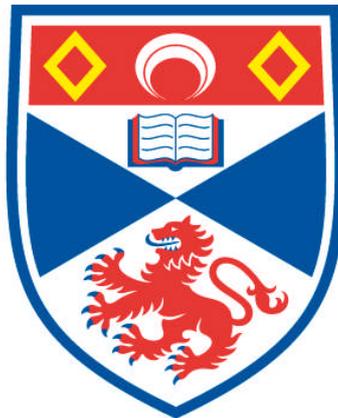


**ESTIMATING ABUNDANCE OF RARE, SMALL MAMMALS:  
A CASE STUDY OF THE KEY LARGO WOODRAT (*NEOTOMA  
FLORIDANA SMALLI*)**

**Joanne Marie Potts**

**A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews**



**2011**

**Full metadata for this item is available in  
Research@StAndrews:FullText  
at:**

**<http://research-repository.st-andrews.ac.uk/>**

**Please use this identifier to cite or link to this item:**

**<http://hdl.handle.net/10023/2068>**

**This item is protected by original copyright**

ESTIMATING ABUNDANCE OF RARE, SMALL  
MAMMALS: A CASE STUDY OF THE KEY LARGO  
WOODRAT (*NEOTOMA FLORIDANA SMALLI*)

Joanne Marie Potts



Thesis submitted for the degree of  
DOCTOR OF PHILOSOPHY  
in the School of Mathematics and Statistics  
UNIVERSITY OF ST ANDREWS  
ST ANDREWS  
DECEMBER 2010

© Copyright by Joanne Marie Potts, 2011

# Abstract

Estimates of animal abundance or density are fundamental quantities in ecology and conservation, but for many species such as rare, small mammals, obtaining robust estimates is problematic. In this thesis, I combine elements of two standard abundance estimation methods, capture-recapture and distance sampling, to develop a method called trapping point transects (TPT). In TPT, a “detection function”,  $g(r)$  (i.e. the probability of capturing an animal, given it is  $r$  m from a trap when the trap is set) is estimated using a subset of animals whose locations are known prior to traps being set. Generalised linear models are used to estimate the detection function, and the model can be extended to include random effects to allow for heterogeneity in capture probabilities. Standard point transect methods are modified to estimate abundance. Two abundance estimators are available. The first estimator is based on the reciprocal of the expected probability of detecting an animal,  $\hat{P}$ , where the expectation is over  $r$ ; whereas the second estimator is the expectation of the reciprocal of  $\hat{P}$ .

Performance of the TPT method under various sampling efforts and underlying true detection probabilities of individuals in the population was investigated in a simulation study. When underlying probability of detection was high ( $g(0) = 0.88$ ) and between-individual variation was small, survey effort could be surprisingly low (*c.* 510 trap nights) to yield low bias (*c.* 4%) in the two estimators; but under certain situations, the second estimator can be extremely biased. Uncertainty and relative bias in population estimates increased with decreasing detectability and increasing between-individual variation.

Abundance of the Key Largo woodrat (*Neotoma floridana smalli*), an endangered rodent with a restricted geographic range, was estimated using TPT. The TPT method compared well to other viable methods (capture-recapture and spatially-explicit capture-recapture), in terms of both field practicality and cost. The TPT method may generally be useful in estimating animal abundance in trapping studies and variants of the TPT method are presented.

# Declarations

I, Joanne Marie Potts, hereby certify that this thesis, which is approximately 51,200 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date: \_\_\_\_\_ signature of candidate: \_\_\_\_\_

I was admitted as a research student in September 2006 and as a candidate for the degree of Doctor of Philosophy in Statistics in September 2006; the higher study for which this is a record was carried out in the University of St Andrews between 2006 and 2010.

date: \_\_\_\_\_ signature of candidate: \_\_\_\_\_

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in Statistics in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date: \_\_\_\_\_ signature of supervisor: \_\_\_\_\_

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Embargo on both all or part of printed copy and electronic copy for the same fixed period of 1 year (maximum five) on the following ground: publication would preclude future publication.

date: \_\_\_\_\_

signature of candidate: \_\_\_\_\_

date: \_\_\_\_\_

signature of supervisor: \_\_\_\_\_

# Table of Contents

<b>Abstract</b>	<b>i</b>
<b>Declarations</b>	<b>ii</b>
<b>Table of Contents</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>xvii</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 The Key Largo woodrat . . . . .	4
1.2.1 Species description and life history . . . . .	4
1.2.2 Study area . . . . .	6
1.2.3 Species management . . . . .	10
1.3 Thesis outline . . . . .	11
<b>2 Methods for estimating small mammal abundance</b>	<b>15</b>
2.1 Introduction . . . . .	15
2.2 Distance sampling . . . . .	16
2.2.1 Line transect sampling . . . . .	17
2.2.2 Point transect sampling . . . . .	19
2.3 Capture-recapture . . . . .	20
2.4 Spatially-explicit capture-recapture . . . . .	26
2.5 Trapping webs . . . . .	30
2.6 Trapping point transects . . . . .	32
2.7 Indirect observations . . . . .	35
2.8 Occupancy modelling . . . . .	37
2.9 Methods applicable to estimating abundance of the Key Largo woodrat . . . . .	38
<b>3 Estimating Key Largo woodrat abundance using a capture-recapture   approach</b>	<b>46</b>
3.1 Introduction . . . . .	46

3.2	Methods . . . . .	49
3.2.1	Survey design . . . . .	49
3.2.2	Trapping and handling . . . . .	49
3.2.3	Model selection . . . . .	51
3.2.4	Abundance estimation . . . . .	56
3.3	Results . . . . .	57
3.4	Conclusion . . . . .	60
<b>4</b>	<b>Estimating Key Largo woodrat abundance using a spatially-explicit capture-recapture approach</b>	<b>63</b>
4.1	Introduction . . . . .	63
4.2	Methods . . . . .	65
4.2.1	Survey design, trapping and handling . . . . .	65
4.2.2	Data analysis . . . . .	65
4.3	Results . . . . .	70
4.3.1	Trapping and handling . . . . .	70
4.3.2	Data analysis . . . . .	70
4.4	Discussion and Conclusion . . . . .	72
<b>5</b>	<b>Estimating abundance using trapping point transects - methodology</b>	<b>75</b>
5.1	Introduction . . . . .	75
5.2	Survey design . . . . .	76
5.2.1	Main survey . . . . .	76
5.2.2	Trials survey . . . . .	77
5.3	Modelling of data . . . . .	79
5.3.1	Fitting the detection function . . . . .	79
5.3.2	Estimating abundance . . . . .	81
5.3.3	Variance estimation . . . . .	83
5.4	Discussion . . . . .	85
<b>6</b>	<b>A simulation study to investigate the performance of TPT under different scenarios</b>	<b>88</b>
6.1	Introduction . . . . .	88
6.2	Methods . . . . .	90
6.2.1	Three detection function scenarios . . . . .	90
6.2.2	Simulating the main survey . . . . .	93
6.2.3	Selecting trial distances . . . . .	94
6.2.4	Estimating abundance . . . . .	98
6.3	Simulation 1: Allocation of effort when underlying between-group heterogeneity in the population <i>is</i> accounted for . . . . .	101
6.3.1	Methods . . . . .	101
6.3.2	Results . . . . .	102

6.3.3	Conclusion of Simulation 1 . . . . .	106
6.4	Simulation 2: Allocation of effort when underlying between-group heterogeneity <i>is not</i> accounted for . . . . .	114
6.4.1	Methods . . . . .	114
6.4.2	Results . . . . .	114
6.4.3	Conclusion of Simulation 2 . . . . .	117
6.5	Simulation 3: Effort required to achieve unbiased abundance estimates . . . . .	123
6.5.1	Methods . . . . .	123
6.5.2	Results . . . . .	123
6.5.3	Conclusion of Simulation 3 . . . . .	128
6.6	Simulation 4: Including a behavioral response to being trapped . . . . .	129
6.6.1	Methods . . . . .	129
6.6.2	Results . . . . .	131
6.6.3	Conclusion of Simulation 4 . . . . .	132
6.7	Simulation 5: The importance of including a random effect . . . . .	137
6.7.1	Methods . . . . .	137
6.7.2	Results . . . . .	138
6.7.3	Conclusion of Simulation 5 . . . . .	139
6.8	General Conclusion . . . . .	142
6.8.1	Recommendations . . . . .	145
<b>7</b>	<b>Estimating Key Largo woodrat abundance using a trapping point transect approach</b> . . . . .	<b>147</b>
7.1	Introduction . . . . .	147
7.2	Methods . . . . .	148
7.2.1	Survey design . . . . .	148
7.2.2	Trapping and handling . . . . .	152
7.2.3	Fitting the detection function . . . . .	152
7.2.4	Estimating abundance . . . . .	154
7.2.5	Estimating variance in abundance . . . . .	156
7.3	Simulation study . . . . .	157
7.4	Results . . . . .	159
7.4.1	Key Largo woodrat survey and abundance estimation . . . . .	159
7.4.2	Simulation study . . . . .	167
7.5	Conclusion and Discussion . . . . .	168
<b>8</b>	<b>General Discussion and Conclusion</b> . . . . .	<b>172</b>
8.1	Estimating Key Largo woodrat abundance . . . . .	172
8.2	Extending the TPT simulation study . . . . .	177
8.3	Generalising the TPT method . . . . .	179
8.4	General conclusion . . . . .	182

<b>Bibliography</b>	<b>183</b>
<b>A Table of results from trapping point transect simulations (Chapter 6)</b>	<b>196</b>

# List of Figures

1.1	Picture of the Key Largo woodrat (photo courtesy of Melissa Olivencia, Disney’s Animal Kindgom). . . . .	5
1.2	Map of the south east United States of America, highlighting Key Largo. . . . .	8
1.3	Map of available habitat in northern Key Largo. . . . .	9
3.1	Map of trapping grid locations, used in the capture-recapture analysis. . . . .	50
3.2	Diagram of trapping grid layout. . . . .	51
4.1	A map of the 32 trapping grid locations (solid black dot), overlaid on the habitat mask (shaded grey) used in the spatially-explicit capture-recapture analysis. . . . .	69
4.2	The estimated density of woodrats per hectare in the survey region, shown as solid lines, and 95% confidence intervals (dashed lines). . . . .	73
5.1	Schematic of how one individual trial survey is conducted in the field. . . . .	78
6.1	Plot of the three detection function scenarios used to simulate unique detection functions for each individual in the main and trial surveys. . . . .	92
6.2	Plot of a single realisation of the main survey, for each of the three detection function scenarios. . . . .	95
6.3	Plot of a single realisation of a trial survey with 40 trials per individual, using the “Uniform” (Panel A), “Stopping Rule 5” (Panel B) and “Stopping Rule 8” (Panel C) method for selecting trial distances. . . . .	99
6.4	Plot of a single realisation of a trial survey with 40 trials per individual, using the “Adaptive” method for selecting trial distances.100	

6.5	Plot of $\hat{N}_2$ and $\hat{N}_3$ with 95% percentile intervals and percentage bias estimates, with varying levels of survey effort allocated to the number of trials per individual, and the number of individuals included in the trials survey when the “High” detection function scenario is used. . . . .	108
6.6	Plot demonstrating that at high average detection probabilities and at small sample sizes, abundance estimates can be positively biased when the “Adaptive” method is used to select trial distances.	109
6.7	Plot demonstrating that at high average detection probabilities and at small sample sizes, abundance estimates can be negatively biased when the “Uniform” or two “Stopping Rule” methods are used to select trial distances. . . . .	110
6.8	Plot of $\hat{N}_2$ and $\hat{N}_3$ with 95% percentile intervals and percentage bias estimates, with varying levels of survey effort allocated to the number of trials per individual, and the number of individuals included in the trials survey when the “Medium” detection function scenario is used and underlying group heterogeneity is accounted for. . . . .	111
6.9	Plot of $\hat{N}_2$ and $\hat{N}_3$ with 95% percentile intervals and percentage bias estimates, with varying levels of survey effort allocated to the number of trials per individual, and the number of individuals included in the trials survey when the “Low” detection function scenario is used and underlying group heterogeneity is accounted for. . . . .	112
6.10	Plot demonstrating that at lower average detection probabilities and at small sample sizes, abundance estimates can be positively biased. . . . .	113
6.11	Plot of $\hat{N}_2$ and $\hat{N}_3$ with 95% percentile intervals and percentage bias estimates, with varying levels of survey effort allocated to the number of trials per individual, and the number of individuals included in the trials survey when the “Medium” detection function scenario is used and underlying group heterogeneity is not accounted for in the analysis. . . . .	119
6.12	Plots of mean abundance estimates from estimators $\hat{N}_2$ and $\hat{N}_3$ for the “Medium” detection function scenario when underlying group heterogeneity was ignored. These results are for the “Adaptive” method of selecting trial distances, when 60 trials were conducted per individual, and 6 individuals were used in the trials survey. .	120
6.13	Plot of $\hat{N}_2$ and $\hat{N}_3$ with 95% percentile intervals and percentage bias estimates, with varying levels of survey effort allocated to the number of trials per individual, and the number of individuals included in the trials survey when the “Low” detection function scenario is used and underlying group heterogeneity was not accounted for in the analysis. . . . .	121

6.14	Plots of mean abundance estimates from estimators $\hat{N}_2$ and $\hat{N}_3$ for the “Low” detection function scenario when underlying group heterogeneity was ignored. These results are for the “Adaptive” method of selecting trial distances, when 60 trials were conducted per individual, and 6 individuals were used in the trials survey. . . . .	122
6.15	Plots of mean abundance estimates from estimators $\hat{N}_2$ and $\hat{N}_3$ (Panel A and B, respectively, shown as a solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort changed for the “Medium” detection function scenario and when the “Adaptive” method was used to select trial distances. True population abundance was 2,000 (shown as a dotted line in Panel A and B). . . . .	125
6.16	Plots of mean abundance estimates from estimators $\hat{N}_2$ and $\hat{N}_3$ (Panel A and B, respectively, shown as a solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort changed for the “Medium” detection function scenario, when the “Adaptive” method was used to select trial distances, and underlying heterogeneity in the probability of detection was ignored during the analysis. True population abundance was 2,000 (shown as a dotted line in Panel A and B). . . . .	127
6.17	Plot of behavioral response when the “High” detection function scenario was used. When an individual has not been captured in the previous trial, the “Mean normal behavior” detection function applies (same parameters as the “High” detection function scenario). However, if the animal was captured in the previous trial, the “Mean behavior immediately after capture” detection function applies, which reduced the probability of capture by 50%. . . . .	130
6.18	Plot of behavioral response when the “Low” detection function scenario was used. When an individual has not been captured in the previous trial, the “Mean normal behavior” detection function applies (same parameters as the “Low” detection function scenario). However, if the animal was captured in the previous trial, the “Mean behavior immediately after capture” detection function applies, which reduced the probability of capture by 50%. . . . .	131
6.19	Plot of abundance estimators $\hat{N}_2$ and $\hat{N}_3$ (and 2.5 and 97.5 percentiles) using the “Low” and “High” detection function scenarios with and without the inclusion of a behavioral response. . . . .	133
6.20	Plot of mean abundance estimate $\hat{N}_2$ for the “High” detection function scenario, when the probability of detection decreased by 50% if the individual was detected in the previous trial. . . . .	135

6.21	Plot of mean abundance estimate $\hat{N}_2$ for the “Low” detection function scenario, when the probability of detection decreases by 50% if the individual was detected in the previous trial. . . . .	136
6.22	Plot of mean abundance estimate $\hat{N}_2$ for the “Low” detection function scenario, when the repeated trial observations on each individual are incorrectly assumed independent, and the underlying group heterogeneity is not accounted for in the model. . . .	141
7.1	Map of main survey trap locations used in the trapping point transect analysis. . . . .	151
7.2	Plot of 80 habitat sample points centered around trap location #47 in the main survey. This trap was located on the western boundary of the survey region, such that five points did not occur within habitat (grey) and 75 did (black). . . . .	156
7.3	A single realisation of the main survey simulation, based on the estimated detection function using woodrat data. The main survey was a 10 trap by 15 trap grid (trap location shown as an ‘x’), with a 250 m horizontal and vertical inter-trap spacing located in a survey area of 3,000 m by 4,300 m. . . . .	160
7.4	Plots of the two mean detection functions with the lowest AIC (Table 7.4) fitted to trial data collected in 2008 (solid line) and 2009 (dotted line). Model 27 included sex and an interaction term between sex and year as a covariate (males and females are shaded black and grey, respectively). See Table 7.4 for full model structure.	161
7.5	Mean detection function plots corresponding to the three models with the lowest AIC fitted to the female trial data collected in 2008 (solid line) and 2009 (dotted line). See Table 7.5 for full model structure. . . . .	164
7.6	Plot of the simple mean detection function fitted to the male trial data, where the probability of detection only depended on distance to trial trap and an individual random effect. . . . .	165
7.7	Plot of the availability of habitat with increasing distance from the main survey point when edge effects are accounted for (dashed line), or not (solid line). . . . .	166
7.8	Plot of the detection function scenario used to simulate unique detection functions for each woodrat in the main and trial surveys.	167
8.1	Within the animal’s home range (denoted by a grey dotted line), is its nest, located at C. The animal was caught (and released) at R, and upon release, is assumed to return to its nest. A trial trap was then set at T. $d$ is the distance between the animal’s point-of-release, and the trial trap. $D$ is the distance between the animal’s start location when the trial trap was set (e.g., its nest) and the trial trap. . . . .	180

8.2 Within the animal's home range (denoted by a grey dotted line), is its nest, located at  $C$ . The animal was caught (and released) at  $R$ , and upon release, is assumed to return to its nest. If a trial trap is placed at distance  $d$  from the point-of-release  $R$ , some trials will be set beyond the animal's home range boundary (dashed line) and some will be set inside the animal's home range boundary (solid line). As  $d$  increases, the proportion of trials occurring outside the animal's home range increases, but some trials will be set very close to the animal's home range center ( $C$ ). . . . . 182

# List of Tables

1.1	Summary of results for previous studies of the Key Largo woodrat.	12
2.1	Three forms of the capture probability model used in spatially-explicit capture-recapture. $d$ is the distance between an animals home range centre and a detector. The parameter $g_0$ is common to all functions and represents the probability of detection at a single detector placed in the centre of the home range; values of the spatial scale parameter $\sigma$ are not comparable between functions (Efford et al., 2009b).	27
3.1	Definition of terms used in model formulae presented in Tables 3.2 and 3.3. There were 50 parameters in the full, time dependent model relating to survival ( $S$ ), migration ( $\gamma''$ and $\gamma'$ ), the two-part mixture process for heterogeneity ( $\pi$ ), and probability of capture ( $p$ ) and recapture ( $c$ ).	53
3.2	Summary of the models fitted to trapping data collected on the Key Largo woodrat. Models included the probability of capture ( $p$ ) and probability of recapture ( $c$ ) (dependent on trapping session, $sess$ , sex of woodrat, and habitat, $hab$ ). Each of these ten models included the probability of temporary emigration being either random ( $\gamma'' = \gamma'$ ) or not (i.e., no movement $\gamma'' = \gamma' = 0$ ) (i.e., 20 models were run in total, but for brevity I report only 10). Survival between sessions was always assumed constant.	54
3.3	Summary of the 11 models fitted to the trapping data collected on the Key Largo woodrat. Models included the probability of capture ( $p$ ) and probability of recapture ( $c$ ) (dependent on trapping session, $sess$ , and sex of woodrat), probability of temporary emigration being random ( $\gamma'' = \gamma'$ ) or not (i.e., no movement $\gamma'' = \gamma' = 0$ ), and heterogeneity ( $g$ ) based on a 2-point mixture process. Survival and the 2-point mixture process was always assumed constant across each session.	55

3.4	Number of woodrat capture events for 3 primary sessions (Mar-Apr, Jul-Sept, Nov-Dec) across 3 habitat strata. The number of individual woodrats captured are given in parentheses. Row totals (i.e., by strata) for the number of woodrats caught are not a direct summation of each row (highlighted by an ‡), as some woodrats were caught across multiple sessions. . . . .	57
3.5	Summary of 29 models were fitted to the Key Largo woodrat capture data. Models included the probability of capture ( $p$ ) and probability of recapture ( $c$ ) (dependent on trapping session, <i>sess</i> , and sex of woodrat), probability of temporary emigration being random ( $\gamma'' = \gamma'$ ) or not (i.e., no movement $\gamma'' = \gamma' = 0$ ), and heterogeneity ( $g$ ) based on a 2-point mixture process. Survival and the 2-point mixture process was always assumed constant across each session. Eight models had a difference in AIC of less than 2 from the top model. AIC weights (wAICc) and the number of parameters (No. Par.) in each model are also provided. . . . .	59
3.6	Key Largo woodrat abundance estimates using the four different methods outlined in Section 3.2.4 to convert the population size in the covered area (i.e., effective area trapped by each grid) to a total population size estimate for the entire survey region (850 ha of habitat). MMDM is mean maximum distance used. . . . .	60
4.1	Three forms of the capture probability model. The probability of capturing an animal if the trap is placed 0 m from its home range centre is given by $g_0$ . The shape of the capture probability model is given by $\sigma$ . $d$ is the distance between an animals home range centre and a detector. $\theta$ is a vector of parameters containing $g_0$ , $\sigma$ and $b$ (if the hazard rate model is specified), for which maximum likelihood estimates are sought. . . . .	67
4.2	Covariates considered during model selection for fitting the “capture probability” model (Efford, 2010). . . . .	67
4.3	Twenty-six models were fitted to the data based on AIC, that incorporated behavior (b and B), session, time (t), sex and habitat (hab) covariates (see Table 4.2 for complete covariate definitions). All models were fitted using the exponential form of the capture function, unless otherwise stated. . . . .	71
6.1	Input parameters for the three detection function scenarios (“High”, “Medium”, “Low”). . . . .	93
6.2	Number of individuals upon which trials were performed, based on four different levels of trials survey effort (360, 720, 1080 and 1440 trap nights), and the number of trials per individual (6, 10, 18, 24, 40, 60). . . . .	124

6.3	$\hat{N}_2$ and $\hat{N}_3$ abundance estimator results with and without a trap shyness behavioral response for the “High” and “Low” detection function scenarios. . . . .	134
6.4	$\hat{N}_2$ and $\hat{N}_3$ abundance estimates when the random effects component of the data structure is ignored (i.e., individual trials are assumed independent of each other, “No RE”) for the “High” and “Low” detection function scenarios (24 trials on 15 individuals, respectively). . . . .	140
6.5	Results for $\hat{N}_2$ and $\hat{N}_3$ abundance estimators when the random effects component of the data structure is ignored (i.e., individual trials are assumed independent of each other) and the group covariate for the “Medium” and “Low” detection function scenarios is also ignored. . . . .	140
6.6	Summary of input parameters for the five simulation studies. . .	143
7.1	Definitions of explanatory variables used in estimating the detection function for the trapping point transect analysis. . . . .	152
7.2	The three random effects structures considered during model selection, presented using R notation . . . . .	154
7.3	The five simulation scenarios used to investigate the consequences of changing the number of visits to each trap location in the main survey ( <i>Main survey visits</i> ) and the number of woodrats upon which trials were performed ( <i>No. woodrats</i> ). . . . .	159
7.4	The 27 different detection functions fitted to the 512 trial observations collected on male and female woodrats in 2008 and 2009. . . . .	162
7.5	The 15 different detection functions fitted to the 288 trial observations collected on female woodrats in 2008 and 2009. . . . .	164
7.6	Abundance estimates and 95% percentile intervals for female and male woodrats in 2008 and 2009 using two different abundance estimators. For consistency, the percentile intervals were reported for male woodrats in 2009 as 0, but these are clearly biased low, and perhaps the minimum number of male woodrats captured in 2009 would better reflect the uncertainty in population size estimates. . . . .	166
7.7	The results for $\hat{N}_2$ and $\hat{N}_3$ abundance estimators when total survey effort is increased and allocation between the main and trials survey changes, by varying the number of times each trap location in the main survey was visited ( <i>Visits</i> ) and the number of individuals that trials were performed on ( <i>WR</i> ). . . . .	169
8.1	Summary of Key Largo woodrat abundance estimates using capture-recapture, spatially-explicit capture-recapture and trapping point transects. . . . .	175

A.1	Input parameters for the three detection function scenarios (“High”, “Medium”, “Low”). . . . .	196
A.2	Summary of input parameters for Simulations 1, 2 and 3. Three detection function scenarios were used (“High”, “Medium”, and “Low”) in Simulations 1 and 2, and only the “Medium” scenario was used in Simulation 3. The fitted detection function (Model) used in the analysis was either correct or not. The number of trials in the trials survey was either 360 (Simulation 1 and 2), or 780, 1080 and 1440 (Simulation 3). The method used to select the trial distances was either all of those considered (i.e., “Uniform”, “Stopping rule 5”, “Stopping rule 8” and “Adaptive”; Simulation 1 and 2) or just “Adaptive” (Simulation 3). . . . .	197
A.3	Table of results for all scenarios in Simulation 1. That is, $\hat{N}_2$ and $\hat{N}_3$ with associated percentage bias (pc.bias), standard deviation estimates (sd), percentile intervals (PI), median and root mean square error (RMSE) when underlying heterogeneity is accounted for, using the “High”, “Medium” and “Low” detection function scenarios, with differing effort allocated per individual (No.tri) and number of individuals (No.ind) in the trials survey. . . . .	198
A.4	Table of results for all scenarios in Simulation 2. That is, $\hat{N}_2$ and $\hat{N}_3$ with associated percentage bias (pc.bias), standard deviation estimates (sd), percentile intervals (PI), median and root mean square error (RMSE) when underlying heterogeneity is <i>not</i> accounted for, using the “High”, “Medium” and “Low” detection function scenarios, with differing effort allocated per individual (No.tri) and number of individuals (No.ind) in the trials survey. . . . .	200
A.5	Table of results for all scenarios in Simulation 3. That is, $\hat{N}_2$ and $\hat{N}_3$ with associated percentage bias (pc.bias), standard deviation estimates (sd), 2.5 and 97.5 percentile intervals (PI), median and root mean square error (RMSE) when underlying heterogeneity is accounted for (Model=Include) or not (Model=Ignore), using the “Medium” detection function scenario, with differing effort allocated per individual (No.tri) and number of individuals (No.ind) in the trials survey. . . . .	202

# Acknowledgements

The work presented in this thesis couldn't have been possible without the following:

- Intellectual guidance from my supervisors, Steve Buckland, Len Thomas and Anne Savage. Countless conversations with Steve and Len, some at particularly last minute notice for which they always made themselves available, were crucial to my understanding about the problems involved in estimating animal abundance. They have shaped my future career and I hope to work with them for many years to come. Anne provided excellent support during my field work and funded the majority of my PhD - without her support, this project would never have existed. Thank you.
- My knowledge about spatially explicit capture-recapture was furthered by conversations with David Borchers and Murray Efford, who always made themselves available and replied to my e-mails very promptly. Thank you.
- Additional funding for my PhD was provided via an Overseas Research Student Award from the University of St Andrews.
- Funding for my field work to collect data presented in Chapter 7 was provided by three sources:
  - Disney's Animal Kingdom. Staff from the Wildlife Tracking Centre and Night Team were field assistants during data collection and I appreciated everyone's excitement and enthusiasm to be involved in the project. I would particularly like to make a special mention to Christy Alligood, Andy Daneault, Mark Bechtel, Melissa Olivencia, Tom Dillenbeck and Rob Carlson who are wonderful people to work with.
  - the US Fisheries and Wildlife Service. I would particularly like to thank Steve Klett and Sandra Sneckenberger for continual support and granting access to Key Largo woodrat habitat, and Pat Wells from the Florida Parks Service. Britta Muiznieks provided training.
  - John Harwood at CREEM provided additional funds to complete my second season of field work.

- Additional field work assistance was provided by Laura Marshall, Lindesay Scott-Hayward, Karen Garrod and Sarah Dewees, and Clay and Ralph DeGayner. They were all fantastic field assistants.
- Data analysed in Chapters 3 and 4 were collected by Dan Greene, assisted by Jessica Demarco, for his Master's project, supervised by Michael Mengak and Steven Castleberry at the University of Georgia.
- Everyone in my personal life! The list is great: my wonderful office mates, Lindesay Scott-Hayward, Laura Marshall and Glenna Evans. My wonderful coffee buddies, Tiago Marques (great friendship and support), Leslie New (cake! and also proof reading some of my chapters), Debbie Russell, Angelika Studeny, Bruno Caneco, Liliana Martins, Rodrigo Wiff, Monique Mackenize, Carl Donovan, Vanessa Cave, John Bishop, Danielle Harris, Joyce Yuan, Charles Paxton, Calum Brown, Theoni Photopoulou, Lorenzo Milazzo, Cornelia Oedekoven, Eric Rexstad, Jill Calder and Tom Brown for listening and various important discussions. Supporting colleagues at CREEM included Rhona Rodgers (for fixing up my field budgets), and Phil le Furve (for fixing my computer). There are also lots of people at the University of Melbourne, especially Mark Burgman and Jane Elith who kick started my interest in statistical ecology, and Brendan Wintle for random discussions on the importance (or lack thereof) of monitoring.
- A special mention of love and thanks goes to Clay and Barb DeGayner. From when I first stepped off the plane and took 3 hours to drive from Miami to Key Largo on the “wrong” side of the road and an in automatic car while you waited anxiously for me... wow. We have come a long way. Meeting you both has been a wonderful blessing.
- My dad and family in Australia. I think you will be as relieved as me, now this mammoth thesis is done. I miss you lots and hope to be living closer to you soon. I especially want to mention my wonderful mum, I would exchange the world for one more hug.
- Two final thanks go to my husband Martin and our daughter Annabelle. Martin, no words can describe my gratitude for your unwavering support in finishing my thesis, proof reading my chapters, raising our daughter, and coping with the loss of mum. You helped me so much, I doubt I could have finished this thesis and survived these past years without you. You are the love of my life. Annabelle, your birth was a blessing. You have provided a welcome and happy distraction to the trials and tribulations of life and taught me there is more to life than stressing over my thesis. Much love.

# Chapter 1

## General introduction

### 1.1 Overview

Monitoring programs designed to estimate animal abundance are required to assess a species' population status or to determine whether management actions, such as increasing population size of the target species (Solari et al., 2002), are having the desired effect. Obtaining accurate abundance estimates can be complicated because the biology of the target species might be poorly understood, the species might be rare (Witmer, 2005), and there are typically financial and practical limitations. Such uncertainties and limitations may result in monitoring programs failing to provide the requisite information to managers to assist with their decision making (Vos et al., 2000).

Although the complete design of a population monitoring program is complex and species specific, some general principles are applicable which will increase the likelihood of the program's success. Firstly, the importance of clearly articulating the objective of the monitoring program, such as estimating animal abundance, so that it is understood by those working on the project, cannot be over-emphasized (Vos et al., 2000; Wilson and Delahay, 2001; Oakley et al., 2003; Wintle et al., 2010). This involves identifying the study area, the population of interest and the time period in which the program will operate.

The second step in developing a monitoring program is to devise a sampling strategy. The geographical area containing the species is divided into units, of which a subset is chosen using a probability based selection method (e.g., simple random sampling, Williams et al., 2002). For the third step, the survey method is selected. The survey method is applied within the sampling units, with effort allocated over space (i.e., increasing the number of sampling units visited) and time (i.e., by visiting sampling units multiple times). Many survey methods are available (e.g., distance sampling, Buckland et al., 2001; or capture-recapture, Pollock et al., 1990). These methods all have their inherent strengths and weaknesses that must be weighed against the logistical constraints of the study area, applicability of the method to the target species and funding limitations.

Once the monitoring program has been undertaken, data must be collated and analysed. At this point, the monitoring program should be assessed to determine if it is meeting its objectives, and if not, the program can be modified in an iterative process (Peterson et al., 1997; Vos et al., 2000; Solari et al., 2002; Schreiber et al., 2004). For example, the monitoring program could be altered to ensure uncertainty in parameter estimates is reduced (McCarthy and Possingham, 2007; Rout et al., 2009) or learning about the system processes is maximised (McDonald-Madden et al., 2010; Moore and McCarthy, 2010).

The most commonly implemented survey designs and monitoring methods to estimate abundance are geared toward moderately abundant to abundant species and do not translate well to rare, elusive or cryptic species (Thompson et al., 1998). However, it is these species that often require monitoring due to low abundance, limited geographic range, and/or limited reliable knowledge about the population status or trend of the species (e.g., risk of extinction). To monitor these species, ingenuity, sound methodology and technology are essential to developing new approaches in monitoring (Thompson et al., 1998).

One such species is the eastern woodrat (*Neotoma floridana*). The endangered Key Largo woodrat (*N. f. smalli*) is restricted to 850 ha of remnant vegetation on northern Key Largo, Florida, USA. Reliable population size and density estimates, together with associated measures of uncertainty, are essential for successful management of the woodrat population (USFWS, 1999). The development of monitoring methods for the woodrat presents particular challenges, since they are nocturnal, cryptic, and sparsely distributed.

Two commonly implemented methods (standard distance sampling and capture-recapture) do not work well when trying to estimate abundance of the woodrat. With regards to distance sampling, visual detection is very low. Hence a key assumption of distance sampling, that detectability of animals on the transect line are always detected, is violated. In addition, for realistic levels of search effort, sample size is likely to be very small. Thus, reliable estimation of the detection function, the probability of detecting an animal as a function of distance from an observer walking a transect, in a distance sampling approach is problematic. Detection is therefore more likely to occur when trapping individuals, and such data is typically analysed within a capture-recapture framework. Robust estimation of abundance using capture-recapture data is problematic, because animals beyond the extent of the trapping array are available for capture, and methods for defining this “effective trapping area” of the array are lacking (Efford et al., 2009b).

In this thesis, I use trapping point transects (TPT) to estimate abundance of the Key Largo woodrat. Trapping point transects combine elements of capture-recapture and distance sampling, thus enabling estimation of the effective trapping area to obtain robust estimates of abundance. Previous attempts at combining distance sampling with capture-recapture have included trapping webs (Anderson et al., 1983), however a large amount of survey effort is required to

implement this survey method (Lukacs et al., 2005). TPT potentially allows abundance of a species to be estimated with fewer resources than other methods (Buckland et al., 2006).

I compare abundance estimates obtained using TPT with two other methods (capture-recapture and spatially-explicit capture-recapture, Borchers and Efford, 2008). By assessing the bias and precision of abundance estimates, together with the practicality and cost efficiency of these three methods, I found the TPT to be the most suitable survey technique for developing a long-term monitoring program of woodrat abundance. The TPT method may be useful in estimating abundance of other animals where detection of individuals relies on trapping, especially in situations where detectability is low.

## **1.2 The Key Largo woodrat**

### **1.2.1 Species description and life history**

The Key Largo woodrat is one of six recognised subspecies of the eastern woodrat (*Neotoma floridana*), first discovered in 1923 (Small, 1923) and later reclassified in 1955 (Sherman, 1955; Schwartz and Odum, 1957). Adult body length ranges between 120-230 mm and is the body covered in grey-brown hair above with white ventral colouration (USFWS, 1999). They are characterised by large ears, protuberant eyes and a short tail (Figure 1.1).



Figure 1.1: Picture of the Key Largo woodrat (photo courtesy of Melissa Olivenca, Disney's Animal Kingdom).

Individual woodrats maintain multiple nest locations within their home range territories (Hersh, 1981; Sherman, 1955; Schwartz and Odum, 1957). Home range size is approximately 0.25 ha, but has been found to vary with sex (females < males), time of year (breeding season v's non-breeding season) (Vestral, 1938; Hersh, 1981; McCleery, 2003; Gore and Loggins, 2005; Lee and Tietje, 2005) and probably population density. Nesting substrate varies (e.g., the root system of upturned trees, rock piles, burrows, cars, machinery and piles of rubbish), but all usually have one common feature: a pile of sticks is constructed by the occupying woodrat on or around its nesting location. Nests are used for sleeping during daylight (as they are nocturnal), protection from inclement weather and predators, caching food, and raising young.

Key Largo woodrats are believed to breed throughout the year, but peaks

in the spring and summer have been noted (English, 1923; Hersh, 1981). Female woodrats can produce two to three litters per year (Gander, 1929; Gore and Loggins, 2005). Litter size ranges between one and four offspring, with two most common (English, 1923). Both sexes take approximately five months to reach maturity (Hersh, 1981). The life expectancy of the Key Largo woodrat is unknown. In the closely related dusky-footed woodrat (*Neotoma fuscipes*), survival probability severely diminishes after two years of age (Lee and Tietje, 2005). From field observations, we know most Key Largo woodrats are not typically recaptured more than two years after first recapture, however one individual has been recaptured five years after first capture (Potts unpubl. data).

### 1.2.2 Study area

The Key Largo woodrat has a highly restricted geographic range, occupying extant, tropical hardwood hammock on northern Key Largo, Florida, USA (Figure 1.2). Key Largo is the first island in the Florida Keys, linking the lower keys to the mainland. It is the largest island of the keys, covering approximately 21 km<sup>2</sup>. Key Largo has an average elevation of 2.4 m above sea level and is formed from the exposed tops of ancient coral reefs (Hersh, 1981). The climate of Key Largo is subtropical. Mean annual rainfall is approximately 102 cm (40 inches), with the majority of this precipitation occurring between June and October (US Fish and Wildlife Service 1999b). Average maximum daily temperature in summer is approximately 29 °C (84 °F), decreasing to approximately 21 °C (70 °F) in winter (Gore and Loggins, 2005).

Although tropical hardwood hammock was once the dominant vegetation type across Key Largo, due to urbanization, approximately only one third remains within two protected reserves that fall under different jurisdiction: the Dagny Johnson Key Largo Hammock Botanical State Park, managed by the Florida Department of Environment and Protection, and the Crocodile Lake

National Wildlife Refuge, managed by the U.S. Fish and Wildlife Service (Figure 1.2). Despite its protected status, the remaining tropical hardwood hammock is highly fragmented with roads, tracks, and abandoned developments. In particular, County Road 905 (a major 2-lane highway, Figure 1.2) separates the two reserves and has been identified as a barrier to Key Largo woodrat movement (McCleery, 2003).

The age and structure of the hardwood hammock is not homogeneous throughout Key Largo (Gore and Loggins, 2005). Ross et al. (1992) and Ross et al. (1995) identified three strata delineated by age: young (disturbed since 1971,  $\approx 87$  ha), medium (disturbed between 1940-71,  $\approx 327$  ha), and old (disturbed before 1940,  $\approx 431$  ha); and a fourth strata based on level of disturbance and urban areas located within the hammock ( $\approx 127$  ha, representing  $\approx 13\%$  of the total habitat, Figure 1.3). Vegetative composition and structure may influence the density and distribution of woodrats by affecting their ability to find food resources, nest materials, and secure cover for travel both in trees and across the forest floor (Barbour and Humphrey, 1982; USFWS, 1999). However, the affinity of woodrats to a particular age class of hammock remains unknown, as previous research is conflicting (Barbour and Humphrey, 1982; McCleery, 2003). The importance of other habitat characteristics such as proximity to artificial nest substrate, surrounding density of feral cats (*Felis catus*), and raccoons (*Procyon lotor*), and density of the canopy have been demonstrated to influence woodrat occupancy (Winchester et al., 2009).

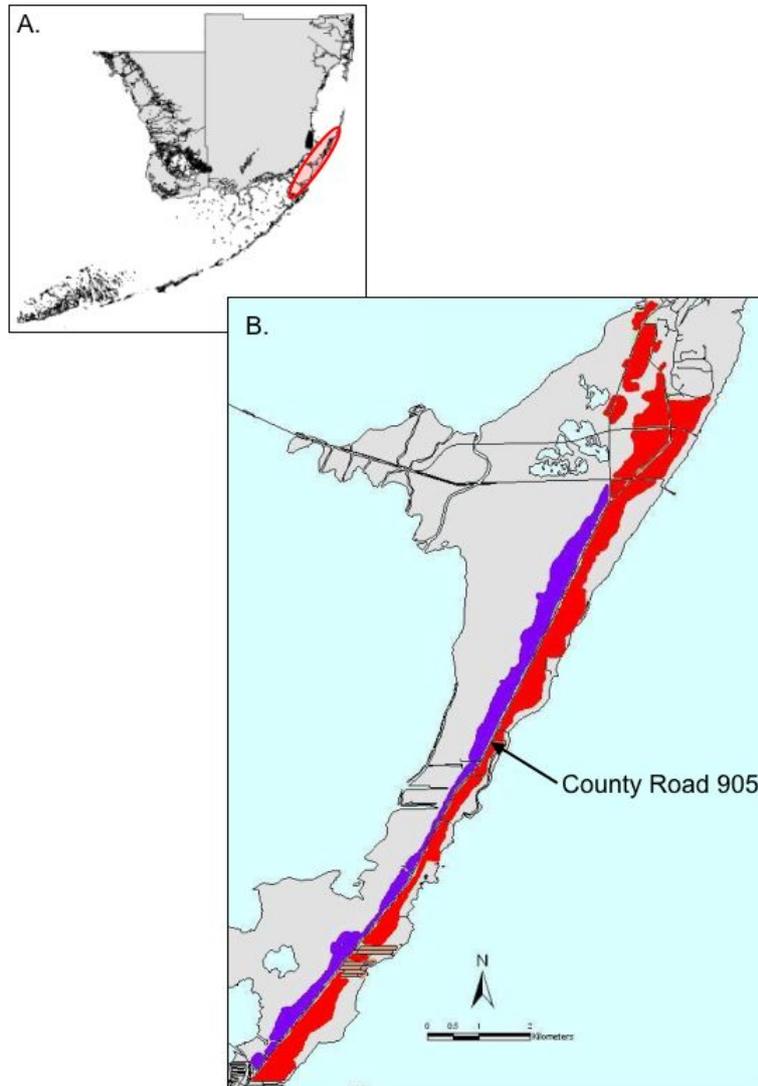


Figure 1.2: A) Map of the south east United States of America, highlighting the northern third of Key Largo (shaded red). B) The Key Largo woodrat occupies extant tropical hardwood hammock protected by two reserves on the northern third of Key Largo. On the NW side of County Road 905 is the Crocodile Lake National Wildlife Refuge (shaded purple, managed by the Federal Fish and Wildlife Service). On the SE side of County Road 905 is the Dagny Johnson Key Largo Hammock Botanical State Park (shaded red, managed by the Florida Department of Environment and Protection). The hammock is surrounded by mangrove (shaded grey) and water (shaded blue). Roads and tracks are shown as a solid black line.

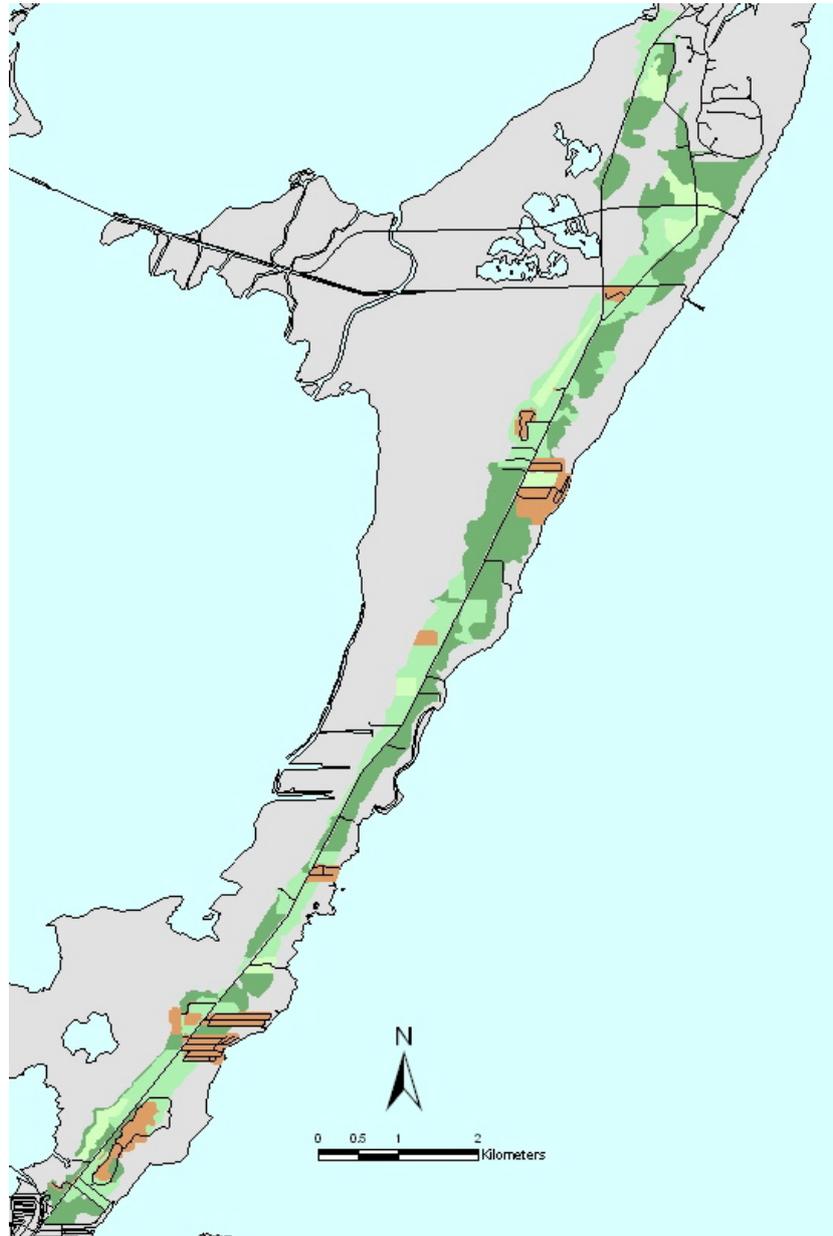


Figure 1.3: Map of available habitat in northern Key Largo. Four types of habitat have been identified: young, medium and old tropical hardwood hammock (shaded light to dark green, respectively) and disturbed/urban habitat (shaded brown). Unsuitable habitat is shaded blue (water) and grey (mangrove swamp). Roads and tracks are shown as a solid black line.

### 1.2.3 Species management

In 1984, the Key Largo woodrat was listed as a threatened subspecies under the Endangered Species Conservation Act of 1969, and is subject to a recovery plan that identifies potential threats and describes management actions to be undertaken (USFWS, 1999). Identified threats include habitat loss and fragmentation (USFWS, 1999); competition with black rats (*Rattus rattus*, Goodyear, 1992; Hersh, 1981; Muiznieks, 2006); dumping of new rubbish (Muiznieks, 2006; USFWS, 1999); disease (raccoon round worm, *Baylisascaris procyonis*, LoGiudice, 2001, 2003; McCleery, 2003; McCleery et al., 2005); and predation from numerous invasive species including fire ants (*Solenopsis invicta*, Forsy et al., 2001), raccoons (Humphrey, 1988), feral cats (USFWS, 1999; Hatley, 2003), and Burmese pythons (*Python molurus bivittatus*, Reed, 2005; Greene et al., 2007). Management actions have included public land acquisition on north Key Largo, through the establishment of two wildlife reserves that have restricted public access (Figure 1.2, USFWS, 1999); relocation (Brown and Williams, 1971; Barbour and Humphrey, 1982; USFWS, 1999; Gore and Loggins, 2005), captive breeding and reintroduction for captive bred individuals (Muiznieks, 2006; Alligood et al., 2008, 2009), and predator control (Muiznieks, 2006) programs; and enhancing habitat with supplemental nest structures (Winchester et al., 2009).

Despite these management efforts, intermittent monitoring over the past 25 years has suggested the woodrat population has suffered an extensive range contraction (occupying less than one-half of its original distribution) and decline in population size (Hersh, 1981; Barbour and Humphrey, 1982; Humphrey, 1988; Gore and Loggins, 2005; Frank et al., 1997; USFWS, 1999; McCleery, 2003; Muiznieks, 2006; Winchester, 2007). A variety of approaches based on direct and indirect observations have been employed with different levels of effort to estimate woodrat abundance and density. The most common method has been

a capture-recapture survey using a grid-based approach (Table 1.1). This survey design is too expensive to implement as part of a long-term monitoring program and has found to be inefficient due to low recapture rates (Winchester, 2007). Methods based on indirect observation, e.g., track detection, have also found to be unsatisfactory due to issues with track identification (Gore and Loggins, 2005). Standard monitoring methods typically have not worked well for estimating the abundance of the woodrat, because it is cryptic, nocturnal, and sparsely distributed.

A new approach to developing a long term monitoring program of woodrat abundance is required, especially to assess the impact of current management actions on the population size of the species. The program needs to consider both survey design and analytical methods, to ensure that estimation of the abundance and associated uncertainty of the woodrat is robust. The program should be inexpensive to maintain over a long period of time and must yield sufficient data from which management decisions can be based. It is also important for the method to be practical, that is, it must be readily implementable by field staff working in difficult terrain.

### **1.3 Thesis outline**

In this introductory chapter, I have explained the importance of obtaining accurate animal abundance estimates and highlighted the problems involved in undertaking such a task. I also introduced a case study species, the Key Largo woodrat, for which a new monitoring program is required. Commonly available methods of estimating abundance are too costly to form part of a long-term monitoring program for this species, as the woodrat is sparsely distributed, cryptic and nocturnal. A new approach is therefore required.

Table 1.1: Summary of results for previous studies of the Key Largo woodrat. Table summarized from Gore and Loggins (2005). A dash (-) indicates information was not reported by the author. Captures is given as number of capture events per 100 trap nights.

Reference	Year study commenced	Duration (yrs)	Area sampled	Number of trap nights	Captures	Animals (per ha)
Brown (1973)	-	1	-	-	-	1.2
Hersh (1981)	1976	2	1 5.6ha grid	1,200	7.5	2.2
Barbour and Humphrey (1982)	1979	1	2 1.6 ha grids	1,696	-	2.8
Goodyear (1985)	1985	2	44 transects	1,848	4.0	-
Humphrey (1988)	1986	2	6 2.6-3.4ha grids	8,956	3.3	7.6
Frank et al. (1997)	1995	2	41 transects and 4 3.8ha grids	6,300 13,824	0.7 0.6	- 0.9
Sasso (1999)	1996	2	4 1.8ha grids	8,400	-	1.8
Keith and Frank (1998)	1997	3	6 transects at 1 site	1,320	0.6	-
US FWS (unpubl. data) (2000)	1997	3	25 transects	4,500	0.4	-
US FWS (unpubl. data) (2002)	2000	2	4 1.8ha grids and 10 transects	10,400	0.4	0.6
McCleery (2003)	2002	1	60 1ha grids	12,316	0.11	0.22
Winchester (2007)	2005	1	40 0.56ha grids	5,256	-	0.22

In the following chapter, I review currently available methods for estimating small mammal abundance, and assess these methods for estimating Key Largo woodrat abundance according to statistical properties, practicality and cost. From this review, I decide three methods warrant further investigation and field testing to determine their potential in forming part of a long-term monitoring program of Key Largo woodrat abundance.

The first of these selected methods that I consider (Chapter 3) is a standard capture-recapture analysis (Pollock et al., 1990). This is currently the most commonly implemented method for estimating small mammal abundance, and has previously been used to estimate Key Largo woodrat abundance (e.g., McCleery, 2003; Winchester, 2007). However, estimating abundance using capture-recapture is problematic due to the edge effect: animals with home range centres off the grid can be trapped on the grid due to movement. Consequently, the effective trapping area of the grid is larger than the grid itself, but its size is difficult to estimate. I demonstrate the use of four different methods to estimate the size of the effective trapping area of the grid, and found the resulting estimates of population abundance to be highly sensitive to which method is used.

