection distance.

We show a histogram of the robin distance data with the above grouping in Fig. 5.5. It is now much easier to assess how detectability changes with distance. There is also a suggestion that some birds on or near the line either are not detected or move further from the line before detection. We will return to this later.

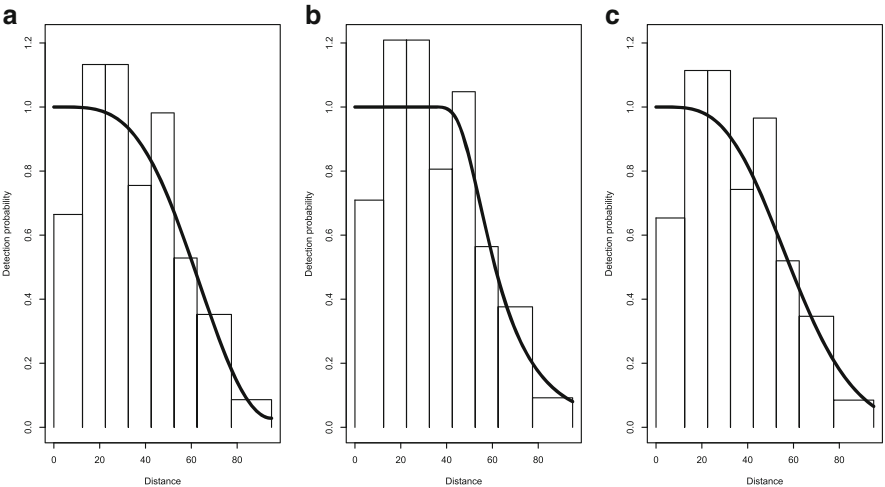
We will now consider fitting models for the detection function to these data. The models we will consider are the uniform key with cosine adjustment terms, the half-normal key with Hermite polynomial adjustment terms, and the hazard-rate key with simple polynomial adjustment terms. We could consider various other combinations of key function and adjustment term, but there is generally little to be gained from this.

The next decision is whether to analyse the distance data as if they are exact, or whether to group them. Of course, if they were collected in the field as grouped data, we have no option. The robin data are not grouped. In that circumstance, it is generally better to analyse the data as exact, unless the amount of rounding is fairly extreme, in which case grouping might be preferred to ameliorate the effect of rounding. For example, if there is a ‘spike’ of detections at zero distance caused by rounding, this can cause bias if we select a model that fits the spike. We might

**Table 5.1** Analysis summary for three detection function models applied to the robin line transect data from the Montrave case study

Key function	Adjustment type	No. adj. terms	$\Delta$ AIC	$\hat{\mu}$	$cv(\hat{\mu})$	95 % CI for $\mu$
Uniform	Cosine	2	0.00	60.4	0.162	(43.9, 83.2)
Half-normal	Hermite polynomial	1	0.64	59.4	0.194	(40.5, 87.1)
Hazard-rate	Simple polynomial	0	0.57	64.5	0.078	(55.2, 75.3)

Units of the effective strip half-width  $\hat{\mu}$  are metres, and density estimates are male birds per hectare



**Fig. 5.6** Estimated detection functions for the line transect data for robins. Fits of (a) uniform key with cosine adjustments, (b) hazard-rate model and (c) half-normal key with Hermite polynomial adjustments are shown

therefore choose to group the data, with the first group width being chosen to be large enough that all distances recorded as zero are likely to have come from the first group, had exact distances been recorded. We do not have this level of rounding, and so we will analyse the data as exact. (The Montrave case study description on the book web site clarifies how to analyse the data as grouped.)

We summarize estimates from each model in Table 5.1. We see that AIC indicates that there is very little to choose between the three models, with a slight preference for the uniform key function with cosine adjustments. For this model and the half-normal key with Hermite polynomial adjustments, estimates are very similar. However, the hazard-rate model gives a slightly higher estimated effective strip half-width, with a much smaller coefficient of variation. We can gain some insight into these differences by examining plots of the estimated detection functions (Fig. 5.6). We see that the first two model fits are very similar, while the hazard-rate model gives a wider, flatter shoulder to the fitted detection function. Generally, the wider and flatter the shoulder, the better the precision. However, if the true detection

**Table 5.2** Observed numbers of robins detected by distance interval (line transect sampling), together with expected numbers under each of the fitted models

	Distance interval (m)							$\chi^2$ (df)
	0–12.5	12.5–22.5	22.5–32.5	32.5–42.5	42.5–52.5	52.5–62.5	62.5–95	
Observed freqs	11	15	15	10	13	7	9	
Exp., uniform-cosine	16.55	13.15	12.74	11.75	10.00	7.59	8.22	3.80 (4)
Exp., hn-Hermite	16.82	13.31	12.72	11.41	9.40	7.01	9.32	4.20 (4)
Exp., hazard-rate	15.51	12.41	12.41	12.37	11.08	7.57	8.66	3.23 (4)

function does not have such a wide shoulder, then the hazard-rate estimate will be biased.

Of the models available in the *Distance* software, only the hazard-rate model is derived from a model of the detection process. However that detection process model does not incorporate any heterogeneity across individuals in probability of detection, except for that due to distance of the animal from the line. If in reality there are other sources of heterogeneity, then this causes the shoulder to become more rounded. This is illustrated in Fig. 11.2 of Buckland et al. (2004).

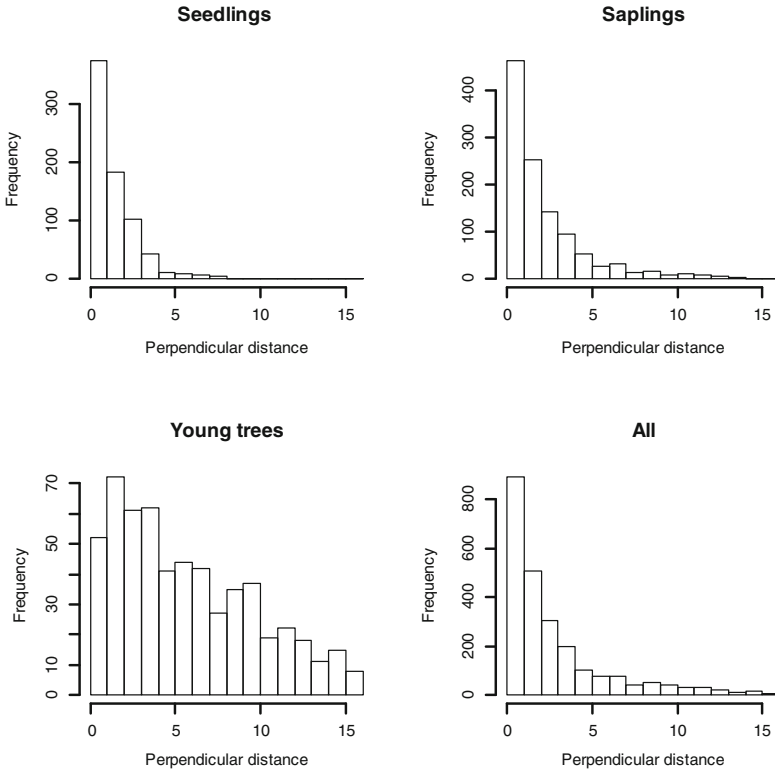
To further assess which model to use, we can examine the results of the goodness-of-fit tests. The output from *Distance* reveals that for the Kolmogorov–Smirnov test and both the weighted and the unweighted versions of the Cramér–von Mises test, and for all three models,  $p$ -values all exceed 0.2, indicating that all models provide good fits to the data. We show the results of the  $\chi^2$  goodness-of-fit tests for each model in Table 5.2. For these tests, we have pooled the last two distance categories (62.5–77.5 m and 77.5–95 m) to avoid expected values less than five (Sect. 5.1.4.1). Again, all three models provide an excellent fit to the data, as judged by this test. (The  $\chi^2$  statistic should be significantly higher than its degrees of freedom to find evidence of poor fit, and for all three models, the statistic is close to the degrees of freedom.)

We see from Table 5.2 that only 11 robins were detected within 12.5 m of the line, whereas under these models, around 16 or 17 were ‘expected’. However, we now see that this discrepancy can easily arise by chance, so that there is insufficient evidence here to suggest that an assumption may have failed (either failure to detect all birds on the line, or movement of birds away from the line prior to detection).

We thus have three excellent model fits, and need to decide which model to use. Generally, we would simply choose a criterion, probably AIC, and go with the best model as judged by that criterion (in this case, the uniform key with cosine adjustments).

#### 5.2.2.4 Cork Oaks: An Illustration of Pooling Robustness

We use data from a study to investigate differences in density of cork oak (*Quercus suber*) seedlings, saplings and young trees between 13 conservation zones and 14 management zones located in eight different estates. In each zone, three to five 50 m



**Fig. 5.7** Perpendicular distances of detected cork oak seedlings, saplings and young trees from the line

lines were surveyed, and distances from the line of detected seedlings, saplings and young trees recorded. Here, we analyse the distance data to illustrate pooling robustness (Sect. 5.1.1).

We see from Fig. 5.7 that, as might be expected, seedlings are less detectable than saplings, which are in turn less detectable than young trees. When data are pooled (bottom right plot of Fig. 5.7), the histogram shows the characteristics of a dataset with substantial heterogeneity in detection probabilities: a spiked histogram, with frequencies dropping rapidly with distance interval, together with a long upper tail, because a few, highly visible individuals are detected at relatively large distances. Indeed, the seedling and sapling histograms also show these characteristics.

Preliminary analyses revealed that, of the models considered, the hazard-rate model provided the best fit to these data, and so we show results for this model only, with the addition of a simple polynomial adjustment if this improved the AIC value.



**Table 5.3** Estimated tree densities  $\hat{D}$  (numbers per hectare) together with coefficients of variation  $cv(\hat{D})$  for the separate categories of seedling, sapling and young tree, and for the three categories combined

	Seedlings	Saplings	Young trees	All
$w$	4	5	15	
$\hat{\mu}$	1.35	1.73	8.78	
$\hat{D}$	602	676	72	1350
$cv(\hat{D})$	0.149	0.127	0.152	0.092

A hazard-rate model was fitted to the data for each tree type. Also shown for each analysis is truncation distance  $w$  (m) and effective strip half-width  $\hat{\mu}$  (m). The density for all categories combined was found by summing the three estimates of density by tree type

In this example, we have large sample sizes, so that we can fit separate detection functions for each category of tree. Resulting density estimates are given in Table 5.3. We will use these to compare estimates obtained below. All models provided satisfactory fits, as judged by  $\chi^2$ , Kolmogorov–Smirnov and Cramér–von Mises goodness-of-fit tests.

Very often, sample sizes are too small to allow independent estimation of the detection function for each stratum. We now consider the implications when data are pooled across strata before fitting the detection function. The distribution for the pooled data has a long tail (Fig. 5.7); for such datasets, substantial truncation gives analyses with greater robustness. Preliminary analyses suggested a truncation distance of  $w = 4$  m was reasonable for these data. To verify that detectability does vary across strata, we will also analyse the data, fitting a separate detection function to each stratum, but this time with truncation at  $w = 4$  m for each stratum, allowing us to use AIC to compare a pooled fit with separate fits by stratum. Results are shown in Table 5.4. (We cannot use AIC to select between the stratified analyses of Table 5.3 and the pooled analyses of Table 5.4 because the truncation distance is not the same in every analysis, hence the need for the stratified analyses reported in Table 5.4.)

AIC shows a clear preference for the stratified analysis, and the stratified results are almost identical to those of Table 5.3. It is perhaps surprising that the estimated density of young trees is unchanged, and the estimated precision very similar, despite having reduced the truncation distance from  $w = 15$  m to  $w = 4$  m for this tree type, which reduced sample size from  $n = 543$  to  $n = 247$ . With a truncation distance of  $w = 4$  m, the model estimates that all young trees in the strip are detected.

We can now see the effect of pooling robustness. If we ignore the large variability in detectability among strata, we estimate total density as 1320 individuals per hectare, a change of a little more than 2 % from the estimate allowing for that variability. However, the stratum-specific estimates do not benefit from pooling robustness, and if we compare density estimates by tree type, we see that seedling density reduces by nearly 20 % and young tree density increases by nearly

**Table 5.4** Estimated tree densities  $\hat{D}$  (numbers per hectare) together with coefficients of variation  $cv(\hat{D})$  for the separate categories of seedling, sapling and young tree, and for the three categories combined

		Seedlings	Saplings	Young trees	All
Pooled	$\hat{\mu}$				1.67
	$\hat{D}$	486	662	172	1320
	$cv(\hat{D})$	0.129	0.116	0.167	0.089
	AIC				4660.3
Stratified	$\hat{\mu}$	1.35	1.63	4.00	
	$\hat{D}$	602	678	72	1352
	$cv(\hat{D})$	0.149	0.124	0.158	0.091
	AIC	1550.8	2304.4	688.8	4544.0

For the pooled analysis, a hazard-rate model was fitted to the data pooled across tree type. For the stratified analysis, a separate hazard-rate model was fitted for each tree type. The truncation distance for all analyses was  $w = 4$  m. Also shown for each analysis is effective strip half-width  $\hat{\mu}$  (m). For the stratified analysis, the density for all categories combined was found by summing the three estimates of density by tree type

140 % when we ignore variability among strata. This is because we overestimate detectability of seedlings, and underestimate detectability of young trees, leading to underestimation and overestimation of density respectively. The message is clear: do not rely on pooling robustness if stratum-specific estimates are important, unless you have evidence that detectability varies little if at all among strata.

We will return to this example in Sect. 5.3.2.3, to show how the heterogeneity in the detection probabilities can be modelled.

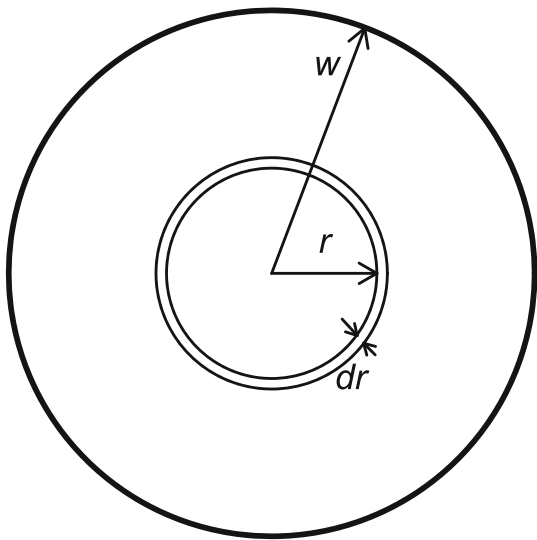
5.2.3 Point Transect Sampling

5.2.3.1 Exact Distance Data

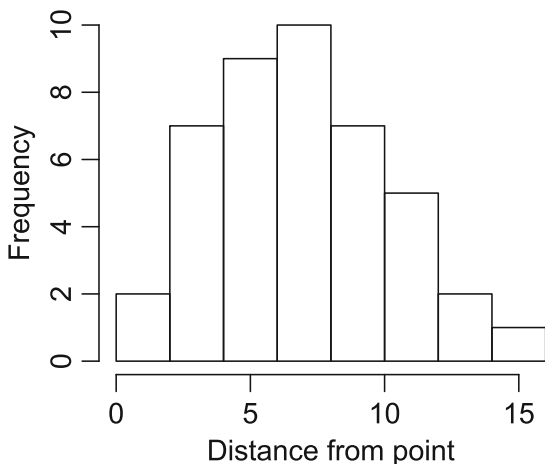
The mathematics is not quite so straightforward for point transect sampling as for line transect sampling. As for line transect sampling, we assume that  $g(0) = 1$ , which in this case equates to assuming that an animal at the point is certain to be detected. We now use the recorded distances  $r$  of detected animals from the point (*i.e.* animal-to-observer distances) to model the detection function  $g(r)$ , which is the probability of detecting an animal that is at distance  $r$  ( $0 \leq r \leq w$ ) from the point.

We assume that points are placed independently of animal locations. Animals available for detection now increase linearly with distance — for example we expect twice as many animals to be present on average at 20 m from a point than at 10 m (Fig. 5.8). This means that animals have a triangular distribution with respect to distance from the point. We therefore expect numbers of detections to increase

**Fig. 5.8** In point transect sampling, the area in an annulus with internal radius  $r$  and external radius  $r + dr$  is approximately  $2\pi r dr$ . Thus if we double the distance from the point, the area in the annulus doubles, and we expect twice as many animals to be present. In mathematical terms, the distribution of animals with respect to distance from the point is triangular:  $\pi(r) = Cr$  where  $C$  is a constant. The honesty condition  $\int_0^w \pi(r) dr = 1$  gives  $C = 2/w^2$



**Fig. 5.9** In point transect sampling, the number of animals available for detection at a given distance increases with distance. Hence histograms of detection distances from the point initially increase roughly linearly. At larger distances, many animals are not detected, so that the initial increase in frequencies slows, then reverses, giving rise to a histogram whose maximum frequency is typically at mid-distance

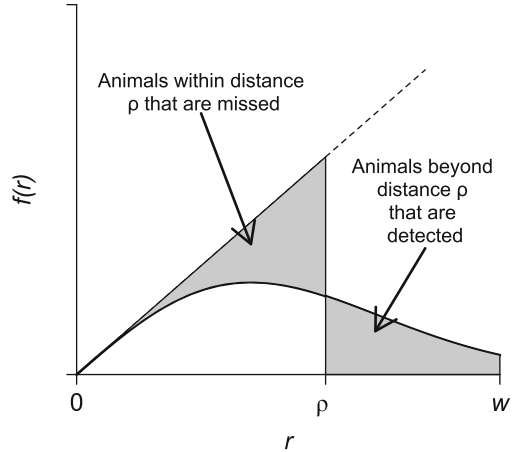


with distance from the point at small distances. At larger distances, the increasing numbers of animals available for detection are offset by the decreasing probability of detection (Fig. 5.9).

As for line transect sampling, we model the relative frequencies of observed detection distances by fitting a probability density function  $f(r)$  to the distances  $r$ ,  $0 \leq r \leq w$ . We still have  $f(r)$  proportional to  $\pi(r)g(r)$ , but now,  $\pi(r)$  is triangular:  $\pi(r) = 2r/w^2$ . Hence  $f(r)$  is proportional to  $r \cdot g(r)$ . Again using the result  $\int_0^w f(r) dr = 1$ , we find:

$$f(r) = \frac{\pi(r)g(r)}{\int_0^w \pi(r)g(r) dr} = \frac{rg(r)}{\int_0^w rg(r) dr} \quad (5.27)$$

**Fig. 5.10** The effective radius  $\rho$  is the distance for which as many animals are detected at distance greater than  $\rho$  (but less than  $w$ ) as are missed closer to the point than  $\rho$ . Thus the two shaded areas have the same size, and  $\rho$  is the distance for which, if you were able to do a complete count of the circular plot of radius  $\rho$ , you would expect to detect the same number of animals as were detected within a distance  $w$  of the point



which can be written as

$$f(r) = \frac{2\pi r g(r)}{\nu} \quad (5.28)$$

where

$$\nu = 2\pi \int_0^w r g(r) dr. \quad (5.29)$$

The rearrangement above is helpful as the quantity  $\nu$  is the effective area surveyed per point, and  $\nu = \pi\rho^2$  where  $\rho$  is the effective radius — the expected number of animals detected beyond distance  $\rho$  (but within  $w$ ) equals the expected number of animals missed within a distance  $\rho$  of the point (Fig. 5.10).

If we denote the distances from the point of the  $n$  detected animals by  $r_1, r_2, \dots, r_n$ , then the likelihood function (conditional on  $n$ ) is:

$$\mathcal{L}_r = \prod_{i=1}^n f(r_i) = \left[ \frac{2\pi}{\nu} \right]^n \prod_{i=1}^n r_i g(r_i). \quad (5.30)$$

We can expect that good models for the detection function  $g(x)$  in line transect sampling will also provide good models for  $g(r)$  in point transect sampling, which allows us to derive corresponding models for  $f(r)$  using Eq. (5.28). As for line transect sampling, we substitute our chosen model into the likelihood, and maximize it, to obtain estimates of the model parameters (Buckland et al. 2001, pp. 61–68).

The proportion  $P_a$  may be expressed as the effective area searched divided by the covered area:

$$P_a = \frac{K\nu}{K\pi w^2} = \frac{\nu}{\pi w^2}. \quad (5.31)$$

We have from Eq. (5.28)

$$f(r)/r = \frac{2\pi g(r)}{\nu} . \quad (5.32)$$

We also know that  $g(0) = 1$ , and so

$$\lim_{r \rightarrow 0} \frac{f(r)}{r} = \frac{2\pi}{\nu} . \quad (5.33)$$

The left-hand side of this equation is the slope of  $f(r)$  at  $r = 0$ ; we denote this by  $h(0)$ . Hence

$$\nu = \frac{2\pi}{h(0)} . \quad (5.34)$$

Thus, having fitted our model for  $f(r)$ , we can estimate  $h(0)$  (by  $\hat{h}(0)$ ), from which

$$\hat{\nu} = \frac{2\pi}{\hat{h}(0)} \quad (5.35)$$

and

$$\hat{P}_a = \frac{2}{w^2 \hat{h}(0)} . \quad (5.36)$$

We use an estimate of the information matrix to estimate the variance of  $\hat{h}(0)$ , and hence of  $\hat{\nu}$  and  $\hat{P}_a$ . See Buckland et al. (2001, pp. 61–68).

### 5.2.3.2 Grouped Distance Data

If distances are grouped, instead of Eq. (5.30), we get a multinomial likelihood (Buckland et al. 2001, p. 63):

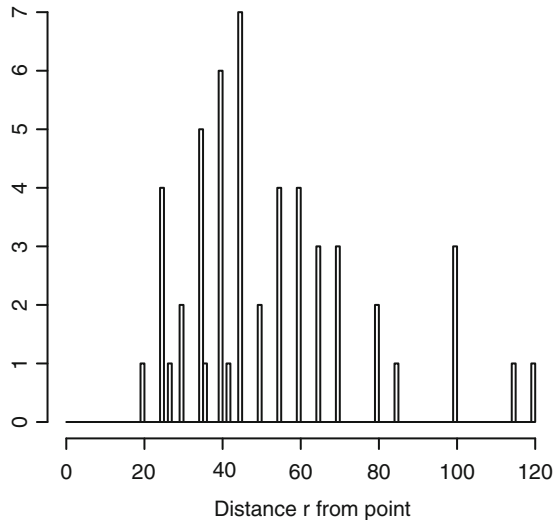
$$\mathcal{L}_m = \frac{n!}{m_1! \dots m_u!} \prod_{j=1}^u \pi_j^{m_j} \quad (5.37)$$

where  $m_j$  is the number of animals counted in distance interval  $j$ ,  $j = 1, \dots, u$ ,  $n = \sum_{j=1}^u m_j$ , and  $\pi_j = \int_{c_{j-1}}^{c_j} f(r) dr$ , where  $c_0 = 0, c_1, \dots, c_u = w$  are the cutpoints between bins.

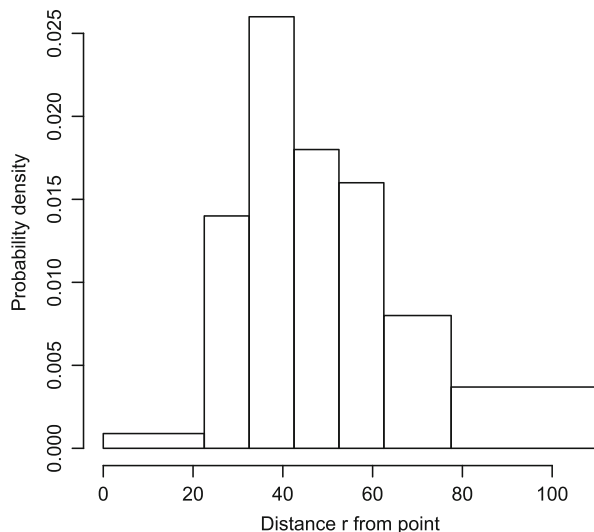
### 5.2.3.3 The Montrave Case Study: Point Transect Sampling

We use the robin data from the snapshot method to illustrate point transect sampling. In total 52 robins were detected, with distance estimates ranging from 20 to 120 m. Note the lack of detections close to the point. This may be partly due to avoidance of the observer, but is primarily because in point transect sampling, the area surveyed and hence the number of birds available to be detected close to the point is low. As for line transect sampling, Fig. 5.11 shows clear evidence of rounding to the nearest 5 m. We should therefore choose cutpoints that avoid these favoured rounding distances to assess model fit. We take the following cutpoints: 0, 22.5, 32.5, 42.5, 52.5, 62.5, 77.5, 110 m. Note the wider intervals at both small and large distances, where smaller numbers of birds are detected, thus ensuring that we avoid small expected values in the  $\chi^2$  goodness-of-fit test (Sect. 5.1.4.1). Our choice of truncation distance,  $w = 110$  m, means that we have truncated just two detections of the 52 (Fig. 5.11). When heterogeneity in probability of detection due to factors other than distance is strong, typically substantially more observations should be truncated for robust estimation from point transect data, perhaps 10% or more (Buckland et al. 2001, p. 151). However, for the robin data, distance was the only factor that significantly affected detectability of a singing male, so that choice here is not influential. If substantial heterogeneity is, or might be, present, exploratory analyses should be carried out with a range of truncation distances. A suitable choice should allow the data to be well-modelled with just one or two parameters, or perhaps three if sample size is large. (The total number of parameters

**Fig. 5.11** Estimated distances of the 52 robin detections from the point for snapshot point transect sampling, plotted by 1 m interval



**Fig. 5.12** Distribution of robin detections by distance (snapshot point transect sampling). Note that probability density is plotted on the y-axis rather than counts, because interval width varies, so that untransformed counts are not a valid guide to the shape of the probability density function



is equal to the number of key parameters — none for the uniform, one for the half-normal and two for the hazard-rate — plus the number of adjustment terms fitted.)

Figure 5.12 shows a histogram of the robin distance data with the above grouping. Unlike for line transect sampling, this histogram is not a guide to the shape of the detection function, because the detection function  $g(r)$  and the probability density function  $f(r)$  do not have the same shape (Sect. 5.2.3.1). However, we note that there are few detections in the first interval (in fact, Fig. 5.11 shows that there is just one observation of less than 22.5 m), and a relatively large number between 32.5 and 42.5 m.

We will fit the same three detection function models as for line transect sampling: the uniform key with cosine adjustments, the half-normal key with Hermite polynomial adjustments, and the hazard-rate key with simple polynomial adjustments. We will also analyse the data as exact; again, the degree of rounding in these data is not sufficient to warrant grouping the data for analysis. (Remember that the groups defined above are purely for presenting the data in the form of a histogram, and for goodness-of-fit testing using  $\chi^2$ ; they are not used for fitting the model, unless we decide that they should be.)

We summarize estimates from each model in Table 5.5. In contrast with the line transect analysis, AIC gives a clear preference for the hazard-rate model. The hazard-rate model gives the highest estimate of effective radius, although the differences are not large. However, one of the models, the half-normal with Hermite polynomial adjustments, gives a much higher coefficient of variation and much wider confidence interval than do the other models. Again, plots of our model fits help to interpret these results. For point transect sampling, we need two sets of plots. In Fig. 5.13, we show the fitted probability density functions and the corresponding detection functions. In the latter plots, the frequency counts have

**Table 5.5** Analysis summary for three detection function models applied to the robin snapshot point transect data from the Montrave case study

Key function	Adjustment type	No. adj. terms	$\Delta\text{AIC}$	$\hat{\rho}$	$\text{cv}(\hat{\rho})$	95 % CI for $\rho$
Uniform	Cosine	1	2.78	61.8	0.032	(58.0, 65.8)
Half-normal	Hermite polynomial	1	2.68	63.1	0.287	(35.9, 111.0)
Hazard-rate	Simple polynomial	0	0.00	64.5	0.085	(54.4, 76.6)

Units of the effective radius  $\hat{\rho}$  are metres

been scaled, making model fit more difficult to assess, so we use the plots of the probability density functions to assess the fits of our models to the data, and the detection function plots to see the shape of the fitted detection functions.

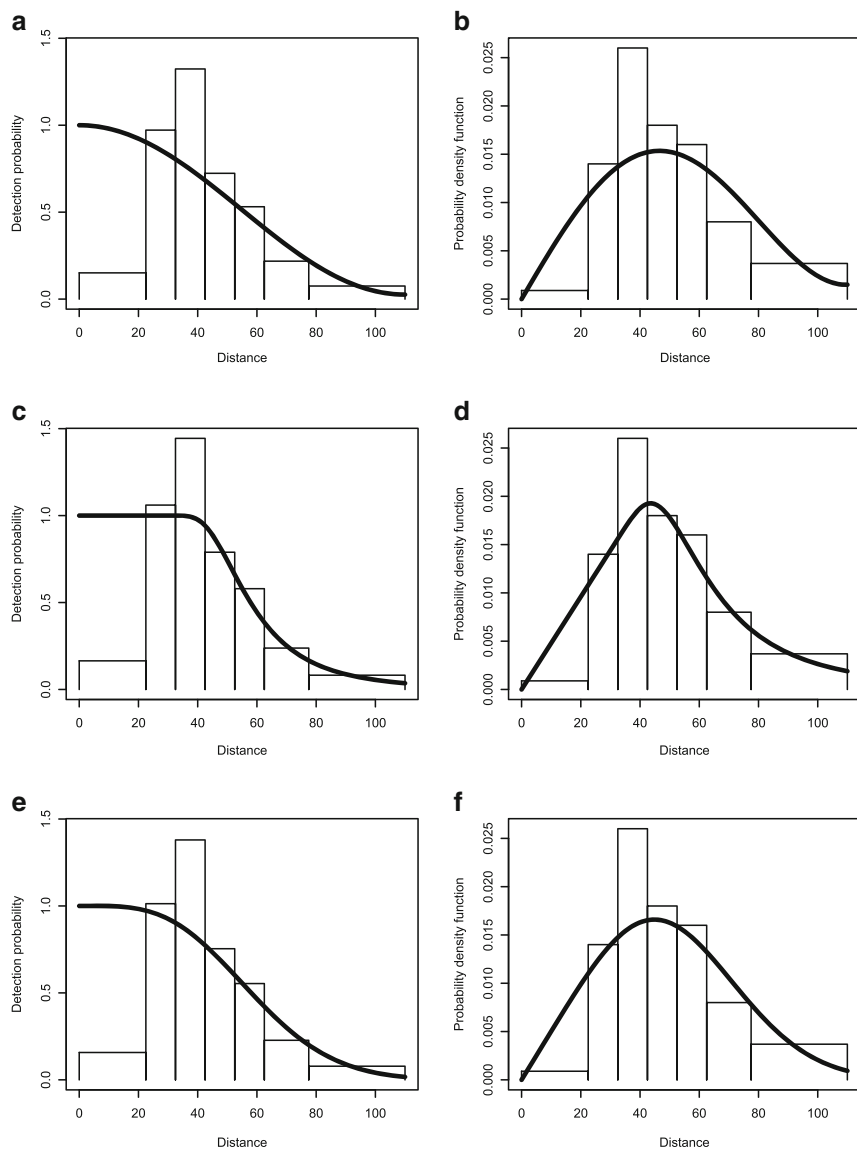
We see from these plots that there appear to be too few detections in the first interval (0–22.5 m), and too many in the third interval (32.5–42.5 m). This suggests that avoidance may have occurred, with birds initially close to the point moving to beyond 32.5 m. The hazard-rate model can fit a very flat, wide shoulder to the detection function, and so is able to fit these data most successfully.

We need to examine the goodness-of-fit tests to draw firmer conclusions. Perhaps surprisingly, Distance output shows that for all three models, the Kolmogorov–Smirnov test and both the weighted and the unweighted versions of the Cramér–von Mises test all yield  $p$ -values that exceed 0.2, indicating that all models provide good fits to the data. The  $\chi^2$  goodness-of-fit tests for each model are summarized in Table 5.6. The  $p$ -values corresponding to the  $\chi^2$  statistics are 0.069 for the uniform-cosine model, 0.079 for the half-normal-Hermite model, and 0.173 for the hazard-rate model. Thus no model is rejected when testing at the 5 % level, although the uniform-cosine and half-normal-Hermite models give significance levels only slightly above 5 %.

Table 5.6 shows that only one robin was detected within 22.5 m of the point, whereas each model suggests that around six should have been detected. Further, 13 were detected between 32.5 and 42.5 m, whereas the models suggest that around 7–9 should have been detected. The  $\chi^2$  goodness-of-fit results therefore give some evidence of avoidance behaviour of robins, although it is not conclusive.

The choice of model is not easy here. AIC favours the hazard-rate model (which has an AIC weight of 0.66, compared with around 0.17 for each of the other two models). The  $\chi^2$  goodness-of-fit statistics also give support, albeit inconclusive, for this model. We therefore choose to base our inference on the results from this model, although we note that, if there was avoidance behaviour, then this may be the reason that the hazard-rate model is favoured, and in practice, this model may underestimate density a little.





**Fig. 5.13** Estimated detection functions (*left*) and probability density functions (*right*) for the snapshot point transect data for robins. Fits of the uniform key with cosine adjustments (**a**, **b**), the hazard-rate model (**c**, **d**), and the half-normal key with Hermite adjustments (**e**, **f**) are shown

**Table 5.6** Observed numbers of robins detected by distance interval (snapshot point transect sampling), together with expected numbers under each of the fitted models

	Distance interval (m)							$\chi^2$ (df)
	0–22.5	22.5–32.5	32.5–42.5	42.5–52.5	52.5–62.5	62.5–77.5	77.5–110	
Observed freqs	1	7	13	9	8	6	6	
Exp., uniform-cosine	6.30	6.15	7.30	7.64	7.17	8.49	6.95	10.24 (5)
Exp., hn-Hermite	6.30	6.50	7.91	8.20	7.35	7.91	5.82	8.37 (4)
Exp., hazard-rate	6.08	6.60	8.87	9.22	7.05	6.45	5.74	6.37 (4)

## 5.2.4 Summary

We use  $y$  to represent either distance  $x$  from the line (line transect sampling) or distance  $r$  from the point (point transect sampling), and we denote the distribution of distances of animals (whether detected or not) from the line or point by  $\pi(y)$ , where for line transect sampling and random line placement,  $\pi(y) = 1/w$ , and for point transect sampling with random point placement,  $\pi(y) = 2r/w^2$ . The detection function  $g(y)$  is the probability that an animal at distance  $y$  from the line or point is detected,  $0 \leq y \leq w$ . Then  $P_a$  is the expected probability of detection for an animal within distance  $w$  of the line or point, where expectation is over distance  $y$ :

$$P_a = \int_0^w g(y)\pi(y) dy. \quad (5.38)$$

Further, the effective area surveyed around line or point  $k$  is  $\nu = a_k P_a$  where  $a_k = 2wl_k$  for line transect sampling, where  $l_k$  is the length of line  $k$ , and  $a_k = \pi w^2$  for point transect sampling. The corresponding effective half-width of the strip centred on a line is  $\mu = wP_a$ , and the effective radius around a point is  $\rho = w\sqrt{P_a}$ .

We also have  $P_a = 1/[wf(0)]$  for line transect sampling, where  $f(0)$  is the probability density function of distances  $y$  evaluated at  $y = 0$ , and  $P_a = 2/[w^2 h(0)]$  for point transect sampling, where  $h(0)$  is the slope of the probability density function of distances  $y$ , evaluated at  $y = 0$ . Statistical tools allow us to fit the probability density function to observed distances  $y_i$ ,  $i = 1, \dots, n$ , and hence we can estimate  $P_a$  and any function of  $P_a$ , such as effective area.

## 5.3 Multiple-Covariate Distance Sampling

### 5.3.1 Adding Covariates to the Detection Function Model

So far, we have assumed that probability of detection for an animal depends only on its distance from the line or point. In many circumstances, this leads to reliable estimates of abundance even when we know that some animals are inherently more

detectable than others (see Sect. 5.1.1). However, it can be useful to model the dependence of detectability on covariates such as habitat, animal behaviour, cluster size, observer, *etc.*

Multiple-covariate distance sampling (MCDS) is covered in detail by Marques and Buckland (2003, 2004). We give a summary of those methods here.

In Sect. 5.2, we defined three key functions for modelling the detection function  $g(y)$ . Two of these, the half-normal detection function (Eq. (5.10)) and the hazard-rate detection function (Eq. (5.11)) each has a scale parameter  $\sigma$ . (The negative exponential model also has a scale parameter, but we will not consider that model here.) A natural extension of the modelling of the previous section is to model the scale parameter as a function of covariates. The scale parameter must always be positive, and so a natural model for the scale parameter, ensuring that this constraint is respected, is as follows.

$$\sigma(\mathbf{z}_i) = \exp \left( \alpha + \sum_{q=1}^Q \beta_q z_{iq} \right) \quad (5.39)$$

where  $\mathbf{z}_i = (z_{i1}, z_{i2}, \dots, z_{iQ})'$  is a vector of covariate values recorded for the  $i^{\text{th}}$  detected animal, and  $\alpha, \beta_1, \dots, \beta_Q$  are coefficients to be estimated.

We can now write the half-normal detection function for detection  $i$  as

$$g(y_i, \mathbf{z}_i) = \exp \left[ \frac{-y_i^2}{2\sigma^2(\mathbf{z}_i)} \right] \quad 0 \leq y_i \leq w \quad (5.40)$$

and the hazard-rate detection function as

$$g(y_i, \mathbf{z}_i) = 1 - \exp \left[ (-y_i/\sigma(\mathbf{z}_i))^{-b} \right] \quad 0 \leq y_i \leq w. \quad (5.41)$$

Given an appropriately randomized design, we know the distribution of  $y$  in the population: it is uniform on  $(0, w)$  for line transect sampling, and triangular for point transect sampling. Typically however, we do not know the distribution of covariate  $z_{iq}$  for  $q = 1, \dots, Q$ . Hence for MCDS, in addition to conditioning on  $n$  when specifying the likelihood, we also condition on the observed  $z_{iq}$ :

$$\mathcal{L}_{y|\mathbf{z}} = \prod_{i=1}^n f_{y|\mathbf{z}}(y_i|\mathbf{z}_i) \quad (5.42)$$

where  $f_{y|\mathbf{z}}(y_i|\mathbf{z}_i)$  is the probability density function of  $y_i$  conditional on the covariates  $\mathbf{z}_i$  (and on  $n$ ). For line transect sampling, distance  $y_i$  becomes  $x_i$ , the shortest distance from the line of the  $i^{\text{th}}$  detected animal, and

$$f_{x|\mathbf{z}}(x_i|\mathbf{z}_i) = \frac{g(x_i, \mathbf{z}_i)}{\mu(\mathbf{z}_i)} \quad (5.43)$$

where

$$\mu(\mathbf{z}_i) = \int_0^w g(x, \mathbf{z}_i) dx . \quad (5.44)$$

As  $g(0, \mathbf{z}_i) = 1$  by assumption, if we set  $x_i = 0$  in Eq. (5.43), we can write

$$\mu(\mathbf{z}_i) = \frac{1}{f_{x|z}(0|\mathbf{z}_i)} . \quad (5.45)$$

For point transect sampling,  $y_i$  becomes  $r_i$ , the distance from the point of the  $i^{\text{th}}$  detected animal, and

$$f_{r|z}(r_i|\mathbf{z}_i) = \frac{2\pi r_i \cdot g(r_i, \mathbf{z}_i)}{\nu(\mathbf{z}_i)} \quad (5.46)$$

where

$$\nu(\mathbf{z}_i) = 2\pi \int_0^w r \cdot g(r|\mathbf{z}_i) dr . \quad (5.47)$$

Again noting that  $g(0, \mathbf{z}_i) = 1$  by assumption, we find from Eq. (5.46) that

$$\nu(\mathbf{z}_i) = \frac{2\pi}{h(0, \mathbf{z}_i)} \quad (5.48)$$

where

$$h(0, \mathbf{z}_i) = \lim_{r \rightarrow 0} \frac{f_{r|z}(r|\mathbf{z}_i)}{r} . \quad (5.49)$$

We can now maximize the conditional likelihood to give estimates of the coefficients  $\alpha, \beta_1, \dots, \beta_Q$ . As for conventional distance sampling, we can add series adjustment terms to the detection function if model fit is poor, although in software *Distance*, there are no monotonicity constraints when using the *mcds* analysis engine, so implausible fits may occur.

### 5.3.2 MCDS Case Studies

#### 5.3.2.1 Hawaii Amakihi

We illustrate fitting MCDS models to point transect data using a bird survey in Hawaii. The survey was implemented to assess a translocation experiment of an endangered Hawaiian honeycreeper on the island of Hawaii (Fancy et al. 1997).

**Fig. 5.14** The Hawaii amakihi, one the most abundant native Hawaiian honeycreepers. Photo: Jack Jeffrey

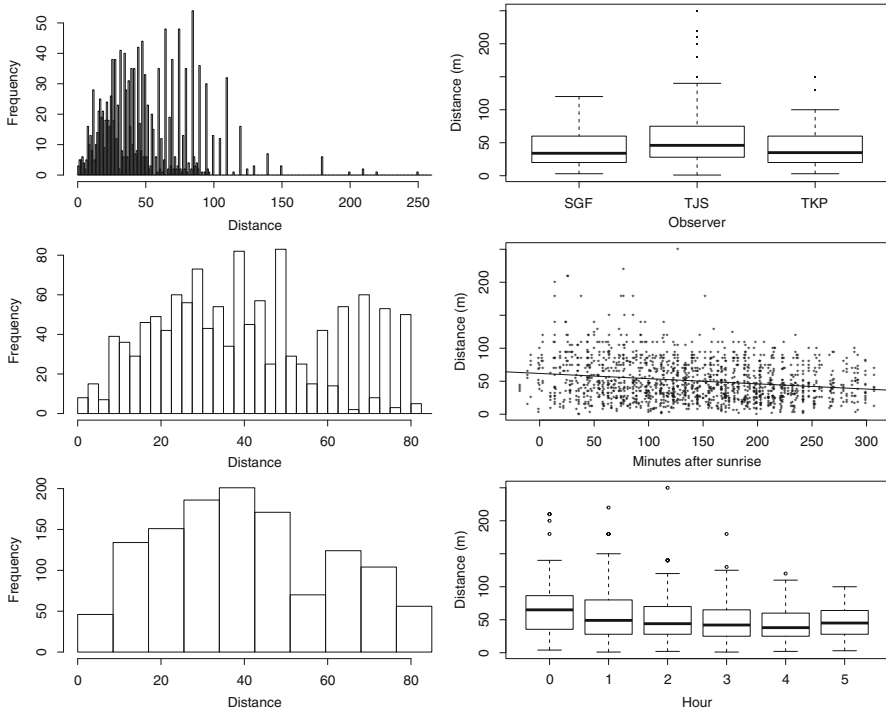


For illustration, we select data on an abundant species, the Hawaii amakihi (*Hemignathus virens*, Fig. 5.14). The same data set, which is available as a sample project in software Distance, was analysed by Marques et al. (2007), a tutorial-like paper for MCDS, where further details about the survey and data analysis are described.

The survey was conducted over seven time periods, which we consider as temporal strata. In each period, counts were made at up to 41 points, leading to a total of 1485 amakihi detections. Here we pool data across time periods and concentrate on modelling the effects of covariates on the detection function. (In Sect. 6.4.3.4, we build on this to estimate density per time period.)

The available covariates include observer (a factor with three levels) and time of day. Time of day was coded both as a factor covariate (hour, six levels) and as the number of minutes after sunrise. *A priori*, these covariates were expected to have an influence on detectability: more experienced observers might be expected to detect more birds, and as birds become less active later in the day, they are also likely to become less detectable.

It is important to gain a good understanding of the data prior to fitting the detection function. Exploratory data analysis can reveal problems in the data and anticipate problems in the analysis, guiding the remainder of the modelling process.



**Fig. 5.15** The detection distances for the Hawaii amakihi survey with different bins and truncation (*left column*), and as a function of the available covariates (*right column*). The *black line* in the middle right plot is a simple regression line

The detection distances are summarized in Fig. 5.15. Also shown is the relationship between the distances and each of the available covariates.

We can now draw some initial conclusions. First, although the largest detection distance is 250 m, there are relatively few beyond 100 m. Unusually large detection distances can be difficult to model satisfactorily, and contribute little to fit of the model at small distances, which is where fit is important. As with CDS models, truncating these distances will facilitate model fitting, for example by reducing or eliminating the need for adjustment terms (Sect. 5.2.1) to achieve a satisfactory fit. Despite reducing sample size, such truncation often results in more precise estimation. See Sects. 5.2.2.3 (line transect sampling) and 5.2.3.3 (point transect sampling) for more discussion of truncation.

The distance data exhibit a mode at mid-distances, as expected for point transects (Sect. 5.2.3.1). If we plot the data using a large number of bins, heaping to the nearest 5 m is evident, especially between 60 and 80 m (Fig. 5.15, top and middle left panels). Although we could proceed with an analysis based on binned data, choosing bins judiciously to solve the problem (Sect. 4.2.2.1), we proceed with the analysis of exact data; modest amounts of heaping have little impact on analyses.

However, we still need to select bin cutpoints for conducting a goodness-of-fit test, and in the presence of heaping, those cutpoints need to be selected with care, to ensure that most distances, after rounding, remain in the correct bin (Sect. 5.1.4.1).

Figure 5.15 suggests that all covariates have an effect on the detection distances, and hence we can anticipate that these will be useful for explaining detectability.

After conducting this exploratory analysis and evaluating the sensitivity of results to several truncation distances, we selected  $w = 82.5$  m. This corresponds to truncating about 16 % of the data. Note that this truncation choice avoids a preferred rounding distance. This is reasonable as we expect that distances of between 77.5 and 82.5 m would tend to be rounded to 80 m (and hence included in the analysis), while distances between 82.5 and 87.5 m would tend to be rounded to 85 m (and hence excluded); one should never choose for truncation a distance to which rounding was evident. Alternatively, one might prefer to consider a more severe truncation, perhaps somewhere between 50 and 60 m, which would correspond to 42.7 % and 35.4 % of the data, respectively. This would avoid the poor fit due to the apparent lack of detections at around 50–60 m relative to greater distances (bottom left panel of Fig. 5.15), but would not affect estimates greatly.

We begin by implementing a conventional distance sampling (CDS) analysis over the pooled data. The candidate detection function models were the uniform key with cosine adjustments, the half-normal key with cosine adjustments, and the hazard-rate key with simple polynomial adjustments. The best model according to AIC is the half-normal with four cosine adjustment terms, followed by the uniform with two cosine terms and the hazard-rate without adjustment terms (Table 5.7).

**Table 5.7** Summary details for models fitted to the amakihi data: key function used (HN is half-normal, HR is hazard-rate, Uni is uniform); number of parameters in the model (Pars);  $\Delta$ AIC values; effective detection radius (EDR); and  $p$ -values for  $\chi^2$ , Cramér–von Mises (CvM) and Kolmogorov–Smirnov (KS) goodness-of-fit tests

Analysis	Covariates	Key function	Pars	$\Delta$ AIC	EDR	$\chi^2$	CvM	KS
MCDS	<i>obs + mas</i>	HR	5	0.00	45.54	0.000	>0.3	0.036
MCDS	<i>obs</i>	HR	4	1.73	45.30	0.000	>0.1	0.003
MCDS	<i>obs + hour</i>	HN	10	3.61	47.97	0.000	>0.5	0.515
MCDS	<i>obs + mas</i>	HN	6	5.53	47.92	0.000	>0.4	0.410
MCDS	<i>obs + hour</i>	HR	9	5.62	46.01	0.000	>0.3	0.020
MCDS	<i>obs</i>	HN	6	7.12	41.97	0.000	>0.5	0.261
CDS	–	HN	5	21.40	43.85	0.000	>0.5	0.320
MCDS	<i>mas</i>	HN	4	24.56	47.97	0.000	>0.5	0.456
MCDS	<i>hour</i>	HN	8	24.57	47.96	0.000	>0.6	0.619
CDS	–	Uni	2	25.11	44.30	0.000	>0.4	0.169
MCDS	<i>mas</i>	HR	3	29.18	47.92	0.000	>0.1	0.015
CDS	–	HR	2	29.90	47.28	0.000	>0.3	0.062
MCDS	<i>hour</i>	HR	7	31.62	47.80	0.000	>0.2	0.016

Covariate *obs* is observer, a factor with 3 levels; *mas* is minutes after sunrise; *hour* is a factor with six levels corresponding to number of hours after sunrise

As we expect detectability to be affected by the covariates, the next step is to include these in the modelling of the detection function. We therefore consider both the hazard-rate and the half-normal model with various combinations of covariates. (We do not consider models that include both the continuous covariate *mas* and the factor *hour* as the two variables are highly correlated.) Results appear in Table 5.7. None of the models with the hazard-rate key require adjustment terms, while the half-normal models always require cosine terms. The covariate representing observer seems to be very important in explaining detectability; all models that include it have a considerably lower AIC than any of the CDS models or any of the other MCDS models without this covariate. Overall, detectability, as reflected by the estimates of the effective detection radius, is relatively constant across models, suggesting that the analysis is robust to model choice. Models that include time recorded as a continuous variable (*mas*) are usually preferred to those that include time as a factor (*hour*). The  $\chi^2$  tests indicate poor fits, while none of the Cramér–von Mises tests do so. This probably indicates heaping in the data (Sect. 4.2.2.1) rather than mis-fit of the model; the  $\chi^2$  test is sensitive to such heaping, unless distance intervals can be identified that ensure most detections are recorded in the correct interval, while the Cramér–von Mises test is not.

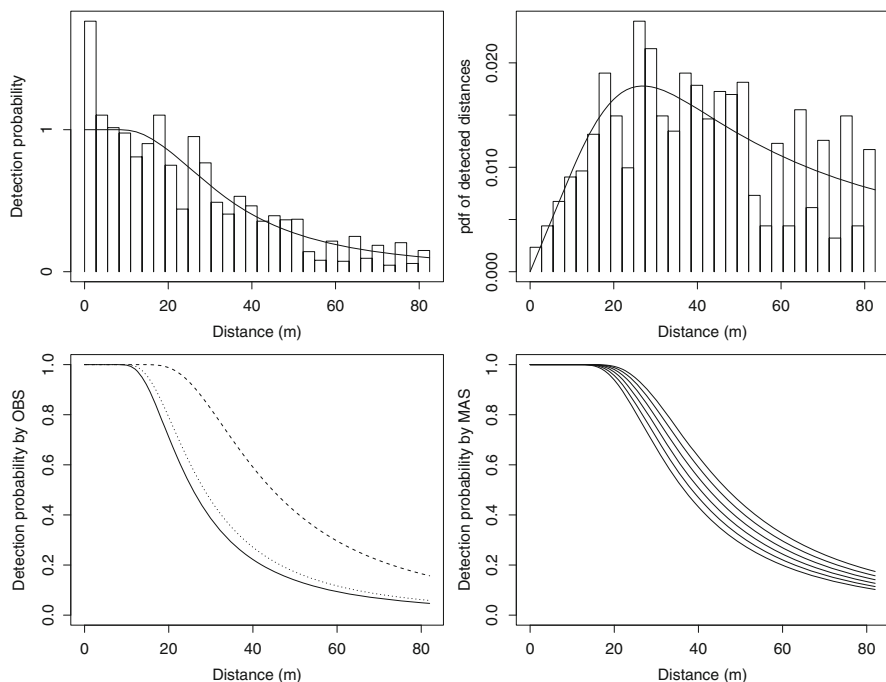
Here, we draw inference based on the MCDS model with the lowest AIC, which includes observer and the continuous time covariate. Output from this model is shown in Fig. 5.16. The fit of the model seems reasonable. The apparent bad fit due to the spike in detections for the first interval in the top left plot is mis-leading, as the histogram bars have been rescaled, exaggerating the mis-fit at small distances relative to that at larger distances; we should judge model fit from the top right plot of the probability density function (Sect. 5.2.3.3), which indicates good fit close to the point.

Figure 5.16 indicates that one of the three observers was able to detect birds at larger distances. We also see that detectability decreases as expected with time since sunrise. Marques et al. (2007) also considered detection functions fitted separately to each of the seven sampling periods, assessing whether a stratified detection function would be better than one including the covariates. AIC showed clearly that this was not the case, suggesting that the available covariates were the predominant sources of heterogeneity in detectability over time. This shows how MCDS can be a parsimonious way of allowing detection probabilities to vary across strata, by fitting a single MCDS detection function across strata. In our case, the strata were the time periods, but they could equally well be spatial strata. Specific software implementation details are provided on the online amakihi case study.

### 5.3.2.2 The Montrave Case Study

In this case study, sample sizes for one species, the great tit, tended to be small. In particular, for the snapshot point transect method, the number of detections within a distance  $w = 110$  m was just  $n = 18$ . Thus we have insufficient data for reliable modelling of the detection function, and we can expect poor precision. A possible





**Fig. 5.16** The model selected by AIC for the Hawaii amakihi survey data was a hazard-rate model with no adjustments, and including observer (*obs*) and minutes since sunrise (*mas*) as covariates. *Top left* shows the estimated detection function, *top right* shows the estimated probability density function, *bottom left* is the detection function as a function of observer (*each line* represents one observer), and *bottom right* is the detection function as a function of minutes since sunrise for the best observer (the *top line* represents zero minutes from sunrise and *subsequent lines* represent hourly increments up to 5 h from sunrise)

solution is to pool data across species, and conduct an MCDS analysis, with species as a factor-type covariate. The Montrave case study on the book website explains how to do this.

In Table 5.8, we show estimates of the effective radius for each species, both for independent analyses by species and for the above MCDS analysis. AIC favoured the hazard-rate model over the half-normal model for the MCDS analysis. For comparability, for the independent analyses, just the half-normal and hazard-rate models were fitted, with no adjustment terms.

We see that point estimates of effective radius are sensitive to choice of model, but not to whether a pooled MCDS analysis is used. The pooled analysis gives a smaller coefficient of variation for all but one of the eight comparisons (four species, two models for each species). Thus the MCDS analysis gives better precision, at the expense of having to assume the same model (hazard-rate or half-normal) for all species. We suggest that data should be pooled across species when multi-species surveys are carried out, but that the pooling should be restricted to sets of species

**Table 5.8** Estimates  $\hat{\rho}$  of effective detection radius (m) for the snapshot point transect method for independent analyses of each species, and for MCDS with species as a factor. (Coefficients of variation in parentheses.)

Species	Independent analyses		MCDS	
	$\hat{\rho}$ , hazard-rate	$\hat{\rho}$ , half-normal	$\hat{\rho}$ , hazard-rate	$\hat{\rho}$ , half-normal
Common chaffinch	86 (3.9 %)	67 (7.8 %)	81 (3.8 %)	67 (4.5 %)
European robin	65 (8.5 %)	56 (8.7 %)	64 (5.8 %)	56 (6.2 %)
Great tit	84 (7.6 %)	65 (17.2 %)	79 (8.7 %)	65 (10.3 %)
Winter wren	75 (4.8 %)	64 (6.5 %)	74 (3.4 %)	64 (3.8 %)

**Table 5.9** Density estimates  $\hat{D}$  for each tree type under hazard-rate models for the detection function

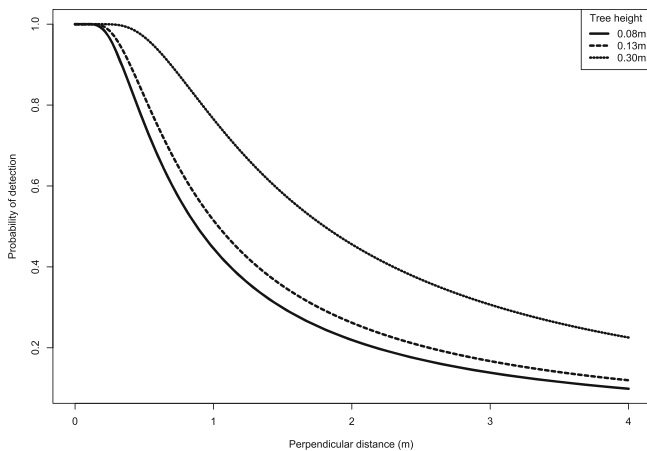
Model	$\Delta AIC$	$\hat{D}$ (seedlings)	$\hat{D}$ (saplings)	$\hat{D}$ (yng trees)	$\hat{D}$ (all)
<i>distance</i>	138.67	493	672	174	1340
<i>distance + height</i>	0.54	623	621	77	1321
<i>distance + treetype</i>	31.28	591	650	72	1313
<i>distance + plottype</i>	138.87	479	659	170	1309
<i>distance + height + treetype</i>	0.69	588	671	78	1337
<i>distance + height + plottype</i>	0.67	618	618	77	1313
<i>distance + treetype + plottype</i>	32.23	586	647	72	1305
<i>distance + height + treetype + plottype</i>	0.00	586	667	78	1331

Also shown are  $\Delta AIC$  values. The possible covariates are *distance* (distance from the line, included in every model), *height* (tree height), *treetype* (seedling, sapling or young tree) and *plottype* (conservation or management)

with similar characteristics in terms of their detectability, so that the detection function can be expected to have a similar shape. Those characteristics might reflect song frequency, behaviour, whether the species typically occupies the canopy or the ground stratum, *etc.*

5.3.2.3 Cork Oaks

In Sect. 5.2.2.4, we saw that for line transect sampling of cork oaks, detectability varied by tree type (seedling, sapling or young tree). Here, we seek to model heterogeneity in detection probabilities using MCDS. Tree type is an obvious covariate to include, and is a factor with three levels, corresponding to the three tree types. Also recorded was tree height, a continuous covariate which may more directly relate to detectability than does tree type. We also consider whether the type of plot (‘conservation’ or ‘management’, and so a factor with two levels) affects detectability. Using a truncation distance of  $w = 4$  m, the hazard-rate model (without adjustment terms) gave substantially better fits than the half-normal model, and we show AIC values for each possible hazard-rate model in Table 5.9.

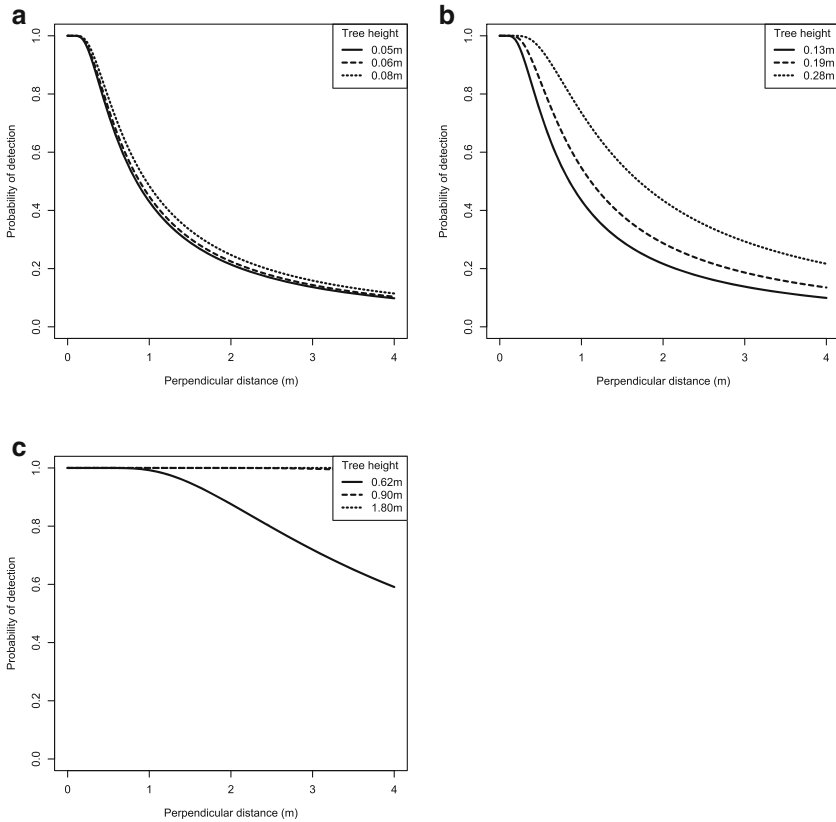


**Fig. 5.17** Fitted detection function corresponding to the median (13 cm) and lower (8 cm) and upper (30 cm) quartile heights of trees pooled across categories (seedlings, saplings and young trees)

We see that all models give good estimates of overall density (a consequence of pooling robustness), and, apart from those with neither height nor tree type, all models give estimates of the density of each tree type that are in reasonable agreement with those of Table 5.3. The models with small  $\Delta\text{AIC}$  values all have height in them. While there is some evidence for including either tree type or plot type, or both, as well as height, the model with just height seems adequate. We show the fitted detection functions at the median height, and also at the lower and upper quartiles, in Fig. 5.17. As there are relatively few young trees (far less than 25 %), these plots do not give an indication of the fitted detection function for young trees. In Fig. 5.18, we show the comparable plots for each tree type, obtained from the model with both height and tree type. We see that the model estimates that almost all young trees of height 90 cm or more and within 4 m of the transect are detected.

From the cork-oak study, we construct a small simulation study to assess the robustness of the estimate of overall abundance if the presence of the three types of trees (seedling, saplings and young trees) is ignored by the analyst. We fitted the hazard-rate detection function to the cork oak data with tree type as a factor using program Distance with truncation distance  $w = 4$  m, giving values of the scale parameter  $\sigma$  of 13.13, 0.90 and 0.65, corresponding to young trees, saplings and seedlings respectively, and a common shape parameter of 1.21.

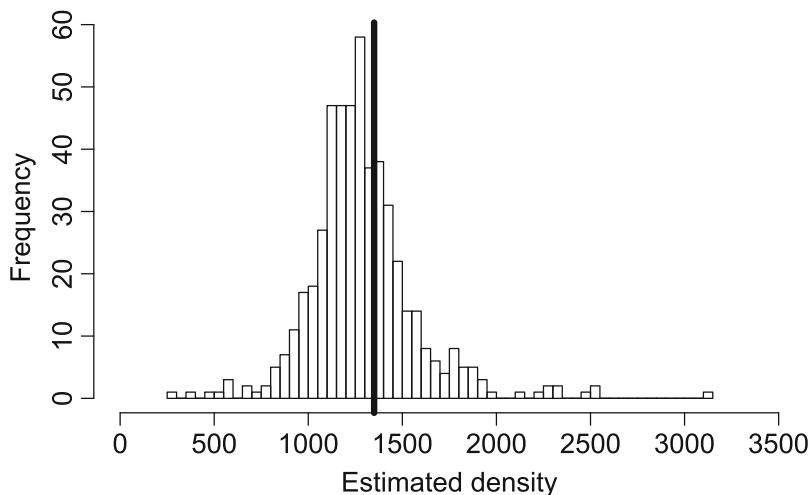
We performed 500 replicates of the simulation with a true density of  $1350 \text{ trees ha}^{-1}$ . Simulated trees were assigned one of the three age-specific values of  $\sigma$  with equal probabilities. Candidate models (ignoring size classes whether as strata or as a factor covariate with three levels) were fitted to each replicate set of detection perpendicular distances. The candidate models were hazard-rate key with Hermite adjustment terms or half-normal key with cosine adjustments. Detection



**Fig. 5.18** Fitted detection function corresponding to the median and lower and upper quartile heights of trees for each tree type. Panel (a) depicts seedlings, panel (b) saplings and panel (c) young trees

distances in the simulated data were truncated at  $w = 4$  m. This resulted in a percent relative bias of  $-4.4\%$ , confidence interval coverage of  $92\%$ , and a distribution of estimates shown in Fig. 5.19.