In Chapter 4, I use an extension of capture-recapture analysis, spatially-explicit capture-recapture (SECR, Borchers and Efford, 2008). The overall aim of the analysis is to estimate the density of animal home range centres,  $\hat{D}$ , existing in the survey area using spatial information contained in the pattern of recapture events of individual animals, and also prior knowledge of where traps are placed. I use maximum likelihood estimation to estimate the parameters of a ‘capture probability’ model that models the observation process of detecting an animal at a trap, given the trap is placed some distance from an animal’s (unknown) home range centre. Using the capture probability model, estimation of  $\hat{D}$  follows using a Horvitz-Thompson-like estimator (Borchers and Efford,

2008).

In Chapter 5, I present an alternative abundance estimation technique, trapping point transects (TPT, Buckland et al., 2006). The TPT method combines elements of both capture-recapture and distance sampling (Buckland et al., 2001), whereby the “detection function”,  $g(r)$  (i.e., the probability of capturing an animal, given it is  $r$  m from a trap when the trap is set), is estimated using a representative subset of animals whose locations are known prior to the traps being set. I modify standard point transect methods to estimate animal abundance. In Chapter 6, I undertake a simulation study to assess bias in abundance estimates when the assumptions of the TPT method are violated, and when different sampling strategies are used to estimate the detection function.

In 2008 and 2009, a field study was undertaken during which the TPT method was used to estimate the abundance of the Key Largo woodrat. Results of this field study are presented in Chapter 7, along with a simulation study that investigated how survey design might be changed and additional survey effort allocated to increase precision of abundance estimates.

I conclude this thesis with a general discussion of how monitoring of the Key Largo woodrat abundance should continue in the future and what the implications are for management actions (Chapter 8). I also discuss future research directions, especially with regard to extending the trapping point transect method, and how these extensions relate to spatially explicit capture-recapture (Borchers and Efford, 2008).

## Chapter 2

# A review of methods for estimating small mammal abundance, and their applicability to the Key Largo woodrat

### 2.1 Introduction

Numerous reviews exist on survey methods to estimate abundance of animals (Seber, 1986, 1992; Schwarz and Seber, 1999; Buckland et al., 2000), some are restricted to certain taxa (e.g., carnivores, Wilson and Delahay 2001; and land birds, Rosenstock et al. 2002) and others to survey techniques (e.g., genetic sampling, Kohn and Wayne 1997). It seems few attempts have been made for reviewing methods applicable to small mammals, especially when they are rare, e.g., the Key Largo woodrat, as certain survey methods are no longer applicable. In this chapter, I review potential survey methods for small mammals. I do not consider the selection of sampling units, from which the parameter of interest is observed. I assume enough sampling units have been selected according to a randomized scheme, from which extrapolation to the broader survey area is reliable.

Broadly speaking, survey methods can be separated into two categories. Census based approaches assume perfect detection, i.e., all animals are seen with certainty, and result in a complete count of all individuals in the sampling units. A census may be efficient for sessile objects (e.g., plants) or when densities are high (e.g., photo counts of seal colonies). For example, Dique et al. (2004) did a strip transect search for the koala (*Phascolarctos cinereus*) in various habitat types in eastern Australia and by assuming they detected all koalas present, they estimated density as the number of individuals detected per hectare of habitat searched. However, since the detection of rare, small mammals is likely to be imperfect, I consider census based approaches to be inappropriate. Instead, I concentrate on survey approaches.

A *survey* is a partial or incomplete count of the population within a defined sampling unit, where not all individuals within the sampling unit are counted (Thompson et al., 1998; Wilson and Delahay, 2001). These surveys can be based on indirect (e.g., field sign) or on direct observation of individuals. I conclude this chapter with a discussion on how feasible these methods may be in forming a long-term monitoring program of Key Largo woodrat abundance.

## 2.2 Distance sampling

Distance sampling is a commonly implemented method used to provide reliable estimates of abundance and density (Buckland et al., 2001). In distance sampling the observer measures the distance from a line or point to each detected individual animal, group of animals, or cue (e.g., nests or dung, see Section 2.7). The distribution of observed distances is used to estimate a ‘detection function’ (i.e., the probability of detecting an individual as a function of distance from a line or point). Given not all individuals are detected in the area surveyed, the detection function is used to estimate the proportion of individuals detected

( $\hat{P}_a$ ). Density is then estimated as:

$$\hat{D} = \frac{n}{a \cdot \hat{P}_a} \quad (2.1)$$

where  $n$  is the number of individuals detected and  $a$  is the area surveyed (Buckland et al., 2001). The development of the free software Distance (<http://www.ruwpa.st-and.ac.uk/distance/>) and two comprehensive books written by Buckland et al. (2001, 2004) have greatly increased the ease with which distance sampling studies can be planned, conducted and analysed.

### 2.2.1 Line transect sampling

In line transect sampling, the observer records the perpendicular distance from the line to each animal detected (Buckland et al., 2001). In the standard approach ('conventional distance sampling'), all animals on or near the line should be detected, but a proportion of animals within distance  $w$  of the line may be missed. The perpendicular distance,  $x_i$  (where  $i$  indicates the  $i$ th animal,  $i = 1, \dots, n$ ), from the line may be measured directly, or calculated from the radial distance  $r_i$  and sighting angle  $\theta_i$ , i.e.,  $x_i = r_i \sin(\theta_i)$ .

Line transects have many advantages. Firstly, the method does not require all animals to be detected within the covered strips, a common feature of survey data. Secondly, efficiency is increased as a wider strip can be searched when it is not necessary to detect all animals in the strip. Thirdly, distance sampling methods also have an additional property of pooling robustness not possessed by capture-recapture methods (see section 2.3). Models of detectability are pooling robust if the data can be pooled over many factors that affect detection probability (e.g., vegetation cover) and still yield a reliable estimate of density (Buckland et al., 2004). This is a very useful feature of the method because not all factors

affecting detection probability (e.g., environmental variables) are known and can be measured. In distance sampling, because of pooling robustness, this need not cause a significant bias in the estimate of abundance.

Conventional line transect methods require some important assumptions (all of which can be relaxed in more advanced variants of the methods, Buckland et al. 2004):

1. Transect lines are located randomly with respect to the distribution of animals. Given random line placement, one can safely assume that animals are uniformly distributed with respect to perpendicular distance from the line (or distributed according to a triangular distribution with distance from a point in point transect sampling). In addition, random line placement ensures lines are representative of habitat conditions throughout the survey region, such that results can be extrapolated to the whole study area, and not just the area surveyed;
2. All animals on the transect line are detected with certainty. Any violation of this assumption translates directly into a negative bias in the density estimate. For example, if detection probability on the transect line is 0.9, the density estimate will show a 10% negative bias. Design of surveys must fully consider ways to ensure that this assumption is met, or to allow  $g(0)$  to be estimated;
3. Animals are detected at their initial location, i.e., a ‘snap shot’ is taken at time of survey. Movement in response to the observer, either an attraction toward or away from the line, should not occur. Generally, animal movement independent of the observer is problematic, unless the mean speed of movement of the animal is slow relative to the speed of the observer; and
4. Distance measurements are exact. When distances and/or angles have

been rounded (especially to zero), grouping data into intervals can help, assuming that on average, the estimates are accurate (i.e., errors are unbiased and not too large). A histogram of data can reveal heaping and outliers.

Two other less critical assumptions are that animals are identified correctly and observations are independent (i.e., detecting one individual does not influence whether another will be detected).

### 2.2.2 Point transect sampling

Instead of walking a line transect, an observer may remain stationary at a point for a fixed period of time, and measure the sighting (radial) distance from the point to each of the animals detected. This may be useful when, for example, there is rough terrain that makes it difficult for the observer to walk along the transect safely while also concentrating on detecting animals.

Such a procedure can be advantageous over line transect sampling, as the observer has time to observe animals close to the point, and different species can be studied simultaneously (Seber, 1986). However, measurement errors generate substantially more bias in density estimates than do errors of similar magnitude in line transect sampling (Buckland et al., 2004). In addition, any movement of animals during the count period can also generate substantial bias in density estimates, even if the movement is independent of the observer (Buckland, 2006).

The success of applying line or point transect sampling to small mammals has been variable. Hounscome et al. (2005) walked line transects and spotlighted for badgers (*Meles meles*) in south east England and found similar results to capture-recapture estimates, with less survey effort. Healy and Welsh (1992) also successfully used line transects to estimate density of the grey squirrel (*Sciurus*

*carolinensis*) in Massachusetts.

However, when detectability of animals is low, the assumption that detectability on the line or point is perfect is often violated. In addition, with low detectability, a large amount of survey effort is required to obtain enough detections to reliably model the detection function. In these instances, distance sampling tends to perform badly. For example, Morrison and Kennedy (1989) walked line transects to estimate density of chipmunks (*Tamias* sp.) in New Mexico concluding that  $g(0) < 1$  and that by walking the transects, individuals were being flushed away from the transect line. Also, Gitzen et al. (2001) drove line transects at night and used spotlights to detect jackrabbits (*Lepus californicus*) in Washington State. However, the detectability of jackrabbits using this approach was so low, too few individuals were detected to estimate density using distance sampling theory. In these instances, detectability of individuals typically relies on trapping and data are analysed within a capture-recapture framework.

### 2.3 Capture-recapture

Capture-recapture is probably the most common monitoring method applied in small mammal surveys (Pollock et al., 1990; Solari et al., 2002). The simplest capture-recapture study involves just two sampling periods. In the first sampling period,  $M$  animals are captured from the target population of unknown size  $N$ , marked and released back into the population. In the second sampling period, a sample of size  $n$  is taken, containing  $m$  previously marked individuals. Inference can be made about the unknown population size,  $N$ , using the Lincoln-Petersen estimator:  $\hat{N} = nM/m$  (Thompson et al., 1998; Lukacs, 2008). This approach requires batch marks only (i.e., individuals need to be marked, but not uniquely).

Three key assumptions underlie the Lincoln-Petersen estimator (Bailey et al., 2004): 1) the population is closed (i.e., no births, deaths, immigration or emigration); 2) marks are not lost or misread by the observer; and 3) all individuals have the same non-zero probability of capture within each sample. Good field methods can assist in meeting two of these assumptions, e.g., double-marking animals with an ear and passive integrator transponder (PIT) tag can reduce chances of tag loss, and ensuring sampling occasions are close together in time can increase the likelihood that the closed population assumption will be met. However, the equal capturability assumption is seldom met in practice (Williams et al., 2002).

Research into developing capture-recapture models over the past few decades that permit various modifications to these assumptions has been substantial, especially when individual ‘capture histories’ have been recorded. That is, a string of 0s and 1s indicating for each sampling period whether the individual was captured (1) or not (0). For example, a capture history of 1011 indicates that the individual was captured on the first, third and fourth sampling periods but not on the second sampling period, either because the individual was present and not detected, or the individual was not present.

Closed population models can now relax the assumption of equal capture probability by allowing capture probability to vary as a result of time ( $M_t$ , i.e., probability changes between trap periods, but is constant within a trap period), behavior (or trap) response ( $M_b$ , i.e., all animals initially have the same probability of capture, but after the first capture event, probability of recapture may increase, trap happy, or decrease, trap shy), and individual heterogeneity ( $M_h$ , initial capture probability is fixed according to an individual characteristic, e.g., sex or age), or any combination thereof (Otis et al. 1978; Rexstad and Burnham 1991; Chao et al. 1992; Menkens and Anderson 1988; Bailey et al.

2004).

In studies where demographic closure is not met, open population models can estimate population size  $N$  at each sampling period, and survival probability between sampling periods (Pollock et al., 1990; Bailey et al., 2004). Inference can also be made about recruitment (e.g., McDonough and Rexstad, 2005, combined survival rate estimates with population size estimates of red-backed voles, *Clethrionomys rutilus*, in Alaska to obtain estimates of reproductive recruitment and immigration). The most common open population model is probably the Jolly-Seber model (Jolly, 1965; Seber, 1965) and has three assumptions (Pollock et al., 1990): 1) all animals in the population at time  $i$  are equally likely to be captured; 2) every marked animal present in the population at time  $i$  has the same probability of survival between  $i$  and  $i + 1$ ; and 3) marks are not lost or misread by the observer. An implicit assumption of the Jolly-Seber model is that all emigration from the population is permanent (i.e., no temporary emigration is permitted, Lee and Tietje 2005). The Jolly-Seber model has some parameter estimation limitations, e.g., population size for the first and last sampling periods cannot be estimated, and the population size estimator is not robust to variation in capture probabilities (other than time variation, Kendall and Pollock, 1992; Bailey et al., 2004).

The robust design model (Pollock, 1982) combines a closed population model with the open population, Jolly-Seber model (Nichols et al., 1984; Kendall et al., 1995, 1997). Under this design, there are multiple primary sampling periods, separated by a longer period of time during which the population is open to gains (births and immigration) and losses (deaths and emigration). Data within each primary period are pooled to estimate survival probability ( $\phi$ ) between primary sampling periods. Within each primary sampling period, there are several secondary sampling periods separated by a short period of time during which

the population is assumed closed. Data from each secondary sampling period is used to estimate capture ( $p$ ) and recapture ( $c$ ) probabilities, and population size ( $N$ ) for each primary sampling period. Data from both primary and secondary sampling periods are used to estimate the probability of emigration ( $\gamma'$ ) and the probability of remaining away from the study area, given the individual had previously emigrated ( $\gamma''$ , Kendall et al. 1995, 1997; Bailey et al. 2004).

The robust survey design is particularly useful for estimating temporary emigration probabilities. Temporary emigration off and on to the trapping grid is primarily associated with animals whose home ranges were only partially on the grid (Lee and Tietje, 2005). Kendall et al. (1997) recommend fixing various migration parameters to ensure these parameters remain identifiable (e.g., no migration, Markovian migration and random migration models).

Robust design capture-recapture models have been used with success to estimate small mammal abundance. Lee and Tietje (2005) used Pollock's robust design to investigate how survival and temporary emigration of the dusky-footed woodrat changed in response to prescribed burning in California. They conducted nearly 41,000 trap nights at 10 trapping grids in a nine year period. Through model selection based on Akaike's Information Criterion, AIC (Burnham and Anderson, 2003), they concluded the most parsimonious model had survival probability as a function of age and date, emigration was temporary and a function of age and season, and capture and recapture probabilities were equal and constant within primary trapping periods but varied between trapping periods. Hadley and Wilson (2004) conducted 16,800 trapping nights across four trapping grids in a four-year period, also using Pollock's robust design to determine the effect of a ski-run development on a variety of small mammals in Colorado.

Despite the large body of research into capture-recapture models, some issues

still remain. Most importantly, capture-recapture approaches are particularly prone to problems with heterogeneity in capture probability caused by unequal access to traps (Otis et al., 1978; Menkens and Anderson, 1988; Eberhardt, 1990; Royle et al., 2009). If ignored, substantial over-confidence and negative bias in estimates of population size may result (Anderson et al., 1983; Seber, 1992; Link, 2003).

Also, capture-recapture methods can only provide an estimate of animal density by converting estimated abundance based on the effective trapping area (i.e., density = number caught/trapping area). Estimation of the effective trapping area is problematic due to the edge effect (Corn and Conroy, 1998; Efford et al., 2004), and is caused by two reasons. Firstly, animal home ranges may lie on the boundary of the trapping grid and therefore the animal only spends a portion of its time inside the grid. Secondly, animals with entire home ranges outside of the grid may still be detected on the grid because they are attracted to the grid via baiting (Otis et al., 1978; Parmenter et al., 1989). This results in an unusually large number of captures in the outermost traps compared to the innermost traps of the grid (or webs, Section 2.5). The smaller the grid relative to the mean home range size of the study species, the greater the edge effect (Mendel and Vieira, 2003). The simplest solution is to ignore the edge effects, and assume the effective trapping area of the grid is simply the area of the grid,  $A_G$ . This naïve density estimator (i.e.,  $D = N/A_G$ ) is known to be upwardly biased but still is commonly implemented (e.g., McCleery, 2003).

In the two studies discussed above, Lee and Tietje (2005) used the minimum number of animals captured to represent abundance, and Hadley and Wilson (2004) adjusted the effective trapping area using the mean maximum distance moved (MMDM, Wilson and Anderson, 1985a). Both of these solutions are

ad hoc and also bias-prone. Firstly, using the actual number of animals captured may not be representative of the overall abundance as the proportion of uncaptured animals remains unknown and may not be constant through time. Estimating the MMDM is also problematic, as the estimate is generally based on trapping data and is consequently a function of trap spacing, animal home range size, capture probability (Efford et al., 2004), and also density. For example, Parmenter et al. (2003) found correcting the effective trapping area by the MMDM improved confidence interval coverage, root mean square error, accuracy and bias in estimating density, however caution against its general use in small-mammal trapping studies due to theoretical reasons. Instead of relying on trapping data to estimate MMDM, radio telemetry data has been used but found to depend on the number of fixes used per individual and also found to vary with season (Tioli et al., 2009).

Other solutions to estimate the effective trapping area exist, but are not inerrant and none have been implemented widely (Efford, 2004; Efford et al., 2009b). For example, White and Shenk (2001) proposed estimating the average proportion of time trappable animals spend on the grid,  $\bar{P}_A$ , and relate the naïve density estimate by this factor. A similar approach was also proposed by Eberhardt (1990). However, precise estimation of  $\bar{P}_A$  is expensive, and a large number of representative samples is required, usually via radio tracking. Also, this method is biased towards individuals with a higher probability of being caught (Efford, 2004). Few studies are available that compare different methods of estimating the effective trapping area (e.g., Tanaka, 1972, 1980; Mendel and Vieira, 2003; Soisalo and Cavalcanti, 2006; Tioli et al., 2009), and even fewer against a known, true population density (e.g., Parmenter et al., 2003). However, all conclude that estimates of the effective trapping area greatly influences estimates of abundance.

One particular advantage of capture-recapture approaches is that if individual capture histories are recorded, estimates of fecundity and survival can be made. Also, instead of trapping and handling individual animals, camera stations have been used extensively in capture-recapture surveys of species with unique pelage patterns (e.g., Karanth and Nichols 2002; Royle et al. 2009). Camera stations have several advantages over other monitoring methods, such as reducing human disturbance as the cameras are relatively non-intrusive and the animals do not need to be captured, and cameras can be operational for extended periods of time, and hence fieldwork requirements can be reduced (Gore and Loggins, 2005). However, the initial cost for purchasing a sufficient number of cameras is great, and they can malfunction. I am unaware of any studies that have successfully used remote camera stations to uniquely identify small mammals for analysis within a capture-recapture framework. Small mammals can be uniquely identified using hair trapping and DNA analysis, and this data type has been analysed with a capture-recapture framework (Lukacs and Burnham, 2005; Kéry et al., 2011).

## 2.4 Spatially-explicit capture-recapture

Capture-recapture studies with fixed trap locations have a spatial component: animals close to traps are more likely to be caught than animals further away. This is not addressed in standard capture-recapture analyses and without this spatial component, rigorous estimates of density cannot be obtained (Conn et al., 2006; Borchers and Efford, 2008; Borchers, 2011). The incorporation of spatial information into capture-recapture analyses was first proposed by Efford (2004) and Efford et al. (2004). This approach avoids estimation of the effective trapping area, instead using the known location of the traps and the pattern of

recapture events to estimate the density of home range centers and a capture function. The capture function is conceptually consistent with distance sampling approaches (c.f. section 2.2), in that the probability of detecting an individual is assumed to be a radially decreasing function of the distance between the center of the animal's home range and the distance to the trap. More formally, the capture probability model is the probability that an animal located at  $\mathbf{X}$  (a vector specifying the coordinates of the home range center) is detected by trap  $k$  on occasion  $s$ . Three forms of the capture probability model are available (Table 2.1), that typically require two parameters:  $g_0$ , the probability of being trapped if the animal's home range is centered at a trap (i.e.,  $d(\mathbf{X}) = 0$ ); and  $\sigma$ , the spatial scale (the hazard-rate function also has an additional parameter,  $b$ , Table 2.1).

Table 2.1: Three forms of the capture probability model used in spatially-explicit capture-recapture.  $d$  is the distance between an animals home range centre and a detector. The parameter  $g_0$  is common to all functions and represents the probability of detection at a single detector placed in the centre of the home range; values of the spatial scale parameter  $\sigma$  are not comparable between functions (Efford et al., 2009b).

Detection function	Equation	Parameters in vector $\theta$
Halfnormal	$g_0 \exp(\frac{-d^2}{2\sigma^2})$	$g_0, \sigma$
Exponential	$g_0 \exp(\frac{-d}{\sigma})$	$g_0, \sigma$
Hazard rate	$g_0 [1 - \exp\{-\frac{d}{\sigma}\}^{-b}]$	$g_0, \sigma, b$

In early implementations of this approach, simulation and inverse prediction were used to jointly estimate density,  $\hat{D}$ , and the parameters of the capture function (Pledger and Efford, 1998). For example, Efford et al. (2005) found density estimates of brushtail possums to be similar to those based on removal methods, and Efford (2004) conclude density estimates were virtually unbiased and relatively precise when applied to simulated data. Despite some success, using inverse prediction was limited with respect to model selection and the

inclusion of covariates (Borchers and Efford, 2008). Recent developments now enable spatially-explicit capture-recapture models to be fitted within both maximum likelihood (Borchers and Efford, 2008; Efford et al., 2009b) and Bayesian frameworks (Royle and Young, 2008; Royle et al., 2009).

Requisite data for a SECR analysis are individual capture histories, similar to standard capture-recapture analyses but instead of recording a string of 1s and 0s if the individual was detected or not, respectively, when the individual is captured, the location of capture is recorded. That is, a capture history of “0,10,14,0” represents an individual not captured during the first or last sampling occasion, and was captured at trap location 10 and 14 on the second and third sampling occasions, respectively. Since the survey design and therefore trap locations are known, the distance between recapture events (i.e., between trap locations 10 and 14) are known. Given the capture history  $\omega_i$  of individual  $i$  and an estimate of the capture function, the probability density of the location of the home range center of this individual can be estimated (Borchers and Efford, 2008).

The maximum-likelihood estimation method of this approach is readily implementable using the ‘secr’ package in program R (Efford, 2010), which allows model selection methods to be based on Akaike’s Information Criterion, AIC (Burnham and Anderson, 2003). The inclusion of covariates in the capture-probability function, such as trapping occasion, whether an individual has been trapped before and additional heterogeneity not attributable to distance of trap (e.g., individual covariates such as sex, and unobservable covariates accounted for using a two-point mixture model of Pledger (2000)), is straight-forward.

The maximum-likelihood based SECR approach assumes the population is closed (i.e., no births, deaths, immigration or emigration), although recent developments within the Bayesian framework can permit open-population models

(Gardner et al., 2010). As per standard capture-recapture, SECR assumes that tags are not lost and the identity and location of capture of each individual is recorded accurately. Other assumptions of the method are:

- Animals occupy circular home ranges,
- Home range centers are distributed in space according to a Poisson point-process,
- Detectors are operated at known locations for a fixed period of time,
- Animals are detected independently of each other, and
- Detector placement is random with respect to location of animal home range centers.

The method is remarkably robust to violation of the first two of these assumptions (non-circular home ranges and a clumped distribution of individuals), causing increased estimates of variation, rather than significant bias in density estimates (Efford, 2004; Efford et al., 2005). The other assumptions can be assured by good survey design (e.g., by placing traps randomly in the survey area). One particular strength in this approach is its flexibility in terms of accommodating any spatial arrangement (regular or irregular spatial arrangement of lines, webs and random placement, or otherwise) of traps (Efford, 2004), and traps can include physical capture of an individual, or a proximity detector such as a camera trap, or passive detector array (e.g., microphones, Efford et al., 2009a or hydrophones, Marques et al., 2011).

## 2.5 Trapping webs

Trapping webs combine capture-recapture studies with distance sampling theory (Anderson et al., 1983). A trapping web consists of lines of regularly spaced traps radiating from a central, randomly chosen point. Here, ‘detection’ by an observer is replaced by animals being caught in traps at a known distance from the center of the trapping web. Animals at the center of the web should be captured with probability one. Trapping continues for  $t$  occasions, and data from either the initial capture of each animal or all captures and recaptures are analyzed (however the latter approach remains contentious).

Lukacs et al. (2005) suggested that 60-80 captures may be sufficient to provide reasonable estimates and recommended a trapping web design consisting of greater than 90 traps, with 5-7 trapping occasions. To estimate density over a wider area, several randomly located webs are required (ideally greater than 15, Lukacs et al., 2005). The web can be set up with varying trap spacing, number of traps and trapping occasions. The simulation program WebSim (Lukacs, 2002) offers a framework to explore trapping web design for a particular scenario. Trapping webs can become very labor intensive and expensive; however, they have been implemented successfully in the field on small sized mammals (e.g., mongoose, *Herpestes javanicus*, in West Indies, Corn and Conroy 1998).

Trapping webs are considered a special case of point transects (Lukacs et al., 2005), and the same assumptions apply. Namely, all individuals at the center of the web were captured by the end of the trapping period; there is no movement of animals with respect to the location of the trapping web, and distances were measured accurately. Given that trapping can be conducted across multiple occasions, information on recaptures can be used in the estimation of the detection probability (Lukacs et al., 2005).

Edge effects can still be an important issue in trapping webs, because animals

from outside the web may be attracted to the trap baits in the outer rings of the web. It may often be necessary to delete captures from the outermost one or two rings of the web through truncation, which can make the method inefficient (Anderson et al., 1983; Parmenter et al., 1989; but see Link and Barker, 1994).

Trapping line transects is another approach, similar to trapping webs, but traps are placed each side of a transect line with a declining density (Lukacs et al., 2005). The distance from the transect line to the trap containing an animal is the detection distance. This method is considered to be a special case of line transect sampling, and the same assumptions apply. The length of the transect, maximum distance from the transect line at which traps are placed, and the distribution at which trap density declines away from the transect line need to be considered (Lukacs et al., 2005).

Density estimates from trapping webs are very prone to upward bias as animals move toward the center of the web (Parmenter et al., 2003; Efford et al., 2005). This upward bias may be counteracted by the downward bias caused because detection at the center of the web,  $g(0)$ , is less than one. The size of the web should be chosen so that all animals near the web center are caught, yet the method requires that home ranges vary little in size and are small relative to the size of the web, to avoid substantial bias due to movement. Without separate studies to estimate movement and  $g(0)$ , the resulting bias in final estimates caused by these two factors is not known. Wilson and Anderson (1985b) concluded via a simulation study that density estimates were negatively biased with low capture probability ( $\bar{p} = 0.09$ ) because detection at the center of the web was less than one. However, when  $\bar{p}$  increased, this bias decreased as the chance of capturing all animals at the center of the web increased. Also, Parmenter et al. (1989) found little or no detectable bias in estimates of density of ground dwelling darkling beetles (*Eleodes* sp.), using artificially stocked, enclosed areas

where true population size was known.

## 2.6 Trapping point transects

Trapping point transects are another extension of distance sampling, where, similar to trapping webs (Section 2.5), ‘detection’ of an animal by an observer is replaced by catching the animal in a ‘trap’ (whereby ‘trap’ can be a physical trap capturing the animal, or an observer sighting the animal). The concept behind trapping point transects is if the location of a representative subset of animals is known prior to the trap being set, these can be used to estimate the ‘detection function’. Standard point transect methods can be modified to estimate abundance (Buckland et al., 2006).

In this approach, two surveys are conducted. In the *main survey*, trapping locations are randomly selected such that they are independent of each other (i.e., distance between traps is great enough to ensure capturing an animal at one trapping location does not influence whether a neighbouring trap will capture an animal). Traps are established at each trap location and checked during the trapping period for caught animals. It is not necessary to individually mark animals in the main survey, but recording covariate information (e.g., sex) of captured animals may be useful at the analysis stage. This main survey provides an encounter rate, that is, how many capture events occur per trap night. It is difficult, if not impossible, to convert this encounter rate into an estimate of density, as the effective survey area sampled by a trap is unknown. Consequently, a separate survey (the *trial survey*) is required to estimate a detection function, and using this detection function, the encounter rate can be corrected for the number of animals missed.

To conduct an individual trial, a known-location animal is required (e.g., via

searching). A trial ‘trap’ (e.g., a physical trap, or an observer using a lure to sight the individual) is established at a predetermined distance away from where the animal is located. After a set period of time, the trial terminates, after which, it is discovered whether the known-location animal has been detected at that trial distance. This is repeated for many distances (e.g., 10m intervals between 10 and 100m). Thus, a binomial response variable is recorded, informing whether each known-location individual was detected in a trap, together with its distance from the trap when it was set. The detection function can be fitted using, e.g., a logistic regression.

The design of the trial survey should ensure that there is a range of distances from the known-location animals to the nearest trap, to make the logistic regression more reliable. This is very important. Because distances in the main survey cannot be truncated (as it remains unsure how far unmarked animals have traveled to be trapped), some of the traps in the trial survey must be set at distances where capture is unlikely. This ensures the tail of the logistic function is fitted well. In addition, spatial configuration of traps used for capturing animals to be marked with radio collars should be independent of that for traps used in the main survey.

This method does not require that detection at distance zero from the trap,  $g(0)$ , be one. However, since  $g(0)$  will usually be less than one, the pooling robustness property of distance sampling no longer applies. Therefore, information on all important sources of heterogeneity (e.g., sex) in the detection probability, and using a survey design that minimizes unmodelled heterogeneity (e.g., conducting the main and trial surveys at the same time and location), is very important.

This method was first proposed by Buckland et al. (2006) and was applied to the Scottish crossbill (*Loxia scotia*). They used a point transect lure survey,

whereby flocks of crossbills were located by one observer, and another observer some predetermined distance away from the known-location flock initiated the lure (a tape of excitement calls). Whether the known-location flock responded to the lure was one trial, of which many trials were conducted on many flocks of crossbills. Thus, enough data were collected to which a detection function could be fitted.

A similar idea has been applied to lions (*Panthera leo*) and hyaenas (*Crocuta crocuta*) in Kenya (Ogutu and Dublin, 1998; Ogutu et al., 2005; Kiffner et al., 2007), however these authors did not go so far as to estimate a detection function dependent on distance. Ogutu and Dublin (1998) and Ogutu et al. (2005) broad-casted animal vocalisations to determine the proportion of individuals within a fixed distance that responded to the lure. This proportion was used as a multiplier to correct the total number of animals observed for those that were present but not detected when the vocalisation was played. Kiffner et al. (2007) investigated how different factors such as moon phase, time of play back, direction of play back, sex and age of individual influenced whether the individual responded to the lure, but did not measure distance at which the individual responded. Mills et al. (2001) estimated the abundance of the spotted hyaena in Kruger National Park, South Africa, by first using a play back vocalisation to determine the number of individuals that responded. Similar to Buckland et al. (2006), an independent experiment ('trial' survey) was conducted whereby one observer stayed with an individual hyaena and the other observer some distance away played a vocalisation lure to determine if the known-location individual responded to the call, but they assumed response probability within a fixed radius of the calling station was constant.

To my knowledge, the trapping point transect method has not been implemented in the field to monitor small mammal abundance via trapping surveys.

In such instances, known-location animals may be acquired via radio tracking. Alternatively, it may be possible to use the point-of-release of a trapped animal to obtain the ‘known-location’, instead of radio tracking animals. This may not be appropriate if there is an effect of trap experience (i.e., animals released at point of capture may avoid a particular area and this will affect their capture probability).

## 2.7 Indirect observations

Animal signs (“cue”, e.g., nest, dung or tracks) can provide a useful measure of population density, assuming that the intensity or frequency of field signs is related in some way to the number of animals present (Wilson and Delahay, 2001). Monitoring of field sign can be used if it is more efficient to measure the cue than animal abundance, either because the cue is cheaper or more accurate to measure, or the cue has an earlier response (Vos et al., 2000). Such approaches first require an estimate of the cue density, as obtained e.g., using distance sampling (e.g., walk a line transect and measure distance between transect and cue, section 2.2). Then an estimate of the conversion rate between number of individuals present and density of the cue is required. This can be problematic because cue rate may be density dependent and/or the decay rate of the cue needs to be estimated (Seber, 1986; Wilson and Delahay, 2001). Such conversion rates are sensitive to changes in environmental conditions and behavior of individual animals. Therefore, a change in the index may not necessarily reflect changes in population size. Also, given that indirect methods are based on signs produced by animals, there is a loss of information on how many animals are present. Consequently, estimates from these methods will have less precision (Seber, 1992).

Cues produced by small mammals include dung, nests and tracks. There has been considerable success using dung surveys to estimate density of larger mammals (e.g., ungulates, Marques et al. 2001; and elephants) but typically dung produced by small mammals is too small to readily detect (Dique et al., 2004). However, Velazquez (1994) successfully estimated population density of the volcano rabbit (*Romerolagus diazi*) in Mexico using plot-based surveys of rabbit pellets and found a high correlation with abundance of rabbits as estimated using line transects. In addition, Karels et al. (2004) found pellet counts strongly correlated with the abundance of the Hoary marmot (*Marmota caligata*), and concluded that counts of fecal pellets could form the basis of a long-term monitoring program. On the contrary, St-Laurent and Ferron (2008) found dung surveys were a poor index of relative abundance for the southern red-backed vole (*Myodes gapperi*) in Canada.

Surveys of nests (e.g., dusky footed woodrat, *Neotoma fuscipes* in California, Vreeland and Tietje 1999; Hamm et al. 2002; and red squirrel, *Tamiasciurus hudsonicus* in Yellowstone National Park, Mattson et al. 1996) and track detection (e.g., various small mammals in Artic tundra, Boonstra et al. 1992, North Dakota, Wiewel et al. 2007, and Nova Scotia, Nams and Gillis 2003; and rats (*Rattus rattus*) and mice (*Mus musculus*) in New Zealand, Brown et al. 1996) have had mixed success when used to estimate abundance.

Alternatively, removal methods exist, where the level of field sign is calculated before and after a known number of animals are removed (e.g., Adkins 2003), however a high proportion of animals needs to be removed to have any confidence in the resulting density estimates (Seber, 1986). Some capture-recapture survey designs (including trapping webs) consider only the first capture event of each animal (and thus can be considered removal methods), but they can be inefficient as not all data are used in the analysis. Since I am focusing on rare or endangered

small mammals, removal methods are not appropriate and will not be considered further.

## 2.8 Occupancy modelling

Instead of converting indirect observations (i.e., cues, section 2.7) into an estimate of abundance via a conversion factor, the data can be analyzed within an occupancy modelling framework (MacKenzie et al., 2006). In these models, the proportion of sites occupied by a species is estimated. These methods can be implemented more easily and less expensively than other methods available to estimate abundance (MacKenzie et al., 2002). Occupancy modelling can be used to estimate abundance, by treating abundance as a random variable with a parametric form (e.g., a Poisson model, Royle and Nichols 2003); and can reliably model changes in range of the species.

Since my aim is to develop methods to estimate abundance, I do not consider occupancy modelling further. However, should an alternative, inexpensive, practical and reliable method not be found to estimate abundance of the Key Largo woodrat, presence/absence surveys can be implemented to monitor changes in, e.g., species range. It is not advised, however, to assume that a change in the range of the species may directly correlate with a corresponding change in abundance.

## 2.9 Methods applicable to estimating abundance of the Key Largo woodrat

To develop a long-term monitoring program to estimate population size of the Key Largo woodrat, it is necessary to consider the currently available monitoring methods that might be applicable and assess them according to statistical properties, practical needs and cost. In terms of equipment cost, methods based on field sign are cheaper to implement than, e.g., trapping and radio collaring approaches, as such approaches do not require any or minimal field equipment. Field sign produced by the woodrat includes nest, dungs and tracks.

Nest counts have been used to estimate density of the Key Largo woodrat (Linsdale and Tevis, 1956; Sakai and Noon, 1993; Gore and Loggins, 2005) and also other woodrat species (e.g., the dusky-footed woodrat, *Neotoma fuscipes*, Sakai and Noon 1993; Hamm et al. 2002; Vreeland and Tietje 1999; Laudenslayer and Fargo 2002). Nest counting using distance sampling is an appealing method of monitoring the woodrat because it is not labor intensive, and one can survey a larger portion of an area more efficiently and quickly than trapping in grids (Sakai and Noon, 1993). Nest surveys can be undertaken as plot surveys or line transect surveys. Once the number of nest sites in an area has been estimated, a conversion rate to calculate the abundance of woodrats can be made. Determining an appropriate conversion rate is difficult. Using nest surveys to estimate woodrat abundance has two problems. Firstly, there is a poor correlation between the number of nests found and the estimated abundance described through live trapping because woodrats can construct and maintain multiple stick nests, and the number of nests per woodrat appears to be variable (Linsdale and Tevis, 1956; Vreeland and Tietje, 1998; Gore and Loggins, 2005; Laudenslayer and Fargo, 2002). Radio collaring will help to estimate how many nests the radio-collared individual is maintaining, by radio-tracking during the

day time when the nocturnal woodrat should be in its nest.

Secondly, traditional stick nests on Key Largo are now rarely found (Gore and Loggins, 2005), and the population may have altered its behavior to nest in other materials (e.g., rubbish piles, Winchester et al. 2009) which are more difficult to detect. Therefore, estimating the number of nest sites might be biased to the extent that cryptic nests go uncounted. For example, Sakai and Noon (1993) captured woodrats in traps, when no nests or signs of foraging were observed. In such circumstances, radio-tracking woodrats may help identify the proportion of cryptic nests that exist. A line transect approach may alleviate this issue, however estimates will still be biased to the extent that detection on the line may be less than one (violating an assumption of distance sampling).

Other sign produced by woodrats is dung and tracks. Dung surveys are considered inappropriate because although the droppings of woodrats are distinctive (Vestral, 1938), they are very small ( $< 10\text{mm}$ ). Consequently, meeting the assumption in line transect sampling that detection at zero perpendicular distance is perfect, is unlikely. In plot sampling of dung, plots would have to be small to ensure that all droppings within the plot are detected. This will most likely cause a large number of plots to have zero droppings, so a very large sample of plots will be required, and thus making this approach very labor intensive. Due to issues with sample size and detection, I would not recommend dung surveys for estimating the density of the woodrat. One additional problem with dung surveys is it is unknown how many woodrats actually contribute to any dung piles that may be detected. Therefore, a conversion index would be required to estimate how many dung piles are produced by how many woodrats, and this could be difficult to ascertain.

Track detection via tracking tubes was proposed by Gore and Loggins (2005) as the most efficient method for monitoring the presence of the woodrat, and

tracking tubes have been used to detect the presence of other small mammals in the area (e.g., beach mice, *Peromyscus polionotus*, Loggins et al., 2010). Tracking tubes are water proof, inexpensive to operate and require little maintenance. However, like trapping, animals can respond to the tracking tubes in a positive way (e.g., tracking tubes may provide food and shelter) or in a negative way (e.g., animals may be wary of new items appearing in their home range, Nams and Gillis, 2003; Gore and Loggins, 2005). Similarly to nest and dung surveys, track detection requires a conversion rate to be estimated to translate the number of tracks observed into an estimate of abundance. When monitoring track counts over time at a location, the ability to detect real changes in abundance is typically hampered by differences in the probability of finding tracks between sampling visits. Observed differences may be due to factors that are not related to population size, such as differences in activity, weather or food availability (Wilson and Delahay, 2001). Also, there are practicality issues associated with these methods, e.g., despite tracking tubes being proposed by Gore and Loggins (2005) as “the most efficient method for monitoring” woodrat presence, after field trials they concluded “we cannot confidently distinguish between tracks of woodrats and black rats” (p.20), and note that tracks obtained in areas where black rats are known or suspected, track identification should be treated with caution. Given that black rats are thought to occur throughout the range of the woodrat (albeit, at low densities), this may cause track identification issues. Using the number of tracks to determine density or abundance is unreliable, as there may be a non-linear relationship between tracks and density.

Cues such as nest, dung and tracks can be analysed within an occupancy modelling framework, however one can only obtain reliable estimates of range and not species abundance. Hence, unless an alternative method is not found, I do not recommend occupancy modelling. However, camera stations have been

used to monitor Key Largo woodrat behavior (C. Alligood, unpub. data). An individual with an ear tag has been seen on film, but the individual's ear tag number could not be distinguished (C. DeGayner, pers. comm.). As technology improves, it may be feasible to use camera stations in a mark-recapture survey.

To date, all studies undertaken on the Key Largo woodrat (Table 1.1, on page 12) have used some form of capture-recapture design based on trapping individual woodrats. The two most recent and comprehensive trapping surveys undertaken on the woodrat have employed several intensely-trapped grids at randomly selected locations throughout the woodrat habitat (McCleery, 2003; Winchester, 2007). Since the density of the woodrat is low, recapture rates have been small, and hence reliable estimates of population size using this approach have not been achieved (c.f. section 1.2). In addition, since the population occupies such a large area, for good estimates to be obtained, a large financial investment is needed (e.g., a field crew working year-round running a trapping program).

Efficiency of capture-recapture studies of the Key Largo woodrat might be improved by integrating less expensive occupancy modelling methods (e.g. tracking tubes) within a CR framework (e.g. Conroy et al. 2008).

Inefficiencies also arise when set traps capture non-target species (e.g., it is common to capture the Key Largo cotton mouse, *Peromyscus gossypinus allapaticola*, when surveying for the woodrat), or are disturbed such that no animals are caught. For example, Humphrey (1988) reported a large trap disturbance when trapping for the woodrat caused by raccoons, and ended up trapping for raccoons and woodrats concurrently to increase the efficiency in sampling for the woodrat.

Once captured, the standard method of marking woodrats has been to double tag them with passive intergrator transponder (PIT) tags and ear tags. PIT

tagging offers a great technological advancement to other marking methods, because there is a high reliability in tag detection (95–100%) and reading accuracy ( $\approx 100\%$ ) (Gibbons and Andrews, 2004). This technology is considerably more expensive than, e.g., ear tagging (PIT tag  $\approx$  USD 7, ear tag  $\approx$  USD 0.20), however the benefits of reduced tag loss and correct identification of animals (a very important assumption of capture-recapture analysis) outweighs the cost. There have been no reports of problems associated with double tagging woodrats using these techniques (e.g., decreased survival of double tagging).

A pertinent issue of capture-recapture is heterogeneity in capture probability (c.f. section 2.3). To reduce individual heterogeneity in capture probabilities, traps would need to be independently redistributed through the whole study area between each sampling occasion, but this would result in far too few recaptures. Alternatively, several intensively-trapped plots would be needed, so that density can be estimated at representative locations. However, without strong assumptions, it is very difficult to estimate the effective area surveyed for each plot. Regardless, since capture-recapture is the most common technique for monitoring small mammals (Solari et al., 2002), this is the first method I chose to investigate in estimating Key Largo woodrat abundance in the following chapter. In addition, a base-line of the current standard method is required with which newer methods can be compared, to assess any improvements in efficiency.

A recent development of standard capture-recapture methods has been the incorporation of spatial capture information, in spatially-explicit capture-recapture. Spatially explicit capture-recapture estimation has been applied to a variety of bird species (Efford et al., 2004), the house mouse (*Mus musculus*, Efford 2004), and the brushtail possum (*Trichosurus vulpecula*, Efford et al. 2005). Since data collected in a standard capture-recapture manner from Chapter 3 can also be

analysed within a spatially-explicit capture-recapture framework without additional field costs, I also investigate this method in Chapter 4.

The other category of sampling methodology is distance sampling. Despite its appeal, distance sampling does not seem an appropriate method to form part of a long-term monitoring program for woodrat abundance. Firstly, vegetation on Key Largo is very dense and with the woodrat being nocturnal, detecting individuals would rely on eye-shine reflection using a spotlight. Consequently, it may be difficult to detect animals and/or they may exhibit a response to an observer moving through the vegetation due to noise. Also, the safety of the observer must take priority, and it would be unsafe for an observer to undertake a distance sampling survey at night, in the dark, through difficult terrain. Secondly, it seems likely that detection on the line or at the point will be imperfect for the woodrat, and it is likely to prove difficult to estimate the corresponding probability of observing individuals on the line. Finally, sample size of detected woodrats is likely to be very small for realistic levels of search effort. Distance sampling works well when animals are readily visible, however for the woodrat, this is not the case. Consequently, the detection of a woodrat usually relies on trapping it. Traditionally such data is analyzed within a capture-recapture framework (c.f. Section 2.3). However, there are methods available that combine trapping studies with distance sampling, trapping webs and trapping point transects.

The use of trapping webs does not seem cost efficient because so many traps are required (e.g., using 15 webs with 90 traps in each equates to 1,350 traps being set for multiple nights, thus requiring a large amount of field time, Lukacs et al., 2005). Therefore, trapping webs are not considered further.

From other recent studies on the woodrat, we know that small sample sizes and low recapture rates are problematic (McCleery, 2003; Winchester, 2007).

Consequently, problems with applying the trapping point transect approach to the woodrat will be sample size issues. The woodrat currently has a sparse density, so obtaining individuals on which to conduct trials will be difficult. Then once a known-location woodrat has been obtained, recapture rates might be quite low. These small sample sizes in the trial survey may result in a poorly fitted detection function. Issues may also arise with the radio collaring aspect of the approach, such as technological difficulties with radio transmitters, inability to re-detect radio-collared individuals due to thick forest cover, and the initial expense of purchasing radio collars.

Trapping point transects potentially allow abundance estimates of species that can be trapped, with fewer resources needed than trapping webs and conventional capture-recapture methods (Buckland et al., 2006). However, I recognize the monetary cost of radio transmitters ( $\approx$  USD 224 per collar), relative to the cost of a trap ( $\approx$  USD 33 per trap), is great. It is possible to conduct repeated trials on individuals, and/or rotate radio collars between individuals, to maximise the data obtained for minimal cost. Also, if telemetry work is going to be conducted for other purposes (e.g., habitat use studies), the use of the trapping point transect approach could be carried out at little additional expense by using existing radio-collared animals. In terms of initial financial outlay to establish the TPT survey, one would need to purchase as many radio collars that would be deployed simultaneously and traps that would be set simultaneously (considering both the trials and main survey). For example, if ten radio collars were deployed simultaneously, ten radio collars must be purchased along with 20 traps to conduct the trials survey; plus additional traps to conduct the main survey. To my knowledge, the trapping point transect method has not been implemented in the field to monitor small mammal abundance via trapping. As such, a detailed methodology is provided in Chapter 5, a simulation exploring

survey effort is given in Chapter 6, and a case study as applied to the Key Largo woodrat is in Chapter 7.

TPT and SECR methods do not require estimation of the surveyed area, a problem encountered in capture-recapture analyses to estimate population size, and hence associated problems that arise from the edge effect do not occur. In addition, these methods explicitly account for detectability, an important factor for the woodrat which is considered to have low density and demonstrate cryptic behavior that results in low detectability. However, since capture-recapture is so commonly applied to monitor small rodents, I also investigate this method in the following chapter.

## Chapter 3

# Estimating Key Largo woodrat abundance using a capture-recapture approach

### 3.1 Introduction

To date, most studies designed to estimate Key Largo woodrat abundance have used capture-recapture statistics with a grid-based design (see Table 1.1 on page 12). The two most recent, and by far the most comprehensive surveys to estimate woodrat abundance, were undertaken by McCleery (2003) and Winchester (2007).

McCleery (2003) randomly placed 20 1 ha trapping grids in each of three habitat strata present in the woodrat's geographic range (young, medium and old growth hammock, Section 1.2.2). Each grid was a 5 by 5 array, with one single-catch Sherman trap placed at each trap station (i.e., 25 traps per grid) with a 25 m trap spacing. Grids were trapped for four consecutive nights. This yielded 16 woodrat capture events of 13 individual woodrats in the 6,000 trap nights undertaken between March and September, 2002. Since capture rates were so low, only naïve density estimates could be calculated, i.e., number of individuals captured divided by 20 ha of habitat trapped in each stratum (Otis et al., 1978). The population size of the entire study area was estimated to

be 106 individuals (95% CI: 30-182 individuals). Naïve density estimates are known to be biased upwards as the effective area trapped by each grid is much larger than the actual area of the grid (i.e., 20 ha). This is because animals with home range centres not on the grid are still detected on the grid due to animal movement (Wilson and Anderson, 1985a).

Despite such a large survey effort and investment of financial resources by McCleery (2003), due to low capture rates, no modelling of the data within a capture-recapture framework could be undertaken. Winchester (2007) recognised capture rates might be increased by changing the survey design, and compared the efficiency of adaptive cluster sampling (ACS) to stratified-random sampling (SRS) survey designs for estimating abundance of the Key Largo woodrat. In the SRS survey, 40 0.5 ha trapping grids were established in each of the three habitat strata. Each grid was a 3 by 3 array, with two single-catch Sherman traps placed at each trap station (i.e., 18 traps per grid) with a 25 m trap spacing. Grids were trapped for 4 consecutive nights between April and June, 2005. In the ACS survey, an additional four grids were placed adjacent to any of the 40 SRS grids where at least one woodrat was captured. This led to an additional 33 grids being sampled in the ACS survey. Of 40 grids trapped in the SRS survey, 11 individuals were captured on seven grids. Twenty-two additional individuals were captured on 33 additionally surveyed grids in the ACS survey.

Due to low number of captures, Winchester (2007) pooled all individuals for the analysis, and a closed population model was assumed (see Section 2.3). Behavior and individual heterogeneity in capture probability were found to be important. Individual covariates (e.g., sex) were not investigated. Like McCleery (2003), Winchester (2007) also used naïve density estimates but corrected this estimate by the proportion of time trappable animals spent on the grid, as estimated using radio tracking data (White and Shenk, 2001). Total population

size was estimated at 321 (95% confidence interval: 13-629) and 323 (95% CI: 0-652) individuals, for the SRS and ACS surveys, respectively.

Despite an 83% increase in sample size using the ACS compared to the SRS survey, the variance estimate of ACS was higher, resulting in a less precise estimate of woodrat abundance. Winchester (2007) estimated 100 0.56 ha trapping grids would be required to estimate abundance of woodrats within 50% of the actual value, with 90% confidence using the SRS survey method and go on to conclude their “results suggest that estimating abundance of the entire Key Largo woodrat population is likely to be cost prohibitive”. The development of new, and more cost-effective, monitoring methods are explored in the following chapters, but in order to compare these new methods to current practice (i.e., capture-recapture using a grid based approach; McCleery, 2003; Winchester, 2007), I undertake a capture-recapture analysis to estimate abundance of the Key Largo woodrat in this chapter.

All data presented in this chapter were collected by Dan Greene (University of Georgia) as part of his Masters project studying the Key Largo cotton mouse (*Peromyscus gossypinus allapaticola*, Greene 2009). As the traps used to capture the cotton mouse were not species-specific, woodrats were captured and essentially considered ‘bi-catch’ to his cotton mouse survey. The cotton mouse maintains a much smaller home range than the woodrat, so trap spacing and grid size were less than previous capture-recapture surveys of the woodrat implemented by McCleery (2003) and Winchester (2007). Nevertheless, the data were collected in a manner that allows us to estimate abundance of the Key Largo woodrat using a capture-recapture analysis.