This simulation was performed using the R package *DSSim* and is included as a case study on the book website. The amount of heterogeneity in detectability in the simulation is evident from comparing the top panel of Fig. 5.18 with the bottom panel: for young trees, probability of detection is still close to one at 4 m, while for seedlings, probability of detection has fallen to around 0.25 at 2 m. As indicated in Sect. 5.1.1, this extreme amount of heterogeneity produces only small bias in the population-level estimation of density. This is in contrast to mark-recapture methods, as described by Link (2003, 2004), where any unmodelled heterogeneity may produce considerable bias in estimates of abundance.



**Fig. 5.19** Distribution of density estimates of cork oaks from 500 replicates when data are simulated from hazard-rate detection functions with differing values of  $\sigma$  for three age classes. True density (shown by vertical line) is  $1350 \text{ trees ha}^{-1}$

## 5.4 Mark-Recapture Distance Sampling

### 5.4.1 Modelling the Detection Function

Mark-recapture distance sampling (MRDS; Laake and Borchers 2004; Borchers et al. 2006) is similar to MCDS, in that probability of detection is modelled as a function of both distance from the line or point and other covariates. However, we now remove the assumption that detection at the line or point is certain ( $g(0) = 1$ ). Burt et al. (2014) provide a non-technical explanation of MRDS methods.

Unfortunately, conventional distance sampling data do not provide information to allow estimation of  $g(0)$ , and so additional data are needed. We consider several ways in which the additional data might be collected below. Each method generates a form of mark-recapture data, which allows probability of detection to be estimated, without having to assume  $g(0) = 1$ . A disadvantage of this approach is that the pooling robustness property (Sect. 5.1.1) does not hold; in common with standard mark-recapture methods, unmodelled heterogeneity in probability of detection generates positive bias in the estimated probability of detection and negative bias in abundance estimates. For this reason, we do not consider MRDS models with no covariates other than distance.

One option for generating the mark-recapture data that we need is to have two observers, searching the same area independently but simultaneously. If one observer detects an animal, then we can conceptualize this as a trial, in which the

second observer either ‘recaptures’ (*i.e.* detects) the animal or not. Thus we need to be able to identify ‘duplicate’ detections: those animals detected by both observers.

In the above arrangement, there is a symmetry: conceptually, each observer sets up trials for the other. For line transect sampling especially, we might instead have one observer setting up the trials (*e.g.* by searching further ahead with optical aids and tracking animals in, Hammond et al. 2002), and the other carrying out standard line transect search. A tracked animal that is detected by the second observer is a recapture.

Mark-recapture trials might be set up by having identifiable animals at known locations (*e.g.* radio-tagged animals), and recording whether a single observer detects these animals. This might be during the course of the main survey, or by setting up trials, so that the observer is directed along a route or to a point such that the known-location animal is within detection range. A model may then be fitted to the trial data, predicting the probability that a given animal is detected as a function of distance and of other relevant covariates, such as animal behaviour, cluster size for clustered populations, weather conditions, observer, *etc.*

In MRDS, it is convenient to scale the function  $g(y, \mathbf{z})$  so that  $g(0, \mathbf{z}) = 1$  by definition. We then denote the detection function by  $p(y, \mathbf{z})$ , which represents the probability that an animal at distance  $y$  from the line or point and with covariates  $\mathbf{z}$  is detected. Thus  $p(0, \mathbf{z}) \leq 1$ .

When there are two observers, we need to distinguish between their respective detection functions. We denote the detection function for observer 1 by  $p_1(y, \mathbf{z})$ , and that for observer 2 by  $p_2(y, \mathbf{z})$ . We will also need further detection functions:  $p_{1|2}(y, \mathbf{z})$  is the probability that observer 1 detects an animal at distance  $y$  and covariates  $\mathbf{z}$  *given that* observer 2 detects it, and the converse:  $p_{2|1}(y, \mathbf{z})$ . Finally, we need to define the probability that at least one observer detects the animal:  $p_{\cdot}(y, \mathbf{z})$ .

These different detection functions are related as follows:

$$\begin{aligned} p_{\cdot}(y, \mathbf{z}) &= p_1(y, \mathbf{z}) + p_2(y, \mathbf{z}) - p_2(y, \mathbf{z})p_{1|2}(y, \mathbf{z}) \\ &= p_1(y, \mathbf{z}) + p_2(y, \mathbf{z}) - p_1(y, \mathbf{z})p_{2|1}(y, \mathbf{z}) . \end{aligned} \quad (5.50)$$

Early models for MRDS assumed *full independence*. Under this assumption, Eq. (5.50) becomes

$$p_{\cdot}(y, \mathbf{z}) = p_1(y, \mathbf{z}) + p_2(y, \mathbf{z}) - p_1(y, \mathbf{z})p_2(y, \mathbf{z}) . \quad (5.51)$$

Unfortunately, this approach is typically biased due to unmodelled heterogeneity in the probabilities of detection. Even though the approach incorporates covariates  $\mathbf{z}$  in an attempt to model heterogeneity, the bias from this source is often greater than the bias that arises by assuming that detection at the line or point is certain. Laake (1999) realised that a weaker assumption is to assume independence at the line or point only, termed *point independence*. At the line or point, probability of detection is typically high, and heterogeneity consequently less problematic. Under point independence, we have:

$$p_{\cdot}(0, \mathbf{z}) = p_1(0, \mathbf{z}) + p_2(0, \mathbf{z}) - p_1(0, \mathbf{z})p_2(0, \mathbf{z}) . \quad (5.52)$$

For  $y > 0$ , we can write

$$\begin{aligned} p_{\cdot}(y, \mathbf{z}) &= p_1(y, \mathbf{z}) + p_2(y, \mathbf{z}) - p_{12}(y, \mathbf{z}) \\ &= p_1(y, \mathbf{z}) + p_2(y, \mathbf{z}) - \delta(y, \mathbf{z})p_1(y, \mathbf{z})p_2(y, \mathbf{z}) . \end{aligned} \quad (5.53)$$

where

$$p_{12}(y, \mathbf{z}) = p_{1|2}(y, \mathbf{z})p_2(y, \mathbf{z}) = p_{2|1}(y, \mathbf{z})p_1(y, \mathbf{z}) \quad (5.54)$$

is the probability that both observers detect an animal at distance  $y$  and with covariates  $\mathbf{z}$ . Hence we can write

$$\delta(y, \mathbf{z}) = \frac{p_{12}(y, \mathbf{z})}{p_1(y, \mathbf{z})p_2(y, \mathbf{z})} = \frac{p_{1|2}(y, \mathbf{z})}{p_1(y, \mathbf{z})} = \frac{p_{2|1}(y, \mathbf{z})}{p_2(y, \mathbf{z})} \quad (5.55)$$

Thus we now need to specify a model for  $\delta(y, \mathbf{z})$ , such that  $\delta(0, \mathbf{z}) = 1$ , and we expect  $\delta(y, \mathbf{z}) \geq 1$  for  $y \geq 0$ .

If we take the above point independence argument to its logical conclusion, then it is clear that independence holds in the limit as probability of detection tends to one; there can be no heterogeneity in probability of detection if all probabilities are one. This is the concept of *limiting independence* (Buckland et al. 2010). In principle, this allows estimation with lower bias, but in practice, the data often provide too little information to allow such models to be fitted reliably. However, the method does provide a framework for assessing formally whether a point independence model or a full independence model provides an adequate fit to the data, for example by comparing AIC values for models with full independence, point independence and limiting independence (Sect. 5.4.2.2).

When formulating the likelihood component corresponding to the detection function for MCDS, we found it convenient to condition on the covariates  $\mathbf{z}$ , as we did not know the distribution of  $\mathbf{z}$  in the population. The same applies here, and so we again have

$$\mathcal{L}_{y|\mathbf{z}} = \prod_{i=1}^n f_{y|\mathbf{z}}(y_i|\mathbf{z}_i) , \quad (5.56)$$

which we can write as

$$\mathcal{L}_{y|\mathbf{z}} = \prod_{i=1}^n \frac{p_{\cdot}(y_i, \mathbf{z}_i)\pi(y_i|\mathbf{z}_i)}{E(p_{\cdot}|\mathbf{z}_i)} \quad (5.57)$$

where  $E(p_{\cdot}|\mathbf{z}_i) = \int p_{\cdot}(y, \mathbf{z}_i)\pi(y|\mathbf{z}_i) dy$ , and  $\pi(y_i|\mathbf{z}_i)$  is the conditional probability density function of  $y$  given  $\mathbf{z}_i$ , evaluated at  $y_i$ . Given a randomized survey design, we can generally assume that  $\pi(y|\mathbf{z}_i) \equiv \pi(y)$ , independent of the covariates  $\mathbf{z}_i$ ; then we have  $\pi(y) = 1/w$  with  $0 \leq y \leq w$  for line transect sampling and  $\pi(y) = 2y/w^2$  for  $0 \leq y \leq w$  for point transect sampling.

However, we now also have a mark-recapture component to the likelihood. We denote the *capture history* of an animal by the vector  $\omega$ , comprising two elements, each of which is zero or one. For an animal detected by observer 1 but not observer 2,  $\omega = (1, 0)$ , while for an animal detected by observer 2 but not observer 1,  $\omega = (0, 1)$ . For an animal detected by both observers,  $\omega = (1, 1)$ . Note that the capture history  $\omega = (0, 0)$  is not observed, and so does not appear in the likelihood. We can now write

$$\mathcal{L}_\omega = \prod_{i=1}^n \Pr(\omega_i | \text{detected}) = \prod_{i=1}^n \frac{\Pr(\omega_i)}{p_{\cdot}(y_i | \mathbf{z}_i)} \quad (5.58)$$

where

$$\begin{aligned} \Pr(\omega_i = (1, 0)) &= p_1(y, \mathbf{z}) (1 - p_{2|1}(y, \mathbf{z})) , \\ \Pr(\omega_i = (0, 1)) &= p_2(y, \mathbf{z}) (1 - p_{1|2}(y, \mathbf{z})) , \\ \Pr(\omega_i = (1, 1)) &= p_1(y, \mathbf{z}) p_{2|1}(y, \mathbf{z}) = p_2(y, \mathbf{z}) p_{1|2}(y, \mathbf{z}) . \end{aligned}$$

By substituting these expressions into Eq. (5.58) and rearranging, we obtain:

$$\begin{aligned} \mathcal{L}_\omega &= \left\{ \prod_{i=1}^n \left[ \frac{p_2(y_i | \mathbf{z}_i)}{p_{\cdot}(y_i | \mathbf{z}_i)} \right]^{\omega_{i2}} \left[ \frac{p_1(y_i | \mathbf{z}_i) (1 - p_{2|1}(y_i | \mathbf{z}_i))}{p_{\cdot}(y_i | \mathbf{z}_i)} \right]^{1 - \omega_{i2}} \right\} \\ &\quad \times \left\{ \prod_{i=1}^{n_2} p_{1|2}(y_i | \mathbf{z}_i)^{\omega_{1i}} (1 - p_{1|2}(y_i | \mathbf{z}_i))^{1 - \omega_{1i}} \right\} \quad (5.59) \end{aligned}$$

where  $n_2$  is the number of animals detected by observer 2, and  $\omega_{1i}$  and  $\omega_{2i}$  are the elements of  $\omega_i$ , with  $\omega_{1i} = 1$  if animal  $i$  was detected by observer 1 and  $\omega_{1i} = 0$  otherwise, and similarly for  $\omega_{2i}$ . The first product is the likelihood arising from whether or not observer 2 detected each of the  $n$  recorded animals, given that at least one of the observers detected them, and the second term is the likelihood arising from whether or not observer 1 detected the  $n_2$  animals detected by observer 2, given that observer 2 detected them.

We can now specify models for the detection functions, assuming full, point or limiting independence, and fit them, for example by maximizing the product of likelihoods  $\mathcal{L}_{y|z} \times \mathcal{L}_\omega$ . See Laake and Borchers (2004) and Borchers et al. (2006) for full details. As noted by Burt et al. (2014), if we wish to use a full independence model, we could base inference on the mark-recapture component  $\mathcal{L}_\omega$  alone, including distance as a covariate.



## 5.4.2 MRDS Case Studies

### 5.4.2.1 Antarctic Crabeater Seals

To illustrate fitting a detection function using double-observer (*i.e.* mark-recapture distance sampling) methods, we consider a crabeater seals dataset collected during an Antarctic survey. The data were analysed by Borchers et al. (2006) to show use of covariates to model heterogeneity in mark-recapture distance sampling.

The survey used a helicopter to fly line transects over the ice. Front and rear observation platforms operated simultaneously and independently. There were two observers on each platform, one on each side of the helicopter. Detection distances up to 800 m were recorded. However, the flat windows of the helicopter prevented effective searching within 100 m of the line. The data have therefore been left-truncated, by removing all detections recorded as  $<100$  m, and subtracting 100 m from the remaining distances. Thus we take  $w = 800 - 100 = 700$  m.

Several covariates in addition to distance were available to model detectability. These included factor covariates observer identity (*obs*, 12 levels), group size (*size*, three levels), visibility (*vis*, three levels), ice cover (*ice*, three levels), platform (two levels), presence of glare (*glare*, two levels) and helicopter side (*side*, two levels). Continuous covariates observer experience (*exp*), time in minutes since start of flight (*ssmi*) and time in minutes on effort since last rest (*fat*) were also recorded.

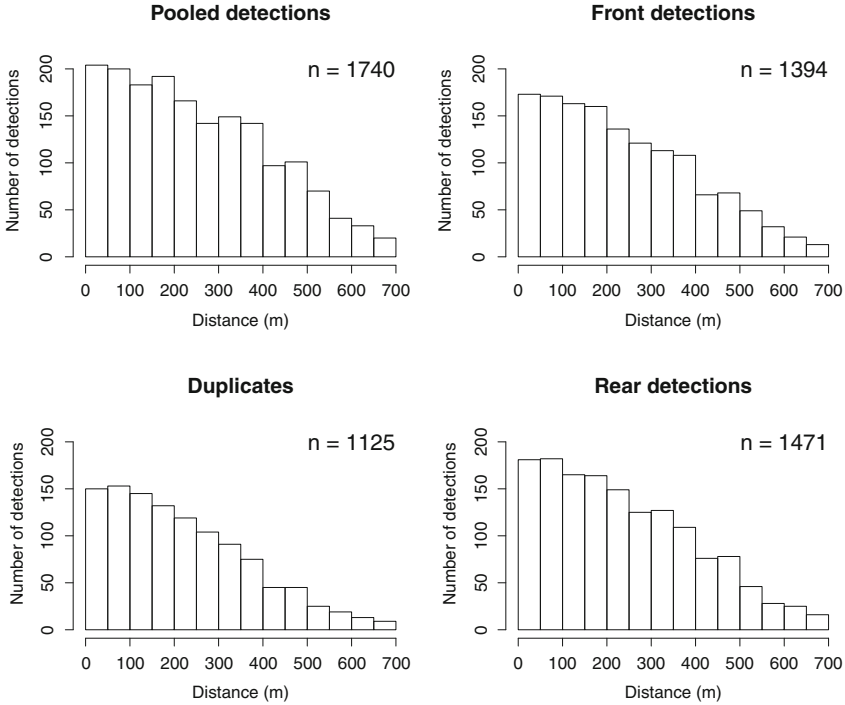
The detection distances are shown in Fig. 5.20. In total, 1740 unique groups were recorded. The front platform detected 1394 groups and the rear platform detected 1471; 1125 groups were detected by both platforms. The distance data appear well-behaved, showing a smoothly decreasing number of detections with distance from the transect line.

A standard line transect sampling analysis assuming  $g(0) = 1$  was carried out. The half-normal model was preferred over the hazard-rate model, and gave a point estimate of 2976 seal groups within the covered strips.

If we ignore the distance data and the other covariates, the two-sample Lincoln-Petersen capture–recapture model yields a much lower estimate:  $(1471 \times 1394)/1125 = 1822$  seal groups. This illustrates the danger from unmodelled heterogeneity when using a mark-recapture approach. Even though we have used a method that does not assume that all animals on the line are detected, we obtain a much smaller abundance estimate than was found using distance sampling and assuming certain detection on the line:  $g(0) = 1$ .

Note that estimates given by Borchers et al. (2006), although said to be estimated numbers of seals, are in fact estimates of the number of seal groups. The mean observed group size was 1.18 animals. There are also some minor differences in AIC values between our results and those reported by Borchers et al. (2006). The data used here are available on the book website and the results obtained can therefore be easily reproduced using the code provided.

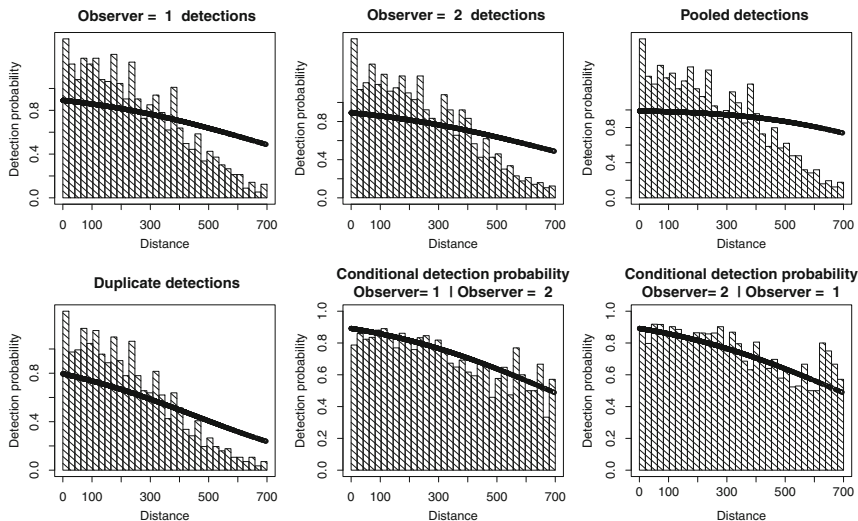
We next explicitly model the double-platform data. If we ignore any available covariates apart from distance, and fit a full independence model (Fig. 5.21), the point estimate for the number of groups in the covered area is 1918 (Table 5.10).



**Fig. 5.20** Perpendicular distances to detected groups of crabeater seals collected during an Antarctic survey

We have substantial underestimation compared with the CDS estimate, again suggesting a severe problem from unmodelled heterogeneity. When a point independence model is used (Fig. 5.22), the estimated number of groups increases to 3012, which is in much better agreement with the CDS estimate. Perhaps not surprisingly, AIC strongly favours the point independence model over the full independence model. Also not surprisingly given the fits shown in Figs. 5.21 and 5.22, the goodness-of-fit tests show a significant lack of fit for the full independence model, while the point independence model provides an adequate fit. Also, the full independence model gives the lowest standard error: we gained precision at the expense of large bias.

For a full independence analysis, the estimated detection function for platform 1, conditional on detection from platform 2, and *vice versa* (Fig. 5.21, bottom right and middle plots) are the same as the unconditional detection functions for platforms 1 and 2 (Fig. 5.21, top left and middle plots). For the point independence models, the conditional detection functions differ from the unconditional plots (Fig. 5.22, bottom right and middle plots compared with top left and middle plots). The distance sampling component estimates the probability of detection of the pooled detections



**Fig. 5.21** Fitted detection functions under the full independence model without covariates. *Top left:* estimated detection function for platform 1, scaled so that  $g(0) = 1$ , plotted on a scaled histogram of distances to seal groups detected from platform 1. *Top middle:* equivalent plot for platform 2. *Top right:* estimated detection function for the two platforms combined, scaled so that  $g(0) = 1$ , plotted on a scaled histogram of distances to seal groups detected by at least one platform. *Bottom left:* estimated probability that a seal group is detected from both platforms as a function of distance, plotted on a scaled histogram of distances to duplicate detections. *Bottom middle:* estimated probability that a seal group detected from platform 2 is also detected from platform 1, plotted on a histogram showing proportion of seal groups detected from platform 2 that were also detected from platform 1. *Bottom right:* equivalent plot for platform 2 given platform 1

(Fig. 5.22, top right plot), while the mark-recapture component defines the height of the distance sampling detection function (which would equal 1 by assumption under CDS, corresponding to distance zero).

Point independence models are strongly favoured over full independence ones. We can investigate which covariates help to explain heterogeneity in detection probabilities. Using a stepwise approach, and including at each step the covariate that gives the biggest reduction in AIC, we obtain a model which includes all the covariates available except *ssmi*; adding *ssmi* too gives a slight increase in AIC (Table 5.10). It is reassuring to see that the results are insensitive to the specific point independence model considered, despite large differences in AIC. Given that the point estimate of  $p(0)$  for both platforms combined is 0.99 (Table 5.10), this is unsurprising, as we can expect pooling robustness (Sect. 5.1.1) to ensure insensitivity, when  $p(0)$  is very close to one. This ties in with the concept of limiting independence (Sect. 5.4.1): the idea that in the limit as probability of detection tends to one, we can assume independence of detections. We illustrate this concept in the next example.

**Table 5.10** The crabeater seal double-observer results

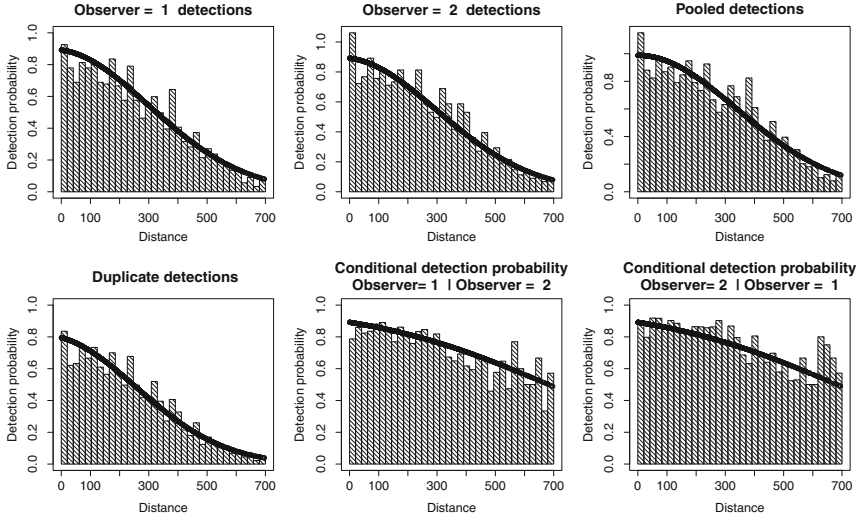
Model	$\Delta\text{AIC}$	$\hat{N}$ (groups)	$\text{se}(\hat{N})$	% CV( $\hat{N}$ )	$\hat{P}_a$	$\hat{p}(0)$
CDS HR	3.7	2518	57.0	2.27	0.69	(1.0)
CDS HN	0	2976	78.4	2.63	0.58	(1.0)
MRDS FI	419.5	1918	24.8	1.29	0.91	0.99
MRDS PI	63.9	3012	79.8	2.65	0.58	0.99
+ <i>obs</i>	27.6	3009	79.8	2.65	0.58	0.99
+ <i>fat</i>	23.9	3009	79.7	2.65	0.58	0.99
+ <i>vis</i>	14.6	3010	79.8	2.65	0.58	0.99
+ <i>side</i>	9.1	3011	79.8	2.65	0.58	0.99
+ <i>glare</i>	6.3	3011	79.8	2.65	0.58	0.99
+ <i>size</i>	3.0	3011	79.8	2.65	0.58	0.99
+ <i>exp</i>	0.0	3011	79.8	2.65	0.58	0.99
+ <i>ssmi</i>	0.2	3011	79.8	2.65	0.58	0.99

HR indicates the hazard-rate model and HN the half-normal model. FI indicates a full independence model, and PI a point independence model. Note that AIC for CDS models cannot be compared with AIC for the other models, as the CDS models do not include a mark-recapture likelihood component. A forward stepwise model selection procedure based on AIC was implemented: rows with a “+” represent the previous row model with the corresponding variable added to it. We show the point estimate for the number of groups, the corresponding standard error and % coefficient of variation, and estimates of  $P_a = \int_0^w g(y) dy$  (where  $g(y)$  is scaled so that  $g(0) = 1$ ) and  $p(0)$  (which, for the MRDS PI models, is an estimate of the average of the intercepts  $p(0, \mathbf{z})$  over the population of  $N$  seal groups in the covered region). For the CDS models,  $p(0) = 1$  by assumption

#### 5.4.2.2 Minke Whales in the North Sea

We use the northern minke whale (*Balaenoptera acutorostrata*) data analysed by Buckland et al. (2010) to illustrate limiting independence models. The second Small Cetacean Abundance in the North Sea and adjacent waters (SCANS II) survey was a multinational survey conducted in 2005 by ship and aircraft. Double-observer line transect survey methods were used because for many species, detection of animals on the trackline was expected to be less than unity. Here we analyse only shipboard survey data on minke whales. The methods used in the SCANS surveys were designed to break up the dependence between the two observers, by ensuring that they are not simultaneously searching the same patch of sea. A tracker scans with high-powered binoculars well ahead of the ship, and tracks detected animals in, to check whether the primary platform, searching with hand-held binoculars and naked eye, detects them.

Using a truncation distance of 700 m, the tracker detected 54 minke groups totalling 62 animals, while the primary platform detected 57 groups totalling 59 animals; 17 groups (19 animals) were detected by both tracker and primary platform.



**Fig. 5.22** Fitted detection functions under the point independence model without covariates. See Fig. 5.21 for explanation of the plots

We take as our full model:

$$p_j(y, z) = \frac{\exp(\lambda_{0j} + \lambda_{1j}y + \lambda_{2j}z)}{1 + \exp(\lambda_{0j} + \lambda_{1j}y + \lambda_{2j}z)} \quad (5.60)$$

where  $j = 1, 2$  indicates platform, and

$$\delta(y, z) = L(y, z) + \delta_0(y, z) [U(y, z) - L(y, z)] \quad (5.61)$$

where

$$L(y, z) = \max \left\{ 0, \frac{p_1(y, z) + p_2(y, z) - 1}{p_1(y, z)p_2(y, z)} \right\}, \quad (5.62)$$

$$U(y, z) = \min \left\{ \frac{1}{p_1(y, z)}, \frac{1}{p_2(y, z)} \right\}, \quad (5.63)$$

and

$$\delta_0(y, z) = \frac{[1 - L(y, z)] \exp(\alpha + \beta y)}{[U(y, z) - 1] + [1 - L(y, z)] \exp(\alpha + \beta y)}. \quad (5.64)$$

The covariate  $z$  is scalar in this example, and represents sea state (Beaufort).

We can use maximum likelihood methods to fit this full model, and a range of reduced models, with  $\lambda_{2j} = 0$  for  $j = 1, 2$  (models with no covariate  $z$ ), with  $\beta = 0$

(also a limiting independence model, but with less flexibility), with  $\alpha = 0$  and  $\beta \neq 0$  (a point independence model), and with  $\alpha = \beta = 0$  (a full independence model). We can also remove the dependence of the  $\lambda$ 's on  $j$ , so that the same detection function model is assumed for both platforms. We return to this case study in Sect. 6.4.4.4, where we show abundance estimates under each model.

### 5.4.2.3 Double-Observer Point Transects for Songbirds

In the same way that perfect detectability on the transect line can be violated, so too can animals be missed at distance zero from point transects. Double-observer methods can be employed with point transects to overcome this violation of classical point transect estimation. The analysis framework for double-observer point transects is the same as for line transects. There is a distance sampling element and a mark-recapture component to the likelihood. Covariates are included in the mark-recapture component so as to account for heterogeneity in the detection process. There are also models that accommodate full and point independence between the observers.

Double-observer methods for point transect data that incorporate only distance as a covariate do not adequately model heterogeneity in the detection process and hence tend to produce estimates of density with negative bias, as occurs when  $g(0) < 1$  but is assumed to equal one (Laake et al. 2011).

In songbird surveys conducted by Collier et al. (2013), 453 point transects were sampled with a fixed truncation distance of 100 m and a 5-min sampling visit to each point during spring of 2011 on the Fort Hood military base in Texas. Aural detections of singing males were recorded in binned distance categories. Data from endangered golden-cheeked warblers (*Setophaga chrysoparia*) and black-capped vireos (*Vireo atricapilla*) were recorded.

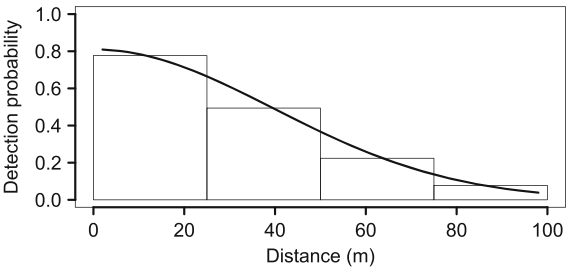
Several candidate models were fitted to the two datasets; we present a subset of models fitted for the black-capped vireos. Full independence models (with only a mark-recapture component) and point independence models were fitted to the black-capped vireo data. For point independence models, hazard-rate and half-normal key functions with no covariates were considered. In both cases, distance and observer were included as covariates in the mark-recapture model. For the mark-recapture model, main effects of observer and distance along with an interaction between observer and distance were considered in the candidate model set (Table 5.11).

The goodness-of-fit  $\chi^2$  statistic for the selected model (half-normal, distance $\times$ observer) shown in Fig. 5.23 was  $\chi^2 = 0.59$  with 2 df ( $p = 0.75$ ) for the distance sampling component and  $\chi^2 = 7.88$  with 4 df ( $p = 0.10$ ) for the mark-recapture component.

**Table 5.11** Model selection results for double-observer point transect survey of black-capped vireos

Model description	Parameters	AIC	$\Delta$ AIC
Half-normal, <i>distance</i> $\times$ <i>observer</i>	5	1875.61	0.00
Half-normal, <i>distance</i>	3	1876.56	0.95
Half-normal, <i>distance</i> + <i>observer</i>	4	1877.86	2.26
Hazard-rate, <i>distance</i>	4	1880.06	4.46
Hazard-rate, <i>distance</i> + <i>observer</i>	5	1881.37	5.76
Half-normal, <i>distance</i> (FI)	2	2043.49	167.88
Half-normal, <i>distance</i> + <i>observer</i> (FI)	3	2044.89	169.29
Half-normal, <i>distance</i> $\times$ <i>observer</i> (FI)	4	2046.57	170.96

Non-italicized terms are key functions of the distance sampling models, italicized terms define the mark-recapture model. FI indicates full independence. Data and analysis from Collier et al. (2013)



**Fig. 5.23** Unconditional detection function for pooled observers fitted to point transect data of black-capped vireo males from Fort Hood in 2011. The detection function was point independence and mark-recapture model with main effects of distance and observer, with an interaction of distance and observer. Note that the y-intercept corresponds to  $\hat{g}(0) = 0.81$ . From Collier et al. (2013)

# Chapter 6

## Design-Based Estimation of Animal Density and Abundance

### 6.1 Introduction

Conventional distance sampling may be thought of as an extension of plot sampling (Sect. 1.4), to allow imperfect detection of animals on the plots. Standard plot sampling is purely design-based: we use a randomized design to ensure that the density on plots, which we observe, can be taken as an estimate of density in the wider study area. By contrast, conventional distance sampling is a mix of design-based and model-based methods. Models are proposed for the detection function, and are fitted to the distance data using maximum likelihood methods. This component of estimation is therefore model-based. However, the likelihood maximized is not the full likelihood, but a conditional likelihood: the likelihood of the distances  $y$  from the line or point, conditional on the number  $n$  of animals detected.

In Chap. 7, we will explore modelling both the distances  $y$  and the sample size  $n$ , using separate likelihoods in a two-stage model-based approach, and in Chap. 8, we will consider methods that use an integrated likelihood, allowing both components to be modelled simultaneously.

Conceptually, we can think of a *full likelihood* approach for conventional distance sampling using the product  $\mathcal{L}_n \times \mathcal{L}_y$  of likelihood components, where for line transect data,  $\mathcal{L}_y$  is as given in Eq. (5.13) (exact distances) or Eq. (5.26) (grouped distances). The equivalent likelihood components for point transect sampling are given by Eqs. (5.30) and (5.37). The other component,  $\mathcal{L}_n$ , is the likelihood corresponding to the count  $n$  of animals. If we ignore this component, then we use the design-based methods of this chapter to estimate abundance: we use properties of the design to extrapolate from the surveyed plots to estimate density and abundance for the whole study area.

For MCDS, the full likelihood is given by  $\mathcal{L}_n \times \mathcal{L}_z \times \mathcal{L}_{y|z}$  where  $\mathcal{L}_z$  is the component of the likelihood corresponding to covariates  $\mathbf{z}$ , and  $\mathcal{L}_{y|z}$  is the contribution to the likelihood of the distances  $y$ , given the covariates  $\mathbf{z}$  of Eq. (5.42). However, there is generally no benefit in attempting to model the unknown distri-



bution of covariates  $\mathbf{z}$ , and so for a model-based approach, we use the likelihood  $\mathcal{L}_n \times \mathcal{L}_{y|z}$ , while for a design-based approach, we use the conditional likelihood  $\mathcal{L}_{y|z}$  of Eq. (5.42). Similarly, for MRDS, in the model-based approach, we usually use the likelihood  $\mathcal{L}_n \times \mathcal{L}_\omega \times \mathcal{L}_{y|z}$ , and for the design-based method of estimating abundance, we use  $\mathcal{L}_\omega \times \mathcal{L}_{y|z}$ .

In this chapter, we explore the design-based approaches. These tend to be more robust for estimating abundance and quantifying precision than the model-based approaches, but offer less flexibility and fewer options. The design-based methods have also been in existence far longer than model-based methods; the latter largely date from Borchers et al. (2002), Hedley and Buckland (2004) and Royle et al. (2004), although Högmänder (1991) made early progress in developing spatial distance sampling methods.

## 6.2 Plot Sampling

In Sect. 1.4.1, we showed that for strip transect sampling,  $\hat{D} = \frac{n}{2wL}$  and  $\hat{N} = \frac{nA}{2wL}$ , where  $n$  is the number of animals detected,  $w$  is the strip half-width,  $L$  is the total length of the strips, and  $A$  is the size of the study area. For circular plot sampling with  $K$  plots, we obtained  $\hat{D} = \frac{n}{K\pi w^2}$  and  $\hat{N} = \frac{nA}{K\pi w^2}$ .

### 6.2.1 Quantifying Precision

In the above equations, every quantity is assumed to be a known, fixed constant, except  $n$ : if the survey were repeated, we would expect  $n$  to change, but other quantities would remain fixed. Thus it is variation in  $n$  that we wish to quantify, to allow us to quantify the precision of estimated density or abundance. It is convenient instead to quantify precision of the *encounter rate*, defined to be  $n/L$  for strip transect sampling and  $n/K$  for circular plot sampling. We adopt design-based estimates of encounter rate standard errors; these are based on sample standard errors of encounter rate, where the sampling units are the plots. In the case of strip transect sampling, we use a weighted standard error, where weights are equal to line lengths  $l_k$ :

$$\text{se}\left(\frac{n}{L}\right) = \sqrt{\frac{K}{L^2(K-1)} \sum_{k=1}^K l_k^2 \left(\frac{n_k}{l_k} - \frac{n}{L}\right)^2} \quad (6.1)$$

where  $n_k$  is the number of animals on plot  $k$  (Fewster et al. 2009). The equivalent expression for circular plot sampling is simpler because all points have equal weight:

$$\text{se}\left(\frac{n}{K}\right) = \sqrt{\frac{1}{K(K-1)} \sum_{k=1}^K \left(n_k - \frac{n}{K}\right)^2}. \quad (6.2)$$

It follows that, for strip transect sampling,

$$\text{se}(\hat{D}) = \frac{\text{se}(n/L)}{2w} \quad (6.3)$$

and for circular plot sampling,

$$\text{se}(\hat{D}) = \frac{\text{se}(n/K)}{\pi w^2}. \quad (6.4)$$

In both cases,  $\text{se}(\hat{N}) = A \cdot \text{se}(\hat{D})$ .

We could obtain an approximate 95 % confidence interval for  $D$  as  $\hat{D} \pm 1.96 \text{se}(\hat{D})$ . However, when precision is low, this symmetric interval can be a poor approximation. We obtain a better interval by assuming that  $\hat{D}$  is log-normally distributed (Buckland et al. 2001, p. 77):

$$(\hat{D}/C, \hat{D} \cdot C) \quad (6.5)$$

where

$$C = \exp \left[ 1.96 \text{se} \left( \log_e \hat{D} \right) \right] \quad (6.6)$$

and

$$\text{se} \left( \log_e \hat{D} \right) = \sqrt{\log_e \left[ 1 + \left( \text{cv}(\hat{D}) \right)^2 \right]}. \quad (6.7)$$

The coefficient of variation of  $\hat{D}$  is  $\text{cv}(\hat{D}) = \frac{\text{se}(\hat{D})}{\hat{D}}$ . Especially when  $K$  is small (say under 20), confidence interval coverage can be improved by replacing the upper 2.5 percentile of the standard normal distribution (1.96) by the upper 2.5 percentile of Student's  $t$  distribution with  $K - 1$  degrees of freedom.

## 6.2.2 The Montrave Case Study: Plot Sampling

The Montrave case study provides data to illustrate plot sampling. A key assumption of plot sampling is that all animals on the plot are detected. As we saw in Sects. 5.2.2.3 and 5.2.3.3, it is reasonable to assume that all, or almost all, male European robins within 35 m of the line or point are detected. Thus to illustrate circular plot sampling and strip transect sampling, we will truncate distances from the line or point at 35 m.

### 6.2.2.1 Strip Transect Sampling

We use the line transect robin data to illustrate strip transect sampling. Each line was walked twice, and the sightings data pooled. Thus the effort expended (see

**Table 6.1** Summary of European robin counts and line lengths by plot for strip transect sampling from the Montrave study

Line	1	2	3	4	5	6	7	8	9	10
Line length (m)	208	401	401	299	350	401	393	405	385	204
Count	2	5	9	2	6	4	2	2	1	0

Line	11	12	13	14	15	16	17	18	19	
Line length (m)	39	47	204	271	236	189	177	200	20	
Count	0	0	2	3	3	0	4	2	0	

Sect. 6.3.3.1) equates to twice the line length, and we need to double the plot size to account for this. This is not the same as doubling the number of lines; we have 19 lines, and estimated precision of encounter rate is obtained by quantifying variability in encounter rate among those 19 lines.

The data appear in Table 6.1. These can be obtained from Table 1.1 by counting for each line the number of robins detected within 35 m of the line. Plot size for line  $k$  is the length of that line multiplied by twice the half-width:  $2wl_k = 0.07l_k \text{ km}^2$ . Total area covered is thus  $0.07L \text{ km}^2$ , where  $L = \sum_k l_k = 4.83 \text{ km}$ . We need to double this to take account of the two visits to each line, so that Eq. (1.1) becomes:

$$\begin{aligned}\hat{D} &= \frac{n}{4wL} = \frac{2 + 5 + 9 + \cdots + 0}{4 \times 0.035 \times (0.208 + 0.401 + \cdots + 0.020)} \\ &= \frac{47}{4 \times 0.035 \times 4.83} = 69.51\end{aligned}\quad (6.8)$$

and so we estimate that there are around 70 robin territories  $\text{km}^{-2}$ .

From Eq. (6.3), and remembering to double the plot size to take account of the two visits, we find

$$\text{se}(\hat{D}) = \frac{1.6887}{4 \times 0.035} = 12.06. \quad (6.9)$$

Applying Eq. (6.5) gives a 95 % confidence interval of approximately (50, 97) territories  $\text{km}^{-2}$ . (If instead of 1.96, we use the critical value of Student's  $t$  distribution with 18 degrees of freedom, we get a confidence interval of (49, 99) territories.) This compares with an estimate of 84 from territory mapping. These results may be reproduced in Distance as explained in the Montrave case study on the book website.

### 6.2.2.2 Circular Plot Sampling

We use the robin data from the snapshot point transect sampling method to illustrate circular plot sampling. Each plot was visited twice for the snapshot method. Thus counts, summed across visits, are on average twice what they would be, if there

**Table 6.2** Summary of European robin counts by plot for circular plot sampling from the Montrave study

Number of animals on plot	0	1	2	$\geq 3$
Frequency	22	7	3	0

were just a single visit to each plot. We address this by defining the effort to be 2 for each point (Sect. 6.3.3.2). We then multiply the plot size by the effort. Note that precision is quantified from variation in counts across plots, not from variation in the two counts from the same plot; those two counts cannot be assumed independent.

We show the data in Table 6.2. These data are obtained from Table 1.2 by counting the number of points that have 0, 1 or 2 birds detected within 35 m of the point. (No points had more than two robins within 35 m.) Plot size is the size of a circle of radius  $w = 35$  m:  $\pi w^2 = 0.00384845$  km<sup>2</sup>. Applying Eq. (1.3) with  $n = 7 \times 1 + 3 \times 2 = 13$  robins,  $K = 32$  points, and multiplying the plot size by the effort 2, we obtain

$$\hat{D} = \frac{13}{32 \times 2 \times 0.00384845} = 52.78. \quad (6.10)$$

Hence we estimate that there are around 53 robin territories km<sup>-2</sup>.

From Eq. (6.2), we obtain  $\text{se}(n/K) = 0.1175984$ . Substituting this into Eq. (6.4), and remembering to double the plot size to take account of the two visits, we find

$$\text{se}(\hat{D}) = \frac{0.1175984}{2 \times 0.00384845} = 15.27867. \quad (6.11)$$

Applying Eq. (6.5) gives a 95 % confidence interval of approximately (30, 92) territories km<sup>-2</sup>. (Using the critical value from Student's t distribution with 31 degrees of freedom gives (30, 94) territories.) The Montrave case study on the book website explains how to reproduce these analyses in the Distance software.

Territory mapping located 28 territories, equating to a density of around 84 territories km<sup>-2</sup>. This lies towards the upper end of the above confidence interval.

## 6.3 Conventional Distance Sampling

### 6.3.1 Line Transect Sampling

#### 6.3.1.1 Estimating Density and Abundance

Having fitted a model for the detection function, we can use Eq. (1.5) to estimate density:

$$\hat{D} = \frac{n}{2wL\hat{P}_a} \quad (6.12)$$

where  $\hat{P}_a$  is the maximum likelihood estimate of the proportion  $P_a$  (Sect. 5.2.2). From Eqs. (5.24) and (5.25), we obtain:

$$\hat{D} = \frac{n}{2wL\hat{P}_a} = \frac{n}{2\hat{\mu}L} = \frac{n\hat{f}(0)}{2L}. \quad (6.13)$$

Thus, having fitted our model for  $f(x)$  to the observations, we can estimate the function at  $x = 0$ , giving us  $\hat{f}(0)$ , and hence  $\hat{D}$ . As before,  $\hat{N} = \hat{D} \cdot A$ .

In the case of our example with a half-normal detection function from Sect. 5.2.2,  $\hat{f}(0) = \frac{1}{\hat{\mu}} = \sqrt{\frac{2n}{\pi \sum_{i=1}^n x_i^2}}$  and  $\hat{D} = \sqrt{\frac{n^3}{2\pi L^2 \sum_{i=1}^n x_i^2}}$ .

### 6.3.1.2 Quantifying Precision

Estimating the standard error of  $\hat{D}$  is not as straightforward as for plot sampling, as two components of Eq. (6.12) are estimates:  $n$  and  $\hat{P}_a$ . We use the approximate result that squared coefficients of variation add when we have a product of independent estimates (Buckland et al. 2001, p. 66). Hence

$$[cv(\hat{D})]^2 = [cv(n/L)]^2 + [cv(\hat{P}_a)]^2. \quad (6.14)$$

Using the same method of approximation, we also have

$$cv(\hat{P}_a) = cv(\hat{\mu}) = cv\{\hat{f}(0)\} \quad (6.15)$$

so that

$$se(\hat{D}) = \hat{D} \sqrt{[cv(n/L)]^2 + [cv\{\hat{f}(0)\}]^2} \quad (6.16)$$

where standard maximum likelihood theory is used to estimate  $se\{\hat{f}(0)\}$  and hence  $cv\{\hat{f}(0)\}$  (Buckland et al. 2001, p. 62), and  $cv(n/L)$  is obtained by dividing Eq. (6.1) by the encounter rate  $n/L$ . As before,  $se(\hat{N}) = A \cdot se(\hat{D})$ . Confidence intervals are found as in Sect. 6.2.1, except that, if percentiles of the Student's *t* distribution are used in preference to percentiles of the standard normal distribution, the degrees of freedom are more complex, using a Satterthwaite correction (Buckland et al. 2001, p. 78). This correction is used by *Distance*.

A more direct method of variance estimation is to use the bootstrap. By taking lines as the sampling unit, this method has the advantage that variance in encounter rate and in estimated detectability are simultaneously addressed, without the need to assume that these quantities are independent. The method also automatically incorporates variance in estimated mean cluster size for clustered populations.

To implement the bootstrap with lines as sampling units, we number the lines from 1 to  $K$ . To generate a single resample, we draw numbers *with replacement* from these, until we again have  $K$  lines. This creates variation from one resample to another because some lines will appear in the resample more than once, while others will not appear at all. Having generated a resample, it is analysed exactly as for the real data. Suppose we denote the density estimate for this resample by  $\hat{D}_1$ . We now generate another resample and another, until we have  $B$  bootstrap estimates,  $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_B$ , where  $B$  is large, typically of the order of 1000. Our bootstrap standard error of  $\hat{D}$  is simply the square root of the sample variance of the  $\hat{D}_b$ :

$$\text{se}(\hat{D}) = \sqrt{\frac{\sum_{b=1}^B (\hat{D}_b - \bar{D})^2}{K - 1}} \quad (6.17)$$

where  $\bar{D} = (\sum_{b=1}^B \hat{D}_b) / B$  is the mean of the bootstrap estimates. Further, if we order the bootstrap estimates, denoting the ordered list by  $\hat{D}_{(1)}, \hat{D}_{(2)}, \dots, \hat{D}_{(B)}$ , where  $\hat{D}_{(1)}$  is smallest and  $\hat{D}_{(B)}$  the largest, the percentile method (Buckland 1984) gives a simple but robust  $100(1 - 2\alpha)\%$  confidence interval for  $D$  as  $(\hat{D}_{(r)}, \hat{D}_{(s)})$ , where  $r = (B + 1)\alpha$  and  $s = (B + 1)(1 - \alpha)$ . For example, if we take  $B = 999$  and  $\alpha = 0.025$ , then an approximate 95% confidence interval is given by  $(\hat{D}_{(25)}, \hat{D}_{(975)})$ .

Note that the above bootstrap implementation assumes that the lines are a simple random sample of possible lines within the study area, whereas usually, we adopt a systematic design. Typically, this leads to overestimation of the true variance — systematic samples tend to give higher precision than do random samples, due to better coverage of the study area — but usually, the overestimation is modest.

We can also use the bootstrap for a simple approach to model averaging. We generate resamples as above, but when analysing each resample, we reassess which model is best, so that a different model might be used to analyse the resample as was selected for the original data. Apart from this, the process proceeds exactly as before; now, our bootstrap standard errors and confidence intervals will incorporate model uncertainty. We can also take the average of the bootstrap estimates as a model-averaged estimate of any parameter of interest, such as density  $D$ . The software *Distance* allows the user to implement this approach. Its default method for selecting the best model for each resample is to use AIC.

### 6.3.1.3 Estimation When Animals Occur in Clusters

Animals often occur in groups which we term clusters. Provided that these are well-defined, then distances  $x_1, \dots, x_n$  of the group centres from the line should be recorded, along with group sizes  $s_1, \dots, s_n$ . Equation (6.13) now gives an estimate of the density of animal clusters. To convert this to an estimate of animal density,

we need to multiply by an estimate of mean cluster size in the population, which we denote by  $\hat{E}(s)$ . The simplest estimate is the sample mean of sizes of detected clusters:  $\hat{E}(s) = \bar{s}$ . However, there are two reasons why this is not necessarily a good estimate. Firstly, larger clusters are typically more detectable than smaller ones, especially at larger distances. Hence the mean size of *detected* clusters tends to be larger than the mean size of all clusters — an example of *size-biased sampling*. Secondly, although we assume that cluster size is accurately recorded, in practice, this may not be the case. Typically, the size of clusters far from the line tends to be underestimated. These biases operate in different directions, reducing the bias in using  $\bar{s}$ . However, we prefer the default option in `Distance`, which uses regression to estimate the mean size of detected clusters on the line, by regressing the logarithm of cluster size on the estimated probability of detection, given the cluster's distance from the line. Provided the sizes of clusters on the line are recorded without error, and all clusters on the line are detected (which will be true when  $g(0) = 1$ ), then this method corrects for both sources of bias. This and other methods of estimating mean cluster size and the corresponding standard error are described in Buckland et al. (2001, pp. 71–76).

Given our estimate  $\hat{E}(s)$  of mean cluster size, animal density is now estimated as

$$\hat{D} = \frac{n\hat{f}(0)\hat{E}(s)}{2L} \quad (6.18)$$

with standard error

$$\text{se}(\hat{D}) = \hat{D} \sqrt{[\text{cv}(n/L)]^2 + [\text{cv}\{\hat{f}(0)\}]^2 + [\text{cv}\{\hat{E}(s)\}]^2}. \quad (6.19)$$

For estimating  $\text{cv}\{\hat{E}(s)\}$ , see Buckland et al. (2001, p. 74).

### 6.3.1.4 The Montrave Case Study

We show estimated densities, with corresponding coefficients of variation and confidence intervals, corresponding to the line transect data for robins in Table 6.3. We saw previously (Table 5.1) that the hazard-rate model gives a slightly higher estimated effective strip half-width, and we see now that this results as expected in a correspondingly lower estimate of density. The coefficient of variation of density for this model is also smaller. Overall, the density estimates are rather below the territory mapping estimate of 84 territories  $\text{km}^{-2}$ . This value lies comfortably inside the confidence interval for both the uniform-cosine and half-normal-Hermite models, but just outside that for the hazard-rate model.

If we wished to take account of model uncertainty, we could evaluate AIC weights for each model (Eq. (5.4)), then take a weighted average of the density estimates. Thus for the uniform-cosine model, we have  $w_1 = \exp(0/2)/[\exp(-0/2) + \exp(-0.64/2) + \exp(-0.57/2)] = 0.404$ . Similarly, the

**Table 6.3** Density estimates corresponding to the three detection function models applied to the robin line transect data from the Montrave case study

Model	$\Delta\text{AIC}$	$\hat{D}$	$\text{cv}(\hat{D})$	95 % CI for $D$
Uniform-cosine	0.00	68.5	0.193	(46.9, 100.2)
Half-normal-Hermite	0.64	69.7	0.220	(45.2, 107.4)
Hazard-rate	0.57	64.2	0.131	(49.4, 83.5)

Units of density estimates are territories  $\text{km}^{-2}$

weights for the half-normal-Hermite and hazard-rate models are  $w_2 = 0.293$  and  $w_3 = 0.303$  respectively. This gives a weighted estimate of density of 67.5 territories  $\text{km}^{-2}$ . Estimating precision of the weighted estimate is not straightforward (Buckland et al. 1997; Burnham and Anderson 2002). However, as noted in Sect. 6.3.1.2, the bootstrap offers a simple method of model-averaging for which precision is readily estimated. When this was carried out in *Distance* using the above three models and with 999 resamples, the mean bootstrap estimate of density was 68.0 territories  $\text{km}^{-2}$ , close to the above weighted average. The corresponding 95 % confidence interval was (47.0, 98.6) when a log-normal confidence interval was calculated, based on the bootstrap standard error, and (44.4, 97.4) when the percentile method (Buckland 1984) was used.

When using standard distance sampling methods, model-averaging is seldom worthwhile, unless different plausible models give appreciably different estimates. This seldom arises given good quality data for line transect sampling, but is more likely to occur for point transect sampling.

## 6.3.2 Point Transect Sampling

### 6.3.2.1 Estimating Density and Abundance

Given an estimate  $\hat{P}_a$  of  $P_a$ , we use Eq. (1.7) to estimate density:

$$\hat{D} = \frac{n}{K\pi w^2 \hat{P}_a}. \quad (6.20)$$

Thus, having fitted our model for  $f(r)$ , we can estimate  $h(0)$  (by  $\hat{h}(0)$ ), and, using Eqs. (5.35) and (5.36), we obtain

$$\hat{D} = \frac{n}{K\pi w^2 \hat{P}_a} = \frac{n}{K\hat{v}} = \frac{n\hat{h}(0)}{2\pi K}. \quad (6.21)$$

Again,  $\hat{N} = \hat{D} \cdot A$ .



For clustered populations, we simply multiply the estimated density of clusters by an estimate of the mean cluster size in the population, exactly as for line transect sampling.

### 6.3.2.2 Quantifying Precision

We again use the approximate result that squared coefficients of variation add when we have a product of independent estimates (Buckland et al. 2001, p. 66). Hence

$$[\text{cv}(\hat{D})]^2 = [\text{cv}(n/K)]^2 + [\text{cv}(\hat{P}_a)]^2 \quad (6.22)$$

and

$$\text{cv}(\hat{P}_a) = \text{cv}(\hat{\nu}) = \text{cv}\{\hat{h}(0)\} \quad (6.23)$$

so that

$$\text{se}(\hat{D}) = \hat{D} \sqrt{[\text{cv}(n/K)]^2 + [\text{cv}\{\hat{h}(0)\}]^2} \quad (6.24)$$

where standard maximum likelihood theory is used to estimate  $\text{se}\{\hat{h}(0)\}$  and hence  $\text{cv}\{\hat{h}(0)\}$  (Buckland et al. 2001, p. 62), and  $\text{cv}(n/K)$  is obtained by dividing Eq. (6.2) by the encounter rate  $n/K$ . We again have  $\text{se}(\hat{N}) = A \cdot \text{se}(\hat{D})$ . Confidence intervals are found as in Sect. 6.2.1. Again, if percentiles of Student's  $t$  are preferred to those of the standard normal distribution, then a Satterthwaite correction can be used (Buckland et al. 2001, p. 78). Software `Distance` uses this correction.

For clustered populations, the squared coefficient of variation for mean cluster size is added to the expression inside the square root of Eq. (6.24) exactly as for line transect sampling.

The bootstrap can be implemented as for line transect sampling, except now, we resample points instead of lines. If the survey design has clusters of points (Sect. 2.1), then those clusters should be taken as the resampling unit. Similarly, if points are placed along lines, with smaller separation of neighbouring points on the same line than for neighbouring points on different lines (left-hand plot of Fig. 2.2), then the lines should be resampled, as points on the same line cannot be considered independent.

### 6.3.2.3 The Montrave Case Study

Density estimates, coefficients of variation and confidence intervals for three models for the robin snapshot point transect data are shown in Table 6.4. As judged by AIC, the hazard-rate model is best. We see that the density estimates are slightly lower

**Table 6.4** Density estimates corresponding to the three detection function models applied to the robin snapshot point transect data from the Montrave case study

Model	$\Delta\text{AIC}$	$\hat{D}$	$\text{cv}(\hat{D})$	95 % CI for $D$
Uniform-cosine	2.78	65.1	0.164	(46.8, 90.6)
Half-normal-Hermite	2.68	62.5	0.593	(20.8, 187.8)
Hazard-rate	0.00	59.7	0.229	(38.1, 93.6)

Units of density estimates are territories  $\text{km}^{-2}$

than those obtained using line transect sampling, with the hazard-rate model again giving the lowest estimate, and the half-normal-Hermite model gives poor precision. The territory mapping estimate of 84 territories  $\text{km}^{-2}$  is towards the upper end of the confidence intervals for the uniform-cosine and hazard-rate models.

The hazard-rate AIC model weight is 0.66, with the other two models each having a weight of around 0.17. If we use the bootstrap in *Distance* to estimate density and calculate confidence intervals allowing for model uncertainty, we obtain a model-averaged estimated of 59.3 territories  $\text{km}^{-2}$ , which is actually slightly lower than the hazard-rate estimate of Table 6.4. The corresponding confidence interval is (39.3, 89.3) using the log-based interval, and (35.2, 85.1) using the percentile method. These are again close to the hazard-rate estimates in Table 6.4, and just include the territory mapping estimate of 84 territories  $\text{km}^{-2}$ .

### 6.3.3 Generalizing the Estimation Equations

#### 6.3.3.1 Effort in Line Transect Sampling

In the above derivations, we have assumed that *effort* expended is equal to the total line length  $L$ . Suppose that we now denote effort by  $T$ , where  $T$  is the total length of line travelled. Thus  $T = \sum_{k=1}^K t_k l_k$ , where  $t_k$  is the number of times that line  $k$  was surveyed, and  $l_k$  is the length of line  $k$ . (If line  $k$  is surveyed only once, and just one side of the line is recorded, then  $t_k = 0.5$  for that line.)

Density is now estimated as

$$\hat{D} = \frac{n\hat{f}(0)}{2T} \quad (6.25)$$

where  $n$  is the total number of animals detected. The standard error of  $\hat{D}$  is found as before, except that Eq. (6.1) is replaced by:

$$\text{se}\left(\frac{n}{T}\right) = \sqrt{\frac{K}{T^2(K-1)} \sum_{k=1}^K (t_k l_k)^2 \left(\frac{n_k}{t_k l_k} - \frac{n}{T}\right)^2}. \quad (6.26)$$

If the  $t_k$  vary by line, then the above estimates assume that the number of times that a line is surveyed is independent of animal density.

### 6.3.3.2 Effort in Point Transect Sampling

We assumed that effort  $T = K$  in the above point transect derivations. We now take  $T = \sum_{k=1}^K t_k$ , where  $t_k$  is the number of times that point  $k$  was surveyed. (If say only one quarter of point  $k$  was surveyed once, then set  $t_k = 0.25$ .)

Density is now estimated as

$$\hat{D} = \frac{n\hat{h}(0)}{2\pi T} \quad (6.27)$$

where  $n$  is the total number of animals detected. The standard error of  $\hat{D}$  is found as before, except that Eq. (6.2) is replaced by:

$$\text{se}\left(\frac{n}{T}\right) = \sqrt{\frac{K}{T^2(K-1)} \sum_{k=1}^K t_k^2 \left(\frac{n_k}{t_k} - \frac{n}{T}\right)^2}. \quad (6.28)$$

If the  $t_k$  vary by point, then we have assumed that the number of times that a point is surveyed is independent of animal density.

## 6.4 Horvitz–Thompson-Like Estimators

### 6.4.1 The Horvitz–Thompson Estimator

The Horvitz–Thompson estimator (Horvitz and Thompson 1952) is a general but simple design-based estimator. In the context of estimating animal abundance, suppose that for each animal  $i$  in the sample, we knew the probability  $p_i$  that it would be one of the  $n$  animals from the population of size  $N$  to be recorded. Then we can estimate abundance  $N$  as

$$\hat{N} = \sum_{i=1}^n \frac{1}{p_i}. \quad (6.29)$$

For clustered populations, if  $n$  animal clusters are recorded, Eq. (6.29) gives the abundance of clusters, and animal abundance is estimated by

$$\hat{N} = \sum_{i=1}^n \frac{s_i}{p_i} \quad (6.30)$$

where  $s_i$  is the size of the  $i^{\text{th}}$  detected cluster.

For plot sampling, we know the probability  $p_i$  that any given animal is recorded: it is simply the probability that the animal is on one of the sample plots, and given a random sample of plots, this is simply  $a/A$  for each animal  $i$  in the sample, where  $a$  is the total area of sample plots and  $A$  is the size of the study area. Substituting this into Eq. (6.29), we obtain

$$\hat{N} = \frac{nA}{a} . \quad (6.31)$$

For strip transect sampling, the total area of the strips is  $a = 2wL$ , where  $w$  is the strip half-width and  $L$  is the total length of strips. Substituting this value into Eq. (6.31), we obtain Eq. (1.2). Similarly, for circular plot sampling with  $K$  plots of radius  $w$ , we have  $a = K\pi w^2$ , giving Eq. (1.4).

In distance sampling, we have the complication that the probabilities  $p_i$  are not known, and must be estimated. In this case, the estimator is termed ‘Horvitz–Thompson-like’ (Borchers 1996). These probabilities now have two components: one is the probability that the animal is on a sample plot, which, given a randomized design, is known; and the other is the probability of detection, given that the animal is on the plot, which we must estimate.

The Horvitz–Thompson estimator is unbiased. However, when we must estimate the  $p_i$ , the estimator is no longer unbiased in general, although it remains asymptotically unbiased (*i.e.* the bias tends to zero as sample size tends to infinity) because we estimate detectability using maximum likelihood methods, which give asymptotically unbiased estimators of the  $p_i$ . However, for small or even moderate sample sizes, bias may be substantial when the  $p_i$  are very variable. This is because small values of  $p_i$  lead to much larger contributions to  $\hat{N}$  in Eq. (6.29) than do larger values. When the  $p_i$ ’s are known, this may lead to poor precision in  $\hat{N}$  but does not cause bias, but when they are estimated, relatively small errors in estimating small  $p_i$  can generate large bias in  $\hat{N}$ .

### 6.4.2 The Horvitz–Thompson-Like Estimator for Conventional Distance Sampling

For conventional distance sampling, perhaps the most obvious estimate of  $p_i$  is  $\hat{p}_i = \frac{a\hat{g}(y_i)}{A}$  where  $\hat{g}(y_i)$  is the estimated probability of detection of animal  $i$ , which is at distance  $y_i$  from the line or point. However, the  $\hat{g}(y_i)$  may be very variable, especially if a large truncation distance  $w$  is chosen. Thus by using this estimator, we risk high bias. Fortunately, we have a better option. We know the distribution of  $y_i$  in the population (uniform for line transect sampling and triangular for point transect sampling), and this allows us to replace the estimated detection probabilities  $\hat{g}(y_i)$  by an estimate of their mean. For line transect sampling, we have

$$P_a = \int_0^w g(x) \frac{1}{w} dx = \frac{\mu}{w} \quad (6.32)$$

where  $\mu = \int_0^w g(x) dx$ , and hence  $\hat{P}_a = \hat{\mu}/w$ , which we saw in Eq. (5.24). Thus

$$\hat{p}_i = \frac{a\hat{P}_a}{A} \quad (6.33)$$

for each detected animal  $i$ , and substituting this into Eq. (6.29), we obtain

$$\hat{N} = \frac{nA}{a\hat{P}_a} \quad (6.34)$$

which is the same estimator of  $N$  that we obtained by multiplying  $\hat{D}$  from Eq. (6.13) by size  $A$  of the study area.