## 3.2 Methods

### 3.2.1 Survey design

Using a stratified random design with proportional allocation, 32 grids were established in the study area across three habitat strata (young, medium, and old, c.f. Section 1.2.2; a fourth stratum, disturbed/urban, was excluded from the analysis because it was too small to accommodate randomly located grids<sup>1</sup>). Sampling units were selected in proportion to the area available in each stratum, such that each sampling unit had an equal inclusion probability. The 32 grids are mapped in Figure 3.1 and were distributed as follows: Young hammock (disturbed after 1971) - 4 grids, Medium hammock (disturbed between 1940 and 1971) - 12 grids, and Old hammock (disturbed before 1940) - 16 grids.

### 3.2.2 Trapping and handling

Each of the 32 established grids were trapped according to Pollock's robust survey design (Pollock et al., 1990, and Section 2.3). There were three primary sampling sessions: March - April, July - September, and November - December, 2007, and within each primary session, each grid was trapped for four secondary sessions (i.e., four consecutive nights). Each grid was a 7 by 7 array, with one trap single-catch Sherman trap<sup>2</sup> placed at each trap station (i.e., 49 traps per grid) with a 10 m trap spacing (Figure 3.2). Traps were opened and baited with whole rolled oats in the late afternoon and checked the following morning within the first two hours after sunrise. All captured woodrats were double-marked with passive integrated transponders (PIT)<sup>3</sup> and ear<sup>4</sup> tags, and sex was recorded.

---

<sup>1</sup>As the disturbed/urban stratum was excluded from the analysis, inference was based on capture events in the three other stratum included in the analysis.

<sup>2</sup>10.2 x 11.4 x 38.1 cm, vented, Sherman trap with a raccoon-proof door, model PXL15, H. B. Sherman Traps Inc., Tallahassee, Florida.

<sup>3</sup>AVID, Norco, California.

<sup>4</sup>#1005 Monel ear tags, National Band and Tag, Newport, Kentucky.

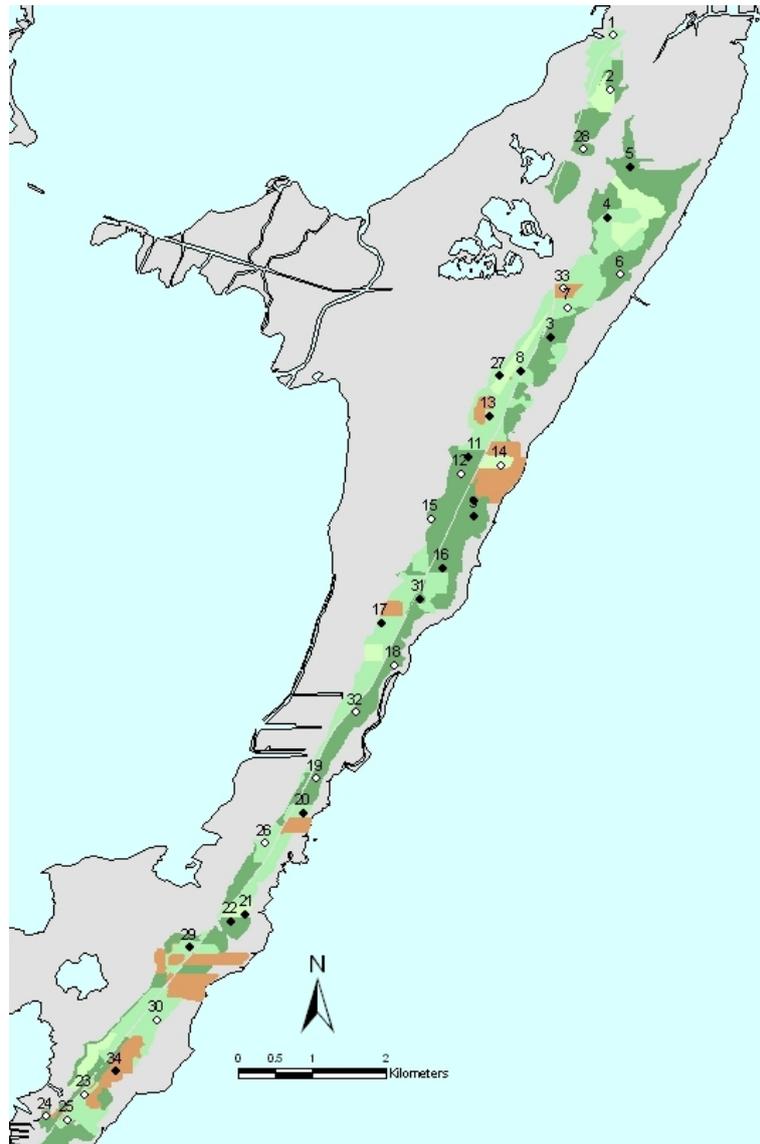


Figure 3.1: A stratified random design with proportional allocation was used to locate 32 trapping grids in three habitat strata delineated by age (shaded light to dark green representing young to old habitat, respectively). Two additional grids (#33 and 34) located within disturbed/urban habitat (shaded brown) were omitted from the analysis. Black and white circles represent grids where woodrats were captured or not, respectively, at some point during data collection (i.e., any primary session). Unsuitable habitat is shaded blue (water) and grey (mangrove swamp).

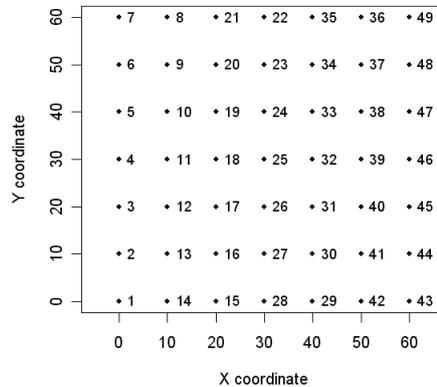


Figure 3.2: Diagram of trapping grid layout. Each grid was a 7 by 7 trapping square, with a 10 m spacing between traps.

### 3.2.3 Model selection

Data were analysed using program MARK (v.5.2, White and Burnham, 1999). Huggins' full heterogeneity model with a robust design was specified that used an open population, Jolly-Seber model between primary sessions (Kendall, 2008). Within each primary session, a model that combines Huggins' closed population model (Huggins, 1989, 1991) with the mixture models for heterogeneity of Pledger and Efford (1998); Pledger (2000) and Norris and Pollock (1995) was specified. This closed population model is conditioned on each animal that is caught at least once in each primary trapping period (i.e., population size,  $N$ , is not included in the likelihood). This allows capture probability,  $p$ , and recapture probability,  $c$ , to be modelled as functions of individual covariates (here, sex; White, 2008). Since the data were naturally sparse (i.e., few captures across the 32 grids, despite a large survey effort), grid could not be used as a covariate in the model, and consequently the probability of capture (and recapture, where appropriate) is assumed equal across all grids. As population size is not

included in the likelihood, it is derived using a Horvitz-Thompson estimator (i.e.,  $\hat{N} = \sum_{i=1}^n 1/p_i$ , where  $p_i$  is the probability of first capture for individual  $i = 1, \dots, n$ ; calculated for each primary session).

Given there are three primary sessions and four secondary sessions, the full, time dependent model has 50 parameters (Table 3.1), and cannot be fitted due to parameter identifiability issues (White, 2005). Staying in the realms of biological plausibility, a set of 31 models were chosen *a priori*. Twenty of these models did not include individual heterogeneity (i.e.,  $\pi_1 = 1$ ) and investigated the effect of session (*sess*) and sex on capture ( $p$ ) and recapture ( $c$ ) probability, and whether temporary emigration was not present ( $\gamma'' = \gamma' = 0$ ), or was random ( $\gamma'' = \gamma'$ , Table 3.2). A Markovian emigration model could not be fitted as with only three primary sessions,  $\gamma'$  is confounded with survival. A further eleven models were fitted that investigated heterogeneity not already modelled (i.e., due to individual, excluding a behavioral response due to capture or affect of sex), using a two-part mixture (i.e., each variable was estimated using the mixture of two distributions, where  $g$  was the group in the mixture model). Due to parameter identifiability issues, all heterogeneity models assumed there was no migration, except one (model  $\{p(g + sex).c(g + sex)\}$ , where random migration was also investigated, Table 3.3). Akaike's Information Criterion (AIC) was used to select models (Burnham and Anderson, 2003).

Table 3.1: Definition of terms used in model formulae presented in Tables 3.2 and 3.3. There were 50 parameters in the full, time dependent model relating to survival ( $S$ ), migration ( $\gamma''$  and  $\gamma'$ ), the two-part mixture process for heterogeneity ( $\pi$ ), and probability of capture ( $p$ ) and recapture ( $c$ ).

Symbol	No. of parameters	Definition
$S$	2	Survival between primary session 1 and 2, and survival between session 2 and 3.
$\gamma''$	2	Probability of emigrating away from the study area in either session 1 or session 2.
$\gamma'$	1	Probability of remaining away from the study area between sessions 2 and 3, given animal has emigrated in the previous time step (session 1).
$\pi$	3	Probability of mixture distribution, one for each primary session.
$p$	24	Apparent encounter probability, which is conditional on the probability that the animal is alive and available for recapture (2 mixture distributions x 4 secondary sessions x 3 primary sessions).
$c$	18	Probability of recapture (2 mixture distributions x 3 secondary sessions x 3 primary sessions).

Table 3.2: Summary of the models fitted to trapping data collected on the Key Largo woodrat. Models included the probability of capture ( $p$ ) and probability of recapture ( $c$ ) (dependent on trapping session,  $sess$ , sex of woodrat, and habitat,  $hab$ ). Each of these ten models included the probability of temporary emigration being either random ( $\gamma'' = \gamma'$ ) or not (i.e., no movement  $\gamma'' = \gamma' = 0$ ) (i.e., 20 models were run in total, but for brevity I report only 10). Survival between sessions was always assumed constant.

Model notation	Definition
$\{p(\cdot) = c(\cdot), \gamma'' = \gamma'\}$	Null model, capture probability is constant for duration of study.
$\{p(sess) = c(sess), \gamma'' = \gamma'\}$	Probability of capture changes between the three primary trapping sessions.
$\{p(sex) = c(sex), \gamma'' = \gamma'\}$	Probability of capture is different for male and female woodrats.
$\{p(sess + sex) = c(sess + sex), \gamma'' = \gamma'\}$	Probability of capture is different across sessions, and affect of sex is additive.
$\{p(sess * sex) = c(sess * sex), \gamma'' = \gamma'\}$	Probability of capture is different for male and female woodrats across sessions.
$\{p(sess * hab) = c(sess * hab), \gamma'' = \gamma'\}$	Probability of capture is different for habitat stratum across sessions.
$\{p(\cdot).c(\cdot), \gamma'' = \gamma'\}$	Capture probability is constant for duration of study. Once an individual has been captured, its probability of recapture changes (i.e., a behavioral response, such as trap happiness or trap shyness).
$\{p(sess).c(sess), \gamma'' = \gamma'\}$	Probability of capture changes between the three primary trapping sessions. Once an individual has been captured, its probability of recapture changes (i.e., a behavioral response), depending on session.
$\{p(sex).c(sex), \gamma'' = \gamma'\}$	Probability of capture and recapture is different for male and female woodrats.
$\{p(sess + sex).c(sess + sex), \gamma'' = \gamma'\}$	Probability of capture and recapture is different across sessions, but affect of sex is additive.

Table 3.3: Summary of the 11 models fitted to the trapping data collected on the Key Largo woodrat. Models included the probability of capture ( $p$ ) and probability of recapture ( $c$ ) (dependent on trapping session,  $sess$ , and sex of woodrat), probability of temporary emigration being random ( $\gamma'' = \gamma'$ ) or not (i.e., no movement  $\gamma'' = \gamma' = 0$ ), and heterogeneity ( $g$ ) based on a 2-point mixture process. Survival and the 2-point mixture process was always assumed constant across each session.

Model notation	Definition
$\{p(g) = c(g), \gamma'' = \gamma' = 0\}$	Probability of capture is constant for duration of study, but is separated into two groups (i.e., a high capture probability group and a low capture probability group). No temporary emigration between sessions.
$\{p(g + sex) = c(g + sex), \gamma'' = \gamma' = 0\}$	Probability of capture is separated into two groups with an additive sex effect. No temporary emigration between sessions.
$\{p(sess + g) = c(sess + g), \gamma'' = \gamma' = 0\}$	Probability of capture differs across session, and group. No temporary emigration between sessions.
$\{p(g * sex) = c(g * sex), \gamma'' = \gamma' = 0\}$	Probability of capture is separated into two groups with an interactive sex effect. No temporary emigration between sessions.
$\{p(sess + g + sex) = c(sess + g + sex), \gamma'' = \gamma' = 0\}$	Probability of capture differs across session, with an additive effect of sex and group. No temporary emigration between sessions.
$\{p(sess + g * sex) = c(sess + g * sex), \gamma'' = \gamma' = 0\}$	Probability of capture differs across session, with an interactive effect of sex and group. No temporary emigration between sessions.
$\{p(sess * sex + g) = c(sess * sex + g), \gamma'' = \gamma' = 0\}$	Probability of capture differs across session and sex, with an additive effective of group. No temporary emigration between sessions.
$\{p(g).c(g), \gamma'' = \gamma' = 0\}$	Probability of capture is constant for duration of study, but depends on group. Probability of recapture differs (i.e., behavioural effect) and depends on group. No temporary emigration between sessions.
$\{p(g + sex).c(g + sex), \gamma'' = \gamma'\}$	Probability of capture with an additive sex effect. Once an individual has been caught, its probability of recapture changes (i.e., a behavioral response). Random temporary emigration between sessions.
$\{p(g + sex).c(g + sex), \gamma'' = \gamma' = 0\}$	As per previous model, with no temporary emigration.
$\{p(sess * g).c(sess * g), \gamma'' = \gamma' = 0\}$	Probability of capture and recapture differs across session and group. No temporary emigration between sessions.

### 3.2.4 Abundance estimation

When using Huggins' conditional likelihood model, abundance estimates for each primary session are derived using a Horvitz-Thompson estimator:  $\hat{N}_c = \sum_{i=1}^n 1/p_i$ , where  $p_i$  is the probability of first capture for individual  $i = 1, \dots, n$ . This estimate of  $\hat{N}_c$  is the estimated population size in the covered survey region,  $A_c$  (i.e., effective trapping area of the grid). To incorporate model selection uncertainty, the estimate of  $\hat{N}_c$  was obtained by model averaging over all 29 models based on the AIC weight (Burnham and Anderson, 2003). To estimate the total population size of the region ( $\hat{N}_t$ ),  $\hat{N}_c$  was converted into density (i.e.,  $D = \hat{N}_c/A_c$ ), and multiplied by the total survey region (i.e., 850 ha of hammock). Estimating  $A_c$  is difficult and numerous methods exist. Three methods were investigated that relied on estimating  $A_c$  (naïve density estimates, home range radius and mean maximum distance moved):

**Naïve density estimates:** Assume the effective trapping area of each grid is equal to the area of the trapping grid (here, 0.36 ha).

**Home range radius:** Dice (1938) proposed adding a strip the width of one home range radius to the area of the grid. Assuming home ranges of woodrats are circular (0.25 ha, see section 1.2), the radius of the home range is 28.2 m. Consequently, the effective area trapped by each grid was 1.29 ha.

**MMDM:** Wilson and Anderson (1985a) proposed using the mean maximum distance moved (MMDM) for each individual recaptured at least twice. In this data set, the MMDM by each recaptured woodrat was 25.7 m, 14.3 m, and 17.0 m, in session 1, 2, and 3, respectively. Consequently, the effective area trapped by each grid was 1.18 ha, 0.77 ha, and 0.86 ha in session 1, 2, and 3, respectively.

A fourth method was used that corrected the estimated animal abundance by the proportion of time animals spent on the grid (radio tracking):

**Radio tracking:** White and Shenk (2001) proposed estimating the average proportion of time individuals are available for capture on the grid,  $\bar{P}_A$ , and related the naïve density estimator to this factor accordingly ( $\hat{D} = (\hat{N} * \bar{p})/A_c$ ). Naïve estimates for Key Largo woodrats are upwardly biased by 20% (Winchester, 2007), and this same correction factor was used here.

### 3.3 Results

In total, there were 18,816 trapping nights (= 32 grids x 49 traps x 4 nights x 3 sessions), during which there were 144 capture events of 51 unique woodrats (Table 3.4). This equates to an encounter rate of 0.76 woodrats per 100 trap nights. No woodrat was trapped on more than one grid, even between primary trapping sessions, however 13 unique woodrats were caught on the same grid across multiple primary sessions.

Table 3.4: Number of woodrat capture events for 3 primary sessions (Mar-Apr, Jul-Sept, Nov-Dec) across 3 habitat strata. The number of individual woodrats captured are given in parentheses. Row totals (i.e., by strata) for the number of woodrats caught are not a direct summation of each row (highlighted by an †), as some woodrats were caught across multiple sessions.

Strata	Session			Total
	1	2	3	
Old	26 (9)	16 (9)	45 (21)	87 (31 <sup>†</sup> )
Medium	10 (4)	14 (7)	16 (7)	40 (14 <sup>†</sup> )
Young	4 (2)	4 (2)	9 (4)	17 (6 <sup>†</sup> )
Total	40 (15)	34 (18)	70 (32)	144 (51)

The model with the lowest AIC included capture probability as a function of session and sex (Table 3.5). A further 8 models had a difference in AIC of less than 2 (and therefore are considered equally plausible models, Burnham and

Anderson, 2003, Table 3.5). Of these top nine models, a behavioral response was not included (i.e.,  $p = c$ ), eight included sex, seven included session and six included heterogeneity. Eight of these nine models assumed no temporary emigration (i.e.,  $\gamma'' = \gamma' = 0$ ), but this is more likely to be a function of the models fit (i.e., heterogeneity models could not include temporary or Markovian emigration due to parameter identifiability).

Model averaged estimates of population size in the covered region increased with session from 0.55 woodrats per grid (se = 0.14) in session 1, to 0.84 (se = 0.26) woodrats per grid in session 2 and 1.36 (se 0.39) woodrats per grid in session 3. However, depending on which method was used to estimate the effective area trapped on the grid (i.e., naïve density estimate, half home range radius, mean maximum distance moved, and a radio tracking correction factor), the estimate of total population size varied widely (by as much as a factor of 3.5, Table 3.6).

Table 3.5: Summary of 29 models were fitted to the Key Largo woodrat capture data. Models included the probability of capture ( $p$ ) and probability of recapture ( $c$ ) (dependent on trapping session,  $sess$ , and sex of woodrat), probability of temporary emigration being random ( $\gamma'' = \gamma'$ ) or not (i.e., no movement  $\gamma'' = \gamma' = 0$ ), and heterogeneity ( $g$ ) based on a 2-point mixture process. Survival and the 2-point mixture process was always assumed constant across each session. Eight models had a difference in AIC of less than 2 from the top model. AIC weights (wAICc) and the number of parameters (No. Par.) in each model are also provided.

Model	AICc	$\Delta$ AICc	wAICc	No. Par.
$\{p(sess * sex) = c(sess * sex), \gamma'' = \gamma' = 0\}$	415.40	0.00	0.14	7
$\{p(sess * sex + g) = c(sess * sex + g), \gamma'' = \gamma' = 0\}$	415.62	0.22	0.12	9
$\{p(sess + g + sex) = c(sess + g + sex), \gamma'' = \gamma' = 0\}$	415.73	0.33	0.12	7
$\{p(sess + g) = c(sess + g), \gamma'' = \gamma' = 0\}$	415.97	0.57	0.10	6
$\{p(g + sex) = c(g + sex), \gamma'' = \gamma' = 0\}$	416.49	1.10	0.08	5
$\{p(sess + g * sex) = c(sess + g * sex), \gamma'' = \gamma' = 0\}$	416.69	1.29	0.07	8
$\{p(sess * sex) = c(sess * sex), \gamma'' = \gamma'\}$	416.80	1.40	0.07	8
$\{p(g * sex) = c(g * sex), \gamma'' = \gamma' = 0\}$	417.15	1.75	0.06	6
$\{p(sess + sex) = c(sess + sex), \gamma'' = \gamma' = 0\}$	417.17	1.77	0.06	5
$\{p(g) = c(g), \gamma'' = \gamma' = 0\}$	418.30	2.90	0.03	4
$\{p(sess + sex) = c(sess + sex), \gamma'' = \gamma'\}$	419.02	3.62	0.02	6
$\{p(g + sex).c(g + sex), \gamma'' = \gamma' = 0\}$	419.17	3.77	0.02	8
$\{p(sess * g) = c(sess * g), \gamma'' = \gamma' = 0\}$	419.92	4.52	0.01	8
$\{p(sess + sex).c(sess + sex), \gamma'' = \gamma' = 0\}$	420.26	4.86	0.01	9
$\{p(sess).c(sess), \gamma'' = \gamma' = 0\}$	420.73	5.33	0.01	7
$\{p(sess) = c(sess), \gamma'' = \gamma' = 0\}$	420.94	5.54	0.01	4
$\{p(sex) = c(sex), \gamma'' = \gamma' = 0\}$	421.03	5.63	0.01	3
$\{p(g + sex).c(g + sex), \gamma'' = \gamma'\}$	421.16	5.76	0.01	9
$\{p(sess).c(sess), \gamma'' = \gamma'\}$	421.49	6.09	0.01	8
$\{p(sess + sex).c(sess + sex), \gamma'' = \gamma'\}$	421.62	6.23	0.01	10
$\{p(sex) = c(sex), \gamma'' = \gamma'\}$	421.77	6.37	0.01	4
$\{p(g).c(g), \gamma'' = \gamma' = 0\}$	421.90	6.50	0.01	6
$\{p(sess) = c(sess), \gamma'' = \gamma'\}$	422.42	7.02	0.00	5
$\{p(sess * hab) = c(sess * hab), \gamma'' = \gamma'\}$	423.07	7.67	0.00	5
$\{p(sex).c(sex), \gamma'' = \gamma' = 0\}$	423.71	8.31	0.00	5
$\{p(sess * hab) = c(sess * hab), \gamma'' = \gamma' = 0\}$	424.59	9.20	0.00	6
$\{p(sex).c(sex), \gamma'' = \gamma'\}$	425.00	9.60	0.00	6
$\{p(.) = c(.), \gamma'' = \gamma'\}$	425.54	10.14	0.00	3
$\{p(.) = c(.), \gamma'' = \gamma' = 0\}$	425.65	10.26	0.00	2
$\{p(.) . c(.), \gamma'' = \gamma' = 0\}$	425.88	10.48	0.00	3
$\{p(.) . c(.), \gamma'' = \gamma'\}$	426.65	11.25	0.00	4

Table 3.6: Key Largo woodrat abundance estimates using the four different methods outlined in Section 3.2.4 to convert the population size in the covered area (i.e., effective area trapped by each grid) to a total population size estimate for the entire survey region (850 ha of habitat). MMDM is mean maximum distance used.

Session	Naïve	Home range	MMDM	Radio tracking
1	1291.95	360.54	394.15	1033.56
2	1982.92	553.37	927.08	1586.34
3	3203.18	839.91	1340.86	2562.55

### 3.4 Conclusion

The weakest aspect of using a capture-recapture framework is estimating the effective trapping area of each grid, such that one can extrapolate the abundance estimate in the covered region, to that of the entire survey region. I used four commonly implemented methods: naïve density estimate, half home range radius, mean maximum distance moved, and estimating proportion of woodrats with nests on the grid, and results varied considerably (Table 3.6). So what was the population size of Key Largo woodrats in 2007? The naïve population estimate is upwardly biased: since animals move on and off the grid, the effective area trapped by the grid is larger than the grid itself. Assuming these two areas are equal will cause population estimates to be upwardly biased. Therefore, we can safely say that an estimate of 1291 woodrats in March - April of 2007 is too high (and similarly for the July - September and November - December estimates of 1982 and 3203 woodrats, respectively). Despite this method providing upwardly biased results, it is still commonly implemented and has been used to estimate abundance of the Key Largo woodrat in recent studies (e.g., McCleery, 2003).

Correcting the effective area trapped by the radius of the average home range

size is also problematic because estimating the home range size is not without error and variation (e.g., different woodrats may maintain different home range sizes) and there is no reason to believe half the home range size is actually the correct value to use (Wilson and Anderson, 1985a). In addition, I used an estimate of home range size from a study conducted in 2002 (McCleery, 2003). If woodrat density has changed since this study, home range size may also have changed and this estimate will be incorrect (and bias remains unknown).

The proposal by Wilson and Anderson (1985a) to use the mean maximum distance moved between recaptured individuals is commonly implemented and was found by Parmenter et al. (2003) to provide the best performance when compared with a known population size. However, despite their findings, Parmenter et al. (2003) caution against using correction factors based on capture-recapture data as estimates of MMDM are dependent on trap spacing (here, the maximum observable distance using a 60 m square grid is  $\approx 85$  m), animal home range size, and also biased with density (at high population densities, population estimates biased upwards and at low population densities, estimates biased downwards). Hence, the true level of bias in this estimate is unknown.

The last method I investigated used radio tracking to estimate the proportion of individuals on the trapping grid. As no concurrent radio tracking was undertaken during this trapping survey, the correction factor estimated by Winchester (2007) was used. This has problems. Firstly, this estimate was based on a different grid size (0.5 ha), which will cause an abundance estimate that is upwardly biased as the estimate of proportion of time spent on the grid is over-estimated. Secondly, this estimate is from a study conducted in 2006, which may now be out of date if e.g., home range size has changed (Royle and Young, 2008). Thirdly, there is a risk of positive bias because the sample of animals upon which radio-collars were attached and monitored will be biased towards those animals that

are inherently more catchable and which tend to spend more time closer to the centre of the grid, and hence have a higher than average probability of capture (Efford, 2004). Hence, it would seem this estimate is also upwardly biased, but to what extent, remains unknown.

In addition to considering the consequences that different effective trapping areas has on abundance estimates, we also need to consider model selection. Six of the top nine models fitted (with a  $\Delta\text{AIC}<2$ ) included a 2-point mixture process to incorporate individual heterogeneity (Table 3.5). Unmodelled heterogeneity in capture probability can cause over-confidence and negative bias in estimates of population size (Anderson et al., 1983; Seber, 1992; Link, 2003). Unmodelled sources of heterogeneity include pregnant and/or lactating females that are possibly exposed to traps less often while they remain inside nests caring for offspring (e.g., Smith, 1968). Also, male woodrats can be nomadic and wander large distances (e.g., distances up to 680 m between nest locations has been observed, Potts unpubl. data). Consequently, these individuals can have variable capture probabilities and the fact that model selection favoured models that included a 2-point mixture process emphasises that within these data, there is strong heterogeneity in the capture probability.

A final consideration is that capture-recapture studies with fixed trap locations have a spatial component: animals close to traps are more likely to be caught than those animals further away. This is not addressed in standard capture-recapture studies, and without the spatial component, rigorous estimates of density cannot be obtained (Borchers and Efford, 2008). In the next chapter, I look at incorporating this spatial component using a spatially-explicit capture-recapture analysis (Borchers and Efford, 2008; Efford et al., 2004).

## Chapter 4

# Estimating Key Largo woodrat abundance using a spatially-explicit capture-recapture approach

### 4.1 Introduction

In the previous chapter, we saw there was a considerable amount of heterogeneity in the capture probability of fitted models when using a standard capture-recapture analysis to estimate Key Largo woodrat abundance. Based on AIC model selection, six of the top nine models fitted included a 2-point mixture process to incorporate individual heterogeneity (see Table 3.5 on page 59). This was despite using numerous different model parameterisations and incorporating individual covariates, such as sex, to account for differences in capture probability.

Another cause of heterogeneity in capture probability not accounted for in the standard capture-recapture analysis of the previous chapter is the proximity of animals to traps. That is, animals close to traps are more likely to be caught than animals further away, and therefore, their capture probabilities differ. Without the incorporation of this spatial information, rigorous estimates

of density cannot be obtained as the effective trapping area of the grid remains unknown (e.g., Table 3.6, page 60; Borchers and Efford, 2008). In this chapter, I use the same data set analysed in the previous chapter and expand on the capture-recapture analysis by incorporating spatial information within a maximum likelihood estimation framework (Section 2.4; Efford et al., 2004; Borchers and Efford, 2008; Borchers, 2011). The data were collected by Dan Greene (the University of Georgia) as part of his Masters project (c.f. Section 3.1, Greene 2009).

In the capture-recapture analysis of the previous chapter, abundance within the area “effectively trapped” by the trapping grid was estimated, and extrapolated upwards to the entire survey area. Here, the overall aim of the analysis was to estimate the density of home range centers,  $\hat{D}$ , existing in the survey area using spatial information contained in the pattern of recapture events of individual animals, and also prior knowledge of where traps were placed. The estimated density of home range centers,  $\hat{D}$ , can of course be converted into an estimated abundance,  $\hat{N}$ , if the area  $A$  of the survey region is known ( $\hat{N} = \hat{D}.A$ ).

Home range centers can be summarised by the coordinates of a point, and I assume these are fixed for the duration of the study and closed to births, deaths, immigration and emigration. I model the distribution of home range centers as a homogeneous spatial Poisson process. I use maximum likelihood estimation to estimate the parameters of a “capture probability” model that models the observation process of detecting an animal at a trap, given the trap is placed some distance from an animal’s (unknown) home range center; and I also assume that traps “compete” for animals. Despite using single-catch traps to collect data (c.f. Section 3.2.2), I assume traps can catch at least one individual. For consequences of this assumption, please see the Discussion and Conclusion section at the end of this chapter. The capture probability model is analogous

to the “detection function” in distance sampling (c.f. Section 2.2). Using the capture probability model, estimation of  $\hat{D}$  follows using a Horvitz-Thompson-like estimator.

## 4.2 Methods

### 4.2.1 Survey design, trapping and handling

A detailed description of the survey design and data collection is provided in Chapter 3, Sections 3.2.1 and 3.2.2. To summarise briefly here, a stratified random design was used to proportionally allocate 32 trapping grids to three habitat strata (young, medium and old hammock, c.f. Section 1.2.2; a fourth stratum, disturbed/urban, was excluded from the analysis because it was too small to accommodate randomly located grids). The spatial location of the 32 grids are shown in Figure 3.1 (page 50) and were distributed across habitat strata as follows: Young hammock - 4 grids, Medium hammock - 12 grids, Old hammock - 16 grids. Each grid was a 7 by 7 array, with one single-catch Sherman trap placed at each trap station (i.e., 49 traps per grid) with a 10 m trap spacing.

Data were collected according to Pollock’s robust design (Pollock et al., 1990), with three primary sessions (March - April, July - September, and November - December, 2007) and four secondary (and consecutive) sessions. Traps were opened and baited with whole rolled oats in the late afternoon and checked the following morning within the first two hours after sunrise. All captured woodrats were double-marked with PIT and ear tags, and its sex and capture location was recorded.

### 4.2.2 Data analysis

#### Capture probability model

The capture probability model,  $p_s(d_k(\mathbf{X}))$ , models the probability of detecting an animal at trap  $k$  on occasion  $s$ , when the animal’s home range center is  $\mathbf{X}$  (as

defined by the coordinates of a point). Three forms of the capture probability model are available: the half-normal, the exponential and the hazard rate (see Table 4.1 for model formulae). Following Efford et al. (2009b), I first investigated the form of the capture probability model (i.e., null models were fitted whereby the parameters of the capture probability models were not dependent on any covariates) and then covariates were added using forward-stepwise model selection based on Akaike's Information Criterion, AIC. When the difference in AIC between two competing models was not large enough to allow clear model selection (i.e.,  $\Delta\text{AIC} < 2$ , Burnham and Anderson, 2003), adjacent models were investigated. The covariates considered in model selection are described in Table 4.2. Since the data were naturally sparse (i.e., few captures across the 32 grids, despite a large survey effort), grid could not be used as a coefficient in the model, and consequently the probability of capture at distance zero ( $g_0$ ), and also the shape parameter  $\sigma$  were assumed equal across all grids. Sparse data cause imprecise estimates, with precision improving once approximately 20 recapture events occur (Efford et al., 2009b). Not all models could be fitted with all combinations of covariates, as the number of recapture events was small and this caused parameter identifiability issues (Gimenez et al., 2004).

Table 4.1: Three forms of the capture probability model. The probability of capturing an animal if the trap is placed 0 m from its home range centre is given by  $g_0$ . The shape of the capture probability model is given by  $\sigma$ .  $d$  is the distance between an animals home range centre and a detector.  $\theta$  is a vector of parameters containing  $g_0$ ,  $\sigma$  and  $b$  (if the hazard rate model is specified), for which maximum likelihood estimates are sought.

Detection function	Equation	Parameters in vector $\theta$
Halfnormal	$g_0 \cdot \exp(\frac{-d^2}{2\sigma^2})$	$g_0, \sigma$
Exponential	$g_0 \cdot \exp(\frac{-d}{\sigma})$	$g_0, \sigma$
Hazard rate	$g_0 \cdot [1 - \exp\{-(\frac{-d}{\sigma})^{-b}\}]$	$g_0, \sigma, b$

Table 4.2: Covariates considered during model selection for fitting the “capture probability” model (Efford, 2010).

Variable	Description	Notes
session	factor	three sessions were conducted throughout the year, each in a different season
t	factor	time, one level for each occasion
sex	individual level covariate	sex of the individual captured (male or female)
hab	individual level covariate	habitat type of where individual was captured
b	learned trap response	step change in the parameter after first detection of animal
B	transient trap response	parameter depends on detection at previous occasion (Markovian response)
h2	2-class mixture	finite mixture model with 2 latent classes (i.e., each variable was estimated using the mixture of two distributions)

### Estimating density

Since a conditional likelihood model was specified, whereby density of home range centers were assumed to be distributed in space by a homogeneous Poisson process,  $\hat{D}$  is estimated using a Horvitz-Thompson-like estimator (Borchers and Efford, 2008):

$$\hat{D} = \sum_{i=1}^n \frac{1}{\hat{a}(\hat{\theta}, z_i)} \quad (4.1)$$

where  $\hat{a}(\hat{\theta}, z_i)$  is the probability of detecting an animal  $i$  (for  $i = 1, \dots, n$ ) given the parameters of the capture probability model defined in  $\hat{\theta}$  (e.g.,  $g_0$  and  $\sigma$  for a halfnormal detection function, Table 4.1) and a set of individual-based observed covariates  $z_i$ , regardless of its home range center (i.e.,  $\int p(\mathbf{X})d\mathbf{X}$ , where  $p(\mathbf{X})$  is the probability of detecting an animal on any occasion in any detector). Density was estimated as a model averaged estimate across all fitted models proportional to AIC weight.

### Habitat mask

A habitat mask must be specified to bound the integration in calculating  $\hat{a}(\hat{\theta}, z_i)$  (i.e.,  $\int p(\mathbf{X})d\mathbf{X}$ ). The habitat mask is essentially a set of points that define a region around the traps from which animals may be detected.  $p(\mathbf{X})$  is evaluated for each point  $\mathbf{X}$  in the habitat mask and summed together (Efford et al., 2009b). Specifying a habitat mask is also important to define “habitat” from “non-habitat”, as treating habitat as non-habitat can cause density to be underestimated because same number of animals are estimated to be present, but in a larger area.

I created a habitat mask using the boundary between known habitat (i.e., hammock) and non-habitat (i.e., mangrove swamp), available as a ESRI shapefile (<http://www.esri.com/>). Thirty-two thousand points were placed within the

habitat mask to ensure the integration over  $\mathbf{X}$  was at a fine scale (Figure 4.1).

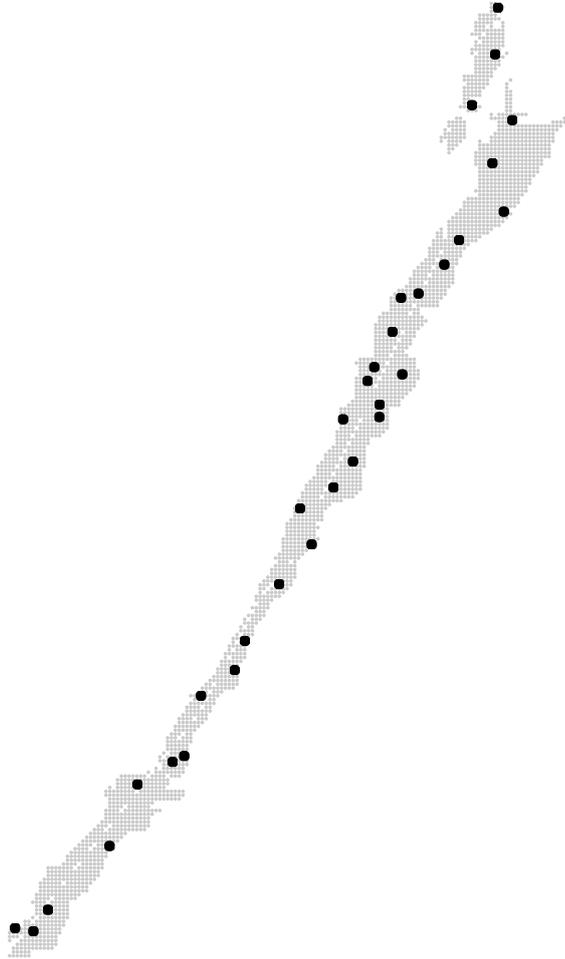


Figure 4.1: A map of the 32 trapping grid locations (solid black dot), overlaid on the habitat mask (shaded grey) used in the spatially-explicit capture-recapture analysis.

## Software

All modelling was undertaken using the ‘secr’ package (version 1.4.0, Efford, 2010) in statistical program R (version 2.11.1).

## 4.3 Results

### 4.3.1 Trapping and handling

Please see Table 3.4 on page 57 for a summary of the woodrat capture information.

### 4.3.2 Data analysis

Of the three forms of detection function available (half-normal, exponential, and hazard-rate), the exponential had the lowest AIC and was used as the form for all subsequent analyses (Table 4.3). Twenty-three further models were fitted that incorporated covariate information in to the capture probability model (Table 4.3). The model with the lowest AIC modelled  $g_0$  (the probability of detection if a trap was placed on a home range center) as a single intercept (i.e.,  $g_0 \sim 0$ ), and the shape parameter,  $\sigma$ , varied by session and habitat (i.e.,  $\sigma \sim \text{session} + \text{hab}$ ). Estimates of density derived using the model with the lowest AIC for the three trapping sessions were 0.05 (se= 0.02), 0.25(se=0.07) and 0.45 (se=0.09) woodrats per hectare, in trapping sessions 1 (March - April), 2 (July - September) and 3 (November - December), respectively. Based on a total survey area of 850 ha, the estimated abundance and 95% confidence intervals for the first, second and third sampling sessions were: 43 (31, 116), 214 (140, 391), 382 (243, 543), respectively. The density estimates for the three trapping sessions are shown in Figure 4.2.

Table 4.3: Twenty-six models were fitted to the data based on AIC, that incorporated behavior (b and B), session, time (t), sex and habitat (hab) covariates (see Table 4.2 for complete covariate definitions). All models were fitted using the exponential form of the capture function, unless otherwise stated.

	model	npar	logLik	AIC	AICc	$\Delta$ AICc	AICwt
1	$g0 \sim 1, \sigma \sim \text{session} + \text{hab}$	6	-743.52	1499.04	1500.49	0	0.55
2	$g0 \sim B, \sigma \sim \text{session} + \text{hab}$	7	-743.02	1500.04	1502.01	1.52	0.26
3	$g0 \sim b, \sigma \sim \text{session} + \text{hab}$	7	-743.4	1500.81	1502.78	2.29	0.17
4	$g0 \sim 0, \sigma \sim \text{session} + \text{sex}$	7	-749.39	1508.78	1509.8	9.31	0.005
5	$g0 \sim 1, \sigma \sim \text{session}$	4	-750.67	1509.33	1510	9.51	0.005
6	$g0 \sim 1, \sigma \sim \text{session} + B$	5	-750.44	1510.88	1511.9	11.41	0
7	$g0 \sim B, \sigma \sim \text{session}$	5	-750.52	1511.04	1512.06	11.57	0
8	$g0 \sim h2, \sigma \sim 1$	4	-751.74	1511.49	1512.16	11.67	0
9	$g0 \sim b, \sigma \sim \text{session}$	5	-750.66	1511.31	1512.33	11.84	0
10	$g0 \sim 1, \sigma \sim \text{session} + b$	5	-750.66	1511.32	1512.33	11.84	0
11	$g0 \sim 1, \sigma \sim h2$	4	-751.95	1511.9	1512.57	12.08	0
12	$g0 \sim t, \sigma \sim \text{session}$	7	-749.27	1512.54	1514.51	14.02	0
13	$g0 \sim 1, \sigma \sim \text{session} + t$	7	-750.37	1514.73	1516.7	16.21	0
14	$g0 \sim \text{sex}, \sigma \sim 1$	3	-755.28	1516.57	1516.96	16.47	0
15	$g0 \sim \text{session}, \sigma \sim 1$	4	-754.44	1516.87	1517.54	17.05	0
16	$g0 \sim 1, \sigma \sim \text{hab}$	4	-755.48	1518.97	1519.63	19.14	0
17	$g0 \sim 1, \sigma \sim 1$	2	-757.93	1519.87	1520.06	19.57	0
18	$g0 \sim 1, \sigma \sim \text{sex}$	3	-757.45	1520.89	1521.29	20.8	0
19	$g0 \sim B, \sigma \sim 1$	3	-757.65	1521.31	1521.7	21.21	0
20	$g0 \sim 1, \sigma \sim b$	3	-757.78	1521.56	1521.96	21.47	0
21	$g0 \sim 1, \sigma \sim B$	3	-757.9	1521.79	1522.19	21.7	0
22	$g0 \sim b, \sigma \sim 1$	3	-757.93	1521.86	1522.25	21.76	0
23	$g0 \sim t, \sigma \sim 1$	5	-755.79	1521.57	1522.59	22.1	0
24	$g0 \sim 1, \sigma \sim t$	5	-757.34	1524.69	1525.71	25.22	0
halfnormal	$g0 \sim 1, \sigma \sim 1$	2	-761.16	1526.32	1526.51	26.02	0
hazard	$g0 \sim 1, \sigma \sim 1$	3	-761.91	1529.83	1530.22	29.73	0

## 4.4 Discussion and Conclusion

Density (and abundance) of woodrats showed a clear and increasing trend over the course of the data collection period (2007) using the spatially-explicit capture-recapture approach demonstrated in this chapter. This is similar to the trend observed when data were analysed using standard capture-recapture (c.f. previous chapter). Overall abundance estimates using the SECR analysis were lower than the four methods (naïve, home range radius, mean maximum distance moved and radio tracking) used in the standard capture-recapture analysis, and were closest to the home range estimation method.

To assess reliability of the SECR method, we need to consider variance about, and bias in, the density estimates. In order to assess bias, we need to consider the assumptions of the model, how well they are met, and the likely effect of any violations. A critical assumption of this approach is that animals occupy home ranges. This method cannot be assumed to work when a high proportion of the individuals are nomadic, and its robustness in this circumstance has to be investigated (Efford, 2004). Woodrats tend to occupy home ranges. Home range size may vary by sex and season (McCleery, 2003). Some woodrats, in particular males, seem to be nomadic, searching far distances probably looking for female woodrats to mate with. Even these woodrats have specific nest locations they maintain over large distances (e.g., distances  $>600$  m between nest locations maintained by the same male woodrat have been observed in the field, Potts, unpubl. data). Any consequence that this individual variation might have in violating the circular home-range assumption will affect precision, rather than cause bias, in estimates of density. Similarly, violating other assumptions (clumped distribution of animals, and individual variation in  $g(0)$  and  $\sigma$ ) influences precision, rather than bias, of estimates (Efford, 2004).

This method assumes the home range centers of animals are distributed in the

survey area according to a homogeneous Poisson process. In reality, animals do not distribute themselves in space according to a homogeneous Poisson process, however modelling underlying density in this manner may be reasonable over small areas, particularly when multiple, randomly located sets of traps are used to infer density over a larger area (e.g., the survey design used for the Key Largo woodrat; Efford et al., 2005; Borchers and Efford, 2008).

In this analysis, I assumed that traps could catch at least one individual. Single-catch traps (e.g., the Sherman traps used in this study) are able to catch only one animal at a time, and the capture probability is affected by the presence of other animals that may compete for traps. Capture of an animal disables a trap and immediately reduces the capture probabilities of neighbouring animals. This assumption is clearly violated by the method of trapping used in this study, but any bias is negligible unless trap saturation is very high ( $> 86\%$ , Efford et al.

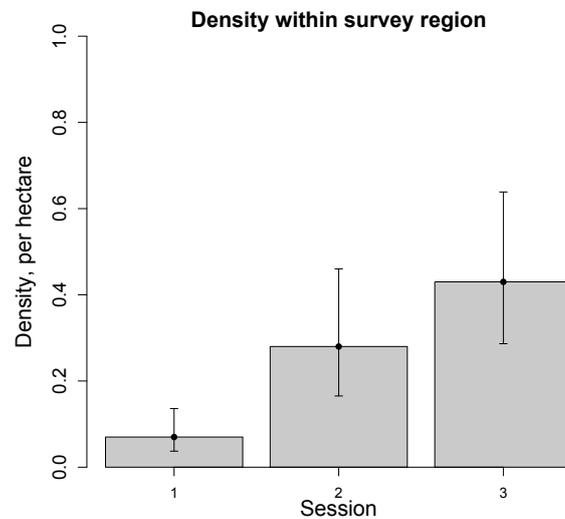


Figure 4.2: The estimated density of woodrats per hectare in the survey region, shown as solid lines, and 95% confidence intervals (dashed lines).

2009b). Since the woodrat is at extremely low densities, levels of trap saturation at which density estimates will be biased will never be attained. A likelihood model for single-catch traps is considerably more complicated than for multi-catch traps and remains to be developed (Efford et al., 2009b).

A disadvantage of using SECR to estimate abundance is the large trapping effort required to collect data. The data presented in this chapter were collected by two full-time field staff over the course of a year (with total survey effort being approximately seven months), and cannot realistically form the basis of an on-going monitoring program of Key Largo woodrat abundance. Since SECR is agnostic to trap configuration, it may be possible to decrease the cost by changing the trap configuration. Using ‘hollow’ grids (where only a circle or square of traps is set, with no traps at the center of the grid) may achieve this, as field time will decrease since fewer traps are set, and checking time can be more efficient as field assistants are not required to back-track as much. However, setting fewer traps will naturally decrease the number of recapture events. Instead, in the next two chapters, I begin to investigate distance sampling methods that incorporate trapping data, and apply these new methodologies to estimate the Key Largo woodrat abundance in Chapter 7.

## Chapter 5

# Estimating abundance using trapping point transects - methodology

### 5.1 Introduction

Capture-recapture is probably the most common monitoring method applied in small mammal surveys (Pollock et al., 1990; Solari et al., 2002). However, robust estimation of population size using capture-recapture is difficult because an estimate of the effective trapping area of the survey is required (Efford, 2004). For example, we saw in Chapter 3 how different methods to estimate the effective trapping area gave rise to considerable discrepancy in estimates of Key Largo Woodrat population size.

An alternative to capture-recapture is distance sampling, where density is estimated directly from the number of animals detected and the average probability of detecting those animals (Buckland et al., 2001). Consequently, a robust estimate of population size can be made. However, when detectability of animals is low (as is the case for most small mammals when detections are made visually), the assumption that animals at distance zero from the transect line or point are always detected is often violated. In addition, with low detectability, a large amount of survey effort is required to obtain enough visual detections to

model reliably the detection function (i.e., the probability of detecting an animal as a function of distance, and possibly other covariates, e.g., sex). For these reasons, distance sampling based on visual detection is not commonly implemented in small mammal surveys.

Research has consequently focused on methods that combine trapping studies with distance sampling theory, leading to the development of trapping webs (see Section 2.5 for discussion) and more recently, spatially-explicit capture-recapture, SECR (Borchers and Efford, 2008). In Chapter 4, I demonstrated the SECR approach to estimating abundance of the Key Largo woodrat. In this Chapter, I explore an alternative approach: trapping point transects (Buckland et al., 2006).

In the trapping point transect approach, two surveys are conducted. Data collected during the *trial survey* are used to estimate a ‘detection function’. Then for each animal detected during a separate *main survey*, its probability of detection is calculated using the detection function, and a Horvitz-Thompson-like estimator is used to estimate overall abundance. An assumption of the method is that the detection function estimated during the trial survey applies to animals in the main survey. The best way to achieve this is if the main and trial surveys are held concurrently in space and time, and use the same field methods. For the remainder of this chapter, I present the survey design and methods to analyse data collected using the trapping point transect approach, and I conclude the chapter with a discussion of the assumptions made.

## 5.2 Survey design

### 5.2.1 Main survey

In the main survey, sample locations are selected according to some randomized scheme, e.g., simple random sampling, or a systematic sampling grid with a

random start location. If systematic sampling is used, the distance between neighbouring sample points on the grid should be large enough so that one can assume the probability of detection at a given location is not influenced by detection events at other locations. I assume enough sampling locations to allow reliable extrapolation to the broader survey area.

It is not necessary to individually mark animals in the main survey, but recording covariate information (e.g., sex) of detected animals may be useful at the analysis stage. Various field methods can be used to detect individuals at each sample location, e.g., an observer might stand at each sample location for a fixed period of time and visually detect individual animals, or a trap may be set for a fixed period of time that physically captures (i.e., detects) individuals.

### **5.2.2 Trials survey**

Various field methods can be used to undertake the trial survey, although the concept behind each method is the same: a known-location animal is acquired, and at some pre-determined distance and random direction away from the known-location animal, a ‘trap’ is set. After a set period of time, the ‘trap’ is checked to see if the known-location animal was captured or not (Figure 5.1). Different methods can be used to obtain the known-location animals, e.g., visual detection, or radio tracking; and different methods of ‘trapping’ can also be used, e.g., physically capturing an animal in a trap, or placing an observer at the trial trap location to visually detect the animal. Most importantly, the method chosen in this trial survey should be the same as the main survey to ensure the estimated detection function is consistent and applicable; and trials should be conducted on a representative sample of animals from the population.

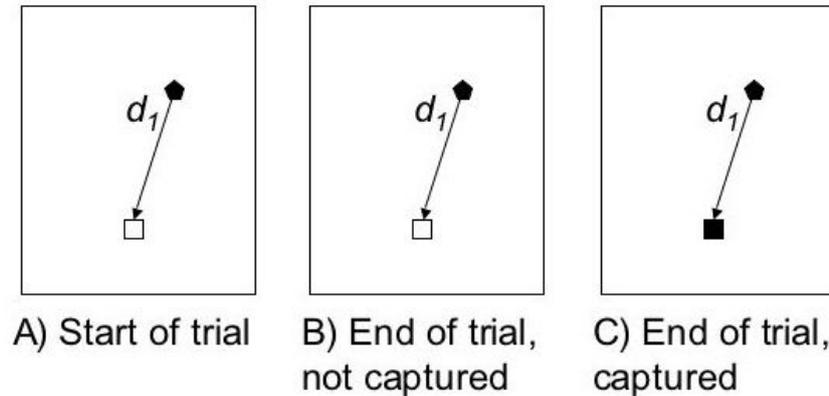


Figure 5.1: Schematic of how one individual trial survey is conducted in the field. Panel A) To start the trial, a known-location animal is located (black hexagon). At some pre-determined distance,  $d_1$ , and random direction away from the animal, a ‘trap’ is placed (white square). A response to the trial is observed after a set period of time, during which e.g., a second observer plays a lure to attract animals to the location for visual detection, or a trap is set and the animal has sufficient time to encounter the trap and be captured. If the known-location animal was not captured (Panel B, white square) a “0” is recorded, or if the animal was captured (Panel C, black square) a “1” is recorded. This process is repeated for various trial distances, and for numerous known-location animals, thus ensuring enough binomial response data is collected for the detection function to be estimated. (see Section 5.3.1).

In Buckland et al. (2006), flocks of Scottish crossbill (*Loxia scotica*) were located by the first observer, and at some pre-determined distance away from the known-location flock the second observer initiated a lure for a fixed period of time (a tape of excitement calls). Whether the known-location flock responded to the lure (i.e., was detected by the second observer) was the result of one trial, and many trials were conducted at various distances and random directions on many different flocks of crossbills.

Alternatively, the known-location animal can be acquired through radio tracking (i.e., by attaching a radio-collar to the animal during a previous capture event). A trial trap is established at a pre-determined distance and random

direction away from where the animal is located via radio telemetry, and after a set period of time (e.g., one day) the trial terminates. It is then discovered whether the known-location animal was captured at that trial distance.

It may be possible to acquire known-location animals using their point of release. That is, if an animal is captured in a trap, upon releasing the animal a second trap is set some pre-determined distance away from the animal's initial capture point. After some set period of time, it is recorded whether the animal is captured in the second trap. (See the 'Further Research' section on page 179.)

## 5.3 Modelling of data

### 5.3.1 Fitting the detection function

A detection function,  $g(r|z)$ , specifying the probability of detecting an animal at distance  $r$  from a trap location given a set of covariates  $z$ , can be estimated using the binomial response data collected from the trial surveys. The modelling technique used to estimate the detection function,  $\hat{g}(r|z)$ , changes depending on the method used to locate the known-location individual (e.g., detecting individuals visually versus radio-tracking). If known-location animals are detected from a visual search, it is likely that only one trial will be conducted per individual, because: 1) it is unlikely the animal will be uniquely identifiable from the distance it is observed (e.g., unique pelage patterns are rare in small mammals), and 2) the search effort required to re-locate the same individual so repeat trial surveys can be conducted on it will be resource intensive. A single trial conducted per individual is the simplest approach to conducting the trial survey, in which the binomial response variable,  $y_i$  (captured, or not), can be modelled as a logistic regression including individual-based covariates, and the general detection function is given by:

$$g(r|z) = \frac{\exp\left(\alpha + \beta_0 r + \sum_{j=1}^J \beta_j z_j\right)}{1 + \exp\left(\alpha + \beta_0 r + \sum_{j=1}^J \beta_j z_j\right)} \quad (5.1)$$

where  $\alpha$  is the intercept,  $\beta_0$  is the coefficient (slope) for the explanatory variable distance ( $r$ ), and  $\beta_j$  are the coefficients for each of the  $J$  explanatory variables ( $z_j$ ).

However, in some instances repeat trial surveys can be conducted on the same individual. For example, if radio-transmitter collars are attached to animals and individuals are located via radio-telemetry, multiple trial surveys can be conducted on each radio-collared animal for the duration in which the collar is attached and operational. In this instance, trials cannot be considered independent. One possible framework for accounting for this is a generalized linear mixed effects model with observations grouped on individual animal (Brown and Prescott, 2004), where the binomial response variable  $y_i$  (whether or not individual  $i$  is captured or not) is modelled as:

$$E(y_i) = p_i \quad \text{say} \quad = \frac{\exp\left(\alpha + \beta_0 r_i + \sum_{j=1}^J \beta_j z_{ij} + b_i\right)}{1 + \exp\left(\alpha + \beta_0 r_i + \sum_{j=1}^J \beta_j z_{ij} + b_i\right)} \quad (5.2)$$

where  $\alpha$  is the intercept,  $\beta_0$  is the coefficient (slope) for the explanatory variable distance ( $r$ ),  $\beta_j$  are the coefficients for each of the  $J$  explanatory variables of individual  $i$  ( $z_{ij}$ ), and  $b_i$  are the random effects grouped on individual animal that are assumed Normally distributed with mean 0 and variance  $\sigma_b^2$ . It follows that, in general, the detection function is:

$$g(r|z, b) = \frac{\exp\left(\alpha + \beta_0 r + \sum_{j=1}^J \beta_j z_j + b\right)}{1 + \exp\left(\alpha + \beta_0 r + \sum_{j=1}^J \beta_j z_j + b\right)} \quad (5.3)$$

Generalised additive mixed models is an alternative framework that could be used to model the detection function. Regardless of which form of the detection function that is fit (Equation 5.1 or 5.3), some type of model selection process (e.g., based on Akaike's Information Criterion, AIC, Burnham and Anderson, 2003) should be conducted to ensure the fitted detection function is parsimonious.

### 5.3.2 Estimating abundance

In the case where one trial is conducted per individual, standard equations from distance sampling theory can be used to estimate abundance (Buckland et al., 2001). By using the fitted detection function, the probability of detecting each animal captured in the main survey given that it is within distance  $w$  of the sample point ( $\hat{p}_i$ ) can be calculated (Buckland et al., 2006). To calculate  $\hat{p}_i$  for each animal detected in the main survey, observed values of each explanatory variable  $z_{ij}$  can be substituted into the fitted model (Equation 5.1), but the distance,  $r_i$ , remains unknown. Consequently,  $\hat{p}_i$  is a function of the unknown  $r_i$ :  $\hat{p}_i = \hat{p}(r; z_{i1}, \dots, z_{iJ})$  for  $0 \leq r \leq w$  where  $w$  is the truncation distance (i.e., where the probability of detecting an animal at distance  $w$  is very close to 0). Therefore, the estimated probability of detecting the  $i$ th animal captured in the main survey unconditional of its distance from the trapping sample point,  $\hat{P}(z_{i1}, \dots, z_{iJ})$ , can be estimated by integrating over the unknown  $r$ :

$$\hat{P}(z_{i1}, \dots, z_{iJ}) = \int_0^w \pi(r) \hat{p}(r; z_{i1}, \dots, z_{iJ}) dr \quad (5.4)$$

where  $\pi(r)$  is the probability density function of distances of individuals from the trapping sample point (i.e., the availability function). In point transect sampling, it is critical that the points be placed randomly with respect to the distribution of animals so that  $\pi(r) = 2r/w^2$  (Buckland et al., 2001), and Equation 5.4 becomes:

$$\hat{P}(z_{i1}, \dots, z_{iJ}) = \frac{2}{w^2} \int_0^w rg(r|z) dr \quad (5.5)$$

Since the value of  $\hat{P}(z_{i1}, \dots, z_{iJ})$  is dependent on  $w$  (i.e., as  $w$  increases to infinity,  $\hat{P}(z_{i1}, \dots, z_{iJ}) \rightarrow 0$ ) it might be easier conceptually to use the effective area of detection as a measure of detectability. The effective area of detectability,  $\nu$ , is the area for which as many animals are detected outwith this area as are missed within the area, and is estimated using:

$$\hat{\nu} = \frac{2\pi g(0)}{h(0)} \quad (5.6)$$

where  $g(0)$  is the probability of detecting an animal at zero distance, and  $h(0)$  is the derivative of the probability density function,  $f(r)$ , evaluated at distance zero:  $f(r) = rg(r) / \int_0^{w=\infty} rg(r)dr$ , such that  $\int_0^{w=\infty} f(r)dr = 1$ . All of what follows can be expressed in terms of  $\nu$  or  $\hat{P}(z_{i1}, \dots, z_{iJ})$ , but I have chosen  $\hat{P}(z_{i1}, \dots, z_{iJ})$  for consistency.

Once  $\hat{P}(z_{i1}, \dots, z_{iJ})$  has been estimated for each animal  $i = 1, \dots, n$  detected in the main survey, a Horvitz-Thompson-like estimator can be used to estimate abundance in the survey region,  $\hat{N}$  (Borchers et al., 1998):

$$\hat{N} = \frac{A}{A_c} \sum_{i=1}^n \frac{1}{\hat{P}(z_{i1}, \dots, z_{iJ})} \quad (5.7)$$

where  $A$  is the area of the survey region, and  $A_c$  is the area of the covered region ( $A_c = k\pi w^2$ , where  $k$  is the number of sample points in the main survey).

In the case where multiple trials are conducted on the same individual,  $\hat{p}_i$  is now a function of unknown distance,  $r_i$ , and a random effect  $b_i$ :  $\hat{p}_i = \hat{p}(r, b; z_{i1}, \dots, z_{iJ})$  for  $0 \leq r \leq w$ , and Equation 5.4 must be modified to include the random effect term:

$$\hat{P}(z_{i1}, \dots, z_{iJ}) = E_r[\hat{p}(r, b; z_{i1}, \dots, z_{iJ})] = \int_0^w \pi(r) \cdot \hat{p}(r, b; z_{i1}, \dots, z_{iJ}) dr \quad (5.8)$$

To estimate abundance in the main survey, Equation 5.7 must be modified, since  $\hat{P}(z_{i1}, \dots, z_{iJ})$  now includes an unknown random effect variance which must be integrated out. This integration can be done by simulation (e.g., take 10,000 samples of  $\hat{P}(z_{i1}, \dots, z_{iJ})$  from the distribution of  $b_i$  for each individual  $i$  detected in the main survey, thus implicitly including uncertainty in the estimate of the variance of the random effect) and two abundance estimators are available:

$$\hat{N}_2 = \frac{A}{A_c} \sum_{i=1}^n \frac{1}{E_b[\hat{P}(z_{i1}, \dots, z_{iJ})]} = \frac{A}{A_c} \sum_{i=1}^n \frac{1}{\int_0^w \pi(b) \hat{P}(z_{i1}, \dots, z_{iJ}) db} \quad (5.9)$$

$$\hat{N}_3 = \frac{A}{A_c} \sum_{i=1}^n E_b \left[ \frac{1}{\hat{P}(z_{i1}, \dots, z_{iJ})} \right] = \frac{A}{A_c} \sum_{i=1}^n \int_0^w \frac{1}{\pi(b) \hat{P}(z_{i1}, \dots, z_{iJ})} db \quad (5.10)$$

where  $\pi(b)$  is the probability density function of the random effect distribution, taken to be Normally distributed with mean 0 and variance  $\sigma_b^2$ .

### 5.3.3 Variance estimation

Variance in the estimated abundance can be due to two sources. Firstly, there is variance associated with the encounter rate in the main survey, and secondly, variance associated with estimating the detection function. Buckland et al. (2006) present a bootstrap approach to estimate variance for the simpler case when the detection function is based on one trial survey per individual (Equation

5.1, i.e., trial surveys are assumed independent of each other). That is, a resample of the  $k$  main survey points along with their data (i.e., how many animals were detected at each point) is taken with replacement; and a resample of the individual trial surveys is taken from the total number of trial surveys conducted, also with replacement. For each resample of the trials survey, a new detection function is fitted via model selection, thus incorporating uncertainty due to model structure of the detection function. Using this ‘new’ detection function, abundance for the resampled data set is estimated using the methods presented in the previous section (i.e., within each resample of the main survey data, the probability of detecting each animal is calculated using the ‘new’ detection function, and a Horvitz-Thompson-like estimator used to estimate abundance for that particular resampled data set). By repeating this resampling process a large number of times (e.g., 999 times), the sample variance of the bootstrap estimates for  $\hat{N}$  can be obtained and an approximate confidence interval for  $\hat{N}$  estimated (Davison and Hinkley, 1997).

In the case where repeated trial surveys are conducted on the same individual, they can no longer be assumed independent of each other, and the bootstrap resampling must account for this ‘blocking’ structure. That is, when performing the bootstrap resample of the trials survey, the unit of resampling is not at the trial level, it is on the animal on which many trial surveys were performed. If the number of trials conducted on each animal varied, the total number of trials in each resample may vary (but the number of animals selected, with replacement, will remain constant for each bootstrap resample). A ‘new’ detection function is fitted to this resampled data set that incorporates a random effect grouped on each resampled individual (i.e., Equation 5.3). Abundance estimation methods follow as per the previous section.

Alternatively, if the above equations were reparameterised in terms of density

$\hat{D}$  (i.e.,  $\hat{D} = \hat{N}/A$ ) then the delta method can be used to estimate variance of  $\hat{D}$  as follows:

$$v\hat{a}r(\hat{D}) = \hat{D}^2 \left( \frac{v\hat{a}r(n)}{n^2} + \frac{v\hat{a}r(a\hat{P}_a)}{a\hat{P}_a^2} \right) \quad (5.11)$$

where  $n$  is the number of animals trapped in the main survey,  $a$  is the area surveyed,  $\hat{P}_a$  is the proportion of woodrats detected,  $v\hat{a}r(n)$  is the variance of the encounter rate and  $v\hat{a}r(a\hat{P}_a)$  is the variance of the proportion detected (multiplied by the survey area  $a$ ). A bootstrap can be used to estimate  $v\hat{a}r(a\hat{P}_a)$  (as above). The variance of the encounter rate,  $v\hat{a}r(n)$ , is equal to  $k^2t^2\text{var}(n)/kt$ ; where  $k$  and  $t$  is the number of points visited in the main survey, and the number of times each main survey point is visited, respectively (Fewster et al., 2009).

## 5.4 Discussion

To date, the only published study using the trapping point transect (TPT) approach relied on visual detection and lures to estimate abundance of the Scottish crossbills (Buckland et al., 2006). Since one trial was performed on each flock of crossbills, the first form of the detection function was used (Equation 5.1). Given that individual flocks of birds were not uniquely identifiable across periods of time, multiple trials may have been unknowingly performed on the same flock of birds (or in another instance, more than one individual animal). So long as the population is large and trial survey effort has been allocated throughout the survey region, repeat surveys on the same flocks of birds would rarely occur.

In this chapter, I have explored the application of the TPT method to trapping surveys, whereby detecting individual animals relies on passively trapping them, rather than an active search. Attaching radio-collars to animals is one method to obtain known-location animals for the trial surveys, however the

detection function must include a random effect grouped on individual. Alternatively, it may be possible to use the point-of-release of an animal to its point-of-recapture as the distance of each trial survey. (See ‘Further Research’ section on page 179 for more discussion on this matter.)

Unlike conventional distance sampling, this method does not require all animals at distance zero to be detected with certainty. This is because ‘true’ zeros are obtained in the data collection process, so that a binomial response variable detection function can be fitted (i.e., since the location of the animal is known, a ‘zero’ observation is the animal known to be present, and yet not detected). This is important because it is likely that even at short distances, detection of some small mammals is less than perfect. In this sense, there are parallels with double observer methods in distance sampling, where observers set up trials for one another (Buckland et al., 2004).

There are some assumptions of the TPT method that should be noted:

1. Similar to standard capture-recapture studies, all individuals in the population with the same covariate values are assumed to have equal probability of being caught (i.e., the catchability of animals in the trial and main survey are the same); and
2. The estimated detection function is assumed to be correct. For example, if a covariate is observable and known to influence detectability (e.g., sex, body size, pelage colour), the covariate should be modelled in the detection function. Bias in population estimates may arise when the covariate is not known, or not adequately represented in the subset of ‘known-location’ animals.

The first assumption can be met by good survey design, for example, by holding the trial and main survey concurrently. That is, one cannot conduct the

main survey and return in e.g., another season to conduct the trial survey, as detectability may have changed and would not be applicable to the main survey data. Also the same field method used in the main survey (e.g., visual detection, luring or trapping) to detect animals must be used in the trial survey and the duration over which the field method is used to detect animals must be constant for both the main and the trial survey.

The second assumption can be difficult to achieve if the underlying factor influencing detectability is not known (e.g., some animals are inherently more trappable than others). We can also never be sure that all relevant covariates influencing detectability have been included in the model. In such instances it is likely that the subset of animals on which trials are performed will contain more “trappable” animals than less “trappable” animals. Under certain scenarios, this may result in bias in the estimated population abundance and associated variance. We need to understand the consequences of violating these assumptions, and this will be investigated using a simulation study in the next chapter.

## Chapter 6

# A simulation study to investigate the performance of TPT under different scenarios

### 6.1 Introduction

Simulation studies are a useful way to assess how different sampling strategies and levels of survey effort influence variance and bias in estimates of abundance. In this chapter, I use a simulation study to evaluate the performance of the trapping point transect (TPT) method detailed in Chapter 5. Both sampling and modelling of the data were simulated to reflect how a TPT survey would be undertaken in the “real world”. A TPT survey requires traps to be set in a survey region at random locations (the *main survey*) containing a population of unknown abundance,  $N$ . In a separate *trials survey*, a number of “trials” are performed on a number of individual animals to estimate the detection function. For example, when animals obtained for the trials survey have a radio collar attached to them, repeat trials can be performed on the same individual across a number of days (i.e., one trial per individual per day).

In this simulation study, each animal in the main and trial survey is allocated a unique detection function, drawn randomly from an underlying “true” mean detection function with a specified random variation about the intercept term

(i.e., probability of detecting the animal if a trap is placed 0 m from its location). The underlying true detection function from which trial data are sampled is assumed to be unknown, and different detection function scenarios are investigated to determine how best to sample from this unknown detection function so it can be estimated reliably. Using data obtained from the trial and main simulated surveys, an estimate of abundance,  $\hat{N}$ , is calculated using methods presented in Chapter 5. I use three different detection function scenarios to investigate:

1. how a fixed amount of trial survey effort (360 trap nights) should be allocated and trial distances selected when any underlying heterogeneity in the population *is* accounted for;
2. how a fixed amount of trial survey effort (360 trap nights) should be allocated and trial distances selected when any underlying heterogeneity in the population *is not* accounted for;
3. how much total trial survey effort would be required to ensure estimates of abundance are unbiased;
4. how a behavioral response to the trapping experience (i.e., when animals become trap shy after first capture) influences estimates of abundance under the “optimal” method of selecting trial distances as identified in simulations 1 and 2; and
5. how using a generalised linear model to estimate the detection function, that does not take into account the repeated observations taken on the same individual (instead of a generalised linear mixed model), can influence estimates of abundance.

Parameters in the three detection function scenarios, and levels of sampling effort in the simulation studies were chosen to match what is realistic for the

Key Largo woodrat (c.f. Chapter 7), and other small mammal species that have relatively low detectability. I end this chapter with a set of general conclusions regarding choice of method to select trial distances and the sample sizes required to obtain (nearly) unbiased estimates of abundance.

## 6.2 Methods

### 6.2.1 Three detection function scenarios

I use three different detection function scenarios (“High”, “Medium”, and “Low”), from which a unique detection function for each individual in the main and trial surveys is sampled (Figure 6.1). All the detection functions specified were of similar form to Equation 5.2 on page 80, but as there were  $t$  multiple trials on each individual  $i$  that belonged to group  $g$ , the probability of observing a successful response (i.e., the radio collared animal was recaptured,  $y_{tig} = 1$ ), was:

$$\text{logit}(E[y_{tig}]) = \text{logit}(p_{tig}) = \alpha_g + \beta r_{tig} + b_{ig} \quad (6.1)$$

where  $\alpha_g$  is the intercept for an individual in group  $g$  ( $g = 1, 2$  or  $3$  for an individual in the high, medium or low group, respectively);  $\beta$  is the coefficient (slope) for the explanatory variable distance ( $r$ ); and  $b_{ig}$  is a random effect due to individual  $i$  in group  $g$  with an independent Normal distribution with mean 0 and variance  $\sigma_{bg}^2$ . The three scenarios are (Table 6.1):

**“High” detectability.** A single group of individuals with the same mean detection function ( $\alpha_1 = 2$ ,  $\beta = -0.15$ ; effective trapping radius,  $\rho = 17.67$  m), with a small individual random effect variance on the intercept of the detection function (i.e.,  $b_{ig} \sim N(0, 0.1)$ ,  $g = 1$ ; Figure 6.1A).

**“Medium” detectability.** Two underlying groups are present in the population, group 1 has high detectability (as previous,  $\rho = 17.67$  m), group 2

has medium detectability ( $\alpha_2 = 1$ ,  $\beta = -0.15$ ,  $\rho = 12.67$  m), and both groups have a small individual random effect variance on the intercept ( $b_{ig} \sim N(0, 0.1)$ ,  $g = 1, 2$ ; Figure 6.1B).

**“Low” detectability.** Two groups present in the population, but with a large difference in detectability (group 1: as previous,  $\rho = 17.67$  m; group 3:  $\alpha_3 = -0.3$ ,  $\beta = -0.15$ ,  $\rho = 7.51$  m). The random effect variance on the intercept for individuals in the high group ( $b_{ig} \sim N(0, 0.1)$ ,  $g = 1$ ) was smaller than individuals in the low group ( $b_{ig} \sim N(0, 0.3)$ ,  $g = 3$ ; Figure 6.1C).

The group covariate in the “Medium” and “Low” scenarios can be thought of as an observable covariate (e.g., sex), or an unobservable covariate (e.g., some animals are inherently more trappable and choose to go in traps and be detected, regardless of previous trap experience, others do not).

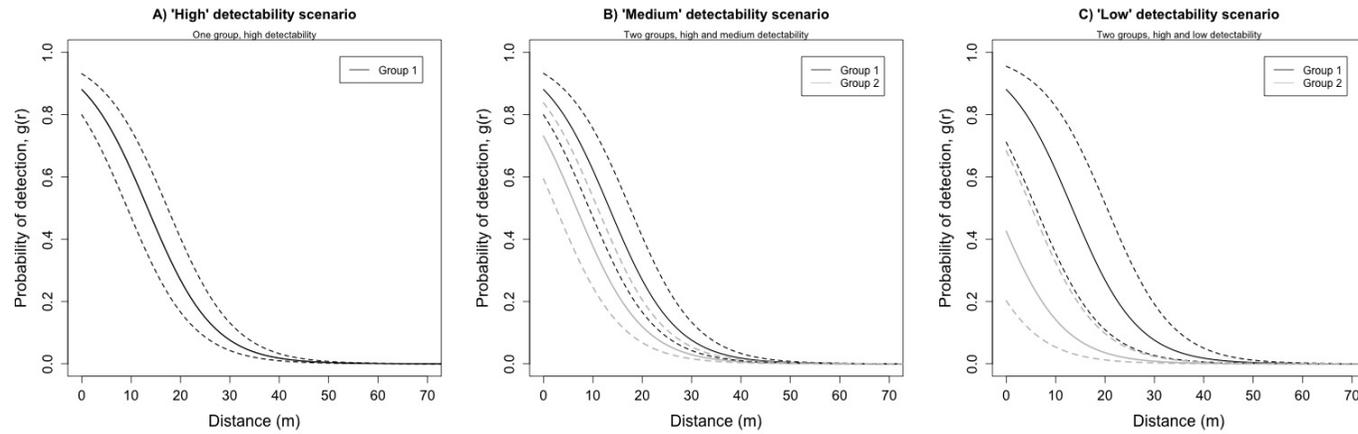


Figure 6.1: Plot of the three detection function scenarios used to simulate unique detection functions for each individual in the main and trial surveys. The solid line indicates the mean detection function, and the dashed lines indicate the 2.5th and 97.5th percentiles (as calculated based on 10,000 samples of the random effects distribution). Panel A) A single group of animals with high detectability. Panel B) Two groups of animals, the first with high detectability, the second with medium detectability, and both groups have a small common random variance on the intercept parameter. Panel C) Two groups of animals, the first with high detectability, the second with low detectability and both groups have a larger common random variance on the intercept parameter. See main text in Section 6.2.1 for model parameter values.

Table 6.1: Input parameters for the three detection function scenarios (“High”, “Medium”, “Low”), where  $\alpha_g$  is the intercept term for an individual in group  $g$  ( $g = 1, 2, 3$ ),  $\beta$  is the coefficient (slope) parameter for the explanatory variable distance  $r$ , and  $b_{ig}$  is a random effect due to individual  $i$  in group  $g$ , in equation 6.1.

Scenario	Group	$\alpha_g$	$\beta$	$b_{ig}$
High	1	2	-0.15	$N(0, 0.1)$
Medium	1	2	-0.15	$N(0, 0.1)$
	2	1	-0.15	$N(0, 0.1)$
Low	1	2	-0.15	$N(0, 0.1)$
	3	-0.3	-0.15	$N(0, 0.3)$

### 6.2.2 Simulating the main survey

For each simulation, a population of 2,000 individuals was randomly located on a rectangular survey area (dimensions: 3,000 by 4,300 m), giving a density of 1.5 individuals per hectare. A *main survey* trapping grid was overlaid onto the survey area with dimensions of 10 by 15 traps (in the X and Y dimension, respectively) with a 250 m trap spacing. If multiple groups of animals were present in the population (i.e., detection function scenarios “Medium” and “Low”), 50% of the individuals were randomly allocated to the “High” group, and the remainder were allocated to the other group (“Medium” or “Low”). Each individual in the main survey was allocated a unique two-parameter detection function with the parameters drawn from a random Normal distribution with a mean detection function specified by its group (i.e.,  $\alpha_1 = 2$ , and  $\beta = -0.15$  for an individual in group 1, the high detectability group) and a variance specified by a known random effect variance (e.g.,  $b_{i1} \sim N(0, 0.1)$  for an individual in the high group). The distance between each individual and its closest main survey trap location was calculated. Individuals were selected at random, and firstly it was checked whether its nearest trap was “open” (i.e., hadn’t caught another individual). If the closest trap was shut, that animal could not be caught, and a new animal was selected. If the trap was open, a Bernoulli trial was performed to determine

if the individual was detected, given its distance to the nearest trap and a probability specified by the individual's unique detection function parameters. The Bernoulli trial was only performed once per animal, because given the large trap spacing in the main survey, the probability of detecting the animal at any trap other than that nearest was essentially zero, regardless of the detection function scenario (Figure 6.1). The resulting simulated data from the main survey was the number of individuals detected in the main survey, and if appropriate, a record of group membership. In Figure 6.2, I have provided an example simulation of the main survey from each of the three detection function scenarios in Figure 6.1.

### 6.2.3 Selecting trial distances

Recall in Section 5.2.2 of Chapter 5, the trials survey is conducted by positioning a trial trap some pre-determined distance and random direction away from a known-location animal. After a set period of time (e.g., one night), the trap is checked to see if the known-location animal was captured or not. One trial is set per individual, per time period (e.g., night). The animals upon which trials are performed were selected in proportion to their expected frequency in the main survey data (i.e., highly detectable individuals were over-represented in the trials survey). Although this is a realistic scenario to what would occur when implementing the method in the field, it will lead to bias in population estimates when underlying between-group heterogeneity is ignored (see Simulation 2, page 114). Since the underlying true detection function is unknown to the field observer, different methods will perform differently under different detection function scenarios. Four methods for selecting trial distances were investigated:

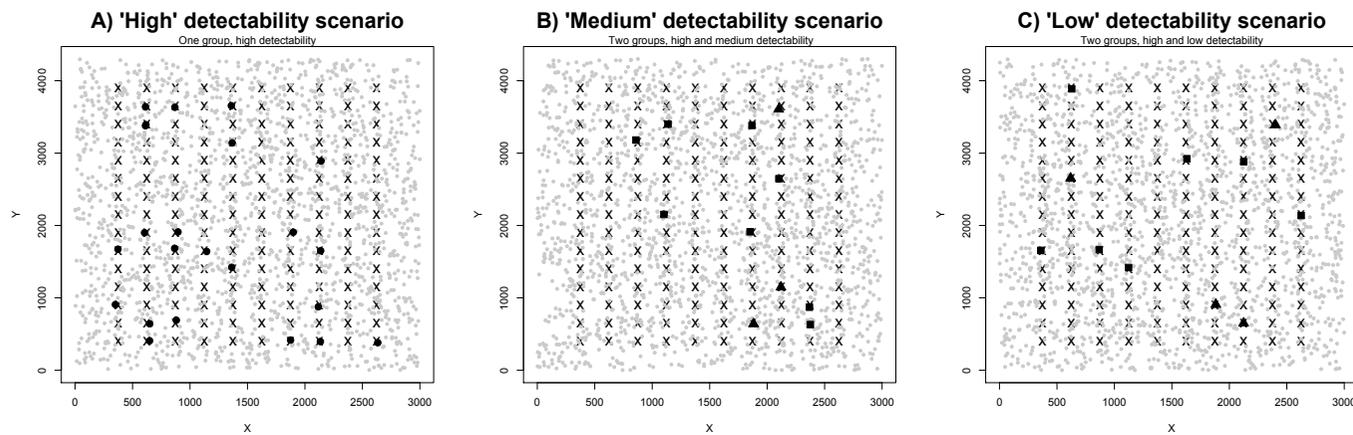


Figure 6.2: Plot of a single realisation of the main survey simulation, based on three detection function scenarios. The main survey was a 10 trap by 15 trap grid (trap location shown as an “x”), with a 250 m horizontal and vertical inter-trap spacing, located in a survey area of 3,000 m by 4,300 m. The total number of individuals in the main survey was 2,000 (equal ratio between groups of animals where applicable, i.e., scenarios “Medium” and “Low”). Panel A) All individuals had a single “High” probability detection function, and 22 individuals were detected (solid circle). Panel B) Individuals were separated into two groups, the first with “High” detectability (eight individuals detected, solid square), the second with “Medium” detectability (three individuals detected, solid triangle). Panel C) Individuals were separated into two groups, the first with “High” detectability (seven individuals detected, solid square), the second with “Low” detectability (four individuals detected, solid triangle). See Section 6.2.1 and Figure 6.1 for detectability group definition. Undetected individuals are shown as a small solid grey circle.

**Uniform:** Trial distances are selected uniformly between 1 m and  $w$  (the truncation distance, here 125 m). For example, if 40 trials are conducted per individual (i.e., one trial every night for 40 nights), the distance between each trial trap would be  $\approx 3.2$  m (Figure 6.3A). If the true underlying detection probability decreases rapidly with increasing distance from the trap, and the probability of detection is essentially zero at distances much shorter than  $w$  (i.e., trials at distances  $> 60$  m in Figure 6.3A), a large proportion of trials may be with traps placed in the “tail” of the detection function. This is an inefficient use of survey effort as most trials at these long distances will be unsuccessful, and the trial surveys will provide little information for estimating the detection function.

**Stopping Rule with 5 consecutive unsuccessful trials:** Trial distances are distributed as per the “Uniform” method, except that they start at 1 m, and continually increase in distance (e.g., a 3.2 m interval) until five consecutive trials were unsuccessful on the individual, or the distance  $w$  was reached. Trials on that individual then stop, and a new individual is selected (Figure 6.3B). The average number of trials per individual is 12, 11, and nine for the “High”, “Medium” and “Low” detection function scenarios, respectively<sup>1</sup>. Depending on the underlying true detection function, the method may perform badly if the stopping rule is too short to capture the tail of the detection function (i.e., not enough trials at longer distances). Also, the cost (effort) of the survey prior to starting can only be approximated, because it is not exactly known how many trials will be completed on each individual.

---

<sup>1</sup>Average number of trials per individual for each detection function scenario were calculated using simulation. Trials were performed on 10,000 individuals and the average number of trials per individual was calculated.

**Stopping Rule with 8 consecutive unsuccessful trials:** As per the previous method, however eight consecutive unsuccessful trials were required before the survey was stopped and a new individual was selected (Figure 6.3C). The average number of trials per individual is 15, 14, and 13 for the “High”, “Medium” and “Low” detection function scenarios, respectively. Similar issues to “Stopping Rule 5” occur (i.e., can only approximate cost of survey prior to starting, and the stopping rule used may be inappropriate for the true detection function).

**Adaptive:** Half the trial survey effort was allocated using the “Uniform” method, and a preliminary detection function was fitted (Figure 6.4A). The remaining survey effort was then allocated between 0 and  $w$  m, according to the cumulative distribution function (CDF) of detection distances, estimated using the data from the initial uniform trials. Values systematically distributed between 0 and 1 were back-transformed through the fitted CDF to generate trial distances (Figure 6.4B). This results in more trials at shorter distances, and fewer trials at distances in the tail of the detection function. The detection function used in the final abundance estimation is based on all trials conducted in both the uniform and adaptive phases of the trial survey. For example, if 40 trials are conducted per individual, 20 are allocated using the uniform method, and 20 are allocated based on the CDF of the detection distances, and trial results are combined to estimate the detection function (Figure 6.4C).

### 6.2.4 Estimating abundance

Using data obtained from the trial and main survey simulations, two abundance estimators were used as per methods detailed in Section 5.3.2 on page 81. They are:

$$\hat{N}_2 = \frac{A}{A_c} \sum_{i=1}^n \frac{1}{E_b[\hat{P}(g_i)]}$$

$$\hat{N}_3 = \frac{A}{A_c} \sum_{i=1}^n E_b \left[ \frac{1}{\hat{P}(g_i)} \right]$$

where  $A$  is the area of the survey region ( $=3,000 \times 4,300 \text{ m}^2$ ),  $A_c$  is the area of the covered region ( $= k\pi w^2$ , where  $k$  is the number of sample points in the main survey and  $w = 125$ ), and  $\hat{P}(g_i)$  is the estimated probability of detecting the  $i$ th animal captured in the main survey, given it is in group  $g$ . If the group covariate is unknown, then the detection function is assumed to depend only on distance, and no other covariates. Variance estimation followed using a bootstrap, as presented in Section 5.3.3.

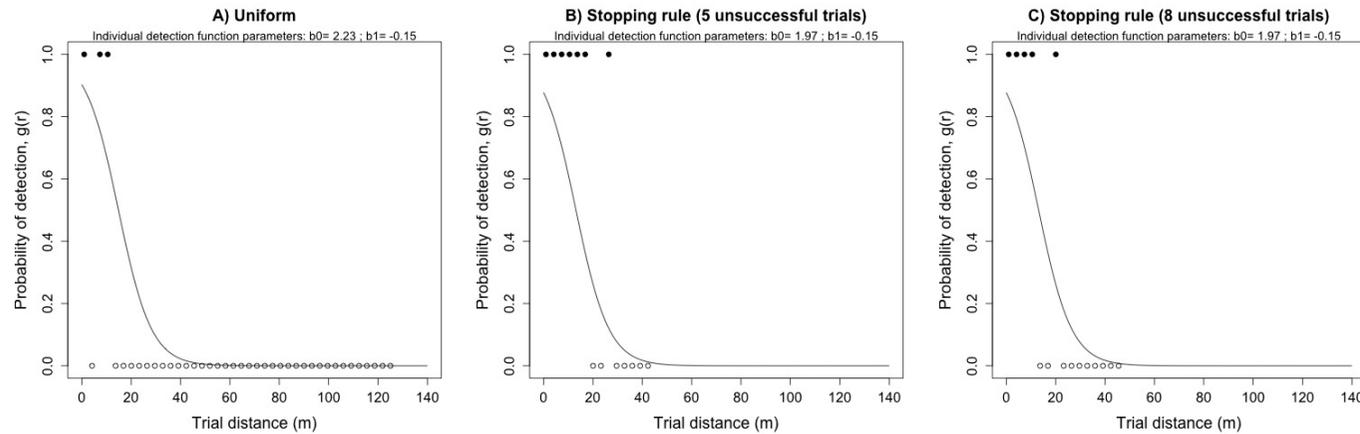


Figure 6.3: Plot of a single realisation of one set of simulated trial surveys (solid and hollow circles are successful and unsuccessful trials, respectively) for one individual, using the “Uniform” (Panel A), “Stopping Rule 5” (Panel B) and “Stopping Rule 8” (Panel C) method for selecting trial distances. The detection function was drawn from an underlying “true” detection function with parameters specified by the “High” detection function scenario. In this example, a maximum of 40 trials were performed per individual. Panel A) Trial distances were selected uniformly between 1 m and 125 m. Three were successful. Panel B) Trials started at distance 1 m, and increased until five consecutive trials were unsuccessful. In this instance, 14 trials were undertaken (seven were successful) before the survey ceased. On average, 12 trials are conducted per individual using this method, and for these detection function parameters. Panel C) As per B) but eight consecutive trials were required before the trial survey ceased. In this instance, 15 trials were undertaken (five were successful) before the survey ceased. On average, 15 trials are conducted per individual using this method, and for these detection function parameters.

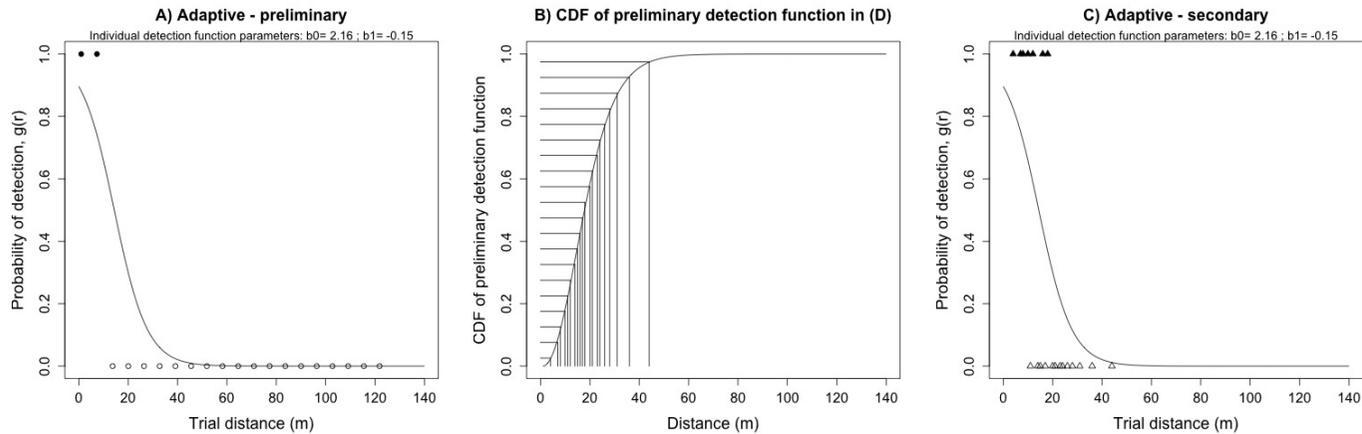


Figure 6.4: Plot of a single realisation of one set of simulated trial surveys for one animal, using the “Adaptive” method for selecting trial distances. The detection function was drawn from an underlying “true” detection function with parameters specified by the “High” detection function scenario, and a maximum of 40 trials were performed on the individual. Panel A) Half of the trial survey effort (i.e., here, 20 trials) were uniformly positioned between 1 m and  $w$ . Two trials (solid circles) were successful (18 trials, hollow circles, were unsuccessful). A preliminary detection function is fitted to the preliminary trial survey data (from multiple individuals) and a new set of trials is selected uniformly between 0 and 1, that were back-transformed through the fitted CDF of detection distances (Panel B). Most trials in the second phase of the survey are at shorter trial distances (Panel C, seven trials shown as solid triangles were successful, 13 were unsuccessful).

## 6.3 Simulation 1: Allocation of effort when underlying between-group heterogeneity in the population *is* accounted for

### 6.3.1 Methods

For a fixed level of survey effort, a trade-off exists between conducting a greater number of trials per individual, but on fewer individuals, or conducting fewer trials per individual, but on more individuals. I investigated this trade-off, as well as different methods for selecting trial distances, using a simulation study. I fixed the amount of trial survey effort to 360 trap nights, as this was believed to be a realistic amount of survey effort two people could undertake during the eight week field season to conduct a TPT survey on the Key Largo woodrat (Chapter 7). The number of trials per individual was varied: 6, 10, 18, 24, 40, and 60 (forcing the number of individuals in the survey to be 60, 36, 20, 15, 9 and 6, respectively)<sup>2</sup>. A simulation was completed for each of the three detection function scenarios, and the four methods of selecting trial distances. In total, 72 simulation scenarios were completed, and each simulation was run 999 times.

In all simulations, the correct form of the detection function was fitted during data analysis (i.e., underlying between-group heterogeneity was accounted for). Repeat captures of the same individual were assumed to not influence detectability (i.e., there was no behavioral effect included in the simulation, so that capture events within an individual were assumed to only depend on distance between animal's location and the trial trap, and its unique detection function specified by its group membership and random effect distribution). A behavioral effect was investigated in Section 6.6).

I calculated the percentage bias ( $=100 * (\bar{N} - N)/N$ ), where  $\bar{N}$  was the mean abundance for all 999 simulations and  $N$  was true abundance), root mean square

---

<sup>2</sup>When the number of trials per individual was 18 or more, the two "Stopping Rule" methods did not use as much survey effort as the other methods (c.f. Section 6.2.3).

error ( $\text{RMSE} = \sqrt{(\sum (N^{\hat{[i]}} - N)^2)/B}$ , where  $N^{\hat{[i]}}$  was the abundance estimate of the  $i$ th simulation and  $B$  was the total number of simulations), and report the 2.5 and 97.5 percentiles for the two abundance estimators  $\hat{N}_2$  and  $\hat{N}_3$  (referred to as percentile intervals for the remainder of the chapter).

### 6.3.2 Results

#### High detectability scenario

Once a minimum number of trials were performed on each individual ( $\geq 10$  trials), the percentage bias in  $\hat{N}_2$  and the 2.5 and 97.5 percentiles did not change substantially for all methods tested, when average detectability was high and there were no underlying groups present in the population (Figure 6.5). When the  $\hat{N}_3$  estimator was used, percentage bias decreased until  $\geq 18$  trials were performed per individual, after which conducting more trials per individual did not change percentage bias; and the “Adaptive” method tended to perform better (i.e., had a lower percentage bias and narrower 2.5 and 97.5 percentiles), compared to the other three methods tested (Figure 6.5).

Regardless of which method was used to select trial distances, estimates of  $\hat{N}_2$  were biased and extremely variable when the number of trials per individual was very low (i.e., 6 trials per individual). Similarly, when the number of trials per individual was  $\leq 10$ , estimates of  $\hat{N}_3$  were biased and extremely variable (Figure 6.5). The direction of bias in  $\hat{N}_2$  changed depending on the method used to select trial distances (i.e., the “Adaptive” method was positively biased, and the “Uniform” and two “Stopping Rule” methods were negatively biased, Figure 6.5). When the “Adaptive” method was used with a small number of trials per individual, if few trials were successful and they were at very short distances, the estimated detection function is “spiked” at zero (Figure 6.6A). When the detection function is “spiked” at zero, the estimated probability of detection is underestimated (Figure 6.6B), and abundance is overestimated (Figure

6.6C). The magnitude of bias at small sample sizes is worse when the “Adaptive” method is used because when the number of trials per individual is small (e.g., 6 trials), the preliminary detection function is estimated on only half as many trials per individual (e.g., 3 trials) spread uniformly between 1 m and  $w$  m. Consequently, the preliminary detection function can be a very poor approximation of the true detection function. Hence, the second set of trials can be badly placed with respect to distance of the trial (e.g., all trials might be placed at short distances with no trials at greater distances so the estimated detection function is “spiked”).

When the “Uniform” or two “Stopping Rule” methods were used to select trial distances with a small number of trials per individual,  $\hat{N}_2$  was negatively biased (Figure 6.5). Figure 6.7 shows the 999 simulation results for the “Uniform” method when 6 trials were performed on 60 individuals. Notice the two groups of detection functions. The first group reflects the same shape as the true detection function from which data were generated, but the second group has a very high intercept and a wide shoulder of high detectability before a rapid decline in detectability. This second group of detection functions represent models where the intercept and random effects variance were greatly over-estimated in the model fitting (e.g., in some instances, the random effects variance was  $> 2800$ , when the true variance was 0.1). In these instances, the average probability of detection was biased high, so  $\hat{N}_2$  was biased low. But due to the high variability in the random effect variance,  $\hat{N}_3$  was biased high: since  $\hat{N}_3$  was based on the average of  $1/p_i$ , a single estimate of  $p$  near zero can cause high bias. This same problem occurred for the two “Stopping Rule” methods.

Regardless of which method was used to select trial distances, the abundance estimator  $\hat{N}_3$  was always positively biased (Figure 6.5), and this bias increased

with a small number of trials conducted on each individual, for the same reasons explained above (i.e., “spiked” detection function). When the “Adaptive” method was used to select trial distances, the 2.5 and 97.5 percentiles were narrower than other methods investigated using the same level of survey effort and relatively constant once  $> 18$  trials were performed per individual. See Table A.3 in Appendix A for complete results table.

Note the percentile intervals for the “Stopping Rule” methods are slightly wider than the “Adaptive” and “Uniform” methods, but use less total survey effort. The average number of trials per individual when the “Stopping Rule” method was used was typically less than 15. Once the number of trials per individual exceeded approximately 15, this additional survey effort was superfluous (Section 6.2.3). Therefore, precision could be increased for the two “Stopping Rule” methods by sampling more individuals for the same total level of survey effort as the other two methods, but the positive bias would remain.

### **Medium detectability scenario**

When there are two underlying groups present in the population, one of which has high average detectability and the other medium average detectability (Figure 6.1B), percentage bias in  $\hat{N}_2$  and  $\hat{N}_3$  is closest to zero, and the 2.5 and 97.5 percentiles are narrowest, when the “Adaptive” method was used to select trial distances, and  $\geq 18$  trials are conducted on each individual (Figure 6.8). Percentage bias in  $\hat{N}_2$  and  $\hat{N}_3$  increases if the number of trials performed on each individual is very high, and across few individuals (i.e., 60 trials on six individuals). This is worse for the two “Stopping Rule” methods (Figure 6.8) than either the “Uniform” or “Adaptive” methods. Since the number of trials performed on each individual using the “Stopping Rule” methods is substantially lower than 60 (typically less than 15, c.f Section 6.2.3), with the decrease in detectability, the fitted detection function is not being estimated as well using these methods

and the “Medium” scenario, compared to the “High” scenario. Percentage bias in  $\hat{N}_3$  is always greater than  $\hat{N}_2$ , regardless of the method used to select trial distances.

When compared to the “High” detection function scenario, percentile intervals for the “Medium” scenario are wider. For example, when 60 trials were performed on 6 individuals, the percentile interval increased from (1136, 3404) when the “High” detection function scenario was specified, to (1031, 3914) when the “Medium” detection function scenario was used (see Table A.3 in Appendix A for complete results table.)

### **Low detectability scenario**

When there are two underlying groups present in the population, one of which has a high average probability of detection and the other a low average probability of detection (i.e., Figure 6.1C), mean estimates of  $\hat{N}_2$  and  $\hat{N}_3$  can be extremely variable and on average, very highly biased, irrespective of the method used to select trial distances – e.g., when 18 trials were performed on 20 individuals, the mean of  $\hat{N}_2$  was  $7.5 \times 10^7$  (see Table A.3 in Appendix A). Consequently, to improve readability in Figure 6.9, I presented the median rather than the mean estimates of  $\hat{N}_2$  and  $\hat{N}_3$ . Similar to the “High” and “Medium” detection function scenarios, fitted detection curves can be “spiked” at zero, causing estimated probabilities of detection for the high group to be slightly underestimated and consequently abundance can be slightly overestimated. More problematic is that detection probabilities for the low group can be severely underestimated, leading to severe overestimation of abundance. For example, when the intercept of the estimated detection function is close to zero, detection probabilities can be severely underestimated (Figure 6.10).

Except for the “Adaptive” method when few trials ( $\leq 10$ ) were conducted on many individuals, mean abundance estimates did not even lie within percentile

intervals (see Table A.3 in Appendix A). This is because sometimes the intercept of the estimated detection function was close to zero (as explained above), and abundance was extremely over-estimated (e.g., estimates of over a million individuals). In such instances, using the median probability of detection produces more “realistic” results (as presented in Figure 6.9, see Table A.3 in Appendix A for mean estimates). Severe underestimation of the detection function was an artifact of the trials survey sampling strategy and estimation procedure, but tended to be less severe when the “Adaptive” method was used (see Table A.3 in Appendix A).

### 6.3.3 Conclusion of Simulation 1

When underlying detectability was “High”, all methods performed similarly, even though the two “Stopping Rule” methods used less survey effort than the “Uniform” and “Adaptive” methods (Figure 6.5). It could be investigated if increasing the survey effort of the two “Stopping Rule” methods (e.g., by varying the stopping rule so the average number of trials per individual was similar to the levels of the others tested, i.e., 6, 10, 18, 24, 40 and 60 trials) improved bias and uncertainty, however as the cost of implementing this method remains unknown prior to starting field work, implementing the “Stopping Rule” method might not be practical.

When underlying detectability was “Medium” (Figures 6.8), in terms of the lowest percentage bias and narrowest percentile interval, the “Adaptive” method outperformed the other methods tested.

In both the “High” and “Medium” detection function scenarios, a minimum of 10 trials per individual were required to achieve the lowest percentage bias in  $\hat{N}_2$  and  $\hat{N}_3$ . Estimates remained positively biased, but the percentage bias in  $\hat{N}_2$  was less than  $\hat{N}_3$  (between 4 and 7% for  $\hat{N}_2$ , and 8 and 12% for  $\hat{N}_3$ ). Conducting a greater number of trials per individual typically did not increase

precision in abundance estimates, when this meant a corresponding decrease in the total number of individuals trials were performed on (since the total survey effort was held constant).

Regardless of detection function scenario, in some instances, the detection probability is overestimated, which causes underestimation of  $\hat{N}$ . As detection probability tends to 1, estimated abundance merely becomes the number of animals captured in the main survey ( $<$  truth). This problem affects the mean estimate of  $\hat{N}$  less than underestimating detection probability, which can cause a single estimate of  $\hat{N}$  to be over a million individuals and severely bias the mean  $\hat{N}$ . In such instances, the median may be a more appropriate metric to use to obtain a better impression of the performance of the simulation. This problem arose more frequently in the “Low” detection function scenario, when mean abundance estimates of  $\hat{N}_2$  and  $\hat{N}_3$  for all methods of selecting trial distances tended to perform badly. However, the median abundance estimates of both estimators in the “Low” detection function scenario performed satisfactorily, for all methods, and hence suggests the TPT method can still be used in these situations where the probability of detection is extremely low. If someone ever obtained such an abundance estimate that was severely overestimated, it would be easy to spot in practise (i.e. abundance estimates would be an order of magnitude larger than expected), and it is recommended that more survey data be collected.

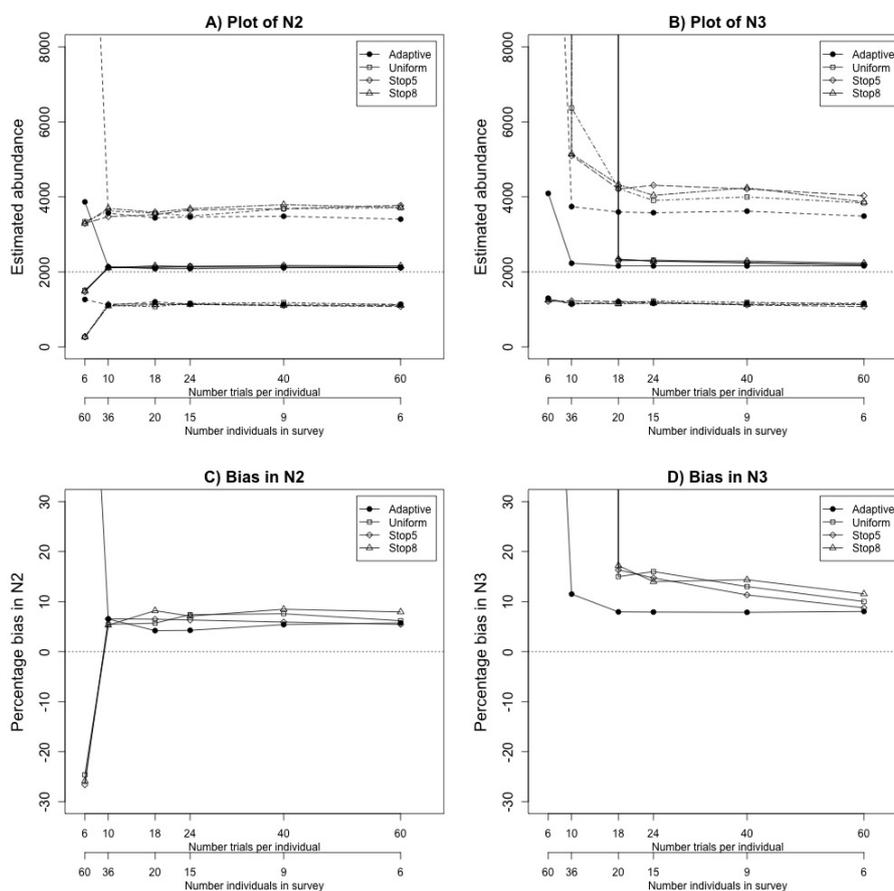


Figure 6.5: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort was 360 trap nights for the “High” detection function scenario. True population abundance was 2,000 (shown as a dotted line in Panel A and B). Note, the number of trials per individual for the two “Stopping Rule” methods is, on average, < 15 trials per individual (see Section 6.3.1 for explanation), so the total survey effort of these two methods is typically less than the “Adaptive” and “Uniform” methods.