Similarly for point transect sampling,

$$P_a = \int_0^w g(r) \frac{2r}{w^2} dy = \frac{\nu}{\pi w^2} \quad (6.35)$$

where  $\nu = 2\pi \int_0^w r.g(r) dr$ , and hence  $\hat{P}_a = \hat{\nu}/(\pi w^2)$ , which agrees with Eq. (5.31). Again

$$\hat{p}_i = \frac{a\hat{P}_a}{A} \quad (6.36)$$

for each detected animal  $i$ , which we substitute into Eq. (6.29), as for line transect sampling. Thus conventional distance sampling is just a special case of the general Horvitz–Thompson-like estimator.

For clustered populations, Eq. (6.30) reduces to

$$\hat{N} = \frac{A \sum_i s_i}{a\hat{P}_a} = \frac{nA\bar{s}}{a\hat{P}_a} \quad (6.37)$$

which is conventional distance sampling in which the mean cluster size in the population is estimated by the sample mean — which is subject to size bias if larger clusters are more detectable than smaller clusters.

### 6.4.3 The Horvitz–Thompson-Like Estimator for Multiple-Covariate Distance Sampling

#### 6.4.3.1 Estimating Abundance

The probability  $P_a$  now depends on covariates, and so is no longer the same for each animal. We write

$$P_a(\mathbf{z}_i) = E[g(y_i, \mathbf{z}_i)|\mathbf{z}_i] = \int_0^w g(y, \mathbf{z}_i)\pi(y|\mathbf{z}_i) dy \quad (6.38)$$

where  $\mathbf{z}_i$  are the values of the covariates for detection  $i$ . As we have seen, if the design is random, then  $\pi(y|\mathbf{z}_i) \equiv \pi(x|\mathbf{z}_i) = 1/w$  (independent of  $\mathbf{z}_i$ ) for line transect sampling, and  $\pi(y|\mathbf{z}_i) \equiv \pi(r|\mathbf{z}_i) = 2r/w^2$  (again independent of  $\mathbf{z}_i$ ) for point transect sampling. If the design is not random, we will need data that allow us to estimate  $\pi(y|\mathbf{z}_i)$ ; we will meet examples of this more general formulation in Sects. 9.2 and 11.1.

Thus for line transect sampling with random line placement, we have

$$P_a(\mathbf{z}_i) = \frac{1}{w} \int_0^w g(x, \mathbf{z}_i) dx. \quad (6.39)$$

Using Eqs. (5.44) and (5.45), we can thus write

$$P_a(\mathbf{z}_i) = \frac{\mu(\mathbf{z}_i)}{w} = \frac{1}{wf(0|\mathbf{z}_i)}. \quad (6.40)$$

The inclusion probabilities  $p_i$  of the Horvitz–Thompson estimator are thus  $p_i = P_a(\mathbf{z}_i) \cdot a/A$ . We estimate  $f(0|\mathbf{z}_i)$  and hence  $P_a(\mathbf{z}_i)$  using maximum likelihood methods (Sect. 5.3), and hence the Horvitz–Thompson-like estimator of abundance is

$$\hat{N} = \frac{A}{a} \sum_{i=1}^n \frac{1}{\hat{P}_a(\mathbf{z}_i)}. \quad (6.41)$$

For point transect sampling with random point placement,

$$P_a(\mathbf{z}_i) = \frac{2}{w^2} \int_0^w r \cdot g(r, \mathbf{z}_i) dr. \quad (6.42)$$

From Eqs. (5.47) and (5.48), we obtain

$$P_a(\mathbf{z}_i) = \frac{\nu(\mathbf{z}_i)}{\pi w^2} = \frac{2}{w^2 h(0|\mathbf{z}_i)}. \quad (6.43)$$

Abundance estimation for point transect sampling proceeds as for line transect sampling, except that  $\hat{P}_a(\mathbf{z}_i)$  is obtained by substituting the maximum likelihood estimates  $\hat{h}(0, \mathbf{z}_i)$  in Eq. (6.43).

### 6.4.3.2 Estimating Variances and Confidence Intervals

Denoting abundance on the covered strips (line transect sampling) or circles (point transect sampling) by  $N_c$ , we have

$$\hat{N}_c = \sum_{i=1}^n \frac{1}{\hat{P}_a(\mathbf{z}_i)}, \quad (6.44)$$

so that Eq. (6.41) may be written

$$\hat{N} = \frac{A}{a} \hat{N}_c . \quad (6.45)$$

Variance of  $\hat{N}_c$  can be estimated by

$$\text{var}(\hat{N}_c) = S^2(\hat{\boldsymbol{\theta}}) + \hat{\mathbf{d}}' \hat{\mathbf{I}}^{-1} \hat{\mathbf{d}} \quad (6.46)$$

where

$$S^2(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \frac{1 - \hat{P}_a(\mathbf{z}_i)}{\{\hat{P}_a(\mathbf{z}_i)\}^2} , \quad (6.47)$$

$$\hat{\mathbf{d}} = \frac{d\hat{N}_c}{d\boldsymbol{\theta}} |_{\hat{\boldsymbol{\theta}}} , \quad (6.48)$$

and  $-\hat{\mathbf{I}}$  is the matrix of second derivatives of the log likelihood, evaluated at  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ , the vector of parameter estimates. Although expressed differently, the above expression is equivalent to Eqn 3.24 of Marques and Buckland (2004).

The variance of the overall abundance estimate  $\hat{N}$  is estimated as (Buckland et al. 2010)

$$\text{var}(\hat{N}) = \left(\frac{A}{a}\right)^2 \left\{ L \sum_{k=1}^K \frac{l_k \left( \hat{N}_{ck}/l_k - \hat{N}_c/L \right)^2}{K-1} + \hat{\mathbf{d}}' \hat{\mathbf{I}}^{-1} \hat{\mathbf{d}} \right\} , \quad (6.49)$$

where  $\hat{N}_{ck} = \sum_{i=1}^{n_k} \frac{1}{\hat{P}_a(\mathbf{z}_i)}$  is estimated abundance for strip  $k$ , which has half-width  $w$  and length  $l_k$ , where  $\sum_{k=1}^K l_k = L$ .

The log-based method of Sect. 6.2.1 may be used to set confidence intervals.

A simple but computer-intensive alternative is to use the nonparametric bootstrap (see Sects. 6.3.1.2 and 6.3.2.2), in which resamples are generated by sampling with replacement from the sampling units (usually the individual lines for line transect sampling or points for point transect sampling, although they may be clusters of lines or points if the design is based on cluster sampling).

### 6.4.3.3 Estimation for Clustered Populations

As for conventional distance sampling, if animals occur in clusters, Eq. (6.41) gives the estimated number of clusters in the population, and animal abundance is then estimated as

$$\hat{N} = \frac{A}{a} \sum_{i=1}^n \frac{s_i}{\hat{P}_a(\mathbf{z}_i)} \quad (6.50)$$

where  $s_i$  is size of the  $i^{\text{th}}$  detected cluster. Note that  $s_i$  might also be one of the covariates in  $\mathbf{z}_i$ .

Similarly,  $N_c$  is estimated as

$$\hat{N}_c = \sum_{i=1}^n \frac{s_i}{\hat{P}_a(\mathbf{z}_i)}. \quad (6.51)$$

To calculate the variance, we now take

$$s^2(\hat{\theta}) = \sum_{i=1}^n \frac{\left\{1 - \hat{P}_a(\mathbf{z}_i)\right\} s_i^2}{\left\{\hat{P}_a(\mathbf{z}_i)\right\}^2}, \quad (6.52)$$

and the derivative is calculated using the expression of Eq. (6.51) for  $\hat{N}_c$ .

The variance of  $\hat{N}$  is now obtained as in the previous section, except that  $\hat{N}_{ck} = \sum_{i=1}^{n_k} \frac{s_i}{\hat{P}_a(\mathbf{z}_i)}$ .

Note that we can estimate  $N_s$ , the number of clusters in the study region, by

$$\hat{N}_s = \frac{A}{a} \sum_{i=1}^n \frac{1}{\hat{P}_a(\mathbf{z}_i)}. \quad (6.53)$$

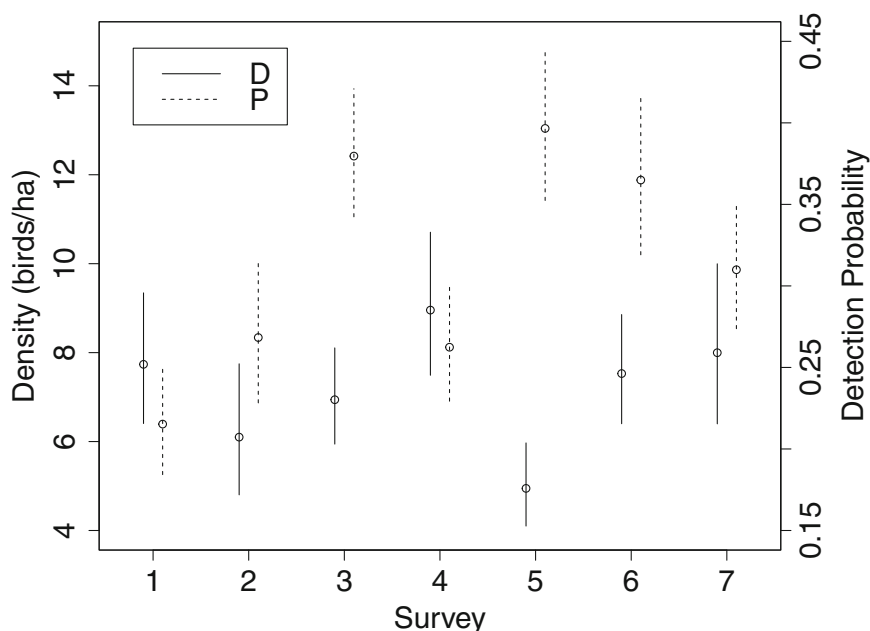
Further, we can estimate  $E(s)$ , the mean cluster size in the population, by taking the ratio of individual abundance to cluster abundance:  $\hat{E}(s) = \hat{N}/\hat{N}_s$ . Unlike the sample mean of detected cluster sizes, this estimate does not suffer from size bias, provided the effect of cluster size on probability of detection has been modelled appropriately. The estimates  $\hat{N}$  and  $\hat{N}_s$  are not independent, complicating estimation of the variance of  $\hat{E}(s)$  if an analytic approach is used, but it is straightforward to estimate this variance using the nonparametric bootstrap in which lines or points are resampled.

#### 6.4.3.4 The Hawaii Amakihi Case Study

To illustrate how to obtain point transect design-based density estimates when additional covariates are available, we continue the analysis presented in Sect. 5.3.2.1. Recall that, besides distance itself, the best model for the detectability of the Hawaii amakihi included observer and minutes from sunrise as covariates (Table 5.7).

We were interested in estimating density for each of the seven survey periods, which can be considered temporal strata. We can therefore estimate a detection probability for each survey period, conditioning on the observed covariates. Intuitively, it seems natural that counts conducted under poor conditions (*e.g.* made by a less effective observer and late in the day) will need to be scaled up more than those conducted under optimal conditions (*e.g.* made by the best observer at sunrise).





**Fig. 6.1** Hawaii amakihi estimated detection probability (P) and density (D) for each of the seven survey periods, using the best model for the detection function which includes observer and minutes since sunrise. The corresponding 95 % confidence intervals are also shown

If observer A detected twice as many birds as observer B, then, all else being equal, survey counts for which only observer B was present should have a scale factor that is double the scale factor for observer A.

The estimated detection probabilities and densities per survey period, conditional on the best model, are given in Fig. 6.1. Detectability was more variable than density. Considering the seven point estimates as representative of an underlying average, the CV of density was 18.4 % while the CV of the detection probability was 21.8 %; the CV on the uncorrected counts was 20.6 %. The intrinsic variability of detection probability is the largest of the three, and accounting for detectability reduces the variability of density compared to the raw counts. This is perhaps unsurprising: part of the variability on the raw counts seems to be due to variation in detectability, rather than fluctuations in density. Therefore, by allowing for variable detectability, we see less variability in the density estimates than in the raw counts.

### 6.4.4 *The Horvitz–Thompson-Like Estimator for Mark-Recapture Distance Sampling*

#### 6.4.4.1 Estimating Abundance

There are several different options for how Horvitz–Thompson-like estimators can be used in the case of MRDS. These are detailed in Laake and Borchers (2004). The appropriate choice depends on several factors, including whether lines or points are randomly located with respect to the animals, and on the field methods adopted by the two observers (Sect. 4.1.2.4). If both observers operate in the same way, then there is symmetry, and either observer can ‘recapture’ an animal first detected by the other. In this case, the summation in the Horvitz–Thompson-like estimator is over all animals detected by at least one of the observers. Often however, one observer searches further ahead, either from a different platform or using high-powered binoculars. This observer then sets up trials, tracking detected animals in, until either the second observer detects it (a ‘recapture’) or it passes the platform undetected by the second observer. In this case, summation in the Horvitz–Thompson-like estimator is only over those animals detected by the second observer. The trials set up by the first observer are used only for estimating the probability of detection for those animals recorded by the second observer. Although this method fails to utilise all the information in the data, it can give a more robust estimate of abundance, for example if there is responsive movement before the second observer detects animals.

If  $n$  is the number of animals to be included in the sum (which corresponds to the total number of animals detected by at least one observer for the first approach described above, and to the number of detections made by the second observer only in the second approach), then a Horvitz–Thompson-like estimator of the number of animals  $N_c$  in the covered region is simply

$$\hat{N}_c = \sum_{i=1}^n \frac{1}{\hat{E}[p.(z_i)]} \quad (6.54)$$

where  $E[p.(z_i)] = \int_0^w p.(y, z_i) \pi(y) dy$  is estimated by maximizing the product of selected components of the likelihood (Sect. 5.4.1) to give  $\hat{E}[p.(z_i)]$ . The function  $\pi(y)$  represents the distribution of available animals as a function of distance from the line or point, and is usually taken to be uniform ( $\pi(y) = 1/w$ ) for line transect sampling and triangular ( $\pi(y) = 2y/w^2$ ) for point transect sampling.

Having obtained  $\hat{N}_c$ , then we estimate abundance in the entire region in the usual way:

$$\hat{N} = \frac{A}{a} \hat{N}_c \quad (6.55)$$

#### 6.4.4.2 Estimating Variances and Confidence Intervals

We proceed much as for MCDS, except we now write the equivalent expression to  $P_a(\mathbf{z}_i)$  by  $E[p.(\mathbf{z}_i)]$ . Thus we can estimate the variance of  $\hat{N}_c$  by (Buckland et al. 2010)

$$\text{var}(\hat{N}_c) = S^2(\hat{\boldsymbol{\theta}}) + \hat{\mathbf{d}}' \hat{\mathbf{I}}^{-1} \hat{\mathbf{d}} \quad (6.56)$$

where

$$S^2(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \frac{1 - \hat{E}[p.(\mathbf{z}_i)]}{\{\hat{E}[p.(\mathbf{z}_i)]\}^2}, \quad (6.57)$$

$$\hat{\mathbf{d}} = \frac{d\hat{N}_c}{d\boldsymbol{\theta}}|_{\hat{\boldsymbol{\theta}}}, \quad (6.58)$$

and  $-\hat{\mathbf{I}}$  is the matrix of second derivatives of the log likelihood, evaluated at  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ .

The variance of the overall abundance estimate  $\hat{N}$  is estimated as (Buckland et al. 2010)

$$\text{var}(\hat{N}) = \left(\frac{A}{a}\right)^2 \left\{ L \sum_{k=1}^K \frac{l_k \left( \hat{N}_{ck}/l_k - \hat{N}_c/L \right)^2}{K-1} + \hat{\mathbf{d}}' \hat{\mathbf{I}}^{-1} \hat{\mathbf{d}} \right\}, \quad (6.59)$$

where  $\hat{N}_{ck} = \sum_{i=1}^{n_k} \frac{1}{\hat{E}[p.(\mathbf{z}_i)]}$  is estimated abundance for strip  $k$ , which has half-width  $w$  and length  $l_k$ , where  $\sum_{k=1}^K l_k = L$ .

Again, the log-based method of Sect. 6.2.1 may be used to set confidence intervals.

Alternatively, we may estimate variances and confidence intervals by using the nonparametric bootstrap as described in Sect. 6.4.3.2.

#### 6.4.4.3 Estimation for Clustered Populations

For clustered populations, with the  $i^{\text{th}}$  cluster having size  $s_i$ , we have

$$\hat{N}_c = \sum_{i=1}^n \frac{s_i}{\hat{E}[p.(\mathbf{z}_i)]}. \quad (6.60)$$

To calculate the variance, we now take

$$S^2(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \frac{\left\{ 1 - \hat{E}[p.(\mathbf{z}_i)] \right\} s_i^2}{\left\{ \hat{E}[p.(\mathbf{z}_i)] \right\}^2}, \quad (6.61)$$

and the derivative is calculated using the expression of Eq. (6.60) for  $\hat{N}_c$ .

The variance of  $\hat{N}$  is now obtained as in the previous section, except that  $\hat{N}_{ck} = \sum_{i=1}^{n_k} \frac{s_i}{E[p_{\cdot}(\mathbf{z}_i)]}$ .

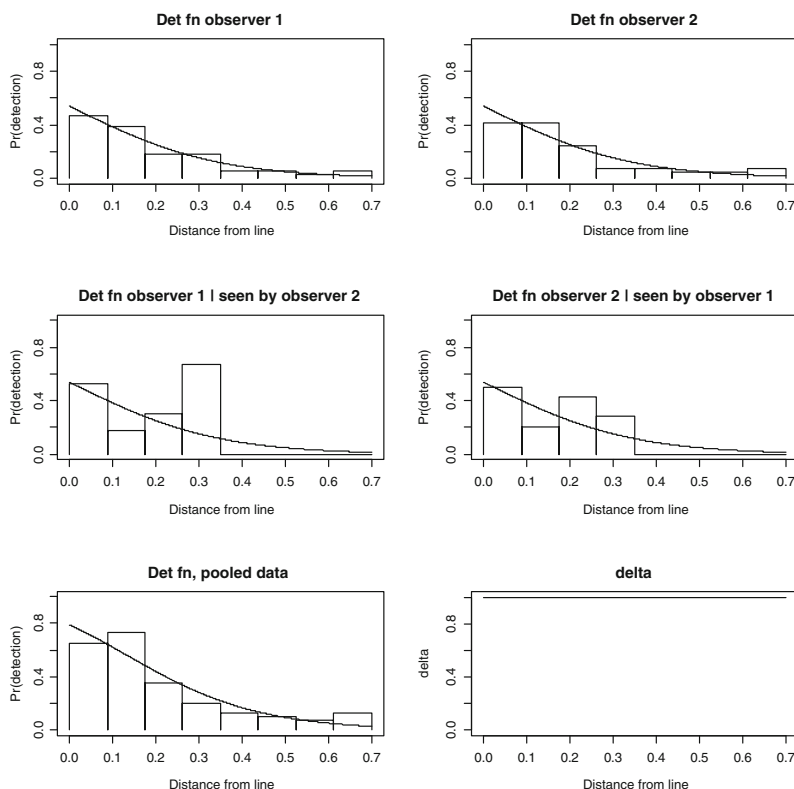
#### 6.4.4.4 The Minke Whale Case Study

In this case study, the two observers were not searching the same area of sea at the same time, in an effort to reduce or even eliminate the dependence in the detection probabilities for the two observers. The benefits of this strategy are apparent from Table 6.5. AIC favours models with full independence ( $\alpha = \beta = 0$ ), and selects the model with identical detection functions for the two observers, and sea state as a covariate. Estimation is largely unaffected by whether we assume full independence or point independence. If we also relax the assumption of point independence, AIC values are larger, but estimation is not greatly affected, with the exception of model 2. High correlation between parameter estimates prevents estimation of abundance under some of the limiting independence models, but we are still able to calculate AIC for those models.

**Table 6.5** Models fitted to the minke whale survey data

Model	Dependence	Form of model	No. parameters	$\hat{N}$	$se(\hat{N})$	$\Delta\text{AIC}$
1	LI	<i>observer</i> ×( <i>dist</i> + <i>beau</i> )	8	16,912	6,663	7.4
2	LI ( $\beta = 0$ )	<i>observer</i> ×( <i>dist</i> + <i>beau</i> )	7	26,453	18,422	7.0
3	PI	<i>observer</i> ×( <i>dist</i> + <i>beau</i> )	7	17,156	5,359	5.4
4	FI	<i>observer</i> ×( <i>dist</i> + <i>beau</i> )	6	17,935	5,436	3.6
5	LI	<i>observer</i> × <i>dist</i>	6	–	–	13.6
6	LI ( $\beta = 0$ )	<i>observer</i> × <i>dist</i>	5	–	–	11.7
7	PI	<i>observer</i> × <i>dist</i>	5	16,192	4,446	12.3
8	FI	<i>observer</i> × <i>dist</i>	4	15,195	3,311	10.4
9	LI	<i>dist</i> + <i>beau</i>	5	16,837	7,130	3.9
10	LI ( $\beta = 0$ )	<i>dist</i> + <i>beau</i>	4	16,091	4,365	1.9
11	PI	<i>dist</i> + <i>beau</i>	4	17,337	5,647	1.9
12	FI	<i>dist</i> + <i>beau</i>	3	18,173	5,534	0.0
13	LI	<i>dist</i>	4	–	–	9.8
14	LI ( $\beta = 0$ )	<i>dist</i>	3	–	–	7.9
15	PI	<i>dist</i>	3	16,209	4,447	8.5
16	FI	<i>dist</i>	2	15,202	3,313	6.6

The full model, denoted here by LI (Limiting Independence), *observer*×(*dist*+*beau*) (indicating different detection functions for the two observers, each of which is a function of both distance and sea state), is defined by Eqs. (5.60) and (5.61). The estimates  $\hat{N}$  and  $se(\hat{N})$  are obtained as described in Sect. 6.4.4, although in both cases, the method was extended to accommodate the stratified design used. PI indicates Point Independence ( $\alpha = 0$ ) and FI denotes Full Independence ( $\alpha = \beta = 0$ ). ‘–’ indicates that abundance estimation was not possible due to very high correlation between estimated parameters



**Fig. 6.2** Estimated detection functions for minke whales, model 12 (see Table 6.5). The *top left* plot is the estimated unconditional detection function for observer 1, and *top right* is the estimated unconditional detection function for observer 2. The corresponding conditional detection functions are shown in the centre. Under this model, all four of these detection functions are identical, but the data in each plot differ. The estimated detection function for the two observers combined is shown at the *bottom left*. The bars are: relative frequencies of detections made by observer 1 (*top left*), relative frequencies of detections made by observer 2 (*top right*), proportion of observer 2 detections made by observer 1 (*middle left*), proportion of observer 1 detections made by observer 2 (*middle right*), and relative frequencies of detections made by at least one observer. Reproduced from Buckland et al. (2010). ©The International Biometric Society

AIC favours model 12, and corresponding fits are shown in Fig. 6.2. It is surprising that AIC favours models which assume the same detection function for the two observers, given that the tracker is searching much further ahead of the ship than the primary platform. However, estimation is barely affected by whether we make this assumption or not. The distribution of distances from the line of detections from the two platforms is clearly very similar out to the truncation distance of 700 m (Fig. 6.2), although beyond this distance, the tracker detects more animals than the primary platform.

# Chapter 7

## Model-Based Distance Sampling: Two-Stage Models

### 7.1 Introduction

In Chap. 5, we considered modelling frameworks for estimating the detection function. Having fitted such a model, we can correct counts of animals on the surveyed plots for detectability. A natural way to analyse the data therefore is to develop a count model, assuming perhaps a Poisson, quasi-Poisson or negative binomial model, while including estimated probability of detection as an *offset*. This allows abundance, or equivalently density, on the surveyed plots to be predicted. We term this approach ‘two-stage’ because detectability is estimated in the first stage, and then the counts are modelled in the second stage.

Having predicted abundance on the surveyed plots, we can extrapolate to the entire study area to estimate total abundance. This can be done in a design-based way, relying on the design to ensure that the plots are representative, or, if the covariates in the count model are recorded throughout the study area, and not just on the surveyed plots, we can predict abundance on a grid, and sum the predictions to estimate total abundance.

Sometimes, we are not interested in estimating total abundance, but instead wish to test whether densities differ among sets of plots. For example, we may wish to test the effectiveness of a management regime, and we design a survey so that some plots are treatment plots, to which the regime is applied, and some are control plots (Sect. 3.2). In this case, there is no need to extrapolate to a wider study area, but we do need to include a treatment effect in our count model.

Disadvantages of the above approach are as follows. First, the count model treats the estimated offset as if it is a known constant. Solutions to this problem include using a bootstrap to propagate uncertainty from the first stage to the second (Sect. 7.2.2.1), maximizing the full, integrated likelihood in a single step (Sect. 8.3), or adopting a Bayesian approach, again using the full likelihood (Sect. 8.4). Second, the data are analysed as counts. This means that an approximation is needed to accommodate covariates that vary by individual. Further, in the case of line transect

sampling, covariates may vary appreciably over the length of the line, so that the line must be arbitrarily divided into segments, and the segment counts (which will typically not be independent) modelled.

## 7.2 Plot Count Models

One approach to model-based distance sampling is to model the counts  $n_k$  of animals recorded on each plot, where plot  $k$  ( $k = 1, \dots, K$ ) is a strip of half-width  $w$  and length  $l_k$  in the case of line transect sampling, and a circle of radius  $w$  for point transect sampling. We denote the true but unknown abundance on plot  $k$  by  $N_k$ , with corresponding expectation  $E(N_k)$ , and plot area by  $a_k$ , with  $a_k = 2wl_k$  for line transect sampling and  $a_k = \pi w^2$  for point transect sampling. We define  $D_k = N_k/a_k$ , and if we take expectations over plot abundance,  $E(D_k) = E(N_k)/a_k$ . If we denote the probability that an animal on plot  $k$  is detected (unconditional on its distance from the line or point) by  $P_k$ , then we can write

$$\lambda_k = E(N_k) \times P_k = E(D_k) \times a_k \times P_k \quad (7.1)$$

where  $\lambda_k = E(n_k)$  is the expected number of animals detected on plot  $k$ .

For two-stage count models,  $P_k$  is estimated in stage one, and conditional on those estimates, the counts are modelled in stage two.

### 7.2.1 Stage One: Estimating Probability of Detection

If we do not have covariates  $\mathbf{z}$ , then we may use the methods of Sect. 5.2 to estimate  $P_a$ , the probability of detection of an animal that is on a plot. In this case,  $P_k$  of Eq. (7.1) is equal to  $P_a$  for all  $k$ , and we estimate it by  $\hat{P}_a$ , given in Eq. (5.24) for line transect sampling and in Eq. (5.36) for point transect sampling.

If we have covariates, then the probability of detection depends on those covariates. For the  $i^{\text{th}}$  animal detected, we denote this probability by  $P_a(\mathbf{z}_i)$ . If the covariates  $\mathbf{z}$  are recorded at the plot level or, for a stratified design, at the stratum level, then all detections on a given plot have the same values of  $\mathbf{z}$ , so that each plot has a single value of  $P_k$ , estimated from the fitted MCDS model by substituting parameter estimates into Eq. (6.40) (line transect sampling) or Eq. (6.43) (point transect sampling).

If at least one of the covariates in  $\mathbf{z}$  is recorded at the level of individual detections, then in general it varies among individuals on a given plot. We cannot simply average over the animals detected, because the more detectable animals are over-represented (an example of size-biased sampling). However, we can exploit the Horvitz–Thompson estimator. Plot abundance  $N_k = D_k a_k$ , and if the  $P_a(\mathbf{z}_i)$  were known, we could write

$$\hat{N}_k = \sum_{i=1}^{n_k} \frac{1}{P_a(\mathbf{z}_i)} \quad (7.2)$$

where  $n_k$  is the number of animals detected on plot  $k$ . In practice we estimate these probabilities using our fitted MCDS model, and so obtain a Horvitz–Thompson-like estimator (Sect. 6.4.3):

$$\hat{N}_k = \sum_{i=1}^{n_k} \frac{1}{\hat{P}_a(\mathbf{z}_i)} . \quad (7.3)$$

If  $n_k = 0$ , then  $\hat{N}_k = 0$ ; in this case, we might take  $\hat{P}_k$  to be the mean value for plots with at least one detection, or the mean value for neighbouring plots with at least one detection.

We can now estimate  $P_k$  in Eq. (7.1) by  $\hat{P}_k = n_k / \hat{N}_k$ . Note that  $1/\hat{P}_k$  is the sample mean of  $1/\hat{P}_a(\mathbf{z}_i)$  for the plot. Buckland et al. (2009) used this approach (although formulated a little differently).

If animals occur in clusters, with  $s_i$  animals in the  $i^{\text{th}}$  detected cluster, then we have

$$\hat{N}_k = \sum_{i=1}^{n_k} \frac{s_i}{\hat{P}_a(\mathbf{z}_i)} . \quad (7.4)$$

In this case, Eq. (7.3) gives an estimate of the number of clusters in the population.

## 7.2.2 Stage Two: Modelling the Counts

In some studies, we are interested only in comparing densities on the plots, for example to assess whether densities on treatment plots differ from those on control plots. In this case, we can use generalized linear or additive models, perhaps including random effects, to model the counts. In other studies, we are interested in modelling density across the study region, and perhaps also in estimating total abundance in the region. We then need a spatial model of some form for the counts. We consider the two cases separately below.

### 7.2.2.1 Designed Experiments

Conditional on the  $\hat{P}_k$ , we can model the counts  $n_k$  as Poisson or negative binomial, with a log link function and an offset of  $\log_e(a_k \hat{P}_k)$ :

$$E(n_k) = \lambda_k = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q + \log_e(a_k \hat{P}_k) \right) . \quad (7.5)$$



From Eq. (7.1), we have  $\lambda_k = E(D_k) \times a_k P_k$ . Replacing  $P_k$  by its estimate, we obtain

$$\lambda_k = E(D_k) \times a_k \hat{P}_k = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q + \log_e(a_k \hat{P}_k) \right) \quad (7.6)$$

so that our model is equivalent to modelling density, as the effective area surveyed per plot,  $a_k \hat{P}_k$ , cancels:

$$E(D_k) = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q \right). \quad (7.7)$$

The design of our experiment determines the  $x_{qk}$ . For example, if we had a randomized blocks design with three blocks and two treatments, we might take  $x_{1k} = 1$  for all  $k$ ;  $x_{2k} = 1$  if plot  $k$  is in block 2, and zero otherwise;  $x_{3k} = 1$  if plot  $k$  is in block 3 and zero otherwise (so that  $x_{2k} = x_{3k} = 0$  indicates that plot  $k$  is in block 1); and  $x_{4k} = 1$  if plot  $k$  has treatment 2, and zero otherwise. In this case,  $\beta_4$  corresponds to the effect of treatment 2 relative to treatment 1; if the two treatments are equivalent with respect to animal density, then  $\beta_4 = 0$ . Hence we are primarily interested in the estimate of  $\beta_4$ , and whether it differs significantly from zero.

We might also wish to include additional covariates, not determined by the design, in the  $x_{qk}$ . In the above example, we might have  $x_{5k}$  defined to be a measure of habitat suitability for example, which might be a continuous variable or correspond to a single level of a factor. If a continuous variable, it might enter the model as a linear term or as a smooth in the exponent (*i.e.* a generalized additive model). Random effects (Sect. 8.2.5.1) might also be included in the exponent. It is possible for the same covariate to be one of the  $x_q$  covariates in Eq. (7.5) and one of the covariates in  $\mathbf{z}$  used to model probability of detection.

If we assume a Poisson model, then the likelihood corresponding to the plot counts  $n_k$  is given by

$$\mathcal{L}_{\{n_k\}} = \prod_{k=1}^K \frac{\lambda_k^{n_k} \exp[-\lambda_k]}{n_k!} \quad (7.8)$$

with expectation  $\lambda_k$  given by Eq. (7.5). We use the notation  $\mathcal{L}_{\{n_k\}}$  to indicate that this is the product of count likelihoods, corresponding to the set of plot counts  $n_k$ . A quasi-Poisson model could be used if the counts are overdispersed or underdispersed. For overdispersed data, another option is a negative binomial model.

The formulation of Eq. (7.5) is for when each plot count  $n_k$  corresponds to a single visit to one line or point on the plot. If there are repeat visits and counts are summed before analysis, then the effort  $e_k$  (number of visits to plot  $k$ ) should be included in the offset:  $\log_e(a_k e_k \hat{P}_k)$ . If there are multiple lines or points on each

plot, and counts are summed across them before analysis, then  $a_k$  is the total area of sampled strips or circles on the plot.

An advantage of this two-stage approach is that we can use standard generalized linear (or additive) modelling software to fit the count model. If we also have random effects, we can use standard generalized linear (or additive) mixed modelling software.

If we estimate precision using our fitted count model, then we have ignored a source of variance, as our model treats the  $\hat{P}_k$  as known constants. Buckland et al. (2009) used bootstrap methods to propagate uncertainty from stage one through to stage two. For designs with an adequate number of replicate lines or points on each plot (ideally ten or more), we can simply use the nonparametric bootstrap to resample with replacement points on each plot. The resulting resamples are then analysed exactly as for the real data. If a point is selected for a resample, the data from all visits to that point are included; this allows for any dependence in counts from repeat visits. This method also accommodates the uncertainty in estimating the offset of the count model.

If there is insufficient replication of lines or points on each plot for the above method to be viable, then the bootstrap may be implemented in two stages. In the first, bootstrap resamples of distances  $y$  are generated. If there are no covariates  $\mathbf{z}$ , this is most simply done by nonparametric resampling of the  $y$ 's. If there are covariates  $\mathbf{z}$ , we can either use the naïve bootstrap, generating nonparametric resamples by resampling the  $y$ 's along with their associated  $\mathbf{z}$ 's, or generate a parametric resample of  $n$   $y$ 's from the estimated density  $\hat{f}_{y|\mathbf{z}}(y|\mathbf{z})$ . For each of these resamples, we then generate a parametric bootstrap resample from the fitted count model. If our count model is a quasi-Poisson model, then we need to generate resamples with preserving any overdispersion; one means of doing this is the probability integral transform resampling method (Bravington 1994; Borchers et al. 1997).

Williams et al. (2011) describe a less computer-intensive method of propagating uncertainty in the offset. Alternatively, the full likelihood methods of Chap. 8 might be adopted, to avoid this difficulty.

### 7.2.2.2 Spatial Two-Stage Count Models

We again model the plot counts. However, in the case of line transect sampling, the plots may be sufficiently long that spatial covariates, and animal density, vary appreciably within the plots. Hence it may be necessary to split the plots into short segments (Hedley and Buckland 2004).

We can again fit a model of the form of Eq. (7.5), but typically the variables  $x_{gk}$  will not represent the design. Instead, they will be spatial covariates (continuous and/or factors). Usually, we would take  $x_{1k} = 1$  for all  $k$ , so that  $\beta_1$  is the intercept term. For continuous covariates, we might replace linear terms by smooth terms, giving a generalized additive model. The analysis should allow for spatial autocorrelation.

### 7.3 Plot Abundance Methods

Equation (7.3) (or Eq. (7.4) for clustered populations) gives us plot abundance estimates. We could treat these as our responses, and fit a model of the form

$$E(\hat{N}_k) = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q + \log_e(a_k) \right) . \quad (7.9)$$

We should assume an error distribution that restricts the range of  $\hat{N}_k$  to be non-negative, such as the log-normal, gamma or quasi-Poisson distributions. Compare Eq. (7.9) with Eq. (7.5). An advantage of this approach is that the offset term  $\log_e(a_k)$  is now a known constant. The disadvantage is that we are now modelling estimates derived from the observations rather than the observations themselves. If we observed the plot abundances  $N_k$ , then Poisson or negative binomial models would be natural choices, but as they are estimated, we instead assume an arbitrary distribution for them, for which the variance-mean relationship may be inappropriate. This disadvantage is less compelling when we have covariates  $\mathbf{z}$  that vary by individual, because there is then less justification for assuming that plot counts have a Poisson or negative binomial distribution.

We may replace linear terms in Eq. (7.9) by smooths, giving a generalized additive model, and random effects can be added, to give a generalized linear (or additive) mixed model.

Hedley and Buckland (2004) compared three spatial methods for analysing line transect data on southern minke whales (*Balaenoptera bonaerensis*). One of those was a plot abundance model, in which estimated number of minke pods per segment of line was assumed to follow a quasi-Poisson model. They found that the method performed well when compared with more advanced methods.

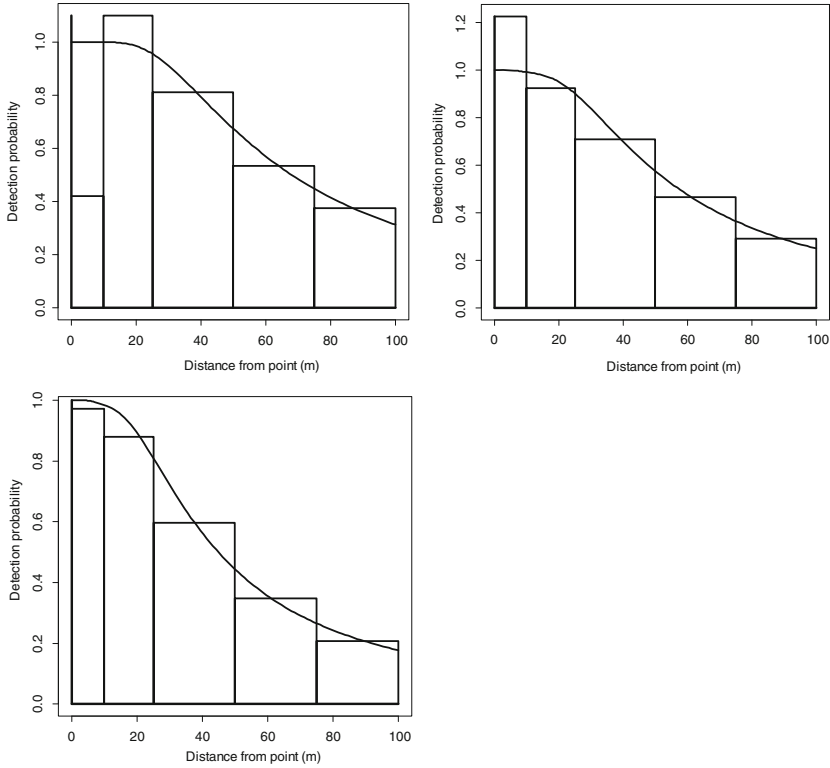
## 7.4 Examples

### 7.4.1 The BACI Case Study

In this case study, we are interested in whether prescribed burning of ponderosa pine (*Pinus ponderosa*) forest reduces breeding densities of Grace's warbler (*Dendroica graciae*) and the yellow-rumped warbler (*Dendroica coronata*). Our analyses are taken from Buckland et al. (2009). As both species are expected to show a similar response to burning, we pool the data from the two species to increase power.

#### 7.4.1.1 Modelling the Detection Function

A truncation distance of  $w = 100$  m was found to be adequate, leaving  $n = 2300$  detections. The multiple covariate distance sampling engine of Distance was



**Fig. 7.1** Estimated detection function by year for Grace's and yellow-rumped warblers, pooling over values of continuous covariates tree and snag density. Years 2003 (*top left*), 2004 (*top right*) and 2005 (*bottom left*). Note that the histogram bars represent number of detections divided by distance from the point

used to fit a number of models. Potential continuous covariates were shrub stem density, large tree density, snag density, and a burn severity index, while factor covariates considered for inclusion were year at three levels (2003, 2004, 2005) and species at two levels (Grace's warbler, yellow-rumped warbler). A stepwise procedure was adopted for inclusion of covariates; the covariate giving rise to the greatest improvement in AIC was selected at each step, until no additional covariate reduced the AIC score (Table 7.1). The fitted detection function (Fig. 7.1) corresponding to the model selected by AIC was  $\hat{g}(y) = 1 - \exp \left[ - (y/\hat{\sigma})^{-1.63} \right]$  where  $\hat{\sigma} = 56.2 \exp [-0.0024t_i - 0.0100u_i + 0.37c_i + 0.22d_i]$ , and  $t_i$  is tree density,  $u_i$  is snag density,  $c_i$  is 1 if observation  $i$  was for year 2003 and 0 otherwise, and  $d_i$  is 1 if observation  $i$  was for year 2004 and 0 otherwise. Median tree density was  $140 \text{ km}^{-2}$ , and median snag density  $2.5 \text{ km}^{-2}$ . Note that the covariate species was not selected; the two warbler species gave nearly identical estimated detection functions. Estimated detection probabilities were lower at points with high tree density, and slightly lower at points with high snag density.

**Table 7.1** AIC scores for selected models for the detection function for Grace's warbler and yellow-rumped warbler

Key function	Continuous covariates	Factor covariates	$\Delta$ AIC
Half-normal	None	None	25.26
Half-normal	None	Year	20.63
Half-normal	Tree	Year	16.94
Half-normal	Tree, Snag	Year	16.89
Hazard-rate	None	None	16.57
Hazard-rate	Tree	None	10.14
Hazard-rate	Tree, Snag	None	4.69
Hazard-rate	Tree, Snag	Year	0.00

#### 7.4.1.2 Modelling the Plot Counts

The `glm` command of the statistical package R was used to fit a Poisson model to the warbler plot counts (Table 7.2). The counts analysed were the sum of counts across the two species, the points on the plot, and the visits made within a year, giving rise to 24 counts (four sites by two plots per site by 3 years). To estimate the offset, the hazard-rate detection function model with year as a factor and tree and snag density as continuous covariates was used. The number of visits to each point (*i.e.* the effort) was also included in the offset. AIC selected the count model with all two-factor interactions but no three-factor interaction (Table 7.2). We also tried adding covariates (averaged across the points within the plot for each year) — tree density, snag density and burn index — to this model, but none led to a reduction in the AIC score, and so they were not included. (Thus covariates were included in the model for detectability, and hence in the offset of the count model, but were not included in the linear predictor of the count model.) When a model with the same structure but with a negative binomial error distribution was fitted (using the `glm.nb` command from the MASS package in R), the limiting form corresponding to Poisson regression was obtained, so results were unaffected. Similarly, as the residual deviance under this model (6.04, Table 7.2) is almost equal to the degrees of freedom (6), inference is unaffected if a quasi-Poisson regression is conducted instead of the Poisson regression.

The above analysis makes no allowance for the estimation of the offset, so that both the analytic standard errors and AIC may be unreliable guides to model selection. We therefore calculated bootstrap standard errors and 95 and 99 % percentile confidence intervals for parameters based on 999 bootstrap resamples. There was adequate replication in the design (25–50 points on each plot) to adopt the simple nonparametric bootstrap, in which points were resampled within plots to generate resamples, and each resample was analysed using the same methods as for the real data. The bootstrap analyses allow for any dependence between counts at the same point across years, and for estimation of the offset, and we see that the standard errors are larger on average than the analytic ones (Table 7.3).

**Table 7.2** Analysis of deviance of warbler counts, treating the values of the offset as known constants, Poisson error model

	df	Deviance	Resid. df	Resid. dev.	AIC
<i>Null model</i>			23	1211.69	
+ <i>site</i>	3	923.34	20	288.35	
+ <i>year</i>	2	69.13	18	219.23	
+ <i>treatment</i>	1	11.33	17	207.89	
+ <i>site</i> × <i>year</i>	6	81.64	11	126.25	
+ <i>site</i> × <i>treatment</i>	3	109.51	8	16.74	191.1
+ <i>year</i> × <i>treatment</i>	2	10.70	6	6.04	184.4
+ <i>site</i> × <i>year</i> × <i>treatment</i>	6	6.04	0	0.00	190.4

The factor *treatment* refers to whether a plot is a prescribed fire treatment plot or a control plot

If prescribed burning affected bird density, we would hope to detect a treatment by year interaction, because the treatment covariate distinguished between treatment and control plots, but did not reflect whether the treatment had yet been applied. Just two of the bootstrap standard errors of Table 7.3 are (slightly) smaller than the analytic ones, but these both correspond to the treatment by year interaction. Inference is therefore unaffected in this case; we have strong evidence of reduced bird densities on burned plots in the year following burning (the *year2*×*treatment2* term of Table 7.3), and moderate evidence that the effect continues into a second year (the *year3*×*treatment2* term of Table 7.3). Our results provide evidence that the among-year differences were not consistent across sites, and that an interaction between type of plot (treatment or control) and site existed. Although the analytic results gave strong evidence of a difference between treatment and control plots (see the *treatment2* term in Table 7.3), when uncertainty arising from estimation of the offset and possible dependence in repeat counts at a site across years were taken into account, the 95 % bootstrap confidence interval for the difference between treatment and control plots (just) included zero.

### 7.4.2 Density Surface Modelling: Spotted Dolphins in the Gulf of Mexico

This analysis stems from efforts to describe the spatial distribution of pantropical spotted dolphins (*Stenella attenuata*) in the Gulf of Mexico. Data were derived from National Marine Fisheries Service survey cruises conducted during 1992–2001. Full details of the analysis can be found in Miller et al. (2013) and this dataset is distributed with the GUI version of Distance.

Survey transect effort was divided into 387 segments, each of which was roughly 20 km in length. In total, 47 dolphin schools were detected. Model-based inference is used to (a) scale up number of dolphins recorded to an estimate of number of

**Table 7.3** Poisson model: estimated coefficients and standard errors

Term	Estimate	Analytic std. error	Bootstrap std. error
<i>Intercept</i>	−11.999	0.212**	0.317**
<i>site2</i>	1.485	0.223**	0.310**
<i>site3</i>	0.607	0.257*	0.345*
<i>site4</i>	1.176	0.230**	0.336**
<i>year2</i>	0.729	0.248**	0.311*
<i>year3</i>	0.947	0.236**	0.273**
<i>treatment2</i>	−0.523	0.199**	0.264
<i>site2</i> × <i>year2</i>	0.004	0.259	0.305
<i>site3</i> × <i>year2</i>	−0.591	0.290*	0.329
<i>site4</i> × <i>year2</i>	−0.041	0.268	0.323
<i>site2</i> × <i>year3</i>	−0.209	0.246	0.264
<i>site3</i> × <i>year3</i>	−1.680	0.297**	0.340**
<i>site4</i> × <i>year3</i>	−0.637	0.258*	0.292*
<i>site2</i> × <i>treatment2</i>	1.105	0.188**	0.268**
<i>site3</i> × <i>treatment2</i>	1.618	0.232**	0.314**
<i>site4</i> × <i>treatment2</i>	0.387	0.199	0.289
<i>year2</i> × <i>treatment2</i>	−0.378	0.116**	0.111**
<i>year3</i> × <i>treatment2</i>	−0.242	0.116*	0.111*

Linear predictor: *intercept* + *site* + *year* + *treatment* + *site*×*year* + *site*×*nt* + *year*×*nt*. Levels of site are Apache-Sitgreaves (*site1*), Coconino (*site2*), Gila (*site3*) and Kaibab (*site4*); levels of year are 2003 (*year1*), 2004 (*year2*) and 2005 (*year3*); levels of treatment are control (*treatment1*) and burn (*treatment2*)

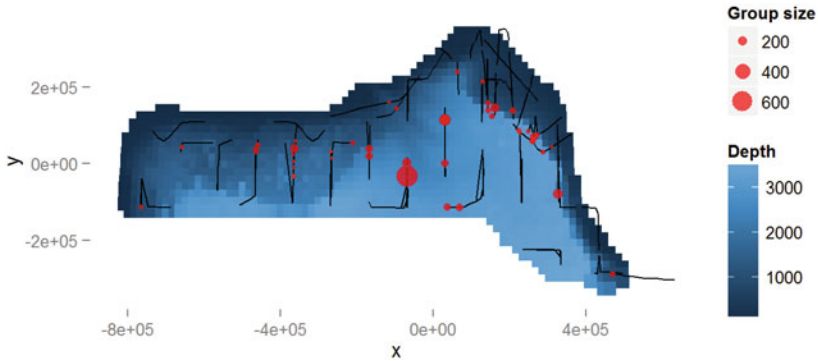
\*Estimate differs significantly from zero (5 % level) using z-test (analytic standard error) or percentile CI (bootstrap)

\*\*Estimate differs significantly from zero (1 % level) using z-test (analytic standard error) or percentile CI (bootstrap)

dolphins in each survey segment and (b) derive a relationship between estimated dolphin numbers per segment and environmental covariates to predict number and distribution of dolphins within the study area. The study area, transect lines (which were surveyed repeatedly over the period of the study), observed school sizes and bottom depth within the study area are shown in Fig. 7.2.

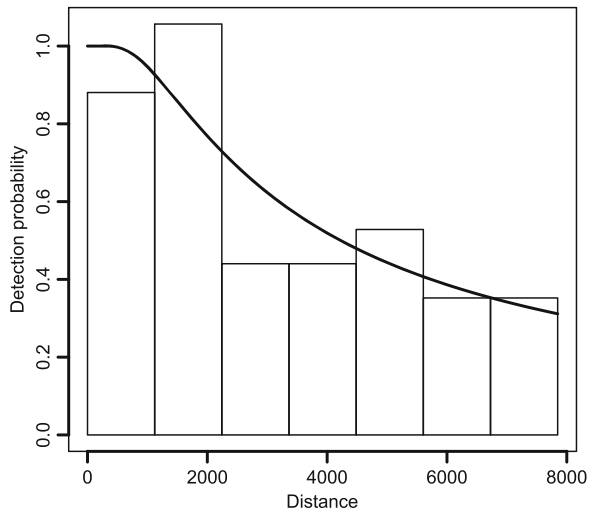
**7.4.2.1 Modelling the Detection Function**

In practice, we would select from a set of competing detection function models (Sect. 5.1), using for example AIC. However we omit details of the detection function modelling, present the fitted hazard-rate model (Fig. 7.3), and move on to the fitting of the density surface model.



**Fig. 7.2** Study area within Gulf of Mexico for pantropical spotted dolphin surveys. Plot of depth within the study area with transects and detections overlaid. Point size is proportional to group size. Map axes show  $x$  and  $y$  coordinates in metres, scaled such that  $(0, 0)$  is roughly in the centre of the study area

**Fig. 7.3** Histogram of detection distances for spotted dolphin schools. Also shown is the fitted hazard-rate detection function



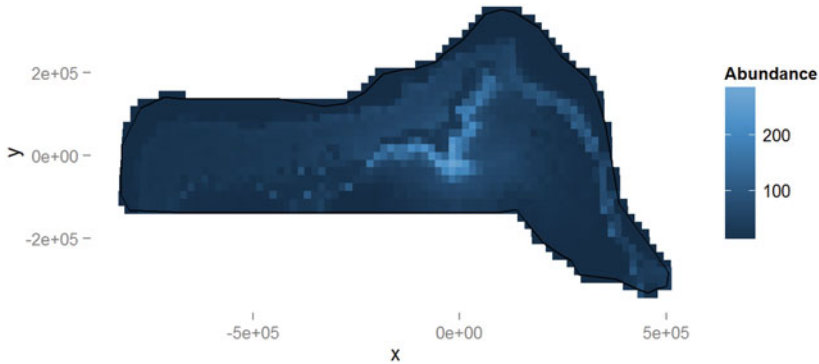
### 7.4.2.2 Building the Density Surface Model

We adopt a plot abundance model (Sect. 7.3), in which we first estimate the abundance for each segment, using the Horvitz–Thompson-like estimator of Eq. (7.4).

A generalized additive model was fitted to the resulting segment abundance estimates, using a log link function and a quasi-Poisson error structure. An offset, equal to log segment area, was also included as in Eq. (7.4) to account for variable segment size.

A model of the form of Eq. (7.9) was used with covariates of ocean depth and spatial location  $(x, y)$  and a quasi-Poisson error structure. Because the bivariate smooth is isotropic (equivalent in both  $x$  and  $y$  directions), latitude and longitude was





**Fig. 7.4** Predicted density surface across study area with depth and bivariate smooth for  $x$  and  $y$  location as predictors. Map axes show  $x$  and  $y$  in metres, scaled such that  $(0, 0)$  is roughly in the centre of the study area. Scale of density shown in legend is in number of dolphins within prediction grid cells of size  $444 \text{ km}^2$

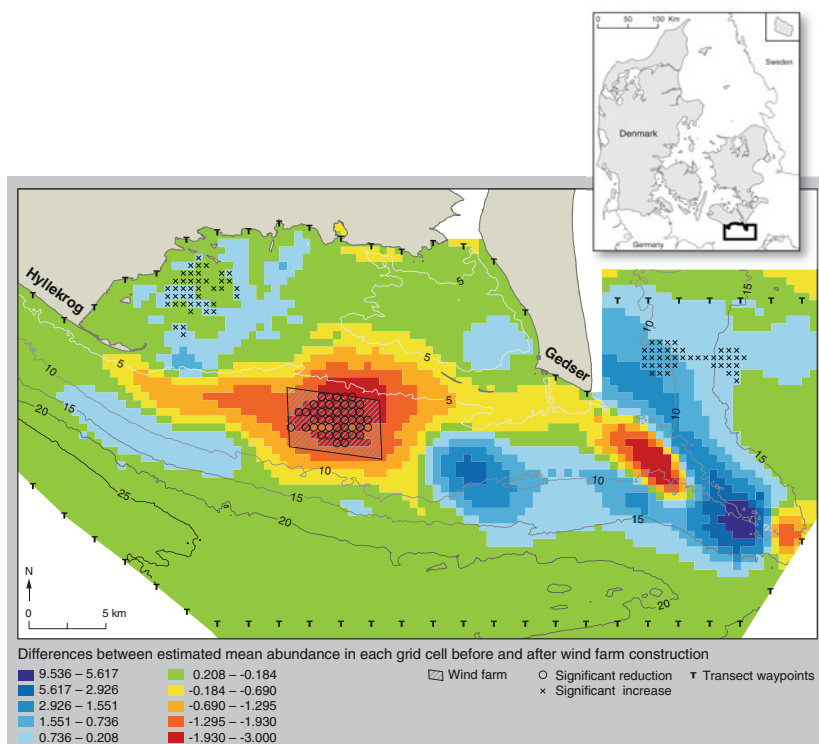
transformed to Universal Transverse Mercator (UTM) coordinates and standardized such that the centre of the study area had coordinate  $(0,0)$ . Predictions from a model using depth and a bivariate smooth of  $x$ - and  $y$ -coordinates are shown in Fig. 7.4.

### 7.4.3 The Nysted Windfarm Example

As noted in Sect. 3.3, pre- and post-impact line transect surveys were conducted over an area spanning the footprint of the Nysted windfarm and surrounding waters. The numbers of detected long-tailed ducks were recorded in three distance intervals, with the line offset by 44 m to allow for observers' inability to see below the aircraft. Total abundance in the study area was estimated using conventional distance sampling methods. The fine-scale abundance patterns over time were modelled by combining an MCDS detection function model with a spatially-adaptive generalized additive model using generalized estimating equations (Hardin and Hilbe 2003). This approach allows assessment both of whether abundance has changed post-impact and whether birds have redistributed themselves spatially.

For the spatio-temporal analysis, lines were divided into segments each 0.5 km long. MCDS methods were used to estimate the detection function, as described in Sect. 7.2.1. The detection function model selected by AIC included covariates for Beaufort sea state, estimated flock size, behaviour (flocks sitting, swimming, flying or flushing) and observer. The Horvitz–Thompson-like estimator of Eq. (7.4) was used to estimate abundance  $N_k$  in each segment.

The abundance estimates  $\hat{N}_k$  were modelled as described in Sect. 7.3, except that the model was more advanced than the generalized linear model of Eq. (7.9). A spatially-adaptive quasi-Poisson generalized additive model with an offset was



**Fig. 7.5** Map of the Nysted offshore windfarm study area. The map shows changes in predicted differences in long-tailed duck numbers before and after construction. Negative differences indicate fewer individuals after construction than before, positive differences indicate more. The  $\circ$  symbols indicate significant decreases post-construction, and the  $\times$  symbols indicate significant increases. The position of the windfarm is indicated by the hatched polygon and transect waypoints are indicated by 'T'

used, incorporating a one-dimensional smooth term for depth, and a spatially-adaptive two-dimensional smooth term for the  $(x, y)$  coordinate-based spatial surface. These smooth relationships were allowed to differ before and after construction in both shape and magnitude through the inclusion of interaction terms. Non-independence of repeat surveys of the same line and of neighbouring line segments covered on the same day was incorporated in the model. Full details are given by Petersen et al. (2011).

The predicted density surface is shown in Fig. 7.5. Also shown are the areas for which a significant increase or decrease in long-tailed duck numbers was observed post-construction. This shows that significant decreases occurred through most of the footprint of the windfarm, but there were no significant decreases outside the footprint. Significant increases were seen to the north-west and to the east of the study area. Overall abundance did not change significantly; the response of the birds

to the construction of the windfarm appears to have been to redistribute themselves within the study area. However, given the lack of replication, we cannot be sure that the redistribution is in response to the windfarm construction. If similar studies elsewhere find a similar effect, then confidence in assigning the redistribution to windfarm construction will be greater.

# Chapter 8

## Model-Based Distance Sampling: Full Likelihood Methods

### 8.1 Introduction

In the following, we consider the case that animals are not in clusters except when we state otherwise. In Sect. 6.1, we noted that for conventional distance sampling (CDS, Sect. 5.2), we can define the full likelihood to be

$$\mathcal{L}_{n,y} = \mathcal{L}_n \times \mathcal{L}_y. \quad (8.1)$$

For multiple-covariate distance sampling (MCDS, Sect. 5.3), the full likelihood also includes a component for the likelihood of the covariates  $\mathbf{z}$ :

$$\mathcal{L}_{n,z,y} = \mathcal{L}_n \times \mathcal{L}_z \times \mathcal{L}_{y|z}. \quad (8.2)$$

For mark-recapture distance sampling (MRDS, Sect. 5.4), we have a further component corresponding to the mark-recapture data:

$$\mathcal{L}_{n,\omega,z,y} = \mathcal{L}_n \times \mathcal{L}_\omega \times \mathcal{L}_z \times \mathcal{L}_{y|z}. \quad (8.3)$$

Borchers and Burnham (2004) give an excellent summary of these formulations.

In Chap. 5, we addressed maximization of the likelihood component  $\mathcal{L}_y$  (CDS),  $\mathcal{L}_{y|z}$  (MCDS) or  $\mathcal{L}_\omega \times \mathcal{L}_{y|z}$  (MRDS), and in Chap. 6, we showed how the design-based Horvitz–Thompson estimator can be adapted, using model-based estimates of the probability of detection, to estimate abundance. This mixture of design-based and model-based approaches is very effective in practice, but is conceptually unappealing, at least to statisticians.

In Chap. 7, we exploited the fact that we have the methods of Chap. 5 to maximize the component of the likelihood for distances  $y$  (together with the mark-recapture component for MRDS), and we have standard modelling software to maximize the likelihood component  $\mathcal{L}_n$  (or, in the case of plot count methods,

$\mathcal{L}_{\{n_k\}}$ , Eq. (7.8)). Thus the methods of Chap. 7 are relatively easy to implement. However, if we separately maximize different components of the full likelihood, we do not in general obtain the maximum likelihood estimates corresponding to the full likelihood, and quantifying precision of the parameter estimates is also problematic.

In this chapter, we address full-likelihood methods, for which all parameter estimates are obtained simultaneously, and for which precision is estimated using standard methods. We first consider likelihood formulations, then summarize two inference approaches, having specified the likelihood: maximum likelihood estimation and Bayesian inference. In earlier chapters, we have used maximum likelihood methods, applied to components of the full likelihood. When we apply them to the full likelihood, we cannot use standard software to fit our models; in these circumstances, there may be advantages to using Bayesian methods instead. When using maximum likelihood methods, we will use AIC for model selection and AIC weights for model averaging. For the Bayesian approach, we describe a reversible-jump Markov chain Monte Carlo (RJMCMC) algorithm which allows us to conduct model-averaging across a large number of possible models.

## 8.2 Formulating the Likelihood

### 8.2.1 Conventional Distance Sampling

#### 8.2.1.1 Exact Distance Data

We first consider the model-based equivalent of conventional distance sampling, for which the detection function is a function of distance from the line or point alone, and for which the standard assumptions of Sect. 1.7 hold. In this case, the two components from Eq. (8.1) are given by:

$$\mathcal{L}_n = \binom{N}{n} (\pi_c P_a)^n (1 - \pi_c P_a)^{N-n} \quad (8.4)$$

and

$$\mathcal{L}_y = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \frac{g(y_i)\pi(y_i)}{P_a}. \quad (8.5)$$

Here,  $\pi_c$  is the probability that an animal, detected or not, falls in one of the surveyed plots (so that for a simple random design,  $\pi_c = a/A$ , where  $a$  is the total area of surveyed plots and  $A$  is the size of the whole study area);  $N$  is the total population size;  $P_a = \int_0^w g(y)\pi(y) dy$  is the expected probability of detection for an animal that is on one of the surveyed plots, where expectation is over distance  $y$  ( $0 \leq y \leq w$ ); assuming random placement of lines or points,  $\pi(y) = 1/w$  for

line transect sampling and  $\pi(y) = 2y/w^2$  for point transect sampling. Thus  $P_a$  is given by  $\int_0^w g(y)dy/w = \mu/w$  for line transect sampling (see Sect. 5.2.2.1) and by  $2 \int_0^w yg(y)dy/w^2 = \nu/(\pi w^2)$  for point transect sampling (Sect. 5.2.3.1).

We can now write the full likelihood as

$$\mathcal{L}_{n,y} = \mathcal{L}_n \times \mathcal{L}_y = \binom{N}{n} (\pi_c)^n (1 - \pi_c P_a)^{N-n} \prod_{i=1}^n g(y_i) \pi(y_i). \quad (8.6)$$

Thus we substitute  $\pi(y_i) = 1/w$  (line transect sampling) or  $\pi(y_i) = 2y_i/w^2$  (point transect sampling), and specify a model for the detection function. This formulation is given by Borchers et al. (2002, p. 141). For the case of line transect sampling, Royle and Dorazio (2008, p. 232) give the same formulation, and refer to it as individual-based modelling, because each detected animal  $i$  has its own detection distance  $y_i$ .

Using maximum likelihood (Sect. 8.3) or Bayesian (Sect. 8.4) methods, we fit our model to the data. As population size  $N$  is one of the unknown parameters of the full likelihood, this allows us to estimate  $N$ , and to quantify the precision of our estimate.

In the case of clustered populations, an additional likelihood component corresponding to a model for cluster sizes may be required. In the case of size bias where larger clusters have higher detection probabilities, this model may consist of a regression model where the logarithm of cluster size is regressed on the estimated probability of detection, given the cluster's distance from the line or point (see Sect. 6.3.1.3). If detection probabilities depend on cluster size, we refer the reader to Sect. 8.2.2. When detection probabilities are independent of cluster sizes, we may estimate mean cluster size in the population by the sample mean of observed cluster sizes. Then,  $N$  from Eq. (8.6) represents cluster abundance and individual abundance may be obtained by multiplying  $N$  by the mean cluster size.

### 8.2.1.2 Grouped Distance Data

If the distance data are grouped (*i.e.* binned), with intervals defined by the cutpoints  $c_0 = 0, c_1, \dots, c_u = w$ , then we replace  $\mathcal{L}_y$  of Eq. (8.1) by the multinomial likelihood of Eq. (5.26):

$$\mathcal{L}_m = \left( \frac{n!}{\prod_{j=1}^u m_j!} \right) \prod_{j=1}^u f_j^{m_j} \quad (8.7)$$

where  $m_j$  is the number of detections in distance interval  $j$ , with  $\sum_{j=1}^u m_j = n$ , and

$$f_j = \int_{c_{j-1}}^{c_j} f(y) dy = \frac{\int_{c_{j-1}}^{c_j} g(y) \pi(y) dy}{P_a}. \quad (8.8)$$

We can again use  $\mathcal{L}_n$ , the likelihood for the binomial model for counts from Eq. (8.4), and the full likelihood is now  $\mathcal{L}_{n,m} = \mathcal{L}_n \times \mathcal{L}_m$ . As for the exact distances case, we now substitute in the appropriate form for  $\pi(y)$ , and specify a model for the detection function  $g(y)$ . It then remains to use maximum likelihood or Bayesian methods to fit the model to the data.

In the case of clustered populations, the same considerations hold as for the equivalent model using exact distance data from Sect. 8.2.1.1.

## 8.2.2 Multiple-Covariate Distance Sampling

Royle and Dorazio (2008, p. 232) referred to their distance sampling model as an individual-based model, because each individual  $i$  has associated with it a distance  $y_i$ . MCDS methods give much greater flexibility for developing individual-based models. In this section, we consider only exact-distance models, because if distances are grouped, then we do not have distances to detected individuals, and cannot develop an individual-based model. However, if covariates are associated with the line or point, or in the case of stratified designs, the stratum, rather than with the individual, then we may need MCDS models when the distances are grouped. We will address this in Sect. 8.2.4.2.

If we extend the above model-based setting for CDS to MCDS, then the full likelihood consists of three components: the likelihood for the count model (Eq. (8.4)), the likelihood for the distribution of covariates  $\mathbf{z}$  that are part of the detection model, and the likelihood for the observed distances  $y$  given covariates  $\mathbf{z}$ :  $\mathcal{L}_n \times \mathcal{L}_z \times \mathcal{L}_{y|\mathbf{z}}$ . Given random line or point placement, we can assume that the joint distribution  $\pi_{y,\mathbf{z}}(y, \mathbf{z}) = \pi(y)\pi_z(\mathbf{z})$ ; that is that the distribution of distances  $y$  for all animals on sample plots (whether detected or not) is independent of that of covariates  $\mathbf{z}$ . This allows us to factorize the joint likelihood  $\mathcal{L}_{z,y}$  as the product  $\mathcal{L}_z \times \mathcal{L}_{y|\mathbf{z}}$  (Borchers and Burnham 2004, p. 20). We now have

$$\mathcal{L}_n = \binom{N}{n} (\pi_c P_a)^n (1 - \pi_c P_a)^{N-n} \quad (8.9)$$

$$\mathcal{L}_z = \prod_{i=1}^n \frac{P_a(\mathbf{z}_i) \pi_z(\mathbf{z}_i)}{P_a} \quad (8.10)$$

and

$$\mathcal{L}_{y|z} = \prod_{i=1}^n f(y_i | \mathbf{z}_i) = \prod_{i=1}^n \frac{g(y_i, \mathbf{z}_i) \pi(y_i)}{P_a(\mathbf{z}_i)} \quad (8.11)$$

where

$$P_a(\mathbf{z}_i) = \int_0^w g(y, \mathbf{z}_i) \pi(y) dy \quad (8.12)$$

and

$$P_a = \int_{\mathbf{z}} P_a(\mathbf{z}) \pi_{\mathbf{z}}(\mathbf{z}) d\mathbf{z} . \quad (8.13)$$

Thus we have the full likelihood

$$\mathcal{L}_{n,z,y} = \mathcal{L}_n \times \mathcal{L}_z \times \mathcal{L}_{y|z} = \binom{N}{n} (\pi_c)^n (1 - \pi_c P_a)^{N-n} \prod_{i=1}^n \pi_{\mathbf{z}}(\mathbf{z}_i) g(y_i, \mathbf{z}_i) \pi(y_i) . \quad (8.14)$$

In general, the integral of Eq. (8.13) will be a multiple integral. Inference is now more problematic because  $\pi_{\mathbf{z}}(\mathbf{z})$ , unlike  $\pi(y)$ , is unknown, and so we need to specify a model for it. Where we have just a single covariate, and one for which we can specify a suitable model (such as a negative binomial model if the covariate is a count, such as cluster size), then this approach may be useful. However, with multiple covariates, and the need to specify a model for their joint distribution, this approach is unappealing. In this circumstance, we can instead restrict inference to the conditional likelihood  $\mathcal{L}_{y|z}$ . To estimate abundance, we revert to the Horvitz–Thompson-like estimators of Sect. 6.4, using  $P_a(\mathbf{z}_i)$  to calculate the inclusion probability for detection  $i$  (Borchers and Burnham 2004, p. 21). We estimate  $P_a(\mathbf{z}_i)$  by maximizing  $\mathcal{L}_{y|z}$ . When detections are of individual animals, we use the Horvitz–Thompson-like estimator for estimating abundance  $N$ :

$$\hat{N} = \sum_{i=1}^n \frac{1}{\pi_c P_a(\mathbf{z}_i)} . \quad (8.15)$$

For clustered populations, if  $n$  animal clusters are recorded, Eq. (8.15) gives the abundance of clusters, and animal abundance is estimated by

$$\hat{N} = \sum_{i=1}^n \frac{s_i}{\pi_c P_a(\mathbf{z}_i)} \quad (8.16)$$

where  $s_i$  is the size of the  $i^{\text{th}}$  detected cluster.



### 8.2.3 Mark-Recapture Distance Sampling

We can extend MCDS full-likelihood methods to MRDS simply by including an additional likelihood component,  $\mathcal{L}_\omega$  from Eq. (5.58). Then the full likelihood is given by  $\mathcal{L}_{n,\omega,z,y} = \mathcal{L}_n \times \mathcal{L}_\omega \times \mathcal{L}_z \times \mathcal{L}_{y|z}$ . Again, if we wish to avoid the problem of specifying a suitable model for  $\pi_z(\mathbf{z})$ , we estimate abundance using a Horvitz–Thompson-like estimator (Sect. 6.4).

### 8.2.4 Plot Count Models

#### 8.2.4.1 Exact Distance Data

So far, we have ignored the spatial information in our plot counts. For simplicity we begin with detections of single animals and denote the unknown number of animals on plot  $k$  ( $k = 1, \dots, K$ ) by  $N_k$ , and the number of animals detected on plot  $k$  by  $n_k$ , where  $\sum_{k=1}^K n_k = n$ . Here we define plot  $k$  to mean the strip of half-width  $w$  and length  $l_k$  centred on line  $k$  (line transect sampling) or the circle of radius  $w$  centred on point  $k$  (point transect sampling).

We could extend the binomial model of Eq. (8.4), or the equivalent MCDS model of Eq. (8.9), by a multinomial model with  $K + 1$  categories, where  $K$  is the number of plots, and category  $K + 1$  is for undetected animals (either because they are not on one of the plots or because they are, but are not detected). This would allow direct inference on abundance  $N$ . If we were interested only in comparing plot abundances or densities, and not in estimating abundance in a wider study area, then we could exclude the category of animals that are not on any sample plot, so that category  $K + 1$  corresponds to animals on a plot, but which are undetected. This approach is developed by Buckland et al. (submitted), but we do not pursue it here. Instead, we concentrate on Poisson models.

If we do not have covariates  $\mathbf{z}$  in our model for the detection function, then the full likelihood is  $\mathcal{L}_{\{n_k\},y} = \mathcal{L}_{\{n_k\}} \times \mathcal{L}_y$ , where  $\mathcal{L}_{\{n_k\}}$  is the Poisson likelihood of Eq. (7.8), and  $\mathcal{L}_y$  is as given in Eq. (8.5):

$$\mathcal{L}_{\{n_k\},y} = \mathcal{L}_{\{n_k\}} \times \mathcal{L}_y = \prod_{k=1}^K \frac{\lambda_k^{n_k} \exp[-\lambda_k]}{n_k!} \times \prod_{i=1}^n \frac{g(y_i) \pi(y_i)}{P_a}. \quad (8.17)$$

If we do have covariates in the detection function, then the full likelihood is  $\mathcal{L}_{\{n_k\},z,y} = \mathcal{L}_{\{n_k\}} \times \mathcal{L}_z \times \mathcal{L}_{y|z}$  where  $\mathcal{L}_z$  is given by Eq. (8.10) and  $\mathcal{L}_{y|z}$  by Eq. (8.11). It can be difficult to specify a good model for the covariates  $\mathbf{z}$ , especially when there are several potential covariates. In this circumstance, provided these covariates are recorded at the plot level or higher (as distinct from the individual animal level), then we can simplify the modelling simply by omitting the component  $\mathcal{L}_z$  from the likelihood, and maximizing the remainder, conditional on  $\mathbf{z}$ .

As in Sect. 7.2.2, we can specify models for  $\lambda_k$  of the form of Eq. (7.5), where the covariates  $x_{qk}$  might define the design in the case of a designed distance sampling experiment, or might be spatial covariates for a spatial model, or might simply be any explanatory variables that are potentially useful for modelling animal density. The difference from Sect. 7.2.2 is that we do not now maximize the two likelihood components separately, but instead maximize the full likelihood, or use Bayesian methods to draw inference on all unknown parameters. The parameter  $P_k$ , which is the probability of detection of an animal on plot  $k$  unconditional on its distance from the line or point (Sect. 7.2), appears in the offset; for the two-stage approach, it was estimated in stage 1, then treated as known in stage 2, while now it is estimated along with all other parameters in a single step, so that our model for  $\lambda_k$  is

$$\lambda_k = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q + \log_e(a_k P_k) \right) \quad (8.18)$$

where  $\hat{P}_k$  in the offset of Eq. (7.5) has been replaced by the true but unknown value  $P_k$ , which is a function of the parameters of the detection function (below).

As for two-stage methods, linear terms in the above expression may be replaced by smooth terms, and random effects can be added (Sect. 8.2.5.1). Oedekoven et al. (2014) proposed the above approach, with the inclusion of a random effect for location in the model for  $\lambda_k$  (see Sect. 8.5.2).

If counts are summed across repeat visits, the offset term is multiplied by the effort, where effort is defined to be the number of repeat visits; and if counts are summed across replicate plots, plot size  $a_k$  is the combined size of the plots whose counts have been combined.

The product  $a_k P_k$  is the effective area surveyed on plot  $k$ . If the detection function is assumed to be a function of distance alone, then  $P_k = P_a$  for every point. When covariates  $\mathbf{z}_k$  are recorded at the plot level or higher, then

$$P_k = \int_0^w g(y, \mathbf{z}_k) \pi(y) dy \quad (8.19)$$

where  $\pi(y) = 1/w$  for line transects and  $\pi(y) = 2y/w^2$  for point transects. If there are additional covariates say  $\mathbf{u}_{ki}$  recorded at the individual level, we can integrate them out, provided that we can specify a joint distribution  $\pi_{u_k}(\mathbf{u}_k)$  for them:

$$P_k = \int_{u_k} \int_0^w g(y, \mathbf{z}_k) \pi(y) \pi_{u_k}(\mathbf{u}_k) dy d\mathbf{u}_k. \quad (8.20)$$

This approach is not straightforward to implement. In practice when we have individual covariates, we might prefer to adopt the less elegant but much simpler option of a plot abundance model, as described in Sect. 7.3.

Note that neither population size  $N$  nor plot abundances  $N_k$  appear as parameters in the likelihood. For designed distance sampling experiments, we only wish to

compare densities on the plots, and have no interest in estimating  $N$  in a wider study area. For spatial distance sampling models, we can predict density throughout the study area, and so can use numerical integration under the fitted density surface to estimate abundance either for the full study area or for any subset of it. We can also estimate  $N_k$  by  $\hat{N}_k = \hat{\lambda}_k / \hat{P}_k$ .

When detections are of clusters of animals, we have several options. One is to define  $n_k$  from Eq. (8.17) to be the number of clusters detected on plot  $k$  and model  $\lambda_k$ , the expected number of clusters on plot  $k$ , using Eq. (8.18). We then add a model for cluster size to Eq. (8.17). If probability of detection is thought to be unaffected by cluster size, this may be a simple count model such as Poisson or negative binomial with the zero class truncated, and the observed cluster sizes are substituted into the assumed probability function to form the corresponding likelihood component. If we suspect that larger clusters are more likely to be detected, then a size-biased regression model (Sect. 6.3.1.3) might be used. If we are modelling density as a function of spatial covariates, then it would be preferable to model cluster size similarly (Cañadas and Hammond 2006).

An alternative to the above strategies is to define  $n_k$  from Eq. (8.17) to be the number of individuals detected on plot  $k$  and model the expected number of individuals  $\lambda_k$  on plot  $k$  using the covariate model from Eq. (8.18). Buckland et al. (submitted) ignores the fact that animals occur in clusters, and may not work well when cluster size is highly variable.

#### 8.2.4.2 Grouped Distance Data

For the case without covariates, let  $m_{jk}$  be the number of detections in distance interval  $j$  on plot  $k$ , with  $\sum_{j=1}^u m_{jk} = n_k$ . Adopting a Poisson model for these counts, and given  $E(n_k) = \lambda_k$ , then  $E(m_{jk}) = \lambda_k f_j$  for  $j = 1, \dots, u$ , where

$$f_j = \int_{c_{j-1}}^{c_j} f(y) dy = \frac{\int_{c_{j-1}}^{c_j} g(y) \pi(y) dy}{P_a} . \quad (8.21)$$

We can now write the full likelihood as

$$\prod_{k=1}^K \prod_{j=1}^u \frac{(\lambda_k f_j)^{m_{jk}} \exp(-\lambda_k f_j)}{m_{jk}!} . \quad (8.22)$$

For detection function covariates  $\mathbf{z}$  recorded at the plot level, or at the stratum level if the design is stratified (but not at the individual level), we can define

$$f_{jk} = \int_{c_{j-1}}^{c_j} f_{y|z}(y|\mathbf{z}_k) dy = \frac{\int_{c_{j-1}}^{c_j} g(y, \mathbf{z}_k) \pi(y) dy}{P_a(\mathbf{z}_k)} \quad (8.23)$$

giving the full likelihood

$$\prod_{k=1}^K \prod_{j=1}^u \frac{(\lambda_k f_{jk})^{m_{jk}} \exp(-\lambda_k f_{jk})}{m_{jk}!} . \quad (8.24)$$

Royle et al. (2004) adopted what appears to be a different strategy for grouped distance data. Again  $m_{jk}$  is the count of detected animals in distance interval  $j$  on plot  $k$ , with  $\sum_{j=1}^u m_{jk} = n_k$ . We define the proportion  $P_j$  of plot abundance that was observed within distance interval  $j$ :

$$P_j = \int_{c_{j-1}}^{c_j} g(y) \pi(y) dy \quad (8.25)$$

where  $g(y)$  and  $\pi(y)$  represent the detection function and the distribution of distances from the line or point in the population as before. The sum of proportions  $P_j$  over all  $u$  distance intervals gives the average detection probability  $P_a$ , i.e.  $\sum_{j=1}^u P_j = P_a$ , where  $P_a = \int_0^w g(y) \pi(y) dy$  (Eq. (5.38)).

The  $P_j$  represent the proportion of plot abundance  $N_k$  that was both located in distance interval  $j$  and detected, while the  $f_j$  represent the proportion of detected animals  $n_k$  that were located in distance interval  $j$ . Hence we have the relationship  $f_j = P_j/P_a$ .

As we do not observe the true abundances on the plot, we set  $E(N_k) = \kappa_k$  and model these using a log-linear Poisson model:

$$\kappa_k = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q \right) . \quad (8.26)$$

The observed counts in distance interval  $j$  are then modelled as a Poisson random variable,  $m_{jk} \sim \text{Poisson}(\kappa_k \times P_j)$ . The likelihood for this model is:

$$\mathcal{L}_{n,m} = \prod_{k=1}^K \prod_{j=1}^u \frac{(\kappa_k P_j)^{m_{jk}} \exp(-\kappa_k P_j)}{m_{jk}!} . \quad (8.27)$$

By noting that  $\lambda_k = \kappa_k P_a$ , we see that the above model for  $\kappa_k$  is equivalent to our model for  $\lambda_k$ , provided plot sizes are all the same, and arbitrarily set as  $a_k = 1$ :

$$\lambda_k = \kappa_k P_a = \exp \left( \sum_{q=1}^Q x_{qk} \beta_q + \log(P_a) \right). \quad (8.28)$$

Thus the Poisson rate corresponding to count  $m_{jk}$  is  $\lambda_{kj} = \lambda_k P_j / P_a = \kappa_k P_j$ , so that the likelihood of Eq. (8.27) is equivalent to that of Eq. (8.22). Here, we use Eq. (8.22), as it allows plot area  $a_k$  to vary. Note that when using the Poisson model for the expected abundances for this approach, we do not have separate components for the  $\mathcal{L}_{\{n_k\}}$  and  $\mathcal{L}_m$ . Oedekoven et al. (2014) adopted a different strategy which is essentially the grouped-data equivalent of Eq. (8.17), and which does have separate components for  $\mathcal{L}_{\{n_k\}}$  (using a generalized version of Eq. (7.8)) and  $\mathcal{L}_m$  (using Eq. (8.7)). However, given the equivalence of the plot count and plot abundance models, Eq. (8.22) is a special case of the more general likelihood of Oedekoven et al. (2013).

When we have covariates recorded at the individual detection level, such as cluster size, full model-based approaches for grouped distance data are not straightforward to implement. For factor-type covariates, the data become counts within each distance band and at each level of the covariate (or combination of levels for more than one covariate). Thus we can again model these counts assuming say a Poisson model as above.

## 8.2.5 Model Extensions

The advantage of adopting a model-based approach to distance sampling is that it gives much greater flexibility. Instead of focusing on estimating abundance, we can investigate and model various issues of interest. We consider here some ways in which our models can be extended.

### 8.2.5.1 Random Effects

We can add random effects to our models. Such models are often called mixed models (incorporating both fixed and random effects) or hierarchical models (because of their hierarchical error structure). In distance sampling, there are several possible uses of random effects (Oedekoven 2013). One example is when repeat visits are made to plots. For some purposes, we can just pool data from the repeat visits, but sometimes we may wish to model the separate plot counts. We can then define plot to be a random effect, which allows for the fact that repeat counts on a

given plot are likely to be correlated. Our model for expected plot count, Eq. (7.5), is readily extended:

$$\lambda_{kt} = \exp \left( \sum_{q=1}^Q x_{qkt} \beta_q + b_k + \log_e(a_k P_{kt}) \right) \quad (8.29)$$

where the subscript  $t$  indicates visit  $t$  to plot  $k$ . (If there are no time-varying covariates, we can drop the  $t$  subscript from this expression.) Typically, the random effects are assumed to be normally distributed:

$$b_k \sim N(0, \sigma_k^2). \quad (8.30)$$

Compared to the Poisson likelihoods given for counts above (e.g. Eq. (8.17)), the likelihood for the count model now includes a normal density for the random effects (Oedekoven et al. 2014) and is given by:

$$\mathcal{L}_{\{n_{kt}\}} = \prod_{k=1}^K \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_k} \frac{\lambda_{kt}^{n_{kt}} \exp[-\lambda_{kt}]}{n_{kt}!} \right\} \times \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp \left[ -\frac{b_k^2}{2\sigma_k^2} \right] db_k, \quad (8.31)$$

where  $T_k$  represents the total number of visits to plot  $k$ .

A similar example is when a sampling location comprises more than one line or point; the conservation buffer case study is an example, where each location has two points, one treatment and one control. Again, for some purposes, we can just pool the data for the location. However, we may wish to model the separate plot counts (as in the conservation buffer case study, so we can test for a treatment effect). In this case, we can define a random effect for location, to allow for correlation across plots within a single location:

$$\lambda_{kl} = \exp \left( \sum_{q=1}^Q x_{qkl} \beta_q + b_l + \log_e(a_{kl} P_{kl}) \right) \quad (8.32)$$

where  $l$  indicates location, and  $b_l \sim N(0, \sigma_l^2)$ . In this case, the likelihood for the counts becomes:

$$\mathcal{L}_{\{n_{kl}\}} = \prod_{l=1}^L \int_{-\infty}^{\infty} \left\{ \prod_{k=1}^K \frac{\lambda_{kl}^{n_{kl}} \exp[-\lambda_{kl}]}{n_{kl}!} \right\} \times \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp \left[ -\frac{b_l^2}{2\sigma_l^2} \right] db_l, \quad (8.33)$$

where  $L$  is the total number of locations.

We could also (or instead) make a coefficient  $\beta_q$  a location random effect, which would mean that the effect of the corresponding covariate  $x_q$  varies by location.

Random effects can also be included in the model for the detection function, to model any heterogeneity not accounted for by any covariates  $\mathbf{z}$  included in the

model (Oedekoven 2013; Oedekoven et al. 2015). For example if we have an MCDS model, the model for the scale parameter  $\sigma$ , given in Eq. (5.39), may be extended as:

$$\sigma(\mathbf{z}_i) = \exp \left( \alpha + t_i + \sum_{q=1}^Q \beta_q z_{qi} \right) \quad (8.34)$$

where  $i$  indicates individual  $i$ ,  $i = 1, \dots, n$ , and  $t_i \sim N(0, \sigma_t^2)$ .

The likelihood corresponding to the detection function is now

$$\mathcal{L}_g(\beta, \sigma_t | \mathbf{z}) = \prod_{i=1}^n \frac{\int_{-\infty}^{\infty} g(y_i | \mathbf{z}, t_i) \pi(y_i) N(t_i, 0, \sigma_t) dt}{\int_{-\infty}^{\infty} \int_0^w g(u | \mathbf{z}, t_i) \pi(u) du N(t_i, 0, \sigma_t) dt} \quad (8.35)$$

where  $\pi(y)$  represents the distribution of animals (whether detected or not) with increasing distance from the line or point (Oedekoven et al. 2015). It is given by  $\pi(y) = 1/w$  for line transects and, because this is independent of  $y$ , it cancels top and bottom, and so can be omitted from Eq. (8.35). For point transects,  $\pi(y) = 2y/w^2$ .

We define  $N(t_i, 0, \sigma_t) = \exp \left[ -0.5 \left( \frac{t_i}{\sigma_t} \right)^2 \right] (\sqrt{2\pi}\sigma_t)^{-1}$ . Again, we could make a coefficient  $\beta_q$  a random effect, so that the effect of the corresponding  $x_q$  varies by individual.

### 8.2.5.2 More on Spatial Modelling

Spatial distance sampling models have several advantages over design-based extrapolation from the plots to a region. First, the methods allow extrapolation when the plots (strips or circles) are not laid down according to a randomized design. Thus the methods are useful for marine surveys conducted from platforms of opportunity, or terrestrial surveys conducted at points or along lines where access is possible. Second, abundance can be estimated for any part of the study area simply by integrating under the appropriate region of the fitted density surface. This is useful if management of a population is based on estimates of abundance within defined management areas which might not correspond to strata in the design. Third, the models allow the analyst to explore ecological relationships: how density is related to habitat, climate variables, altitude, and so on.

We have proposed relatively simple ways to fit spatial distance sampling models. For line transect sampling in particular, unless lines are very short, these methods can be unsatisfactory, as longer lines must be arbitrarily divided into short segments, so that spatial covariates do not vary appreciably within a segment. A more satisfactory approach is to model the data as a spatial point process. We treat the locations of detected animals as the realization of a spatial point process, that has been thinned by the detection process, leaving just the detected animals. When animals occur in clusters, the process becomes a marked point process, where

the mark is the cluster size. Högmader (1991) was the first to consider such an approach. Hedley and Buckland (2004) laid down the theory for such models, as applied to line transect sampling. Johnson et al. (2010) generalized the approach, and implemented their method in software `DSpat` (Johnson et al. 2009), which is available as an R library. Buckland et al. (submitted) note how, for any of our Poisson point count models, we can formulate a spatial nonhomogeneous Poisson process model, in which the Poisson parameters  $\lambda_k$ , representing expected count on plot  $k$ , are represented as an integral over the plot of the animal density surface, thinned by the detection function.

## 8.3 Model Fitting: Maximum Likelihood Methods

### 8.3.1 Maximum Likelihood Estimation

Maximum likelihood methods in principle are very straightforward, once we have written down our likelihood. The standard method for maximizing a mathematical function is to differentiate it with respect to each unknown parameter, set each derivative to zero, and solve the simultaneous equations, giving the maximum likelihood estimates of the parameters. Thus for any full likelihood  $\mathcal{L}$  of this chapter, we write it as a function of the parameters. We denote the parameters of the count model by  $\beta$  and the parameters of the detection function model by  $\theta$ . We can then write the full likelihood as  $\mathcal{L} = \mathcal{L}(\beta, \theta)$  where  $\mathcal{L} = \mathcal{L}_{n,y}$  for CDS,  $\mathcal{L} = \mathcal{L}_{n,z,y|z}$  for MCDS, and  $\mathcal{L} = \mathcal{L}_{n,\omega,z,y|z}$  for MRDS.

We would now like to differentiate  $\mathcal{L}(\beta, \theta)$  with respect to each of the parameters in vectors  $\beta$  and  $\theta$ . However,  $\mathcal{L}(\beta, \theta)$  is formed as a product of terms. By taking the log of the likelihood, we obtain a sum of terms. A sum is simpler to differentiate than a product, and as logarithm is a one-to-one transformation, exactly the same values of the parameters maximize both the likelihood and the log-likelihood. Thus we define the log-likelihood as  $\ell(\beta, \theta) = \log_e\{\mathcal{L}(\beta, \theta)\}$ . Differentiating, we have

$$\frac{d\ell(\beta, \theta)}{d\beta} = 0 \quad \text{and} \quad \frac{d\ell(\beta, \theta)}{d\theta} = 0. \quad (8.36)$$

Note that, if there are say three parameters in the  $\beta$  vector, then  $\frac{d\ell(\beta, \theta)}{d\beta}$  signifies that the log-likelihood is differentiated with respect to each parameter in turn, and each derivative set to zero. We now solve the simultaneous equations in Eq. (8.36) to obtain the maximum likelihood estimates of the parameters.

In the case that random effects are contained in the model, we can no longer log-transform the likelihood in the same manner. Now, we first need to integrate out the random effect leaving the terms within the integral untransformed. For example



if we include random effects for plot in the count model as in Eq. (8.31), the log-likelihood becomes:

$$\begin{aligned} \ell(\boldsymbol{\beta}) &= \log_e \mathcal{L}(\boldsymbol{\beta}) \\ &= \sum_{k=1}^K \log_e \left[ \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_k} \frac{\lambda_{kt}^{n_{kt}} \exp[-\lambda_{kt}]}{n_{kt}!} \right\} \times \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp \left[ -\frac{b_k^2}{2\sigma_b^2} \right] db_k \right]. \end{aligned} \quad (8.37)$$

Here, the random effect standard deviation  $\sigma_b$  is contained in  $\boldsymbol{\beta}$ . In the case that we wish to log-transform Eq. (8.35) where the random effect is included in the detection model, the random effect standard deviation  $\sigma_t$  is contained in  $\boldsymbol{\theta}$ .

### 8.3.2 Estimating Precision

We use the information matrix to estimate precision when using maximum likelihood methods. Suppose we have just a single parameter  $\beta$  in the count model, and a single parameter  $\theta$  in the detection function model. (The methods extend in the obvious way if we have more parameters.) Then the information matrix is given by

$$I(\beta, \theta) = \begin{bmatrix} E \left[ \left\{ \frac{d\ell}{d\beta} \right\}^2 \right] & E \left[ \frac{d\ell}{d\beta} \frac{d\ell}{d\theta} \right] \\ E \left[ \frac{d\ell}{d\theta} \frac{d\ell}{d\beta} \right] & E \left[ \left\{ \frac{d\ell}{d\theta} \right\}^2 \right] \end{bmatrix} = - \begin{bmatrix} E \left[ \frac{d^2\ell}{d\beta^2} \right] & E \left[ \frac{d^2\ell}{d\beta d\theta} \right] \\ E \left[ \frac{d^2\ell}{d\theta d\beta} \right] & E \left[ \frac{d^2\ell}{d\theta^2} \right] \end{bmatrix} \quad (8.38)$$

where  $\ell \equiv \ell(\beta, \theta)$  is the log-likelihood and  $E[\cdot]$  indicates expectation over the observations. Thus we can obtain an estimate  $\hat{I}(\beta, \theta)$  of the information matrix either by taking products of first derivatives or by taking second derivatives of the log-likelihood. In either case, we substitute in maximum likelihood estimates of the parameters to obtain our estimate  $\hat{I}(\beta, \theta)$ . The diagonal elements of the inverse of the estimated information matrix are estimates of the variances of the corresponding parameters. We can also estimate the variance of functions of the parameters, by using the appropriate Jacobian, which is found by differentiating the function of parameters with respect to each of the parameters. See for example Buckland et al. (2001, p. 64).

Algorithms such as the R routines `optim` and `nlm` may be used to maximize the likelihood and estimate the information matrix. In the conservation buffer case study for northern bobwhite coveys (available on the book website), we show how to estimate density of animals on the plot and the corresponding variance using a plot count model.

We can incorporate model uncertainty in our estimates of precision using model averaging as noted in Sect. 5.1.2.

## 8.4 Model Fitting: Bayesian Methods

We describe a Bayesian approach to fitting distance sampling models which is applicable to any of the full likelihood approaches of this chapter. The methods described here are summarized from Oedekoven et al. (2014). The parameter space is explored using a Metropolis–Hastings (MH) updating algorithm so that different likelihoods and prior distributions for the parameters are easily implemented. A reversible-jump Markov chain Monte Carlo (RJMCMC) algorithm allows model uncertainty to be incorporated. This may include different key functions for the detection function model and, in the case of the covariate models, different combinations of covariates for both the detection function and the count models.

### 8.4.1 MCMC Algorithm

We use an MCMC algorithm to explore the posterior distribution of the parameters given the data. Here we focus on the MH-update (Hastings 1970; Metropolis et al. 1953) as some of the likelihood functions that may be used to form the posterior conditional distributions of parameters are non-standard (*e.g.* a half-normal or hazard-rate detection function which may include a covariate model for the scale parameter).

We use a single-update random-walk MH algorithm with normal proposal density, where we cycle through each parameter in the full likelihood. To use a simple scenario, we assume a CDS likelihood  $\mathcal{L}_{n,y}$ , which is a function of parameters  $\{\beta_0, \theta\}$ . During each iteration of the MCMC algorithm, we update each of these parameters by proposing to move to a new state and accepting this move with some probability. To update, say, parameter  $\beta_0$  at iteration  $t$  with current value  $\beta_0^t$ , we propose to move to a new state,  $\beta_0'$ , with  $\beta_0' \sim N(\beta_0^t, \sigma_{\beta_0}^2)$  (Hastings 1970; Davison 2003), where  $\sigma_{\beta_0}^2$  is the proposal variance for  $\beta_0$ . This newly proposed state is accepted as the new state with probability  $\alpha(\beta_0'|\beta_0^t)$  given by:

$$\alpha(\beta_0'|\beta_0^t) = \min \left( 1, \frac{\mathcal{L}_{n,y}(\beta_0', \theta^t) p(\beta_0') q(\beta_0^t|\beta_0')}{\mathcal{L}_{n,y}(\beta_0^t, \theta^t) p(\beta_0^t) q(\beta_0'|\beta_0^t)} \right). \quad (8.39)$$

where  $p(\beta_0^t)$  is the prior distribution for  $\beta_0$  evaluated at  $\beta_0 = \beta_0^t$  and  $q(\beta_0^t|\beta_0')$  denotes the proposal density of the newly proposed state  $\beta_0^t$  given that the current state is  $\beta_0'$ . We note that the terms  $q(\beta_0^t|\beta_0')$  and  $q(\beta_0'|\beta_0^t)$  cancel in the acceptance probability as we use a symmetrical proposal distribution (normal). Proposal variances can be chosen using pilot-tuning where we aim to obtain acceptance rates between 20 and 40 % (Gelman et al. 1996). Increasing the variance on average results in decreasing acceptance probabilities and *vice versa*. If prior knowledge on the parameters exists, it may be incorporated via the prior distributions. If no prior knowledge on the parameters exists, uninformative priors, such as uniform, may be placed on all parameters.

We note that in the case of fitting random effects (Sect. 8.2.5.1) using MH-updating, the random effects are not integrated out analytically. Instead, we use a data augmentation scheme where the individual random effects coefficients are included as parameters (or auxiliary variables) in the model and updated at each iteration of the MCMC algorithm.

### 8.4.2 Model Averaging: Reversible-Jump MCMC

Using the MH approach from Sect. 8.4.1 with a single model, we conditioned on the model being correct. To include model uncertainty in inference using Bayesian methods, we decide on a set of plausible models and implement a reversible-jump MCMC algorithm. As will become evident below, the number of competing models can be very large with relatively little extra computing time.

To discriminate between competing models, we treat the model itself as a parameter and form the joint posterior distribution over both parameters and models. The joint posterior distribution of models and parameters is given by:

$$\pi_{n,y}(\beta_m, \theta_m, m) \propto \mathcal{L}_{n,y}(\beta_m, \theta_m, m) p(\beta_m, \theta_m | m) p(m), \quad (8.40)$$

where  $\mathcal{L}_{n,y}(\beta_m, \theta_m, m)$  denotes the probability density function of the data given current parameter values  $\beta_m$  (from the count model) and  $\theta_m$  (from the detection model) and model  $m$ ,  $p(\beta_m, \theta_m | m)$  the prior distribution for model parameters  $\beta_m$  and  $\theta_m$ , and  $p(m)$  the prior probability of model  $m$ . The symbol  $\propto$  means ‘is proportional to’, and indicates that we have omitted a constant (*i.e.* a term that is not a function of the parameters) from the expression. This approach may be applied to any full likelihood function of this chapter. We use the Poisson likelihood from Sect. 8.2.4, where counts are modelled using covariates  $\beta_m$ , to illustrate Eq. (8.40).

In this example, model  $m$  represents a combination of a count model and a detection model. This approach can easily be extended to use the binomial likelihood for counts given in Sect. 8.2.1.1 instead of the Poisson likelihood, or the MCDS likelihood of Sect. 8.2.2, or the MRDS likelihood of Sect. 8.2.3. It cannot be used, however, to determine which of these approaches, the Poisson, the Poisson with random or mixed effects, the binomial or the MRDS likelihoods, provides the best fit to the data. It can be used to average predictions from the fits of different detection functions, *e.g.* the half-normal and the hazard-rate which may also include different covariate models for the scale parameter using MCDS methods.

The RJMCMC algorithm is used to explore simultaneously the parameter and model space (Green 1995). For this algorithm, each iteration involves two steps: a within-model move and a between-model move. During the within-model move, the Metropolis–Hastings (MH) algorithm is used to update the parameters given the model (as described above in Sect. 8.4.1). During the between-model move, the reversible-jump (RJ) step, model  $m$  conditional on the current parameter values is updated. This move involves a proposal to update the model itself.

We illustrate this move with a simple example: suppose the chain is in model  $m$  and we propose to move to model  $m'$ . A bijective function describes the relationship between the current and proposed parameters and is used to convert parameters from model  $m$  to parameters for model  $m'$ . In a simple scenario, say where model  $m$  contains parameters  $\beta = \{\beta_0, \beta_1\}$  and model  $m'$  contains parameters  $\beta' = \{\beta'_0, \beta'_2\}$ , the bijective function might be expressed as an identity function:

$$\beta'_0 = \beta_0 \quad u' = \beta_1 \quad \beta'_2 = u. \quad (8.41)$$

Here  $u$  and  $u'$  are random samples from some proposal distributions for the respective parameters. The acceptance probability may then be expressed as:

$$A = \frac{\pi_{n,y}(\beta', m') P(m|m') q'(u')}{\pi_{n,y}(\beta, m) P(m'|m) q(u)} |J|, \quad (8.42)$$

where  $P(m'|m)$  denotes the probability of proposing to move to model  $m'$  given that the chain is in model  $m$ ,  $q(u)$  and  $q'(u')$  are the proposal densities of  $u$  and  $u'$ , and  $|J|$  is the Jacobian (which equals one if the bijective function is the identity function).

For the RJ step, two main strategies may be followed. In cases where models differ only in the combination of covariates included, a single RJ step may involve going through each covariate and proposing to delete or add it depending on whether it is in the current model or not. This involves generating a value for the new parameter from a proposal distribution (if we propose to add it) or setting it to zero (if we propose to delete it) and calculating the acceptance probability each time we propose to add or delete a parameter.

In those cases where all parameters of the newly-proposed model change, one RJ step involves generating new values for all parameters of the new model and accepting or rejecting the new model based on the above acceptance probability. A proposed move from a half-normal detection function model to a hazard-rate model represents a simple example for this scenario.

Posterior model probabilities are estimated as the proportion of time the chain spent in a particular model after the burn-in.

### 8.4.3 Bayesian Inference

To obtain summary statistics of parameters of interest, we use the respective marginal posterior distributions. The marginal posterior distributions for the parameters of interest are obtained by integration of the joint posterior distribution (King et al. 2009). Instead of integrating analytically, we use our MCMC (excluding the burn-in phase) as simulated samples from the joint posterior distribution. Often-used summary statistics include the mean, standard deviation and central 95 % credible intervals. For the case that plot abundances,  $N_k$  are of interest, these need to be calculated during each iteration of the MCMC algorithm.

In the case of model averaging, model probabilities are given by the proportion of time that the RJMCMC algorithm spends in the respective model after the burn-in phase. Model parameters are now conditional on the model. Hence, for obtaining summary statistics of parameters  $\beta_m$  of model  $m$ , we use only those iterations of the chain for which the algorithm was in model  $m$ .

## 8.5 Conservation Buffer Case Study

### 8.5.1 *Indigo Buntings: A Full Likelihood Approach Using Maximum Likelihood Methods*

Here, we summarize analyses carried out by Oedekoven et al. (2013). The data were of indigo buntings, a common passerine found throughout the eastern United States in the northern summer where it breeds in brushy borders to deciduous woodland. The objective of the study was to assess whether conservation buffers at field edges increased bird densities (Sect. 3.2.2.1). During the breeding seasons of 2006–2007, point transect surveys were conducted from one point per treatment field, located in the buffer at the edge of the field. For each treatment point, an unbuffered control point was located on the edge of a field of the same agricultural use, located 1–3 km away. The matched points were surveyed concurrently to ensure similar conditions. Each pair of adjacent treatment and control points was considered a site, and each site was surveyed at least one and at most four times per survey year.

The indigo bunting is a sexually dimorphic species: males are bright blue, particularly during breeding, while females are dull brown. Males engage in territorial song during the breeding season, making them easy to detect. Hence, observers recorded all male indigo buntings detected visually or aurally in a 10-min period in predetermined distance intervals (0–25, 25–50, 50–100, 100–250, 250–500, >500 m). The data were truncated at  $w = 100$  m, as few were detected at the larger distances. Only those sites surveyed at least once in each of the two survey years were included. An additional criterion was that each state included in the analysis contained >50 detections. The 446 sites satisfying these criteria were located in nine states.

With just three distance intervals, and allowing a degree of freedom for assessing model fit, only one-parameter models were considered for the detection function: the half-normal model and the hazard-rate model with fixed shape parameter. Heterogeneity was modelled using three possible stratification factors: year (2006 or 2007), type (buffer or control field), or state (nine levels). When a stratification factor was included in the model, a separate detection function model was specified for each level of the factor; the likelihood component  $\mathcal{L}_y$  was thus a product of the stratum-specific components. AIC was used to select appropriate values for the fixed shape parameters for the global or the stratified hazard-rate functions.

The methods of Sect. 8.2.4.2 were used to model the data, with the addition of site random effects (Sect. 8.2.5.1). Possible covariates were year, type, date (taken as a continuous covariate) and state. In these models, the parameter of interest was the covariate type. A significant type term in the plot abundance model would indicate a difference in expected bird densities between control plots and plots with buffers. The site random effects term was assumed to be normally distributed with mean zero. Analytical standard errors were obtained from the Hessian matrix.

During the two survey years included in this study, 2006–2007, in total 2924 counts were made at 446 sites, and 3785 male indigo buntings were detected in the three innermost distance intervals. The data were analysed by maximizing the full likelihood (Sect. 8.3).

A forward stepwise procedure using AIC selected a hazard-rate model for the detection function, fitted separately for each state, and a count model with type (buffer or control) and state as factors and date as a continuous covariate. For comparison, the same data were analysed with the two-stage approach of Sect. 7.2 as well as with the Metropolis–Hastings algorithm described in Sect. 8.4.1 using the full likelihood from Sect. 8.3. For these two approaches, we did not include model selection in the analyses but conditioned on the best model determined by maximizing the full likelihood. We show parameter estimates for the three methods in Table 8.1.

The full likelihood intercepts of Table 8.1 are not comparable with the two-stage intercept. For the two-stage model, the intercept corresponds to the logarithm of birds per square metre, while the full likelihood intercepts correspond to the logarithm of birds per plot area  $a = \pi w^2 = 31416$  square metres; if we take the full likelihood ML intercept  $-0.99$  and multiply by  $\log_e 31416$ , we obtain  $-10.25$ , close to the intercept for the two-stage model. The full likelihood intercept of the MH-update was smaller compared to the other two approaches ( $-0.63 \times \log_e 31416 = -6.52$ ). Central 95 % credible intervals for this parameter ( $-1.32, 0.12$ ) overlapped the value of the ML full likelihood intercept.

The analytic standard errors for the count model for the two-stage method fail to incorporate uncertainty in estimating the detection function parameters, and so it is perhaps unsurprising that the two-stage analytic standard errors for the state parameters in the count model are smaller than for the full likelihood methods. The bootstrap propagates the uncertainty for the detection function model through to stage two, but the bootstrap standard errors for the count model parameters are larger than those for the maximum likelihood method, and larger than the standard deviations of the marginal posterior distributions for the MH-update. We might expect the full likelihood methods to give better precision, as they use full information from the data for estimating all parameters.

The parameter of interest here is listed in Table 8.1 as *type buffer*. We see that the estimate is the same under all three approaches, although the parameter is estimated with higher precision under the full likelihood methods. The parameter estimate is highly significantly different from zero, and the positive value indicates that bird densities are higher along field margins when conservation buffers are present. We estimate that the increase in density is around 35 % ( $\exp(0.30) = 1.35$ ).

**Table 8.1** Maximum likelihood estimates (MLE) and analytic (ASE) and bootstrap (BSE, two-stage approach only) standard errors corresponding to detection function and count model parameters for the full-likelihood and two-stage approaches

	ML full likelihood		ML two-stage model			MH full likelihood		Fixed shape parameter
	MLE	ASE	MLE	ASE	BSE	Mean	SD	
Detection model parameters								
Scale GA	45.85	9.33	37.27	7.72	8.12	39.23	8.31	2
Scale IL	36.03	3.21	34.42	2.86	3.17	34.84	2.87	2.5
Scale IN	27.31	2.66	24.34	2.35	4.92	24.49	2.42	2
Scale KY	29.63	1.23	27.75	1.13	1.52	27.77	1.14	2.5
Scale MO	41.31	3.03	37.78	3.14	2.86	38.01	3.12	2
Scale MS	38.50	3.31	38.73	3.33	4.21	39.18	3.33	2
Scale OH	27.15	2.16	24.59	1.97	1.94	24.69	1.97	2
Scale SC	57.79	4.40	56.30	4.13	6.85	56.62	4.15	3
Scale TN	26.11	0.04	21.08	1.74	3.51	21.19	1.80	2
Count model parameters								
Random effect std dev	0.50	0.02	0.49	–	0.04	0.51	0.03	
Intercept	−0.99	0.28	−10.91	0.29	0.43	−0.63	0.39	
Type buffer	0.30	0.02	0.30	0.03	0.04	0.30	0.03	
Date	0.0053	0.0005	0.0046	0.0017	0.0018	0.0048	0.0017	
State IL	1.46	0.32	1.20	0.18	0.38	1.22	0.34	
State IN	1.50	0.32	1.34	0.18	0.49	1.39	0.34	
State KY	2.00	0.29	1.79	0.17	0.36	1.83	0.32	
State MO	0.53	0.29	0.32	0.16	0.36	0.35	0.33	
State MS	1.25	0.31	0.91	0.17	0.37	0.94	0.34	
State OH	1.11	0.30	0.92	0.17	0.37	0.95	0.34	
State SC	0.59	0.31	0.28	0.18	0.39	0.32	0.33	
State TN	2.14	0.28	2.12	0.17	0.44	2.17	0.34	

In addition, posterior mean and standard deviation (SD) of model parameters obtained using a Metropolis–Hastings algorithm (MH Update) are shown. States are Georgia (GA), Illinois (IL), Indiana (IN), Kentucky (KY), Missouri (MO), Mississippi (MS), Ohio (OH), South Carolina (SC) and Tennessee (TN)

### 8.5.2 Northern Bobwhite Surveys: A Bayesian Approach

The matched-pairs design to assess the effects of conservation buffers (Sect. 3.2.2.1) was also used to evaluate covey densities of northern bobwhite quail (*Colinus virginianus*) on crop fields with and without field buffers (Evans et al. 2013). Thirteen US states spanning nine ecological regions participated in the study. Sample points were visited in autumn of each year 2006–2008, and the locations of calling bobwhite coveys were marked on satellite images, so that distances from the points could later be measured. No attempt was made to estimate the size of each covey, so that the study compares covey densities, not bird densities.

We summarize here the Bayesian analysis of the bobwhite data reported by Oedekoven et al. (2014). Distances to detected coveys were measured without grouping, and so the methods of Sect. 8.2.4.1 were used. Covariates were considered

for inclusion in the detection function. Rather than specify a model for  $\mathcal{L}_z$ , this component of the likelihood was omitted (see Sect. 8.2.2).

Of the 13 states in the study, there were fewer than 50 detections of coveys in two, which were excluded from the analysis, leaving 11 states: Georgia, Iowa, Illinois, Indiana, Kentucky, Missouri, Mississippi, North Carolina, South Carolina, Tennessee and Texas. Within these states, 447 sites were visited 1–3 times in each survey year. A truncation distance of  $w = 500$  m was selected, resulting in 2545 detections from 2534 counts.

Covariates considered for inclusion in the detection function and count models were the factor covariates year (three levels: 2006, 2007, 2008), type (two levels: buffer or control plot), state (11 levels) and the continuous covariate date, which was centred around its mean for use in the count model. One random effect coefficient per site was included in the model, so that the likelihood had a very similar form to that of Eqn 8.33. (The same random effect coefficient applies to both points within a site, and to repeat counts at each point.)

Uniform priors were specified on all parameters, with large bounds selected for unbounded parameters to ensure that there was little impact on posterior distributions. Preliminary investigation of the distance data indicated that the hazard-rate detection function provided better fits than the half-normal. Hence, we included eight different hazard-rate models as choices for the probability density function of observed distances: one global (with no covariates) and seven multiple-covariate models.

The model for the expected count included an intercept and combinations of the four covariates, together with a random effect for site. There was no evidence of overdispersion, so that a Poisson model was assumed. RJMCMC (Sect. 8.4.2) was used to search model space.

In total, 100,000 iterations were carried out, with the first 10,000 being used as burn-in; model probabilities and summary statistics for parameters were extracted from the remaining 90,000.

A two-stage plot count model (Sect. 7.2) was fitted, for comparison. A site random effect was added to the model for comparability, and the bootstrap was used to estimate standard errors and confidence intervals. For each bootstrap resample, AIC was used to select the best model, thus allowing for model uncertainty.

In addition, we fitted a full likelihood model equivalent to the Bayesian model using maximum likelihood (ML) methods. Forwards stepwise model selection was carried out based on minimum AIC. Here, the first model consisted of the hazard-rate detection function without covariates and an intercept-only model with site random effects for counts. Covariates were added one at a time, first to the detection model, then to the count model.

For the Bayesian full likelihood approach, the hazard-rate detection function model with covariates year, type and state had a model probability of 1.00 to two decimal places. The same model was selected by AIC in 81 % of the bootstrap resamples for the two-stage approach. AIC forwards stepping also favoured it for the ML full likelihood model. For the count model, two models dominated the RJMCMC algorithm: the model with covariates type, date and state (model probability of 0.89) and the full model (year, type, date and state), with a model



probability of 0.11. For the two-stage model, the latter model was selected in 89 % of resamples, while the former was selected in 10 % of resamples. For the full likelihood model fitted with maximum likelihood methods, AIC forwards stepping favoured the same count model as for the Bayesian approach, *i.e.* with covariates type, date and state. Note, however, that fitting complex full likelihood models using maximum likelihood methods may be problematic in terms of calculating standard errors. For our preferred full likelihood model, we were unable to obtain standard error estimates for the shape parameter of the detection function and the date coefficient for the count model. Hence, we fixed these parameters at the maximum likelihood estimates obtained using the same algorithm.

**Table 8.2** Bobwhite quail estimates

Parameters	Bayesian full likelihood			Two-stage			ML full likelihood		
	Mean	SD	95 % CRI	Estimate	BSE	95 % CI	Estimate	ASE	95 % CI
<b>Det.model</b>									
Shape	3.30	0.16	(3.00, 3.63)	3.01	0.27	(2.68, 3.41)	3.32	—	—
Intercept	153.0	9.0	(135.5, 170.1)	138.6	16.1	(112.3, 163.8)	153.0	5.5	(142.2, 163.8)
Year 2006	0.06	0.03	(0.01, 0.12)	0.10	0.06	(−0.05, 0.14)	0.08	0.02	(0.03, 0.13)
Year 2007	−0.11	0.04	(−0.19, −0.05)	−0.15	0.05	(−0.25, −0.10)	−0.10	0.02	(−0.15, −0.06)
No buffer	0.13	0.04	(0.05, 0.20)	0.15	0.05	(0.05, 0.23)	0.14	0.02	(0.10, 0.18)
GA	0.38	0.07	(0.23, 0.52)	0.42	0.17	(0.05, 0.54)	0.40	0.08	(0.25, 0.55)
IA	0.17	0.10	(−0.01, 0.36)	0.21	0.16	(−0.12, 0.30)	0.16	0.09	(−0.02, 0.34)
IL	0.66	0.09	(0.48, 0.85)	0.70	0.17	(0.35, 0.76)	0.67	0.09	(0.50, 0.85)
IN	0.62	0.10	(0.43, 0.81)	0.66	0.14	(0.34, 0.72)	0.62	0.09	(0.45, 0.79)
KY	0.58	0.08	(0.43, 0.74)	0.64	0.12	(0.35, 0.68)	0.59	0.07	(0.44, 0.73)
MO	0.62	0.06	(0.51, 0.73)	0.69	0.09	(0.46, 0.71)	0.61	0.05	(0.51, 0.71)
MS	0.55	0.07	(0.41, 0.70)	0.61	0.10	(0.37, 0.64)	0.57	0.07	(0.44, 0.70)
NC	0.60	0.09	(0.43, 0.79)	0.66	0.12	(0.35, 0.70)	0.61	0.09	(0.44, 0.78)
SC	0.01	0.08	(−0.14, 0.16)	0.00	0.14	(−0.29, 0.12)	0.03	0.07	(−0.12, 0.17)
TN	0.44	0.10	(0.25, 0.63)	0.47	0.12	(0.19, 0.54)	0.44	0.04	(0.36, 0.51)
<b>Count model</b>									
Ran. eff. sd	0.82	0.05	(0.73, 0.91)	0.78	0.04	(0.69, 0.81)	0.29	0.01	(0.27, 0.31)
Intercept	−13.10	0.18	(−13.43, −12.73)	−13.23	0.33	(−13.91, −12.87)	−13.05	0.15	(−13.36, −12.75)
Year 2007	—	—	—	0.17	0.13	(−0.16, 0.37)	—	—	—
Year 2008	—	—	—	0.17	0.11	(−0.12, 0.31)	—	—	—
Type buffer	0.62	0.07	(0.48, 0.75)	0.63	0.12	(0.36, 0.71)	0.65	0.03	(0.59, 0.71)
Date	−0.013	0.002	(−0.018, −0.008)	−0.011	0.003	(−0.017, −0.005)	−0.012	—	—
IA	−0.81	0.29	(−1.38, −0.23)	−0.74	0.44	(−1.65, −0.24)	−0.82	0.25	(−1.32, −0.33)
IL	−0.59	0.27	(−1.12, −0.06)	−0.53	0.38	(−1.25, −0.07)	−0.57	0.22	(−1.00, −0.14)
IN	−1.24	0.27	(−1.79, −0.71)	−1.18	0.41	(−1.99, −0.70)	−1.23	0.22	(−1.67, −0.80)
KY	−0.47	0.25	(−0.98, 0.03)	−0.44	0.34	(−1.07, −0.02)	−0.46	0.17	(−0.79, −0.13)
MO	0.01	0.22	(−0.44, 0.42)	0.05	0.34	(−0.63, 0.46)	0.03	0.18	(−0.32, 0.37)
MS	−0.43	0.25	(−0.92, 0.05)	−0.37	0.34	(−1.04, 0.05)	−0.41	0.20	(−0.81, −0.01)
NC	−1.39	0.26	(−1.88, −0.87)	−1.31	0.36	(−1.99, −0.87)	−1.38	0.23	(−1.83, −0.94)
SC	0.01	0.27	(−0.53, 0.53)	0.08	0.42	(−0.76, 0.56)	0.02	0.22	(−0.41, 0.45)
TN	−1.10	0.28	(−1.65, −0.57)	−1.03	0.38	(−1.80, −0.60)	−1.10	0.15	(−1.40, −0.79)
TX	1.74	0.18	(1.33, 1.99)	1.46	0.29	(0.99, 1.81)	1.73	0.23	(1.27, 2.19)

Posterior mean and standard deviation (SD) and 95 % credible interval (CRI) for parameters of the best Bayesian full likelihood model; estimate, bootstrap standard error (BSE) and 95 % percentile confidence interval (CI) for parameters of the best two-stage model; and estimate, analytical standard errors and 95 % percentile CI for parameters of the best maximum likelihood (ML) full likelihood model. States are Georgia (GA), Iowa (IA), Illinois (IL), Indiana (IN), Kentucky (KY), Missouri (MO), Mississippi (MS), North Carolina (NC), South Carolina (SC), Tennessee (TN) and Texas (TX). State scale parameters in the detection function are relative to Texas, and state effects in the count model are relative to Georgia. Note that for the ML full likelihood model, parameter values for *Shape* and *Date* were fixed (see text for more details)

We show parameter estimates for the favoured model for each approach in Table 8.2. Generally, parameter estimates are similar from the three approaches, but the full likelihood models (Bayesian and maximum likelihood) tend to give higher precision than does the two-stage model.

The parameter of interest in these models was the coefficient for the level buffer of the type covariate in the count model. This was 0.62 (SD = 0.07), 0.63 (BSE = 0.12) and 0.65 (ASE = 0.03) for the Bayesian full likelihood, the two-stage and the ML full likelihood approach, respectively, indicating an increase in covey densities on treatment plots by 85 % ( $\exp(0.62) = 1.85$ ), 88 % ( $\exp(0.63) = 1.88$ ) or 92 % ( $\exp(0.65) = 1.92$ ) for the respective methods.

**Part III**  
**Distance Sampling Variations,**  
**Special Issues and Assumptions**

## Chapter 9

# Variations on a Theme

There are several variants of distance sampling, developed to extend the applicability of the methods to a wider range of populations. Several involve the use of traps or lures. The first use of traps in a distance sampling context was for trapping webs (Anderson et al. 1983), in which a series of points across the study region is selected, and a ‘web’ of perhaps 80 or more traps established around each point. In the case of trapping line transects, traps are distributed within a strip about the centreline (Lukacs et al. 2004). Both methods are very labour-intensive, and, while they depend on animal movement if animals are to encounter traps, that same movement generates bias. In trapping (or lure) point transects (Buckland et al. 2006), again a sample of points is selected, ideally a grid of points spanning the entire study area, and a single trap (or lure) is placed at each point for a fixed period. The number of animals trapped (or lured) in that period is recorded. Thus the number of traps required is appreciably less than for trapping webs or trapping line transects, and animal movement does not bias the method. However, as the distances animals have moved to reach the trap are unknown, separate trials are required on a sample of known-location animals to allow estimation of the detection function. Trapping point transects are intended for when trap separation is sufficient that at most a single trap is likely to be accessible to any given animal during the period that the traps are set. If trap separation is substantially less, so that many animals are caught at more than one location, then spatially-explicit capture–recapture methods (Efford 2004; Borchers and Efford 2008) may be used. In these, information on movement of marked animals allows estimation of mean size of home range, from which animal density can be estimated. Alternatively, if presence of animals is inferred by listening for calls, and the same call can be heard from more than one point (corresponding to ‘recapture’), then those data can be analysed using spatially-explicit capture–recapture methods (Borchers 2012).

Indirect surveys are surveys of animal sign, usually dung or, in the case of apes, nests. Standard line transect methods are used to estimate the density of sign. To convert this to animal density, we need to estimate the deposition rate — the mean

number of signs generated per animal per day. For the ‘standing crop’ method, we also need to estimate decay rate — the reciprocal of the mean time to decay of sign. Alternatively, the sample plots should be cleared of sign at the outset (the ‘clearance plot’ method), and the plots surveyed before any newly-deposited sign decays.

Indirect surveys typically involve long-lasting sign. Cue counting is designed for short-lived signs or ‘cues’ of animal presence, such as whale blows or bird songbursts. The mean cue rate per unit area per unit time is converted to an estimate of animal density by separately estimating a cue rate, which has a close parallel with deposition rate for indirect methods, and is the mean number of cues per animal per unit of time.

Three-dimensional distance sampling surveys are starting to see some use. Typically, we cannot assume that animals are distributed independently of the line or point in the vertical dimension, so that distribution of animals vertically must be modelled. Applications include shipboard sonar surveys of fish (Cox et al. 2011) and radar surveys of migrating birds.

Acoustic surveys have mostly been used for cetaceans, and for monitoring relative abundance of bats, but they also have potential for other sound-producing taxa like insects, amphibians, fish, birds and primates (Marques et al. 2013). We do not have a separate section for them because they can potentially be analysed by several standard methods:

1. Conventional line transect sampling (Sects. 5.2.2 and 6.3.1), for example using an acoustic array towed by a ship for marine surveys (*e.g.* Barlow and Taylor 2005). A means of estimating distance from the line is required, which may be achieved by triangulation.
2. Conventional point transect sampling (Sects. 5.2.3 and 6.3.2). Distances of acoustic cues from the sensor must be estimated, which requires either an array to allow triangulation or a means of estimating distance from a single sensor (Marques et al. 2011). As sensors typically remain in place for some time, it is likely that animal movement will need to be modelled.
3. Trapping point transect sampling (Sect. 9.2). This method requires trials carried out on a sample of known-location animals, so that the detection function can be estimated using logistic regression (Marques et al. 2009).
4. Cue counting (Sect. 9.4.2). As with conventional point transect sampling, this method also needs estimates of distances of cues from sensors (*e.g.* Buckland 2006). However, animal movement independent of the sensor does not bias this method.
5. Spatially-explicit capture–recapture methods (Borchers 2012). For this method to be useful, cues should be potentially detectable from more than one point (sensor), and it should be possible to identify as a ‘multiple capture’ a cue that is detected from multiple points (*e.g.* Martin et al. 2012). The exact times of arrival of cues at sensors are useful both for identifying those cues detected from multiple points and for estimating the location of animals that give such cues.

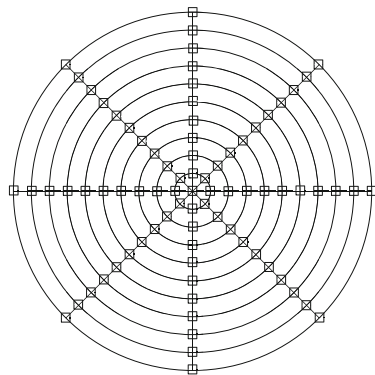
## 9.1 Trapping Webs and Trapping Line Transects

The concept underlying the trapping web is that, instead of an observer actively searching for animals, traps are placed throughout a circular plot, and animals select their own distance from the point at the plot's centre, by entering a trap. To ensure that probability of detection (capture) is at or close to one in the vicinity of the point (so that we can assume  $g(0) = 1$ ), trap density is high in the vicinity of the point, and decreases with increasing distance from the point. An example of the layout of a trapping web is shown in Fig. 9.1. In the case of a trapping line transect, traps are most dense near the line, and reduce in density with distance from the line.

For each captured animal, the distance of the trap that caught it from the centre point or line is recorded, and the data are analysed as for standard point transect sampling (trapping web) or line transect sampling (trapping line transect). We do not cover the methods in detail here, because of their limitations. Firstly, to draw inference on a population occupying a wider area, a large number of points or lines should be surveyed; for trapping webs, Lukacs et al. (2004) recommend at least 15, and say that '25 is better'. They also advise at least 90 traps in each web, and so considerable resources are needed if the method is to be implemented effectively.

The second serious limitation of the methods is that they are dependent on sufficient animal movement during the period that the traps are set to ensure that almost all animals close to the point or line are detected, yet that movement biases animal density estimates. For example, for trapping webs, some animals initially off the web will move onto it, and encounter a trap, leading to a recorded distance that is biased low. If the data are plotted in a histogram, the excess of detections at the rim of the web is typically evident. This bias can be addressed by truncating outer rings of traps, but this means that a substantial proportion of the data is discarded, and the corresponding effort 'wasted'. However, the bias is not limited to the outer rings of traps. Consider an animal that is initially on the web. If it wanders around, then it is more likely to encounter a trap if it moves towards the centre, where trap density is high, than if it moves away from the centre, where trap density is low. Thus there

**Fig. 9.1** An example of a trapping web with eight lines, each of ten traps, together with a single central trap, giving 81 traps in total. If for example the concentric rings of traps are at radii that are multiples of 5 m, then animals caught in the centre trap are recorded as in the distance interval 0–2.5 m, those caught in the first ring as in the distance interval 2.5–7.5 m, and so on



is a tendency to record distances that are closer to the web centre than would be the mean location of animals initially on the web. This bias towards shorter distances results in positive bias in density estimates.

Bias is especially large if animals are not restricted to home ranges. If they are, and if the home range size does not vary too much among animals, then bias from movement can be modest given a good choice of trap separation across the web. Lukacs (2002) describes simulation software WebSim to help the user design a survey to give low bias.

Bias can also arise if a trap becomes unavailable once it captures an animal, as animals near the centre point or line might not be caught because the traps in their home ranges are unavailable.