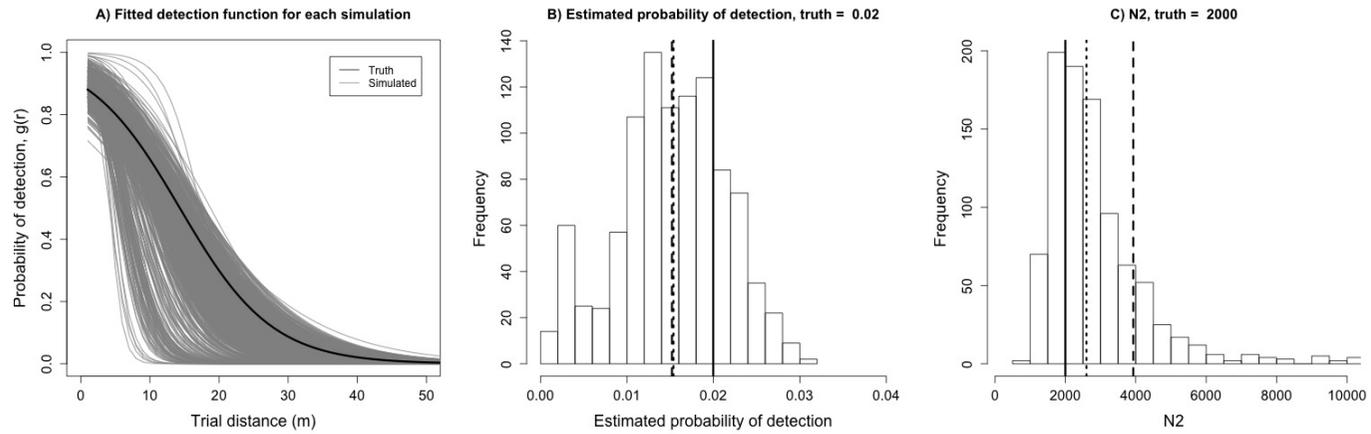


Figure 6.6: Results from 999 simulations of the “Adaptive” method for a “High” detection function scenario when survey effort was set to six trials performed on 60 individuals. Abundance estimates can be positively biased at smaller sample sizes, because some estimated detection functions are “spiked” at zero distance (Panel A), which causes some probabilities of detection to be underestimated (Panel B, the true and estimated mean and median probability of detection is shown as a solid, big dashed and small dashed line, respectively) and consequently, some abundance estimates can be severely overestimated (Panel C, the true and mean and median of the estimated population size is shown as a solid, big dashed and small dashed line, respectively).

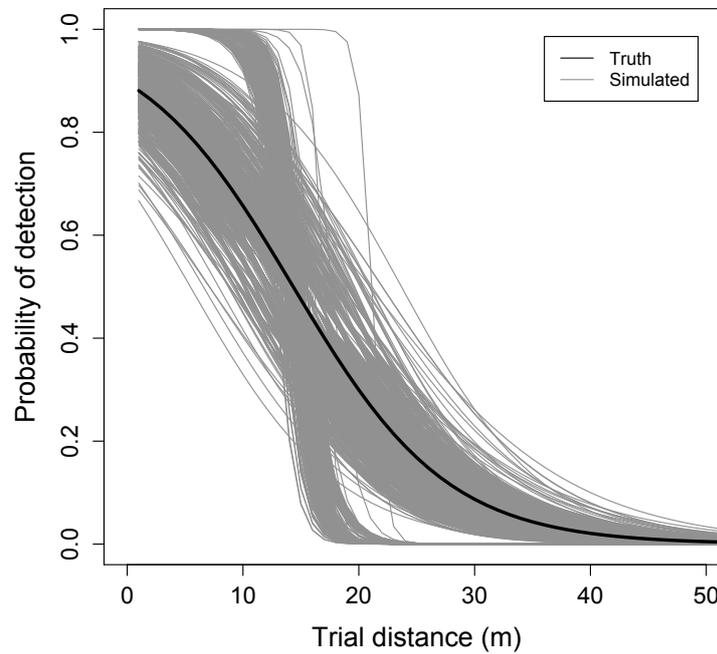


Figure 6.7: Results from 999 simulations of the “Uniform” method for a “High” detection function scenario when survey effort was set to 6 trials being performed on 60 individuals. Abundance estimates can be negatively biased at smaller sample sizes, because some estimated detection functions have a high intercept and wide shoulder, causing estimated probability of detection to be biased high. The same problem occurs for the two “Stopping Rule” methods of selecting trial distances when few trials are performed per individual.

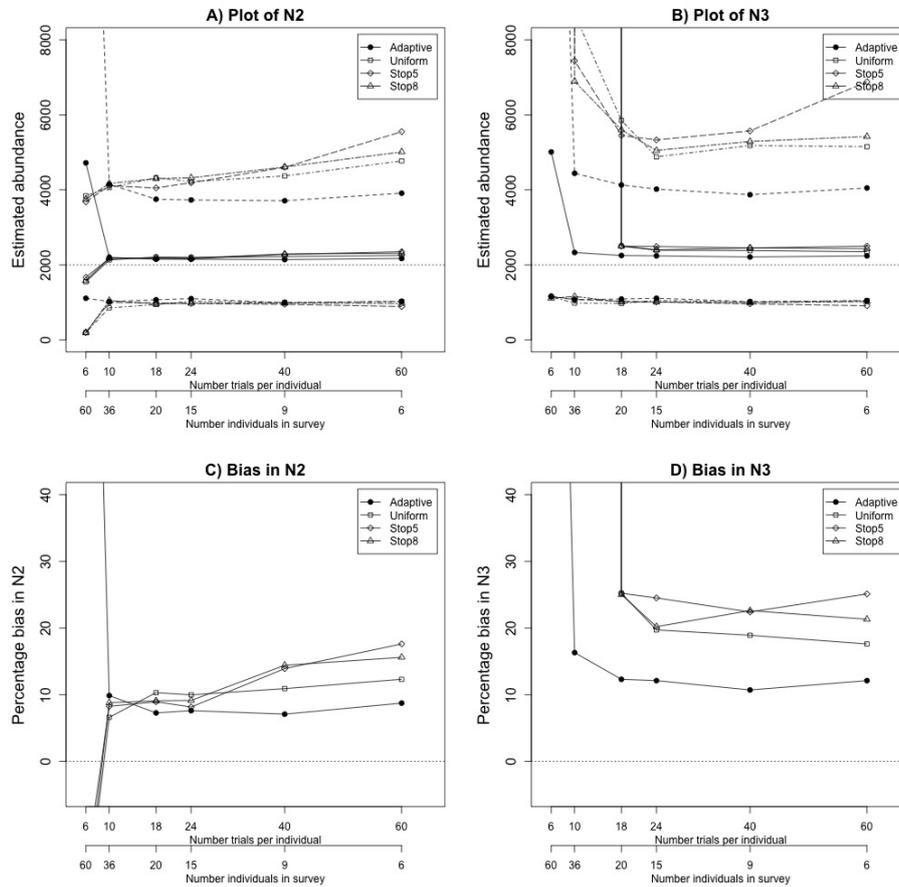


Figure 6.8: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort (360 trap nights) changes for the “Medium” detection function scenario. True population abundance was 2,000 (shown as a dotted line in Panel A and B). Note, the number of trials per individual for the two “Stopping Rule” methods is, on average, < 15 trials per individual (see Section 6.3.1 for explanation), so the total survey effort of these two methods is typically less than the “Adaptive” and “Uniform” methods.

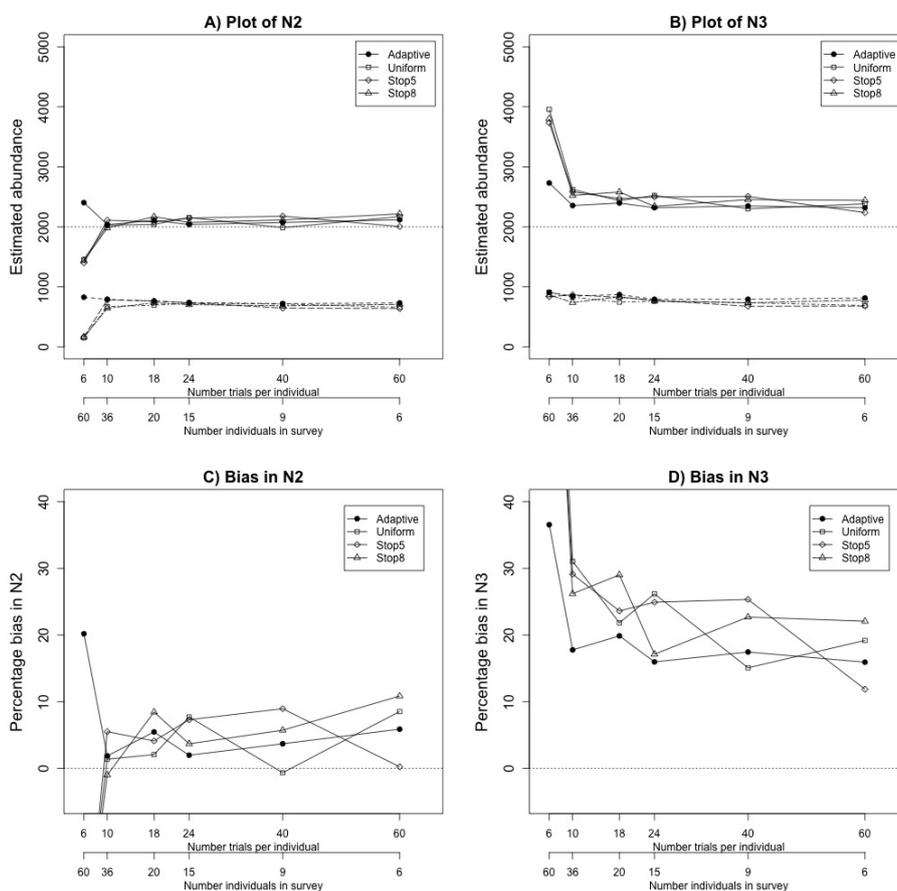


Figure 6.9: Plots of *median* abundance estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort (360 trap nights) changes for the “Low” detection function scenario (see text for explanation of why the median results are presented instead of the mean). True population abundance was 2,000 (shown as a dotted line in Panel A and B). The upper percentile intervals of  $\hat{N}_2$  and  $\hat{N}_3$  for the many simulations were extremely high, and are omitted from these plots. Full results are provided in Table A.3 in Appendix A.

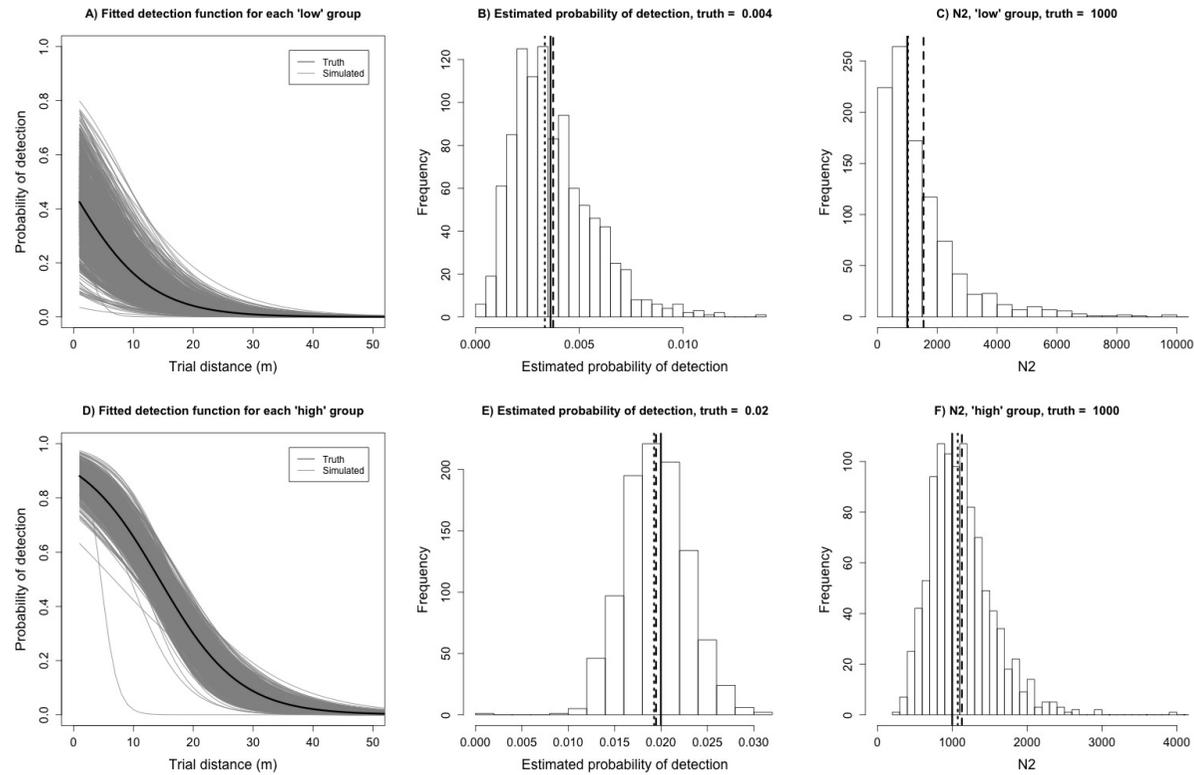


Figure 6.10: Results from 999 simulations of the “Adaptive” method for a “Low” detection function scenario when survey effort is set to 10 trials being performed on each of 36 individuals. The positive bias detected in  $\hat{N}_2$  is caused by estimated probabilities of detection near zero, which causes positive bias in abundance estimates. Plots in the top and bottom lines are results for individuals in the low and high groups, respectively. Plots A and D show the fitted and true mean detection curves (grey and solid lines, respectively). Plots B and E show the true and estimated mean and median probability of detection (solid, big dashed and small dashed lines, respectively). Plots C and F show the true, and mean and median estimate of  $\hat{N}_2$  (solid, big dashed and small dashed line, respectively).

## 6.4 Simulation 2: Allocation of effort when underlying between-group heterogeneity *is not* accounted for

### 6.4.1 Methods

Usually all sources of heterogeneity in capture probability are not known. The aim of Simulation 2 was to determine the consequences of not accounting for a source of underlying heterogeneity on abundance estimates. The same simulations relating to the “Medium” and “Low” detectability scenarios were undertaken as in Simulation 1, but instead of fitting a model that accounts for group heterogeneity (i.e., using R-like linear model formulation:  $\text{Capture} \sim \text{Distance} + \text{Group} + (1|\text{Woodrat})$ ) the group covariate was ignored (i.e.,  $\text{Capture} \sim \text{Distance} + (1|\text{Woodrat})$ ). The “High” detection function scenario was omitted here, since it does not contain any underlying group heterogeneity. Consequently, 48 simulation scenarios were run, each 999 times and the same parameters were calculated as in Simulation 1 (i.e., percentage bias, root mean square error, and the 2.5 and 97.5 percentiles of  $\hat{N}_2$  and  $\hat{N}_3$ ).

### 6.4.2 Results

#### Medium detectability scenario

When there are two underlying groups present in the population, one of which has high average probability of detection and the other medium (i.e., Figure 6.1B), and this underlying heterogeneity is unknown,  $\hat{N}_2$  is negatively biased and  $\hat{N}_3$  is positively biased. These biases in  $\hat{N}_2$  occur irrespective of the number of trials conducted per individual and which method is used to select trial distances (Figure 6.11). The percentage bias in  $\hat{N}_2$  is similar across the four methods, but slightly closer to zero when the “Stopping Rule 8” method is used and a large number of trials are conducted on fewer individuals. This method, however,

has the greatest uncertainty (Figure 6.11). The variance of  $\hat{N}_2$  and  $\hat{N}_3$  using the “Adaptive” method to select trial distances is consistently smaller when compared to the other methods (Figure 6.11).

The cause of negative bias in  $\hat{N}_2$  and positive bias in  $\hat{N}_3$  is due to how each estimator is calculated. Abundance estimator  $\hat{N}_2$  is based on  $1/\bar{p}_{ig}$ , where  $\bar{p}_{ig}$  is the average probability of detecting individual  $i$  in group  $g$ , calculated using 10,000 samples of  $\hat{p}_{ig}$  from the random effect variance of individual  $i$  in group  $g$ . The true probability of detection for the medium and high group is  $p_{.2} = 0.01$  (i.e.,  $1/p_{.2} = 97.3$ ) and  $p_{.1} = 0.02$  (i.e.,  $1/p_{.1} = 50$ ), respectively. Since heterogeneity is ignored,  $1/p_{i.}$  for the medium and high group is equal ( $1/p_{i.} = 66.08$ ). The magnitude of the negative bias for the medium group (Figure 6.12A) is larger than the positive bias of the high group (Figure 6.12B), and hence the overall estimate of  $\hat{N}_2$  is negatively bias (Figure 6.11). Since underlying heterogeneity is unknown, in practise one would never know what the population estimate is for each group. Note that these calculations are based on the the “Adaptive” method of selecting trial distances, when 60 trials were performed on 6 individuals. A similar pattern in bias of  $\hat{N}_2$  is observed for all methods of selecting trial distances.

Abundance estimator  $\hat{N}_3$  is based on the average of  $1/(p_{igs})$  where  $p_{igs}$  is the probability of detecting individual  $i$  in group  $g$  for sample  $s$  from the random effects distribution. As above, one divided by the true probability of detection is 97.3 and 50 for the medium and high group, respectively. Since heterogeneity is ignored, the mean of  $1/p_{i.s}$  for the medium and high group is equal ( $=73.67$ ). The magnitude of negative bias for the low group is slightly smaller (Figure 6.12C) that then positive bias for the high group (Figure 6.12D), and hence  $\hat{N}_3$  is slightly positively biased. In this simulation,  $\hat{N}_3$  was positively biased, but its magnitude was smaller (e.g., 5.5% when 60 trials were performed on

six individuals) than the positive bias also found in  $\hat{N}_3$  when heterogeneity was accounted for (e.g., 12.1% when 60 trials were performed on six individuals, c.f. Simulation 1).

Similarly, RMSE was largest when six trials were conducted on 60 individuals and either the “Uniform” (35.0), “Stopping rule 5” (35.7) or “Stopping rule 8” (34.8) method was used to select trial distances. RMSE was lowest when the “Adaptive” method (20.2) was used to select trial distances, and continued to decrease when the number of trials conducted per individual increased (Table A.4 in Appendix A).

### Low detectability scenario

When there are two underlying groups present in the population, one of which has high average probability of detection and the other low (i.e., Figure 6.1C), and this heterogeneity is unknown,  $\hat{N}_2$  is negatively biased. When compared to the “Medium” scenario that also ignores heterogeneity, the bias and uncertainty in  $\hat{N}_2$  is greater (Figure 6.13). The amount of bias in  $\hat{N}_3$  is comparatively lower than  $\hat{N}_2$  but the percentile intervals are wider. Percentile intervals are narrower when the “Adaptive” method was used to select trial distances (Figure 6.13).

The negative bias in  $\hat{N}_2$  and positive bias in  $\hat{N}_3$  (Figure 6.14) occurred for the same reasons as above. That is, the true probability of detection for the low and high group is  $p_L = 0.003$  (i.e.,  $1/p_L = 276.4$ ) and  $p_H = 0.02$  (i.e.,  $1/p_H = 50$ ), respectively. Since heterogeneity is ignored,  $1/\bar{p}_i$  is equal for both groups (84.7). The magnitude of the negative bias for the low group (Figure 6.14A) is much greater than the small positive bias for the high group (Figure 6.14B), and hence the overall estimate of  $\hat{N}_2$  is biased low. However, the mean of  $1/p_{i.s}$  for the medium and high group is equal to 163.2. The magnitude of the negative bias for the high group is slightly lower (Figure 6.14C) than the positive bias for the high group (Figure 6.14D) and the overall estimate of  $\hat{N}_3$  is underestimated.

Since underlying heterogeneity is unknown, in practise one would never know what the population estimate is for each group. Note that these calculations are based on the the “Adaptive” method of selecting trial distances, when 60 trials were performed on 6 individuals. A similar pattern in bias of  $\hat{N}_2$  is observed for all methods of selecting trial distances.

Similar to the “Medium” detectability scenario, RMSE was largest when fewer trials were conducted on many individuals, but overall was higher with decreasing detectability (Table A.4 in Appendix A). Also, with few trials conducted on many individuals, combined with low detectability, RMSE using the “Adaptive” method to select trial distances has a comparatively high RMSE compared to the other three methods (“Uniform”: 39.0, “Stopping Rule 5”: 38.7, “Stopping Rule 8”: 38.6, “Adaptive”: 148.0, Table A.4 in Appendix A). With  $\geq 6$  trials conducted per individual, the RMSE using the “Adaptive” method was lowest.

### 6.4.3 Conclusion of Simulation 2

Since individuals in the trials survey were selected proportional to their expected frequency in the main survey (as is the most realistic situation to how trials would be allocated in the “real” world), ignoring group membership leads to biased abundance estimators (as shown in this simulation). The “Adaptive” method for selecting trial distances outperformed the other methods tested, in terms of lowest percentage bias and uncertainty, regardless of underlying detectability. When detectability was “Medium”, a minimum of 18 trials per individual was required to achieve the lowest percentage bias in  $\hat{N}_2$  and  $\hat{N}_3$  (albeit, estimates remained negatively biased for  $\hat{N}_2$ ,  $-9.95\%$ , and positively biased for  $\hat{N}_3$ ,  $6.13\%$  when 18 trials were performed on 20 individuals). Conducting a greater number of trials per individual typically did not increase precision in abundance estimates. Conducting too few (i.e., six) or too many (i.e., 60) trials on a large

(i.e., 60) or small (i.e., six) number of individuals caused moderate bias (13%) and uncertainty in  $\hat{N}_3$ . When heterogeneity was ignored (i.e., the incorrect model for detectability was used),  $\hat{N}_3$  was less biased than when the correct model was fitted (as in Simulation 1). The reason the bias in  $\hat{N}_3$  is relatively small in Simulation 2 (when underlying heterogeneity was ignored) compared to Simulation 1 (when underlying heterogeneity was accounted for) is because the negative bias in estimating abundance of the medium-detectability group is approximately equal to the positive bias of the high-detectability group (compare Fig 6.12c with Fig 6.12d). When underlying heterogeneity is accounted for,  $\hat{N}_3$  is extremely positively biased. This finding was surprising, and warrants further investigation to explore whether this result can be generalized or is specific to the scenario in question.

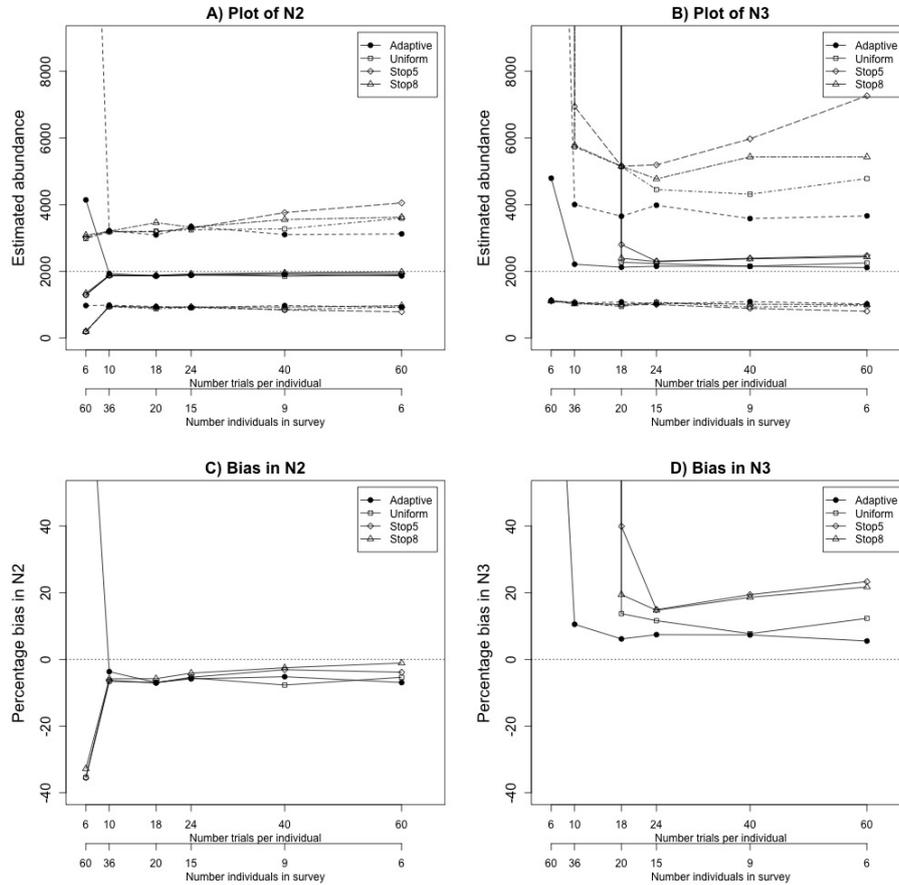


Figure 6.11: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort (360 trap nights) changes for the “Medium” detection function scenario when underlying heterogeneity in the population was ignored. True population abundance was 2,000 (shown as a dotted line in Panel A and B). Note, the number of trials per individual for the two “Stopping Rule” methods is, on average, < 15 trials per individual (see Section 6.3.1 for explanation), so the total survey effort of these two methods is typically less than the “Adaptive” and “Uniform” methods.

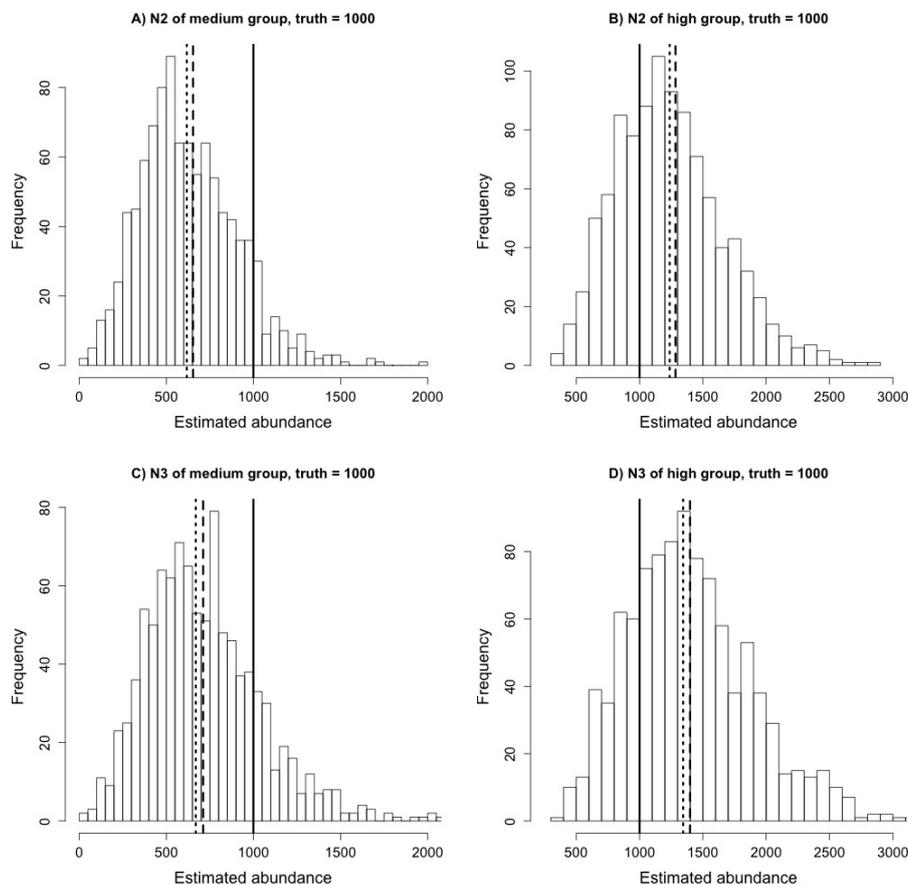


Figure 6.12: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  for the “Medium” detection function scenario when underlying group heterogeneity was ignored. These results are for the “Adaptive” method of selecting trial distances, when 60 trials were conducted per individual, and 6 individuals were used in the trials survey. Plots A) and B) Results of 999 simulations to estimate  $\hat{N}_2$ , which is based on  $1/\bar{p}_{ig}$ , for  $g = 1, 2$ , for the medium and high group, respectively. Plots C) and D) Results of 999 simulations to estimate  $\hat{N}_3$ , which is based on the average of  $1/(p_{i.s})$ , for the medium and high group, respectively. Solid, big dashed and small dashed lines are the true, and mean and median estimated abundances, respectively. Since underlying heterogeneity is unknown, in practise one would never know what the population estimate is for each group.

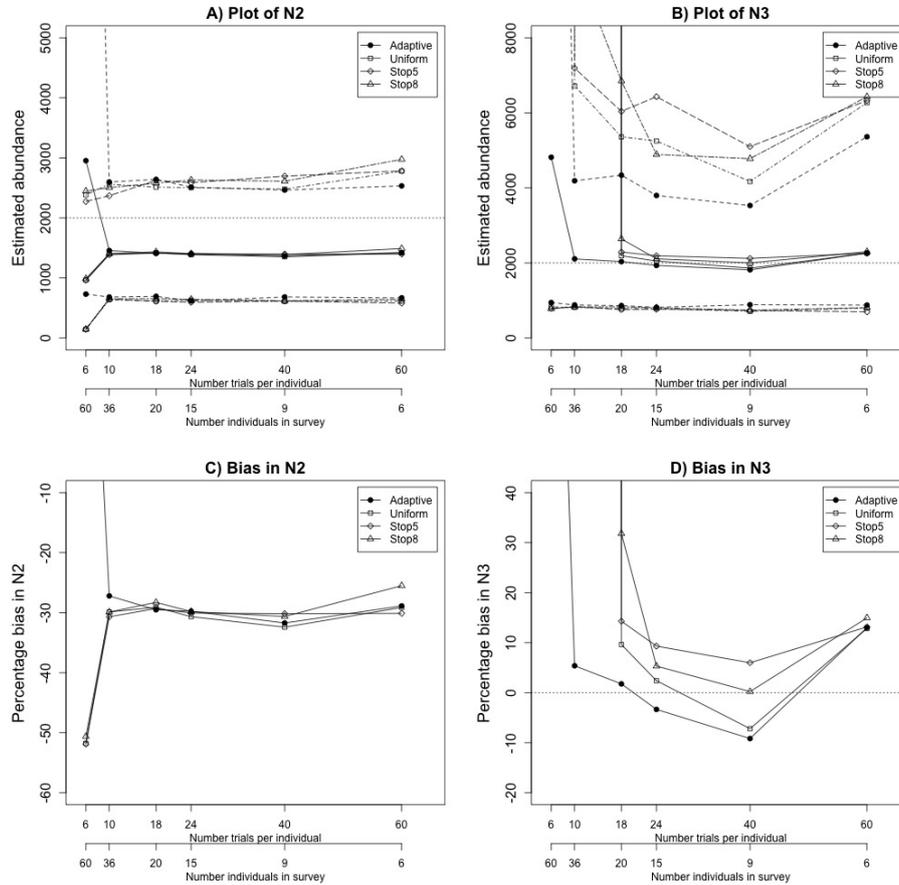


Figure 6.13: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort (360 trap nights) changes for the “Low” detection function scenario when underlying heterogeneity in the population is ignored. True population abundance was 2,000 (shown as a dotted line in Panel A and B). Note, the number of trials per individual for the two “Stopping Rule” methods is, on average, < 15 trials per individual (see text for explanation) so total survey effort of these two methods was typically less than the other methods. Also note, the scale of the Y-axes change.

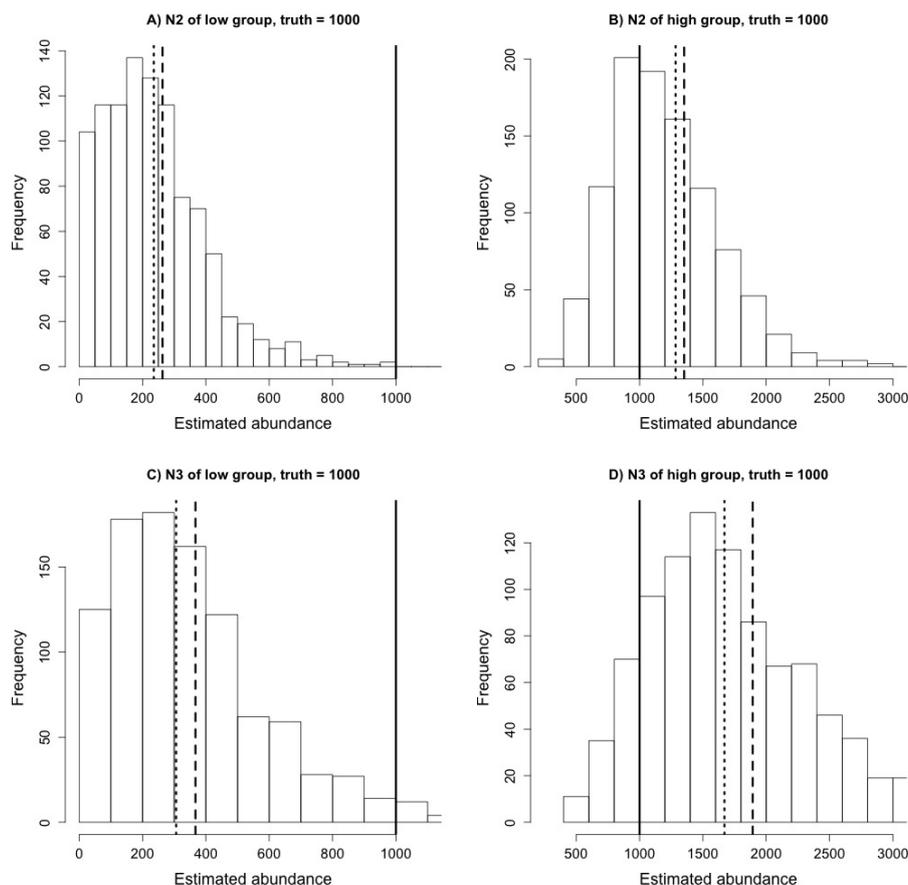


Figure 6.14: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  for the “Low” detection function scenario when underlying group heterogeneity was ignored. These results are for the “Adaptive” method of selecting trial distances, when 60 trials were conducted per individual, and 6 individuals were used in the trials survey. Plot A) Estimated probability of detection for an individual in a single simulation, that sampled the detection function 10,000 times based on the random effects distribution. Plots A) and B) Results of 999 simulations to estimate  $\hat{N}_2$ , which is based on  $1/\bar{p}_{i,g}$ , for the low and high group, respectively. Plots C) and D) Results of 999 simulations to estimate  $\hat{N}_3$ , which is based on the average of  $1/(p_{i,s})$ , for the low and high group, respectively. Solid, big dashed and small dashed lines are the true, and mean and median estimated abundances, respectively. Since underlying heterogeneity is unknown, in practise one would never know what the population estimate is for each group. Plot D was truncated at 3000, to allow easy comparison with Plot B (same x-axes). 9% of observations in Plot D were omitted from the figure (the maximum estimated abundance in the simulation was 9723 individuals).

## 6.5 Simulation 3: Effort required to achieve unbiased abundance estimates

From simulations 1 and 2, I concluded the “Adaptive” method performed best (in terms of percentage bias and uncertainty). However, with a low total survey effort (i.e., 360 trap nights), both estimators were still biased.  $\hat{N}_2$  was biased when underlying heterogeneity was not accounted for;  $\hat{N}_3$  is biased whether underlying heterogeneity is accounted for, or not. In this simulation, I investigate how much total survey effort is needed to obtain unbiased estimates of  $\hat{N}_2$  and  $\hat{N}_3$ .

### 6.5.1 Methods

As per simulation 1 and 2, I set the number of trials per individual to 6, 10, 18, 24, 40, and 60; and changed the number of individuals in the trials survey such that total survey effort was 780, 1080 and 1440 trap nights (Table 6.2). I ran these three additional survey effort simulations for the “Medium” detection function scenario only, both incorporating and ignoring group heterogeneity (i.e., six simulations in total). I make inferences about the likely consequences that increased total survey effort will have on abundance estimates under the “High” and “Low” detection function scenarios.

### 6.5.2 Results

#### Including heterogeneity

As per simulations 1 and 2, increasing the number of trials conducted on each individual to more than approximately ten did not increase precision in abundance estimates, regardless of total survey effort (Figure 6.15). However, as would be expected, percentile intervals of  $\hat{N}_2$  and  $\hat{N}_3$  were narrower when total survey effort increased (regardless of whether the effort was allocated to more trials or more individuals). This increase in precision came at considerable survey cost:

Table 6.2: Number of individuals upon which trials were performed, based on four different levels of trials survey effort (360, 720, 1080 and 1440 trap nights), and the number of trials per individual (6, 10, 18, 24, 40, 60). For example, if the trial survey effort is 720 trap nights and 40 trials are to be conducted per individual, trials will need to be performed on 18 individuals.

Number trials per individual	Number of trap nights			
	360	720	1080	1440
6	60	120	180	240
10	36	72	108	144
18	20	40	60	80
24	15	30	45	60
40	9	18	27	36
60	6	12	18	24

doubling survey effort from 360 to 720 trap nights when 18 trials were performed on each individual, decreased the percentile interval by approximately 18% for  $\hat{N}_2$  from (1073, 3754) to (1100, 3275) (see full results Table A.5 in Appendix A).

Percentage bias in  $\hat{N}_2$  decreased with increasing survey effort, and the lowest percentage bias (=0.95%) was achieved when 40 trials were conducted on 36 individuals (1440 trap nights), however this bias was comparable with those obtained with less sampling effort. For example, the bias when 40 trials were conducted on 27 individuals (1080 trap nights) was 1.06%, and when 18 trials were conducted on 40 individuals (720 trap nights), the bias in  $\hat{N}_2$  was 1.57%. Once approximately ten trials were conducted on each individual, bias was relatively stable. As found in Simulation 1, estimates of  $\hat{N}_3$  remain positively biased when underlying heterogeneity was accounted for in the analysis.

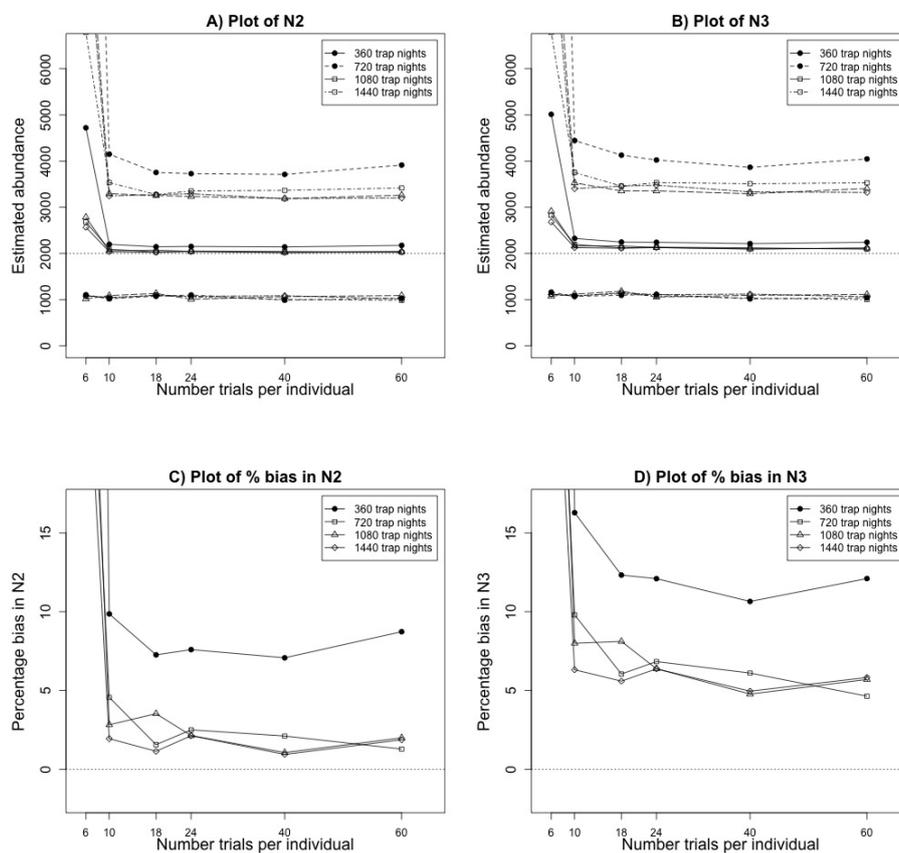


Figure 6.15: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as a solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort changed for the “Medium” detection function scenario and when the “Adaptive” method was used to select trial distances. True population abundance was 2,000 (shown as a dotted line in Panel A and B).

### Ignoring heterogeneity

As would be expected, variability in  $\hat{N}_2$  and  $\hat{N}_3$  decreased with increasing total survey effort, however the increase in certainty (repeatability) was minimal considering the cost, Section 6.16). For example, when 40 trials were conducted per individual and survey effort was 360 trap nights, the difference between the 2.5 and 97.5% percentiles of  $\hat{N}_2$  was 2127. When survey effort was quadrupled to 1440 trap nights, the difference between the 2.5 and 97.5 percentiles of  $\hat{N}_2$  decreased to 1732. Similar results were found with the  $\hat{N}_3$  estimator (see Table A.5 in Appendix A). As per Simulation 1 and 2, once approximately ten trials were conducted per individual, percentile intervals remained relatively constant with increasing the number of trials per individual, as total survey effort increased (Figure 6.16).  $\hat{N}_2$  was always negatively biased, regardless of survey effort (as per the results from Simulation 2). The bias in  $\hat{N}_3$  was closest to zero when 60 trials were conducted on 24 individuals (=3.00%), however when survey effort was halved (i.e., 40 trials on 18 individuals), percentage bias was comparable (=3.18%).

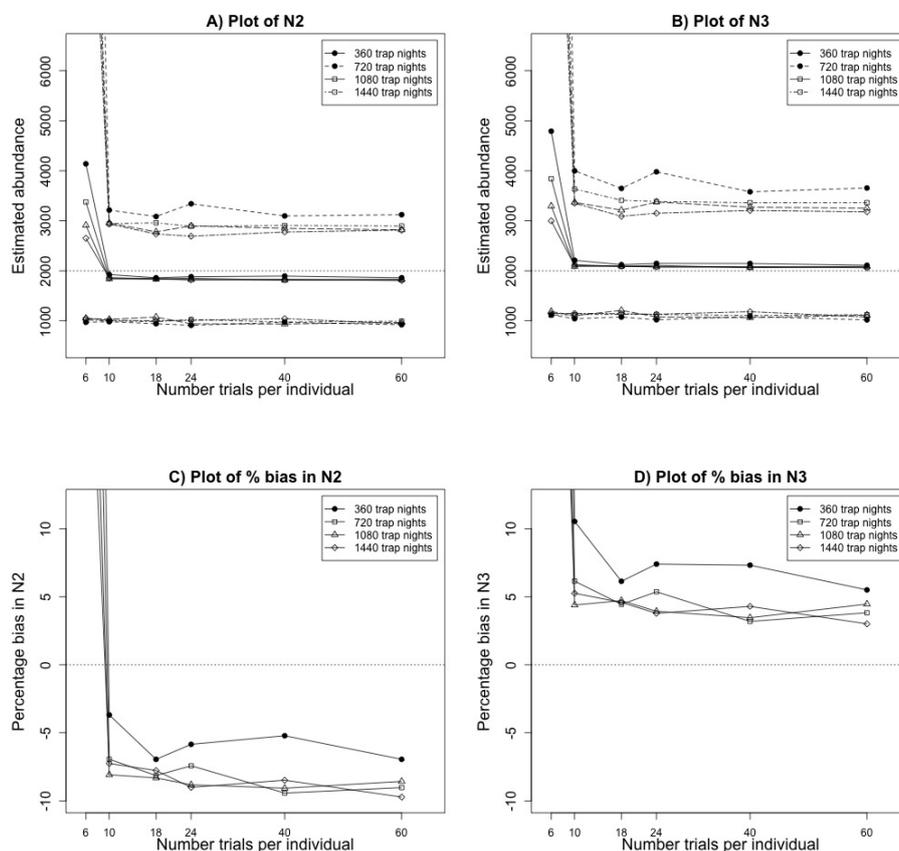


Figure 6.16: Plots of mean abundance estimates from estimators  $\hat{N}_2$  and  $\hat{N}_3$  (Panel A and B, respectively, shown as a solid line) with 2.5 and 97.5 percentiles (dashed lines), and percentage bias estimates (Panel C and D) when the allocation of total trial survey effort changed for the “Medium” detection function scenario, when the “Adaptive” method was used to select trial distances, and underlying heterogeneity in the probability of detection was ignored during the analysis. True population abundance was 2,000 (shown as a dotted line in Panel A and B).

### 6.5.3 Conclusion of Simulation 3

As per Simulation 1 and 2, once a minimum number of trials were conducted per individual (approximately 10), little reduction in bias and increase in precision of abundance estimators occurred. The biggest factor influencing bias and precision in  $\hat{N}_2$  and  $\hat{N}_3$  was increasing total survey effort (i.e., the number of individuals upon which trials were performed). The most noticeable decrease in bias and increase in precision occurred when survey effort doubled from 360 trap nights (i.e., the survey effort used in Simulations 1 and 2) to 720 trap nights. Survey effort greater than 720 trap nights did not cause a substantial reduction in bias or increase in precision, regardless of whether underlying heterogeneity was accounted for, or not.

With increased detectability (i.e., the “High” scenario), total survey effort might be reduced so requisite levels of certainty and percentage bias could be achieved. With decreased detection probability and increased between individual variation (i.e., the “Low” scenario), much larger survey effort would be required and percentage bias might be closer to zero when the number of trials per individual is increased (above 18 trials per individual).

## 6.6 Simulation 4: Including a behavioral response to being trapped

Simulations 1 and 2 assumed capture events of animals were independent. That is, if an animal was captured in the previous trial, its probability of capture in the next trial depended only on distance to the trial trap and its unique detection function specified by its group membership (i.e., “High”, “Medium” or “Low”) and its random effect distribution. Any behavioral response due to the previous trapping experience was ignored. behavioral responses of animals to trapping fall into two categories. Trap happiness occurs when the animal is attracted to the shelter and/or food provided by the trapping experience and if it sees a trap in the future, will be more inclined to enter (i.e., capture probability increases after first capture event). Trap shyness is the opposite response where capture probability is lower once animals have been caught once (Jones et al., 1996). In many instances, there is a behavioral response to the trapping experience, and the consequences of ignoring any behavioral response if it occurs will be investigated here. There are countless different degrees of responses to the trapping experience. Responses can also be transient (e.g., a 50% reduction in capture probability the day immediately after capture, and thereafter returns to “normal”) or permanent. I investigated the former scenario (temporary “trap shyness”).

### 6.6.1 Methods

For the purposes of brevity, only the “Adaptive” method to choose trial distances was considered. The “Adaptive” method was shown in Simulation 1 and 2 to provide abundance estimates with the least bias when a minimum number of trials per individual were conducted. Survey effort was fixed at 24 trials per individual and 15 individuals were used in the trials survey which gave a total

trials survey effort of 360 trap nights. Two detection function scenarios that specified the mean behavior of an individual that had *not* been captured in the previous trial were used (the “High” and “Low” detection function scenarios, Section 6.2.1). In this simulation though, whenever an individual was captured in the previous trial, a new detection function applied that had 50% lower average probability of detection, indicating the animal became trap shy. This new detection function applied for only the next trial, after which, if the animal was not captured, its behavior returned to normal (see Figure 6.17 and 6.18 for the “High” and “Low” detection function scenarios, respectively).

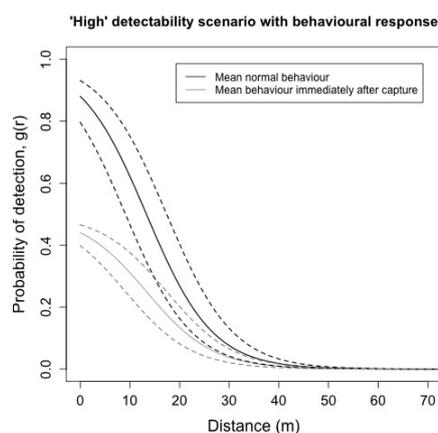


Figure 6.17: Plot of behavioral response when the “High” detection function scenario was used. When an individual has not been captured in the previous trial, the “Mean normal behavior” detection function applies (same parameters as the “High” detection function scenario). However, if the animal was captured in the previous trial, the “Mean behavior immediately after capture” detection function applies, which reduced the probability of capture by 50%.

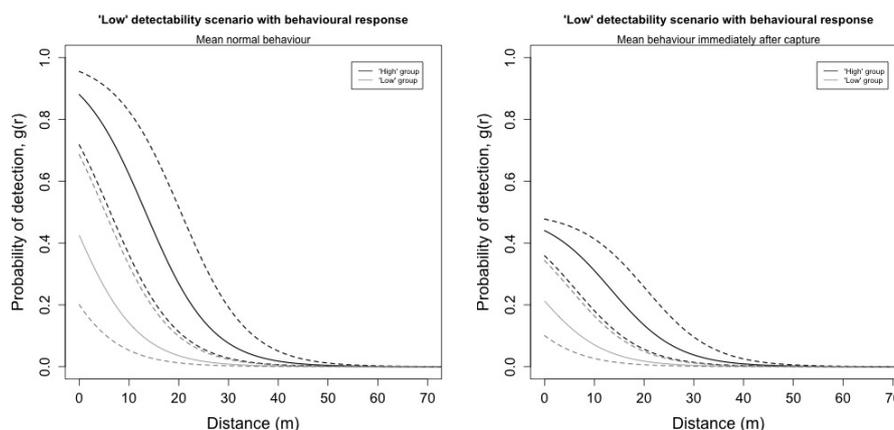


Figure 6.18: Plot of behavioral response when the “Low” detection function scenario was used. When an individual has not been captured in the previous trial, the “Mean normal behavior” detection function applies (same parameters as the “Low” detection function scenario). However, if the animal was captured in the previous trial, the “Mean behavior immediately after capture” detection function applies, which reduced the probability of capture by 50%.