## 9.2 Trapping and Lure Point Transects

The main survey for trapping and lure point transects comprises a set of points through the study area, ideally on a randomly-positioned grid. At each point, a trap or lure is set. For animals subsequently recorded in the trap or at the lure, we typically do not know their initial location, and so cannot estimate the effective area that has been sampled around the point. To address this, trials are established in which the location of a sample of animals is determined, and a trap or lure set at a predetermined distance from each animal. This provides a series of binary trials, leading to success (animal is successfully trapped or lured) or failure (animal is not recorded in the trap or at the lure). These binary data may be modelled using logistic regression, for which one of the covariates is initial distance from the trap or lure.

The concept of estimating the effective area sampled from trials on animals at known locations was used by Turchin and Odendaal (1996) for monitoring population density of the southern pine beetle (*Dendroctonus frontalis*), by Legare et al. (1999) to estimate density of black rails (*Laterallus jamaicensis*), by Acosta and Perry (2000) for sampling crayfish (*Procambarus alleni*) populations, by Mills et al. (2001) for estimating the size of spotted hyaena (*Crocuta crocuta*) populations, and by Okot Omoya et al. (2014) to estimate the population sizes of lions (*Panthera leo*) and spotted hyaenas in Uganda's savannah national parks. The terms trapping point transects and lure point transects were coined by Buckland et al. (2006), who developed their methods so that the population size of Scottish crossbills (*Loxia scotica*) could be estimated (Summers and Buckland 2011). Grids of sample points were established in suitable habitat throughout the species' range, and excitement calls were played at each point to lure birds in. Separately, trials were set up, in which birds were located, and the calls played at a pre-determined distance from them. We base our lure point transect example on this study.

Potts et al. (2012) used trapping point transects to estimate the size of the Key Largo woodrat (*Neotoma floridana smalli*) population. In that study, woodrats were radio-tagged to enable trials to be set up, and multiple trials were conducted on each radio-tagged animal to generate the data from which the effective area sampled

could be estimated. Random effects corresponding to animals were included in the logistic regression model, to allow for the non-independence inherent in carrying out repeat trials on the same animal. This provides our example of trapping point transect sampling.

A variation on the lure point transect was used by Baccaro and Ferraz (2012) to estimate density of ant nests. They placed lures at a grid of points to attract ants. They later returned to these lures, and for those that had attracted ants, they followed the trail of ants back to the entrance to their nest. Thus for each nest detected, it was possible to measure its distance from the lure, so no trials were needed. The only significant problem for the method is that nests may have more than one entrance, in which case the nearer entrance to the lure is the more likely to be recorded, resulting in downward bias in distance estimates and upward bias in density estimates.

### ***9.2.1 Lure Point Transect Case Study: Scottish Crossbills***

This case study reports the study of Summers and Buckland (2011), in which the lure point transect method was used to estimate the population size of Britain's only endemic bird species, the Scottish crossbill. The computer code for analysing the data is available on the book website. Traditional distance sampling methods do not work well with crossbills. Birds are often recorded in flight. Inclusion of such detections may cause substantial upward bias in standard line and point transect sampling estimates (Sect. 10.2.1); exclusion leads to downward bias unless a correction is made for proportion of time in flight (Sect. 10.2.1). Further, crossbills often remain quiet in the tops of trees, so they are overlooked — that is, detection at the line or point is uncertain, contrary to the assumptions of standard point transect sampling (Sect. 10.1.3). This does not bias the lure point transect method because the use of trials for which we know the location of both detected and undetected birds allows estimation of the detection function without assuming certain detection at the point.

Crossbills respond to playbacks of an excitement call by flying to the lure and calling back. This allows counts to be made at each survey point, and the calls are recorded to confirm the species of crossbill (Fig. 9.2), as three similar species occur in Scotland.

#### **9.2.1.1 The Trials**

Trials were conducted at a number of sites in northern Scotland during 2002–2007 prior to the main survey. One observer located birds, while a second played the lure at pre-determined distances, and recorded whether the birds were lured in. If two or more birds were present, and at least one, but not all, responded, the birds were analysed as if they were two groups, one of which responded, and one of which did not.





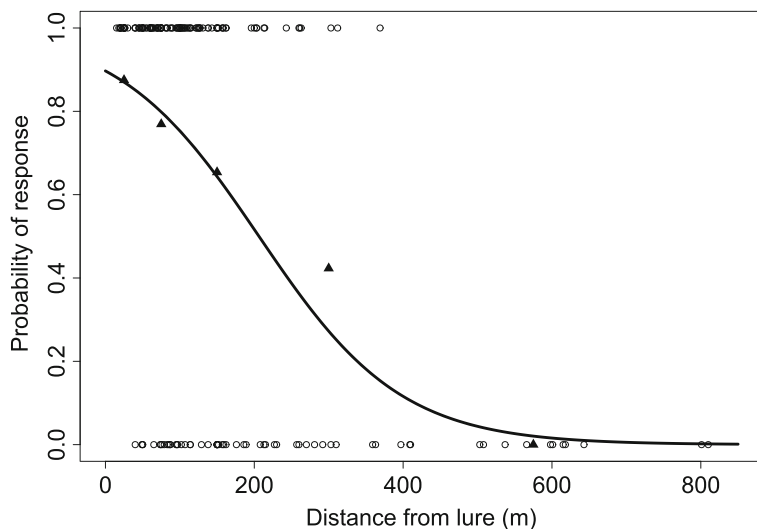
**Fig. 9.2** An observer playing excitement calls to lure crossbills to the point, and ready to record calls of any that respond, to allow identification of species. Photo: Ron Summers

Several detection function models were fitted to the data from 175 trials. Covariates considered for inclusion in addition to distance were: species of crossbill, date, time of day, size of flock, behaviour and habitat. AIC resulted in selecting the model in which probability of a response is a function of distance from the point alone. The fitted model is illustrated in Fig. 9.3. A  $\chi^2$  goodness-of-fit test indicated a good fit of the detection function model to the data:  $\chi^2_3 = 3.001$ . The effective radius of detection over which crossbills were lured was 265.3 m, corresponding to an effective area per point of 22.1 ha.

### 9.2.1.2 The Survey

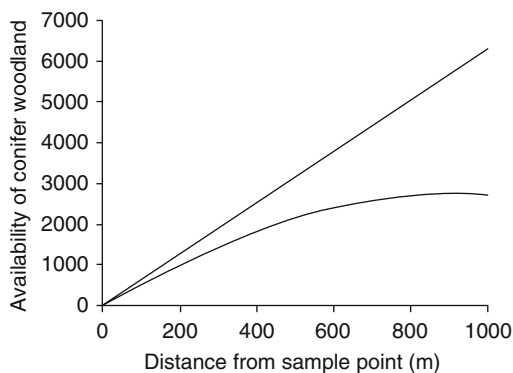
The survey was carried out in conifer woodland in mainland Scotland north of  $56^\circ 50'N$ , comprising almost all of the range of the Scottish crossbill. A systematic sample of 889 survey points was selected. When visited, 34 survey points were re-classified as 'non-conifer' because there were no conifers within 50 m of the survey point. The remaining points were surveyed during January to April 2008.

To model the woodland availability function, additional points were selected at 10 m intervals from each survey point along the eight main compass bearings to a distance of 1 km. Each point was classed as woodland or non-woodland, to give estimates of the proportion of woodland at increasing distances from the



**Fig. 9.3** Plot of response (1 = birds responded; 0 = no response) against distance of birds from the point for 175 trials. Also shown is the estimated probability of a response, as a function of distance from the point. The mean response is shown by a *solid triangle*, plotted at the mean distance of responses from the point, for each of the following distance intervals: 0–50 m; 50–100 m; 100–200 m; 200–400 m; 400–750 m

**Fig. 9.4** Modelled availability of conifer woodland as a function of distance from the point, out to 1 km. The *straight line* shows the availability that would occur if woodland cover was complete, whilst the *curved line* shows actual availability



survey point. These proportions were modelled using logistic regression, giving the availability function shown in Fig. 9.4.

In total, 442 Scottish crossbills were recorded, with at least one bird recorded at 144 of the 852 survey points. In principle, it is now simple to estimate abundance: an effective area of 22.1 ha is estimated to have been surveyed around each point, and so 144 birds were recorded in an effective area surveyed of  $852 \times 22.1$  ha, or roughly  $188.3 \text{ km}^2$ , equating to a density of  $442/188.3 = 2.3474 \text{ birds km}^{-2}$ . With a total area of conifer woodland estimated at  $3505.8 \text{ km}^2$ , this gives an abundance estimate of around 8,230 birds. However, care is needed, as the total woodland area

of course excludes habitat that is not conifer woodland, whereas many of the survey plots of radius 1 km include such habitat. Using the availability function of Fig. 9.4 to correct the effective area surveyed around each point to effective area of conifer woodland surveyed around each point, we obtained a population abundance estimate of 12,900 birds, with bootstrap confidence interval (7720, 21,300). Summers and Buckland (2011) report further corrections, to assign proportionally unidentified birds to species, to remove early-fledging juveniles from the estimate, to correct for birds flying over plots that were not in reality responding to the lure, and to account for incubating females, and their final abundance estimate was 13,600 birds with bootstrap confidence interval (8130, 22,700).

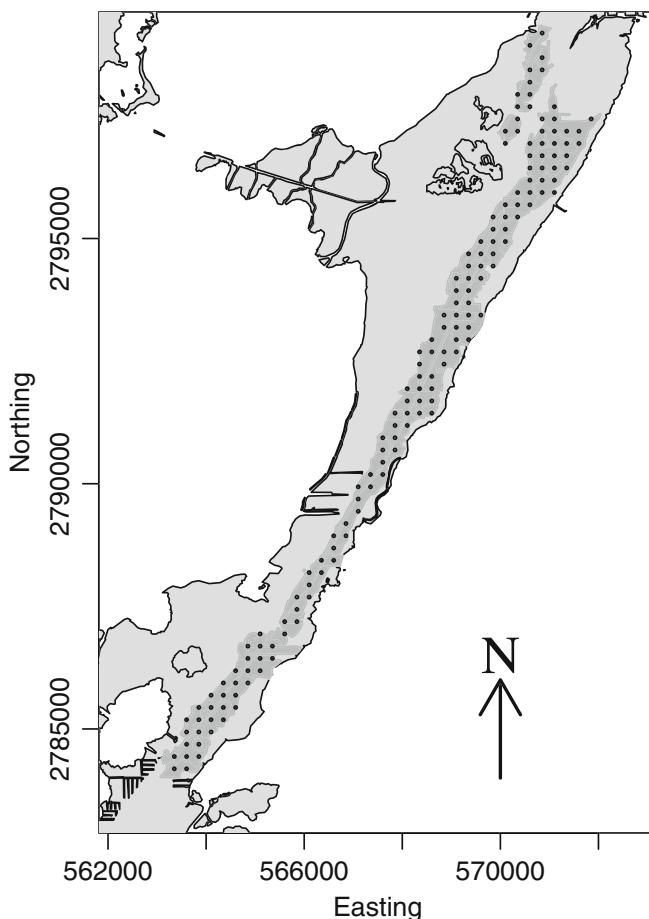
### 9.2.2 *Trapping Point Transect Case Study:* *Key Largo Woodrats*

The endangered Key Largo woodrat is one of six recognised subspecies of the eastern woodrat (*Neotoma floridana*), and is restricted solely to the tropical hardwood hammock on northern Key Largo, Florida. A reliable population abundance estimate, together with an associated measure of uncertainty, is essential to evaluate management decisions for the woodrat population. Since woodrats are nocturnal, cryptic, and sparsely distributed, standard population monitoring methods do not work well. It was decided therefore to try the trapping point transect method (Potts et al. 2012).

Key Largo covers approximately 21 km<sup>2</sup>. Although tropical hardwood hammock was once the dominant vegetation type across Key Largo, due to urbanization, approximately one third remains within two protected reserves. Despite its protected status, the remaining tropical hardwood hammock is highly fragmented with roads, tracks and abandoned developments.

Woodrats typically occupy home ranges of approximately 0.23 ha, forage at night, and rest in one of several nests within their home range during the day.

Trapping point transect surveys were undertaken during the same season (February to April) in each of 4 years, 2008–2011. The main survey comprised a randomly-placed systematic grid of 136 sample points with a 250 m trap spacing throughout suitable habitat within the range of the subspecies (Fig. 9.5). The large trap spacing was used so that detection events at neighbouring sample points could be assumed independent. The distance was chosen based on the home range size of the woodrat. In 2011, the main survey was repeated three times within the 10-week field season. The original 136 points were surveyed and then traps were moved 125 m east and the survey was repeated (142 points), and then traps were moved 125 m north and the survey was repeated again (142 points). In total, 420 main survey points were trapped in 2011. Two Sherman traps were placed within a metre of one another at each trap location. Each trap location was trapped for three consecutive nights. All woodrats captured in the 3-day period were marked.



**Fig. 9.5** Map showing the main survey design, in which 136 sample points were distributed throughout the suitable habitat (*shaded dark grey*) based on a randomly-placed systematic grid with a 250 m trap spacing (*black dots*). Traps were not placed in unsuitable habitats of water (*white*) and mangrove swamp (*light grey*). Reproduced from Potts et al. (2012). ©British Ecological Society

For the trials, radio-transmitters were attached to a subset of woodrats captured during the main survey and during additional dedicated trapping. Since the woodrat is nocturnal, radio-collared individuals were located during the daytime at their nests. An individual trial was conducted on a radio-collared woodrat by setting two trial traps (as in the main survey) at some pre-determined distance and random direction away from the nest. The trial traps were checked the following morning and captured animals were released at point of capture. The trial survey recommenced immediately if the woodrat was not recaptured, or after a rest period of typically 2 days otherwise.

Each radio-tagged woodrat was exposed to trial surveys at various distances. Typically, three trials were performed at each 5 m interval between 5 and 60 m. Some trials were performed at very long distances for which capture was believed to be extremely unlikely (up to 320 m), to ensure that the tail of the detection function would be accurately estimated. Some woodrats were exposed to repeated, very short trial distances (1 m), in order to retrieve the collar.

### 9.2.2.1 Data Analysis

The detection function was estimated from trial data by carrying out a logistic regression with random effects — an example of a generalized linear mixed model. The random effect models the variability among animals, which allows use of multiple trials on the same animal, without having to assume that all trials are independent. (It does however assume that animals do not become trap-shy or trap-happy, having been caught once.) Separate detection functions were fitted to the male and female data. Due to small sample sizes collected during the trial survey on male woodrats, data were pooled across year and a simple detection function was fitted that assumed average probability of detection depended only on distance from trap and a random effect about the intercept term. Covariates considered for inclusion in the detection function for females were distance from trap and year of survey. A simple random effect about the intercept term was assumed and distance was centred by subtracting the mean. Forward model selection based on AIC was used.

The models were fitted using the `glmer` function in the `lme4` package of R. Observed values of each explanatory variable were substituted into the fitted detection function for each woodrat captured in the main survey, and both distance  $r$  and the animal random effect were integrated out (Potts et al. 2012).

For trapping point transects, we cannot truncate the distance data because we do not know which animals trapped in the main survey were originally beyond the truncation distance. A consequence of this is that the method can have substantial bias due to edge effects (Sect. 2.1). To remove the bias, instead of assuming that distances of animals from points are distributed according to a triangular distribution, we estimated the availability function  $\pi(r)$  (Sect. 5.2.3), by taking 10 points evenly distributed between 0 and a distance at which probability of detection can be assumed to be zero, laid along each of the four cardinal (N, S, E, W) and four intercardinal (NE, SE, SW, NW) directions of the compass. Thus 80 surrounding points were checked for each main survey trap location, to determine whether or not they fell in woodrat habitat. A binary logistic regression model was then fitted to define habitat availability as a function of distance from the main survey trap location.

Estimates of woodrat abundance were calculated by sex and year, using a Horvitz–Thompson-like estimator (Sect. 6.4). Variance in the estimated abundance has two sources: variance associated with the encounter rate (*i.e.* animals caught per trap) in the main survey, and variance associated with estimating the detection

function in the trial survey. We used a nonparametric bootstrap as follows. First, generate a random resample with replacement of the trap locations in the main survey. Second, generate a random resample with replacement of the individuals in the trial survey. Fit a new detection function model to the resampled trials data, repeating the model selection procedure used in the original analysis. Third, use the new detection function in conjunction with the new main survey data to generate a bootstrap abundance estimate. These three steps were repeated a large number of times, and the sample variance of the bootstrap estimates of abundance taken as an estimate of the variance of the abundance estimate. Approximate confidence intervals for abundance were obtained using the percentile method (Buckland 1984).

9.2.2.2 Results

The capture rate of female and male woodrats (per 100 trap nights) declined substantially between 2008 and 2010, and then increased slightly in 2011 (Table 9.1). In 2010, only a single woodrat of each gender was caught in the 816 trap-nights of effort in the main survey (Table 9.1).

In total, 1190 individual trials were conducted on 34 female (645 trials) and 22 male (545 trials) woodrats between 2008 and 2011. Of these trials, 176 (66 trials on males and 110 on females) were successful (Table 9.2). The maximum trial distance for which a radio-collared woodrat was recaptured was 80 and 60m for a female and male woodrat, respectively (Table 9.2).

For females, the detection function model with lowest AIC was that with trial distance and an individual-level random effect. The model with the next-lowest AIC included a year effect, but gave an appreciably worse AIC value ( $\Delta AIC = 4.07$ ), hence abundance estimates were based only on the model that included trial distance but not year. For males, the detection function was modelled as a function of trial distance and a random individual-level effect. The detection functions fitted to the female and male trial data are shown in Fig.9.6. The availability function took account of edge effects, and so increased less quickly than a triangular availability distribution.

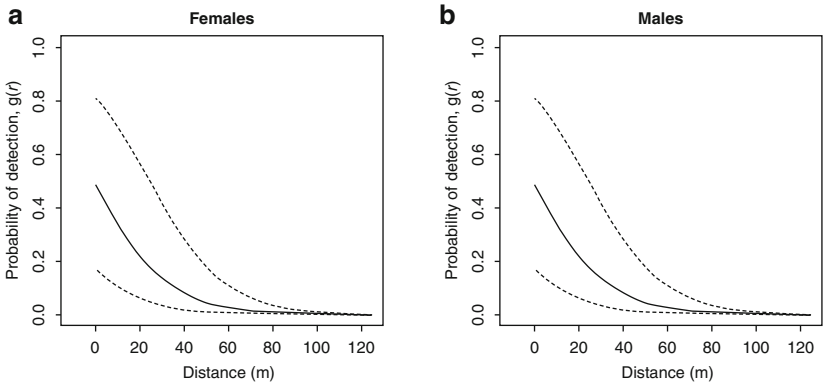
**Table 9.1** The total number of unique female and male woodrats captured, and the capture rate per 100 trap nights (given in parentheses), in the main survey

Year	Female	Male	No. pts
2008	6 (4.4)	8 (9.5)	136
2009	4 (2.9)	2 (1.5)	136
2010	1 (0.7)	1 (0.7)	136
2011	6 (2.4)	7 (2.6)	420

Also shown is the number of main survey points (No. pts). Two traps were placed at each main survey point, and set for three consecutive nights (e.g. 136 main survey points equated to 816 trap nights). Main survey effort in 2011 was trebled

**Table 9.2** The number of individuals for which radio collars were attached and trial surveys performed (Individuals), along with the total number of trials (Trials), the percentage of trials that were successful (Success), and the maximum distance of a successful trial (Max dist) for female and male woodrats in the 4-year survey period

Year	Females				Males			
	Individuals	Trials	Success (%)	Max dist (m)	Individuals	Trials	Success (%)	Max dist (m)
2008	11	119	23	41	4	69	22	30
2009	7	169	12	40	5	148	7	60
2010	5	117	18	80	5	120	17	36
2011	11	240	18	64	8	208	8	30



**Fig. 9.6** Mean detection function plot (*solid line*) corresponding to the best model for the female (**a**) and male (**b**) trial data. *Dashed lines* indicate the 2.5 and 97.5 percentiles (as calculated based on 10,000 samples of the estimated random effects distribution). Reproduced from Potts et al. (2012). ©British Ecological Society

**Table 9.3** Abundance estimates, coefficients of variation and 95 % percentile intervals for female, male and total population of woodrats between 2008 and 2011

Year	Female			Male			Total		
	$\hat{N}_F$	CV	CI	$\hat{N}_M$	CV	CI	$\hat{N}_T$	CV	CI
2008	235	0.65	(63, 623)	458	0.59	(83, 1164)	693	0.45	(253, 1462)
2009	177	0.67	(28, 422)	71	0.84	(0, 201)	248	0.55	(60, 538)
2010	43	1.19	(0, 169)	35	1.11	(0, 139)	78	0.82	(0, 233)
2011	130	0.65	(29, 341)	126	0.52	(31, 273)	256	0.43	(94, 518)

Estimated abundance for both male and female woodrats declined between 2008 and 2010, but increased again in 2011 (Table 9.3). However, uncertainty for sex-specific abundance estimates was large, especially when capture rates were low (*e.g.* in 2010 when only one male and one female woodrat were captured, coefficients of variation for female and male woodrats were 1.19 and 1.11, respectively, Table 9.3).

Coefficients of variation for the abundance estimates of the entire population ranged between 0.43 in 2011 when substantial survey effort was undertaken, to 0.82 in 2010 when capture rates were especially low (Table 9.3).

### 9.3 Indirect Surveys

Indirect surveys are surveys of sign, usually dung or nests, produced by the animal of interest (Buckland et al. 2001, pp. 183–189). Usually, line transect sampling is preferred to point transect sampling. Standard line transect methods give estimated density of sign, and it is conversion of this to estimates of animal density or abundance that makes the approach different from standard methods. To convert sign density to animal density, there are two basic approaches. The first is termed the ‘clearance plot’ method, for which an estimate of the daily production of sign per animal is required, and the second is the ‘standing crop’ method, which also needs an estimate of the mean time to decay of sign.

Advantages of indirect sampling over direct sampling are as follows.

- In dense habitat, animals often move away from the observer without being detected, causing downward bias in abundance estimates. Signs do not move, and so this bias is avoided.
- In surveys of apes for example, very few animals would be encountered in line transect surveys, so that reliable estimation of density, if possible at all, requires substantial survey effort. The nests they produce are far more numerous, and these can be adequately surveyed with appreciably less effort.

Disadvantages are:

- The daily production of sign must be estimated, and it can be difficult to ensure that the animals monitored are representative of the population to be estimated.
- Either plots must be cleared ahead of the main survey (clearance plot method) or an experiment is needed to estimate sign decay rate (standing crop method).

Most examples of indirect distance sampling surveys refer to estimating abundance of apes through surveying their nests (*e.g.* Plumptre and Reynolds 1996, 1997; Kouakou et al. 2009), or of ungulates through surveying their dung piles (*e.g.* Barnes et al. 1995; Barnes 2001; Marques et al. 2001).

In line transect surveys of dung, distances from the line are typically small, so that field procedures are needed to ensure distances are accurately measured. The line should be marked, *e.g.* by one person pulling a rope or chain, while a second person searches for dung and measures distances. Measurements should be to the centre of the dung, not to the nearest part of the dung pile. If a single defecation can result in several pellets or clumps of dung, clear instructions should be developed for determining the centre of the dung to which measurements should be made. In the absence of clearly defined field procedures and adequate observer training, observers tend to record zero distance for dung piles close to, or intersecting, the line, making subsequent analysis difficult and unreliable.



In line transect surveys of ape nests, several nests often occur in close proximity. If it is not easy to determine the centre of the group of nests, or to ensure that all nests in the group are detected, then it is better to record each nest individually, together with its distance from the line. Although this may represent a violation of the assumption that detections are independent events, the standard methods are very robust to failures of this assumption.

### 9.3.1 Clearance Plot Method

The clearance plot method is for when it is practical to clear *all* sign from the strips to be surveyed. In the case of nests, this may involve marking nests, so that they can be excluded from the subsequent survey. The survey follows the clearance of plots, allowing sufficient time for new signs to accumulate but not for them to decay. The main survey can be conducted using line transect sampling, although the need to clear strips in advance suggests using relatively narrow strips, for which it may be practical to conduct a strip transect survey (Sect. 6.2.2.1), counting all new sign appearing on the strips. Sign density  $R$  is estimated using standard line or strip transect methods, giving  $\hat{R} = n/(2\hat{\mu}L)$  for line transect sampling (Eq. (6.13)), and  $\hat{R} = n/(2wL)$  for strip transect sampling (Eq. (1.1)). To convert sign density to animal density, we need to estimate a single additional quantity: the mean daily production of sign per animal. This daily production rate  $\eta$  is typically estimated using observational studies. If the observed production (number of signs produced per day) is recorded for each of a representative sample of animals, the sample mean can be taken as the estimate  $\hat{\eta}$  of  $\eta$ , and the sample standard error gives  $se(\hat{\eta})$ .

The mean production rate can be estimated by placing captive animals in an enclosure for a period that is too short for sign to decay, and count the number of sign at the end of the period. This number is divided by the number of animal-days (number of animals multiplied by the number of days in the enclosure), to estimate the mean production per day per animal. There are two problems with this approach: first, captive animals may not have typical production rates; and second, we do not have estimates for individual animals (unless they are individually penned), so that we cannot calculate the standard error of our estimate.

Given estimates  $\hat{R}$  and  $\hat{\eta}$ , we estimate animal density as

$$\hat{D} = \frac{\hat{R}}{d\hat{\eta}} \quad (9.1)$$

where  $d$  is the number of days between plot clearance and the survey. The corresponding standard error is

$$se(\hat{D}) = \hat{D} \sqrt{\left\{ \frac{[se(\hat{R})]^2}{\hat{R}^2} + \frac{[se(\hat{\eta})]^2}{\hat{\eta}^2} \right\}}. \quad (9.2)$$

### 9.3.2 The Standing Crop Method

The standing crop method is appropriate for when it is impractical to clear plots. In this case, both the daily production of sign per animal and the mean time to decay of sign must be estimated. To estimate the latter, it is critical that a clear definition of when sign has ‘decayed’ is used, and that this definition is applied in exactly the same way in the experiment to estimate decay rate as in the line transect survey. For dung for example, if dung is marked and revisited later to determine whether it has decayed, it may be possible to see dung, but if it has decayed to the point that it would not have been recorded as dung in the main survey, then it should be recorded as having decayed in the experiment.

The decay rate of interest is the one that applies in the lead-up to the main survey (Laing et al. 2003). This is more easily estimated using a ‘retrospective’ decay experiment, in which sign is located and marked over a period ahead of the line transect survey and revisited at the time of the survey to determine whether it has decayed or not, than using a ‘prospective’ experiment, where fresh sign is identified, then monitored until it decays. If decay rate did not vary by date, either method would work, but if decay rates are very variable, and the prospective method is used, then modelling may be required to convert to the retrospective decay rate needed. The decay rate has often been assumed to follow an exponential model, but this has been shown to give substantial bias (Barnes and Barnes 1992).

We denote the mean time (in days) to decay by  $\xi$ . We will show how  $\xi$  may be estimated using the retrospective method in the red deer case study below. Given an estimate  $\hat{\xi}$ , we now estimate animal density by

$$\hat{D} = \frac{\hat{R}}{\hat{\xi}\hat{\eta}} \quad (9.3)$$

with

$$se(\hat{D}) = \hat{D} \sqrt{\left\{ \frac{[se(\hat{R})]^2}{\hat{R}^2} + \frac{[se(\hat{\xi})]^2}{\hat{\xi}^2} + \frac{[se(\hat{\eta})]^2}{\hat{\eta}^2} \right\}}. \quad (9.4)$$

Note that Eq. (9.3) can be obtained in two steps. We can first estimate the number of sign produced per day per unit area by

$$\hat{S} = \frac{\hat{R}}{\hat{\xi}}. \quad (9.5)$$

We now estimate the number of animals that produce  $\hat{S}$  signs per unit area by

$$\hat{D} = \frac{\hat{S}}{\hat{\eta}} = \frac{\hat{R}}{\hat{\xi}\hat{\eta}}. \quad (9.6)$$

**Fig. 9.7** The size of the red deer population in Abernethy Forest was estimated in 2001 using a dung survey. Photo: Steve Buckland



### 9.3.3 Red Deer Dung Decay Rate Experiment

We use a decay rate experiment conducted at Abernethy Forest in Scotland as a case study (Laing et al. 2003). Pellet groups of both red deer (*Cervus elaphus*, Fig. 9.7) and roe deer (*Capreolus capreolus*) were monitored in 2001, and we use the red deer data to illustrate the retrospective method of estimating mean time to decay.

Fifteen plots were cleared of dung in January 2000, and regular visits were made to the plots over subsequent months, and any new pellet groups appearing on the plots were marked. In total, 54 red deer pellet groups were recorded. The date of deposition was estimated as the mid-point between the date of the visit at which the pellet group was marked and the date of the previous visit. A pellet group was judged to have decayed when fewer than six pellets were identifiable. The same criterion was adopted for a line transect survey of the site, conducted in May 2001. Pellet groups that had decayed by May 15th 2001 were recorded as 0, and those that had not decayed by that date as 1. Three forms of model were considered for these data. Model 1 was a simple logistic model, left-truncated at zero:

$$E(y_i|x_i) = p(x_i) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x_i)} \quad (9.7)$$

where  $y_i$  is 0 if pellet group  $i$  had decayed, and 1 otherwise,  $x_i$  is the age of the pellet group on May 15th 2001,  $p(x_i)$  is the probability that a pellet group of age  $x_i$  has not yet decayed, and  $\beta_0$  and  $\beta_1$  are parameters to be estimated. This model

is unsatisfactory in the sense that probability that the pellet group has not decayed should be one at age zero, but  $p(0) \neq 1$ . Model 2 ensures  $p(0) = 1$ :

$$E(y_i|x_i) = p(x_i) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 \log x_i)} . \tag{9.8}$$

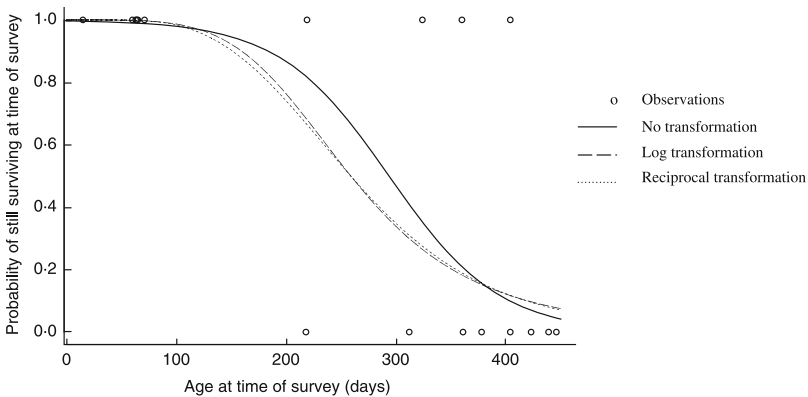
However, the log transformation can cause the model to have a thick upper tail; if there are insufficient large values of  $x_i$  in the data, model 2 can give positively biased estimates of the mean time to decay. Model 3 ensures  $p(0) = 1$ , but has similar tail behaviour to Model 1:

$$E(y_i|x_i) = p(x_i) = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x_i - \beta_2/x_i)} \tag{9.9}$$

where  $\beta_2$  is an additional parameter to be estimated.

For each model, habitat (a factor-type covariate with three categories) was considered for inclusion, but was not found to be significant.  $\Delta AIC$  values for the covariate models ranged from 1.35 to 3.02.

We show the three fitted models in Fig. 9.8. In this case, the estimate  $\hat{p}(0)$  of  $p(0)$  under Model 1 is very nearly one, so that we might expect Model 1 to be satisfactory. Estimates of mean time to decay are given in Table 9.4. We see that Model 2 has the smallest AIC, while that for Model 1 is not much greater. The model-averaged estimate of the mean time to decay, using AIC weights (Buckland et al. 1997), is 284 days.



**Fig. 9.8** The logistic regression curves fitted to the red deer pellet group data. The *open circles* show the observed data, which are 1 for pellet groups surviving to 15 May 2001 and 0 otherwise. Reproduced from Laing et al. (2003). ©British Ecological Society

**Table 9.4** Estimated mean time to decay of red deer pellets, with corresponding standard error and confidence interval, under each model

Model	Mean time (days)	Std error	Log-normal 95 % CI	$\Delta$ AIC	AIC wts
1	295	31	(240, 362)	0.70	0.33
2	280	39	(213, 369)	0.00	0.47
3	275	42	(204, 371)	1.77	0.20

Also shown are  $\Delta$ AIC values and AIC weights

9.3.4 Chimpanzee Nest Survey

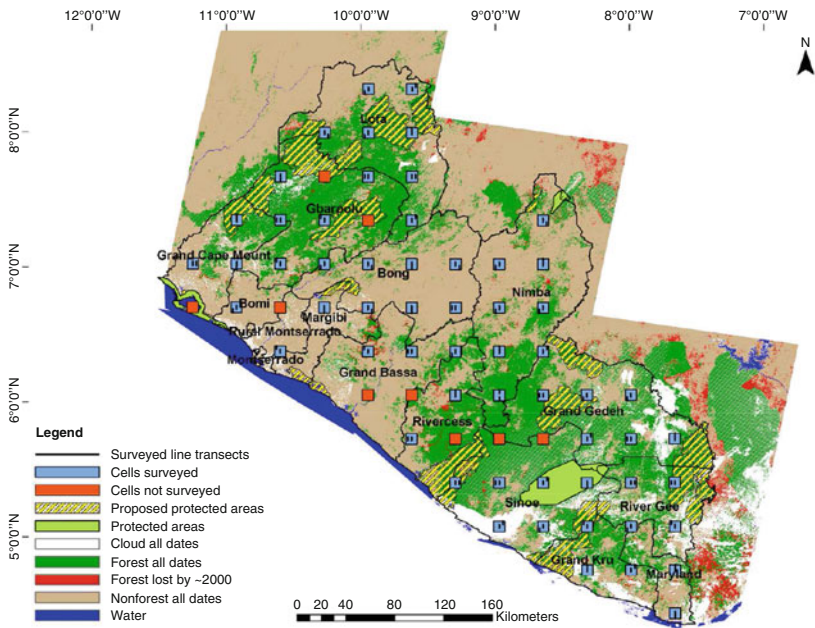
We use as a second example of an indirect survey the Liberian nationwide survey of West African chimpanzees (*Pan troglodytes verus*) conducted in 2010–12 (Tweh et al. 2015). Liberia holds the second largest population of West African chimpanzees, and is one of its last strongholds, making it a priority for conservation. We consider here the three elements needed to estimate abundance using the standing crop method: the line transect survey, conducted throughout the range of the population of interest, to estimate nest density; the mean time to decay of nests, allowing the nest density estimate to be converted to an estimate of the number of nests constructed per unit area per day; and the mean number of nests constructed per day per animal, allowing the estimated number of nests constructed per unit area per day to be converted to animal density and hence abundance.

9.3.4.1 Line Transect Survey

A systematic sample of 68 grid squares, each of size  $9 \times 9$  km, was taken (Fig. 9.9). Nine of these squares were not surveyed due to access problems. Each selected grid square was divided into nine  $3 \times 3$  km<sup>2</sup>. A 3 km transect was placed through the central  $3 \times 3$  km square, running north/south, and a second 3 km transect was similarly placed in one of the remaining eight  $3 \times 3$  km<sup>2</sup>, selected at random. Some transects could not be fully surveyed, again due to access problems, resulting in total transect length surveyed of  $L = 326.18$  km, and 113 nests were detected. Fig. 9.10 shows the survey team measuring the distance from the line of a detected nest.

Tweh et al. (2015) fitted a half-normal detection function to grouped distance data, truncated at  $w = 40$  m, so that truncated sample size was  $n = 108$  (Fig. 9.11). The resulting estimate  $\hat{\mu}$  of the effective strip half-width was 22.56 m, or 0.02256 km. Thus from Eq. (6.13), we can estimate nest density in the study region:

$$\hat{D} = \frac{n}{2\hat{\mu}L} = \frac{108}{2 \times 0.02256 \times 326.18} = 7.338 \text{ nests km}^{-2}. \tag{9.10}$$



**Fig. 9.9** Map of Liberia showing the systematic survey design used to survey chimpanzee nests. Reproduced from Tweh et al. (2015)

**9.3.4.2 Decay Rate Experiment**

From August to December 2010, reconnaissance paths were walked to locate fresh chimpanzee nests such as the one shown in Fig. 9.12. Each such nest was marked, and revisited once, after varying lengths of time, to record whether or not the nest had decayed. The left-truncated logistic model (Eq. (9.7)) was used, and the mean time to decay of nests was estimated to be  $\hat{\xi} = 164.38$  days.

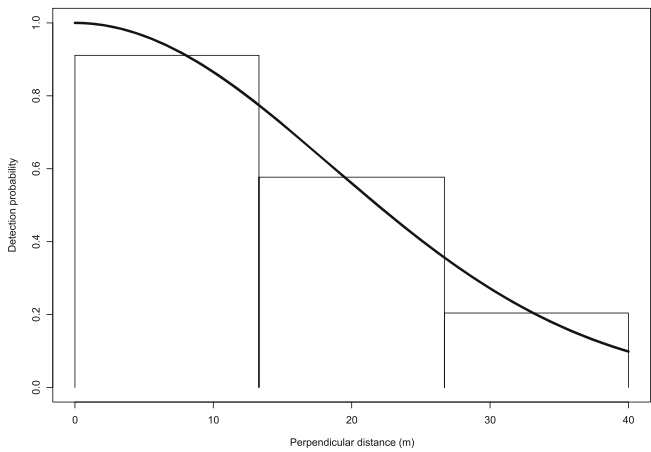
**9.3.4.3 Nest Production Rate**

Kouakou et al. (2009) estimated the nest production rate for nest builders as 1.143, while Plumptre and Cox (2006) estimated that the proportion of nest builders in the population was 0.83. Thus the estimated nest production rate for the population as a whole is  $\hat{\eta} = 1.143 \times 0.83 = 0.94869$ .





**Fig. 9.10** The survey team measuring the distance from the line of a detected nest. Note that the line is marked, allowing accurate measurement of detection distances. Photo: Hjalmar S. Kühl



**Fig. 9.11** Fitted half-normal detection function for the chimpanzee nest data



**Fig. 9.12** The *left-hand photo* shows a fresh chimpanzee nest, while the *right-hand photo* shows the same location after the nest has decayed. It is critical to have clear criteria that can be consistently applied to determine whether a nest has ‘decayed’. Photos: Celestin Y. Kouakou

#### 9.3.4.4 Animal Density and Abundance

We now substitute our estimates into Eq. (9.3):

$$\hat{D} = \frac{\hat{R}}{\hat{\xi}\hat{\eta}} = \frac{7.338}{164.38 \times 0.94869} = 0.0471 \text{ chimpanzees km}^{-2}. \quad (9.11)$$

The size of the survey region was approximately 107,000 km<sup>2</sup>, and our final abundance estimate is 5045 animals. Tweh et al. (2015) provide more details. Their corresponding 95 % confidence interval was (2974, 8559) animals.

## 9.4 Cue Counting

Cue counting (Hiby 1985) was designed to estimate whale abundance. Observers travel on a ship or in an aircraft along a line transect, recording cues (usually whale blows) within a sector ahead of the platform, along with distances of those cues from the platform. The methodology estimates the number of cues produced per unit of time per unit of area. Hence, effort is measured as the time spent observing, not the distance travelled. Thus, the method is equally applicable to point transect surveys, in which the observer does not move at all while recording cues (Buckland 2006). When the cue is aural, such as a songburst for birds in the breeding season, the sector angle can be taken as 360°; that is, all detections are recorded, irrespective of angle. To convert the number of cues per unit of time per unit of area to an estimate of animal density, an estimate of cue rate (average number of cues per unit of time per animal) is required.



Whether conducted along lines or at points, the methodology is closely similar to point transect methodology (Sect. 1.6). For point transect sampling, we had from Eq. (6.21)

$$\hat{D} = \frac{n}{K\hat{\nu}} = \frac{n\hat{h}(0)}{2\pi K}. \quad (9.12)$$

In that expression,  $K$  is the ‘effort’ expended; one visit per point for  $K$  points is a total effort of  $K$ . For cue counting, effort is determined by the total time  $T$  spent recording (regardless of number of points if the design is based on points, or of line length if the design is based on lines). It is also determined by the fraction of the circle centred on the observer that is counted. If the sector angle  $\phi$  is measured in radians, then the fraction of the circle counted is  $\phi/(2\pi)$ . Thus total effort is  $T\phi/(2\pi)$ , and replacing  $K$  by this in the above expression, we obtain an estimate of the mean number of cues per unit area per unit time:

$$\frac{2\pi n}{T\phi\hat{\nu}} = \frac{n\hat{h}(0)}{T\phi}. \quad (9.13)$$

To convert this to animal density, we need to divide by  $\hat{\eta}$ , an estimate of the mean cue rate (cues per animal per unit time) for the population of interest:

$$\hat{D} = \frac{2\pi n}{T\phi\hat{\nu}\hat{\eta}} = \frac{n\hat{h}(0)}{T\phi\hat{\eta}}. \quad (9.14)$$

The parameter  $h(0)$  is estimated exactly as in Sect. 6.3.2.1.

### 9.4.1 Cue Counting for Whales

Cue counting for whales is conducted from a ship or aircraft, using a line transect design. The cue is generally taken to be a blow, but for species without a visible blow, it could be a surfacing. Relative to conventional line transect sampling, cue counting has the following advantages.

- We do not need to assume that a whale that is on the line is certain to be detected — potentially an important advantage for species that are usually submerged. Instead, we assume that a cue is certain to be detected if it is right in front of the observer. (A cue right on the line, but well ahead of the observer, does not have to be detected.)
- Whale movement that is independent of the observer does not bias the method, while such movement for line transect surveys of whales from ships or boats is problematic, generating appreciable bias if mean whale speed exceeds around half the speed of the vessel (Sect. 11.4.1).

- There is no need to determine how many whales are in a group, or to keep track of whales to avoid double-counting.

The disadvantages are as follows.

- It is necessary to monitor a representative sample of whales to estimate their mean cue rate. Such monitoring is problematic. If a ship follows the whale, its cue rate may change. Also, there may be a tendency to sample whales that are more active, giving more cues, and therefore being more easy to locate and monitor — an example of size-biased sampling (Buckland et al. 2001, p. 192).
- Errors in estimating distances to cues generate greater bias in cue counting than do comparable errors in line transect sampling (Buckland et al. 2001, pp. 195–196). It is therefore more important to model measurement error (Sect. 11.3), especially as distances to a short-lived cue at sea are difficult to estimate with accuracy and precision.

Hiby et al. (1989) analysed data from an aerial cue-count survey of northern minke whales (*Balaenoptera acutorostrata*) and fin whales (*Balaenoptera physalus*) in the North Atlantic in 1987. Because blows of northern minke whales are rarely visible, a ‘cue’ was defined to be a dive. Measurement error in distance estimates was modelled (see Sect. 11.3.2). Although double-observer data were collected, and analysis illustrated adopting a mark-recapture distance sampling approach, numbers of duplicate detections were too small to yield reliable estimates unless  $g(0) = 1$  was assumed. Aside from this, the method appeared to work well, although coefficients of variation on abundance estimates for different populations tended to be quite large, ranging from 20 % to over 50 %.

### 9.4.2 Cue Counting for Birds

Cue count surveys for birds are most conveniently conducted from a grid of points, as for point transect sampling. They have potential for single-species surveys of birds such as northern bobwhite quail (*Colinus virginianus*) or kiwis (*Apteryx* spp.) that have distinctive calls, or songbirds, for which the cue is a songburst. Relative to standard point transect sampling, cue counts have the following advantages.

- Movement of the birds during the count does not bias the method, provided the movement is independent of the observer. This allows the observer to spend longer at each point, which is useful for species that call infrequently. The cue count method also has great potential for acoustic surveys, in which an acoustic array is placed around each point, allowing triangulation to estimate distances of calling or singing birds from the point. Such arrays can be left for extended periods in the field, given that bird movement during the recording period does not generate bias.

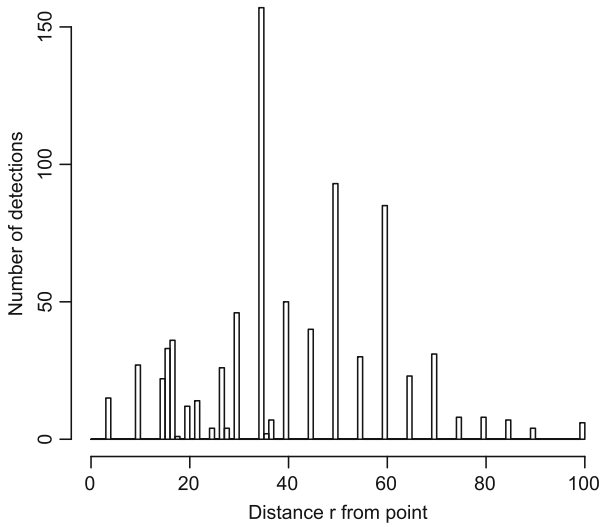
- Related to the above advantage, if a bird is moving around, there is ambiguity about its ‘true’ location while the observer is at the point, while the location of the bird at the time it gives a cue is well-defined.
- We do not need to assume that a bird at the point is certain to be detected. Instead we assume that, if such a bird gives a cue (call or songburst), then that is certain to be detected. In closed habitats, perhaps with a canopy well above the observer, this is a much more plausible assumption.
- The observer does not need to keep track of detected birds. In conventional point transect sampling, this must be done to avoid double-counting.

The disadvantages relative to standard point transect sampling are as follows.

- The observer can get ‘swamped’ in a multi-species songbird survey, with too many cues to allow accurate recording (Buckland 2006). Thus although the location of the bird when it gives a cue may be well-defined, it may also be difficult to measure or estimate accurately.
- Repeated cues from the same bird are not independent, so that analysis methods that are insensitive to the independence assumption should be used. Fortunately, conventional distance sampling methods are insensitive, except that the usual model selection tools, such as goodness-of-fit and AIC, are not, so that care is needed not to overfit the data when modelling the detection function.
- We need to estimate the mean cue rate: the mean number of calls or songbursts per bird per unit of time. This estimation should be done synchronously with the survey, because cue rate can vary with many factors, including weather, habitat, bird density, time of day and season. If two observers are available, then one might conduct the main survey, and the second estimate cue rates of birds detected within a predetermined distance of each survey point. This distance should be relatively small, so that all or almost all birds within that distance are detected, otherwise there is the risk that more vocal birds, with higher than average cue rate, are selectively sampled (Buckland 2006).

We return to the Montrave case study, as cue counting was implemented in that study, and compared with other methods (Buckland 2006). We again use data on the European robin (Fig. 1.2). This species provides a severe test of the cue count method because it does not have simple songbursts, as do most songbirds, but it strings together many short song phrases. The length of the overall song is very variable and poorly defined, so that song phrases must be recorded. A single bird can give around ten of these a minute for several minutes, and in areas of high density, the observer may be ‘swamped’, and fail to record all the phrases detected. Further, with so many song phrases from the same bird, perhaps from the same perch throughout the visit to the point, the robin data also represent an extreme departure from the assumption of independent detections.

The same 32 points were used as for point transect sampling (Fig. 2.1), and each point was again visited twice. For each visit, the observer attempted to record all song phrases detected over a 5-min period, together with estimated distance from



**Fig. 9.13** Estimated distances of the 791 robin song phrases detected from the point for cue counting, plotted by 1 m interval

the point. In addition, a note was made of any birds that remained close enough to the point throughout the 5 min to allow all song phrases to be recorded; cue rate was estimated from these birds.

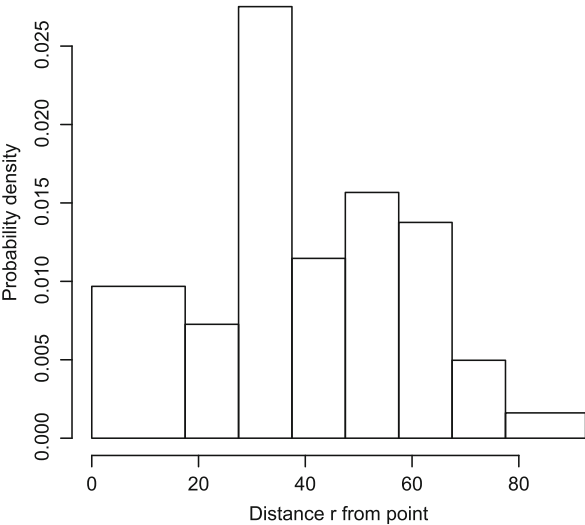
The mean cue rate was estimated at 3.585 song phrases per minute, with a standard error of 0.486, based on a sample size of 26 birds. (This cue rate was more than double that for each of the other three species surveyed, which all had mean cue rates of between 1.44 and 1.64 per minute.)

In total, there were 791 song phrases recorded, ranging from 4 to 100 m (Fig. 9.13). As for line and point transect sampling, there is clear evidence of rounding to the nearest 5 m. However, the variability in the height of the histogram bars is much greater than can be attributed to rounding alone. There is overdispersion present, which arises because multiple song phrases from the same bird are recorded. If the bird remains in the same place for the full 5 min of a count, then this might result in 30 or more song phrases detected at exactly the same distance.

In Fig. 9.14, we have plotted the distance data for the following intervals: 0–17.5 m, 17.5–27.5 m, 27.5–37.5 m, 37.5–47.5 m, 47.5–57.5 m, 57.5–67.5 m, 67.5–77.5 m, 77.5–92.5 m. This truncates just six detections. Overdispersion is still evident in this plot.

If we use routine model selection options for these data, the overdispersion may cause us to overfit the data. It is better therefore to constrain the detection function models so that they do not have the flexibility to fit features of the data that are due to overdispersion. In Table 9.5, we summarize estimation of the effective detection radius for uniform-cosine model with two adjustment terms (and thus two parameters in total), the half-normal-Hermite model with one adjustment term

**Fig. 9.14** Distribution of detected robin song phrases by distance (cue counting). Note that probability density is plotted on the y-axis rather than counts, because interval width varies, so that untransformed counts are not a valid guide to the shape of the probability density function



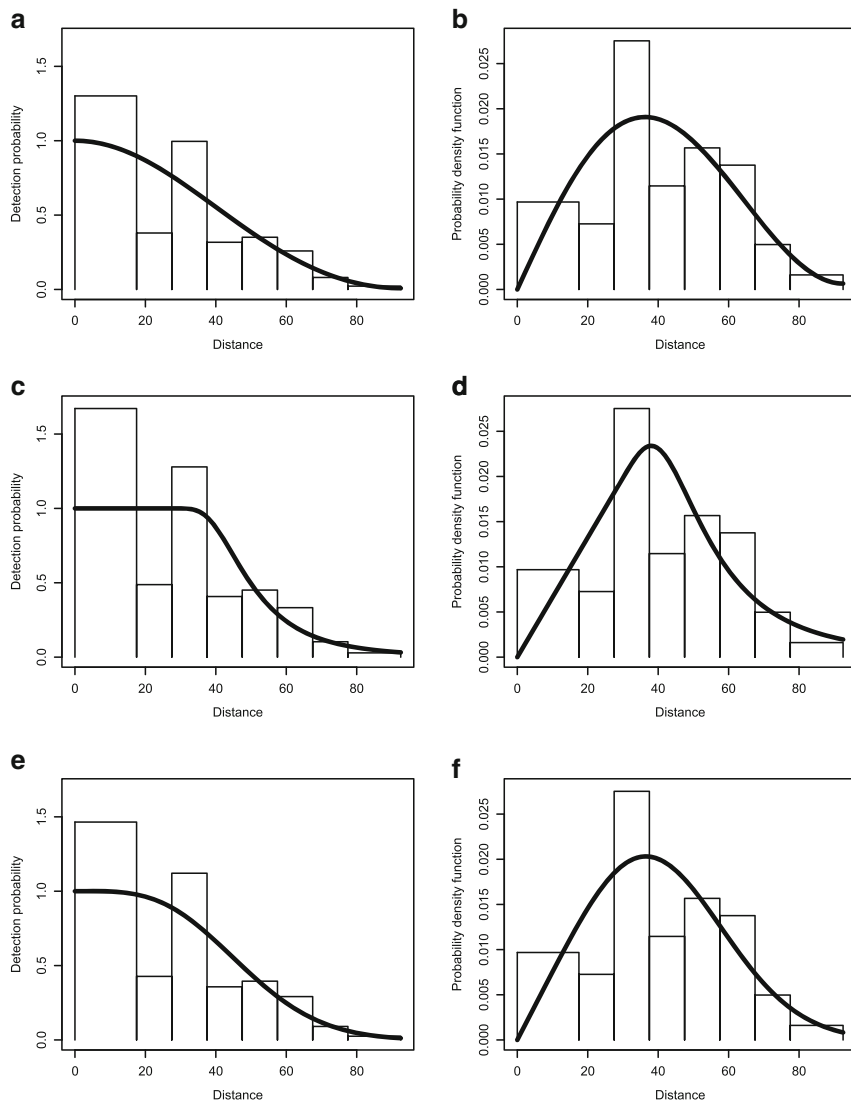
**Table 9.5** Estimates  $\hat{\rho}$  of effective detection radius for three detection function models applied to the robin cue count data from the Montrave case study

Key function	Adjustment type	No. adj. terms	$\Delta AIC$	$\hat{\rho}$ (m)	$cv(\hat{\rho})$	95 % CI for $\rho$
Uniform	Cosine	2	1.8	48.5	0.027	(46.0, 51.2)
Half-normal	Hermite polynomial	1	0.0	51.5	0.051	(46.6, 56.8)
Hazard-rate	None	0	13.9	55.0	0.019	(52.9, 57.1)

(also giving two parameters in total), and the hazard-rate model with no adjustment terms (again two parameters). AIC favours the half-normal-Hermite model; even in the presence of overdispersion, we would generally favour the model that has  $\Delta AIC=0$ . In the absence of overdispersion, we could rule out the hazard-rate model with a  $\Delta AIC$  value of 13.90, but given such strong overdispersion, we should not dismiss it as a possible model.

In Fig. 9.15, we show the estimated probability density functions and detection functions respectively corresponding to the three models. There is no obvious reason from these plots to favour one model over another. The Kolmogorov–Smirnov, Cramér–von Mises and  $\chi^2$  goodness-of-fit tests do not help us here — for all three models,  $p$ -values are all  $<0.005$ , a result of overdispersion rather than poor fit. The half-normal-Hermite model, favoured by AIC, is also intermediate in terms of the shape of the estimated detection function, and is arguably the best choice here.

Density estimates, coefficients of variation and confidence intervals for the three models for the robin cue count data are shown in Table 9.6. We see that estimated density for the model favoured by AIC is very close to the territory mapping estimate of 84 territories  $\text{km}^{-2}$ , although the confidence interval is fairly wide. The other two models also give confidence intervals that easily include the territory mapping estimate.



**Fig. 9.15** Estimated detection functions (*left*) and probability density functions (*right*) for the cue count data for robins. Uniform cosine model (**a**, **b**), hazard-rate model (**c**, **d**) and half-normal-Hermite model (**e**, **f**)

**Table 9.6** Density estimates corresponding to the three detection function models applied to the robin cue count data from the Montrave case study

Model	$\hat{D}$	$cv(\hat{D})$	95 % CI for $D$
Uniform-cosine	92.6	0.223	(59.8, 143.3)
Half-normal-Hermite	82.3	0.239	(51.6, 131.0)
Hazard-rate	72.1	0.219	(46.9, 110.9)

Units of density estimates are territories  $\text{km}^{-2}$

### 9.5 Surveys in Three Dimensions

Wild animal populations exist in four dimensions: latitude, longitude, altitude and time. When sampling for estimating abundance or density, for practical reasons we might decide to ignore some of these, which in practice is translated into some conceptual assumption. As an example, in conventional line transect sampling, we assume that the survey occurred instantaneously, hence removing time from the equation. Here we think about the three spatial dimensions (3D). In fact, as is apparent from Chap. 5, the modelling of distance sampling data usually involves data in a single dimension, the distance from the transect. This means that, besides the temporal dimension, we have collapsed data over two additional dimensions, the “along-transect” dimension and altitude. In this section, we look in particular at why we usually ignore the altitude dimension (*i.e.* height or depth) and explore the consequences of doing so. We also consider the circumstances in which we might want to incorporate altitude explicitly in the detection function modelling and present some pointers about how that might be implemented.

Arguably, all wild animal populations explore and use the vertical dimension of their habitat. However, whether we need to address this depends on the scale at which it occurs. In terms of distance sampling, the key aspect is the scale at which it occurs relative to the scale of the detection process. If animals move little in the vertical dimension (*e.g.* a few metres, as occurs for example with birds in scrubby habitat), while detection distances are typically an order of magnitude greater (*i.e.* tens of metres), then we can safely ignore the vertical dimension, conceptually projecting the animals onto a two-dimensional surface. Thus we would not typically consider the vertical dimension in considering for example ungulates, ground-dwelling insects, plants, or any other taxa that live on or close to the ground. On the other hand, if surveying say chameleons or birds in a dense forest, where typical detection distances might be of the same order as, or even considerably less than, their range of distributions in height, this is something that should be taken into account.

- Surveys for which the third dimension might not be negligible include:
- Bird (*e.g.* Mateos et al. 2010), bat or insect (*e.g.* Chapman et al. 2002; Reynolds and Riley 2002) radar surveys for which the radar is placed on or near the ground and animals exhibit a strong non-uniform distribution with height.
  - Active acoustic 3D surveys for fish and krill, with sensors at or near the surface and strong animal density trends with depth (Cox et al. 2011).

- Tropical forest surveys with transects on foot focussing on animals distributed through the canopy, such as birds (Gale et al. 2009) or chameleons (Jenkins et al. 1999).
- Passive acoustic cetacean surveys with sensors deployed at the bottom, or fixed or towed at the surface, for animals detectable at ranges that may be smaller than the ocean depth (Harris et al. [in prep.](#)).

In these circumstances, the strategy of projecting animal locations onto a two-dimensional surface (which may be the ground or sea surface) may lead to bias if standard distance sampling methods are used. This is because an animal may be at zero horizontal distance from the observer, but at a sufficiently large vertical distance that probability of detection is well below unity. This violates the standard distance sampling assumption that probability of detection of an animal that is on the line or at the point is one. Resulting abundance estimates will then be biased low.

If one could assume that the vertical animal distribution was uniform, then extension of standard methods to three dimensions is relatively straightforward (Burnham et al. 2004, pp. 307–311). In a horizontal plane, we can ensure uniformity on average by distributing a number of lines or points according to a randomized scheme, but in the vertical dimension, this is usually not an option, and the animals themselves generally show a strong trend in density with altitude or depth. Thus methodology that assumes animals are uniformly distributed in the vertical dimension is perhaps theoretically interesting but of little practical use.

Given that animal distribution typically varies with height or depth, and that transects are generally not randomly positioned with respect to height or depth, a strong vertical density gradient of animals with respect to transects might be present. This is problematic and the concept of 3D distance sampling becomes just another example of distance sampling in the presence of density gradients with respect to the transects (Sect. 11.1), albeit a significantly harder version of that problem because now the density gradient exists in the third dimension.

## 9.6 Radar, Sonar and Acoustic Surveys

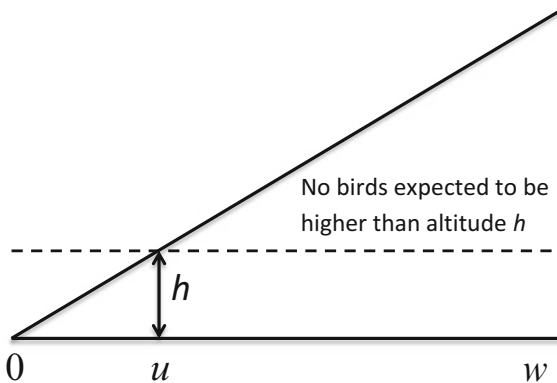
### 9.6.1 Radar Surveys

The use of radar to identify migration and flight behaviour in birds goes back to the 1970s (Bruderer and Steidinger 1972; Able 1977). The use of radar was proposed by Harmata et al. (1999) to detect bird movement in the context of impact assessment. It has been used to estimate distances to birds (Petersen et al. 2006), and in the context of a distance sampling survey to calibrate observer estimates of distance-to-bird measurements (Mateos et al. 2010). Shamoun-Baranes et al. (2014) propose the use of these systems to evaluate bird migrations at a continental scale.

As radars are usually located at ground, we must address the typically non-uniform distribution of animals with respect to altitude. One strategy is to left-



**Fig. 9.16** Diagrammatic representation of bird radar, showing the vertical sector searched. If birds fly no higher than altitude  $h$ , and if we left-truncate the distance data at  $u$ , then we can project the data onto a horizontal plane, and analyse them using two-dimensional distance sampling methods



truncate data, so that only distances at which the entire vertical range of the species is covered are included in the analysis (Fig. 9.16). This allows us to project distances onto the ground, and if the radar is located independently of the birds, for any sweep of the radar, we can consider that the distances of birds from the radar (whether detected or not) follow a triangular distribution on average. Thus any fall-off from the triangular distribution represents a failure to detect all birds at larger distances.

### 9.6.1.1 Sector-Based Analysis

We assume that the detections are of either individual birds or bird flocks, and that, for each bird or flock detected on a given sweep, we have a horizontal distance  $r$  from the radar. In the following, if flocks are recorded rather than individual birds, then density estimates will relate to number of flocks per unit area.

Thus we have two-dimensional point transect sampling data for distances  $r$ , with  $u \leq r \leq w$ .

Typically, radar generates large volumes of data. One method of sub-sampling is to sample sectors from the data for a full sweep. Suppose that  $k$  sectors are counted for each sweep, each sector with horizontal angle  $\theta$ , so that the proportion of the circle counted is  $k\theta/(2\pi)$ . (If the whole of each sweep is counted, then  $k = 1$ .) Suppose that the number of sweeps is  $m$ . Then total effort is  $km\theta/(2\pi)$ .

Let  $g(r)$  be the probability that a bird (or flock) at horizontal distance  $r$  from the radar is detected.

Assuming radar location to be independent of bird flight paths, then the distribution of distance  $r$  irrespective of whether the bird is detected is  $\pi(r) = \frac{2r}{w^2}$ ,  $0 \leq r \leq w$ . With left-truncation at  $r = u$ , this becomes  $\pi(r) = \frac{2r}{w^2 - u^2}$ ,  $u \leq r \leq w$ . Under this assumption, the probability density function of  $r$  is

$$f(r) = \frac{2\pi r g(r)}{\nu} \quad (9.15)$$

where  $\nu = 2\pi \int_u^w rg(r) dr$ . The mean probability that a bird (or flock) that is in a sector is detected is

$$P_a = \int_u^w g(r)\pi(r) dr = \frac{\nu}{\pi(w^2 - u^2)} . \quad (9.16)$$

Bird (or flock) density is then estimated as

$$\hat{D} = \frac{n}{m} \times \frac{2\pi}{k\theta} \times \frac{\pi(w^2 - u^2)}{\hat{\nu}} \times \frac{1}{\pi(w^2 - u^2)} = \frac{n\hat{h}(0)}{mk\theta} \quad (9.17)$$

where  $n$  is number of birds (or flocks) detected (summed over the  $k$  sectors of each sweep and across the  $m$  sweeps), and from Eq. (5.33),

$$\hat{h}(0) = \frac{2\pi}{\hat{\nu}} . \quad (9.18)$$

Thus we estimate the mean density of flocks across the  $m$  sweeps, and we can treat the sweeps as independent replicates for estimating precision. The data may be analysed in *Distance* using point transect sampling options with left-truncation at  $r = u$ , the  $m$  sweeps taken to correspond to points, and a sampling fraction of  $k\theta/(2\pi)$ .

### 9.6.1.2 Radial Analysis

A radial is defined to be a strip of constant width  $z$ , extending from the radar location. It represents an alternative method of sub-sampling a sweep. Under the assumption that the radar is positioned independently of bird flight paths, the number of flocks available for detection along a single radial will be uniformly distributed (since, unlike for a sector, the area covered by a radial does not increase with distance from the radar). With left-truncation at  $r = u$  and right-truncation at  $r = w$ , this gives  $\pi(r) = 1/(w - u)$ , independent of  $r$ . Thus the probability density function of recorded distances is now

$$f(r) = \frac{g(r)}{\mu} \quad (9.19)$$

where  $\mu = \int_u^w g(r) dr$ .

The mean probability that a flock that is within a radial is detected is  $P_a = \int_u^w g(r)\pi(r) dr = \frac{\mu}{w-u}$ .

We again assume that there are  $m$  sweeps, and we assume that data from a single sweep are recorded along  $k$  radials. An annulus with internal radius  $u$  and external radius  $w$  has area  $\pi(w^2 - u^2)$ . A single radial covers an area  $z(w - u)$  within this

annulus. Thus given that a bird (or flock) is within the annulus during a sweep, the probability that it intersects a radial is

$$\frac{kz(w-u)}{\pi(w^2-u^2)} = \frac{kz}{\pi(w+u)} . \quad (9.20)$$

Hence the probability that a flock that is in the annulus at the time of a sweep is recorded is

$$\frac{kzP_a}{\pi(w+u)} = \frac{kz\mu}{\pi(w^2-u^2)} , \quad (9.21)$$

and so we can estimate flock density as

$$\hat{D} = \frac{n\pi(w^2-u^2)}{mkz\hat{\mu}} \times \frac{1}{\pi(w^2-u^2)} = \frac{n}{mkz\hat{\mu}} . \quad (9.22)$$

We can use `Distance`, selecting a line transect analysis with left-truncation, to estimate  $\mu$  by the effective strip half-width of line transect sampling. In fact, we can ‘fool’ `Distance` into doing the full analysis by entering the data as if they had been recorded on  $m$  lines, each of length  $kz/2$ .

### 9.6.1.3 Summary

Schmaljohann et al. (2008) give a detailed account of the problems associated with detectability in bird radar surveys. Dokter et al. (2013) match visual detections with radar tracks, to assess detectability for radar. In the above development, we have assumed that detection is certain for a bird (or flock) at distance  $u$ , and in many circumstances, this assumption will be violated.

As for acoustic data and high-resolution imagery, detection and classification of animals of interest needs to be fully automated, if the methods are to reach their full potential. It is often difficult or impossible to assign echoes to a given species, or to determine the number of individuals. This contrasts with cetacean acoustic data, where the process of automating detectors and classifiers is relatively well-developed. Zaugg et al. (2008) give an example where large groups of animals (*e.g.* birds versus insects) are automatically detected and classified using radar images.

## 9.6.2 Active Sonar Surveys

Cox et al. (2011) present an application of 3D distance sampling methods (Sect. 9.5) in which the third dimension has been explicitly modelled. The data arise from a

sonar survey of krill, and the krill distribution in depth was modelled along with detectability. The additional information required to achieve this was obtained from the angles of detection, via a modification of a procedure developed in Marques et al. (2010) to deal with non-uniform density of hares with respect to roads (Sect. 11.1.2). This highlights a conceptual link between 3D distance sampling in which transects are not located randomly with respect to height or depth, and standard distance sampling in which lines are not placed randomly with respect to animal density gradients (*e.g.* road transects).

### 9.6.3 *Passive Acoustic Surveys*

Over the last few years, there has been a rapid increase in the use of passive acoustic data for animal density estimation. A review of existing methods is presented by Marques et al. (2013). Distance sampling from acoustic data has focussed on cetacean applications, but we predict that these will also be increasingly used for other marine (*e.g.* fish) and terrestrial applications (*e.g.* Dawson and Efford 2009; Thompson et al. 2009; Blumstein et al. 2011). Acoustic arrays offer many options for distance sampling surveys. The resulting data can be analysed by standard distance sampling methods, provided distances of detected animals from a sensor can be estimated, and provided an animal at or near the sensor is sure to be detected. Cue count options are more suited to acoustic data though, as random animal movement does not then generate bias in density estimates. Further, we need only assume that any cue (typically a call or songburst) is detected with certainty for any animal at or near a sensor; it does not matter if we fail to detect a silent animal at the sensor. If an animal can be detected from more than one sensor, then acoustic data can be analysed using spatially-explicit capture–recapture methods (Sect. 12.2.1). Acoustic data can also be used in MRDS methods (Sect. 5.4), in which data from visual observers are analysed together with acoustic data. The advantage is that the two platforms (visual and acoustic) more plausibly generate independent detections than occurs when the two platforms use the same search method. However, it can be difficult to identify duplicate detections (*i.e.* animals detected from both platforms), because for marine mammals at least, typically when an animal is available for acoustic detection, it is unavailable for visual detection, and *vice versa*.

The relationship between 3D acoustic distance sampling, with non-random placement of sensors in the vertical dimension, and non-random placement of lines or points in 2D distance sampling, is explored by Harris et al. ([in prep.](#)), who considered acoustic sensors placed at either the ocean floor or the surface, so that they had to account for the non-uniform animal distribution with respect to depth. A solution is to simplify the approach by conceptually collapsing the dimensions over which the animal distribution might be assumed uniform (*i.e.* in the  $x, y$  plane if sensors are randomly placed in the study area).

# Chapter 10

## Taxon-Specific Issues

In this chapter, we consider some of the problems that may be encountered for specific taxa. We make no attempt to be comprehensive; rather, we have selected a few taxa to illustrate the types of issue that arise in some circumstances. The absence of a taxon from this chapter does not indicate that distance sampling methods are inappropriate for that taxon.

### 10.1 Songbirds

In this section, we discuss some of the issues that arise in breeding season surveys of songbirds. We draw on material from Buckland et al. (2008), and we make frequent reference to the Montrave case study (Buckland 2006).

#### 10.1.1 Line Transect Versus Point Transect Surveys

When both line transect and point transect sampling are feasible, then line transect sampling tends to give higher precision (Buckland 2006). This is because we need to estimate the probability density function (line transect sampling) or the slope of the probability density function (point transect sampling) at distance zero (*i.e.*  $f(0)$  or  $h(0)$  respectively). For line transect sampling, a higher proportion of recorded distances is close to zero than for point transect sampling, simply because there are more birds available to be detected close to a line than to a point, and so we have better information for estimating  $f(0)$  than we do for  $h(0)$ . A second reason for the higher precision of line transect sampling is that a higher proportion of time spent in the field is spent actively searching for birds; time travelling between lines is typically appreciably less than time travelling between points because there tend to

be many more points in a point transect survey than lines in a line transect survey. Further, there tends to be less bias arising from bird movement (Sect. 10.1.2) or from inaccurate distance estimates (Sect. 10.1.4) for line transect sampling than for point transect sampling.

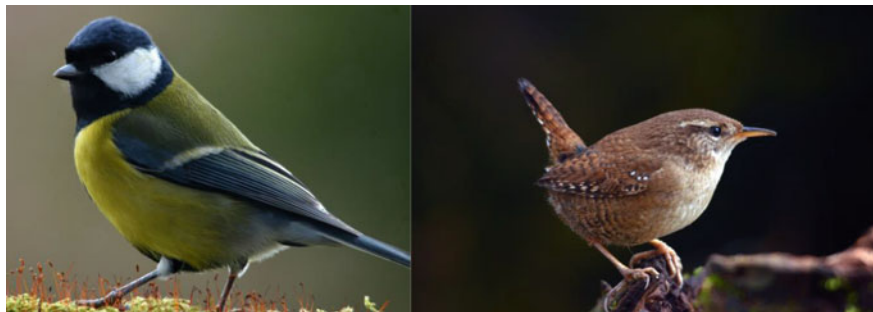
To offset these advantages, there are several disadvantages in using line transect sampling. First, we would like to have a random design for our survey, to avoid bias. In difficult terrain, or study areas with access difficulties, it tends to be easier to reach a random point than to traverse a random line. If a completely randomized design is not feasible, the compromises are likely to be fewer, and hence the bias lower, for point transect sampling than for line transect sampling. Second, if it is important to assess how density depends on habitat, it is much easier to quantify habitat at a point than along a line. Third, for songbird surveys, it is common for many species to be recorded. For observers attempting to detect and record a number of individuals from various species, it is easier if they are stationary at a point than if they also have to concentrate on walking along a randomly-located line in potentially difficult terrain.

## ***10.1.2 Minimizing Bias from Bird Movement***

### **10.1.2.1 Movement Independent of the Observer**

Conceptually, we think of animals frozen in location while we survey them. When this is not the case, density tends to be over-estimated. For line transect sampling, bias is slight provided average speed of the animals is under half that of the observer (see Fig. 11.8). However in standard point transect sampling, for which an observer records from each point for a predetermined length of time (typically between 5 and 20 min), there is potential for substantial bias. In the Montrave case study, for which the observer recorded all male birds detected for a period of 5 min at each visit to a point under the standard method, this movement generated clear bias for one of the four species surveyed. The great tit (Fig. 10.1) tended to keep on the move, so that, if the observer stood at a point long enough, one would eventually move into detection range. Territory mapping indicated that there were 21 territories  $\text{km}^{-2}$ . By contrast, standard point transect sampling gave an estimate of 58, with 95 % confidence interval of (39, 94) territories. A snapshot method eliminated this bias, giving an estimate of 22 territories  $\text{km}^{-2}$ , with a 95 % confidence interval of (12, 39) territories (Buckland 2006).

In the snapshot method, after arriving at the point, the observer notes any birds detected, and as far as possible, keeps track of them. The snapshot moment occurs a pre-determined time after reaching the point, say 2 min. Observers estimate distances to any birds whose location is known at the snapshot moment. They may take as much time as they wish after the snapshot moment to confirm locations of birds, but should not record new birds which may have moved onto the plot after the snapshot moment. If a bird was located either before the snapshot moment or



**Fig. 10.1** Of the four species surveyed in the Montrave case study, the great tit (*left*) showed the greatest tendency to move large distances during the 5-min counts used in the standard point transect method. This generated clear bias in the abundance estimate, which was avoided by using the snapshot method. The winter wren (*right*) exhibited the strongest evidence of responsive movement, probably because it was typically at or near ground level, while other species tended to be in trees. Photos: Steve Buckland

after, or both, but its location was uncertain at the snapshot moment, then it is not recorded. This does not cause bias, unless locations cannot be confirmed for birds at or very near the point, which we assume are certain to be recorded (Sect. 10.1.3). After the snapshot moment, the observer is allowed to move away from the point, to check locations of birds, and to allow triangulation to singing birds, so that distances may be estimated more accurately.

More careful observer training tends to be needed if the snapshot method is adopted, as the rules about how the observer should operate in the field are less rigid and more open to interpretation than those typically adopted for standard point transect sampling based on timed counts.

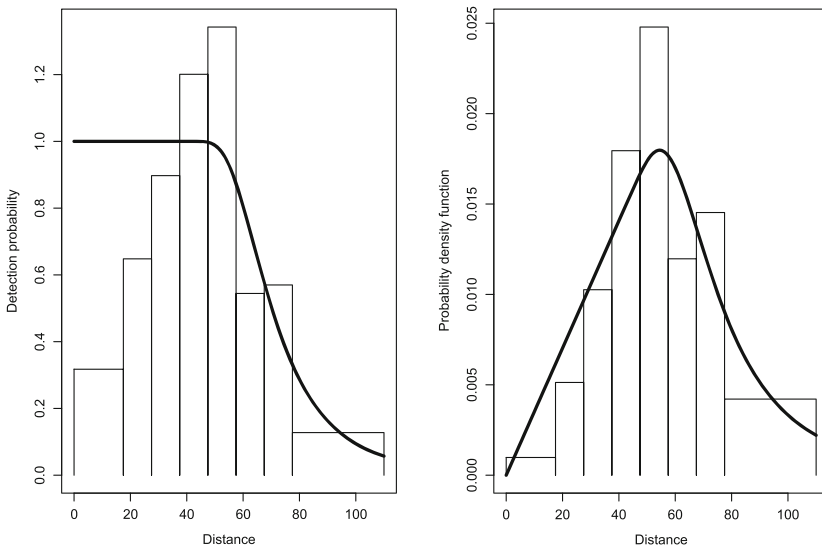
If it is not feasible to implement the snapshot method, but movement independent of the observer is known to be a problem, then the cue-count method (Sect. 9.4.2) should be considered. Because the recording unit for that method is short songbursts, bird movement that is independent of the observer does not cause bias.

### 10.1.2.2 Responsive Movement

As for movement independent of the observer, responsive movement generates bias in both line transect and point transect sampling, and again, the potential for bias is greater for point transect sampling than line transect sampling. Some songbirds will be attracted to an observer; in such cases, if they are not detected prior to responding, upward bias in density estimates will occur. This bias is larger if the observer double-counts a bird that approaches the point, believing it to be two different birds. If birds move away from the observer before detection, under-estimation of density occurs, which might be substantial if many such birds move beyond detection range.

In the Montrave case study, we noted in Sect. 5.2.3.3 an indication that some European robins might have moved away from the point before detection. There was stronger evidence of avoidance behaviour for the winter wren (Fig. 10.1). In the right-hand plot of Fig. 10.2, we show a histogram of estimated distances of detected wrens from the point, together with the fitted probability density function corresponding to the hazard-rate model for the detection function. In the left-hand plot, we show the fitted hazard-rate detection function. There appears to be clear evidence of too few detections near the point, and too many at distances of around 40–50 m. The effect is exaggerated in the bottom plot, due to the rescaling of frequencies; the top plot is the better guide. In fact, goodness-of-fit tests even here do not give compelling evidence of avoidance — the  $\chi^2$  test is not quite significant at the 5 % level, and the Kolmogorov–Smirnov and Cramér–von Mises test statistics are not significant at the 10 % level. Nevertheless, all four survey methods (standard point transect, snapshot point transect, cue count and line transect) gave data showing similar evidence of avoidance; if we pool the evidence, it is clear that some avoidance occurred.

Provided any avoidance does not take birds beyond detection range, bias due to responsive movement away from the line or point tends to be fairly modest. This is because we constrain fitted detection functions to be non-increasing functions. As a consequence, the fitted model tends to average out the lack of detections near the line or point with the excess of detections at mid-distances. This effect is evident in Fig. 10.2.



**Fig. 10.2** *Left*: the fitted hazard-rate detection function, plotted with scaled distance frequencies. *Right*: histogram showing relative frequencies of estimated distances to detected wrens for the snapshot point transect method, together with the fitted probability density function corresponding to the hazard-rate model



### ***10.1.3 Minimizing Bias from Failure to Detect All Birds on the Line or Point***

For most species of songbird during the breeding season, the male sings frequently for at least part of the day, so that for surveys conducted at that time of the day (usually early morning), males near the line or point are readily detected. However, females may be silent, and may be hidden on nests. Thus we cannot expect to detect all females even if they are on the line or point. It is usually preferable therefore to survey just territory-holding males, for which the assumptions of distance sampling are likely to hold. We thus obtain an estimated density or abundance of territories; and for most species, doubling this number is likely to give a better estimate of adult abundance than would a survey in which all detected adults are recorded.

However, some species do not sing, or do not sing frequently enough to ensure that they will be detected through their song if they are at the point. In open habitats, this might not be a problem, if such birds are readily detectable visually. In forest habitats for example, birds may be high up in the canopy, and only detectable if they sing or call. Similarly, birds in dense bush may be undetectable visually. If individuals of such species are frequently silent, then line transect sampling is unlikely to give reliable abundance estimates. In the case of point transect sampling, the time spent at the point might be increased, so that a male is likely to sing at some point during a visit to a point. However, the longer the observer spends at the point, the greater the bias from movement of the bird. A bird moving around will tend to be detected when by chance it passes close to the observer, so that recorded distances tend to be too small. This generates upward bias in abundance estimates, which will cancel with downward bias arising from not detecting all birds at the line or point only if the choice of time spent at each point is fortuitously just right. Rather than rely on this, cue counting (Sect. 9.4.2) is likely to be a better option. In cue counting, there is no requirement for the observer to detect a bird at the line or point with certainty; instead, the song (or call) of such a bird should be detected with certainty. That is a much more plausible assumption for many cryptic species. A further advantage of cue counting is that the observer can remain at the point for as long as necessary; bird movement that is independent of the observer does not bias the method.

We anticipate that many songbird surveys in future will be conducted as acoustic surveys. A suitable acoustic array can be left at a point for a period of time. Distances of singing birds from the point can be estimated by triangulation, given the slightly different times of arrival of a songburst at the different elements of the array. The data may then be analysed using the cue-count methods of Sect. 9.4.2. Alternatively, acoustic data may be analysed using spatially-explicit capture–recapture methods (Sect. 10.3.3.1). For these approaches to be cost-effective, computer software that automatically extracts songburst data and identifies species would be very useful, and is technically feasible. If it is further possible to identify individuals by their song, then such methods may also be an important aid in estimating cue rates (mean number of songbursts per bird per unit of time during the survey period).

### ***10.1.4 Minimizing Bias from Inaccurate Distance Measurement***

Errors in estimating distances are more of an issue for surveys of breeding songbirds than for most distance sampling surveys. There are two reasons for this. First, typically, a high proportion of birds are detected by sound alone, and it is more problematic to estimate distance to a sound than to a visual detection. Second, for most taxa, line transect sampling tends to be preferred to point transect sampling, but because songbird surveys are often multi-species, with a risk of the observer being ‘swamped’ by the number of songbursts of various species in a short time, and because it can be difficult to navigate a random line on foot, point transect sampling is often preferred for breeding songbird surveys. Measurement error creates much greater bias for point transect estimates than for line transect estimates (Sect. 11.3).

For point transect sampling, it is common for observers to record distances from the point in distance intervals: the data are counts by distance interval. This eases the task of the observer if many birds might be detected at a single point. However, it is still necessary to record the correct distance interval for each detected bird, as far as possible.

Laser rangefinders are inexpensive and invaluable tools for estimating distances in point transect surveys. In our view, songbird distance sampling surveys should not be conducted without them. They allow distances to a visual cue to be measured to the nearest metre. While they are less useful for aural cues, they still remove one element of guesswork: to estimate the distance to an aural cue, the observer must first estimate where the bird is, and then estimate the distance to that estimated location. Rangefinders remove the estimation error from the second step. (If the location is not visible, the observer can measure the distance to a visible location that is thought to be around the same distance from the point.) Also, rangefinders are invaluable for checking observers’ abilities to estimate distances in training, in distance estimation experiments, and as feedback during the survey.

Because of the difficulty in estimating distances to birds that can be heard but not seen, often point counts are carried out without recording distances. Nichols et al. (2000) developed a removal estimator to allow for imperfect detection on the plot in the absence of detection distances, in which a primary observer ‘removes’ birds, and a secondary observer records birds that he or she detects that were undetected by the primary observer. Farnsworth et al. (2002) developed a similar removal estimator approach, but divided the time at the point into periods, with new detections in later periods being recorded. The difficulty with this approach is that removal estimators can be badly biased in the presence of unmodelled heterogeneity (differences among birds in the probability of being detected). The main source of unmodelled heterogeneity in this case is due to distance of the bird from the point. Farnsworth et al. (2002) attempted to address heterogeneity by allowing two categories of bird: those with high detectability and those with low. However, heterogeneity arising from distance of the bird from the point will tend to be a

smooth function of distance, with detection probability decreasing with distance. For a more detailed discussion, see Buckland et al. (2004, pp. 352–354).

If detection distances are recorded, then the problem of heterogeneity may be overcome by combining removal estimators with distance sampling methods, as described by Buckland et al. (2004, pp. 354–356). However, bias also arises from bird movement. If birds move around during the count period, the size of the plot being surveyed is not well-defined because birds initially outside the surveyed circle may enter it. The effect of this for the method of Farnsworth et al. (2002) is that too many new detections are recorded in later time periods, leading to overestimation of bird density. For the method of Nichols et al. (2000), both observers tend to record too many detections, and overestimation again occurs.

## 10.2 Seabirds

For colonial seabirds, it is generally simpler to estimate the size of breeding colonies than to estimate numbers at sea. For smaller colonies, complete counts can be attempted; for larger colonies, some form of plot sampling is typically used. However, it is often necessary to estimate numbers at sea. This may be because a species does not nest in well-defined colonies that can easily be surveyed, or because the location of some colonies is unknown. In other cases, it may be because density of birds at sea rather than abundance of a biological population is of interest. For example, if an offshore windfarm is proposed and an impact assessment is required, we wish to quantify the density or abundance of birds within the footprint of the proposed windfarm to assess potential impact. Once such a windfarm is constructed, we would like to assess to what extent numbers have changed, relative to a control area. These issues are discussed in Sects. 3.3 and 7.4.3.

Seabirds at sea are surveyed by observers on board a ship or boat or in an aircraft. Increasingly, high-resolution video or photography taken from aircraft is used. The issues for each of these three options differ, and we treat them separately below.

### 10.2.1 *Shipboard Surveys*

For birds on the sea, it is generally straightforward to conduct a line transect survey. Diving species may be problematic; if so, observers should search using binoculars, so that birds on or near the line should be easily detectable on the surface between dives.

Distance estimation is problematic at sea. Laser rangefinders do not work for measuring to a point on the surface, unless it is possible to ‘hit’ the bird. Further, in high density areas, it is not practical to attempt to record distance to each detected bird. Typically therefore, a count of birds in each of several distance intervals is made. A judgement is still required to place each detected bird in one of the

intervals. Binoculars with reticles, ideally with reticle marks located to correspond to cutpoints between distance intervals, can be very useful. Failing this, a laser rangefinder can be used opportunistically to measure distances to floating objects, so that the observer can calibrate his or her estimates of distance against known distances.

If the target species can occur at very high densities, strip transect sampling (Sects. 1.4.1 and 6.2), with a relatively narrow strip to ensure that all birds within it are detected, might be preferred to line transect sampling.

Seabirds in flight typically have a mean speed that exceeds that of the observer, so that standard line transect methods would substantially over-estimate abundance. If such birds do not respond to the ship, a simple solution is to record birds in flight at their location when they pass abeam of the ship (Buckland et al. (2001, p. 202); see also Sect. 11.4.1). If a previously-detected bird is no longer visible at this moment, it is not recorded. In the case of strip transect sampling, the bird is only counted if it is in the strip when passing abeam.

Another solution to birds in flight is to survey only birds on the sea, then separately estimate the proportion of time spent flying, so that the abundance estimate can be adjusted for flying birds. Satellite tags, which provide data to determine whether a bird is flying, on the sea or on land, make this a more feasible option.

Spear et al. (1992) carried out strip transect sampling, and included detected birds in flight in the counts. They then adjusted their biased estimates, to take account of both speed and direction of travel for birds in flight (Spear and Ainley 1997a,b).

However, the most popular approach for dealing with birds in flight is plot sampling, using the method of Tasker et al. (1984), or a variant of this method. Instantaneous counts of birds in flight within a defined plot ahead of the ship are made at pre-determined points in time. Problems with the method are: the count in practice cannot be instantaneous; it is difficult to determine the boundaries of the plot, to decide which birds are in and which are out; and for smaller species, birds within the plot may pass undetected. Nevertheless, the method is likely to lead to substantially lower bias than would standard line or strip transect sampling in which all detected birds in flight are recorded.

General issues to consider with shipboard surveys are addressed by Buckland et al. (2001, pp. 288–291).

## **10.2.2 Aerial Surveys**

### **10.2.2.1 Visual Surveys**

Birds in flight do not need to be treated separately when carrying out aerial surveys, as the aircraft travels much faster than the mean speed of the birds. However, aerial observers can easily get ‘swamped’ in areas of high density. To reduce the likelihood

of this, distance intervals are defined, and counts of birds in each interval made. This task is made easier if markers are placed on wing struts (see Fig. 7.9 of Buckland et al. (2001, p. 261)), together with markers on the windows, and the observer simply positions him or herself so that the markers are aligned, and counts the number of birds passing between each set of markers.

Typically for aerial surveys, it is necessary to left-truncate the distance data, as birds directly below the aircraft may not be visible, or may be difficult to count. See Fig. 7.8 of Buckland et al. (2001, p. 260).

For aerial visual surveys, the aircraft must fly at low altitude, and this may result in birds taking flight as the aircraft approaches. In this circumstance, observers should ensure that they count birds in the distance interval that they were in before responding to the aircraft.

Relative to shipboard surveys, aerial surveys have the advantages that birds in flight do not generate significant bias to the standard method; birds that respond to the aircraft typically do not have time to move far before detection; survey costs tend to be lower (unless a small boat can be used rather than a ship). Disadvantages are that many areas are not accessible to land-based aircraft; identification of birds is more difficult; observers may be unable to make accurate counts in areas of high density; and for diving birds, some will be underwater and unavailable to be detected when the aircraft passes. In the case of availability, it may be necessary to estimate the proportion of time that birds are unavailable, to correct for birds that are underwater. In this case, it is important that availability is clearly defined, to ensure consistency between the aircraft observers and the study to determine proportion of time available. For example, the aerial observer may detect and count a bird ahead of the aircraft which then dives so that it is underwater at the time the aircraft passes abeam.

For a discussion of issues in conducting visual aerial surveys, see Buckland et al. (2001, pp. 280–287).

### 10.2.2.2 Surveys Using High-Resolution Imagery

Advances in technology and software mean that aerial surveys conducted using high-resolution video or photography can now out-perform visual surveys. The surveys are conducted at substantially higher altitude than visual surveys, which has two advantages: the aircraft does not disturb the birds, so no responsive movement occurs; and in studies to assess the impact of a windfarm on bird densities, the aircraft fly well above the turbines, allowing randomly positioned transects to be followed. For visual surveys, if aircraft are allowed onto the footprint of the windfarm at all, they must stay well away from turbines, compromising the random design.

A further advantage of digital surveys over visual surveys is that the birds counted are on record, and can be validated by different software or experts. As technology advances, resolution improves, allowing more reliable and verifiable species identification relative to visual surveys.

For diving species, availability is more tightly defined than for visual aerial surveys. For digital stills, the bird is available if it is at the surface when the still is taken, and not if it is underwater. For digital video, a line perpendicular to the transect may be superimposed on the video, and if a bird is at the surface when it intersects that line, it is available. In both cases, a simple estimate of proportion of time spent underwater allows the abundance estimate to be corrected for diving birds. This estimate might be obtained by observing a representative sample of birds, or by placing tags on them that provide data on whether or not they are underwater.

Given the altitude that the aircraft fly at, and the set-up of cameras, detectability does not fall off with distance from the trackline, and for birds at the surface, detectability within the surveyed strip can be assumed to be certain. Thus data are analysed using strip transect methods (Sects. 1.4.1 and 6.2). Computer software is used to detect birds on the images. Species identity might be done by the software, with validation by experts, or by experts checking the images of birds detected by the software.

Trials were conducted on the winter flock of common scoter (*Melanitta nigra*) in Carmarthen Bay in the UK, to compare two digital methods (one using high-resolution video and the other using high-resolution stills) with aerial visual surveys (Buckland et al. 2012). For each method, a systematic grid of transect lines was laid over the bay, and surveyed on four occasions. The digital surveys were analysed as standard strip transect data, while the visual survey was analysed using standard line transect methods. In addition to standard design-based methods, a density surface was fitted to the data for each method. We show abundance estimates from Buckland et al. (2012) in Table 10.1. The two digital methods agree very well, while the visual survey estimates are substantially lower. Common scoter are particularly challenging for visual observers because they occur in large numbers concentrated in a small area, and it is clear that the visual observers were unable to count numbers accurately.

**Table 10.1** Design- and model-based abundance estimates for common scoter in Carmarthen Bay, March 2009, using visual aerial surveys, high-resolution video surveys, and high-resolution stills surveys

Survey type	Design-based $\hat{N}$	% cv	Model-based $\hat{N}$	% cv
Visual	6197	22	7976	38
Digital still	17,501	38	16,490	43
Digital video	18,034	32	15,489	23

## 10.3 Cetaceans

Cetacean surveys have several issues in common with seabird surveys. Relative to terrestrial environments, the marine environment is relatively homogeneous from the perspective of designing a survey, and it is generally straightforward to follow random transect lines. Further, the environment is not fixed, unlike terrestrial environments. Thus whether or not repeat surveys follow exactly the same lines is less of an issue; the correlation between counts taken on repeat visits to a transect on the surface of the sea is likely to be weak compared with that for a terrestrial transect.

Many seabirds and all cetaceans are not continuously available for detection at the surface. When dive times are typically short relative to the time that an animal's location is within detection distance of the observer, mark-recapture distance sampling methods (Sects. 5.4 and 6.4.4) can be used to account for this, especially if the two observers do not search the same area at exactly the same time. However, for aerial surveys, and for shipboard surveys of long-diving species, availability should be addressed (Sect. 11.2).

Cetacean surveys also raise some issues that are less relevant to seabirds. They typically occur at lower densities, so that small sample size is often an issue. They may also only be visible for an instant at each surfacing, which for species that occur mostly as single animals or very small groups, has implications for how observers search for them (*e.g.* search close to the line rather than further out; use naked eye search to give a wide field of view; use multiple observers to search simultaneously).

### 10.3.1 Shipboard Surveys

#### 10.3.1.1 Reducing Bias from Responsive Movement

Cetaceans may respond to the ship. It is important to attempt to detect animals before they respond, which may favour binocular search over naked eye search. For species that give brief opportunities for detection (*e.g.* a whale blow, or a brief surfacing of a porpoise), probability of detection can be enhanced by having a team of observers, using low-power binoculars (Fig. 10.3) or naked eye search (or both) to give a wide field of view. For species that usually give a continuous cue, such as dolphins that typically occur in large groups, observers might use large tripod-mounted binoculars to ensure detection at larger distance (Fig. 10.4); this both improves sample size and helps ensure that groups are detected before they respond to the ship.



**Fig. 10.3** Main observation deck with heated booths for marine mammal observers on the Japanese research vessel Kaiko Maru during the SOWER 2009/2010 expedition into Antarctic waters. Note the low-power binoculars with wide field of view, useful for searching for minke whales which tend to occur singly or in small groups Photo: Cornelia Oedekoven

### 10.3.1.2 Estimating Distances and Angles

Distance estimation at sea is problematic, and it is important both to train observers in distance estimation and to provide aids that allow more accurate estimation. Binoculars with reticles (Buckland et al. 2001, pp. 256–258), allowing measurement down from the horizon (Fig. 10.5), are useful if not essential. A laser rangefinder allows observers to assess opportunistically their ability to estimate distance (but does not work for estimating distance to a point on the sea surface, unless there is a floating object). It is seldom possible to estimate perpendicular distances of animals from the line reliably for shipboard surveys, unless the animals are roughly abeam. Thus observer-to-animal distances are estimated, along with sighting angles from the line, from which perpendicular distances from the line can be estimated (Fig. 1.1). Angle boards (Buckland et al. 2001, p. 263) or angle rings on tripods should be provided for angle estimation.

### 10.3.1.3 Dealing with Cetacean Groups

Several issues may arise when estimating cetacean group sizes. The entire group may not be at the surface at the same time, animals may move around quickly making them difficult to track, subgroups of the same group may be spread out





**Fig. 10.4** Flying bridge setup on the NOAA ship David Starr Jordan during a survey off the California, Oregon and Washington coasts. There were three rotating marine mammal observers and one seabird observer. The inner tripod-mounted binoculars were used by the centre marine mammal observer during chases and the seabird observer for flock counts. Tripod-mounted binoculars with 25 $\times$  or 30 $\times$  magnification have a narrow field of view, and so are unsuited to surveys of cetaceans that occur singly or in small groups that do not give many sighting cues, but they are very effective for detecting schools of dolphins at great distances. Photo: courtesy SWFSC/NOAA

over large areas, and individuals may have different surfacing intervals. Group sizes may be very large so that they would be difficult to estimate even if the other issues did not exist. In addition, different types of cetacean behaviour may give rise to different estimated sizes of the same group by the same observer. For example ‘low swimming’ is an evasive type of behaviour observed in dolphins where surfacing is kept to a minimum to avoid detection. In contrast, ‘running’ is a different type of evasive behaviour where dolphins maximize airtime for maximum speed (Mesnick et al. 2002).

When cetacean groups are composed of several subgroups, it is important to define a clear protocol for field personnel whether to treat the group or the subgroup as the recording unit. For the latter case, distances and group sizes need to be recorded for each detected subgroup. During a survey in the Hawaiian Island EEZ, the target species was false killer whales (*Pseudorca crassidens*), a highly mobile species where subgroups may be spread over tens of kilometres (Bradford et al. 2014). Here, visual and acoustic methods were combined for obtaining distances from the trackline and tracking individual subgroups, while visual observers made



**Fig. 10.5** View through large tripod-mounted binoculars. Lining up the 0 reticle reading with the horizon enables conversion of the reticle reading of the sighting to a distance from the observer (which requires knowing the height of the observer above sea level). Image constructed from two photos for illustrative purposes. Photo: Cornelia Oedekoven

group size estimates for each subgroup. Sperm whales also form groups consisting of multiple subgroups which may be spread over several kilometres. These whales are less mobile but dive asynchronously for periods of up to 90 min. Hence, Barlow and Taylor (2005) implemented 90-min counts of sperm whales so that each subgroup could be detected at least once during the count. During these 90-min counts, five or more observers tracked the location and composition of each subgroup. As the recording unit was group, total group size represented the sum of all subgroup sizes.

Dolphins may occur in mixed-species schools (Fig. 10.6), and may have highly variable school sizes possibly containing thousands of individuals. To deal with these issues, observers during eastern tropical Pacific surveys were calibrated by making independent estimates of the size and composition of calibration schools (Gerrodette et al. 2008). These calibration schools generally consisted of single- or mixed-species dolphin groups and were photographed from an aircraft using aerial photogrammetry methods; hence true school size and composition could be established later, to allow errors in estimates to be quantified.



**Fig. 10.6** Mixed school of eastern spinner dolphins (*Stenella longirostris orientalis*) and north-eastern offshore spotted dolphins (*Stenella attenuata attenuata*) sighted during a cruise in the eastern tropical Pacific. Photo: courtesy SWFSC/NOAA

## 10.3.2 Aerial Surveys

### 10.3.2.1 Visual Surveys

Issues for visual aerial surveys of cetaceans are very similar to those for seabirds (Sect. 10.2.2.1). Distance estimation problems are again less if distance data are recorded by interval, with interval cutpoints delineated by aligned markers on the windows and wing struts. Left-truncation is likely to be necessary to allow for failure to detect animals under the aircraft, unless an aircraft with belly and/or bubble windows is used.

For aerial surveys, a diving response is problematic, as the animals may not be visible by the time the observer scans their location. In this circumstance, availability of animals for detection would be lower than would be the case without disturbance, so that any adjustment for availability, calculated from a separate study for example using tagged animals, may introduce bias.

### 10.3.2.2 Surveys Using High-Resolution Imagery

The methods described in Sect. 10.2.2.2 are also useful for cetaceans. The issue of correcting for diving animals is even more important than for seabirds. If dives are typically very short, then multiple images can allow correction for availability. For

example, two cameras might be mounted under the aircraft, one pointing forwards and the other backwards. If both also take overlapping images separated slightly in time, this can help. For longer diving species, it may be necessary to tag a representative sample of animals, to allow estimation of availability (Sect. 11.2).

### 10.3.3 *Acoustic Surveys*

#### 10.3.3.1 Overview

A somewhat recent development is the use of automated recordings of sounds produced by the animals themselves to estimate their abundance or density. These are referred to as passive acoustic density estimation methods. While in principle these methods can be applied to any sound-producing taxa, they have been developed mostly based on cetacean applications. This is natural because many cetacean species are very hard to survey, spending large portions of their time submerged, yet produce loud and easily distinguishable sounds. For an extensive review of this field, including applications to other taxa, we refer the reader to Marques et al. (2013).

Methods for passive acoustic density estimation of cetaceans might include towed surveys or fixed sensors. The former is a form of line transect sampling while the latter corresponds to point transect sampling. The use of slow-moving devices such as drifting buoys or wave gliders (e.g. Klinck et al. 2012) is a current research topic and combines aspects of both.

Towed passive acoustic line transect surveys might be carried out in their own right (e.g. Lewis et al. 2007) or in combination with a more traditional visual survey (Barlow and Taylor 2005). As long as perpendicular distances of detected animals from the trackline can be obtained, the analysis methods used are the same as for visual surveys. A potential problem is when only a slant distance is obtained, yet a distance projected onto the sea surface is required (see Sect. 9.5). An example of when this can be safely ignored is a sperm whale survey. These are detectable at tens of kilometres, which means that, for most detections, the slant and projected distances are approximately the same. For a beaked whale detectable up to at most a few kilometres, and able to dive to deep waters, this might be a problem; the methods of Harris et al. (in prep.) should then be considered.

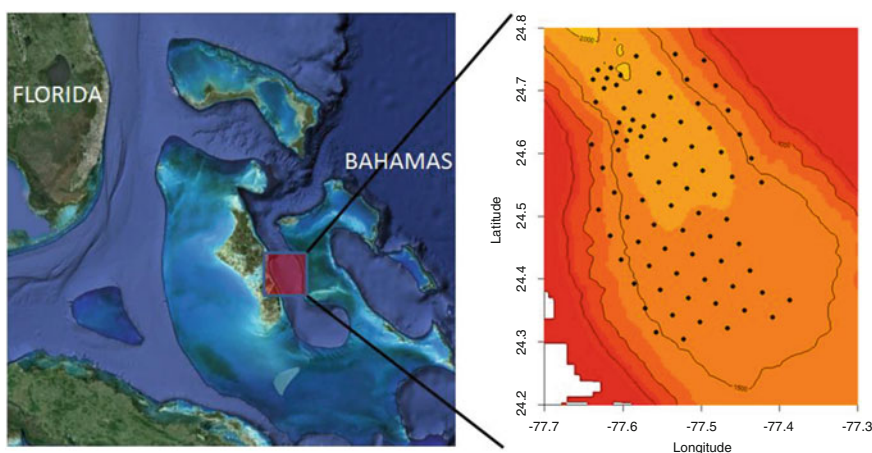
Fixed passive acoustics sensors can be considered as standard point transects, provided distances to vocalizing animals can be estimated based on the sound characteristics alone. This is generally not feasible, although Marques et al. (2011) present an exception with right whales in the Bering Sea. Given the sound propagation characteristics of the shallow continental shelf and the far-travelling right whale ‘up’ calls, distances to detected sounds could be estimated, allowing conventional distance sampling methods to be used. Within a distance sampling setting, if distances cannot be obtained, two options are still available. A theoretical detection function might be derived based on sound propagation models (Kusel et al. 2011), or a detection function might be obtained using auxiliary information, as in

Sect. 9.2. As with any model-based approach, the former method is only as good as the model used. To estimate the detection function using auxiliary information, data might come in a variety of forms from a range of sources (*e.g.* Kyhn et al. 2012). The key aspect to bear in mind is that data should be collected in circumstances representative of survey conditions, otherwise the estimated detection function might not apply, and consequently bias would arise.

Although beyond the scope of this book, acoustic fixed sensors might be used in a spatially-explicit capture–recapture (SECR) approach, where distances to sounds are not required, but identification of the same sound recorded at different sensors is (Marques et al. 2012; Martin et al. 2012). A non-technical overview to SECR is provided by Borchers (2012), while Borchers et al. (2015) provide a conceptual link between conventional capture–recapture methods and distance sampling, which can be viewed as two extremes within a common framework that includes no and full information on animal location, respectively.

### 10.3.3.2 An Example with Beaked Whales

The objective of this study was to estimate the density of Blainville’s beaked whales (*Mesoplodon densirostris*) at a US Navy range facility in the Bahamas (Fig. 10.7). The material in this section is based on Marques et al. (2009), where further details are available. Blainville’s beaked whales are deep-diving odontocetes, which are extremely hard to survey visually, but which produce easily detectable echolocation clicks for foraging. These clicks are produced only during the deep part of their dives. These animals tend to be almost metronomic divers, with deep foraging dives alternating with series of relatively shallow dives during which no clicks are produced (Johnson et al. 2004; Tyack et al. 2006).



**Fig. 10.7** The location of the Atlantic Undersea Test and Evaluation Center (AUTEC) range with an expanded view of the field of 82 hydrophones used for surveying Blainville’s beaked whales

For a period of 6 days, the number of sounds detected and classified as beaked whales in an array of  $K = 82$  bottom-mounted hydrophones was counted. Once pooled across hydrophones, almost three million sounds believed to be beaked whale clicks were detected. Therefore, not surprisingly, the click counting process itself was automated. However, no such automated process is perfect. Hence, a required multiplier was the proportion of false positives (*i.e.* a sound assumed to be a beaked whale click, which in fact was not). Due to difficulties in evaluating some sounds, both optimistic and pessimistic scenarios were considered, *i.e.* assuming all or none of the uncertain sounds were beaked whale clicks respectively. Given an estimated density of clicks per unit time, animal density was obtained by dividing this estimate by an estimated average click production rate, obtained from a sample of animals fitted with acoustic tags (Johnson and Tyack 2003).

The density estimator used was given by

$$D = \frac{n_c(1 - \hat{c})}{K\pi w^2 \hat{P} \hat{T} \hat{r}} \quad (10.1)$$

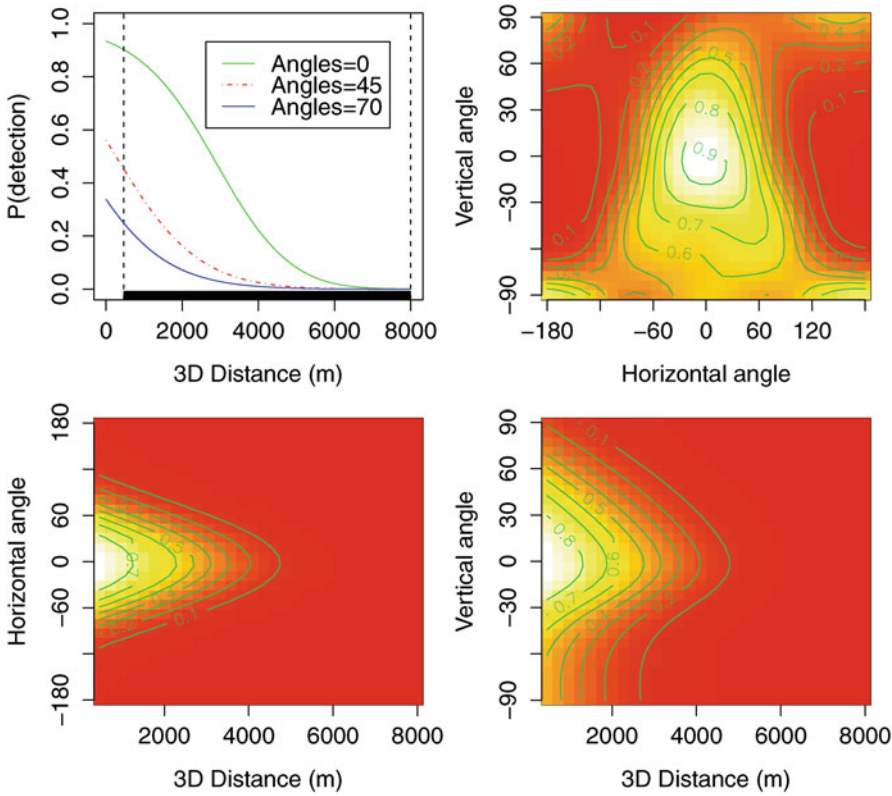
where  $n_c$  is the number of clicks detected,  $w$  is the distance from the hydrophones beyond which no cues are assumed to be detected,  $\hat{P}$  is the estimated average probability of detecting a cue made within a distance  $w$  of a hydrophone,  $\hat{r}$  is the estimated cue production rate,  $\hat{c}$  is the estimated proportion of false positive detections,  $K$  is the number of replicate sensors used, and  $T$  is the time that the hydrophones were operating. Note therefore that  $n_c(1 - \hat{c})$  is the estimated true number of beaked whale clicks detected,  $\hat{r}T$  is the mean number of clicks produced by an animal during the recording period, and  $K\pi w^2 \hat{P}$  is the effective detection area. Additional details, including how the different multipliers were obtained, can be found in Marques et al. (2009). Here we concentrate on describing how detectability was estimated; that is, how the mean probability of detecting a click from a given hydrophone, given that the click was produced within a distance  $w$  of the hydrophone, was estimated.

Because distances could not be obtained for clicks detected on the hydrophones, additional information was required to estimate detectability. A small number of whales was fitted with a DTAG, an archival tag which records depth, 3D accelerometer, 3D magnetometer and acoustic data, at a very high sampling rate (Johnson and Tyack 2003). Using these data, we can both (1) know exactly when the tagged animal produced each one of its clicks and (2) by correlating this timing information with the sounds detected on the field of surrounding hydrophones, we can obtain the three-dimensional position of the animal in space, as well as its orientation. These now provide trials at known locations, allowing us to determine whether or not each click was detected by each of the surrounding hydrophones within a given radius. This radius was defined as the radius beyond which there was a negligible probability of detecting a click. Given the properties of the acoustical setting, this was set at 8 km. These 0/1 (corresponding to undetected/detected) data can be modelled using logistic regression, in which the probability of detecting a



click is modelled as a function of distance and relative horizontal and vertical angle with respect to the hydrophones. Note that unlike what is usual for conventional distance sampling, we modelled detectability as a function of slant 3D distance instead of a distance projected on the surface or on the bottom. The orientation of the whale with respect to the hydrophone was known to be important *a priori* because these clicks are highly directional in their sound intensity (e.g. Zimmer et al. 2005). The logistic regression model was fitted using the `mgcv` R package. The outcome of this modelling exercise was reassuring, returning the patterns expected. The probability of detection of a click decreases with distance to the hydrophone, and with the off-axis horizontal and vertical angle with respect to the hydrophone (Fig. 10.8).

Once this model was obtained, the remaining step was to estimate the mean probability of detection of a click. This involves averaging over the unknown



**Fig. 10.8** Probability of detection as a function of the hydrophone-to-click three-dimensional distance and off-axis angles. Plots are conditional on a given value for the variable(s) not shown: distance and angle were considered to be 0 m and 0°, respectively (*i.e.* maximizing the probability of detection). Vertical dashed lines on the top left plot correspond to the maximum and minimum available distances

variables present in the model (distance, horizontal and vertical angle) which we do not observe for the  $n_c$  detected clicks. To do so, thousands of 3D tracks were simulated by re-running the existing deep-dive 3D trajectories around hydrophones, assuming a random start and orientation for each track. For each such simulated click, its probability of detection was evaluated given the estimated model and the observed covariates. The average probability of detection of a click was then estimated by the sample average of the individual click detection probabilities, and the corresponding variance. The variability in the estimated detectability model was propagated through to the final density estimate using the delta method.

The estimated density was 25.3 or 22.5 animals/1000 km<sup>2</sup>, depending on assumptions about false positive detections, with corresponding 95 % confidence intervals of (17.3, 36.9) and (15.4, 32.9). These values are consistent with previous density estimates for this area (*e.g.* Moretti et al. 2006).

This example, chosen to showcase passive acoustic density estimation, has features in common with two earlier sections. Here, we do not have direct information about detectability in the survey data, and so we conduct trials with animals located at known distances. This approach is also used for trapping and lure point transects (Sect. 9.2). This case study is also an example of cue counting (Sect. 9.4), but one for which we are not able to record distances of cues from the point in the main survey.

## 10.4 Primates

Plumptre et al. (2013) review census methods for primates, including line transect sampling, indirect surveys of nest or dung (Sect. 9.3), and lure point transect surveys (Sect. 9.2). Primate line transect surveys often suffer through lack of replication (a very small number of lines, sometimes just one, each of which is walked on multiple occasions) and randomization (lines placed along convenient routes or close to base), so that there is little or no basis for drawing inference on animal densities in a wider area. Good design practice for primate surveys is addressed by Buckland et al. (2010a).

To implement a randomized survey design, it may be necessary to cut transect lines. In this case, the lines should be cut at least a week before the line is surveyed, to avoid bias from disturbance, and a minimal cut should be used, so as not to create an obvious path for either hunters or animals (Buckland et al. 2010a). Each line may be surveyed more than once (Sect. 6.3.3.1), but the number of replicates for variance estimation is the number of lines in the design; repeat surveys of the same line are not independent, and so cannot be considered replicates.

Estimation of distance of detected animals from the line is often problematic for primate surveys. Primates often occur in groups, which may be spread over some distance. It has been common practice in primate surveys to record the distance of the first animal detected from the line, and then to adopt an analysis of clusters, taking the location of the first animal detected to be the location of the group.



However, the first animal detected tends to be closer on average to the line than other animals in the group, so that distances are biased towards zero (often with many being recorded as exactly zero). This leads to overestimation of density, and some primate surveys have thus been conducted using animal-to-observer distances, together with analysis methods that have no mathematical basis, because it is claimed that such methods have lower bias. Buckland et al. (2010b) review this approach, and clarify the problems with it. A far more satisfactory solution is to obtain better estimates of distances from the line.

### ***10.4.1 Dealing with Primate Groups***

The following is summarized from Buckland et al. (2010a).

There should be a clear protocol so that fieldworkers can determine what constitutes a group for the purposes of the survey. For example, if animals are separated by at least 20 m from the originally detected group, the protocol might state that these should be treated as a second group. This might result in a single large social unit being recorded as many groups. Any of those groups that is detected and whose centre is located within the survey strip of half-width  $w$  should be recorded, and their distance from the line measured or estimated.

Distances of group centres from the line should be measured as accurately as possible. This requires that the position of the line is well defined, so that distances from the line are well defined. Unless distances are sufficiently small to be measured with a tape without undue disturbance or delay, a laser rangefinder should always be used for primate surveys. Vegetation may prevent a direct measurement, but the rangefinder is still an invaluable aid for improving distance estimates (see Sect. 10.1.4).

Primates are often in large, dispersed groups, so that it is difficult to estimate distance except for the individuals first detected. The problem is made worse if the subjects flee from the observer. Where it is impossible to determine location of group centres with sufficient accuracy, but feasible to estimate distances to each detected individual, then you can ignore the existence of groups. Each individual that is detected is recorded, along with its distance from the line (Buckland et al. 2001, pp. 75–76). Standard methods assume that all animals on or very close to the line are recorded, but it does not matter if animals further from the line but in a detected group go undetected.

If it is not feasible to record all detected individuals, together with their distances from the line, then it is important to estimate the size and location of detected groups as accurately as possible. In fact, bias in estimates of the size or location of groups well away from the line need not be problematic, but for groups on or close to the line, bias should be as small as possible. A field protocol should be developed with these issues in mind. For example, if animals do not respond to observers, observing the group for a period of time from different locations on and off the line may allow an accurate assessment of size. If animals do respond, a quick count may be needed,

and multiple observers with slightly different vantage points, and a well-rehearsed protocol for coordinating their count, may be effective.

If neither of these strategies is achievable, it may be necessary to estimate mean group size and spread in a separate study from the line transect survey. In this case, the study should be conducted synchronously with the line transect survey, to ensure that the observed mean size and spread is representative of groups in the survey region at the time of the survey. The criterion for defining what constitutes a group must also be consistent between the line transect survey and the study to estimate group size. Problems with this approach include:

- It may be difficult to achieve an adequate sample size — at least 10, and preferably nearer 20 — especially if group size is very variable.
- If only habituated groups can be monitored in this way, they may not be representative of all groups.
- It is still necessary to estimate the location relative to the line of groups detected during the line transect survey.

To address the last point, it may be necessary to record the distance to the closest individual, whether it is closest to the line or to the observer, and correct either the recorded distances or the effective strip half-width to allow for group spread (Whitesides et al. 1988; Buckland et al. 2010a).

### 10.4.2 *Other Approaches*

Buckland et al. (2010a) suggested other distance sampling strategies for when standard methods fail or are impractical. We summarize those suggestions here.

If it is feasible to record each individual that is detected, together with its distance from the line, but it is thought that some individuals on the line are missed, it may be possible to conduct trials by locating individuals, perhaps using radio collars, and then sending observers who are ignorant of animals' positions past those at a known closest distance of approach. These trials generate binary data which may be modelled using logistic regression, with distance from the line and possibly other variables as covariates, from which the probability of detection on the line may be estimated. This estimate and its standard error may then be included as a multiplier in Distance, when analysing the line transect survey data. Similarly, if groups instead of individuals are recorded, but some groups on the line may be missed, trials might be set up involving the group instead of an individual.

The cue counting approach for birds (Sect. 9.4.2) may also work for primates that call. Cue rate (number of calls per animal per unit time) should be estimated in a synchronous survey, to allow conversion from number of calls per unit area per unit time to estimated animal density. The disadvantages of this approach are that it can be difficult to estimate distances to calling animals, and it is difficult to ensure that a representative sample of animals is monitored to estimate the cue rate.

If animals can be lured in by playing a call, then lure strip transects may be possible, as implemented in a study of cotton-top tamarins (*Saguinus oedipus*) (Savage et al. 2010). Observers simultaneously travel along two parallel transects, luring animals from within the strip between the transects. If the lure causes animals to respond by calling, but does not attract them in, a line transect version of this approach might be workable, with just one transect at each location. If several observers are positioned along the line, distances of responding groups from the line may be estimated by triangulation. Another possibility is lure point transects (Sect. 9.2), in which trials are conducted on subjects with known location, from which a detection function is estimated, and assumed to hold for the main survey, where a lure is played at each of a number of points systematically spaced through the survey region.

## 10.5 Ungulates

In this section, we are primarily concerned with surveys of deer and antelope, although the methods are also relevant to many other large terrestrial herbivores. In remote areas, the methods might be used to estimate livestock abundance, or the abundance of introduced populations of for example camels (*Camelus dromedarius* and *Camelus bactrianus*) and donkeys (*Equus asinus*) in Australia.

Deer and antelope are typically vigilant and can travel at high speeds. Consequently investigators employ a variety of strategies to cope with detecting such wary animals. Many ungulates can occur at low densities ( $< 1 \text{ km}^{-2}$ ) and range over large areas such that classical sampling techniques may be inadequate for producing defensible density estimates.

Challenges with ungulate studies include the following.

- Non-random transect placement.
- Responsive movement before detection.
- Cluster size estimation.

We consider these challenges in the context of ground-based and aerial surveys below.

### 10.5.1 Ground-Based Surveys

Ground-based surveys of ungulates often make use of vehicles on roads or tracks. This leads to violation of the assumption of random placement of transects (Sect. 11.1). Ungulates may have atypical densities along roads, for example due to disturbance, hunting pressure, or availability of a different habitat (the track, or perhaps forest edge). Further, roads are unlikely to be placed independently of topography (they avoid rugged terrain) or habitat (they favour dry ground), which

may create larger-scale differences in density between areas near roads and areas far from roads. A recent study by McShea et al. (2011) suggests bias in estimates of deer density can arise from non-random transect placement. Track-based surveys may need supplementary data, *e.g.* from additional transects perpendicular to tracks (Marques et al. 2013), to allow unbiased estimation of density.

If ungulates are too wary to be detected from the ground during the day, nocturnal surveys with spotlights or thermal imagers are one way of circumventing detection of the observer by the animals prior to detection of the animals by the observer (Gill et al. 1997; Marini et al. 2009; Focardi et al. 2013). Indirect methods (Sect. 9.3) can also be employed to circumvent issues of responsive movement (Sect. 11.4.2): pellet groups are surveyed to give pellet density estimates, which may be converted to estimates of animal density provided we have estimates of deposition and decay rate (see for example Marques et al. (2001); Acevedo et al. (2010); Alves et al. (2013)). The challenge with indirect methods is the need to estimate the deposition and decay rates that apply within the survey region in the lead up to the pellet survey (Laing et al. 2003).

Some species occur in large herds. In this circumstance, it is necessary to develop a field protocol for determining a ‘cluster’. This might be taken to be a sub-group of animals within the larger herd for which a distance from the line can be estimated, along with the sub-group size. There is then no need to detect all sub-groups in the larger herd, although all those on or very close to the line (before responsive movement occurs) should be recorded.

In a study of tiger (*Panthera tigris*) prey in Nepal, Wegge and Storaas (2009) found that, when surveys for ungulates were carried out on foot, unseen and unidentified animals were heard fleeing, violating a key assumption. Surveys carried out by vehicle were biased because of the need to use tracks. By contrast, they found that surveys conducted from the backs of elephants (*Elephas maximus*) worked well.

### 10.5.2 Aerial Surveys

If it is impossible to follow random transects using ground-based surveys, then aerial surveys should be considered. Aerial surveys can also be effective for dealing with the problem of responsive movement, especially if observers have a good forward view, so that they can record the position of detected animals before responsive movement occurs. Either fixed-wing aircraft (Johnson et al. 1991; Guenzel 1997; Whittaker et al. 2003) or helicopters (White et al. 1989; Trenkel et al. 1997) might be used.

Detectability of animals near the flight line may be compromised by the configuration of the aircraft, although bubble or belly windows (Sect. 4.1.2.2) can be employed. Often in aerial surveys, observers search just one side of the aircraft, and the line is offset perhaps 30 m or so, to avoid the difficulty of detecting animals under the aircraft. In this case, the observers should be carefully trained, so that animals detected close to the line are recorded on the correct side. Given some uncertainty over animal location (especially if it is moving), and given observers’ inclination to

err on the side of recording animals over which there is any doubt, failure to define a clear protocol, and to ensure that it is correctly implemented, can generate a spike in the histogram at around zero distance for one-sided or offset transects, which leads to overestimates of density.

Double-observer methods (Sect. 5.4) can prove useful for aerial ungulate surveys, allowing estimation of  $g(0)$  (Fewster and Pople 2008).

Aerial surveys may allow more reliable estimation of cluster sizes than is possible from ground-based surveys, especially if high-resolution imagery is used. However if views of the ground are obstructed by foliage, then methods of estimating the proportion of animals available to be seen (Sect. 11.2) might be needed.

## 10.6 Butterflies

The UK Butterfly Monitoring Scheme (Pollard and Yates 1993) comprises a set of sites that are surveyed by walking a transect, and counting numbers of butterflies by species that are detected within an imaginary box ahead of the observer, making the method a form of strip transect sampling. It is often called a ‘Pollard walk’. Originally, all sites were selected subjectively, at least in part because they contained good butterfly habitat, and the transect was selected to pass through the best habitat in the site. More recently, random sites have been added to the scheme, through the Wider Countryside Butterfly Survey. The field methods of the scheme have been adopted by a number of other butterfly surveys. However, there are potential problems with the method.

The first problem is that sites are chosen because they are good butterfly habitat, and then transects are placed through the best butterfly habitat within those sites. This has the potential for bias either way: butterfly trends in the wider countryside may be less favourable than in the best habitats; and, once defined, the transect route at a site stays constant, while the distribution of butterfly habitat in the site may change over time — as the transect becomes less effective at sampling the best habitat, downward trends in butterfly numbers might be observed. The wider countryside sites are an attempt to address both issues: they are selected according to a stratified random sampling scheme, and the ideal transects are two parallel straight lines, each 1 km in length. However, neither advantage is fully realised. The wider countryside sites were actually selected for surveying breeding birds, and the volunteers who cover the sites have the option of recording butterflies too. Some self-selection can be expected, with volunteers more likely to do the butterfly survey work in better sites for butterflies. Also, it is usually not possible to access all of the ideal transect routes, so a route approximating it is selected. In arable habitats especially, the actual route mostly follows field edges, where butterflies are more abundant. This is less of a problem in the case of breeding birds, because a wide strip, extending well into other habitats, is surveyed.

In many environments, this problem is essentially insoluble. If it simply is not possible to follow random transects, and the observers are forced to follow edge

habitat, then there is little basis for estimating absolute abundance. The hope then is that counts in edge habitat reflect trends in relative abundance over time. Further, even if it were possible to walk through arable crops for example, few butterflies would be recorded, which in turn would reduce the number of volunteers willing to conduct the surveys, and too few data might be gathered to assess trend.

The second problem is that in better conditions, more butterflies will be flying. This both makes them more detectable and generates upward bias in counts for the same reason that fast animal movement causes bias in line transect sampling. For any given butterfly at a site, the faster it moves, the more likely it is to enter the imaginary box ahead of the observer. Provided there is no trend in weather conditions, this will add noise to the counts and introduce bias to estimates of absolute abundance, but will not bias estimated trends in relative abundance. The methods attempt to reduce noise and bias from this source by restricting surveys to when the weather is favourable for butterflies.

In closed habitats, such as rain forest, distance sampling methods are likely to be of limited use, at best giving estimates of relative abundance, and at worst simply failing to detect sufficient butterflies to allow analysis.

If a survey is not dependent on volunteers, if random transects can be walked, and if the habitat is open, it should be possible to carry out more rigorous surveys, allowing absolute abundance to be estimated. Often, brood size is of greater interest than abundance at any one point in time; in this case, some modelling of the abundance estimates is required to convert them to estimated brood size.

Brown and Boyce (1998) developed a line transect protocol for the endangered Karner blue butterfly (*Lycaeides melissa samuelis*). In a report on how to conduct surveys of the Karner blue, Grundel (2008) gave an excellent summary of the pros and cons of line transect sampling for assessing butterfly numbers. Isaac et al. (2011) compared the Pollard walk with standard line transect sampling, implemented at 13 sites in England and Wales. They found that on average, one third of butterflies in the box were not recorded using Pollard walk methods. This proportion varied by species; for the most detectable species, no individuals were missed, while for the least detectable, around three quarters were missed. The authors concluded nevertheless that the Pollard Walk was the more practical method for large-scale volunteer surveys. They did not address bias arising from butterfly movement, or from the non-random placement of transects.