### 6.6.2 Results

The inclusion of a trap shyness behavioral effect caused significant positive bias in  $\hat{N}_2$  (of 32.76%) and  $\hat{N}_3$  (of 33.30%) when the “High” detection function scenario was specified (Table 6.3). When no trap response was present, bias in  $\hat{N}_2$  and  $\hat{N}_3$  was 4.26% and 7.91%, respectively (see Table 6.3 and Figure 6.19; also c.f. Simulation 1 and Figure 6.5 on page 108). When average probability of detection was high, the trap shyness behavior caused the estimated probability of detection to be biased low, and abundance to be overestimated (Figure 6.20).

When the “Low” detection function scenario was specified, percentage bias in  $\hat{N}_2$  increased with the inclusion of a behavioral response (35.56%, versus 23.12% when the behavioral response was not present, Table 6.3). However, the reverse was found with  $\hat{N}_3$ : percentage bias decreased to 38.99% when behavior was included, versus 45.43% when behavior was not present. When a behavioral

response was included in the simulation, the difference in percentage bias between  $\hat{N}_2$  and  $\hat{N}_3$  was less when the “Low” scenario was used, compared to when the “High” detection function scenario was used. It seems the lower the average probability of detection caused by trap shyness (Figure 6.21), the lower the bias in the estimate of the probability of detection (Figure 6.21B and E) yielding a lower bias in the abundance estimate (Figure 6.21C and F).

### 6.6.3 Conclusion of Simulation 4

Trap shyness can be considered a source of underlying heterogeneity in detection probability. Regardless of detectability (“High” or “Low”), percentage bias was slightly less when the  $\hat{N}_2$  estimator was used. This may be because the behavioral response was transient (i.e., detectability decreased immediately following capture and then returned to “normal”).

In some circumstances, it might be possible to explicitly model a behavioral effect in the detection function. In its simplest form, this might be a factor covariate, indicating if the animal was captured before, or not captured in the previous time period. Not all animals in the main survey would have been captured previously. It is difficult to investigate this issue using the radio collaring approach of the TPT survey. Since all animals wearing radio collars must have at one point been captured to attach the radio collar, it is assumed that the detectability of these radio collared animals is representative of the animals in the main survey. Standard capture-recapture methods have been developed that explore a wide variety of behavioral response models (e.g., Chao 1987) and the application of these concepts to the TPT survey method should be investigated. In addition, a goodness-of-fit test could be developed to investigate whether a behavioral response in the data collected using the TPT method is present (e.g., Chao et al. 2000).

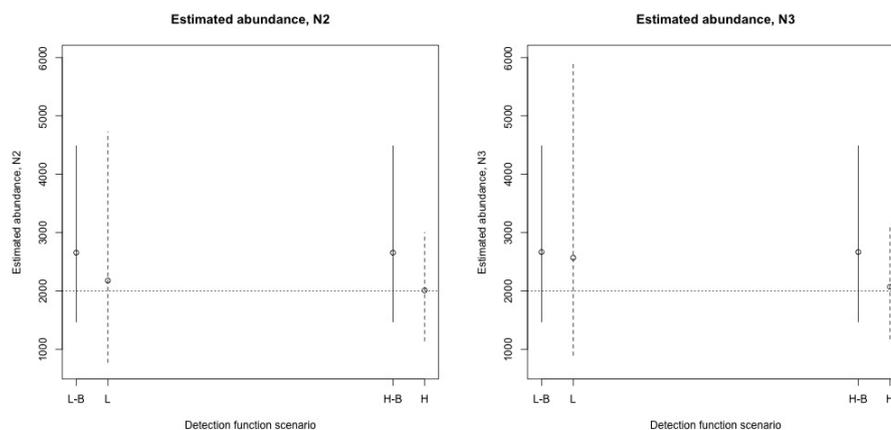


Figure 6.19: Plot of abundance estimators  $\hat{N}_2$  and  $\hat{N}_3$  (and 2.5 and 97.5 percentiles) using the “Low” and “High” detection function scenarios with and without the inclusion of a behavioral response. X-axis labels: L = “Low” detection function scenario, without behavioral response. L-B = “Low” detection function scenario, with a behavioral response. H = “High” detection function scenario, without behavioral response. H-B = “High” detection function scenario, with a behavioral response. True abundance in the main survey was 2,000 individuals (dotted line).

Table 6.3:  $\hat{N}_2$  and  $\hat{N}_3$  abundance estimator results with and without a trap shyness behavioral response for the “High” and “Low” detection function scenarios. Trials survey effort was fixed at 24 trials per individual, and there were 15 individuals (i.e., total survey effort was 360 trap nights). The true abundance in the simulation was 2,000 individuals.

Scenario	Response	mean( $\hat{N}_2$ )	sd( $\hat{N}_2$ )	95% PI( $\hat{N}_2$ )	median( $\hat{N}_2$ )	%bias( $\hat{N}_2$ )	RMSE( $\hat{N}_2$ )
High	Trapshy	2655.14	836.10	(1467.36, 4488.00)	2549.63	32.76	33.60
High	Normal	2085.17	577.90	(1145.67, 3462.51)	2014.22	4.26	18.47
Low	Trapshy	2711.24	1721.26	(854.70, 7647.29)	2249.09	35.56	59.26
Low	Normal	2462.44	1607.33	(737.92, 6751.62)	2026.52	23.12	53.19
Scenario	Response	mean( $\hat{N}_3$ )	sd( $\hat{N}_3$ )	95% PI( $\hat{N}_3$ )	median( $\hat{N}_3$ )	%bias( $\hat{N}_3$ )	RMSE( $\hat{N}_3$ )
High	Trapshy	2665.952	836.091	(1467.36, 4487.99)	2552.74	33.30	33.81
High	Normal	2158.23	614.81	(1159.66, 3577.30)	2077.06	7.91	20.08
Low	Trapshy	2779.81	1778.80	(898.33, 7674.80)	2293.69	38.99	61.80
Low	Normal	2908.67	2101.58	(793.05, 8667.91)	2306.99	45.43	72.81

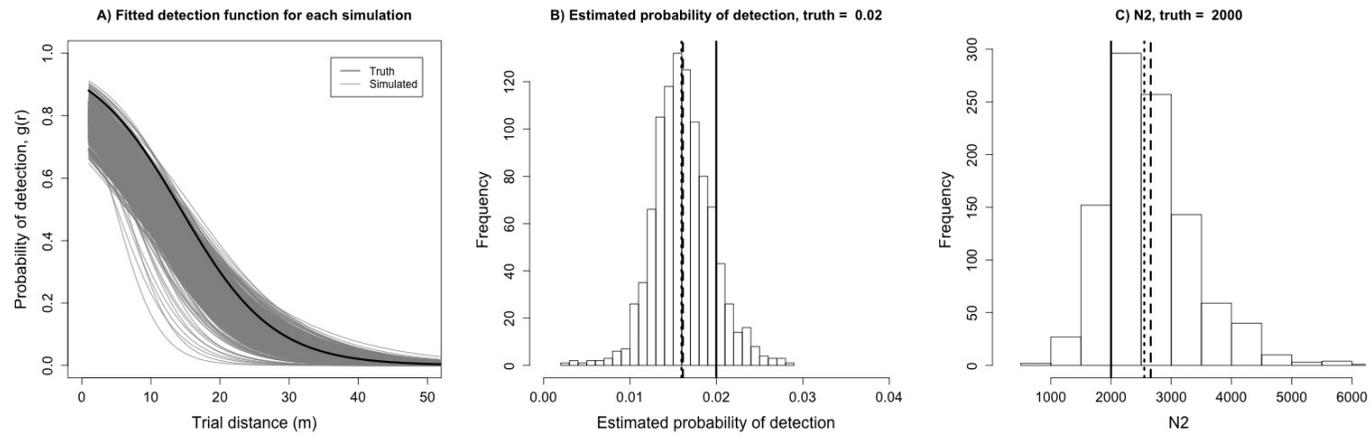


Figure 6.20: Plot of mean abundance estimate  $\hat{N}_2$  for the “High” detection function scenario, when the probability of detection decreased by 50% if the individual was detected in the previous trial. Panel A) shows the 999 fitted detection curves (grey lines) and the true mean detection curve (solid line). The probability of detection for the high group was biased low (Panel B) and hence the abundance estimate  $\hat{N}_2$  was biased high (Panel C). The solid, big dashed and small dashed lines in Panel B and C are the true and estimated mean and median probabilities and abundances estimates, respectively.

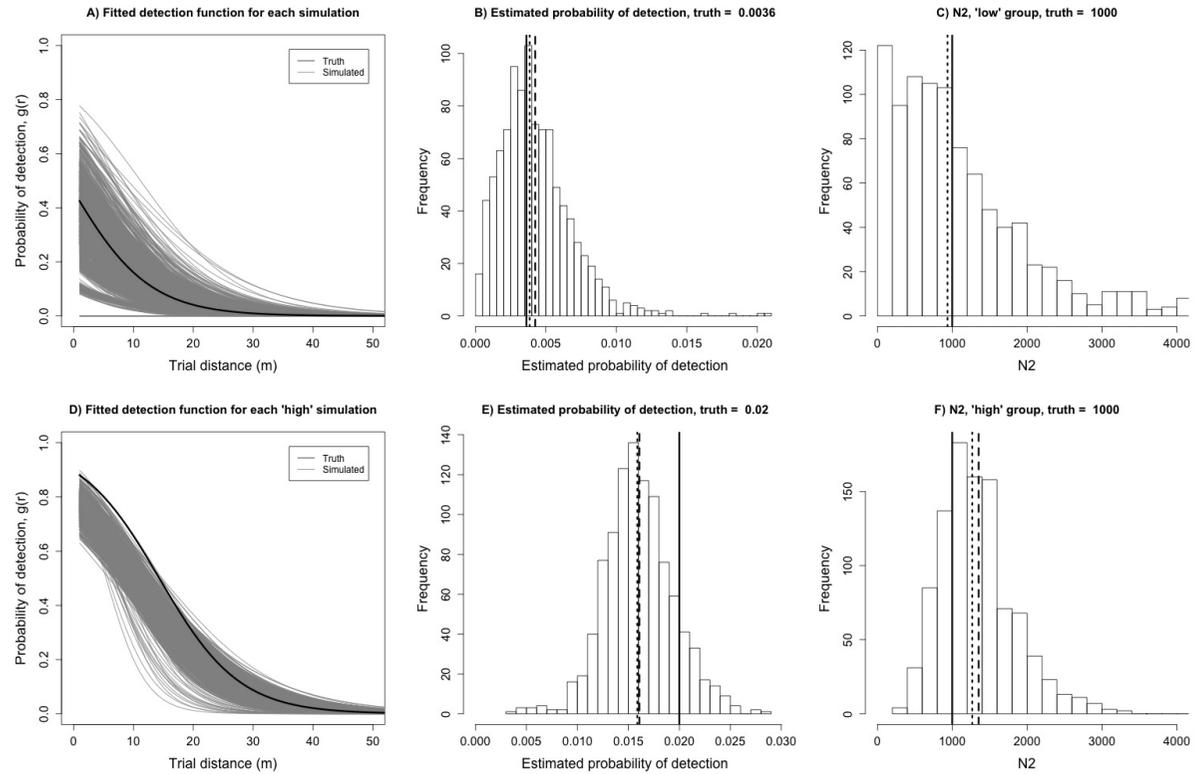


Figure 6.21: Plot of mean abundance estimate  $\hat{N}_2$  for the “Low” detection function scenario, when the probability of detection decreases by 50% if the individual was detected in the previous trial. The probability of detection for the low group is estimated reasonably well (Panel A and B) so that  $\hat{N}_2$  is (nearly) unbiased (Panel C). However, there is negative bias in the detection probability for the high group (Panel D and E). Consequently, the abundance estimate  $\hat{N}_2$  of the high group is biased high (Panel F). Solid line is the true mean probabilities and abundance estimates, the big and small dashed lines are mean and median estimates from the simulation, respectively. The mean estimate is missing on Panel C, because it is biased high and not within the range of the x-axis.

## 6.7 Simulation 5: The importance of including a random effect

Depending on how the trial surveys are conducted, it is foreseeable that repeat trials may be conducted on the same individual, and the detection function needs to be modelled using e.g., a mixed effects model to take account of the non-independence of the trial observations (see Section 5.3.1 on page 79). The inclusion of random effects in the model increases statistical complexity, possibly rendering the analysis of data collected using the trapping point transect field methods arduous for field biologists without input from a statistician. In addition, the inclusion of a random effect increases computational time, especially integrating out the random effect from the detection function, which is the most time consuming aspect of this method<sup>3</sup> (although more efficient integration methods could be implemented, e.g. numerical integration). The consequences of omitting the random effect from the detection function in the analysis are investigated below.

### 6.7.1 Methods

Using the three detection function scenarios and the “Adaptive” method to select trial distances, I simulated trial survey data that included a random effect component. Instead of fitting a detection function that modelled the random effect variance, as in Simulation 1, the random effect was ignored. Total survey effort was fixed at 360 trap nights, 24 trials were conducted per individual, and there were 15 individuals in the survey. Results were compared to those obtained when the random effect component was included (i.e., Simulation 1).

I also ran these two simulations excluding the group covariate and compared

---

<sup>3</sup>Note, conducting 999 simulations did not typically take greater than 16 hours of computer time, and depended on abundance in the main survey, underlying detectability, complexity of the fitted detection function, survey effort and processing power of the computer.

the results to when the group covariate was excluded but the random effect variance was included (i.e., Simulation 2), for the same level of survey effort.

### 6.7.2 Results

When the group covariate was included in the analysis (Table 6.4), omitting the random effect component of the model decreased the percentage bias of  $\hat{N}_2$  when the “High” and “Medium” detectability scenarios were used (from 4.26% to 3.34% when the random effect was included and excluded, respectively for the “High” scenario, and 7.6% to 6.2% for the “Medium” scenario). However, using the “Low” detectability scenario, omission of the random effect increased the percentage bias in  $\hat{N}_2$  (from 23.1%, to 24.4%). The percentile intervals were typically narrower when the random effect was omitted<sup>4</sup>.

When the group covariate was excluded in the analysis (Table 6.5), the consequence of omitting the random effect was extremely noticeable for the abundance estimator  $\hat{N}_3$  (less so for  $\hat{N}_2$  which was already found to be extremely negatively biased when underlying heterogeneity was ignored, c.f. Simulation 2 in Section 6.4). When the random effect was omitted,  $\hat{N}_3$  became more negatively biased, regardless of which detection function scenario was used (Table 6.5). For example, bias in  $\hat{N}_3$  increases from -3.34% to -31.88% when the random effects component was and was not included in the model, respectively, for the “Low” detection function scenario.

Excluding the random effect component from the detection function when heterogeneity was not accounted for increased bias. It seems the probability of detection for the low group present in the population was severely overestimated leading to the abundance estimate being severely underestimated (Figure 6.22, N.B. the same negative bias was found in Simulation 2 when heterogeneity was ignored).

---

<sup>4</sup>Note, with no random effects variance or group covariate  $\hat{N}_2 = \hat{N}_3$ .

### 6.7.3 Conclusion of Simulation 5

Regardless of detection probability and whether underlying heterogeneity in detection probability was accounted for in the analysis, uncertainty in  $\hat{N}_2$  and  $\hat{N}_3$  was larger when the random effect was included in the analysis. This occurred because the random effect adds an additional level of uncertainty regarding individual variability, that a standard GLM does not. When the random effect was omitted for the “High” and “Medium” detection function scenarios the percentage bias in  $\hat{N}_2$  decreased when underlying heterogeneity was accounted for, as in both these scenarios, the individual random variance was small ( $b_{ig} \sim N(0,0.1)$ ,  $g = H$  or  $M$ ). When the random effect variance was larger (i.e., the “Low” scenario,  $b_{iL} \sim N(0, 0.3)$ ), omitting the random effect increased bias in  $\hat{N}_2$ . Similarly, when heterogeneity was excluded, omitting the random effect when the true individual random effect variance was very high (i.e., the “Low” scenario), caused large bias in  $\hat{N}_3$ . When the repeated measures nature of the data (i.e., the random effect) was ignored, there is less bias in  $\hat{N}_3$ , when you know about the important covariates (such as group in this case). When underlying heterogeneity is unknown,  $\hat{N}_3$  is extremely biased if the random effect is excluded, which does at least allow for between-individual variation.

Table 6.4:  $\hat{N}_2$  and  $\hat{N}_3$  abundance estimates when the random effects component of the data structure is ignored (i.e., individual trials are assumed independent of each other, “No RE”) for the “High” and “Low” detection function scenarios (24 trials on 15 individuals, respectively). Models which include the random effects component (“With RE”) from Simulation 1 are provided for comparison. Note, with no random effects variance or group covariate  $\hat{N}_2 = \hat{N}_3$ . True abundance in the main survey was 2,000 individuals.

Model	Scenario	mean( $\hat{N}_2$ )	sd( $\hat{N}_2$ )	95% PI( $\hat{N}_2$ )	median( $\hat{N}_2$ )	% bias( $\hat{N}_2$ )	RMSE( $\hat{N}_2$ )
With RE	High	2085.171	577.90	(1145.67, 3462.51)	2014.222	4.26	18.47
No RE	High	2066.75	543.25	(1165.40, 3334.32)	2023.50	3.34	17.31
With RE	Med	2151.91	699.43	(1101.35, 3728.31)	2062.29	7.60	22.63
No RE	Med	2124.26	678.65	(966.98, 3672.16)	2049.00	6.21	21.82
With RE	Low	2462.44	1607.33	(737.92, 6751.62)	2026.52	23.12	53.19
No RE	Low	2487.99	1753.04	(797.75, 7513.65)	2012.53	24.40	57.98
Model	Scenario	mean( $\hat{N}_3$ )	sd( $\hat{N}_3$ )	95% PI( $\hat{N}_3$ )	median( $\hat{N}_3$ )	% bias( $\hat{N}_3$ )	RMSE( $\hat{N}_3$ )
With RE	High	2158.23	614.81	(1159.66, 3577.30)	2077.06	7.91	20.08
No RE	High	2066.75	543.25	(1165.40, 3334.32)	2023.50	3.34	17.31
With RE	Med	2241.93	750.30	(1114.82, 4024.02)	2129.12	12.10	24.93
No RE	Med	2124.26	678.65	(966.98, 3672.16)	2049.00	6.21	21.82
With RE	Low	2908.67	2101.58	(793.05, 8667.91)	2306.99	45.43	72.81
No RE	Low	2487.99	1753.04	(797.75, 7513.65)	2012.53	24.40	57.98

Table 6.5: Results for  $\hat{N}_2$  and  $\hat{N}_3$  abundance estimators when the random effects component of the data structure is ignored (i.e., individual trials are assumed independent of each other) and the group covariate for the “Medium” and “Low” detection function scenarios is also ignored. A high survey effort was used (24 trials per 15 individuals, respectively). Models which include the random effects component from Simulation 2 are provided for comparison. Note, with no random effects variance or group covariate  $\hat{N}_2 = \hat{N}_3$ . True abundance in the main survey was 2,000 individuals.

Model	Scenario	mean( $\hat{N}_2$ )	sd( $\hat{N}_2$ )	95% PI( $\hat{N}_2$ )	median( $\hat{N}_2$ )	% bias( $\hat{N}_2$ )	RMSE( $\hat{N}_2$ )
With RE	Med	1883.09	626.05	(910.59, 3340.79)	1790.546	-5.85	20.14
No RE	Med	1846.84	519.66	(960.48, 3102.09)	1787.38	-7.66	17.13
With RE	Low	1404.67	477.15	(616.05, 2511.76)	1357.165	-29.77	24.13
No RE	Low	1362.40	416.14	(666.07, 2344.83)	1323.83	-31.88	24.09
Model	Scenario	mean( $\hat{N}_3$ )	sd( $\hat{N}_3$ )	95% PI( $\hat{N}_3$ )	median( $\hat{N}_3$ )	% bias( $\hat{N}_3$ )	RMSE( $\hat{N}_3$ )
With RE	Med	2147.98	750.96	(1020.72, 3979.09)	2011.39	7.40	24.20
No RE	Med	1846.84	519.66	(960.48, 3102.09)	1787.38	-7.66	17.13
With RE	Low	1933.12	765.94	(816.56, 3797.42)	1817.94	-3.34	24.31
No RE	Low	1362.40	416.14	(666.07, 2344.83)	1323.83	-31.88	24.09

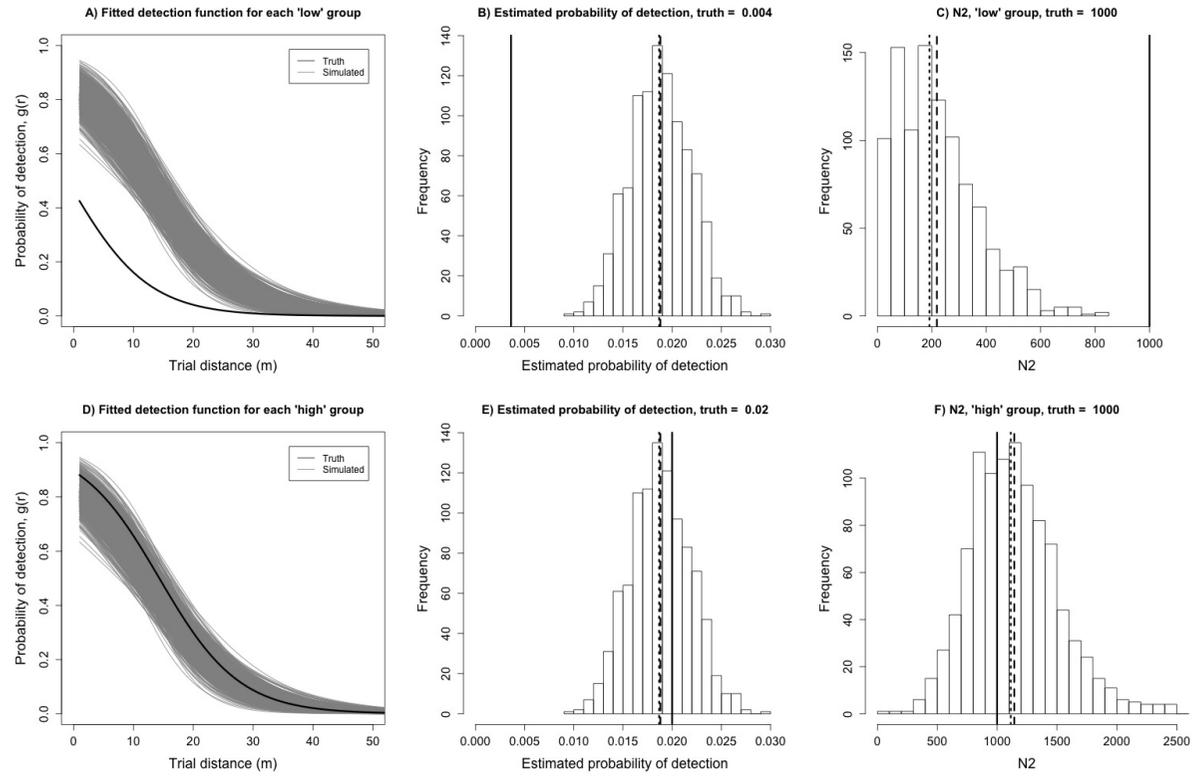


Figure 6.22: Plot of mean abundance estimate  $\hat{N}_2$  for the “Low” detection function scenario, when the repeated trial observations on each individual are incorrectly assumed independent, and the underlying group heterogeneity is not accounted for in the model. The probability of detection for the low group is biased high (Panel A and B), and hence the abundance estimate  $\hat{N}_2$  is biased low (Panel C). The probability of detection of the high group is slightly biased high (Panel D and E), and hence the abundance estimate is slightly biased high (Panel F). The solid, big dashed and small dashed lines in Panel B and C, and E and F are the true, mean and median simulated probabilities and abundances estimates for the high and low groups, respectively.

## 6.8 General Conclusion

It is impossible to fully interrogate all possible sampling strategies and detection function scenarios of a complex, “real-world” problem. As such, I chose what I considered to be a range of survey efforts that could be realistically achieved in the field when conducting a trapping point transect survey. I also considered a range of detection function scenarios from a “best” case scenario with high average probability of detection and little variability within and between individuals (i.e., the “High” detection function scenario), to a “worst” case scenario with high variability within and between individuals (i.e., the “Low” detection function scenario). In the simulation study I presented in this chapter, I investigated the consequences of: 1) performing fewer trials per individual (on more individuals) or more trials per individual (on fewer individuals) for a fixed amount of trial survey effort; 2) changing methods to select trial distances; 3) changing total trial survey effort; 4) ignoring underlying group heterogeneity; 5) ignoring a behavioral response to trapping, and 6) ignoring the repeated measures aspect of the data collection process by fitting a standard generalised linear model as the detection function (instead of a generalised linear mixed model). The simulation studies are summarized in Table 6.6.

When all sources of underlying heterogeneity in the population are modelled in the detection function, with greater levels of survey effort, the bias in  $\hat{N}_2$  becomes small, regardless of which detection function scenario and method of selecting trial distances was used. However,  $\hat{N}_3$  remains positively biased. With decreasing probability of detection, variability about  $\hat{N}_2$  and  $\hat{N}_3$  increased, but is generally lower when the “Adaptive” method is used to select trial distances. However, the total survey effort (and cost) required to achieve unbiased estimates is high, and in some scenarios, greater bias might be tolerated for significant cost savings. For example, if survey effort needed to be quadrupled

Table 6.6: Summary of input parameters for the five simulation studies. Three detection function scenarios were used (“High”, “Medium”, and “Low”) in Simulations 1, 2 and 5, only the “Medium” scenario was used in Simulation 3, and in Simulation 4, two additional scenarios were used that included a behavioural effect (High-B and Low-B). The fitted detection function (Model) used in the analysis was either correct or not. The number of trials in the trials survey was either 360 (Simulation 1, 2, 4 and 5), or 780, 1080 and 1440 (Simulation 3). The method used to select the trial distances was either all of those considered (i.e., “Uniform”, “Stopping rule 5”, “Stopping rule 8” and “Adaptive”; Simulation 1 and 2) or just “Adaptive” (Simulation 3, 4 and 5).

Simulation	Scenario	Model	Trial sample size	Trial distance method
1	High Medium Low	Correct	360	All
2	High Medium Low	Incorrect	360	All
3	Medium	Correct Incorrect	780 1080 1440	Adaptive
4	High High-B Low Low-B	Correct Incorrect	360	Adaptive
5	High Medium Low	Incorrect	360	Adaptive

to achieve a reduction in bias from 2% to 1.5%, the higher bias may be tolerated, especially if management actions would not change with the two different population estimates.

When underlying heterogeneity in the population was not modelled in the detection function,  $\hat{N}_2$  was negatively biased, regardless of which detection function scenario and method of selecting trial distances was used. However,  $\hat{N}_3$  can be nearly unbiased in some circumstances tested in this simulation study

when survey effort was high. With decreasing probability of detection, variability about  $\hat{N}_2$  and  $\hat{N}_3$  increases, but was generally lower when the “Adaptive” method was used to select trial distances. Doubling, tripling and even quadrupling survey effort provided marginal return with regards to increased precision but did reduce bias substantially. When survey effort increased from 360 to 720 trap nights, the magnitude of the bias decreased the most. After survey effort exceeded 760 trap nights, the return in investing more survey effort decreased.

Ignoring a behavioral response to the trapping experience typically increased variability in  $\hat{N}_2$ , and decreased variability in  $\hat{N}_3$ , especially with lower average probability of detection. Percentage bias of  $\hat{N}_2$  and  $\hat{N}_3$  increased when a behavioral response to the trapping experience was ignored, except in the “Low” detection function scenario when percentage bias in  $\hat{N}_3$  decreased. Explicitly modelling a behavioral response in the detection function might be possible (e.g., similar to behavioral models developed for capture-recapture studies) and should be investigated. In addition, a goodness-of-fit test could be developed to investigate whether a behavioral response in the data collected using the TPPT method is present (e.g., Chao et al. 2000). Ignoring the repeated measures aspect of the data and treating all trial observations as independent was found to severely impact abundance estimates, especially when the underlying heterogeneity in the population was not accounted for.

In reality there is no way of knowing if all the relevant covariates to account for heterogeneity have been included in the detection function. There is scope to expand the detection function presented here to be a mixture model (Pledger, 2000; Pledger et al., 2010; Morgan and Ridout, 2009) that might help identify if underlying heterogeneity is present in the data. Nevertheless, great effort should be made by: 1) the field staff collecting the data to ensure potential sources of heterogeneity are recorded (e.g., sex of the animal), and 2) the statistician

analysing the collected data should interpret results of both  $\hat{N}_2$  and  $\hat{N}_3$ , and model the data appropriately (e.g., use a generalised linear mixed model, as presented here, or a generalised additive mixed model).

The level of Monte Carlo error in the simulations might explain some of the small differences observed in bias estimates when simulation scenarios changed. For example, in Table 6.4 on page 140 the percentage bias in  $\hat{N}_2$  for the “High” scenario when the random effect was included, or not, changed from 4.26% to 3.34%. Similarly small differences were observed for the “Medium” and “Low” scenarios. Also, plots of the estimated bias in  $\hat{N}_2$  and  $\hat{N}_3$  when heterogeneity was accounted for and underlying detection probability was low (Figure 6.9 on page 112) was not smooth – as would be expected if Monte Carlo error was minimal. Monte Carlo error tended to be worse when average detectability was low and between-individual heterogeneity was large. In such instances, one single estimate of population size could have been badly biased and this caused the bias and associated variance estimates to be very large. Increasing the number of simulations undertaken could have reduced this issue somewhat.

### 6.8.1 Recommendations

Some broad recommendations can be made using the results of this simulation study that may help guide researchers who might wish to implement the trapping point transect method for their own study. If underlying detectability of the study species is “Medium” or “High”, with little between individual variation, conducting 18 trials per individual and radio collaring 20 individuals should be a minimum investment of survey effort to obtain satisfactory levels of uncertainty and bias in  $\hat{N}_2$  and  $\hat{N}_3$  when underlying heterogeneity is accounted for. With decreasing detectability and increasing between-individual variation, total trial survey effort must be increased. Also in these situations of low detectability and large between-individual variation, the bias is larger in  $\hat{N}_3$  than  $\hat{N}_2$ .

Attempts should always be made to ensure causes of heterogeneity in capture probability (e.g., males versus females) are represented in the proportion of the population on which trial surveys are conducted. If underlying heterogeneity is not accounted for, for example by measuring the relevant covariates at time of capture and including them in the detection function model, the  $\hat{N}_3$  abundance estimator might be less biased than  $\hat{N}_2$ , but this depends on the amount of underlying heterogeneity ( $\hat{N}_3$  is less biased than  $\hat{N}_2$  when underlying detectability is low and between-individual heterogeneity is high) and the number of trials performed on each individual. Modelling the repeated measures aspect of the data collection process should always be done, using e.g., a generalized linear or generalised additive mixed models.

As per any well-designed monitoring program, collected data should be analysed intermittently during the field season (e.g., using the “Adaptive” method to select trial distances), to ensure the setting of trial distances will adequately capture the true shape of the detection function. It might also be possible to use trial survey data to investigate behavioral effects in response to the trapping experience, given that the method has only been applied in one field situation to date (see next chapter).

It is also recommended that a simulation study be undertaken to investigate the amount of bias present in abundance estimates given the survey effort employed, and to determine the level of survey effort required to obtain population estimates with acceptable levels of bias.

## Chapter 7

# Estimating Key Largo woodrat abundance using a trapping point transect approach

### 7.1 Introduction

In Chapters 3 and 4, a large survey effort (i.e., more than 18,000 trap nights over the course of a year) was used to obtain estimates of population size for the Key Largo woodrat using capture-recapture approaches. As briefly described in Section 2.6, and fully detailed in Chapters 5 and 6, the trapping point transect (TPT) approach is a new abundance estimation method, that potentially allows abundance of a species to be estimated with fewer resources than conventional trapping studies (Buckland et al., 2006). Given the issues with estimating Key Largo woodrat abundance using conventional approaches, the TPT method warranted a field-based investigation to assess its effectiveness (see Section 2.9).

A pilot study with minimal survey effort was undertaken between February and April 2007 to determine the practicality, feasibility and cost effectiveness of using the TPT method to estimate Key Largo woodrat abundance. In the *main survey*, a trapping grid of 137 sampling points with a 250 m trap spacing was used, with each sampling point trapped for three consecutive nights. During the *trial survey*, only 33 individual trials were conducted on two individual woodrats

(one male and one female). Due to the small sample size, estimation of the detection function could not be undertaken with any certainty and abundance could not be estimated.

However, in terms of practicality, the TPT method demonstrated potential should greater effort be allocated to the trials survey. Hence, two full-scale surveys were undertaken between February and April in 2008 and 2009. The design of the main survey remained the same as in the pilot study<sup>1</sup>, but a greater number of individual woodrats were radio-collared in the trial survey.

The data collection methods, analysis and results from the 2008 and 2009 field seasons are presented below. I also conduct a small simulation study to investigate how both the trial and main survey designs can be changed to increase precision and decrease bias in abundance estimates.

## 7.2 Methods

### 7.2.1 Survey design

In the *main survey*, a randomly-placed systematic grid of 156 sample points with a 250 m trap spacing was established throughout the suitable habitat (Figure 7.1). The large trap spacing was used so that detection events at neighbouring sample points could be assumed independent. The distance was chosen based on information on the home range size of Key Largo woodrats (see Chapter 1, Section 1.2 and references therein).

Of these 156 sample points, 137 points could actually be surveyed. At the other 19 sample points, either the vegetation was too thick to survey safely, or the point had been misclassified as being in suitable habitat. The habitat misclassification problem occurred due to small inaccuracies in the boundary of the habitat map (Figure 7.1), where points were classified as hammock (habitat)

---

<sup>1</sup>One survey point was removed from the main survey design in 2009 due to habitat modification.

when they were actually mangrove swamp (non-habitat). Also, one main survey sample point in 2009 was removed from the survey due to habitat modification, so 136 main survey sample points were trapped in that year (Figure 7.1). Two traps were placed at each trap location and were set for three consecutive nights.

Due to logistical constraints, all main survey points could not be trapped concurrently. Consequently, the main survey was divided into segments, each containing approximately 20 main survey points; it took approximately eight weeks to completely trap all main survey sampling points.

To conduct the *trial survey*, VHF radio-collars (mass <10g, AVM Instruments, California) were attached to a subset of woodrats captured during the main survey and from additional targeted sampling. During targeted sampling, approximately fifty traps were set in areas near known nest locations of woodrats, where woodrats had been previously caught, or where occupancy was unknown. Targeted trapping typically lasted one or two nights, and captured woodrats were radio-collared.

Since the Key Largo woodrat is nocturnal, each radio-collared woodrat was located at its nest during the day time. An individual trial was conducted on each radio-collared woodrat, by setting two trial traps (i.e., the same number of traps per point as the main survey) at some pre-determined distance and random direction away from the nest. Each woodrat was exposed to trial surveys at various distances. The aim was to perform thirty trials on each individual (5 m intervals between 5 and 50 m). However, since some trials were performed at very long distances for which capture was believed to be extremely unlikely (up to *c.* 320 m), to ensure the tail of the detection function would be accurately estimated and all woodrats were exposed to numerous, very short trial distances (1 m), so the collar could be retrieved, an equal number of trials per individual could not be conducted.

After the trial trap was set, the following morning the traps were checked to see if the radio-collared individual was captured. If the individual was captured, a "1" was recorded for that particular trial distance, and the radio-collared woodrat was released at the point of capture. Trials did not typically recommence on recaptured woodrats for 2 nights (a rest period). If the individual was not captured, a "0" was recorded and the woodrat was re-located to its current nest location and a new trial was established. An individual's nest location may have changed between trials, since woodrats maintain multiple nests within their home ranges (Hersh, 1981).

To abide by permitting regulations, radio-collars were not attached to individual woodrats for more than one month. Permit regulations also required that no trapping took place when the overnight low temperature was forecast to be  $< 60^{\circ}\text{F}$ , so trial surveys could not necessarily be performed every night the collar was attached to a woodrat. Since trapping in the main survey also ceased when the overnight low temperature was forecast to be  $< 60^{\circ}\text{F}$ , no bias in population estimates arose due to weather conditions.

To ensure survey design was consistent between the main and trial surveys, trial traps were not placed less than 250 m apart (the same trap spacing as the main survey), even if multiple radio-collared woodrats were in the same area. In instances where radio-collared woodrats were located within 250 m of one another, one trial trap location was used for all radio-collared woodrats within the area.

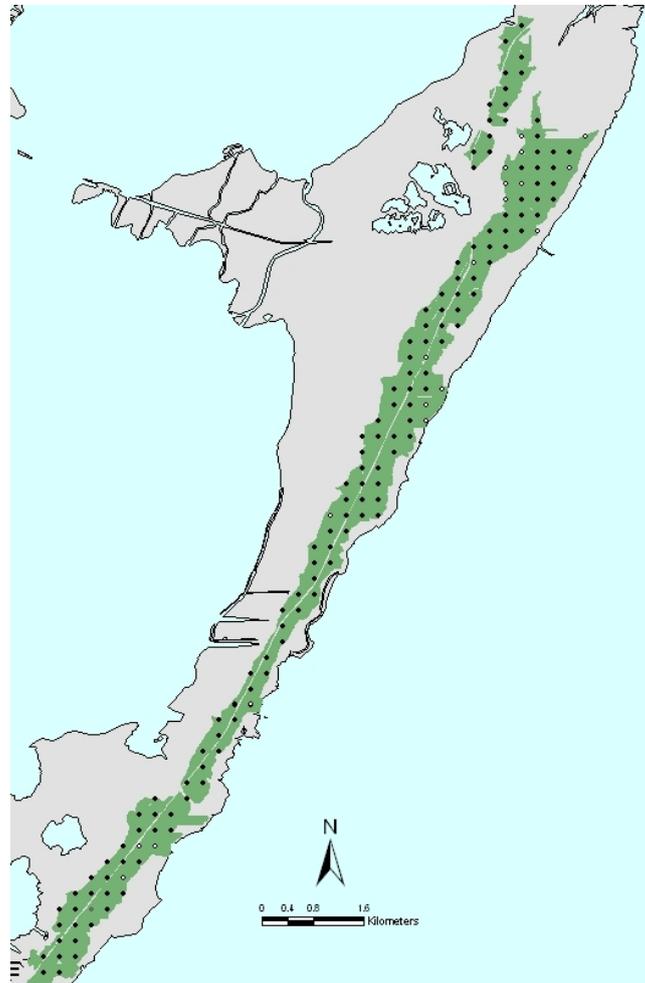


Figure 7.1: Map of the main survey. 156 sample points were distributed throughout the suitable habitat (shaded green) based on a randomly-placed systematic grid with a 250 m trap spacing. Nineteen sample points (white circle with black outline) could not be surveyed and one point (shaded grey) was not surveyed in 2009. Traps were not placed in unsuitable habitats of water and mangrove swamp (shaded blue and grey, respectively).

### 7.2.2 Trapping and handling

At each main and trial survey trapping point, two vented Sherman traps with raccoon-proof door latches (approximately 10.2 x 11.4 x 38.1 cm in size, model PXLF15, H. B. Sherman Traps Inc., Tallahassee, Florida) were placed. Traps were opened and baited with whole rolled oats in the late afternoon and checked the following morning.

All woodrats captured in the three day period were double-marked with passive integrated transponder (PIT) tags (AVID, Norco, California) and #1005 Monel ear tags (National Band and Tag, Newport, Kentucky). Sex and weight were also recorded. All bycatch was released at point of capture.

### 7.2.3 Fitting the detection function

Since multiple trials were performed on individual radio-collared woodrats, each trial observation could not be considered independent. The detection function was specified using a generalized linear mixed effects model (GLMM, Brown and Prescott 2004, c.f. Section 5.3.1 on page 79). Distance from trap, sex of woodrat, and year of survey were covariates considered during model selection (Table 7.1).

Table 7.1: Definitions of explanatory variables used in estimating the detection function for the trapping point transect analysis.

Variable	Definition
Dist	Distance (in metres) between the trial start point (a radio-collared woodrat in its nest) and end point (trial trap location).
Sex	Factor variable indicating sex of woodrat (0=female, 1=male).
WR	Individual woodrat identifier.
Year	Factor variable specifying in which year the trial survey was undertaken: 2008 or 2009.

Three random effects structures were available for investigation (Table 7.2). For each structure, forward model selection was used, whereby the simplest fixed effects model was fitted (i.e., a null model) and each additional explanatory

variable not already in the model was added one at a time. Akaike's Information Criterion (AIC) was used for model selection. Each time a term was added to the model, the AIC was calculated, and this process was repeated until no additional terms could be added to the model without increasing the value of the AIC. In cases where the difference in AIC between two competing models was less than two (Burnham and Anderson, 2003), both models were investigated.

All GLMM modelling was completed using the "glmer" function within the "lme4" package (v0.999375-34, Bates and Maechler 2009) of the statistical free-ware software R (v2.11.0, <http://www.r-project.org/>). Since the likelihood of a GLMM cannot be evaluated analytically (except in the special case of a linear mixed model with a normally distributed response variable, Komarek and Lesaffre 2008), maximum likelihood estimates of the parameters in the detection function were approximated using adaptive Gaussian quadrature (Rabe-Hesketh et al., 2002; Bolker et al., 2009). Greater accuracy in estimating the likelihood can be achieved by using a larger number of quadrature estimation points, although this comes at the expense of greater computation time (Rabe-Hesketh et al., 2002). Also, due to the optimization algorithm used within the "glmer" function, numerical instability arose with a very large number of points (i.e., >50 points, pers. comm. Martin Maechler). It has been suggested that between 5 and 10 quadrature points is adequate to estimate the likelihood with sufficient accuracy for most applications (Lesaffre and Spiessens, 2001). In my analysis, likelihood values were stable (i.e., to within 1 decimal place) when using between 5 and 30 quadrature points. To decrease computation time in fitting the models, 10 quadrature points were used for all models fitted.

Table 7.2: The three random effects structures considered during model selection, presented using R notation

R-like notation	Definition
(1 WR)	The detection function for each individual woodrat included a random intercept term, drawn from a common mean intercept value, with associated variance, applicable to all woodrats. All individuals share a common slope.
(1+Dist WR)	The detection function for each individual woodrat included a random intercept and random slope term, drawn from a common mean intercept and slope, with associated variances, applicable to all woodrats. The values of the random intercept and slope are correlated.
(0+ Dist  WR)+(1 WR)	The detection function for each individual woodrat included a random intercept and random slope term, drawn from a common mean intercept and slope, with associated variances, applicable to all woodrats. The values of the random intercept and slope are uncorrelated.

#### 7.2.4 Estimating abundance

As described in Section 5.3.2 on page 81, two abundance estimators are available:

$$\hat{N}_2 = \frac{A}{A_c} \sum_{i=1}^n \frac{1}{E_b[\hat{P}(z_{i1}, \dots, z_{iJ})]} \quad (7.1)$$

$$\hat{N}_3 = \frac{A}{A_c} \sum_{i=1}^n E_b \left[ \frac{1}{\hat{P}(z_{i1}, \dots, z_{iJ})} \right] \quad (7.2)$$

where  $A$  is the area of the survey region,  $A_c$  is the area of the covered region ( $A_c = vK\pi w^2$ , where  $v$  is the number of visits to the total number of sample points  $K$  in the main survey, and  $w$  is the truncation distance equal to 125 m), and  $\hat{P}(z_{i1}, \dots, z_{iJ})$  is the estimated probability of detecting the  $i$ th animal captured in the main survey unconditional on its distance from the trapping sample point:

$$\hat{P}(z_{i1}, \dots, z_{iJ}) = \int_0^w \pi(r) \cdot \hat{p}(r, b; z_{i1}, \dots, z_{iJ}) dr \quad (7.3)$$

Observed values of each explanatory variable  $z_{i1}, \dots, z_{iJ}$  were substituted into the fitted detection function for each woodrat captured in the main survey. The availability function,  $\pi(r)$ , can be assumed equal to  $2r/w^2$  in the absence of edge effects. However, the study region is very long and narrow, so ignoring edge effects could cause bias in the abundance estimates. Instead,  $\pi(r)$  was calculated (*sensu* Buckland, 2006):

$$\pi(r) = \frac{rq(r)}{\int_0^w rq(r) dr} \quad (7.4)$$

where  $q(r) = [\sum_{k=1}^K q_k(r)]/K$ , for  $0 \leq r \leq w$  and  $q_k(r)$  is the proportion of the circumference of a circle of radius  $r$  centered on the point  $k$  that lies within the survey region, and  $K$  is the number of points in the main survey. If the proportion  $q_k(r)$  is always equal to one, then  $\pi(r) = 2r/w^2$ . I estimated  $q(r)$  by calculating the proportion of 10 points evenly distributed between 0 and  $w$  m (i.e., every 12.5 m), laid on the four cardinal (N, S, E, W) and four intercardinal (NE, SE, SW, NW) directions of the compass that were habitat. This approach meant that 80 surrounding points were checked for each main survey trap location, to determine if they fell in habitat or non-habitat (see Figure 7.2). A logistic regression model was then fitted to define habitat availability as a function of distance from the main survey trap location.

Overall woodrat abundance was calculated using model averaging, where abundance estimates were weighted using AIC weights (Buckland et al., 1997; Burnham and Anderson, 2003; Bolker et al., 2009) for each fitted detection function.

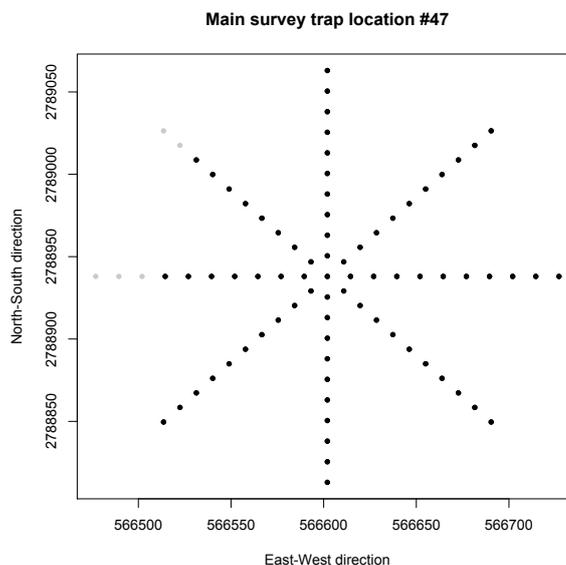


Figure 7.2: Plot of 80 habitat sample points centered around trap location #47 in the main survey. This trap was located on the western boundary of the survey region, such that five points did not occur within habitat (grey) and 75 did (black).

### 7.2.5 Estimating variance in abundance

A non-parametric bootstrap approach was used to estimate the variance in abundance, accounting for the variance associated with encounter rate in the main survey and variance of the estimated detection function (described in Section 5.3.3, page 83). A full model selection procedure could not be undertaken on each bootstrap resample of the trial survey data due to computational time. Instead, a subset of all fitted models was refitted to each bootstrap sample. The model subset contained only models where the AIC differed by less than two from the best model<sup>2</sup> (Burnham and Anderson, 2003) when fitted to the original trial survey data set. Once each model within the subset was fitted to

<sup>2</sup>Alternatively, the model subset could have been chosen to include all of the highest weighted models up to e.g. a cumulative model weight of 0.95.

the bootstrap sample, the model with the lowest AIC was chosen as the detection function. Abundance estimation was conducted as per methods detailed in the previous section. The resampling procedure was repeated 999 times, and confidence intervals were estimated using the quantile method.

### 7.3 Simulation study

In Chapter 6 I used a simulation study to investigate how the number of trials per individual vs. the number of individuals upon which trials were performed influenced bias and precision in the abundance estimates when underlying between-group heterogeneity was accounted for, or not. However, I did not investigate the trade-off that occurs between investing survey effort in the trials survey versus the main survey (e.g., by increasing the number of trapping points in the main survey). Here, I conduct a small simulation study to investigate:

1. given the level of survey effort invested in the 2008 and 2009 surveys, could effort be allocated differently (e.g., visit each main survey point only once, but radio-collar more woodrats?) to achieve better precision?
2. if survey effort was doubled or tripled, how should it be allocated to achieve greater precision and less bias in the abundance estimates?

The simulation study was conducted in a similar manner to that presented in Chapter 6. That is, a simple mean detection function (where the probability of capture depended only on the sex of individual, its distance to the trap when the trap was set and a random effect variance for each individual) was specified with parameters estimated using the woodrat data collected in Section 7.2.3. Each animal in the main and trial survey was allocated a unique two-parameter detection function drawn from this mean detection function, using a specified between-animal variance.

For each simulation, a population of 600 individuals (with an equal sex ratio) was randomly generated on a rectangular survey area (dimensions: 3,000 by 4,300 m). A main survey trapping grid of 10 by 15 traps (in the X and Y dimension, respectively) with a 250 m trap spacing was overlaid onto this survey area (Figure 7.3), and two traps were placed at each trap location. The distance between each woodrat and its closest main survey trap location was calculated. After selecting individuals at random, the simulation checked the number of “open” traps at the the nearest trap location (i.e., the number of traps that hadn’t caught another individual). If all of the closest traps were shut, that animal could not be caught, and a new animal was selected. If the number of closest traps that were open was either 1 or 2, a Bernoulli trial was performed to determine if the individual was detected, given its distance to the nearest trap and a probability specified by the individual’s unique detection function parameters. The resulting simulated data from the main survey was the number of male and female woodrats detected in the main survey. Figure 7.3 provides an example of the main survey simulation.

For the trial survey, the number of trials per individual was fixed at 18, but the number of individuals in the trial survey was varied. Eighteen trials per individual was found to be the approximate number in Chapter 6 beyond which there was little gain in precision. In addition, this was approximately the average number of trials performed per individual woodrat in the trial survey undertaken in 2008 and 2009 (Section 7.4.1). Trial distances were selected using the “Adaptive” method outlined in Section 6.2.3 on page 94. Abundance and variance estimation proceeded as detailed in Chapter 5 (page 75).

In total, five scenarios were simulated, where for a fixed number of trap nights for the total survey (i.e., the main and trial surveys, combined) the number of visits to each trap location in the main survey, and the total number of woodrats

in the trial survey, were varied (Table 7.3).

Table 7.3: The five simulation scenarios used to investigate the consequences of changing the number of visits to each trap location in the main survey (*Main survey visits*) and the number of woodrats upon which trials were performed (*No. woodrats*), for a fixed amount of total survey effort. For example, in Scenario 1, main survey trap effort equalled 900 trap nights (150 trap locations  $\times$  2 traps per location  $\times$  3 visits to each trap location), and the trials survey trap effort equally 540 trap nights (15 individuals  $\times$  18 trials per individual  $\times$  2 traps per location).