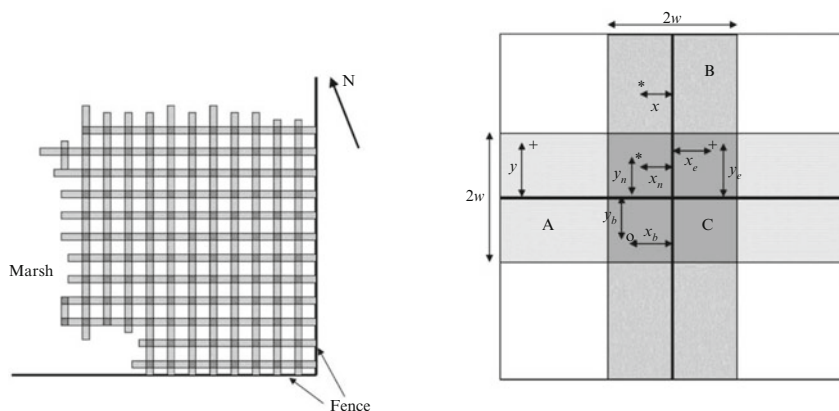
## 10.7 Plants

Most distance sampling surveys of plants are straightforward. Plants do not change their location while the survey is taking place, and it is simple to measure distances from the line or point. Schorr (2013) for example used standard line transect sampling to estimate abundance of Weber's saw-wort (*Saussurea weberi*) at a location in the Rockies. He also discussed the merits of line transect sampling relative to plot sampling.

However, surveys for plants are not always as straightforward. First, it may not be possible to identify individual plants if they spread and intermingle. This may render distance sampling irrelevant, and plot sampling based on percent cover of each species may be more useful. However, for some species, it may be possible and useful to estimate density of flower stalks.

A second issue is that, for many species, it may not be possible to ensure that all plants on the line or point are detected, as some individuals may be small and/or hidden by other vegetation. This problem is made worse because the probability that a plant on (or off) the line or point is detected can be very variable, depending on the size of the individual plant, whether it is flowering, how much other vegetation is present, and so on. A further issue is that surveys are sometimes conducted at small sites holding a colony of plants of interest. In the case of line transect sampling, some sites may be too small to allow an adequate number of non-overlapping transects. Especially because the distribution of many plant species is highly clustered, this can mean that there are too few lines to ensure that plants are uniformly distributed with respect to distance from the line, even when a systematic grid of lines is randomly superimposed on the site.

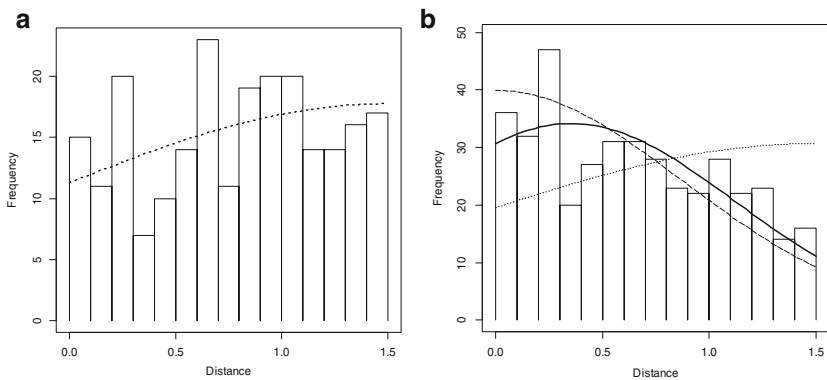
Buckland et al. (2007) proposed a design and analysis to address all but the first of these issues. Two systematic grids of transects were placed at random over the study site, one at right-angles to the other (Fig. 10.9). The key to estimation is then



**Fig. 10.9** *Left*: design of the Fleecefaulds survey, which comprises two systematic grids of strips, one with strips running approximately N/S, and the other with strips running E/W. The strips are of width  $2w$ , and the transect lines run down the middle of each strip. Each grid is randomly superimposed on the survey region. *Right*: schematic enlargement of a crossover of perpendicular transects. Detections made from the E/W line only are indicated by +, from the N/S line only by \*, and from both lines by o. Within region A, distances  $y$  from the E/W line are recorded. Within region B, distances  $x$  from the N/S line are recorded. Within the intersection square C, both  $y$  and  $x$  are recorded; the subscript  $e$  indicates detected from the E/W line only,  $n$  indicates detected from the N/S line only, and  $b$  detected from both lines. The range for each of  $x$  and  $y$  is  $[0, w]$ ; that is, we record absolute distance from each line, and we truncate observations at distance  $w$ . Reproduced from Buckland et al. (2007). ©The International Biometric Society

the squares formed by the intersection of a line from one grid with a line from the other grid. Plants that are detected during the survey from the first grid of lines are marked so that they are uniquely identifiable, but so that the mark does not affect their detectability (*e.g.* a mark on the underside of a leaf), and then plants detected from the second grid of lines are checked for marks. Thus mark-recapture data are available, so that MRDS methods can be used to analyse the data from intersection squares. Further, for any plant detected within an intersection square, whether recorded from both lines or just one, we can calculate its distance from each line (Fig. 10.9). If plants are uniformly distributed with respect to distance from the line, then plants detected from one line should be uniformly distributed with respect to their distances from the line perpendicular to it. If they differ significantly from a uniform distribution, then we can model the actual distribution, and thus eliminate the assumption of uniformity. Further, any covariates recorded on the plant, such as number of flowerheads, plant size, vegetation height, *etc.*, can be used to model heterogeneity in the probability of detection of individual plants.

This approach was used to estimate the density of cowslips (*Primula veris*) at Fleecefaulds Meadow in Fife, Scotland. Full modelling details are given by Buckland *et al.* (2007). Figure 10.10 shows that there is some evidence of non-uniformity of cowslips with respect to distance from the transect, with more plants on average towards the outer edge of the surveyed strips. A summary of density estimates is given in Table 10.2. We see that, if we assume that all plants on the line are detected, it makes very little difference whether or not we include covariates; this



**Fig. 10.10** (a) Histogram of the combined distances from the N/S lines of plants detected from the E/W lines, and distances from the E/W lines of plants detected from the N/S lines for cowslip data. The dotted line is the estimate of  $\pi(\cdot)$  (the availability of plants, as a function of distance from a line), scaled to have the same area as the histogram. (b) Plot of scaled estimated detection probability (dashed line), estimated  $\pi(\cdot)$  (dotted line), and estimated probability density function for observed distances (solid line) for cowslip data. The histogram shows number of detections by distance from the line from which they were detected. All functions have been scaled to have the same area as the histogram. Reproduced from Buckland *et al.* (2007). ©The International Biometric Society



**Table 10.2** Estimates of density (plants  $\text{ha}^{-1}$ ) of cowslips at Fleecefaulds Meadow (95 % confidence intervals in parentheses)

	Estimator	Density estimate	95 % CI
Conventional	$\hat{N}_1$	1048	(689, 1593)
distance sampling	$\hat{N}_2$	873	(568, 1342)
Conventional distance	$\hat{N}_1$	1052	(696, 1591)
sampling with covariates	$\hat{N}_2$	877	(574, 1341)
Allowing for	$\hat{N}_1$	1507	(1105, 2256)
$g(0) < 1$	$\hat{N}_2$	1258	(929, 1920)
Allowing for $g(0) < 1$ and	$\hat{N}_1$	1758	(1219, 2710)
non-uniform availability	$\hat{N}_2$	1467	(1032, 2278)

Estimator  $\hat{N}_1$  assumes that density on the strips is representative of the entire study area, whereas estimator  $\hat{N}_2$  uses data from the cross-strips to correct for any difference in density on and off surveyed strips. See Buckland et al. (2007) for details

is due to the pooling robustness property of conventional distance sampling methods (Sect. 5.1.1). When we allow for missing plants on the line, density estimates increase by around 44 %. When uneven density within the surveyed strips is allowed for, there is a further increase in abundance estimates by around 17 %. However, if we correct for estimated differences in density on surveyed strips and off strips (estimated from the cross-transect data, estimator  $\hat{N}_2$ ), density estimates decrease by around 17 %.

## Chapter 11

# Exchanging Assumptions for Data

Distance sampling has been referred to as a hybrid approach because it typically includes a design-based component, which deals with encounter rate, and a model-based component, which deals with detection probability. For populations occurring in clusters or groups, the group size estimation can be based on design (in the absence of size bias) or on a model (*e.g.* size-bias regression or multiple-covariate distance sampling with group size as a covariate). A model can be thought of as a set of assumptions. If we do not have all the data that we would like, we may need to make additional assumptions. As an example, if we are estimating the abundance of chimpanzees using an indirect method via nest surveys, in the absence of a local estimate of nest decay rate, we might *assume* that the decay rate for another place and time is valid for the current survey. This trade-off between data and assumptions can also be used in the opposite direction. If we have additional data above and beyond that required for conventional distance sampling (CDS) methods, this might allow us to relax our assumptions.

As noted in Sect. 1.7, there are several fundamental assumptions that should hold to a reasonable approximation to allow conventional distance sampling to be used: one key design assumption and three key model assumptions. These four assumptions are:

1. Animals are distributed independently of the lines or points.
2. Animals on the line or at the point are detected with certainty.
3. Distance measurements are exact.
4. Animals are detected at their initial location.

In the following sections, we address each of these in turn, describing how we might collect additional data to address assumption violation.

If it is feasible to design the survey and field methods so that the key assumptions are met, this is generally far more satisfactory than attempting to solve problems arising from assumption failure at the analysis stage. Methods which accommodate

assumption violation are necessarily more complex and harder to implement and, because they rely on estimating additional parameters, resulting density estimates will tend to be less precise than if CDS methods were possible.

## 11.1 Non-random Transect Placement

### 11.1.1 Bias

The advice with respect to number and placement of transects is clear: a large number of transects placed independently of the animal distribution ensures that the distribution of all distances, irrespective of whether or not animals are detected, is known, and hence that CDS methods can be used. The notion of large is not precisely defined because it depends on the animal distribution itself. If density were very high and the animals distributed at random with even density throughout the study area, a single line or point would suffice. Of course, this ideal never occurs in reality. A far more likely scenario is that density is low and animals are distributed patchily, in which case many lines or points might be required. See Sect. 2.4 for additional guidance on this issue. We recommend 12–15 transects as a bare minimum, and for most surveys, above 30 transects should be more than sufficient to ensure that, on average, animals are distributed independently of the lines or points. If a small number of transects is used, for any given realization of the design, the chance is high that the distribution of animals with respect to the line or point deviates markedly from what is assumed, which can generate considerable bias. Buckland et al. (2007) presented an example of such a design with a small number of transects to survey cryptic plants (Sect. 10.7). The novelty there was a crossed design (*i.e.* a regular grid of  $N$ – $S$  transects and  $E$ – $W$  transects). The information along the  $N$ – $S$  transects was used to estimate the distribution of plants with respect to the  $E$ – $W$  transects, and *vice versa*, allowing density estimation by modelling the observed distribution rather than assuming a known distribution of plants with respect to distance from the lines.

If transect locations are not randomized, large bias might result, and the size of the bias will not necessarily decrease as the number of transects increases. If for example lines are placed along some landscape feature, and animal density is atypical along this feature, then CDS methods, which rely on the assumption that animals are located independently of transects, may give very biased estimates of abundance. Only if it simply is not feasible to place transects at random should designs be considered that place transects according to landscape features. The most pervasive example is the placement of transects along roads or tracks (Marques et al. 2010). These might attract animals by providing for example edge habitat, open space and grit, or deter animals due to disturbance, or by providing access for hunting or poaching. Other examples include the use of vantage points (Vargas et al. 2002), land-based transects for surveying marine mammals (which will typically

respond to depth and/or distance to shore) (George et al. 2004; Arranz et al. 2014), and sampling along rivers (Fletcher and Hutto 2006).

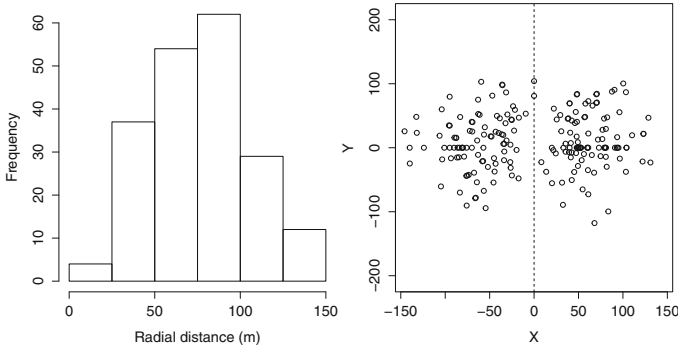
Sampling along such linear features creates two problems:

1. The distribution of animals with respect to distance from the line or point is not uniform (lines) or triangular (points) when density is a function of distance from the corresponding feature. Thus for CDS methods, we obtain a biased estimate of the detection function, and hence of animal density. Below we present approaches that might be helpful under this context, by explicitly modelling density in the vicinity of the feature used for laying transects and incorporating such a model in the distance sampling likelihood.
2. Density in the vicinity of the feature may be unrepresentative of densities well away from it, where we cannot sample. For example, a bird species might increasingly avoid roads as traffic volume and disturbance increase, yet maintain or even increase its density over a large area where no roads exist. A possible way to move forward under such a setting would be to develop a model for density as a function of distance from the features used for laying down transects (and other additional covariates when available). For this alternative to be robust, some sampling away from the features used for the primary sampling, covering the range of relevant covariates, would be required.

Thus inferences drawn from data obtained from transects placed along linear features should be interpreted with extreme caution. It is often not possible to detect non-uniform density with respect to distance from the line or point by examining the observed distances to detected animals. If animals avoid the linear features used for sampling, provided the avoidance is over small distances where probability of detection is high, then a lack of detections at small distances will suggest a problem. However, if animals are attracted to the linear features, the data will give little indication of the problem. This is because the net effect of detectability and availability, as estimated by the histogram of detection distances, will tend to generate a pattern that could be interpreted as having resulted from detectability alone.

### 11.1.2 *Correcting for Bias*

Figure 11.1 shows data from a mountain hare (*Lepus timidus*, Fig. 11.2) survey, where point transects were placed along roads (Marques et al. 2010). The distance data show a lack of detections in the first interval (left-hand plot). The problem is more apparent if we plot the locations of detected animals relative to the points in two-dimensional space (right-hand plot), which was possible here because angles from the road were recorded along with distances. This plot shows clear evidence that these animals avoid the road, and illustrates the merit of the advice offered by Burnham et al. (1980): angle data from point transect surveys can help to identify potential problems.



**Fig. 11.1** Observed radial distances (*left panel*) and sighting locations with respect to points located along roads for a mountain hare survey. The road is represented by the *dashed line*, and the point is at  $(X, Y) = (0, 0)$



**Fig. 11.2** The mountain hare is an example of a species that avoids roads. Photo: Steve Buckland

If point transects are located along a linear feature and angles to detected animals are recorded, as for the hares, we can jointly model these data in a way that allows us to estimate the availability of hares as a function of distance from the road. Provided avoidance of roads is restricted to a distance at which the probability of detection is still markedly above zero, we can separately estimate the availability and detection functions (Marques et al. 2010).

In the case of line transect sampling, we cannot explicitly model the availability as a function of distance from the road given the observed detection distances  $x_i$  ( $i = 1, 2, \dots, n_1$ ) alone, as information on detectability and availability for detection is fully confounded. However, if independent information about the distribution of animals with respect to the survey lines were available, then we could incorporate it in the estimation process to obtain unbiased density estimates.

Using distances  $x_i$  ( $i = n_1 + 1, n_1 + 2, \dots, n_1 + n_2$ ) from the primary transect, either collected along secondary transects perpendicular to the primary transect, or obtained from an independent source (Marques et al. (2013) used a sample of GPS-collared animals), we can directly estimate the distribution of animals with respect to distance from the transects. Here, we consider such a distribution to be indexed by  $\lambda$ , and assuming a model for  $\pi(x|\lambda)$ , we can estimate this parameter by maximizing the following likelihood:

$$\mathcal{L}_\pi = \prod_{j=n_1+1}^{n_1+n_2} \pi(x_j|\lambda). \quad (11.1)$$

Therefore, instead of using a likelihood based on the probability density function defined by Eq. (5.13), we can consider an extended likelihood based on Eq. (5.12) and an additional source of data to estimate  $\pi(x)$ , allowing us to disentangle the product of  $\pi(x)$  and  $g(x)$  in Eq. (5.12). Proper design and survey methods will ensure that  $\mathbf{x}_1 = (x_1, \dots, x_{n_1})'$  and  $\mathbf{x}_2 = (x_{n_1+1}, \dots, x_{n_1+n_2})'$  are independent, allowing the joint estimation of the parameters of both processes:

$$\mathcal{L}(\theta, \lambda | \mathbf{x}_1, \mathbf{x}_2) = \mathcal{L}(\theta | \lambda, \mathbf{x}_1) \mathcal{L}(\lambda | \mathbf{x}_2) = \prod_{i=1}^{n_1} \frac{g(x_i|\theta) \pi(x_i|\lambda)}{\int_0^w g(u|\theta) \pi(u|\lambda) du} \prod_{i=n_1+1}^{n_1+n_2} \pi(x_i|\lambda). \quad (11.2)$$

We can maximize this likelihood to estimate the parameters, and hence estimate the probability  $P_a$ :

$$\hat{P}_a = \int_0^w \hat{g}(u|\theta) \hat{\pi}(u|\lambda) du \quad (11.3)$$

which leads to the corresponding estimator of abundance or density using a Horvitz–Thompson-like estimator (Eq. (1.5)).

### 11.1.3 Example: Eastern Grey Kangaroos

We illustrate the methods for accounting for a density gradient with respect to the transects with an eastern grey kangaroo (*Macropus giganteus*) dataset fully

described and analysed in Marques et al. (2013). These animals were expected to respond to the roads that were used for sampling.

Two characteristics made this dataset particularly useful:

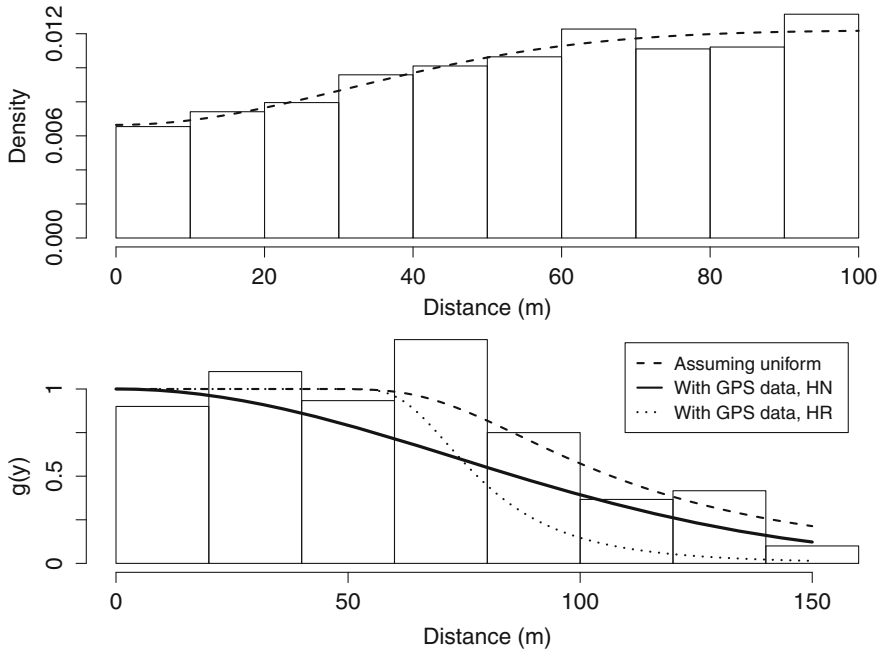
1. Information from a sample of GPS-collared animals was available, allowing us to obtain a direct sample from  $\pi(x)$ , hence providing the means to separate availability from detectability.
2. It corresponds to a closed nature reserve of size 178.6 ha in Australia, for which the true number of animals was available as a total count. Two counts of 516 and 517 animals were made on successive days, corresponding to a density of about 2.89 animals  $\text{ha}^{-1}$ . Therefore, we could directly evaluate the impact of either ignoring or modelling the non-uniform distribution of animals with distance from the road transects.

Adopting a conventional distance sampling approach, a set of three plausible models for the detection function was fitted in software *Distance* with a truncation distance of 150 m, chosen after exploratory data analysis. The best model based on AIC was the hazard-rate with no adjustment terms (Fig. 11.3, bottom panel), and the usual goodness-of-fit measures showed no indication of a bad fit. The estimated probability of detection of animals on the strip of half-width 150 m was 0.73 (95 % CI 0.67, 0.81), with a corresponding density estimate of 2.3 animals  $\text{ha}^{-1}$  (95 % CI 1.60, 3.31). This estimate is rather low, although the confidence interval covers the true density.

Using the GPS data, we modelled the distribution of animals with respect to the road using a function based on the Gaussian distribution suggested by Marques et al. (2010). We note that one could attempt other models, and perform model selection to choose the most parsimonious from a set of candidates. However we do not attempt this here because it is not the main purpose of this case study and the fit obtained with this function looks excellent (Fig. 11.3, top panel).

Using the formulation of Eq.(11.2) to allow for the density gradient, we estimated parameters for the detection function using either a hazard-rate or a half-normal model. Interestingly, when assuming uniform availability, the best detection function required two parameters. By contrast, when accounting for the density gradient, the best detection function was the half-normal, with just one parameter (Fig. 11.3, bottom panel). More animals were located further from the roads, so that it is unsurprising that the detection function fitted by assuming uniform availability has a thicker tail. When the gradient in animal density is accounted for, the detection probability decreases more rapidly and estimated animal density increases. For the preferred half-normal detection function, we obtain a density estimate of 2.82 animals  $\text{ha}^{-1}$ , very close to the true density of 2.89.

See Marques et al. (2013) for a more detailed discussion. Here, we note one additional issue of interest. These animals occur in small clusters, and size bias was apparent. Therefore, using a biased detection function not only affects the estimation of detectability, it potentially affects the estimate of mean cluster size, when we use the regression method to adjust for size bias (Sect. 6.3.1.3). In this case, assuming uniform availability, mean cluster size was estimated to be 1.94 animals, while



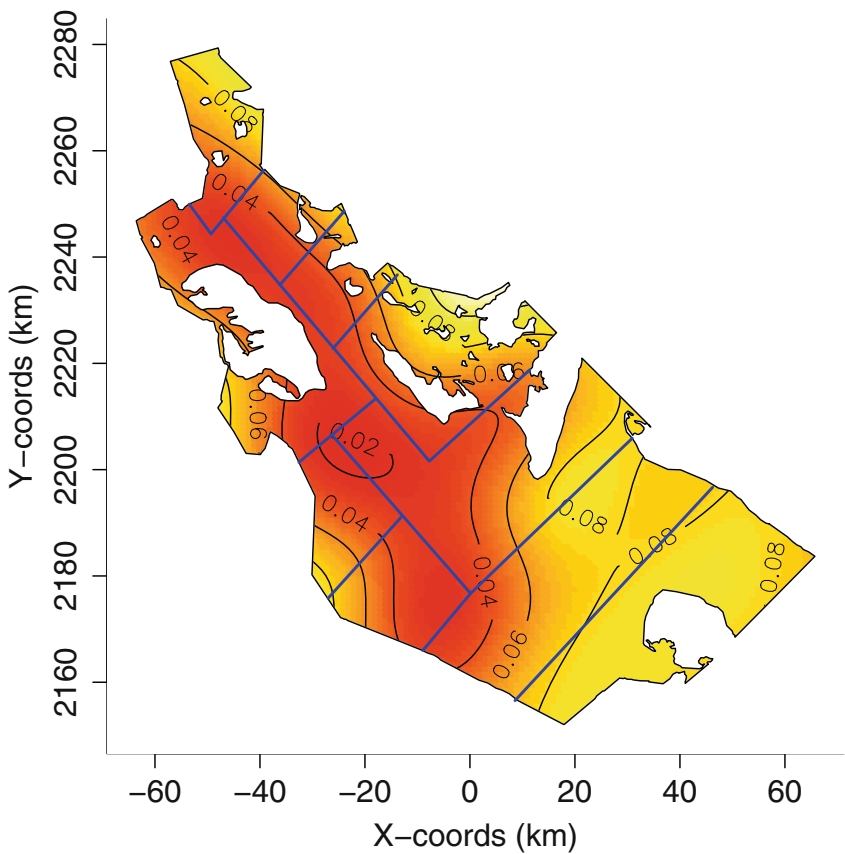
**Fig. 11.3** Estimated distribution of animals with respect to the roads based on a sample of GPS collared animals (*top*); and distances of detected animals from the line (*bottom*). Also shown in the bottom plot are fitted detection functions. One was fitted using CDS (assuming availability is uniform), and the other two accounted for the density gradient, using both a half-normal (HN) and a hazard-rate (HR) model. Note that the histogram bars are scaled such that the area under the bars is the same as the area under the detection function that assumes uniform availability

using the half-normal model accounting for the density gradient, we obtained 1.75 animals per group. The difference in mean cluster size contributed to the difference in density estimates between the two methods.

### 11.1.4 Estimating Bias by Simulation

We return to our simulation example of Sect. 2.5.2 to assess the bias that can arise from using non-random transects. We again generate the population from the density map of Fig. 2.6. Suppose the design comprised paths together with additional transects selected to give relatively even cover through the study area (Fig. 11.4). This design appears good, yet Table 11.1 reveals that the resulting estimates of abundance are biased low by over 15%. From Eq. (2.5), we have  $|\hat{N} - N| = |1265 - 1500| = 235$ , while  $1.96\text{se}(\hat{N}) = 1.96 \times 203 / \sqrt{100} = 40$ . As 235 is much greater than 40, and hence  $N$  is not within the 95% confidence interval, the mean estimate of 1265 differs very significantly from the true abundance of 1500.





**Fig. 11.4** A subjective design comparable with the random designs of Fig. 2.7. Also shown is the density map from which the simulated population was generated

**Table 11.1** Comparison of simulation results for subjective and randomized designs

	Subjective design	Systematic parallel design	Random parallel design	Zigzag design
Mean abundance estimate	1265	1478	1468	1504
Estimated percent bias	−15.6	1.4	2.1	0.2
Mean s.e. of abundance estimate	230	210	205	172
Std dev of abundance estimates	203	208	219	115

All estimates were based on 100 simulations. True abundance was  $N = 1500$

Careful comparison of the transects in Fig. 11.4 with the density map suggests that the darker shades, corresponding to lower animal densities, are oversampled relative to the lighter shades, generating this downward bias. In real examples, we do not know the true distribution of animals; randomized designs ensure that we can avoid bias despite our lack of knowledge.

### 11.1.5 Other Approaches

As noted above, Marques et al. (2010) developed an approach which allowed the estimation of both detectability and availability for detection based on a design consisting of point transects laid along linear features which might affect the distribution of animals. Observers recorded the usual distances from the point of detected animals, but they also recorded the sighting angles with respect to the linear feature, allowing detectability and availability to be separated. Intuitively, a preponderance of small (*i.e.* near  $0^\circ$ ) and large (*i.e.* near  $180^\circ$ ) detection angles with respect to the linear feature indicates a distribution of animals concentrated along the linear feature, whilst an excess of angles near  $90^\circ$  (*i.e.* perpendicular to the road) would suggest avoidance of the feature. Unfortunately, Marques et al. (2010) showed by simulation that the methods are sensitive to the choice of model considered for  $\pi(x)$  (which represents the availability of animals as a function of distance  $x$  from the linear feature). Thus, even in this situation, we recommend that a sample from  $\pi(x)$  is obtained directly from an independent source.

These methods have been extended for two different contexts. The first is for active sonar surveys of krill, for which the animals show a density gradient with respect to depth, while the ship-based radar traverses a transect at the surface (Cox et al. 2011). As with the hares, the angles of detection are the key for disentangling detectability and availability. An additional complication is the need to account for different depths along the track line. This is handled by incorporating in the likelihood an additional model for depth itself. The second context is estimation of the distribution of whales with respect to shore using land-based observations (Arranz et al. 2014), where animal distribution is not independent of distance from the shore. The distribution was modelled as a function of depth, allowing once again the availability and detectability components of the observed distribution to be separated.

We note that these are advanced methods which require bespoke code for implementation. The `nupoint` R package (Cox et al. 2013) allows the implementation of some of these methods. However, because of the added complexity of the methods, the difficulties related to implementation, the requirement of additional assumptions, and the potential sensitivity to models used, our recommendation is that non-random transects should not be used except as a last resort.

We conclude this section with a take-home message. We recommend that surveys are always designed to avoid sampling along linear features that may affect animal density. If this is not possible, then data should be recorded to allow estimation of density gradients, using the methods of this section. If against advice, sampling is conducted along linear features such as roads, and no attempt is made to estimate density gradients, then there should be careful consideration of the degree of bias that may arise, and whether this might compromise the objectives of the study.

## 11.2 Uncertain Detection on the Transect

As we have seen, CDS methods will give abundance estimates that are biased low if animals on the line or point are missed. It is useful to distinguish conceptually between two reasons why animals at distance zero might be undetected, using terms coined by Marsh and Sinclair (1989). Animals might be missed because they were simply not available for detection, perhaps because they were submerged or underground (termed ‘availability bias’); or they might be missed because, despite being available for detection, the observers failed to detect them (‘perception bias’). The distinction is not always clear-cut. For example a cetacean 3 m below the surface may be visible from the air, and hence available for detection, depending on water conditions and light. Nevertheless, separating the two processes is often useful.

In Sect. 5.4, we described mark-recapture distance sampling (MRDS) approaches, for which additional data are collected to avoid having to assume that  $g(0) = 1$ . Usually, this is achieved by having two observers (or teams of observers), each of which collects standard distance sampling data. For each animal detected, we also record whether it was detected by the first observer only, the second observer only, or both. This provides the mark-recapture data, which allows estimation without assuming  $g(0) = 1$ .

Mark-recapture methods for estimating abundance are biased low when some animals are not available to be detected in any sample; only the detectable proportion of the population is estimable. Hence, MRDS only addresses perception bias, not availability bias. For example, there is no information about the proportion of desert tortoises (*Gopherus* spp.) that are underground throughout a survey from data recorded above-ground, even if there are multiple observers. While available animals that are missed can be accounted for, those that are not available to either observer cannot from distance sampling data alone. Thus the probability of detection tends to be overestimated, and consequently density estimates are biased low.

If availability is static, *i.e.* if within the time that animals are at risk of being detected, they are either always available or always unavailable, the usual approach is to estimate the proportion of available animals. This is typically done using an independent source of data, as the distance sampling survey data do not contain any information on this proportion. The same approach might be adopted even if availability is strictly not static, but the observation platform is sufficiently fast that the observation process is essentially instantaneous. An obvious example is an aerial cetacean survey, for which the animals are within detection range for only brief periods. The distance sampling density estimate can simply be divided by the estimated proportion of available animals. The variance of this ratio can be estimated using the delta method. Therefore, instead of the usual equations (Eqs. (1.5) and (1.7)), we can formulate the estimator as

$$\hat{D} = \frac{n}{a\hat{P}_{a|v}\hat{P}_v} \quad (11.4)$$

where  $\hat{P}_{a|v}$  is the estimated probability that an object is detected, given that it is both within  $w$  of a line or point and available (estimable by CDS, MCDS or MRDS), and  $\hat{P}_v$  is the estimated probability that it is available for detection (estimated from an independent data source).

Diefenbach et al. (2007) adopted this approach. They estimated availability by independently estimating the proportion of male birds that sing in a time period equal to the length of a point count, and then only singing males were recorded in the point transect survey, thus addressing the problem that silent males might be missed even when at the point.

Thomson et al. (2012) examined availability patterns of two species of turtles, showing that these might be dependent on covariates such as water temperature and depth. This means that a word of caution is required when estimating availability from an independent data source. The data used should be for the time and place that the survey took place, to ensure representativeness. If this is not possible, a model describing the relationship between availability and relevant covariates should be incorporated in the estimation procedure. This requirement applies to any independently-estimated multiplier used to correct a density estimate, for example cue rate in cue counting (Sect. 9.4), and rates of production and decay for indirect surveys (Sect. 9.3).

Barlow et al. (2013) present estimates of availability for acoustic detection of beaked whales, but here an additional component comes into play. Acoustic detectors operate over extended time periods, during which an individual beaked whale is sometimes available and sometimes not. Correction factors are then considerably harder to estimate, because these depend not only on the proportion of time that animals are available for detection, but also on the patterns of availability. This is intuitively obvious. The same proportion of time that animals are available for detection may arise from two very different scenarios. Suppose an animal is available to be detected 20 % of the time. If it is available one in every 5 s, it is essentially always detectable for any platform, but if it is available for a full day every 5 days, it will be missed at least 80 % of the time for most surveys.

Quite often the temporal scale of the availability pattern is such that an animal might be both available and unavailable for detection during the period that it is within detection range. This was the case for the beaked whale example above, and typically for most marine mammals surveyed from ships or boats. For such cases, the concepts of intermittent and static availability described by Laake and Borchers (2004) become important. Whales that do long dives (unavailable), followed by periods of recovery at the surface (available) might be regarded as having static availability. In intermittent availability, an animal is available for detection at a number of instants. Typical examples of intermittent availability are whales which typically do not dive very deep, but frequently surface to breathe, or birds which produce infrequent discrete calls or songs in closed-canopy forests.

Recent work incorporating availability patterns in the estimation process using hidden Markov models shows promising results (Borchers et al. 2013; Langrock et al. 2013), and we anticipate that such approaches will often be adopted in future.

Using a northern minke whale (*Balaenoptera acutorostrata*) data set, Langrock et al. (2013) illustrate how to incorporate the availability process explicitly in the detection process, assuming that the discrete instants when an animal is available to be detected are well-approximated by a Markov-modulated nonhomogeneous Poisson process (MMPP). They provide additional examples that suggest that long-finned pilot whale (*Globicephala melas*) and Blainville's beaked whale (*Mesoplodon densirostris*) availability patterns might also be well-modelled in this way. The approach assumes that animals may be detected at discrete instants of availability; these instants may occur at different rates, depending on the behavioural state of the animal (e.g. diving or resting at or near the surface). The process is nonhomogeneous because probability of detection decreases with distance from the line. The methods of Borchers et al. (2013) are very similar in nature to Langrock et al. (2013), but consider the detection process in continuous time, giving additional examples of an aerial survey of bowhead whales (*Balaena mysticetus*) and a shipboard survey of Cuvier's beaked whales (*Ziphius cavirostris*).

In addition to the usual perpendicular distances, forward distances contain information about availability. This is accommodated using hazard models in which the probability of detection depends on both distances. Whilst the model parameters can be estimated in the absence of forward distances, Borchers et al. (2013) found that results are extremely model-dependent in this case and advise against it. To fit these models, the authors also require independent information on the surfacing pattern, which might be derived from depth tag data, or recorded by following a representative sample of individuals or groups (so-called 'focal follows'). The authors show that upward bias in detection probability, hence negative bias in density/abundance, arises if a simpler homogeneous Poisson process is assumed for the availability pattern when clustering in availability instants is present.

The two approaches discussed above only consider single-observer data. Borchers and Langrock (2015) develop methods for double-observer surveys, incorporating a Markov-modulated Poisson process model for animal availability, together with MRDS methods to allow for perception bias on the line. Their methods are applicable when (a) multiple detections of the same individual are possible and (b) some or all of the availability process parameters are estimated from the line transect survey itself, rather than from independent data. They found that the recording of multiple detections of individuals improves estimator precision substantially when estimating the availability model parameters from survey data. They applied the methods to estimate detection probability from a double-observer survey of North Atlantic minke whales, and found that double-observer data greatly improve estimator precision.

## 11.3 Measurement Error Models

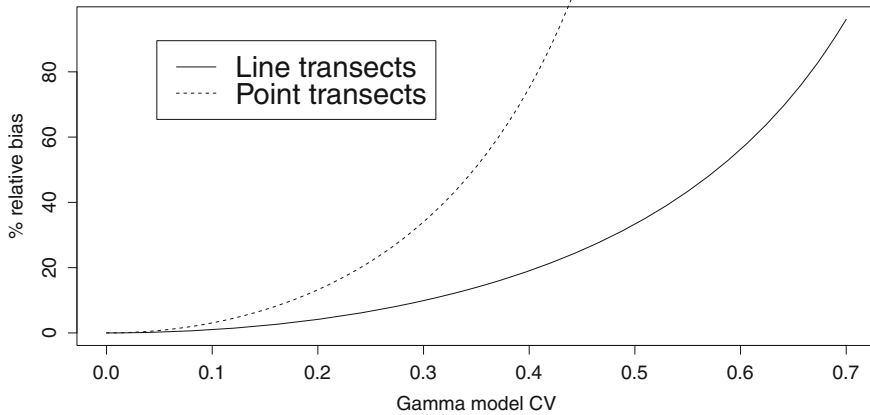
### 11.3.1 *Bias Arising from Measurement Error*

The use of a set of distances to detected animals to model a detection function is a fundamental step in a distance sampling analysis. This leads perhaps to the least surprising of conventional distance sampling assumptions: distances should be measured without error. In the presence of measurement error, bias might arise. The size of the bias is a function of the structure and magnitude of the measurement error.

Assessments of the amount of measurement error in estimating distances during real surveys are scarce. Using an experimental set up with speakers at known distances from a survey point, Alldredge et al. (2007) reported large measurement errors in auditory surveys for birds. However, they did note a 15 % bias reduction in distance estimation following observer training. Strobel and Butler (2014) reported results on an experimental decoy aerial survey where distances were recorded in interval bands. In almost 50 % of trials, the wrong distance band was recorded, a surprisingly high value. This reinforces the advice we share: always train and test observers. This should be done as far as possible under real survey conditions, as the measurement error might be context-dependent. This was found by Williams et al. (2007), who also reported considerable measurement error associated with cetacean boat-based surveys, describing complex non-linear, observer-dependent relationships between estimated distances and true distances.

If distances are systematically overestimated, detection probability is also overestimated, and density underestimated. Correspondingly, if distances are systematically underestimated, detection probability is underestimated, and hence density overestimated. Perhaps more surprisingly, even random errors might lead to bias in the estimated detection probability and hence in density estimates. All else being equal, the effect of measurement error is more pronounced for point transects than for line transects. This can be explained by the geometry of the problem. Considering unbiased random errors in a line transect survey, especially if there is a broad shoulder, random error means that, for a given recorded distance, a similar number of true distances will be above the recorded distance as below. On the other hand, due to the different geometry, for points there are more animals available for detection at greater distances, and so for a given recorded distance, more true distances will exceed the recorded distance than fall below it. This has the effect of increasing the slope of the estimated probability density function of detection distances, overestimating density (see Eq. (1.8)).

In practice, small amounts of measurement error are unlikely to cause problems. Intuitively, the problem depends on the scale of the measurements. While measurement error of the order of metres might be problematic say for a chameleon survey, where animals are detected at most a few metres from the trackline, even errors of the order of tens to low hundreds of metres might be negligible if we are considering large whale surveys, where animals several kilometres from the trackline can be detected.



**Fig. 11.5** Relative percent bias for line and point transect density estimates, as a function of the error model coefficient of variation (CV), assuming an unbiased gamma error model

It is not simple to predict the size of the bias in general. If we consider an unbiased multiplicative gamma model for the errors, *i.e.* if we assume that the estimated distance is  $y = t \times v$ , where  $t$  is the true distance and  $v$  is a gamma-distributed random variable with mean 1 (so that  $\mathbb{E}(y) = t$ ), we can predict the bias as a function of the error coefficient of variation, which provides a good indication of when measurement error matters. This bias is shown for both line and point transects in Fig. 11.5. Additional details are provided in Marques (2007, pp. 57–61).

### 11.3.2 Accounting for Measurement Error

Hiby et al. (1989) probably represents the first attempt to incorporate a measurement error model in the analysis of distance sampling data, adopting a Gaussian error model to analyse cue count data. However, no details of the implementation were provided, and measurement error models were subsequently mostly ignored. Butterworth et al. (1984) and Schweder (1996, 1997) investigated the problem in the context of minke whale surveys in the Antarctic and the North Sea respectively, setting up distance and angle estimation experiments from a ship for which true distance was known. Alpizar-Jara (1997) presented a simulation-based approach to address the problem.

Chen (1998) presented the first explicit and dedicated attempt to model and account for measurement error in distance sampling density estimates, by assuming an additive measurement error model. In such a model, the magnitude of the error is independent of the true distance, which might be adequate say if one is using a GPS to mark the animal location and then the distance is obtained in a GIS. This model was extended by Chen and Cowling (2001) for clustered populations, to incorporate the case of measurement errors in both distances and cluster sizes.

Most often, measurement error increases with true distance, and hence a multiplicative error model seems appropriate. Marques (2004) considered such a model and presented a way to account for measurement error in line transects, by applying a correction factor to the estimation of  $f(0)$  based on an estimated measurement error model.

Section 11.9 of Burnham et al. (2004) provided a first attempt at an integrated approach to the problem of measurement error in distance sampling surveys. Borchers et al. (2010) extended the correction-based approach of Marques (2004) to the case of point transects, verifying that the performance of the correction was poor, especially for points. The reason for this is that the correction factor is based on the impact of the measurement error at distance zero. While that might be considerable, and hence a large correction deemed appropriate, the use of detection functions which are to some extent constrained leads to observed bias that is lower than expected bias; hence the method over-corrects.

The key novelty in Borchers et al. (2010) was the development of an extended likelihood, which includes a model for the measurement error. We define this model as  $f(y|t, \phi)$ , where  $t$  represents the true distance,  $y$  a distance observed with error, and  $\phi$  is a parameter indexing the measurement error model. The usual likelihood, which involves parameters  $\theta$  in the detection function, is extended by integrating out the possible true distance given our measurement error model, allowing estimation of the true detection function based on distances collected with error. Therefore, inferences are still based on the probability density function of observed distances, but we rewrite that function using the error model:

$$f(y|\phi, \theta) = \int_0^{\infty} f(y|t)f(t)dt \quad (11.5)$$

where we note explicitly that the integration is over all possible (true) values of a distance observed with error. In practice we might use a sufficiently large distance instead of  $\infty$ .

Thus in addition to the detection function parameters  $\theta$ , we now have to estimate the nuisance parameter  $\phi$ . Given a sample of  $n_1$  detections with corresponding estimated distances  $y_i$ ,  $i = 1, \dots, n_1$ , this leads to the likelihood

$$\mathcal{L}_y(\theta, \phi) = \prod_{i=1}^{n_1} f(y_i) = \prod_{i=1}^{n_1} \int_0^{\infty} f(y_i|t)f(t)dt . \quad (11.6)$$

The distances measured with error alone do not contain enough information to estimate  $\phi$  and  $\theta$  jointly. Therefore, we need additional information to proceed. If we have separate information about the error model, namely in the form of a set of  $n_2$  pairs of observations for which true and observed distances are available, we can estimate  $\phi$  from the following likelihood:



$$\mathcal{L}_{y,t}(\phi) = \prod_{j=1}^{n_2} f(y_j|t_j) . \quad (11.7)$$

To obtain maximum likelihood estimates of the parameters, we should maximize the product of the two likelihoods (Borchers et al. 2010):

$$\mathcal{L}(\boldsymbol{\theta}, \phi) = \mathcal{L}_y(\boldsymbol{\theta}, \phi) \times \mathcal{L}_{y,t}(\phi) . \quad (11.8)$$

In practice, it is simpler to separate the two components. In other words, we can simply estimate the measurement error model parameters using  $\mathcal{L}_{y,t}(\phi)$  together with a regression model with  $y$  as response and  $t$  as independent variable (possibly with additional covariates). Estimation of the detection function parameters follows, conditional on the estimated measurement error model. Both approaches tend to give similar results.

If distance sampling surveys are formulated using a spatially-explicit capture–recapture framework, we can account for distance measurement error. See Borchers et al. (2015) for details.

### 11.3.3 A Practical Example

As part of an experiment to illustrate acoustic distance sampling methods, an acoustic data set was collected by teams of students during a bioacoustics summer school in Virginia (SEABASS2014). Because true distances and estimated distances were available, we use those data here to illustrate the influence of measurement error in a distance sampling survey with a known population size. The exercise was not set up to evaluate the effects of measurement error, but to illustrate survey methods. Therefore, participants had no special training in estimating distances.

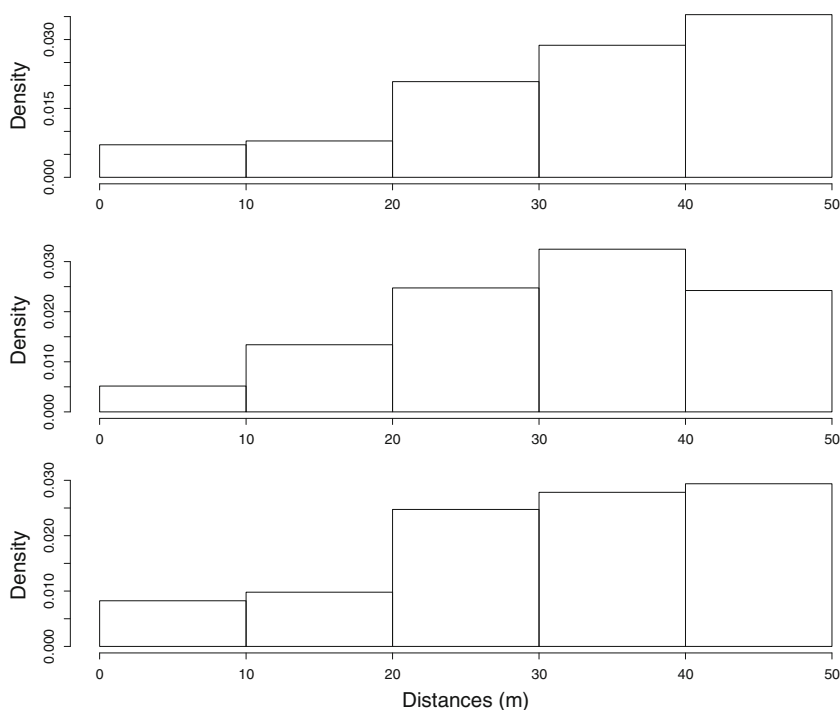
The objective of this artificial survey was to estimate the density of calling animals. Each team comprised four people, one acting as an acoustic detector, two acting as animals (playing a common sound on mobile phones or digital audio players), and one as data recorder.

The survey comprised eight point counts, with a team assigned to each point. For each point, 30 distances were generated from a triangular distribution, up to 50 m from the point. Given that a circle of radius 50 m has area 0.785 ha, this corresponds to a density of 76.4 animals  $\text{ha}^{-1}$ . (Note that the variance across points is unrealistic, being lower than would occur in practice, as the number of available animals from each point was the same.) The ‘animals’ placed themselves at predetermined distances from the point and played the sound. Whether the observer heard the sound or not was recorded by the data recorder. The observer also called out an estimated distance to the sound, hence providing for each detected true distance a distance with measurement error. Note that the observer was aware of the simulation procedure, and hence no distances were estimated as being larger than 50 m. Out of the 240 available animals, 194 were detected.

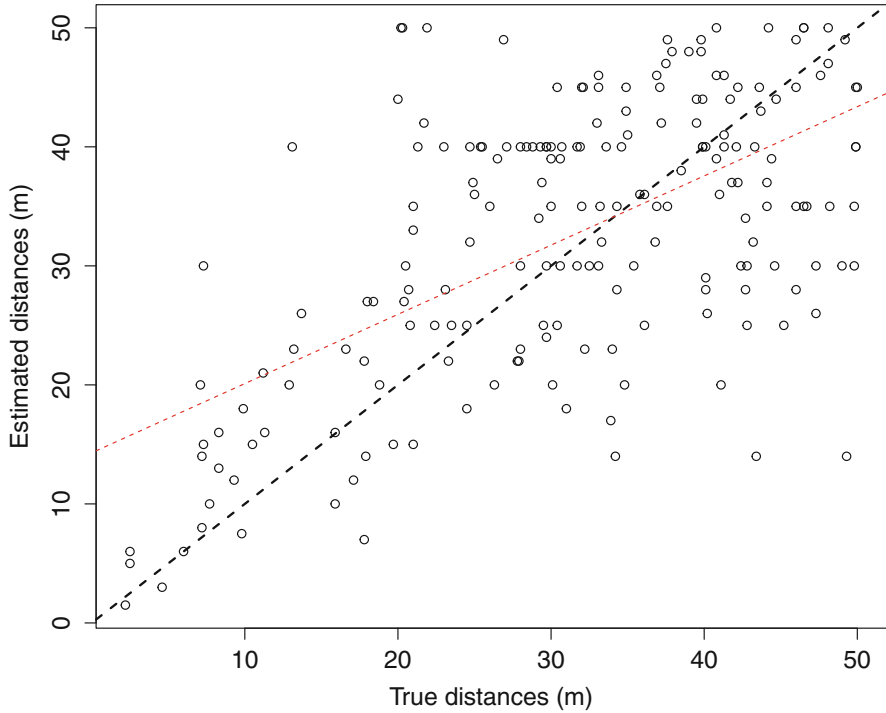
The true and estimated distances are shown in Figs. 11.6 and 11.7. We see a tendency to overestimate small distances and underestimate large distances, with larger variance for larger distances (as would be expected from a multiplicative error model).

We imported the observed and true distances into software `Distance`. We analysed separately the true distances and the estimated distances. We considered a small subset of candidate models for the detection function, namely a uniform and a half-normal, both with cosine adjustments, and a hazard-rate with simple polynomial adjustments.

Using the estimated distances to model the detection function, the model with the lowest AIC was a uniform with no adjustments, implying no decay in detectability with distance. This was inconsistent with the data, as only 81 % of the sounds had been detected. Not surprisingly, this led to underestimation of the true density, with an estimate of 61.3 animals  $\text{ha}^{-1}$  (95 % CI 52.5, 71.7). By contrast, when the true distances were used, the selected model was a half-normal with a cosine adjustment, providing evidence that some of the larger distances had been missed. This led to a point estimate of 82.9 animals  $\text{ha}^{-1}$  (95 % CI 64.4, 106.9).



**Fig. 11.6** Histograms of true distances to all animals (*top panel*), together with estimated (*middle panel*) and true (*bottom panel*) distances to detected animals



**Fig. 11.7** Estimated distances as a function of true distances for the acoustic survey trial. The *dashed black line* represents the 1:1 line (no bias) and the *dashed red line* represents a linear fit to the data

While in a real life survey, there would be no indication in the data that problems had occurred, we note that the confidence interval based on the estimated distances did not contain the true value. On the other hand, the confidence interval using the true distances did.

Because observers knew that no distance exceeded 50m, there was some suppression in variability of estimates corresponding to larger distances, and this will have affected our results when using estimated distances. However, this study does illustrate that, even in an otherwise perfectly-planned survey with a good design, measurement error can generate substantial bias in abundance or density estimates.

### 11.3.4 Take-Home Messages

Humans are often over-confident in their ability to estimate distances, and when tested against known truth, results can be depressing, with large measurement

errors. Therefore, while prevalent in the literature, ‘guestimated’ distances should be avoided wherever possible, and methods adopted to give reliable distance measurements (Sect. 4.1).

The importance of training and testing of observers should not be underestimated, and we strongly recommend that any survey should consider such procedures as a fundamental part of the survey, allocating dedicated time and funds to investigate the extent to which the assumption of no measurement errors might be violated. If measurement error is appreciable, then accounting for it using the above methods should be considered, or field methods should be modified to mitigate the problem. While we can attempt to deal with measurement error at the analysis stage, it is preferable by far to avoid the problem at the data collection stage.

## 11.4 Addressing Bias from Animal Movement

Conceptually, distance sampling is a snapshot method, in which the animals on a survey plot (strip for line transect sampling or circle for point transect sampling) are frozen in position while the observer surveys the plot. Animal movement, whether in response to the observer or independent of the observer, generates bias in distance sampling estimates of abundance.

Advances in tagging technology make it easier to obtain detailed movement data for individuals of many populations. We anticipate that movement models fitted to such datasets will be integrated with distance sampling models before long, allowing abundance estimation free of movement bias. Meantime however, we should seek to use both field methods and analysis methods that reduce bias arising from animal movement.

### 11.4.1 *Movement Independent of the Observer*

Movement independent of the observer is addressed for point transect sampling in Sect. 10.1.2.1. For line transect sampling, Glennie et al. (2015) explored the level of bias that occurs due to non-responsive movement. We summarize their findings here.

We first consider the simpler case of strip transect sampling, and assume that each animal travels in a straight line with the same speed  $u$ , while the observer travels at speed  $v$ . Glennie et al. (2015) show that the percentage bias in the resulting abundance estimate is:

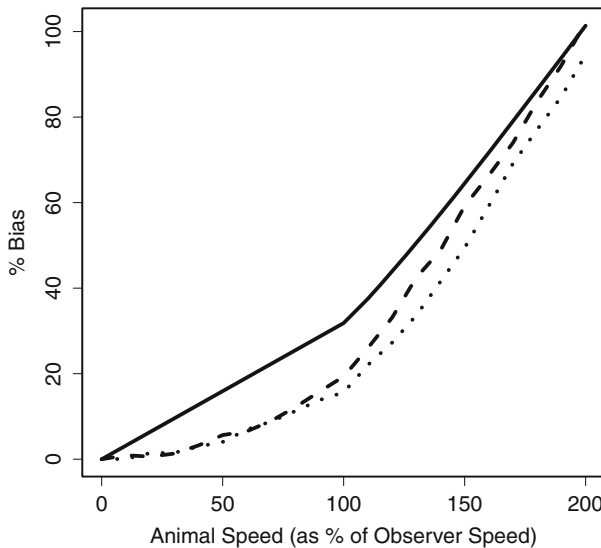
$$100 \times \begin{cases} \frac{ul}{\pi wv} & u \leq v \\ \frac{ul}{\pi wv} + \frac{\sqrt{u^2 - v^2}}{\pi v} - \frac{1}{\pi} \cos^{-1} \left( \frac{v}{u} \right) + \frac{l(u - v)}{2w\pi v} & u > v \end{cases} \quad (11.9)$$

where  $w$  is the strip half-width, and  $l$  is the distance ahead that the observer searches. It is immediately apparent that bias is zero when animal speed  $u = 0$ . Bias is also zero when  $l = 0$ . That is, if the observer searches abeam only, there is no bias (Sect. 10.2.1). If animals show responsive movement however, waiting until animals pass abeam of the observer may not be an option. Instead of using the abeam line, observers can record when detected animals intersect a line perpendicular to the transect but ahead of the observer, at a distance selected so that little responsive movement occurs before animals are recorded.

When animals move faster than the observer, the expression for bias becomes more complex because animals can overtake the observer. The bias is also substantial, but it may be reduced by not recording those animals that overtake the observer, or those animals that enter the 'look ahead' strip of half-width  $w$  and length  $l$  from the side.

Equation (11.9) gives a worst-case scenario. For line transect sampling, movement generates bias both in the count of animals and in the estimate of the detection function, and these two sources of bias interact in a way that reduces overall bias. However, it is not now possible to derive a closed analytic expression for this bias. Thus Glennie et al. (2015) estimated bias in line transect sampling by simulation.

Bias is reduced further if the movement of animals is constrained to lie within a small home range or territory. In Fig. 11.8, we show how bias increases with animal speed for strip and line transect sampling, and for two movement models in the case



**Fig. 11.8** Percentage bias in the abundance estimate for strip transects when  $l = w$  (solid line) and for line transects. For line transect sampling, a hazard-rate detection function was assumed, and bias was estimated by simulation for two movement models: linear movement (dashed line) and movement constrained within a home range (dotted line)

of line transect sampling. We see that bias reaches 10 % for strip transect sampling when animal speed is around one third of observer speed, while for line transect sampling, bias reaches 10 % when animal speed is around 80 % of observer speed.

### ***11.4.2 Responsive Animal Movement***

It is difficult to correct for bias arising from animal movement in response to the observer. A priority should therefore be to adopt field methods so that there is little responsive movement before animals are detected and recorded.

If animals flush in response to the observer, the distance to the point from which they flush should be recorded or estimated. The responsive movement does not then cause bias. More problematic is when responsive movement is not obvious to the observer. It may be necessary to conduct some trials to establish at what distance from the observer responsive movement occurs. Field methods should then be adopted to ensure that animals are detected beyond this response distance. In shipboard surveys, this may require at least some observers searching with binoculars rather than naked eye, or large, tripod-mounted binoculars rather than hand-held binoculars. For aerial surveys, it may mean searching a little ahead of the aircraft rather than abeam, so that animals are detected before they move a significant distance in response to the approaching aircraft.

In some surveys, it is possible to reduce the typical response distance, for example by flying at higher altitude, or by moving more quietly for surveys by foot. In the latter case, this might be achieved by cutting vegetation along the route of transects ahead of the survey. In such cases, the cutting should be minimal, to avoid creating paths for animals or hunters, or creating a ‘tunnel’ which might allow the observer to see animals crossing the transect some distance ahead, which can lead to recording too many animals on the line, and hence to overestimating abundance.

Fewster et al. (2008) used double-observer data to test for the presence of animal movement. They used tests of uniformity, as responsive movement leads to violation of the assumption that animals are uniformly distributed with distance from the line. This is useful for assessing whether responsive movement occurs. If it does, there is still the difficulty of how to remove resulting bias from estimates. Buckland and Turnock (1992) used double-platform data, in which a helicopter set up trials by detecting porpoise ahead of a ship, and recorded whether these animals were subsequently detected by the shipboard observers. This trial mode allows abundance estimates to be obtained that are robust to responsive animal movement.

Sometimes, the effect of responsive movement is obvious from the data. An example of this is the winter wren from the Montrave case study (Sect. 10.1.2.2). In this case, by fitting a model for the detection function that is non-increasing, the lack of detections close to the point ‘averages out’ with the excess of detections at 40–60 m, so that bias from the responsive movement is low. If the movement takes animals further from the line or point, so that no excess of detections is seen at larger distances, abundance is underestimated, and there is no evidence from the data of

the movement. Conversely, if animals move towards the observer, abundance will be overestimated. In this case, the data might show too many detections close to zero distance, and a fitted detection function that falls implausibly fast with increasing distance. If the attraction occurs only for animals already close to the line or point, the distances can be grouped before analysis, with the width of the first group chosen so that animals from the second group are unlikely to respond to the observer. If the responsive movement occurs over larger distances, reliable analysis of the data is unlikely to be possible. In this case, consideration should be given to revising field methods with a view to reducing the problem.

## Chapter 12

### Summary

Distance sampling methods have been successfully applied to estimate density or abundance of populations of a large variety of taxa ranging across many habitats. These are typically very cost-effective methods, especially for species which occur at low densities over large areas. Methods are nonetheless not a panacea, and in many circumstances, assumption violation, and hence bias, might be severe.

Assumption violations are often ignored by practitioners, who might simply report density estimates without considering possible failures of assumptions. We consider that a good distance sampling paper should include in the discussion an assessment of which assumptions hold to a good approximation, and which do not. For those that do not, an assessment of the likely magnitude and direction of the bias should be made.

When planning a survey, the key assumptions should be considered. Are they likely to hold, given what you know about your population and study area? If not, how might you modify your design, or change your field methods, to reduce assumption violations? Do you need to test your methods out in a pilot survey? If at least one assumption remains a serious concern, do you need to resort to more sophisticated methods, as discussed in Chap. 11?

It is far better to address problems when planning the survey than to attempt to rectify problems at the analysis stage. Careful planning can anticipate and negate most problems likely to be encountered, although some populations are inherently difficult to survey in a way that yields reliable density estimates.

In Sect. 12.1, we provide advice on points to consider when planning a distance sampling survey. Although we are not keen on a cookbook approach, we believe that a check-list can be useful. We conclude this book with Sect. 12.2, which deals with technological advances. Although this section is likely to become out-of-date quickly, we feel that it is important to illustrate how technology is not only changing our everyday lives, but also having a tremendous impact on the way we monitor animal populations. We therefore explore how animal survey methods might change in the near future.



## 12.1 Check-List for a Good Survey

We have provided advice to many people engaged in distance sampling studies. Each study has its own nuances, but there are some challenges that occur frequently. We provide here a non-exhaustive list of questions an investigator should consider **prior** to conducting a distance sampling study. This list might be compared with a list of equipment checks by a pilot prior to beginning a flight. Some of these items relate to design of the study, others to conduct of the study, and some to analysis of a completed study.

- Is distance sampling appropriate for your study? If so, which distance sampling method?
  - Do your study animals occur at high density?
    - If this is the case, you might be able to choose a small truncation distance, and meet to a good approximation the assumptions of plot sampling (Sect. 1.4), in which all animals out to the edge of your strip, or to the circumference of your circular plot, are assumed to be detected and counted with certainty. Strip transect or circular plot sampling might then be cost-effective.
  - Is the terrain in your study area difficult to traverse, or is it difficult to estimate distances, perhaps because you are detecting most animals by call or song?
    - Line transects should generally be preferred to point transects (Sect. 1.6). For point transect sampling, relatively few animals are available to be detected at small distances from the point. Consequently, reliably fitting detection functions is more problematic for point transect data. Only consider point transects when it is problematic to traverse random line transects, or when the ability of observers to record high-quality data while moving along a line might be compromised, for example in multi-species songbird surveys, when several individuals of different species might be heard simultaneously.
  - Animals may move more rapidly than observers, or animals may show strong responsive movement before detection (Sect. 11.4), or animals occur at such low densities that it is impractical to detect sufficient numbers.
    - Consider indirect distance sampling (Sect. 9.3), in which density of sign produced by the animals (usually dung or nests) is estimated. The disadvantage of studies based on indirect methods is that they require either two additional components (to estimate sign production rate and sign disappearance rate), or one extra study to estimate sign production rate together with clearance of plots of old sign at the outset of the study. Beyond the additional fieldwork required, the production and disappearance rates are estimates, and the uncertainty in these estimates combines with the uncertainty associated with density of signs, to reduce precision of the resulting estimate of animal density.

- How do animals distribute themselves through the study area at both small and large scales?
  - Is there a large-scale gradient in animal density across the study area?
    - Remember that uncertainty in the estimate of animal density arises as a result of both uncertainty in the estimate of the detection function and variability in encounter rate between point or line transects. The researcher has three choices for reducing the variability in encounter rate. The simplest method is to place transects along the density gradient (*i.e.* perpendicular to animal density contours). This tends to homogenize encounter rates across transects, thus reducing encounter rate variance. The next two bullets deal with other ways of coping with encounter rate variability.
  - Do animals have habitat preferences, with habitats occurring in distinct patches?
    - Statisticians use stratification as a means of variance reduction. Grouping similar things together is a well-used variance reduction technique. Recognising habitat affinities of your study animal and incorporating that information into your survey design, by separately allocating survey effort to distinct habitats, can produce increased efficiency in estimating density (Sect. 2.3). As an added benefit, habitat-specific densities can also be estimated which may be ecologically enlightening.
  - Do animals have habitat preferences, with habitat changing gradually?
    - The third way of coping with variability in encounter rates is to model the changes in encounter rate as a function of covariates. This is the subject of Chaps. 7 and 8. The use of model-based inference methods to extrapolate from the sampled region to the entire study region is an elegant step that may provide biological insight about how the study animal responds to its environment. It also produces estimates that are inherently spatial in nature. The reliability of the estimates depends on the adequacy of the selected model to represent reality.
  - Is there a small-scale gradient in animal density, in the sense that distances to available animals are not distributed according to a uniform (line transect) or triangular (point transect) distribution?
    - The distribution of sightings as a function of distance from the sampler (line or point) is a function of two processes: the distribution of distances of animals from the sampler; and the detectability of animals as a function of distance from the sampler. By using random placement of samplers and a sufficient number of samplers, the first of these two processes is assumed to be correct. This is the so-called design assumption described in Sect. 1.7. Another way to state the assumption is that the study animals are oblivious to the location of the samplers; animals take no notice of sampler location. When this design assumption is not met, because samplers are

placed at locations of significance to the study animals, *e.g.* along fence rows, on roads or tracks, upon headlands looking out to sea, the distribution of distances of animals with respect to the samplers can no longer be assumed and must be estimated. Double-observer methods (Sect. 5.4) or methods that jointly estimate animal density gradient simultaneously with the detection function (Sects. 11.1.2 and 11.1.3) must be used. In either case, additional data need to be collected to compensate for the failure of an assumption. See Chap. 11 for more on the subject of exchanging assumptions for data.