Scenario	Total trap effort	Main survey visits	Main survey trap effort	No. woodrats	Trials survey trap effort
1	1440	3	900	15	540
2	1452	1	300	32	1152
3	2880	6	1800	30	1080
4	2880	3	900	55	1980
5	4140	12	3600	15	540

## 7.4 Results

### 7.4.1 Key Largo woodrat survey and abundance estimation

In 2008, there were 19 capture events of 14 woodrats (six females, eight males) at 13 distinct sample points in the main survey. In the trial survey, 119 and 69 individual trials were conducted on eleven female and four male woodrats, respectively. There were 42 successful trials, at distances ranging between 1 and 41 m. Unsuccessful trial distances ranged between 1 and 239 m.

In 2009, there were six capture events of six unique woodrats (four females, two males) at six sample points in the main survey<sup>3</sup>. In the trial survey, 169 and 155 individual trials were conducted on seven female and six male woodrats, respectively. There were 33 successful trials, at distances ranging between 1 and 60 m. Unsuccessful trial distances ranged between 1 and 319 m.

In total, 512 individual trials were conducted on 10 male and 18 female woodrats in 2008 and 2009, of these 75 were successful. The female woodrats

<sup>3</sup>N.B. only 136 trap locations were surveyed in 2009.

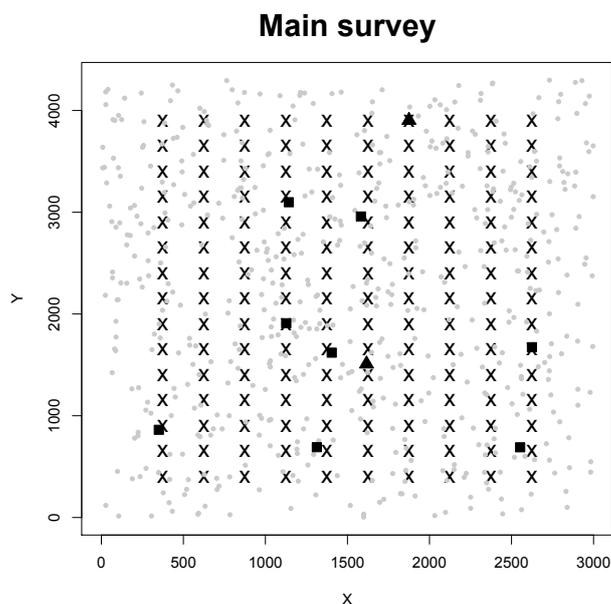


Figure 7.3: A single realisation of the main survey simulation, based on the estimated detection function using woodrat data. Traps are represented as an “x”, while detected females are shown as a solid square, and males as a solid triangle. Undetected individuals are shown as a small solid grey circle.

comprised 288 trials (48 were successful) and the male woodrats comprised 224 trials (27 were successful).

After completing forward model selection on the full trial survey data (512 trials), 27 different detection function models were fitted (Table 7.4). There was one detection function model (model 23) with a difference in AIC of less than 2 from the best model (model 27), and these two models had a combined AIC weight of 0.63 (c.f. Table 7.4). These two models retained distance of trial (as was expected) and year as explanatory variables, but the model with the lowest AIC (model 27) also included sex and an interaction term between sex and year. Both these models included a random effect on the intercept and slope parameters, but they were assumed to be uncorrelated (Table 7.2). Plots of these two detection functions are provided in Figure 7.4. For these two models,

detectability in 2009 was lower than 2008, and the detectability of male woodrats decreased substantially between 2008 and 2009.

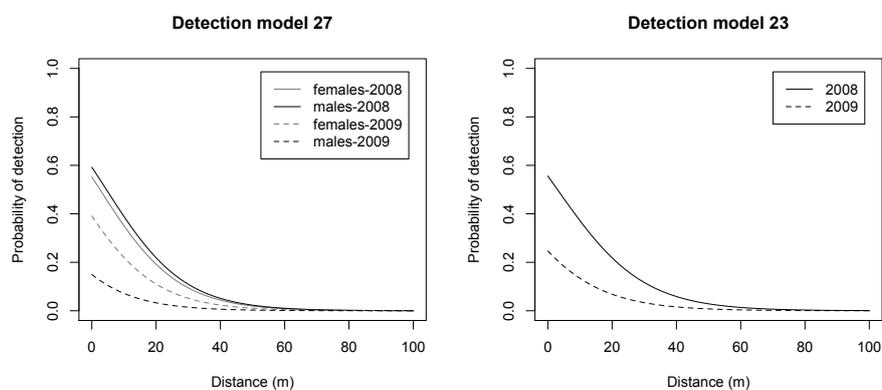


Figure 7.4: Plots of the two mean detection functions with the lowest AIC (Table 7.4) fitted to trial data collected in 2008 (solid line) and 2009 (dotted line). Model 27 included sex and an interaction term between sex and year as a covariate (males and females are shaded black and grey, respectively). See Table 7.4 for full model structure.

Table 7.4: The 27 different detection functions fitted to the 512 trial observations collected on male and female woodrats in 2008 and 2009. The response variable was binomial (capture vs. no capture) and the explanatory variables were distance of trial (Dist), year of survey (Year), sex (Sex) and woodrat (WR) (see Table 7.1 for full explanatory variable definitions). Models are ordered by AIC value. Only one model had a difference in AIC of  $< 2$  from the best model.

No	Parameters	logLik	df	AIC	$\Delta$ AIC	AIC <sub>w</sub>
27	Dist + Year * Sex + (0 + Dist WR) + (1 WR)	-162.51	7	339.03	0	0.37
23	Dist + Year + (0 + Dist WR) + (1 WR)	-164.87	5	339.73	0.7	0.26
19	Dist + Year + Hab + (1 + Dist WR)	-161.74	9	341.47	2.44	0.11
16	Dist + Year + (1 + Dist WR)	-164.83	6	341.66	2.63	0.1
24	Dist * Year + (0 + Dist WR) + (1 WR)	-164.86	6	341.72	2.69	0.1
17	Dist * Year + (1 + Dist WR)	-164.83	7	343.65	4.63	0.04
21	Dist + Sex + (0 + Dist WR) + (1 WR)	-168.15	5	346.31	7.28	0.01
22	Dist * Sex + (0 + Dist WR) + (1 WR)	-167.39	6	346.79	7.76	0.01
25	Dist + Hab + (0 + Dist WR) + (1 WR)	-166.57	7	347.14	8.11	0.01
20	Dist + (0 + Dist WR) + (1 WR)	-169.69	4	347.37	8.34	0.01
18	Dist + Hab + (1 + Dist WR)	-166.51	8	349.02	9.99	0
15	Dist + (1 + Dist WR)	-169.64	5	349.28	10.25	0
26	Dist * Hab + (0 + Dist WR) + (1 WR)	-165.73	10	351.45	12.42	0
13	Dist + Hab + Year + (1 WR)	-172.8	7	359.59	20.56	0
8	Dist + Year + (1 WR)	-176.09	4	360.18	21.15	0
12	Dist + Sex + Year + (1 WR)	-175.33	5	360.66	21.63	0
9	Dist * Year + (1 WR)	-175.98	5	361.95	22.93	0
14	Dist * Hab + Year + (1 WR)	-171.13	10	362.26	23.23	0
7	Dist * Sex + (1 WR)	-178	5	366	26.97	0
6	Dist + Sex + (1 WR)	-179.26	4	366.52	27.5	0
10	Dist + Hab + (1 WR)	-177.3	6	366.6	27.57	0
2	Dist + (1 WR)	-180.52	3	367.04	28.01	0
11	Dist * Hab + (1 WR)	-176.07	9	370.15	31.12	0
3	Year + (1 WR)	-207.16	3	420.33	81.3	0
1	(1 WR)	-211.88	2	427.75	88.73	0
4	Sex + (1 WR)	-211.24	3	428.47	89.45	0
5	Hab + (1 WR)	-210.64	5	431.29	92.26	0

The interaction between sex and year when estimating the detection function (Table 7.4) was surprising: I had no prior expectation that the relative detectability of male and female woodrats would vary differently by year. Since only four male woodrats were radio-collared during the 2008 field season, perhaps the sample was too small to adequately represent the detectability of male woodrats. Consequently, the significance of the interaction term might be an artefact of the small sample size, rather than a true situation where detectability of male and female woodrats show different changes between years. Therefore, data collected on male and female woodrats in the trial and main survey were analysed separately.

For the female woodrat analysis, a full forward model selection procedure could be undertaken on the 288 trials. In total, 15 models were fitted to the trials data collected on 18 female woodrats. Three models had a difference in AIC of less than 2 (Table 7.5). Detectability remained relatively constant between the years, with a slight decrease in detectability in 2009 (Figure 7.5). For the male woodrat analysis, due to small sample sizes, data were pooled across year and only a simple detection function could be fitted that assumed detectability did not change over year or habitat type (i.e.,  $\text{Capture} = \text{Dist} + (1|\text{WR})$ , Figure 7.6).

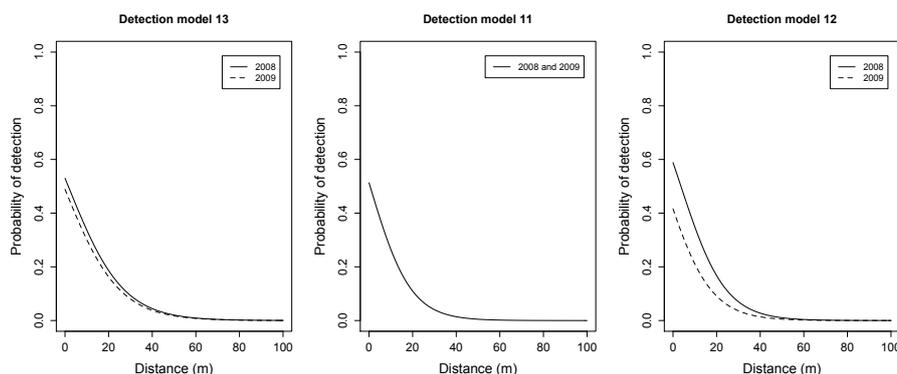


Figure 7.5: Mean detection function plots corresponding to the three models with the lowest AIC fitted to the female trial data collected in 2008 (solid line) and 2009 (dotted line). See Table 7.5 for full model structure.

Table 7.5: The 15 different detection functions fitted to the 288 trial observations collected on female woodrats in 2008 and 2009. The response variable was binomial (capture vs. no capture) and the explanatory variables were distance of trial (Dist), year of survey (Year) and woodrat (WR) (see Table 7.1 for full explanatory variable definitions). Models are ordered by AIC value. Two models had a difference in AIC of  $< 2$  from the best model. Coefficients of the best model were: 0.12 - 0.080 Dist - 0.079 Year2009 -0.086 Dist\*Year2009, with random effects on the intercept and slope distributed as follows:  $N(0, 0.256)$  and  $N(0, 0.003)$ .

No	Parameters	logLik	df	AIC	$\Delta$ AIC	AIC <sub>w</sub>
13	Dist * Year + (0 + Dist WR) + (1 WR)	-95.44	6	202.87	0	0.36
11	Dist + (0 + Dist WR) + (1 WR)	-97.58	4	203.15	0.28	0.32
12	Dist + Year + (0 + Dist WR) + (1 WR)	-96.86	5	203.72	0.85	0.24
14	Dist + Hab + (0 + Dist WR) + (1 WR)	-96.02	7	206.04	3.17	0.07
15	Dist * Hab + (0 + Dist WR) + (1 WR)	-95.27	10	210.54	7.67	0.01
6	Dist * Year + (1 WR)	-103.37	5	216.73	13.86	0
5	Dist + Year + (1 WR)	-104.37	4	216.74	13.87	0
2	Dist + (1 WR)	-105.69	3	217.39	14.52	0
9	Dist + Hab + Year + (1 WR)	-102.22	7	218.44	15.57	0
7	Dist + Hab + (1 WR)	-103.8	6	219.6	16.73	0
10	Dist * Hab + Year + (1 WR)	-101.41	10	222.82	19.95	0
8	Dist * Hab + (1 WR)	-103.12	9	224.23	21.36	0
3	Year + (1 WR)	-128.67	3	263.34	60.47	0
1	(1 WR)	-130.77	2	265.54	62.67	0
4	Hab + (1 WR)	-130.14	5	270.29	67.42	0

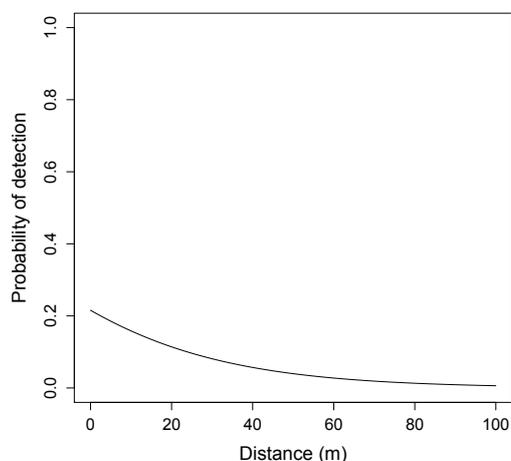


Figure 7.6: Plot of the simple mean detection function fitted to the male trial data, where the probability of detection only depended on distance to trial trap and an individual random effect.

In order to calculate the probability of detecting each woodrat captured in the main survey, the detection function must be multiplied by the availability function (Equation 7.3). The availability function is plotted in Figure 7.7, and shows that by approximately 125 m of each main survey location, the amount of available habitat around each main survey point decreased (e.g., habitat was mangrove rather than hammock). Using the Horvitz-Thompson-like estimator presented in Equations 7.1 and 7.2, abundance was calculated for female and male woodrats in 2008 and 2009 (Table 7.6). There was a large difference in abundance estimates, depending on which estimator ( $\hat{N}_2$  or  $\hat{N}_3$ ) was used. Results from simulation studies undertaken in Chapter 6, suggest that bias in  $\hat{N}_2$  is much lower than  $\hat{N}_3$ , given adequate sample sizes.

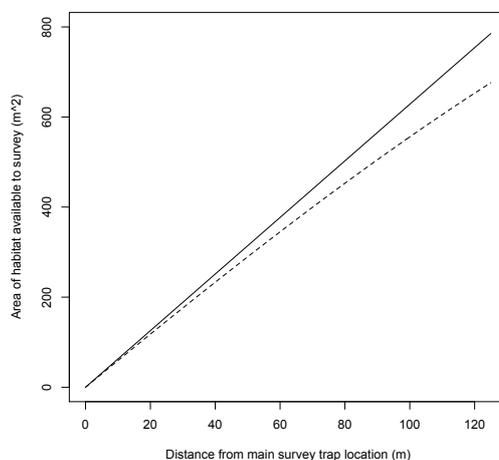


Figure 7.7: Plot of the availability of habitat with increasing distance from the main survey point when edge effects are accounted for (dashed line), or not (solid line).

Table 7.6: Abundance estimates and 95% percentile intervals for female and male woodrats in 2008 and 2009 using two different abundance estimators. For consistency, the percentile intervals were reported for male woodrats in 2009 as 0, but these are clearly biased low, and perhaps the minimum number of male woodrats captured in 2009 would better reflect the uncertainty in population size estimates.

Population	$\hat{N}_2$	$\hat{N}_3$
Females in 2008	185 (54, 689)	868 (143, 11663)
Males in 2008	523 (135, 1402)	1099 (380, 4190)
Females in 2009	144 (32, 531)	680 (82, 7814)
Males in 2009	83 (0, 296)	175 (0, 842)

### 7.4.2 Simulation study

A simple mean detection function curve for each sex was used in the simulation study (model 6 in Table 7.4, Figure 7.8), where detection function coefficients were estimated using the woodrat data collected in Section 7.2.3. The model was chosen for its simplicity: it did not contain spurious interaction terms and provided an adequate representation of the detectability process.

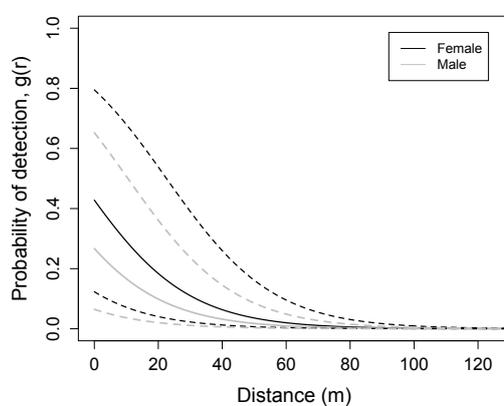


Figure 7.8: Plot of the detection function scenario used to simulate unique detection functions for each woodrat in the main and trial surveys. The solid line indicates the mean detection function, and the dashed lines indicate the 2.5th and 97.5th percentiles (as calculated based on 10,000 samples of the random effects distribution). The mean model for males and females is shown as grey and black, respectively.

When the number of individuals in the trials survey was low (i.e., 15), regardless of how many times each main survey trap location was visited, estimates of  $\hat{N}_2$  and  $\hat{N}_3$  were severely positively biased (Table 7.7). As was seen in Chapter 6, when detectability was low, the estimated probability of detection can be severely biased low and abundance was biased high as a result. When the number of individuals in the trials survey was high (30), the relative bias in  $\hat{N}_2$  was approximately +12%, regardless of how many times each point in the main survey was visited. However, the precision in abundance estimator  $\hat{N}_2$  was higher when each main survey point was visited six times (95% confidence interval was (356, 1277)) compared to visiting each point only once (95% confidence interval was (181, 1456), Table 7.7). When the number of woodrats in the trial survey was extremely high (i.e., 55 individuals), bias in  $\hat{N}_2$  was the lowest of all scenarios tested (+5.8%, Table 7.7).

Similar results were found in the abundance estimator  $\hat{N}_3$ , but estimates remain positively biased (magnitude greater than  $\hat{N}_2$ ) even when total survey effort was increased (Table 7.7). When the number of woodrats in the trials survey was high further increases in the number of trails made little difference to the percentage bias of  $\hat{N}_3$  (when the number of woodrats was 30 and 55, the percentage bias in  $\hat{N}_3$  was 9.1% and 8.3%, respectively), but the precision tended to be slightly higher (Table 7.7).

## 7.5 Conclusion and Discussion

Using the TPT method presented in this chapter, abundance of female woodrats using  $\hat{N}_2$  and  $\hat{N}_3$  may have very slightly decreased between 2008 and 2009, but there was a large amount of uncertainty in the abundance estimates (CV of  $\hat{N}_2$  in 2008 and 2009 and 66.0% and 67.1%, respectively). Estimated abundance of male woodrats appears to have declined between 2008 and 2009, however

Table 7.7: The results for  $\hat{N}_2$  and  $\hat{N}_3$  abundance estimators when total survey effort is increased and allocation between the main and trials survey changes, by varying the number of times each trap location in the main survey was visited (*Visits*) and the number of individuals that trials were performed on (*WR*). True abundance in the main survey was fixed at 600 individuals (equal sex ratio). Eighteen trials were conducted per individual, and there were either 15 (Model 1 and 5), 30 (Model 3), 32 (Model 2) or 55 (Model 4) individuals in the trials survey.

Model	Visits	WR	mean( $\hat{N}_2$ )	sd( $\hat{N}_2$ )	95% PI( $\hat{N}_2$ )	median( $\hat{N}_2$ )	%bias( $\hat{N}_2$ )	RMSE( $\hat{N}_2$ )
1	3	15	8.70E+06	2.75E+08	(302, 2190)	661.92	1.45E+06	8.70E+6
2	1	32	674	324	(181, 1456)	613.24	12.4	10.5
3	6	30	673	237	(356, 1277)	638.51	12.2	7.83
4	3	55	635	190	(323, 1069)	608.96	5.88	6.12
5	12	15	4.33E+07	7.68E+08	(349, 2216)	704.29	7.22E+06	2.43E+07
Model	Visits	WR	mean( $\hat{N}_3$ )	sd( $\hat{N}_3$ )	95% PI( $\hat{N}_3$ )	median( $\hat{N}_3$ )	% bias( $\hat{N}_3$ )	RMSE( $\hat{N}_3$ )
1	3	15	8.71E+06	2.75E+08	(383, 5353)	1002.98	1.45E+06	8.70E+06
2	1	32	1152.3	717	(273, 2952)	973.71	92.0	28.6
3	6	30	1148.68	672	(465, 2575)	981.99	91.4	27.4
4	3	55	1098.89	474	(466, 2359)	1001.48	83.1	21.8
5	12	15	9.77E+07	1.86E+09	(432, 6051)	1085.27	1.63E+07	5.89E+07

there was a very large amount of uncertainty in the abundance estimates. These results assumed detectability for male woodrats remained constant across the two years, and also did not depend on any other explanatory variables.

Both male and female abundance estimates were based on trial survey data collected using small sample sizes (18 and 10 female and male individuals, respectively). Given the results of the simulation studies presented in Chapter 6 and Section 7.4.2, small sample sizes cause positive bias in abundance estimates. It is therefore likely that these estimates of woodrat abundance are biased high, possibly by as much as the 12% found in this simulation study, or perhaps even higher.

Unfortunately, abundance of male woodrats could not be reliably estimated because the sample sizes available to estimate the detection function were too low. Since some male woodrats appear to be nomadic, roaming extremely large distances (e.g., distances greater than 600 m have been observed during radio-collaring, unpubl. data. J. Potts), estimating the detection function and abundance of males is difficult: they have low average detectability and consequently,

the variance estimate will be very high. In addition, obtaining enough successful trials will require a large amount of survey effort. In contrast, the female abundance estimate was relatively stable over the two years, despite the small survey effort. Females tend to have relatively small, well-defined home ranges, staying close to nest locations to rear offspring and do not venture as far as male woodrats. Consequently, estimating the detection function of female woodrats requires less trial survey effort.

To ensure the estimated detection function is representative of all woodrats in the main survey, it is important to ensure the time frame in which the main and trial surveys are conducted is similar (c.f. Chapter 5). However, in this situation these two surveys could not be held at exactly the same time since spacing between traps would have been significantly less than the requisite 250 m (traps set in the main survey would have been interspersed with trial survey traps). In general, trial surveys were performed shortly before or shortly after (a maximum difference of three weeks) the main survey was completed in the same region the radio-collared woodrat occupied. The discrepancy of three weeks was not great enough to cause changes in detectability which may arise due to seasonal differences.

However, is the estimated detection function representative of all woodrats in the main survey? Typically, woodrats in the main survey have not been previously captured. However, individuals in the trial survey have been captured previously, in order to attach the radio-collar. It is therefore assumed that the behaviour of these animals post-capture does not change. Standard capture-recapture models have been developed to incorporate behavioural response to the capture experience, and the TPT method could be extended to include these types of models (see Section 6.6.3).

The data collected in this chapter represent approximately 1000 trap nights

of total survey effort each year. In Chapters 3 and 4, over 18,000 trap nights were used. This is significant for two main reasons:

1. The cost saving using the TPT method compared to capture-recapture approaches is substantial. Although with such small sample sizes, abundance estimates can be very positively biased, by increasing the number of animals upon which trials are performed, bias is reduced.
2. Increasing survey effort for the TPT method to equate to that used in Chapters 3 and 4 is not realistic. A total survey effort of 18,000 trap nights might represent a main survey based on a 250 m grid spacing, where each point is visited 25 times (i.e., 6800 trap nights) and 18 trials are performed on 150 individual woodrats (i.e., 7200 trap nights). Yet this very large survey effort still does not equate to that used in the capture-recapture studies. Implementing these large survey levels is just not realistic: in any one field season, we have not caught more than 70 individual woodrats. It is therefore unreasonable to expect to be able to attach radio collars to 150 individuals.

Based on the simulation study conducted in Section 7.4.2, it is possible to obtain nearly unbiased estimates of population abundance with increased survey effort. In future surveys, greater effort should be allocated to increasing the number of individual radio-collared woodrats for the trial survey. Given the difficulties in capturing males and the dependence of the population on female woodrats, just the female proportion of the population could be monitored. The portion of trials survey effort previously allocated to males could then be allocated to the female woodrats, increasing the trials survey effort for the females (and decreasing bias and increasing precision) without a total increase in survey cost.

## Chapter 8

# General Discussion and Conclusion

### 8.1 Estimating Key Largo woodrat abundance

In this thesis I investigated three methods of estimating the abundance of the Key Largo woodrat, an endangered rodent with a highly restricted geographic range and low abundance. The first method I investigated was capture-recapture (Chapter 3). Although capture-recapture is the most common method of estimating small mammal abundance (Pollock et al., 1990; Solari et al., 2002; Wiewel et al., 2007), estimates of abundance are typically not robust due to imprecise estimation of the effective trapping area (ETA) of the grid (Borchers and Efford, 2008). To date, no standard method of estimating the ETA exists (Efford et al., 2009b). I demonstrated in Chapter 3 that, depending on the method used to estimate the ETA, estimates of abundance can vary widely: from 394 to 1292 individuals, in the first sampling session undertaken in 2007 (Table 8.1).

Spatially-explicit capture-recapture (SECR, Borchers and Efford, 2008) and trapping point transects (TPT, Buckland et al., 2006) were the two other techniques I used to estimate abundance of the woodrat. Both methods combine capture-recapture studies with distance sampling (Buckland et al., 2001). In

Chapter 4, I analysed the same Key Largo woodrat data set presented in Chapter 3 but within an SECR framework. Estimates of woodrat abundance were lower when analysed using SECR, compared to standard capture-recapture (Table 8.1). This suggests the estimates of the effective trapping area used in Chapter 3 was much smaller than that in Chapter 4, and that woodrats can move from distances greater than the estimated home range radius to be captured on the trapping grid.

Also, the size of the trapping grids used in this study was  $60 \times 60$  m ( $7 \times 7$  traps with a 10 m trap spacing), and was originally designed to survey the Key Largo cotton mouse, a co-occurring species with a smaller home range size than the woodrat. Potentially the trap spacing used in this survey was too small for the woodrat and efficiency of the method might be improved using a larger trap spacing. Both McCleery (2003) and Winchester (2007) used a 25 m trap spacing when surveying for the Key Largo woodrat.

In Chapter 7, I used TPT to estimate abundance of the woodrat. The trapping point transect monitoring method was first proposed by Buckland et al. (2006), and applied to flocks of Scottish crossbills. In that study, a single trial was conducted on each flock of birds, and the detection function was fitted using a generalised linear model. In this thesis, I considered an extension of this method as applied to small mammals, where radio-collars were attached to individual woodrats and multiple trials are conducted on each radio-collared animal in the trials survey. The detection function was fitted using generalised linear mixed models.

When the detection function for male and female woodrats was estimated together, an interaction between sex and year was found significant during model selection. The significance of this interaction was surprising: I had no prior expectation that the detectability of male and female woodrats would change

differently for each year. Also, the number of male woodrats upon which trials were performed in 2008 was very small (4), and perhaps the significance of the interaction between sex and year was an artefact of the small sample size. Consequently, I estimated different detection functions for male and female woodrats. Since the behavior of male and female woodrats is quite different, estimating separate detection functions for each sex seems reasonable: male woodrats tend to be nomadic, wandering large distances between many nest locations. In contrast, female woodrats have quite well-defined home ranges, occupying fewer nest locations that are closer together. Detectability of female woodrats was found to depend on year and distance of the trial trap from the nest. Since the sample size for male woodrats was so small, only a simple detection function could be fitted, where detectability depended only on distance of the trial trap from the nest. I used two estimators,  $\hat{N}_2$  and  $\hat{N}_3$ , to estimate female and male woodrat abundance in 2008 and 2009 (Table 8.1).

Differences in abundance estimates for the three methods (capture-recapture, SECR and TPT) may have been caused by differing survey effort and surveys being undertaken at different times of the year. Confidence intervals associated with estimates obtained using the TPT method were wider than those obtained using SECR (Table 8.1), but abundance estimates using SECR were obtained using 18 times more trap nights than the TPT method. This represents a significant cost that is prohibitive to using the SECR method as the basis of a long-term monitoring program of the Key Largo woodrat.

Table 8.1: Summary of Key Largo woodrat abundance estimates using six different estimators. Four methods were based on capture-recapture, where the estimated effective trapping area of the grid changed (Naïve, Home Range (HR), Mean Maximum Distance Moved (MMDM) and Radio tracking (RT), c.f. Section 3.2.4 on page 56). One method was based on spatially-explicit capture-recapture (SECR) and used the same data set as analysed using standard capture-recapture methods. Two methods were based on trapping point transects (TPT), using either the  $\hat{N}_2$  or  $\hat{N}_3$  abundance estimators (see Section 5.3.2 on page 81) and estimated the abundance of females and males in the population separately. 95% confidence intervals are given in parentheses. Trapping effort is the total number of trap nights undertaken in each each.

Session	Naïve	HR	MMDM	RT	SECR	TPT ( $\hat{N}_2$ )	TPT ( $\hat{N}_3$ )	Trapping effort
2007, Spring	1291.95	360.54	394.15	1033.56	43 (31, 116)	-	-	18000
2007, Summer	1982.92	553.37	927.08	1586.34	214 (140, 391)	-	-	18000
2007, Autumn	3203.18	839.91	1340.86	2562.55	382 (243, 543)	-	-	18000
2008 females, Spring	-	-	-	-	-	185 (54, 689)	868 (143, 11663)	941
2008 males, Spring	-	-	-	-	-	523 (135, 1402)	1099 (380, 4190)	891
2009 females, Spring	-	-	-	-	-	144 (32, 531)	680 (82, 7814)	985
2009 males, Spring	-	-	-	-	-	83 (0, 296)	175 (0, 842)	971

Although the number of trap nights used in the capture-recapture methods was substantially greater than the TPT method, this might not reflect the total survey cost. For example, checking an intensive grid of 49 traps in a small geographic area is an efficient use of time, as field staff do not need to walk large distances to check and open traps. Also, the TPT method required radio-collars to be purchased (USD220 each collar), which is not insignificant. In addition, radio-tracking of woodrats, especially male woodrats that roam large distances, is time consuming. It took approximately 8 weeks to complete the TPT field work, whereas the capture-recapture data were collected over a 7 month period (March-April, July-September, and November-December). Within the eight weeks during which the TPT survey was undertaken, approximately 960 man-hours were required. This comprises of approximately 480 hours in the trials survey (two people working approximately 30 hours per week to relocate collared animals) and 480 hours in the main survey (two people working approximately 30 hours per week). By far, the most expensive aspect of the TPT survey was undertaking the trials survey: at most 10 radio collars were deployed simultaneously and they took approximately 5 hours to relocate daily (i.e. approximately 30 mins to relocate each individual, including travel time between locations). The capture-recapture data presented in Chapters 3 and 4, took approximately 1960 hours over the seven month period (pers. comm. Dan Greene). This represented two people working 30 hours per week to check the grids of traps.

In an ideal world, an experimenter would know the true abundance,  $N$ , for a population, and use a variety of field methods and statistical techniques to estimate  $N$  (e.g. Parmenter et al. 2003). By knowing the survey effort of each field method, and the bias and precision associated with each abundance estimate, the experimenter can make a very informed decision about which survey and statistical method provides the most accurate and precise estimate of the

true population abundance  $N$ . Although such studies are desirable, they are rare and perhaps impossible for many species and habitats. Instead people have used artificial stocking of a bounded plot (Rodda and Campbell, 2002), or set-up situations using a known number of objects, and tried to estimate them using a survey methodology (e.g., Bergstedt and Anderson (1990) and Anderson et al. (2001) used distance sampling to estimate a known number of bricks and model tortoises, respectively). Consequently, most studies compare the performance of a new method such as SECR or TPT, with the best currently available alternative (e.g., capture-recapture, Hounsome et al. 2005), based on the economic costs, survey difficulty or time requirements (Rodda and Campbell, 2002). Such validation does not establish the accuracy of a method (Rodda and Campbell, 2002), i.e., how close the estimate is to the true population size (Hounsome et al., 2005). Instead, simulation studies can be used to investigate the statistical properties of a method (see Pearse et al., 2007 and references within). Simulation studies provide a means for investigating how different sampling strategies and levels of survey effort influence bias and variance in estimates of abundance.

## 8.2 Extending the TPT simulation study

In Chapter 6, I used a simulation study for three different detection function scenarios to investigate how the number of trials per individual versus the number of individuals upon which trials were performed influenced bias and precision in abundance estimates when underlying between-group heterogeneity was either accounted for, or not. In Chapter 7 (Section 7.3), I conducted another small simulation study to investigate the tradeoff that occurs between investing survey effort in the trials survey versus the main survey, for example by increasing the number of trapping points in the main survey. All simulation studies are limited to some extent by trying to take a complicated, “real-world” situation, that is

perhaps poorly understood, and summarise it within a few different scenarios of a simulation study. In the following paragraphs I suggest some areas that could be investigated in the future.

All the simulation studies presented in this thesis randomly allocated individuals throughout the area of the main survey according to a homogeneous Poisson process. This is a simple scenario as animals might be clustered together through space and vary their distribution over time. Therefore, the simulation study could be modified to better suit the biology of the target species and if the target species is known to form clusters, and these clusters are quantitatively defined, then this information should be included in the simulation study.

In the simulation studies, I did not investigate precision – instead I investigated between-simulation variability in population estimates. If the variability between simulations is low, and the estimated bias is also low, users of the TPT method will get estimates that are closer to the truth more often. However, I could have also calculated the coverage of the confidence intervals. That is, within each simulation, I could have performed a bootstrap of the simulated data to determine if the estimated confidence intervals contained the true abundance. It was too time consuming to conduct an analysis including coverage of confidence intervals, but this could be pursued in the future.

The section investigating model mis-specification could also be expanded in future studies. In Section 6.4 on page 114, I investigated the effect of not including underlying covariates in the detection function that influence detectability. However, mis-specification could also be investigated by using a different model to generate the detection data than that used to fit the data. For example, detection data could be simulated using a half-normal or hazard rate distribution, and the detection function fitted using the methods presented in this thesis. In these simulation studies, I simulated detection data from an underlying detection

function of the correct form (a generalised linear mixed model).

### 8.3 Generalising the TPT method

In the field work component of this research, I attached radio-collars to individuals and conducted multiple trials on each individual in the trials survey. However, the cost of radio-collars is not insignificant ( $\approx 220$ USD per transmitter) and locating radio-collared individuals is time consuming (*c.* 1 to 2 hours per trial, depending on terrain). In addition, some small mammals are too small to accommodate a radio transmitter. A cost saving measure might therefore be to conduct the trials survey by using the point-of-release of a captured animal as the trial ‘start point’ and by establishing a trial trap some known distance (and random direction) from the point-of-release. This approach would eliminate the reliance on radio collaring and locating individuals via telemetry.

With reference to Figure 8.1, the point-of-release to point-of-recapture trial survey method is as follows: the center of the animal’s home range (e.g., its nest) is located at an unknown point  $C$ . The animal is captured (and released) at point  $R$ , and upon release, it is assumed to return to its nest ( $C$ ). A trial trap is set at  $T$ . Since the location of  $C$  is unknown, the distance of an individual trial,  $d$ , is estimated as the distance between the animal’s point-of-release ( $R$ ) and the location of the trial trap ( $T$ ). However, the “true” trial distance is actually  $D$ , the distance between the animal’s start point ( $C$ ) when the trial trap was set, and the location of the trial trap  $T$ . Whether the discrepancy between  $d$  and  $D$  causes bias in population estimates needs to be explored. I give some preliminary thoughts below.

Consider the simple situation where capturing an animal does not affect the probability of capturing it again. The probability of capturing an animal at location  $R$  is a function of home range size and the probability of capturing an

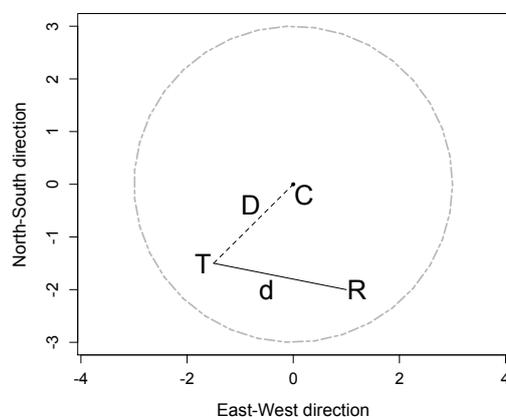


Figure 8.1: Within the animal's home range (denoted by a grey dotted line), is its nest, located at  $C$ . The animal was caught (and released) at  $R$ , and upon release, is assumed to return to its nest. A trial trap was then set at  $T$ .  $d$  is the distance between the animal's point-of-release, and the trial trap.  $D$  is the distance between the animal's start location when the trial trap was set (e.g., its nest) and the trial trap.

animal is greater at its home range centre than at the boundary of its home range. A trial trap ( $T$ ) must then be established at a known distance from the point-of-release ( $R$ ). For the trial to be successful (i.e., recapturing the animal), the first requirement is the trial trap needs to be placed near the animal's activity centre.

Consider the extreme case when  $R$  is near the boundary of the animal's home range (Figure 8.2). Even trial traps placed very short distances from  $R$  will have a low chance of success, because trials will fall outside the home range, and those that fall inside the home range, will be located a long distance from the animal's start point (assuming the animal starts at its home range centre). It would be expected that this would negatively bias the detection probability at short trial distances. As trial distance increases, more trials will be conducted outside the home range boundary. However, some trials will recapture the animal

(i.e., the trial is successful) at larger distances, as they will be located closer to the home range centre. Therefore, the probability of detection will, on average, be overestimated at longer distances, when compared to the “known-position” radio-collared method. Whether these two biases cancel each other, causing the resulting estimate of average detection probability  $\hat{P}$  to be unbiased should be investigated in future research, via a simulation study. Also, as  $R$  tends to  $C$ , the discrepancy between  $d$  and  $D$  tends to zero, and any bias in the estimated probability of detection will also tend to zero. How these two counteracting biases interact, and what might be expected to occur in the presence of a behavioral response to the trapping experience, can be addressed using e.g., a simulation study, and should be investigated, before this technique is implemented in the field.

Conducting the trials survey in this manner is a rather inefficient way to estimate the detection function, as a large number of trials would be required at greater distances so the tail of the detection function could be estimated properly, when the probability of detection becomes very close to zero. Despite the large number of trials required, cost savings may be significant compared to the radio-collaring method, as additional equipment (e.g., radio collars, radio receivers, etc.) is not required.

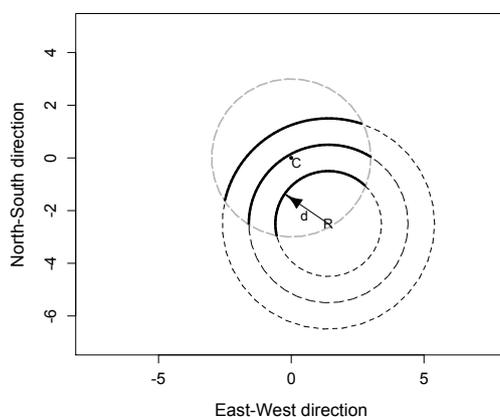


Figure 8.2: Within the animal's home range (denoted by a grey dotted line), is its nest, located at  $C$ . The animal was caught (and released) at  $R$ , and upon release, is assumed to return to its nest. If a trial trap is placed at distance  $d$  from the point-of-release  $R$ , some trials will be set beyond the animal's home range boundary (dashed line) and some will be set inside the animal's home range boundary (solid line). As  $d$  increases, the proportion of trials occurring outside the animal's home range increases, but some trials will be set very close to the animal's home range center ( $C$ ).

## 8.4 General conclusion

The trapping point transect method shows great potential to become a standard monitoring technique in estimating animal abundance, especially when animals are small and rare (e.g., the Key Largo woodrat). The trapping point transect method used considerably less resources than the capture-recapture studies investigated in this thesis, however uncertainty in population estimates was larger. By investing a greater level of TPT survey effort, especially increasing the number of animals upon which trials are performed to estimate the detection function, will decrease bias in population estimates and increase precision.

# Bibliography

- Adkins, J. P. (2003). Experimental predator removal: a response in small mammal communities and relations to duck nest success. Master's thesis, Louisiana State University.
- Alligood, C. A., Wheaton, C. J., Daneault, A. J., Carlson, J. K., and Savage, A. (2009). Behavioral predictors of copulation in captive Key Largo woodrats (*Neotoma floridana smalli*). *Behavioural Processes*, 81(2):337–342.
- Alligood, C. A., Wheaton, C. J., Forde, H. M., Smith, K. N., Daneault, A. J., Carlson, R. C., and Savage, A. (2008). Pup development and maternal behavior in captive Key Largo woodrats (*Neotoma floridana smalli*). *Zoo Biology*, 27(5):394–405.
- Anderson, D., Burnham, K., Lubow, B., Thomas, L., Corn, P., Medica, P., and Marlow, R. (2001). Field testing line transect estimation of desert tortoise abundance. *Journal of Wildlife Management*, 65(3):583–597.
- Anderson, D., Burnham, K., White, G., and Otis, D. (1983). Density estimation of small-mammal populations using a trapping web and distance sampling methods. *Ecology*, 64(4):674–680.
- Bailey, L., Simons, T., and Pollock, K. (2004). Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecological Applications*, 14(3):692–702.
- Barbour, D. B. and Humphrey, S. R. (1982). Status and habitat of the Key Largo woodrat and cotton mouse (*Neotoma floridana smalli* and *Peromyscus gossypinus allapaticola*). *Journal of Mammalogy*, 63(1):144–148.
- Bates, D. and Maechler, M. (2009). *Package 'lme4': Linear mixed-effects models using S4 classes*.
- Bergstedt, R. A. and Anderson, D. R. (1990). Evaluation of line transect sampling based on remotely sensed data from underwater video. *Transactions of the American Fisheries Society*, 119:86–91.
- Bolker, B., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., and White, J. S. S. (2009). Generalized linear mixed models: a

- practical guide for ecology and evolution. *Trends in Ecology and Evolution*, 24(3):127–135.
- Boonstra, R., Kanter, M., and Krebs, C. J. (1992). A tracking technique to locate small mammals at low densities. *Journal of Mammalogy*, 73(3):683–685.
- Borchers, D., Buckland, S., Goedhart, P., Clarke, E., and Hedley, S. (1998). Horvitz-thompson estimators for double-platform line transect surveys. *Biometrics*, 54:1221–1237.
- Borchers, D. and Efford, M. (2008). Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*, 64(2):377–385.
- Borchers, D. L. (2011). A non-technical overview of spatially explicit capture-recapture models. *Journal of Ornithology*. In press.
- Brown, H. and Prescott, R. (2004). *Applied Mixed Models in Medicine*. Wiley Publications, West Sussex, England.
- Brown, K. P., Moller, H., Innes, J., and Alterio, N. (1996). Calibration of tunnel tracking rates to estimate relative abundance of ship rats (*Rattus rattus*) and mice (*Mus musculus*) in a New Zealand forest. *New Zealand Journal of Ecology*, 20(2):271–275.
- Brown, L. and Williams, R. (1971). The Key woodrat and cotton mouse on Lignum Vitae Key, Florida. *Florida Field Naturalist*, 44(3):95–96.
- Brown, L. N. (1973). The Key Largo woodrat (*Neotoma floridana smalli*). Pages 228-229 in Threatened wildlife of the United States. Resource Publication 114. Bureau of Sport Fisheries and Wildlife, US Dept of Interior, Washington D.C.
- Buckland, S. (2006). Point transect surveys for songbirds: robust methodologies. *The Auk*, 123:345–357.
- Buckland, S., Burnham, K., and Augustin, N. (1997). Model selection: An intergral part of inference. *Biometrics*, 53(2):603–618.
- Buckland, S., Goudie, I., and Borchers, D. (2000). Wildlife population assessment: past developments and future directions. *Biometrics*, 65(1):1–12.
- Buckland, S., Summers, R., Borchers, D., and Thomas, L. (2006). Point transect sampling with traps or lures. *Journal of Applied Ecology*, 43:377–384.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2001). *Introduction to Distance Sampling*. Oxford University Press, Oxford.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2004). *Advanced Distance Sampling*. Oxford University Press, Oxford, Oxford.

- Burnham, K. P. and Anderson, D. R. (2003). *Model selection and multi-model inference - A practical information-theoretic approach*. Springer, United States of America, 2nd edition.
- Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, 43:783–791.
- Chao, A., Chu, W., and Hsu, C. H. (2000). Capture-recapture when time and behavioral response affect capture probabilities. *Biometrics*, 56(2):427–433.
- Chao, A., Lee, S.-M., and Jeng, S.-L. (1992). Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal. *Biometrics*, 48:201–216.
- Conn, P., Arthur, A., Bailey, L., and Singleton, G. (2006). Estimating the abundance of mouse populations of known size: promises and pitfalls of new methods. *Ecological Applications*, 16(2):829–837.
- Conroy, M. J., Runge, J. P., Barker, R. J., Schofield, M. R., and Fonnesebeck, C. J. (2008). Efficient estimation of abundance for patchily distributed populations via two-phase, adaptive sampling. *Ecology*, 89(12):3362–3370.
- Corn, J. and Conroy, M. (1998). Estimation of density of mongooses with capture-recapture and distance sampling. *Journal of Mammalogy*, 79(3):1009–1015.
- Davison, A. C. and Hinkley, D. (1997). *Bootstrap Methods and their Application*. Cambridge University Press.
- Dice, L. R. (1938). Some census methods for mammals. *Journal of Wildlife Management*, 2(3):119–130.
- Dique, D. S., Preece, H. J., Thompson, J., and de Villiers, D. L. (2004). Determining the distribution and abundance of a regional koala population in south-east queensland for conservation management. *Wildlife Research*, 31:109–117.
- Eberhardt, L. (1990). Using radio-telemetry for mark-recapture studies with edge effects. *The Journal of Applied Ecology*, 27(1):259–271.
- Efford, M. (2004). Density estimation in live-trapping studies. *Oikos*, 106:598–610.
- Efford, M. (2010). *Package ‘secr’*. University of Otago, [www.otago.ac.nz/density/pdfs/secr-manual.pdf](http://www.otago.ac.nz/density/pdfs/secr-manual.pdf).
- Efford, M., Dawson, D., and Borchers, D. (2009a). Population density estimated from locations of individuals on a passive detector array. *Ecology*, 90(10):2676–2682.

- Efford, M., Dawson, D., and Robbins, C. (2004). Density: software for analysing capture-recapture data from passive detector arrays. *Animal Biodiversity and Conservation*, 27:217–228.
- Efford, M., Warburton, B., Coleman, M., and Barker, R. (2005). A field test of two methods for density estimation. *Wildlife Society Bulletin*, 33:731–738.
- Efford, M. G., Borchers, D. L., and Byrom, A. E. (2009b). Density estimation by spatially explicit capture-recapture: likelihood-based methods. In Thomson, D. L., Cooch, E. G., and Conroy, M. J., editors, *Modeling Demographic Processes In Marked Populations*, pages 255–269. Springer.
- English, P. F. (1923). The dusky-footed wood rat (*Neotoma fuscipes*). *Journal of Mammalogy*, 4(1):1–9.
- Fewster, R., Buckland, S., Burnham, K., Borchers, D., Jupp, P., Laake, J., and Thomas, L. (2009). Estimating the encounter rate variance in distance sampling. *Biometrics*, 65:225–236.
- Forys, E. A., Quistorff, A., and Allen, C. R. (2001). Potential fire ant (Hymenoptera: Formicidae) impact on the endangered Schaus swallowtail (Lepidoptera: Papilionidae). *Florida Entomologist*, 84(2):254–258.
- Frank, P., Percival, F., and Keith, B. (1997). A status survey for Key Largo woodrat (*Neotoma floridana smalli*) and Key Largo cotton mouse (*Peromyscus gossupinus allapaticola*) on North Key Largo, Monroe County, Florida. Technical report, submitted to the US Fish and Wildlife Service, Jacksonville, Florida, USA.
- Gander, F. F. (1929). Experiences with wood rats, *Neotoma fuscipes macrotis*. *Journal of Mammalogy*, 10(1):52–58.
- Gardner, B., Reppucci, J., Lucherini, M., and Royle, J. (2010). Spatially explicit inference for open populations: estimating demographic parameters from camera-trap studies. *Ecology*, 91(11):3376–3383.
- Gibbons, J. and Andrews, K. (2004). PIT tagging: Simple technology at its best. *BioScience*, 54(5):447–454.
- Gimenez, O., Viallefont, A., Catchpole, E. A., Choquet, R., and Morgan, B. J. T. (2004). Methods of investigating parameter redundancy. *Animal Biodiversity and Conservation*, 27(1):561–572.
- Gitzen, R. A., West, S. D., and Trim, B. E. (2001). Additional information on the distributions of small mammals at the Hanford Site, Washington. *Northwest Science*, 75(4):350–362.
- Goodyear, N. C. (1985). Results of a study of Key Largo woodrats and cotton mice. Technical report, Department of Zoology, University of Rhode Island.

- Goodyear, N. C. (1992). Spatial overlap and dietary selection of native rice rats and exotic black rats. *Journal of Mammalogy*, 73(1):186–200.
- Gore, J. A. and Loggins, R. E. (2005). Long term monitoring plan for the Key Largo woodrat and cotton mouse. Technical report, Florida Fish and Wildlife Conservation Commission, FWRI File Code F2160-03-F.
- Greene, D. U. (2009). Trapping study of the endangered Key Largo cotton mouse (*Peromyscus gossypinus allapaticola*): a protocol for long-term population monitoring. Master's thesis, University of Georgia.
- Greene, D. U., Potts, J. M., Duquesnel, J. G., and Snow, R. W. (2007). Geographic distribution: *Python molurus bivittatus* (Burmese python). *Herpetological Review*, 38:355.
- Hadley, G. L. and Wilson, K. R. (2004). Patterns of density and survival in small mammals in ski runs and adjacent forest patches. *Journal of Wildlife Management*, 68(2):288–298.
- Hamm, K. A., Diller, L. V., and Kitchen, D. W. (2002). Comparison of indices to estimate abundance of dusky-footed woodrats. *Wildlife Society Bulletin*, 30(1):64–70.
- Hatley, P. J. (2003). Feral cat colonies in Florida: the fur and feathers are flying. *Journal of Land Use*, 18(2):441–465.
- Healy, W. M. and Welsh, C. J. E. (1992). Evaluating line transects to monitor gray squirrel populations. *Wildlife Society Bulletin*, 20(1):83–90.
- Hersh, S. L. (1981). Ecology of the Key Largo woodrat (*Neotoma floridana smalli*). *Journal of Mammalogy*, 62(1):201–206.
- Hounscome, T., Young, R., Davison, J., Yarnell, R., Trewby, I., Garnett, B., Delahay, R., and Wilson, G. (2005). An evaluation of distance sampling to estimate badger (*Meles meles*) abundance. *Journal of Zoology*, 266(1):81–87.
- Huggins, R. (1989). On the statistical analysis of capture-recapture experiments. *Biometrika*, 76(1):133–140.
- Huggins, R. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics*, 47(2):725–732.
- Humphrey, S. R. (1988). Density estimates of the endangered Key Largo woodrat and cotton mouse (*Neotoma floridana smalli* and *Peromyscus gossypinus allapaticola*), using the nested-grid approach. *Journal of Mammalogy*, 69(3):524–531.
- Jolly, G. (1965). Explicit estimates from capture-recapture data with both death and immigration stochastic model. *Biometrika*, 52(1/2):225–247.