- Do animals occur in clusters?
  - Although there are several ways to incorporate cluster size in analysis, the way we recommend is to use cluster size as a covariate in the detection function (Sect. 5.3).
- Other assumption failures of conventional distance sampling.
  - Detection at distance 0 from the sampler is not perfect ( $g(0) < 1$ ).
    - The exchange of assumptions for data in this situation is direct (Sect. 11.2). Double-observer methods can be used to estimate  $g(0)$ . However, these field methods require: a) an additional observer, hence labour costs increase; b) more elaborate data gathering (to allow identification of duplicate detections between observers) and data collection protocols (with options for observer configuration such as independent observers or tracker and primary observer); and c) more complex data analysis methods. If field methods can be modified to ensure  $g(0) = 1$ , at least to a good approximation, then this is generally a better option than resorting to double-observer methods.
  - Distances of detected animals from the line or point are measured with error.
    - The ideal solution is to improve the field methods. This can be through either high-quality observer training (Sect. 4.2.3) or better technology (*e.g.* rangefinder, clinometer, reticle binoculars; see Chap. 10). If solutions cannot be found for collecting accurate distances, then measurement error models (Sect. 11.3) can be used. The measurement error models require the collection of additional data in the form of field trials in which the true distance is known.
- Final points to consider before engaging in a distance sampling survey.
  - Are you intending to use roads or tracks as transects?
    - The burden of proof is upon the researcher to demonstrate that animal density and use of roadside habitats is representative of the study area in general. This will require additional data, for example from telemetry. A fundamental basis of design-based inference is that areas that receive sampling effort are representative of areas that do not receive sampling

effort. If this assumption is false, then results dependent upon that assumption will be biased (Sect. 4.1). The assumption of representativeness needs to be tested by the investigator as demonstrated by Erxleben et al. (2011); Hilário et al. (2012); McShea et al. (2011).

- Lines or points were not positioned across the study area according to a randomized design (*e.g.* platforms of opportunity were used).
  - If this is the case, then design-based methods cannot be used to draw inference on a wider study region from data on the covered region alone. An alternative is to use model-based inference (Chaps. 7 and 8). However, the model-based approaches require adequate coverage of the range of covariates included in the model, as well as a model that accurately reflects the relationships between the animal and its environment.
- Before conducting your field study, have you conducted a pilot study?
  - There can be no substitute for experience in the field of navigating along lines, detecting animals, recording data and examining patterns in data collected in real time. The virtues of conducting a pilot study are discussed in Sect. 4.2.3.

## 12.2 New Technology

Technological advances have impacts on all aspects of our life, and hence inevitably also lead to changes in the way that surveys are conducted and analysed. Papers with titles like “Robots in Ecology: Welcome to the machine” (Grémillet et al. 2012) suggest that the world of wildlife surveys in the twenty-first century might look very different from what we were used to just a few years ago. We summarize here recent changes, and anticipate some of the future changes.

### 12.2.1 *Using Technology to Detect Animals*

Distance sampling surveys generally rely on human observers to detect animals visually or aurally. This is likely to change in years to come, as cameras and acoustic recorders become more widely used. These will include infrared and high resolution images, as well as both hydrophones and microphones. This will be a game changer for two reasons. First, while these methods have the potential to extend the temporal and spatial scale over which surveys take place, automating the processing of the resulting data becomes fundamental. One can leave a set of cameras or microphones recording for a year or more, but it is not feasible to have operators processing the resulting data manually. Therefore, automated detection and classification of animals from images (*e.g.* Zitterbart et al. 2013; Yang et al. 2014) and acoustic data

(*e.g.* Gillespie et al. 2013) is becoming an important research area. Second, when recording from fixed locations, typically these recordings will occur over long time periods. If animal-based methods are used, accounting for animal movement will therefore be essential. Random animal movement does not generate bias for cue-based methods, and so we can expect to see greater interest in developing and using such methods.

It can be difficult in aerial surveys for the observers to record data that meet the key assumptions of line transect sampling. This might be resolved by replacing observers by high-resolution cameras, fixed to the underside of the aircraft, as described by Buckland et al. (2012). Aircraft can then fly at higher altitude, allowing a wider strip to be surveyed and eliminating disturbance, and detection of animals on the line is not hampered. Detectability typically does not drop across the strip, so that strip transect rather than line transect methods can be used, simplifying analysis. Availability remains an issue. For example, in marine surveys, a proportion of animals may be underwater and undetectable, or in terrestrial surveys, some animals may be hidden by vegetation. Thus methods may be needed to estimate the proportion of the population that is unavailable for detection, to allow estimation of absolute abundance (Sect. 11.2).

The safety problem of aerial or marine surveys may be addressed by using ‘unmanned vehicles’: drones or ‘unmanned aerial vehicles’ (UAVs) for aerial surveys, and gliders or ‘autonomous underwater vehicles’ (AUVs) for underwater surveys. As the costs of operating these fall, and their capabilities improve, we can anticipate their wide use for wildlife surveys. General references to the use of UAVs for wildlife surveys include Watts et al. (2010), Martin et al. (2012) and Lisein et al. (2013). Such methods have been used for surveying geese (Chabot and Bird 2012), elephants (Vermeulen et al. 2013) and dugongs (Hodgson et al. 2013). Current limitations of these devices include low autonomy due to short battery life and cost (*e.g.* Vermeulen et al. 2013), but both of these are likely to become less of a hurdle as technology develops further and becomes more widely used, and hence cheaper and more reliable.

Similarly autonomous underwater vehicles such as gliders are starting to be used for passive acoustic surveys of both cetaceans (*e.g.* Baumgartner and Fratantoni 2008; Klinck et al. 2012; Baumgartner et al. 2013) and fish (Wall et al. 2012).

The kind of images used can improve the detection rate; for example, thermal imagers can increase the detection rate of animals from the air when there is partial canopy cover. Thermal imagers can also be useful for observers on foot or in a vehicle, operating in closed habitats (Gill et al. 1997). Infrared cameras have been used in combination with aerial surveys of deer (Storm et al. 2011). Infrared video cameras can also be used to complement visual sightings for boat surveys (Zitterbart et al. 2013) and will probably see more use in coming years as the cost of such technology drops.

Systems to detect and localize animals automatically from passive acoustic data are being developed for both marine (Hayes et al. 2000) and terrestrial (Hobson et al. 2002; Mennill et al. 2006) environments, which will allow the implementation of distance sampling.

Camera traps are being routinely used for capture–recapture studies (Rowcliffe et al. 2008; Royle and Gardner 2011). It is possible to use these in a similar way to implement distance sampling surveys. Rowcliffe et al. (2011) proposed a way to estimate the area effectively monitored by cameras, using distance sampling theory. To convert such estimates to estimated density or abundance, animal movement rates must be estimated or modelled. However, unlike for capture–recapture, it is not necessary to identify individual animals. We anticipate that methods will soon be available that do not need the ability either to identify individuals or to quantify movement rates.

The widespread use of camera traps is prompting methodological advances. For example Borchers et al. (2015) develop a class of models that synthesizes capture–recapture and distance sampling methods. It accommodates a spectrum of models ranging from non-spatial capture–recapture models (with no information on animal locations) through to distance sampling and mark-recapture distance sampling models, in which animal locations are observed without error. Between these lie spatially-explicit capture–recapture models that include only capture locations, and a variety of models with fewer location data than are typical of distance sampling surveys, but more than are normally used on spatially-explicit capture–recapture surveys. The class also allows measurement error in distance sampling to be accommodated (Sect. 11.3.2).

### ***12.2.2 Data Recording***

Systems for streamlining and automating data recording will become more common, reducing errors (*e.g.* Southwell et al. 2002). A move away from using observers and towards high-resolution images and acoustic surveys will speed up this process.

### ***12.2.3 Distance Estimation***

Accurate estimation of distance is a problem for many surveys. Laser rangefinders have significantly reduced or even eliminated the problem for many terrestrial surveys. However, low-tech methods are typically used to estimate observer-to-animal distances in shipboard surveys (*e.g.* reticles, Sect. 4.1.2.3), or animal-to-line distances in aerial surveys (*e.g.* markers on wing struts, Sect. 4.1.2.2). Similarly, angles from the line in shipboard surveys are typically measured using angle boards or angle rings on the tripod if tripod-mounted binoculars are used (Sect. 4.1.2.3).

Often, no aids at all are used for measuring distances and angles, despite evidence from many studies that resulting bias in measurements, and hence in density estimates, can be large (*e.g.* Alldredge et al. 2007). Hopefully, technology will soon make accurate estimation in marine and aerial surveys as easy as it is for terrestrial surveys in open habitat.

### ***12.2.4 Aerial Surveys***

We expect to see much wider use of aerial surveys in future, over both land and water. Unlike observers on foot or in a surface vehicle, aircraft can readily follow random tracklines. They can also cover far more ground. For marine surveys, they are substantially less expensive than shipboard surveys, and can again cover a larger area in a given time. There are two key problems with aerial surveys. The first is that observers find it difficult to record line transect data that satisfy the assumptions of the method, in part because of the speed of the platform, combined with the need to be at low altitude to allow detection and identification of animals, and in part due to obstructed visibility, especially below the aircraft. The second is that safety concerns restrict where aircraft can operate. Technology can now address both problems, by using drones together with high resolution imagery.

### ***12.2.5 Satellites and GPS***

The idea that satellites could be used to count animals has been around for a few decades (Schwaller et al. 1984). In recent years, several examples have come to light showcasing that ability. Fretwell et al. (2012) demonstrate how technology is affecting the way we conduct surveys, reporting for the emperor penguin the first global census of a widely distributed species using high-definition satellite imagery. Lynch and LaRue (2014) present a similar approach, this time for the Adélie penguin, while Fretwell et al. (2014) report on a survey of southern right whales using the same technology.

Satellite images are unlikely to be used in a distance sampling context, as the notion of the probability of detection being a function of distance is absent in a digital image. Nonetheless, satellite and GPS data can provide auxiliary information useful for distance sampling surveys. This is particularly so when extending conventional methods, and examples include proportion of time animals are available for detection (Langrock et al. 2013), and the distribution of animals with respect to non-randomly allocated transects (Marques et al. 2013). Addressing animal movement in distance sampling surveys is work in progress (Sect. 11.4), but obtaining independent information about animal movement seems a necessity, if we are to account for it adequately. Using satellite/GPS data to quantify animal speed and movement seems a useful strategy (*e.g.* Dujon et al. 2014).

GPS has been used as a way to reduce distance measurement error in helicopter surveys (Marques et al. [2006](#)).

Satellites potentially can be used to survey animal populations, but for these to become widely used for animal surveys, further methodological developments are required, to allow reliable detection and identification of animals.



## Erratum

### Distance Sampling: Methods and Applications

S.T. Buckland, E.A. Rexstad, T.A. Marques, and C.S. Oedekoven

S.T. Buckland et al., *Distance Sampling: Methods and Applications*,  
Methods in Statistical Ecology, DOI 10.1007/978-3-319-19219-2,  
© Springer International Publishing Switzerland 2015

---

#### DOI 10.1007/978-3-319-19219-2\_13

Some original tables were inadvertently published with colons instead of decimal points for values. The corrected tables are listed below:

Table 5.7, 5.9, 5.10, 5.11, 7.2, 8.1, 8.2, and 11.1

---

The online version of the original book can be found at  
<http://dx.doi.org/10.1007/978-3-319-19219-2>

© Springer International Publishing Switzerland 2015  
S.T. Buckland et al., *Distance Sampling: Methods and Applications*,  
Methods in Statistical Ecology, DOI 10.1007/978-3-319-19219-2\_13

# References

- Able, K. P. (1977). The flight behaviour of individual passerine nocturnal migrants: A tracking radar study. *Animal Behaviour* 25, 924–935.
- Acevedo, P., J. Ferreres, R. Jaroso, M. Durán, M. Escudero, J. Marco, and C. Gortázar (2010). Estimating roe deer abundance from pellet group counts in Spain: An assessment of methods suitable for Mediterranean woodlands. *Ecological Indicators* 10, 1226–1230.
- Acosta, C. A. and S. A. Perry (2000). Effective sampling area: a quantitative method for sampling crayfish populations in freshwater marshes. *Crustaceana* 73, 425–431.
- Allredge, M. W., K. Pacifici, T. R. Simons, and K. H. Pollock (2008). A novel field evaluation of the effectiveness of distance and independent observer sampling to estimate aural avian detection probabilities. *Journal of Applied Ecology* 45, 1349–1356.
- Allredge, M. W., K. H. Pollock, T. Simons, and S. A. Shriner (2007). Multiple-species analysis of point count data: a more parsimonious modeling framework. *Journal of Applied Ecology* 44, 281–290.
- Allredge, M. W., T. R. Simons, and K. H. Pollock (2007). A field evaluation of distance measurement error in auditory avian point count surveys. *Journal of Wildlife Management* 71, 2759–2766.
- Alpizar-Jara, R. (1997). *Assessing Assumption Violation in Line Transect Sampling*. Ph. D. thesis, North Carolina State University, Raleigh.
- Alves, J., A. A. da Silva, A. M. Soares, and C. Fonseca (2013). Pellet group count methods to estimate red deer densities: Precision, potential accuracy and efficiency. *Mammalian Biology — Zeitschrift für Säugetierkunde* 78, 134–141.
- Anderson, D. R., K. P. Burnham, G. C. White, and D. L. Otis (1983). Density estimation of small-mammal populations using a trapping web and distance sampling methods. *Ecology* 64, 674–680.
- Arranz, P., D. L. Borchers, N. A. de Soto, M. P. Johnson, and M. J. Cox (2014). A new method to study inshore whale cue distribution from land-based observations. *Marine Mammal Science* 30, 810–818.
- Baccaro, F. B. and G. Ferraz (2012). Estimating density of ant nests using distance sampling. *Insectes Sociaux* 60, 103–110.
- Barlow, J. and B. Taylor (2005). Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. *Marine Mammal Science* 21, 429–445.
- Barlow, J., P. L. Tyack, M. P. Johnson, R. W. Baird, G. S. Schorr, R. D. Andrews, and N. A. de Soto (2013). Trackline and point detection probabilities for acoustic surveys of Cuvier's and Blainville's beaked whales. *The Journal of the Acoustical Society of America* 134, 2486–2496.

- Barnes, R. F. W. (2001). How reliable are dung counts for estimating elephant numbers? *African Journal of Ecology* 39, 1–9.
- Barnes, R. F. W. and K. L. Barnes (1992). Estimating decay rates of elephant dung-piles in forest. *African Journal of Ecology* 30, 316–321.
- Barnes, R. F. W., A. Blom, M. P. T. Alers, and K. L. Barnes (1995). An estimate of the numbers of forest elephants in Gabon. *Journal of Tropical Ecology* 11, 27–37.
- Baumgartner, M. F. and D. M. Fratantoni (2008). Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. *Limnology and Oceanography* 53, 2197–2209.
- Baumgartner, M. F., D. M. Fratantoni, T. P. Hurst, M. W. Brown, T. V. N. Cole, S. M. Van Parijs, and M. Johnson (2013). Real-time reporting of baleen whale passive acoustic detections from ocean gliders. *The Journal of the Acoustical Society of America* 134, 1814–1823.
- Becker, E. F. and P. X. Quang (2009). A gamma-shaped detection function for line-transect surveys with mark-recapture and covariate data. *Journal of Agricultural, Biological, and Environmental Statistics* 14, 207–223.
- Blumstein, D. T., D. J. Mennill, P. Clemins, L. Girod, K. Yao, G. Patricelli, J. L. Deppe, A. H. Krakauer, C. Clark, K. A. Cortopassi, S. F. Hanser, B. McCowan, A. M. Ali, and A. N. G. Kirschel (2011). Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology* 48, 758–767.
- Borchers, D. L. (1996). *Line Transect Estimation with Uncertain Detection on the Trackline*. Ph. D. thesis, University of Cape Town, Cape Town.
- Borchers, D. L. (2012). A non-technical overview of spatially explicit capture-recapture models. *Journal of Ornithology* 152, 435–444.
- Borchers, D. L., S. T. Buckland, I. G. Priede, and S. Ahmadi (1997). Improving the precision of the daily egg production method using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences* 54, 2727–2742.
- Borchers, D. L., S. T. Buckland, and W. Zucchini (2002). *Estimating Animal Abundance*. London: Springer.
- Borchers, D. L. and K. P. Burnham (2004). General formulation for distance sampling. pp. 6–30 in S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Borchers, D. L. and M. G. Efford (2008). Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics* 64, 377–385.
- Borchers, D. L., J. L. Laake, C. Southwell, and C. G. M. Paxton (2006). Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* 62, 372–378.
- Borchers, D. L. and R. Langrock (2015). Double-observer line transect surveys with Markov-modulated Poisson process models for animal availability. *Biometrics*. doi: 10.1111/biom.12341
- Borchers, D. L., T. A. Marques, T. Gunnlaugsson, and P. E. Jupp (2010). Estimating distance sampling detection functions when distances are measured with errors. *Journal of Agricultural, Biological, and Environmental Statistics* 15, 346–361.
- Borchers, D. L., B. Stevenson, D. Kidney, L. Thomas, and T. Marques (2015). A unifying model for capture-recapture and distance sampling. *Journal of the American Statistical Association* 110, 195–204.
- Borchers, D. L., W. Zucchini, M. P. Heide-Jørgensen, A. Cañadas, and R. Langrock (2013). Using hidden Markov models to deal with availability bias on line transect surveys. *Biometrics* 69, 703–713.
- Bradford, A. L., K. A. Forney, E. M. Oleson, and J. Barlow (2014). Accounting for subgroup structure in line-transect abundance estimates of false killer whales (*Pseudorca crassidens*) in Hawaiian waters. *PLoS ONE* 9, e90464.
- Bravington, M. (1994). *The Effects of Acidification on the Population Dynamics of Brown Trout in Norway*. Ph. D. thesis, Imperial College, London.

- Brown, J. A. and M. S. Boyce (1998). Line transect sampling of Karner blue butterflies (*Lycaeides melissa samuelis*). *Environmental and Ecological Statistics* 5, 81–91.
- Bruderer, B. and P. Steidinger (1972). Methods of quantitative and qualitative analysis of bird migration with a tracking radar. pp. 151–167 in S. Galler, K. Schmidt-Koenig, G. Jacobs, and R. Belleville (Eds.), *Animal Orientation and Navigation*. NASA, Washington, D.C.
- Buckland, S. T. (1984). Monte Carlo confidence intervals. *Biometrics* 40, 811–817.
- Buckland, S. T. (2006). Point transect surveys for songbirds: robust methodologies. *The Auk* 123, 345–357.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake (1993). *Distance Sampling: Estimating Abundance of Biological Populations*. London: Chapman and Hall.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford: Oxford University Press.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (2004). *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Buckland, S. T., D. L. Borchers, A. Johnston, P. A. Henrys, and T. A. Marques (2007). Line transect methods for plant surveys. *Biometrics* 63, 989–998.
- Buckland, S. T., K. P. Burnham, and N. H. Augustin (1997). Model selection: an integral part of inference. *Biometrics* 53, 603–618.
- Buckland, S. T., M. L. Burt, E. A. Rexstad, M. Mellor, A. E. Williams, and R. Woodward (2012). Aerial surveys of seabirds: the advent of digital methods. *Journal of Applied Ecology* 49, 960–967.
- Buckland, S. T., J. L. Laake, and D. L. Borchers (2010). Double-observer line transect methods: levels of independence. *Biometrics* 66, 169–177.
- Buckland, S. T., S. J. Marsden, and R. E. Green (2008). Estimating bird abundance: making methods work. *Bird Conservation International* 18, S91–S108.
- Buckland, S. T., C. S. Oedekoven, and D. L. Borchers (submitted). Model-based distance sampling.
- Buckland, S. T., A. J. Plumptre, L. Thomas, and E. A. Rexstad (2010a). Design and analysis of line transect surveys for primates. *International Journal of Primatology* 31, 833–847.
- Buckland, S. T., A. J. Plumptre, L. Thomas, and E. A. Rexstad (2010b). Line transect sampling of primates: can animal-to-observer distance methods work? *International Journal of Primatology* 31, 485–499.
- Buckland, S. T., R. E. Russell, B. G. Dickson, V. A. Saab, D. N. Gorman, and W. M. Block (2009). Analysing designed experiments in distance sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 14, 432–442.
- Buckland, S. T., R. W. Summers, D. L. Borchers, and L. Thomas (2006). Point transect sampling with traps or lures. *Journal of Applied Ecology* 43, 377–384.
- Buckland, S. T. and B. J. Turnock (1992). A robust line transect method. *Biometrics* 48, 901–909.
- Burnham, K. P. and D. R. Anderson (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. Springer-Verlag.
- Burnham, K. P., D. R. Anderson, and J. L. Laake (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monographs* 72, 1–202.
- Burnham, K. P., S. T. Buckland, J. L. Laake, D. L. Borchers, T. A. Marques, J. R. B. Bishop, and L. Thomas (2004). Further topics in distance sampling. pp. 307–392 in S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Burt, L., D. L. Borchers, K. J. Jenkins, and T. A. Marques (2014). Using mark-recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution* 5, 1180–1191.
- Butterworth, D. S., P. B. Best, and D. Hembree (1984). Analysis of experiments carried out during the 1981/82 IWC/IDCR Antarctic minke whale assessment cruise in Area II. *Report of the International Whaling Commission* 34, 365–392.
- Cañadas, A. and P. Hammond (2006). Model-based abundance estimates for bottlenose dolphins off southern Spain: implications for conservation and management. *Journal of Cetacean Research and Management* 8, 13–27.

- Chabot, D. and D. M. Bird (2012). Evaluation of an off-the-shelf unmanned aircraft system for surveying flocks of geese. *Waterbirds* 35, 170–174.
- Chapman, J. W., D. R. Reynolds, A. D. Smith, J. R. Riley, D. E. Pedgley, and I. P. Woiwod (2002). High-altitude migration of the diamondback moth *Plutella xylostella* to the U.K.: a study using radar, aerial netting, and ground trapping. *Ecological Entomology* 27, 641–650.
- Chen, S. X. (1998). Measurement errors in line transect surveys. *Biometrics* 54, 899–908.
- Chen, S. X. and A. Cowling (2001). Measurement errors in line transect surveys where detectability varies with distance and size. *Biometrics* 57, 732–742.
- Collier, B. A., S. L. Farrell, A. M. Long, A. J. Campomizzi, K. B. Hays, J. L. Laake, M. L. Morrison, and R. N. Wilkins (2013). Modeling spatially explicit densities of endangered avian species in a heterogeneous landscape. *The Auk* 130, 666–676.
- Cox, M. J., D. L. Borchers, D. A. Demer, G. R. Cutter, and A. S. Brierley (2011). Estimating the density of Antarctic krill (*Euphausia superba*) from multi-beam echo-sounder observations using distance sampling methods. *Applied Statistics* 60, 301–316.
- Cox, M. J., D. L. Borchers, and N. Kelly (2013). nupoint: An R package for density estimation from point transects in the presence of nonuniform animal density. *Methods in Ecology and Evolution* 4, 589–594.
- Davison, A. C. (2003). *Statistical Models*. Cambridge: Cambridge University Press.
- Dawson, D. K. and M. G. Efford (2009). Bird population density estimated from acoustic signals. *Journal of Applied Ecology* 46, 1201–1209.
- Dickson, B. G., B. R. Noon, C. H. Flather, S. Jentsch, and W. M. Block (2009). Quantifying the multi-scale response of avifauna to prescribed fire experiments in the southwest United States. *Ecological Applications* 19, 608–621.
- Diefenbach, D. R., M. R. Marshall, J. A. Mattice, and D. W. Brauning (2007). Incorporating availability for detection in estimates of bird abundance. *The Auk* 124, 96–106.
- Dokter, A. M., M. J. Baptist, B. J. Ens, K. L. Krijgsveld, and E. E. van Loon (2013). Bird radar validation in the field by time-referencing line-transect surveys. *PLoS ONE* 8, e74129.
- Dujon, A. M., R. T. Lindstrom, and G. C. Hays (2014). The accuracy of Fastloc-GPS locations and implications for animal tracking. *Methods in Ecology and Evolution* 5, 1162–1169.
- Efford, M. G. (2004). Density estimation in live-trapping studies. *Oikos* 106, 598–610.
- Erxleben, D., M. Butler, W. Ballard, M. Wallace, M. Peterson, N. Silvy, W. Kuvlesky, D. Hewitt, S. DeMaso, J. Hardin, and M. Dominguez-Brazil (2011). Wild turkey (*Meleagris gallopavo*) association to roads: implications for distance sampling. *European Journal of Wildlife Research* 57, 57–65.
- Evans, K. O., L. W. Burger, C. S. Oedekoven, M. D. Smith, S. K. Riffell, J. A. Martin, and S. T. Buckland (2013). Multi-region response to conservation buffers targeted for northern bobwhite. *The Journal of Wildlife Management* 77, 716–725.
- Fancy, S. G., T. J. Snetsinger, and J. D. Jacobi (1997). Translocation of the palila, an endangered Hawaiian honeycreeper. *Pacific Conservation Biology* 3, 39–46.
- Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer (2002). A removal model for estimating detection probabilities from point-count surveys. *The Auk* 119, 414–425.
- Fewster, R. M., S. T. Buckland, K. P. Burnham, D. L., Jupp, P. E., Laake, J. L. and Thomas, L. (2009). Estimating the encounter rate variance in distance sampling. *Biometrics* 65, 225–236.
- Fewster, R. M. and A. R. Pople (2008). A comparison of mark-recapture distance sampling methods applied to aerial surveys of Eastern grey kangaroos. *Wildlife Research* 35, 320–330.
- Fewster, R. M., C. Southwell, D. L. Borchers, S. T. Buckland, and A. R. Pople (2008). The influence of animal mobility on the assumption of uniform distances in aerial line transect surveys. *Wildlife Research* 35, 275–288.
- Fletcher, J. R. and L. R. Hutto (2006). Estimating detection probabilities of river birds using double surveys. *The Auk* 123, 695–707.

- Focardi, S., B. Franzetti, and F. Ronchi (2013). Nocturnal distance sampling of a Mediterranean population of fallow deer is consistent with population projections. *Wildlife Research* 40, 437–446.
- Fretwell, P. T., M. A. LaRue, P. Morin, G. L. Kooyman, B. Wienecke, N. Ratcliffe, A. J. Fox, A. H. Fleming, C. Porter, and P. N. Trathan (2012). An emperor penguin population estimate: the first global, synoptic survey of a species from space. *PLoS ONE* 7, e33751.
- Fretwell, P. T., I. J. Staniland, and J. Forcada (2014). Whales from space: counting southern right whales by satellite. *PLoS ONE* 9, e88655.
- Gale, G. A., P. D. Round, and A. J. Pierce (2009). A field test of distance sampling methods for a tropical forest bird community. *The Auk* 126, 439–448.
- Gates, C. E., W. H. Marshall, and D. P. Olson (1968). Line transect method of estimating grouse population densities. *Biometrics* 24, 135–145.
- Gelman, A., G. O. Roberts, and W. R. Gilks (1996). Efficient Metropolis jumping rules. pp. 599–607 in J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Bayesian Statistics*. Oxford University Press, Oxford.
- George, J., J. Zeh, R. Suydam, and C. Clark (2004). Abundance and population trend (1978–2001) of western Arctic bowhead whales surveyed near Barrow. *Marine Mammal Science* 20, 755–773.
- Gerrodette, T., G. Watters, W. Perryman, and L. T. Ballance (2008). Estimates of 2006 dolphin abundance in the eastern tropical Pacific, with revised estimates from 1986–2003. Technical report, U.S. Department of Commerce, Southwest Fisheries Science Center NOAA Technical Memorandum NMFS NOAA-TM-NMFS-SWFSC-422.
- Gill, R. M. A., M. L. Thomas, and D. Stocker (1997). The use of portable thermal imaging for estimating deer population density in forest habitats. *Journal of Applied Ecology* 34, 1273–1286.
- Gillespie, D., M. Caillat, J. Gordon, and P. White (2013). Automatic detection and classification of odontocete whistles. *The Journal of the Acoustical Society of America* 134, 2427–2437.
- Glennie, R., S. T. Buckland, and L. Thomas (2015). The effect of animal movement on line transect estimates of abundance. *PLoS One* 10, e0121333.
- Green, P. J. (1995). Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82, 711–732.
- Grémillet, D., V. Puech, V. Garçon, T. Boulonier, and Y. Maho (2012). Robots in ecology: Welcome to the machine. *Open Journal of Ecology* 2, 49–57.
- Grundel, R. (2008). A guide to the use of distance sampling to estimate abundance of Karner blue butterflies. Technical report, US Fish and Wildlife Service.
- Guenzel, R. J. (1997). Estimating pronghorn abundance using aerial line transect sampling. Technical report, Wyoming Game and Fish Dept., 5400 Bishop Blvd., Cheyenne, WY.
- Hammond, P. S., P. Berggren, H. Benke, D. L. Borchers, A. Collet, M. P. Heide-Jørgensen, S. Heimlich, A. R. Hiby, M. F. Leopold, and N. Øien (2002). Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters. *The Journal of Applied Ecology* 39, 361–376.
- Hardin, J. and J. Hilbe (2003). *Generalized Estimating Equations*. Boca Raton: Chapman and Hall/CRC.
- Harmata, A. R., K. M. Podrutzny, J. R. Zelenak, and M. L. Morrison (1999). Using marine surveillance radar to study bird movements and impact assessment. *Wildlife Society Bulletin* 27, 44–52.
- Harris, D., L. Thomas, L. Matias, and T. Yack (in prep.). Dealing with depth: adapting distance sampling methods to estimate densities of diving marine mammals using passive acoustic recordings.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 57, 97–109.
- Hayes, R. J. and S. T. Buckland (1983). Radial-distance models for the line-transect method. *Biometrics* 39, 29–42.

- Hayes, S. A., D. K. Mellinger, D. A. Croll, D. P. Costa, and J. F. Borsani (2000). An inexpensive passive acoustic system for recording and localizing wild animal sounds. *The Journal of the Acoustical Society of America* 107, 3552–3555.
- Hedley, S. L. (2000). *Modelling Heterogeneity in Cetacean Surveys*. Ph. D. thesis, University of St Andrews, St Andrews.
- Hedley, S. L. and S. T. Buckland (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 9, 181–199.
- Hedley, S. L., S. T. Buckland, and D. L. Borchers (1999). Spatial modelling from line transect data. *Journal of Cetacean Research and Management* 1, 255–264.
- Hiby, A. R. (1985). An approach to estimating population densities of great whales from sighting surveys. *IMA Journal of Mathematics Applied in Medicine and Biology* 2, 201–220.
- Hiby, L. and P. Lovell (1998). Using aircraft in tandem formation to estimate abundance of harbour porpoise. *Biometrics* 54, 1280–1289.
- Hiby, L., A. Ward, and P. Lovell (1989). Analysis of the North Atlantic sightings survey 1987: Aerial survey results. *Report of the International Whaling Commission* 39, 447–455.
- Hilário, R., F. Rodrigues, A. Chiarello, and Í. Mourthé (2012). Can roads be used as transects for primate population surveys? *Folia Primatologica* 83, 47–55.
- Hobson, K. A., R. S. Rempel, H. Greenwood, B. Turnbull, and S. L. Van Wilgenburg (2002). Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. *Wildlife Bulletin* 30, 709–720.
- Hodgson, A., N. Kelly, and D. Peel (2013). Unmanned aerial vehicles (UAVs) for surveying marine fauna: A dugong case study. *PLoS ONE* 8, e79556.
- Högmander, H. (1991). A random field approach to transect counts of wildlife populations. *Biometrical Journal* 33, 1013–1023.
- Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47, 663–685.
- Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54, 187–211.
- Isaac, N. J. B., K. L. Cruickshanks, A. M. Weddle, J. M. Rowcliffe, T. M. Brereton, R. L. H. Dennis, D. M. Shuker, and C. D. Thomas (2011). Distance sampling and the challenge of monitoring butterfly populations. *Methods in Ecology and Evolution* 2, 585–594.
- Jenkins, R. K. B., L. D. Brady, K. Huston, J. L. D. Kauffmann, J. Rabearivony, G. Raveloson, and J. M. Rowcliffe (1999). The population status of chameleons within Ranomafana National Park, Madagascar, and recommendations for future monitoring. *Oryx* 33, 38–46.
- Johnson, B. K., F. G. Lindzey, and R. J. Guenzel (1991). Use of aerial line transect surveys to estimate pronghorn populations in Wyoming. *Wildlife Society Bulletin* 19, 315–321.
- Johnson, D., J. L. Laake, and J. M. Ver Hoef (2009). Package DSpat. Technical report.
- Johnson, D. S., J. L. Laake, and J. M. Ver Hoef (2010). A model-based approach for making ecological inference from distance sampling data. *Biometrics* 66, 310–318.
- Johnson, M., P. T. Madsen, W. M. X. Zimmer, N. A. de Soto, and P. L. Tyack (2004). Beaked whales echolocate on prey. *Proceedings of the Royal Society B: Biological Sciences* 271, S383–S386.
- Johnson, M. P. and P. L. Tyack (2003). A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering* 28, 3–12.
- King, R., B. J. T. Morgan, O. Gimenez, and S. P. Brooks (2009). *Bayesian Analysis for Population Ecology*. Boca Raton: CRC Press.
- Klinck, H., D. K. Mellinger, K. Klinck, N. M. Bogue, J. C. Luby, W. A. Jump, G. B. Shilling, T. Litchendorf, A. S. Wood, G. S. Schorr, and R. W. Baird (2012). Near-real-time acoustic monitoring of beaked whales and other cetaceans using a seaglider. *PLoS ONE* 7, e36128.
- Kouakou, C. Y., C. Boesch, and H. Kuehl (2009). Estimating chimpanzee population size with nest counts: validating methods in Tai National Park. *American Journal of Primatology* 71, 447–457.
- Kusel, E. T., D. K. Mellinger, L. Thomas, T. A. Marques, D. Moretti, and J. Ward (2011). Cetacean population density estimation from single fixed sensors using passive acoustics. *The Journal of the Acoustical Society of America* 129, 3610–3622.

- Kyhn, L., J. Tougaard, L. Thomas, L. Duve, J. Steinback, M. Amundin, G. Desportes, and J. Teilmann (2012). From echolocation clicks to animal density — acoustic sampling of harbour porpoises with static dataloggers. *The Journal of the Acoustical Society of America* 131, 550–560.
- Laake, J. L. (1999). Distance sampling with independent observers: reducing bias from heterogeneity by weakening the conditional independence assumption. pp. 137–148 in G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson (Eds.), *Marine Mammal Survey and Assessment Methods*. Rotterdam: Balkema.
- Laake, J. L. and D. L. Borchers (2004). Methods for incomplete detection at distance zero. pp. 108–189 in S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Laake, J. L., S. T. Buckland, D. R. Anderson, and K. P. Burnham (1993). *DISTANCE Users' Guide V2.0*. Fort Collins, CO: Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University.
- Laake, J. L., B. A. Collier, M. L. Morrison, and R. N. Wilkins (2011). Point-based mark-recapture distance sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 16, 389–408.
- Laing, S. E., S. T. Buckland, R. W. Burn, D. Lambie, and A. Amphlett (2003). Dung and nest surveys: estimating decay rate. *Journal of Applied Ecology* 40, 1102–1111.
- Langrock, R., D. L. Borchers, and H. J. Skaug (2013). Markov-modulated nonhomogeneous Poisson processes for unbiased estimation of marine mammal abundance. *Journal of the American Statistical Association* 108, 840–851.
- Legare, M. L., W. R. Eddleman, P. A. Buckley, and C. Kelly (1999). The effectiveness of tape playback in estimating black rail density. *The Journal of Wildlife Management* 63, 116–125.
- Lewis, T., D. Gillespie, C. Lacey, J. Matthews, M. Danbolt, R. Leaper, R. McLanaghan, and A. Moscrop (2007). Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003. *Journal of the Marine Biological Association of the United Kingdom* 87, 353–357.
- Link, W. A. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics* 59, 1123–1130.
- Link, W. A. (2004). Individual heterogeneity and identifiability in capture-recapture models. *Animal Biodiversity and Conservation* 27, 87–91.
- Lisein, J., J. Linchant, P. Lejeune, P. Bouché, and C. Vermeulen (2013). Aerial surveys using an unmanned aerial system (UAS): comparison of different methods for estimating the surface area of sampling strips. *Tropical Conservation Science* 6, 506–520.
- Lukacs, P. M. (2002). Websim: simulation software to assist in trapping web design. *Wildlife Society Bulletin* 30, 1259–1261.
- Lukacs, P. M., A. B. Franklin, and D. R. Anderson (2004). Passive approaches to detection in distance sampling. pp. 260–280 in S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Lynch, H. J. and M. A. LaRue (2014). First global census of the Adélie penguin. *The Auk* 131, 457–466.
- Marini, F., B. Franzetti, A. Calabrese, S. Cappellini, and S. Focardi (2009). Response to human presence during nocturnal line transect surveys in fallow deer (*Dama dama*) and wild boar (*Sus scrofa*). *European Journal of Wildlife Research* 56, 233–237.
- Marques, F. F. C. and S. T. Buckland (2003). Incorporating covariates into standard line transect analyses. *Biometrics* 59, 924–935.
- Marques, F. F. C. and S. T. Buckland (2004). Covariate models for the detection function. pp. 31–47 in S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Marques, F. F. C., S. T. Buckland, D. Goffin, C. E. Dixon, D. L. Borchers, B. A. Mayle, and A. J. Peace (2001). Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. *Journal of Applied Ecology* 38, 349–363.



- Marques, T. A. (2004). Predicting and correcting bias caused by measurement error in line transect sampling using multiplicative error models. *Biometrics* 60, 757–763.
- Marques, T. A. (2007). *Incorporating Measurement Error and Density Gradients in Distance Sampling Surveys*. Ph. D. thesis, University of St Andrews.
- Marques, T. A., M. Andersen, S. Christensen-Dalsgaard, S. Belikov, A. Boltunov, Ø. Wiig, S. T. Buckland, and J. Aars (2006). The use of global positioning systems to record distances in a helicopter line-transect survey. *Wildlife Society Bulletin* 34, 759–763.
- Marques, T. A., S. T. Buckland, R. Bispo, and B. Howland (2013). Accounting for animal density gradients using independent information in distance sampling surveys. *Statistical Methods and Applications* 22, 67–80.
- Marques, T. A., S. T. Buckland, D. L. Borchers, D. Tosh, and R. A. McDonald (2010). Point transect sampling along linear features. *Biometrics* 66, 1247–1255.
- Marques, T. A., L. Munger, L. Thomas, S. Wiggins, and J. A. Hildebrand (2011). Estimating North Pacific right whale (*Eubalaena japonica*) density using passive acoustic cue counting. *Endangered Species Research* 13, 163–172.
- Marques, T. A., L. Thomas, S. G. Fancy, and S. T. Buckland (2007). Improving estimates of bird density using multiple covariate distance sampling. *The Auk* 124, 1229–1243.
- Marques, T. A., L. Thomas, S. W. Martin, D. K. Mellinger, S. Jarvis, R. P. Morrissey, C.-A. Ciminello, and N. DiMarzio (2012). Spatially explicit capture recapture methods to estimate minke whale abundance from data collected at bottom mounted hydrophones. *Journal of Ornithology* 152, 445–455.
- Marques, T. A., L. Thomas, S. W. Martin, D. K. Mellinger, J. A. Ward, D. J. Moretti, D. Harris, and P. L. Tyack (2013). Estimating animal population density using passive acoustics. *Biological Reviews* 88, 287–309.
- Marques, T. A., L. Thomas, J. Ward, N. DiMarzio, and P. L. Tyack (2009). Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales. *The Journal of the Acoustical Society of America* 125, 1982–1994.
- Marsh, H. and D. F. Sinclair (1989). Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *Journal of Wildlife Management* 53, 1017–1024.
- Marshall, L. (2014). *DSsim: Distance Sampling Simulations*. R package version 1.0.1.
- Martin, J., H. H. Edwards, M. A. Burgess, H. F. Percival, D. E. Fagan, B. E. Gardner, J. G. Ortega-Ortiz, P. G. Ifju, B. S. Evers, and T. J. Rambo (2012). Estimating distribution of hidden objects with drones: From tennis balls to manatees. *PLoS ONE* 7, e38882.
- Martin, S. W., T. A. Marques, L. Thomas, R. P. Morrissey, S. Jarvis, N. DiMarzio, D. Moretti, and D. K. Mellinger (2012). Estimating minke whale (*Balaenoptera acutorostrata*) boing sound density using passive acoustic sensors. *Marine Mammal Science* 29, 142–158.
- Mateos, M., G. M. Arroyo, A. Rodriguez, D. Cuenca, and A. De La Cruz (2010). Calibration of visually estimated distances to migrating seabirds with radar measurements. *Journal of Field Ornithology* 81, 302–309.
- McShea, W. J., C. M. Stewart, L. Kearns, and S. Bates (2011). Road bias for deer density estimates at 2 national parks in Maryland. *Wildlife Society Bulletin* 35, 177–184.
- Mennill, D. J., J. M. Burt, K. M. Fristrup, and S. L. Vehrencamp (2006). Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *The Journal of the Acoustical Society of America* 119, 2832–2839.
- Mesnick, S. L., F. I. Archer, A. C. Allen, and A. E. Dizon (2002). Evasive behavior of eastern tropical Pacific dolphins relative to effort by the tuna purse-seine fishery. Technical report, SWFSC/NOAA.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21, 1087–1091.
- Miller, D. L. (2013). *Distance: a Simple Way to Fit Detection Functions to Distance Sampling Data and Calculate Abundance/Density for Biological Populations*. R package version 0.7.3.
- Miller, D. L., M. L. Burt, E. A. Rexstad, and L. Thomas (2013). Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution* 4, 1001–1010.

- Mills, M. G. L., J. M. Juritz, and W. Zucchini (2001). Estimating the size of spotted hyaena (*Crocuta crocuta*) populations through playback recordings allowing for non-response. *Animal Conservation* 4, 335–343.
- Moretti, D., N. DiMarzio, R. Morrissey, J. Ward, and S. Jarvis (2006). Estimating the density of Blainville's beaked whale (*Mesoplodon densirostris*) in the Tongue of the Ocean (TOTO) using passive acoustics. In *Proceedings of the Oceans'06 MTS/IEEE-Boston*, Boston, MA.
- Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund (2000). A double-observer approach for estimating detection probability and abundance from point counts. *The Auk* 117, 393–408.
- Oedekoven, C. S. (2013). *Mixed Effect Models in Distance Sampling*. Ph. D. thesis, University of St Andrews.
- Oedekoven, C. S., S. T. Buckland, M. L. Mackenzie, K. O. Evans, and L. W. Burger (2013). Improving distance sampling: accounting for covariates and non-independency between sampled sites. *Journal of Applied Ecology* 50, 786–793.
- Oedekoven, C. S., S. T. Buckland, M. L. Mackenzie, R. King, K. O. Evans, and L. W. Burger (2014). Bayesian methods for hierarchical distance sampling models. *Journal of Agricultural, Biological, and Environmental Statistics* 19, 219–239.
- Oedekoven, C. S., J. L. Laake, and H. J. Skaug (2015). Distance sampling with a random scale detection function. *Environmental and Ecological Statistics*. [10.1007/s10651-015-0316-9](https://doi.org/10.1007/s10651-015-0316-9)
- Okot Omoya, E., T. Mudumba, S. T. Buckland, P. Mulondo, and A. J. Plumptre (2014). Estimating population sizes of lions *Panthera leo* and spotted hyaenas *Crocuta crocuta* in Uganda's savannah national parks using lure count methods. *Oryx* 48, 394–401.
- Petersen, I. K., T. K. Christensen, J. Kahlert, M. Desholm, and A. D. Fox (2006). Final results of bird studies at the offshore wind farms at Nysted and Horns Rev, Denmark. Technical report, National Environmental Research Institute Report.
- Petersen, I. K., M. L. MacKenzie, E. Rexstad, M. S. Wisz, and A. D. Fox (2011). Comparing pre- and post-construction distributions of long-tailed ducks *Clangula hyemalis* in and around the Nysted offshore wind farm, Denmark: a quasi-designed experiment accounting for imperfect detection, local surface features and autocorrelation. Technical Report 2011-1, CREEM.
- Plumptre, A. J. and D. Cox (2006). Counting primates for conservation: primate surveys in Uganda. *Primates* 47, 65–73.
- Plumptre, A. J. and V. Reynolds (1996). Censusing chimpanzees in the Budongo Forest, Uganda. *International Journal of Primatology* 17, 85–99.
- Plumptre, A. J. and V. Reynolds (1997). Nesting behavior of chimpanzees: implications for censuses. *International Journal of Primatology* 18, 475–485.
- Plumptre, A. J., E. J. Sterling, and S. T. Buckland (2013). Primate census and survey techniques. pp. 10–26 in E. Sterling, N. Bynum, and M. Blair (Eds.), *Primate Ecology and Conservation: A Handbook of Techniques*. Oxford: Oxford University Press.
- Pollard, E. and T. J. Yates (1993). *Monitoring Butterflies for Ecology and Conservation*. London: Chapman and Hall.
- Potts, J. M., S. T. Buckland, L. Thomas, and A. Savage (2012). Estimating abundance of cryptic but trappable animals using trapping point transects: a case study for Key Largo woodrats. *Methods in Ecology and Evolution* 3, 695–703.
- Reynolds, D. and J. Riley (2002). Remote-sensing, telemetric and computer-based technologies for investigating insect movement: a survey of existing and potential techniques. *Computers and Electronics in Agriculture* 35, 271–307.
- Rowcliffe, J. M., J. Field, S. T. Turvey, and C. Carbone (2008). Estimating animal density using camera traps without the need for individual recognition. *Journal of Applied Ecology* 45, 1228–1236.
- Rowcliffe, J. M., C. Carbone, P. A. Jansen, R. Kays, and B. Kranstauber (2011). Quantifying the sensitivity of camera traps: an adapted distance sampling approach. *Methods in Ecology and Evolution* 2, 464–476.
- Royle, J. A., D. K. Dawson, and S. Bates (2004). Modeling abundance effects in distance sampling. *Ecology* 85, 1591–1597.

- Royle, J. A. and R. M. Dorazio (2008). *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities*. Amsterdam: Elsevier, Academic Press. Code for book hosted at <http://www.mbr-pwrc.usgs.gov/pubanalysis/roylebook/chapters.htm>.
- Royle, J. A. and B. Gardner (2011). Hierarchical spatial capture-recapture models for estimating density from trapping arrays. pp. 163–190 in A. F. O'Connell, J. D. Nichols, and K. U. Karanth (Eds.), *Camera Traps in Animal Ecology: Methods and Analyses*. New York: Springer.
- Russell, R. E., J. A. Royle, V. A. Saab, J. F. Lehmkuhl, W. M. Block, and J. R. Sauer (2009). Modeling the effects of environmental disturbance on wildlife communities: avian responses to prescribed fire. *Ecological Applications* 19, 1253–1263.
- Savage, A., L. Thomas, K. Leighty, L. Soto, and F. S. Medina (2010). Novel survey method finds dramatic decline of wild cotton-top tamarin population. *Nature Communications* 1, 1–30.
- Schmaljohann, H., F. Liechti, E. Bachler, T. Steuri, and B. Bruderer (2008). Quantification of bird migration by radar — a detection probability problem. *Ibis* 150, 342–355.
- Schorr, R. A. (2013). Using distance sampling to estimate density and abundance of *Saussurea weberi* Hultén (Weber's saw-wort). *The Southwestern Naturalist* 58, 378–383.
- Schwaller, M. R., W. S. Benninghoff, and C. E. Olson (1984). Prospects for satellite remote sensing of Adélie penguin rookeries. *International Journal of Remote Sensing* 5, 849–853.
- Schwarz, C. J. and G. A. F. Seber (1999). Estimating animal abundance: review III. *Statistical Science* 14, 427–56.
- Schweder, T. (1996). A note on a buoy-sighting experiment in the North Sea in 1990. *Report of the International Whaling Commission* 46, 383–385.
- Schweder, T. (1997). Measurement error models for the Norwegian minke whale survey in 1995. *Report of the International Whaling Commission* 47, 485–488.
- Seber, G. A. F. (1982). *The Estimation of Animal Abundance, 2nd Ed.* London: Griffin.
- Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics* 42, 267–292.
- Seber, G. A. F. (1992). A review of estimating animal abundance II. *International Statistical Review* 60, 129–166.
- Shamoun-Baranes, J., J. Alves, S. Bauer, A. Dokter, O. Huppop, J. Koistinen, H. Leijnse, F. Liechti, H. van Gasteren, and J. Chapman (2014). Continental-scale radar monitoring of the aerial movements of animals. *Movement Ecology* 2, 9.
- Southwell, C., B. De la Mare, M. Underwood, F. Quartararo, and K. Cope (2002). An automated system to log and process distance sight-resight aerial survey data. *Wildlife Society Bulletin* 30, 394–404.
- Spear, L., N. Nur, and D. G. Ainley (1992). Estimating absolute densities of flying seabirds using analyses of relative movement. *The Auk* 109, 385–389.
- Spear, L. B. and D. G. Ainley (1997a). Flight behaviour of seabirds in relation to wind direction and wing morphology. *Ibis* 139, 221–233.
- Spear, L. B. and D. G. Ainley (1997b). Flight speed of seabirds in relation to wind speed and direction. *Ibis* 139, 234–251.
- Storm, D. J., M. D. Samuel, T. R. Van Deelen, K. D. Malcolm, R. E. Rolley, N. A. Frost, D. P. Bates, and B. J. Richards (2011). Comparison of visual-based helicopter and fixed-wing forward-looking infrared surveys for counting white-tailed deer *Odocoileus virginianus*. *Wildlife Biology* 17, 431–440.
- Strindberg, S. and S. T. Buckland (2004). Zigzag survey designs in line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 9, 443–461.
- Strindberg, S., S. T. Buckland, and L. Thomas (2004). Design of distance sampling surveys and geographic information systems. pp. 190–228 in S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford: Oxford University Press.
- Strobel, B. N. and M. J. Butler (2014). Monitoring whooping crane abundance using aerial surveys: influences on detectability. *Wildlife Society Bulletin* 38, 188–195.

- Summers, R. W. and S. T. Buckland (2011). A first survey of the global population size and distribution of the Scottish Crossbill *Loxia scotica*. *Bird Conservation International* 21, 186–198.
- Tasker, M. L., P. Hope Jones, T. Dixon, and B. F. Blake (1984). Counting seabirds at sea from ships: a review of methods employed and a suggestion for a standardized approach. *The Auk* 101, 567–577.
- Thomas, L., S. T. Buckland, E. A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R. Bishop, T. A. Marques, and K. P. Burnham (2010). Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* 47, 5–14.
- Thompson, M. E., S. J. Schwager, K. B. Payne, and A. K. Turkalo (2009). Acoustic estimation of wildlife abundance: methodology for vocal mammals in forested habitats. *African Journal of Ecology* 48, 654–661.
- Thomson, J. A., A. B. Cooper, D. A. Burkholder, M. R. Heithaus, and L. M. Dill (2012). Heterogeneous patterns of availability for detection during visual surveys: spatiotemporal variation in sea turtle dive-surfacing behaviour on a feeding ground. *Methods in Ecology and Evolution* 3, 378–387.
- Trenkel, V. M., S. T. Buckland, C. McLean, and D. A. Elston (1997). Evaluation of aerial line transect methodology for estimating red deer (*Cervus elaphus*) abundance in Scotland. *Journal of Environmental Management* 50, 39–50.
- Turchin, P. and F. J. Odendaal (1996). Measuring the effective sampling area of a pheromone trap for monitoring population density of southern pine beetle (coleoptera: Scolytidae). *Environmental Entomology* 25, 582–588.
- Tweh, C., M. Lormie, C. Kouakou, A. Hillers, H. Kuehl, and J. Junker (2015). Conservation status of chimpanzees (*Pan troglodytes verus*) and other large mammals across Liberia: results from a nationwide survey. *Oryx*. 10.1017/S0030605313001191.
- Tyack, P. L., M. Johnson, N. A. Soto, A. Sturlese, and P. T. Madsen (2006). Extreme diving of beaked whales. *Journal of Experimental Biology* 209, 4238–4253.
- Vargas, A., I. Jimenez, F. Palomares, and M. J. Palacios (2002). Distribution, status, and conservation needs of the golden-crowned sifaka (*Propithecus tattersalli*). *Biological Conservation* 108, 325–334.
- Vermeulen, C., P. Lejeune, J. Lisein, P. Sawadogo, and P. Bouché (2013). Unmanned aerial survey of elephants. *PLoS ONE* 8, e54700.
- Wall, C. C., C. Lembke, and D. A. Mann (2012). Shelf-scale mapping of sound production by fishes in the eastern Gulf of Mexico, using autonomous glider technology. *Marine Ecology Progress Series* 449, 55–64.
- Watts, A. C., J. H. Perry, S. E. Smith, M. A. Burgess, B. E. Wilkinson, Z. Szantoi, P. G. Ifju, and H. F. Percival (2010). Small unmanned aircraft systems for low-altitude aerial surveys. *Journal of Wildlife Management* 74, 1614–1619.
- Wegge, P. and T. Storaas (2009). Sampling tiger ungulate prey by the distance method: lessons learned in Bardia National Park, Nepal. *Animal Conservation* 12, 78–84.
- White, G. C., R. M. Bartmann, L. H. Carpenter, and R. A. Garrott (1989). Evaluation of aerial line transects for estimating mule deer densities. *Journal of Wildlife Management* 53, 625–635.
- Whitesides, G. H., J. F. Oates, S. M. Green, and R. P. Kluberanz (1988). Estimating primate densities from transects in a west African rain forest: a comparison of techniques. *Journal of Animal Ecology* 57, 345–367.
- Whittaker, D. G., W. A. V. Dyke, and S. L. Love (2003). Evaluation of aerial line transect for estimating pronghorn antelope abundance in low-density populations. *Wildlife Society Bulletin* 31, 443–453.
- Williams, R., S. L. Hedley, T. A. Branch, M. V. Bravington, A. N. Zerbin, and K. P. Findlay (2011). Chilean blue whales as a case study to illustrate methods to estimate abundance and evaluate conservation status of rare species. *Conservation Biology* 25, 526–535.
- Williams, R., R. Leaper, A. N. Zerbin, and P. S. Hammond (2007). Methods for investigating measurement error in cetacean line-transect surveys. *Journal of the Marine Biological Association of the United Kingdom* 87, 313–320.

- Yang, Z., T. Wang, A. K. Skidmore, J. de Leeuw, M. Y. Said, and J. Freer (2014). Spotting east African mammals in open savannah from space. *PLoS ONE* 9, e115989.
- Zaugg, S., G. Saporta, E. van Loon, H. Schmaljohann, and F. Liechti (2008). Automatic identification of bird targets with radar via patterns produced by wing flapping. *Journal of The Royal Society Interface* 5, 1041–1053.
- Zimmer, W. M. X., M. P. Johnson, P. T. Madsen, and P. L. Tyack (2005). Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*). *The Journal of the Acoustical Society of America* 117, 3919–3927.
- Zitterbart, D. P., L. Kindermann, E. Burkhardt, and O. Boebel (2013). Automatic round-the-clock detection of whales for mitigation from underwater noise impacts. *PLoS ONE* 8, e71217.

# Index

## A

Acoustic surveys, 189, 205  
     active, 198  
     passive, 199, 258  
 Adjustments, 57, 63  
     cosine, 63  
     Hermite polynomial, 64  
     simple polynomial, 64  
 Aerial surveys, 18, 38, 215, 224, 260  
 AIC, 56, 63  
 Assumptions, 12, 231, 253  
      $g(0)$ , 12, 36, 45, 93, 171, 205, 231, 256, 258  
     independence, 12, 45, 180, 190  
     measurement error, 12, 231, 243, 256  
     movement, 12, 35, 36, 188, 208, 231, 249, 251, 254, 260  
     random distribution of objects, 35, 37, 64, 223, 228, 231, 232, 255  
     shoulder criterion, 61, 63

## B

BACI design, 32, 132  
 Bayesian methods, 155, 160  
 Bias, 25, 232, 237  
      $g(0)$ , 45, 240  
     measurement error, 243  
     non-responsive movement, 37, 249  
     responsive movement, 44, 251  
     size bias in clusters, 46  
 Bootstrap, 110, 114, 131, 177

## C

Case studies

acoustic beaked whale survey, 217  
 amakihi, 4, 84, 121  
 BACI, 32, 132  
 conservation buffer, 4, 31, 158, 160  
 cork oaks, 71, 91  
 crabeater seals, 97  
 decay rate experiment, 182  
 double-observer songbird survey, 102  
 DSsim, 4, 26, 237  
 eastern grey kangaroos, 235  
 lure point transect, 4, 171  
 minke whales, 125  
 Montrave, 4, 6, 8, 68, 78, 112, 114, 190  
 spotted dolphins, 4, 135  
 trapping point transect, 174, 177  
 CDS, *see* Conventional distance sampling  
 Circular plots, 16, 108  
 Clearance plot method, 168, 179, 180  
 Clinometer, 38, 48  
 Clusters  
     ancillary data, 116  
     definition, 5  
 Coefficient of variation, 23  
 Confidence interval  
     bootstrap, 110, 134  
     log-based, 107  
     Satterthwaite df, 110  
 Conventional distance sampling, 4, 61, 109, 115  
 Covariates  
     count model, 134  
     detection function, 82  
     spatial, 131, 137  
 Cramér–von Mises test, 61  
 Cue counting, 187–192

Cut transects, 220  
 Cutpoints, 58, 67, 78

## D

### Data

grouped, 67  
 recording, 42  
 spiked, 43, 62  
 truncation, 5, 68  
 ungrouped, 64, 74

Density surface model, 137

### Design

bird surveys, 15  
 line transect surveys, 17  
 point transect surveys, 15  
 road surveys, 223, 256  
 sample size, 23  
 track surveys, 223, 256  
 zigzag samplers, 18

Design-based estimators, 105, 231

### Detection function

conventional distance sampling, 10, 64  
 multiple-covariate distance sampling, 82  
 shape criterion, 54

### Distance software

R package, 4  
 Windows version, 4, 46, 111

### Distances

measurement of, 48, 206, 212  
 truncation, *see* Data, truncation

Dung surveys, 179, 182

## E

Effective radius, 76  
 Effective strip half-width, 10, 65  
 Encounter rate, 106, 231

## F

Fast-moving animals, 250  
 Field methods, 35–42  
 Fisher information matrix, 67, 154  
 Full likelihood methods, 105, 141, 158

## G

$g(0)$ , 10, 12  
 Generalized additive modelling, 129–132  
 Generalized linear modelling, 129–132  
 Goodness-of-fit tests  
 $\chi^2$ , 59  
 Cramér–von Mises, 61

Kolmogorov–Smirnov, 60

Grouped distance data, *see* Data, grouped

## H

### Half-normal model

key function, *see* Key function, Half-normal  
 multiple covariate detection function, 83

### Hazard-rate model

key function, *see* Key function, Hazard-rate  
 multiple covariate detection function, 83

### Heaping, 43

Hermite polynomial model, *see* Adjustments,  
 Hermite

### Heterogeneity

covariates, 72, 85

## I

Impact assessment, 33  
 Indirect surveys, 179–187  
 Information matrix, *see* Fisher information  
 matrix  
 Interval estimation, *see* Confidence intervals

## K

### Key function

definition, 62  
 half-normal, 62  
 hazard-rate, 62  
 negative exponential, 62  
 uniform, 62

Kolmogorov–Smirnov test, *see* Goodness-of-fit  
 tests, Kolmogorov–Smirnov test

## L

Left-truncation, 196  
 Likelihood function  
 full likelihood methods, *see* Full likelihood  
 methods  
 half-normal, 65  
 Likelihood ratio test, 58  
 Line transect sampling, 10  
 examples, 68, 71  
 field methods, 37–42  
 survey design, 17  
 Lure point transects, 170

## M

Marine surveys, 35  
 Mark-recapture distance sampling, 40, 93–102  
 Maximum likelihood methods, 65, 153

MCDS, *see* Multiple-covariate distance sampling

MCMC, 155

Measurement error, 243–249

Measurements

angles, 48, 212

distances, *see* Distances, measurement

Model-based methods, 129, 137, 144, 146

Models

detection process, *see* Detection function, Conventional distance sampling

half-normal, *see* Key function, half-normal

hazard-rate, *see* Key function, hazard-rate

Hermite polynomial, *see* Adjustments, Hermite

measurement error, 244

plot counts, 128, 146

spatial, 131, 135

MRDS, *see* Mark-recapture distance sampling

Multiple-covariate distance sampling, 19, 82–92

## N

Negative exponential model, *see* Key function, negative exponential

## O

One-sided transects, 224

## P

Pilot study, 48

Plot sampling, 8, 106

Point counts, *see* Point transect sampling

Point transect sampling, 11

cues, 189

examples, 78, 114

field methods, 36

survey design, 15

Poisson models, 129

Polynomials

Hermite, 63

simple, 63

Pooling robustness, 54

Post-stratification, 22

Probability density function, 64, 75

## R

Radar surveys, 195

Rangefinder, 36, 48

Repeated measures, 31

Replicate lines or points

bootstrap, 110

survey design, 23

variance estimation, 110

Reticle, 48

Reversible-jump MCMC, 156

## S

Sample size, *see* Design, sample size

Satterthwaite correction, *see* Confidence interval, Satterthwaite df

Searching behaviour, 38, 39, 41

Shipboard surveys, 39

Simple polynomials, *see* Polynomials, simple

Snapshot method, 37, 202

Spatial models, *see* Models, spatial

Standing crop method, 181

Stratification, 19

Strip transect sampling, 8, 10, 65

Systematic design, 15, 111

## T

Training, 48

Trapping line transects, 169

Trapping point transects, 170

Trapping webs, 169

Truncation, *see* Data, truncation

Two-stage models, 128, 131, 132

## U

Uncertain detection on the line/point, 40, 45

## V

Variance estimation

bootstrap, *see* Bootstrap

Delta method, 110, 114

information matrix, *see* Fisher information matrix

## Z

Zigzag samplers, *see* Design, zigzag samplers