- Jones, C., McShea, W. J., Conroy, M. J., and Kunz, T. H. (1996). *Measuring and monitoring biological diversity - standard methods for mammals*, chapter 8 - Capturing mammals, pages 115–156. Smithsonian Institution.
- Karanth, K. and Nichols, J. (2002). *Monitoring tigers and their prey*. Centre for Wildlife Studies.
- Karels, T. J., Koppel, L., and Hik, D. (2004). Fecal pellet counts as a technique for monitoring an alpine-dwelling social rodent, the hoary marmot (*Marmota caligata*). *Arctic, Antarctic, and Alpine Research*, 36(4):490–494.
- Keith, B. and Frank, P. (1998). North Key Largo Caysfort small mammal inventory. Technical report, submitted to Department of Environmental Protection.
- Kendall, B. E. and Pollock, K. H. (1992). *Wildlife 2001: Populations*, chapter The robust design for capture-recapture studies: a review and evaluation by Monte Carlo simulation, pages 31–43. Elsevier Science Publishers, London.
- Kendall, W. (2008). The “Robust Design”. In Cooch, E. and White, G. C., editors, *A Gentle Introduction*, chapter 15.
- Kendall, W., Nichols, J., and Hines, J. (1997). Estimating temporary emigration using capture-recapture data with Pollock’s robust design. *Ecology*, 78(2):563–578.
- Kendall, W., Pollock, K., and Brownie, C. (1995). A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics*, 51(1):293–308.
- Kéry, M., Gardner, B., Stoeckle, T., Weber, D., and Royle, J. A. (2011). Use of spatial capture-recapture modeling and dna data to estimate densities of elusive animals. *Conservation Biology*, 25(2):356–364.
- Kiffner, C., Waltert, M., Meyer, B., and Muhlenberg, M. (2007). Response of lions (*Panthera leo* Linnaeus 1758) and spotted hyaenas (*Crocuta crocuta* Erxleben 1777) to sound playbacks. *African Journal of Ecology*, 46(2):223–226.
- Kohn, M. H. and Wayne, R. K. (1997). Facts from faeces revisited. *Trends in Ecology and Evolution*, 12:223–227.
- Komarek, A. and Lesaffre, E. (2008). Generalized linear mixed model with a penalized Gaussian mixture as a random effects distribution. *Computational Statistics and Data Analysis*, 52:3441–3458.
- Laudenslayer, W. and Fargo, R. (2002). Small mammal populations and ecology in the Kings River Sustainable Forest Ecosystems Project Area. Technical report, USDA Forest Service General Technical Report.

- Lee, D. E. and Tietje, W. D. (2005). Dusky-footed woodrat demography and prescribed fire in a California oak woodland. *Journal of Wildlife Management*, 69(3):1211–1220.
- Lesaffre, E. and Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random-effects model: an example. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 50(3):325–335.
- Link, W. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics*, 59(4):1123–1130.
- Link, W. A. and Barker, R. J. (1994). Density estimation using the trapping web design: A geometric analysis. *Biometrics*, 50(3):733–745.
- Linsdale, J. and Tevis, L., J. (1956). A five-year change in an assemblage of wood rat houses. *Journal of Mammalogy*, 37(5):371–374.
- Loggins, R. E., Gore, J. A., Brown, L. L., Slaby, L. A., and Leone, E. H. (2010). A modified track tube for detecting beach mice. *Journal of Wildlife Management*, 74(5):1154–1159.
- LoGiudice, K. (2001). Latrine foraging strategies of two small mammals: implication for the transmission of *Baylisacaris procyonis*. *American Midland Naturalist*, 146:369–378.
- LoGiudice, K. (2003). Trophically transmitted parasites and the conservation of small populations: raccoon roundworm and the imperiled Allegheny woodrat. *Conservation Biology*, 17:258–266.
- Lukacs, P. (2002). Websim: simulation software to assist in trapping web design. *Wildlife Society Bulletin*, 30(4):1259–1261.
- Lukacs, P. and Burnham, K. (2005). Review of capture-recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, 14:3909–3919.
- Lukacs, P. M. (2008). Closed population capture-recapture models. In Cooch, E. and White, G. C., editors, *Program MARK: A Gentle Introduction*, chapter 14, <http://www.phidot.org/software/mark/docs/book/pdf/chap14.pdf>.
- Lukacs, P. M., Anderson, D. R., and Burnham, K. P. (2005). Evaluation of trapping-web designs. *Wildlife Research*, 32(2):103–110.
- MacKenzie, D., Nichols, J., Lachman, G., Droege, S., Royle, J., and Langtimm, C. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8):2248–2255.
- MacKenzie, D., Nichols, J., Royle, J., Pollock, K., Bailey, L., and Hines, J. (2006). *Occupancy estimation and modeling*. Academic Press, San Diego, California, USA.

- Marques, F., Buckland, S., Goffin, D., Dixon, C., Borchers, D., Mayle, B., and Peace, A. (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., Thomas, L., Martin, S. W., Mellinger, D. K., Jarvis, S., Morrissey, R. P., Ciminello, C. A., and DiMarzio, N. (2011). Spatially explicit capture–recapture methods to estimate minke whale density from data collected at bottom-mounted hydrophones. *Journal of Ornithology*. In press.
- Mattson, D. J., Reinhart, D. P. Mattson, D. J., and Reinhart, D. P. (1996). Indicators of red squirrel (*Tamiasciurus hudsonicus*) abundance in the whitebark pine zone. *Western North American Naturalist*, 56(3):272–275.
- McCarthy, M. A. and Possingham, H. P. (2007). Active adaptive management for conservation. *Conservation Biology*, 21(4):956–963.
- McCleery, R. A. (2003). Aspects of Key Largo woodrat ecology. Master’s thesis, Texas A & M University.
- McCleery, R. A., Foster, G. W., Lopez, R. R., Peterson, M. J., Forrester, D. J., and Silvy, N. J. (2005). Survey of raccoons on Key Largo, Florida, USA, for *Baylisascaris procyonis*. *Journal of Wildlife Diseases*, 41(1):250–252.
- McDonald-Madden, E., Probert, W. J. M., Hauser, C. E., Runge, M. C., Possingham, H. P., Jones, M. E. Moore, J. L., Rout, T. M., Vesk, P. A., and Wintle, B. A. (2010). Active adaptive conservation of threatened species in the face of uncertainty. *Ecological Applications*, 20(5):1476–1489.
- McDonough, T. J. and Rexstad, E. (2005). Short-term demographic response of the red-backed vole to spruce beetle infestations in Alaska. *Journal of Wildlife Management*, 69(1):246–254.
- Mendel, S. M. and Vieira, M. V. (2003). Movement distances and density estimation of small mammals using the spool-and-line technique. *Acta Theriologica*, 48(3):289–300.
- Menkens, G. E. and Anderson, S. H. (1988). Estimation of small-mammal population size. *Ecology*, 69(6):1952–1959.
- Mills, M., Juritz, J., and Zucchini, W. (2001). Estimating the size of spotted hyaena (*Crocuta crocuta*) populations through playback recordings allowing for non-response. *Animal Conservation*, 4:335–343.
- Moore, A. L. and McCarthy, M. A. (2010). On valuing information in adaptive-management models. *Conservation Biology*, 24(4):984–993.
- Morgan, B. J. T. and Ridout, M. (2009). Estimating N: A robust approach to capture heterogeneity. *Environmental and Ecological Statistics*, 3:1069–1080.

- Morrison, J. L. and Kennedy, P. L. (1989). The use of line transects to evaluate the abundance of diurnal mammalian prey. *Journal of Raptor Research*, 23(4):172–175.
- Muiznieks, B. (2006). Captive propagation and the Key Largo woodrat. *Endangered Species Bulletin*, XXXI(1):32–33.
- Nams, V. and Gillis, E. (2003). Changes in tracking tube use by small mammals over time. *Journal of Mammalogy*, 84(4):1374–1380.
- Nichols, J., Pollock, K., and Hines, J. (1984). The use of a robust capture-recapture design in small mammal population studies. *Journal of Mammalogy*, 64:253–260.
- Norris, J. L. and Pollock, K. H. (1995). A capture-recapture model with heterogeneity and behavioural response. *Environmental and Ecological Statistics*, 2(4):305–313.
- Oakley, K., Thomas, L., and Fancy, S. (2003). Guidelines for long-term monitoring protocols. *Wildlife Society Bulletin*, 31(4):1000–1003.
- Ogutu, J., Bhola, N., and Reid, R. (2005). The effects of pastoralism and protection on the density and distribution of carnivores and their prey in the Mara ecosystem of Kenya. *Journal of Zoology*, 265:281–293.
- Ogutu, J. and Dublin, H. (1998). The response of lions and spotted hyaenas to sound playbacks as a technique for estimating population size. *African Journal of Ecology*, 36:83–95.
- Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, 62:1–135.
- Parmenter, R. R., MacMahon, J. A., and Anderson, D. R. (1989). Animal density estimation using a trapping web design: field validation experiments. *Ecology*, 70(1):169–179.
- Parmenter, R. R., Yates, T. L., Anderson, D. R., Burnham, K. P., Dunnun, J. L., Franklin, A. B., Friggens, M. T., Lubow, B. C., Miller, M., Olson, G. S., Parmenter, C. A., Pollard, J., Rexstad, E., Shenk, T. M., Stanley, T. R., and White, G. C. (2003). Small-mammal density estimation: A field comparison of grid-based vs. web-based density estimators. *Ecological Monographs*, 73:1–26.
- Pearse, A. T., Dorr, B. S., Dinsmore, S. J., and Kaminski, R. M. (2007). Comparison of sampling strategies to estimate abundance of double-crested cormorants in western Mississippi. *Human-Wildlife Conflicts*, 1(1):27–34.

- Peterson, G., De Leo, G. A., Hellmann, J. J., Janssen, M. A., Kinzig, A., Malcolm, J. R., O'Brien, K. L., Pope, S. E., Rothman, D. S., Shevliakova, E., and Tinch, R. R. T. (1997). Uncertainty, climate change, and adaptive management. *Ecology and Society*, 1:1–7.
- Pledger, S. (2000). Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics*, 56:434–442.
- Pledger, S. and Efford, M. (1998). Correction of bias due to heterogeneous capture probability in capture-recapture studies of open populations. *Biometrics*, 54:888–898.
- Pledger, S., Pollock, K. H., and Norris, J. L. (2010). Open capture-recapture models with heterogeneity: II. Jolly-Seber model. *Biometrics*, 66(3):883–890.
- Pollock, K. (1982). A capture-recapture design robust to unequal probability of capture. *Journal of Wildlife Management*, 46:752–757.
- Pollock, K. H., Nichols, J. D., Brownie, C., and Hines, J. E. (1990). Statistical inference for capture-recapture experiments. *Wildlife Monographs*, 107.
- Rabe-Hesketh, S., Skrondal, A., and Pickles, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, 2(1):1–21.
- Reed, R. N. (2005). An ecological risk assessment of nonnative boas and pythons as potentially invasive species in the United States. *Risk Analysis*, 25:753–766.
- Rexstad, E. and Burnham, K. P. (1991). *User's guide for interactive program CAPTURE*. Colorado Cooperative Fish and Wildlife Research Unit, Fort Collins, Colorado, USA.
- Rodda, G. and Campbell, E. (2002). Distance sampling of forest snakes and lizards. *Herpetological Review*, 33(4):271–274.
- Rosenstock, S. S., Anderson, D. R., Giesen, K. M., Leukering, T., and Carter, M. F. (2002). Landbird counting techniques: current practices and an alternative. *The Auk*, 119(1):46–53.
- Ross, M. S., Meeder, J. F., Telesnicki, G., and Weekley, C. (1995). Terrestrial ecosystems of the Crocodile Lake National Wildlife Refuge: the effects of Hurricane Andrew. Technical Report 4, Submitted by the Florida International University, South east Environmental Research Program. Final report to the Department of the Interior, U.S. Fish and Wildlife Service, Vero Beach, Florida.
- Ross, M. S., O'Brien, J. J., and Flynn, L. J. (1992). Ecological site classification of Florida Keys terrestrial habitats. *Biotropica*, 24(4):488–502.

- Rout, T., Hauser, C., and Possingham, H. (2009). Optimal adaptive management for the translocation of a threatened species. *Ecological Applications*, 19(2):515–526.
- Royle, J., Nichols, J., Karanth, K., and Gopalaswamy, A. (2009). A hierarchical model for estimating density in camera-trap studies. *Journal of Applied Ecology*, 46:118–127.
- Royle, J. and Young, K. (2008). A hierarchical model for spatial capture-recapture data. *Ecology*, 89(8):2281–2289.
- Royle, J. A. and Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, 84(3):777–790.
- Sakai, H. and Noon, B. (1993). Dusky-footed woodrat abundance in different-aged forests in northwestern California. *Journal of Wildlife Management*, 57(2):373–382.
- Sasso, C. R. (1999). *Population dynamics, microhabitat use, and competition in the small mammal community of Key Largo*. PhD thesis, University of Miami.
- Schreiber, E., Bearlin, A., Nicol, S., and Todd, C. (2004). Adaptive management: a synthesis of current understanding and effective application. *Ecological Management and Restoration*, 5(3):177–182.
- Schwartz, A. and Odum, E. P. (1957). The woodrats of the Eastern United States. *Journal of Mammalogy*, 38(2):197–206.
- Schwarz, C. and Seber, G. (1999). Estimating animal abundance: Review III. *Statistical Science*, 14(4):427–456.
- Seber, G. (1965). A note on the multiple recapture census. *Biometrika*, 52(1/2):249–259.
- Seber, G. (1986). A review of estimating animal abundance. *Biometrics*, 42(2):267–292.
- Seber, G. (1992). A review of estimating animal abundance II. *International Statistical Review*, 60(2):129–166.
- Sherman, H. B. (1955). Description of a new race of woodrats from Key Largo, Florida. *Journal of Mammalogy*, 36(1):113–120.
- Small, J. (1923). Green deserts and dead gardens. *Journal of the New York Botanical Garden*, 24:215.
- Smith, M. H. (1968). A comparison of different methods of capturing and estimating numbers of mice. *Journal of Mammalogy*, 49:455–462.

- Soisalo, M. and Cavalcanti, S. (2006). Estimating the density of a jaguar population in the Brazilian Pantanal using camera-traps and capture-recapture sampling in combination with GPS radio-telemetry. *Biological Conservation*, 129:487–496.
- Solari, S., Rodriguez, J. J., Vivar, E., and Velazco, P. M. (2002). A framework for assessment and monitoring of small mammals in a lowland tropical forest. *Environmental Monitoring and Assessment*, 76:89–104.
- St-Laurent, M. H. and Ferron, J. (2008). Testing the reliability of pellet counts as an estimator of small rodent relative abundance in mature boreal forest. *Journal of Negative Results*, 5(1):14–22.
- Tanaka, R. (1972). Investigation into the edge effect by use of capture-recapture data in a vole population. *Researches on Population Ecology*, 13(2):127–151.
- Tanaka, R. (1980). Controversial problems in advanced research on estimating population densities of small rodents. *Researches on Population Ecology (Kyoto) Supplement*, 2:1–66.
- Thompson, W. L., White, G. C., and Gowan, C. (1998). *Monitoring Vertebrate Populations*. Academic Press, San Diego, California, USA.
- Tioli, S., Cagnacci, F., Stradiotto, A., and Rizzoli, A. (2009). Edge effect on density estimates of a radiotracked population of yellow-necked mice. *Journal of Wildlife Management*, 73(2):184–190.
- USFWS (1999). Multi-species recovery plan for South Florida - the Key Largo woodrat. Technical report, United States Fish and Wildlife Service.
- Velazquez, A. (1994). Distribution and population size of *Romerolagus diazi* on El Pelado Volcano, Mexico. *Journal of Mammalogy*, 75(3):743–749.
- Vestral, E. H. (1938). Biotic relations of the wood rat (*Neotoma fuscipes*) in the Berkeley Hills. *Journal of Mammalogy*, 19(1):1–36.
- Vos, P., Meelis, E., and Ter Keurs, W. J. (2000). A framework for the design of ecological monitoring programs as a tool for environmental and nature management. *Environmental Monitoring and Assessment*, 61:317–344.
- Vreeland, J. and Tietje, W. (1998). Initial response of woodrats to prescribed burning in oak woodland. *Transactions of the Western Section of the Wildlife Society*, 34:21–31.
- Vreeland, J. and Tietje, W. (1999). Counts of woodrat houses to index relative population abundance. *Wildlife Society Bulletin*, 27(2):337–343.
- White, G. (2005). Correcting wildlife counts using detection probabilities. *Wildlife Research*, 32(3):211–216.

- White, G. and Burnham, K. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study*, 46:120–138.
- White, G. C. (2008). Closed population estimation models and their extensions in program mark. *Environmental and Ecological Statistics*, 15:89–99.
- White, G. C. and Shenk, T. M. (2001). *Radio tracking and animal populations*, chapter Population estimation with radio-marked animals, pages 329–350. Academic Press, San Diego, California, USA.
- Wiewel, A. S., Clark, W. R., and Sovada, M. A. (2007). Assessing small mammal abundance with track-tube indices and mark-recapture population estimates. *Journal of Mammalogy*, 88(1):250–260.
- Williams, B., Nichols, J., and Conroy, M. (2002). *Analysis and Management of Animal Populations*. Academic Press.
- Wilson, G. J. and Delahay, R. J. (2001). A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation. *Wildlife Research*, 28(2):151–164.
- Wilson, K. and Anderson, D. (1985a). Evaluation of two density estimators of small mammal population size. *Journal of Mammalogy*, 66(1):13–21.
- Wilson, K. R. and Anderson, D. R. (1985b). Evaluation of a density estimator based on a trapping web and distance sampling theory. *Ecology*, 66(4):1185–1194.
- Winchester, C. (2007). An evaluation of habitat selection and an abundance estimate for the endangered Key Largo woodrat. Master's thesis, University of Georgia.
- Winchester, C., Castleberry, S. B., and Mengak, M. T. (2009). Evaluation of factors restricting distribution of the endangered Key Largo woodrat. *Journal of Wildlife Management*, 73(3):374–379.
- Wintle, B. A., Runge, M. C., and Bekessy, S. A. (2010). Allocating monitoring effort in the face of unknown unknowns. *Ecology Letters*, 13(11):1325–1337.
- Witmer, G. W. (2005). Wildlife population monitoring: some practical considerations. *Wildlife Research*, 32(3):259–263.

## Appendix A

# Table of results from trapping point transect simulations (Chapter 6)

Summary of the three detection function scenarios is given in Table A.1, and the input parameters for simulations 1, 2 and 3 are given in Table A.2. Results of simulation 1, 2 and 3 are given in Tables A.3, A.4, and A.5, respectively.

Table A.1: Input parameters for the three detection function scenarios (“High”, “Medium”, “Low”), where  $\alpha_g$  is the intercept term for an individual in group  $g$  ( $g = 1, 2, 3$ ),  $\beta$  is the coefficient (slope) parameter for the explanatory variable distance  $r$ , and  $g_{ig}$  is a random effect due to individual  $i$  in group  $g$ , in equation 6.1.

Scenario	Group	$\alpha_g$	$\beta$	$b_{ig}$
High	1	2	-0.15	N(0, 0.1)
Medium	1	2	-0.15	N(0, 0.1)
	2	1	-0.15	N(0, 0.1)
Low	1	2	-0.15	N(0, 0.1)
	3	-0.3	-0.15	N(0, 0.3)

Table A.2: Summary of input parameters for Simulations 1, 2 and 3. Three detection function scenarios were used (“High”, “Medium”, and “Low”) in Simulations 1 and 2, and only the “Medium” scenario was used in Simulation 3. The fitted detection function (Model) used in the analysis was either correct or not. The number of trials in the trials survey was either 360 (Simulation 1 and 2), or 780, 1080 and 1440 (Simulation 3). The method used to select the trial distances was either all of those considered (i.e., “Uniform”, “Stopping rule 5”, “Stopping rule 8” and “Adaptive”; Simulation 1 and 2) or just “Adaptive” (Simulation 3).

Simulation	Scenario	Model	Trial sample size	Trial distance method
1	High Medium Low	Correct	360	All
2	High Medium Low	Incorrect	360	All
3	Medium	Both	780 1080 1440	Adaptive

Table A.3: Table of results for all scenarios in Simulation 1. That is,  $\hat{N}_2$  and  $\hat{N}_3$  with associated percentage bias (pc.bias), standard deviation estimates (sd), percentile intervals (PI), median and root mean square error (RMSE) when underlying heterogeneity is accounted for, using the “High”, “Medium” and “Low” detection function scenarios, with differing effort allocated per individual (No.tri) and number of individuals (No.ind) in the trials survey.

Method	model	No.tri	No.ind	mean $\hat{N}_2$	sd $\hat{N}_2$	95%( $\hat{N}_2$ )	median $\hat{N}_2$	pc.bias $\hat{N}_2$	rmse $\hat{N}_2$	mean $\hat{N}_3$	sd $\hat{N}_3$	95%( $\hat{N}_3$ )	median $\hat{N}_3$	pc.bias $\hat{N}_3$	rmse $\hat{N}_3$
adapt	low	6	60	4.34E+03	8.03E+03	(827.23, 2.16E+04)	2.40E+03	1.17E+02	2.64E+02	5.11E+03	9.15E+03	(910.09, 2.39E+04)	2.73E+03	1.55E+02	3.06E+02
adapt	low	10	36	2.49E+03	2.07E+03	(792.39, 6.45E+03)	2.04E+03	2.46E+01	6.73E+01	3.00E+03	2.93E+03	(846.04, 8.98E+03)	2.36E+03	5.00E+01	9.79E+01
adapt	low	18	20	7.50E+07	1.45E+09	(761.04, 6.39E+03)	2.11E+03	3.75E+06	4.60E+07	9.00E+07	1.82E+09	(871.59, 8.26E+03)	2.40E+03	4.50E+06	5.76E+07
adapt	low	24	15	1.32E+08	1.54E+09	(737.92, 7.99E+03)	2.04E+03	6.58E+06	4.90E+07	1.83E+08	2.23E+09	(793.04, 1.05E+04)	2.32E+03	9.15E+06	7.08E+07
adapt	low	40	9	3.44E+08	2.07E+09	(720.93, 5.83E+09)	2.07E+03	1.72E+07	6.63E+07	4.91E+08	3.06E+09	(794.03, 8.10E+09)	2.35E+03	2.46E+07	9.79E+07
adapt	low	60	6	2.63E+08	3.00E+09	(732.76, 8.62E+03)	2.12E+03	1.32E+07	9.54E+07	3.17E+08	3.34E+09	(814.77, 1.09E+04)	2.32E+03	1.58E+07	1.06E+08
adapt	med	6	60	4.72E+03	6.46E+03	(1105.18, 2.13E+04)	2.52E+03	1.36E+02	2.22E+02	5.01E+03	6.80E+03	(1161.62, 2.26E+04)	2.67E+03	1.51E+02	2.35E+02
adapt	med	10	36	2.20E+03	9.39E+02	(1019.85, 4.15E+03)	2.08E+03	9.86E+00	3.04E+01	2.33E+03	1.03E+03	(1070.23, 4.44E+03)	2.19E+03	1.63E+01	3.40E+01
adapt	med	18	20	2.15E+03	7.36E+02	(1072.99, 3.75E+03)	2.02E+03	7.26E+00	2.37E+01	2.25E+03	8.02E+02	(1093.84, 4.13E+03)	2.11E+03	1.23E+01	2.65E+01
adapt	med	24	15	2.15E+03	6.99E+02	(1101.34, 3.73E+03)	2.06E+03	7.60E+00	2.26E+01	2.24E+03	7.50E+02	(1114.82, 4.02E+03)	2.13E+03	1.21E+01	2.49E+01
adapt	med	40	9	2.14E+03	7.23E+02	(988.39, 3.71E+03)	2.04E+03	7.07E+00	2.33E+01	2.21E+03	7.61E+02	(1016.28, 3.87E+03)	2.09E+03	1.07E+01	2.50E+01
adapt	med	60	6	2.17E+03	7.39E+02	(1031.59, 3.91E+03)	2.05E+03	8.73E+00	2.40E+01	2.24E+03	7.85E+02	(1038.84, 4.05E+03)	2.11E+03	1.21E+01	2.60E+01
adapt	high	6	60	3.86E+03	4.55E+03	(1263.86, 1.78E+04)	2.58E+03	9.32E+01	1.56E+02	4.09E+03	4.86E+03	(1297.74, 1.92E+04)	2.69E+03	1.05E+02	1.67E+02
adapt	high	10	36	2.13E+03	6.70E+02	(1118.46, 3.57E+03)	2.08E+03	6.51E+00	2.16E+01	2.23E+03	7.01E+02	(1137.97, 3.74E+03)	2.16E+03	1.15E+01	2.33E+01
adapt	high	18	20	2.08E+03	5.70E+02	(1201.67, 3.44E+03)	2.02E+03	4.20E+00	1.82E+01	2.16E+03	6.02E+02	(1203.90, 3.60E+03)	2.09E+03	7.94E+00	1.97E+01
adapt	high	24	15	2.09E+03	5.78E+02	(1145.66, 3.46E+03)	2.01E+03	4.26E+00	1.85E+01	2.16E+03	6.15E+02	(1159.66, 3.58E+03)	2.08E+03	7.91E+00	2.01E+01
adapt	high	40	9	2.11E+03	5.87E+02	(1135.25, 3.48E+03)	2.04E+03	5.40E+00	1.89E+01	2.16E+03	6.04E+02	(1161.62, 3.62E+03)	2.08E+03	7.86E+00	1.97E+01
adapt	high	60	6	2.11E+03	5.63E+02	(1136.85, 3.40E+03)	2.06E+03	5.73E+00	1.82E+01	2.16E+03	5.81E+02	(1163.95, 3.49E+03)	2.10E+03	8.03E+00	1.90E+01
unif	low	6	60	1.85E+08	3.12E+09	(151.83, 6478.27)	1.46E+03	9.23E+06	9.90E+07	6.93E+100	1.53E+102	(909.25, 3.91E+89)	3.96E+03	3.46E+99	4.85E+100
unif	low	10	36	1.51E+09	8.13E+09	(672.90, 2.46E+10)	2.03E+03	7.54E+07	2.61E+08	5.02E+75	1.25E+77	(829.78, 1.50E+11)	2.62E+03	2.51E+74	3.94E+75
unif	low	18	20	1.94E+09	7.92E+09	(694.00, 2.81E+10)	2.04E+03	9.69E+07	2.58E+08	4.80E+09	6.01E+10	(745.93, 4.71E+10)	2.44E+03	2.40E+08	1.91E+09
unif	low	24	15	2.12E+09	7.13E+09	(728.31, 2.79E+10)	2.15E+03	1.06E+08	2.35E+08	4.19E+09	2.43E+10	(759.03, 3.99E+10)	2.52E+03	2.09E+08	7.80E+08
unif	low	40	9	2.33E+09	6.20E+09	(714.16, 2.49E+10)	1.99E+03	1.17E+08	2.10E+08	3.30E+09	9.57E+09	(735.78, 3.27E+10)	2.30E+03	1.65E+08	3.20E+08
unif	low	60	6	2.51E+09	7.12E+09	(660.89, 2.63E+10)	2.17E+03	1.25E+08	2.39E+08	3.88E+09	1.46E+10	(690.16, 3.90E+10)	2.38E+03	1.94E+08	4.77E+08
unif	med	6	60	1.55E+03	1.10E+03	(196.99, 3839.90)	1.62E+03	-2.25E+01	3.75E+01	1.21E+122	3.83E+123	(1149.28, 1.95E+85)	3.05E+03	6.06E+120	1.21E+122
unif	med	10	36	2.13E+03	7.96E+02	(853.12, 4056.21)	2.02E+03	6.60E+00	2.55E+01	1.21E+68	3.82E+69	(979.62, 8691.42)	2.26E+03	6.07E+66	1.21E+68
unif	med	18	20	2.21E+03	8.88E+02	(943.20, 4323.30)	2.04E+03	1.03E+01	2.88E+01	2.50E+03	1.82E+03	(971.85, 5848.45)	2.21E+03	2.52E+01	5.98E+01
unif	med	24	15	2.20E+03	1.03E+03	(1033.54, 4216.20)	2.04E+03	9.97E+00	3.31E+01	2.39E+03	1.18E+03	(1050.94, 4877.80)	2.16E+03	1.97E+01	3.95E+01
unif	med	40	9	2.22E+03	9.47E+02	(997.70, 4374.82)	2.07E+03	1.09E+01	3.07E+01	2.38E+03	1.26E+03	(1009.09, 5175.12)	2.15E+03	1.89E+01	4.17E+01
unif	med	60	6	2.25E+03	9.28E+02	(1021.18, 4772.98)	2.06E+03	1.23E+01	3.04E+01	2.35E+03	1.07E+03	(1048.68, 5146.91)	2.13E+03	1.76E+01	3.55E+01
unif	high	6	60	1.51E+03	9.09E+02	(263.07, 3342.32)	1.57E+03	-2.46E+01	3.27E+01	5.83E+128	1.84E+130	(1271.94, 9.84E+90)	3.43E+03	2.92E+127	5.83E+128
unif	high	10	36	2.11E+03	6.28E+02	(1104.05, 3615.70)	2.02E+03	5.46E+00	2.02E+01	3.18E+69	9.01E+70	(1160.13, 6369.41)	2.27E+03	1.59E+68	2.85E+69
unif	high	18	20	2.11E+03	6.18E+02	(1072.31, 3571.44)	2.04E+03	5.69E+00	1.99E+01	2.29E+03	7.88E+02	(1160.63, 4220.86)	2.17E+03	1.47E+01	2.66E+01
unif	high	24	15	2.15E+03	6.04E+02	(1159.34, 3494.04)	2.08E+03	7.39E+00	1.97E+01	2.31E+03	8.00E+02	(1223.10, 3902.79)	2.20E+03	1.57E+01	2.72E+01
unif	high	40	9	2.15E+03	6.54E+02	(1183.74, 3680.60)	2.06E+03	7.58E+00	2.12E+01	2.26E+03	7.58E+02	(1192.11, 3996.23)	2.13E+03	1.30E+01	2.53E+01
unif	high	60	6	2.12E+03	6.55E+02	(1124.99, 3708.48)	2.03E+03	6.17E+00	2.11E+01	2.19E+03	7.18E+02	(1141.40, 3836.94)	2.09E+03	9.68E+00	2.35E+01

Method	model	No.tri	No.ind	mean $\hat{N}_2$	sd $\hat{N}_2$	95%( $\hat{N}_2$ )	median $\hat{N}_2$	pc.bias $\hat{N}_2$	rmse $\hat{N}_2$	mean $\hat{N}_3$	sd $\hat{N}_3$	95%( $\hat{N}_3$ )	median $\hat{N}_3$	pc.bias $\hat{N}_3$	rmse $\hat{N}_3$
SR5	low	6	60	1.41E+08	2.41E+09	(165.91, 7090.93)	1.40E+03	7.03E+06	7.64E+07	9.43E+108	2.98E+110	(837.50, 8.89E+89)	3.73E+03	4.71E+107	9.43E+108
SR5	low	10	36	1.03E+09	6.21E+09	(784.07, 1.43E+10)	2.11E+03	5.13E+07	1.99E+08	6.75E+89	2.13E+91	(866.03, 2.23E+10)	2.58E+03	3.38E+88	6.75E+89
SR5	low	18	20	1.75E+09	6.89E+09	(766.14, 2.56E+10)	2.08E+03	8.76E+07	2.25E+08	3.20E+09	1.66E+10	(826.28, 4.07E+10)	2.47E+03	1.60E+08	5.34E+08
SR5	low	24	15	2.59E+09	8.54E+09	(732.78, 3.43E+10)	2.15E+03	1.29E+08	2.82E+08	3.56E+09	1.25E+10	(777.88, 4.33E+10)	2.50E+03	1.78E+08	4.10E+08
SR5	low	40	9	3.17E+09	7.34E+09	(646.32, 2.13E+10)	2.18E+03	1.59E+08	2.53E+08	8.08E+09	7.74E+10	(676.69, 3.64E+10)	2.51E+03	4.04E+08	2.46E+09
SR5	low	60	6	2.74E+09	6.97E+09	(639.71, 2.63E+10)	2.00E+03	1.37E+08	2.37E+08	3.87E+09	1.19E+10	(680.03, 3.36E+10)	2.24E+03	1.93E+08	3.94E+08
SR5	med	6	60	1.67E+03	2.62E+03	(199.30, 3681.22)	1.62E+03	-1.64E+01	8.36E+01	4.88E+98	1.54E+100	(1162.79, 1.08E+86)	2.99E+03	2.44E+97	4.88E+98
SR5	med	10	36	2.17E+03	7.84E+02	(997.28, 4113.43)	2.02E+03	8.25E+00	2.53E+01	1.12E+60	2.75E+61	(1075.91, 7443.74)	2.28E+03	5.61E+58	8.69E+59
SR5	med	18	20	2.18E+03	8.13E+02	(989.48, 4045.60)	2.04E+03	8.93E+00	2.63E+01	2.50E+03	1.79E+03	(1033.49, 5447.50)	2.22E+03	2.52E+01	5.88E+01
SR5	med	24	15	2.16E+03	8.08E+02	(970.19, 4190.78)	2.04E+03	8.15E+00	2.61E+01	2.49E+03	2.67E+03	(1007.62, 5333.99)	2.19E+03	2.45E+01	8.59E+01
SR5	med	40	9	2.28E+03	9.57E+02	(947.80, 4595.03)	2.05E+03	1.39E+01	3.15E+01	2.45E+03	1.17E+03	(960.39, 5572.00)	2.16E+03	2.24E+01	3.97E+01
SR5	med	60	6	2.35E+03	1.35E+03	(889.37, 5545.93)	2.05E+03	1.76E+01	4.42E+01	2.50E+03	1.63E+03	(911.17, 6876.23)	2.10E+03	2.51E+01	5.41E+01
SR5	high	6	60	1.47E+03	8.93E+02	(265.32, 3292.31)	1.50E+03	-2.66E+01	3.29E+01	3.55E+141	1.04E+143	(1215.04, 7.07E+91)	3.34E+03	1.77E+140	3.30E+141
SR5	high	10	36	2.13E+03	5.94E+02	(1136.31, 3471.16)	2.08E+03	6.55E+00	1.92E+01	9.61E+50	3.04E+52	(1227.02, 5101.37)	2.24E+03	4.81E+49	9.61E+50
SR5	high	18	20	2.13E+03	6.10E+02	(1147.80, 3524.16)	2.07E+03	6.48E+00	1.97E+01	2.33E+03	8.18E+02	(1212.79, 4212.41)	2.18E+03	1.63E+01	2.79E+01
SR5	high	24	15	2.13E+03	6.54E+02	(1142.25, 3649.73)	2.03E+03	6.32E+00	2.11E+01	2.29E+03	8.74E+02	(1194.31, 4306.77)	2.10E+03	1.47E+01	2.92E+01
SR5	high	40	9	2.12E+03	6.84E+02	(1098.38, 3683.49)	2.02E+03	5.91E+00	2.19E+01	2.23E+03	7.80E+02	(1114.58, 4207.15)	2.09E+03	1.13E+01	2.57E+01
SR5	high	60	6	2.11E+03	6.91E+02	(1077.20, 3773.72)	2.01E+03	5.45E+00	2.21E+01	2.18E+03	7.49E+02	(1077.20, 4029.44)	2.05E+03	8.76E+00	2.43E+01
SR8	low	6	60	1.28E+08	2.34E+09	(163.78, 6540.87)	1.44E+03	6.39E+06	7.41E+07	3.25E+137	1.03E+139	(866.04, 2.31E+92)	3.80E+03	1.63E+136	3.25E+137
SR8	low	10	36	1.12E+09	7.64E+09	(643.70, 1.59E+10)	1.98E+03	5.59E+07	2.44E+08	3.52E+167	Inf	(738.65, 1.22E+11)	2.52E+03	1.76E+166	Inf
SR8	low	18	20	2.04E+09	8.29E+09	(735.06, 3.00E+10)	2.17E+03	1.02E+08	2.70E+08	3.81E+09	2.59E+10	(834.97, 4.42E+10)	2.58E+03	1.91E+08	8.28E+08
SR8	low	24	15	1.54E+09	4.99E+09	(704.09, 1.78E+10)	2.07E+03	7.70E+07	1.65E+08	2.91E+09	1.83E+10	(768.43, 2.61E+10)	2.34E+03	1.45E+08	5.87E+08
SR8	low	40	9	3.15E+09	8.17E+09	(684.64, 2.93E+10)	2.11E+03	1.57E+08	2.77E+08	9.48E+23	3.00E+25	(735.36, 5.07E+10)	2.45E+03	4.74E+22	9.48E+23
SR8	low	60	6	2.75E+09	9.05E+09	(708.31, 3.02E+10)	2.22E+03	1.37E+08	2.99E+08	4.22E+09	1.60E+10	(778.63, 4.86E+10)	2.44E+03	2.11E+08	5.22E+08
SR8	med	6	60	1.58E+03	1.04E+03	(188.10, 3750.20)	1.65E+03	-2.09E+01	3.55E+01	5.66E+98	1.77E+100	(1113.66, 4.60E+86)	2.93E+03	2.83E+97	5.61E+98
SR8	med	10	36	2.18E+03	8.16E+02	(1052.90, 4169.64)	2.03E+03	8.79E+00	2.64E+01	1.01E+63	3.20E+64	(1160.90, 6892.64)	2.27E+03	5.06E+61	1.01E+63
SR8	med	18	20	2.18E+03	9.36E+02	(956.14, 4304.38)	2.02E+03	9.07E+00	3.02E+01	2.50E+03	1.85E+03	(1005.77, 5596.82)	2.16E+03	2.50E+01	6.06E+01
SR8	med	24	15	2.18E+03	8.35E+02	(979.49, 4319.71)	2.01E+03	9.13E+00	2.70E+01	2.40E+03	1.17E+03	(1014.91, 5048.49)	2.16E+03	2.02E+01	3.91E+01
SR8	med	40	9	2.29E+03	9.76E+02	(980.03, 4605.96)	2.14E+03	1.44E+01	3.22E+01	2.45E+03	1.46E+03	(991.14, 5291.63)	2.20E+03	2.26E+01	6.49E+36
SR8	med	60	6	2.31E+03	1.07E+03	(970.81, 5010.39)	2.08E+03	1.56E+01	3.52E+01	2.43E+03	1.22E+03	(1021.86, 5415.78)	2.15E+03	2.13E+01	4.10E+01
SR8	high	6	60	1.48E+03	9.15E+02	(265.53, 3289.15)	1.53E+03	-2.59E+01	3.32E+01	1.33E+104	4.21E+105	(1245.17, 7.09E+88)	3.41E+03	6.66E+102	1.33E+104
SR8	high	10	36	2.11E+03	6.38E+02	(1089.93, 3691.78)	2.03E+03	5.30E+00	2.05E+01	1.41E+74	4.44E+75	(1166.76, 5137.89)	2.25E+03	7.03E+72	1.41E+74
SR8	high	18	20	2.16E+03	6.31E+02	(1123.10, 3585.78)	2.08E+03	8.19E+00	2.06E+01	2.34E+03	8.31E+02	(1145.19, 4330.52)	2.19E+03	1.72E+01	2.84E+01
SR8	high	24	15	2.14E+03	6.42E+02	(1129.81, 3684.79)	2.06E+03	7.10E+00	2.08E+01	2.28E+03	7.46E+02	(1165.64, 4044.16)	2.15E+03	1.40E+01	2.52E+01
SR8	high	40	9	2.17E+03	7.03E+02	(1106.15, 3791.73)	2.06E+03	8.47E+00	2.29E+01	2.29E+03	8.09E+02	(1128.53, 4237.65)	2.15E+03	1.44E+01	2.72E+01
SR8	high	60	6	2.16E+03	6.59E+02	(1105.06, 3715.89)	2.09E+03	7.93E+00	2.14E+01	2.23E+03	7.28E+02	(1127.29, 3855.36)	2.12E+03	1.15E+01	2.42E+01

Table A.4: Table of results for all scenarios in Simulation 2. That is,  $\hat{N}_2$  and  $\hat{N}_3$  with associated percentage bias (pc.bias), standard deviation estimates (sd), percentile intervals (PI), median and root mean square error (RMSE) when underlying heterogeneity is *not* accounted for, using the “High”, “Medium” and “Low” detection function scenarios, with differing effort allocated per individual (No.tri) and number of individuals (No.ind) in the trials survey.

Method	model	No.tri	No.ind	mean $\hat{N}_2$	sd $\hat{N}_2$	95%( $\hat{N}_2$ )	median $\hat{N}_2$	pc.bias $\hat{N}_2$	rmse $\hat{N}_2$	mean $\hat{N}_3$	sd $\hat{N}_3$	95%( $\hat{N}_3$ )	median $\hat{N}_3$	pc.bias $\hat{N}_3$	rmse $\hat{N}_3$
adapt	low	6	60	2.95E+03	4.57E+03	(730.35, 17137.05)	1.55E+03	4.77E+01	1.48E+02	4.82E+03	7.59E+03	(943.20, 29916.15)	2.38E+03	1.41E+02	2.56E+02
adapt	low	10	36	1.46E+03	7.42E+02	(681.82, 2598.27)	1.37E+03	-2.72E+01	2.91E+01	2.11E+03	1.16E+03	(879.47, 4189.00)	1.92E+03	5.38E+00	3.69E+01
adapt	low	18	20	1.41E+03	4.77E+02	(693.64, 2641.35)	1.36E+03	-2.95E+01	2.40E+01	2.04E+03	8.81E+02	(862.55, 4338.81)	1.88E+03	1.76E+00	2.79E+01
adapt	low	24	15	1.40E+03	4.77E+02	(616.04, 2511.75)	1.36E+03	-2.98E+01	2.41E+01	1.93E+03	7.66E+02	(816.56, 3797.42)	1.82E+03	-3.34E+00	2.43E+01
adapt	low	40	9	1.37E+03	4.45E+02	(682.97, 2463.69)	1.31E+03	-3.17E+01	2.45E+01	1.82E+03	6.70E+02	(885.51, 3530.38)	1.70E+03	-9.17E+00	2.20E+01
adapt	low	60	6	1.42E+03	4.85E+02	(667.97, 2533.47)	1.36E+03	-2.89E+01	2.38E+01	2.26E+03	1.17E+03	(876.25, 5362.75)	2.00E+03	1.30E+01	3.79E+01
adapt	med	6	60	4.14E+03	6.02E+03	(970.89, 22649.60)	2.17E+03	1.07E+02	2.02E+02	4.79E+03	7.06E+03	(1117.63, 27327.16)	2.51E+03	1.40E+02	2.40E+02
adapt	med	10	36	1.93E+03	6.15E+02	(982.57, 3213.74)	1.85E+03	-3.69E+00	1.96E+01	2.21E+03	7.69E+02	(1042.62, 3998.72)	2.09E+03	1.05E+01	2.52E+01
adapt	med	18	20	1.86E+03	5.49E+02	(942.04, 3086.17)	1.81E+03	-6.95E+00	1.79E+01	2.12E+03	6.72E+02	(1075.92, 3647.69)	2.03E+03	6.13E+00	2.16E+01
adapt	med	24	15	1.88E+03	6.26E+02	(910.59, 3340.78)	1.79E+03	-5.85E+00	2.01E+01	2.15E+03	7.51E+02	(1020.71, 3979.08)	2.01E+03	7.40E+00	2.42E+01
adapt	med	40	9	1.90E+03	5.49E+02	(970.93, 3097.51)	1.86E+03	-5.21E+00	1.77E+01	2.15E+03	6.54E+02	(1085.60, 3579.23)	2.08E+03	7.32E+00	2.12E+01
adapt	med	60	6	1.86E+03	5.52E+02	(922.04, 3122.26)	1.83E+03	-6.95E+00	1.80E+01	2.11E+03	6.62E+02	(1018.70, 3656.03)	2.04E+03	5.50E+00	2.12E+01
unif	low	6	60	9.65E+02	6.67E+02	(145.58, 2385.00)	9.84E+02	-5.18E+01	3.90E+01	2.27E+101	7.18E+102	(825.64, 9.60E+83)	3.32E+03	1.14E+100	2.27E+101
unif	low	10	36	1.40E+03	4.75E+02	(636.71, 2560.12)	1.34E+03	-2.99E+01	2.42E+01	6.04E+69	1.91E+71	(815.00, 6721.11)	1.86E+03	3.02E+68	6.04E+69
unif	low	18	20	1.42E+03	4.78E+02	(612.67, 2512.77)	1.36E+03	-2.91E+01	2.38E+01	2.19E+03	1.74E+03	(780.27, 5360.63)	1.87E+03	9.62E+00	5.53E+01
unif	low	24	15	1.39E+03	4.70E+02	(628.12, 2504.34)	1.34E+03	-3.07E+01	2.44E+01	2.05E+03	1.23E+03	(785.96, 5250.20)	1.75E+03	2.40E+00	3.88E+01
unif	low	40	9	1.35E+03	4.98E+02	(609.47, 2479.56)	1.29E+03	-3.24E+01	2.59E+01	1.86E+03	1.22E+03	(709.42, 4174.29)	1.65E+03	-7.20E+00	3.87E+01
unif	low	60	6	1.42E+03	5.38E+02	(614.48, 2778.85)	1.35E+03	-2.91E+01	2.51E+01	2.26E+03	1.89E+03	(799.47, 6265.69)	1.85E+03	1.28E+01	6.04E+01
unif	med	6	60	1.29E+03	8.51E+02	(185.87, 2977.95)	1.33E+03	-3.54E+01	3.50E+01	1.23E+106	3.89E+107	(1105.48, 5.65E+84)	3.74E+03	6.15E+104	1.23E+106
unif	med	10	36	1.87E+03	5.73E+02	(936.25, 3182.02)	1.80E+03	-6.61E+00	1.86E+01	5.52E+09	1.74E+11	(1049.78, 5726.25)	2.13E+03	2.76E+08	5.52E+09
unif	med	18	20	1.86E+03	5.90E+02	(872.53, 3200.87)	1.81E+03	-6.98E+00	1.92E+01	2.27E+03	1.09E+03	(944.30, 5144.90)	2.06E+03	1.37E+01	3.55E+01
unif	med	24	15	1.89E+03	5.74E+02	(907.30, 3243.96)	1.84E+03	-5.59E+00	1.85E+01	2.23E+03	8.70E+02	(1071.66, 4447.49)	2.09E+03	1.16E+01	2.85E+01
unif	med	40	9	1.85E+03	6.36E+02	(848.40, 3272.29)	1.75E+03	-7.71E+00	2.07E+01	2.15E+03	8.70E+02	(925.21, 4311.18)	2.02E+03	7.70E+00	2.79E+01
unif	med	60	6	1.89E+03	6.57E+02	(925.45, 3593.87)	1.79E+03	-5.38E+00	2.11E+01	2.25E+03	1.19E+03	(968.47, 4784.40)	2.03E+03	1.23E+01	3.83E+01

Method	model	No.tri	No.ind	mean $N_2$	sd $N_2$	95%( $N_2$ )	median $N_2$	pc.bias $N_2$	rmse $N_2$	mean $N_3$	sd $N_3$	95%( $N_3$ )	median $N_3$	pc.bias $N_3$	rmse $N_3$
SR5	low	6	60	9.62E+02	6.50E+02	(144.88, 2273.67)	9.81E+02	-5.19E+01	3.87E+01	3.18E+90	9.18E+91	(784.43, 2.26E+82)	3.24E+03	1.59E+89	2.90E+90
SR5	low	10	36	1.39E+03	4.53E+02	(649.95, 2369.58)	1.33E+03	-3.07E+01	2.41E+01	1.59E+73	4.74E+74	(835.98, 7192.34)	1.87E+03	7.93E+71	1.50E+73
SR5	low	18	20	1.41E+03	5.04E+02	(611.09, 2624.26)	1.34E+03	-2.93E+01	2.45E+01	2.29E+03	2.26E+03	(753.56, 6036.51)	1.83E+03	1.43E+01	7.21E+01
SR5	low	24	15	1.40E+03	4.93E+02	(593.76, 2586.11)	1.33E+03	-3.01E+01	2.46E+01	2.19E+03	1.55E+03	(765.14, 6426.60)	1.81E+03	9.31E+00	4.92E+01
SR5	low	40	9	1.40E+03	5.50E+02	(615.61, 2696.10)	1.31E+03	-3.02E+01	2.58E+01	2.12E+03	3.47E+03	(732.93, 5101.81)	1.66E+03	5.98E+00	1.10E+02
SR5	low	60	6	1.40E+03	5.76E+02	(579.54, 2783.09)	1.29E+03	-3.01E+01	2.64E+01	2.26E+03	3.29E+03	(695.86, 6331.70)	1.70E+03	1.32E+01	1.05E+02
SR5	med	6	60	1.29E+03	8.78E+02	(191.47, 3018.60)	1.33E+03	-3.55E+01	3.57E+01	7.36E+97	2.33E+99	(1108.14, 2.83E+85)	3.74E+03	3.68E+96	7.36E+97
SR5	med	10	36	1.87E+03	5.88E+02	(944.01, 3197.58)	1.78E+03	-6.31E+00	1.90E+01	1.35E+61	4.25E+62	(1074.68, 6942.28)	2.10E+03	6.73E+59	1.35E+61
SR5	med	18	20	1.86E+03	5.84E+02	(916.96, 3185.07)	1.80E+03	-7.06E+00	1.90E+01	2.80E+03	1.60E+04	(986.12, 5147.26)	2.07E+03	3.99E+01	5.06E+02
SR5	med	24	15	1.89E+03	6.25E+02	(932.69, 3287.72)	1.80E+03	-5.36E+00	2.01E+01	2.30E+03	1.05E+03	(1000.17, 5185.66)	2.09E+03	1.49E+01	3.44E+01
SR5	med	40	9	1.94E+03	7.51E+02	(839.61, 3760.21)	1.80E+03	-3.11E+00	2.38E+01	2.39E+03	1.61E+03	(885.01, 5968.99)	2.08E+03	1.94E+01	5.23E+01
SR5	med	60	6	1.92E+03	8.22E+02	(783.80, 4047.65)	1.77E+03	-3.87E+00	2.61E+01	2.47E+03	2.31E+03	(800.48, 7260.32)	1.97E+03	2.33E+01	7.45E+01
SR8	low	6	60	9.87E+02	6.83E+02	(144.45, 2447.05)	9.78E+02	-5.06E+01	3.86E+01	2.93E+95	9.27E+96	(777.28, 1.59E+82)	3.29E+03	1.47E+94	2.93E+95
SR8	low	10	36	1.40E+03	4.77E+02	(652.27, 2504.98)	1.34E+03	-2.99E+01	2.42E+01	1.24E+66	3.92E+67	(823.73, 10114.4)	1.91E+03	6.20E+64	1.24E+66
SR8	low	18	20	1.43E+03	4.88E+02	(653.26, 2576.91)	1.36E+03	-2.83E+01	2.36E+01	2.64E+03	8.50E+03	(829.48, 6851.78)	1.89E+03	3.18E+01	2.70E+02
SR8	low	24	15	1.40E+03	4.98E+02	(642.76, 2631.95)	1.32E+03	-2.98E+01	2.46E+01	2.11E+03	1.81E+03	(803.90, 4891.95)	1.83E+03	5.32E+00	5.74E+01
SR8	low	40	9	1.39E+03	5.03E+02	(619.11, 2610.54)	1.33E+03	-3.06E+01	2.51E+01	2.00E+03	1.50E+03	(740.20, 4775.72)	1.70E+03	2.23E-01	4.74E+01
SR8	low	60	6	1.49E+03	5.71E+02	(644.75, 2975.07)	1.39E+03	-2.55E+01	2.42E+01	2.30E+03	1.56E+03	(806.88, 6426.17)	1.93E+03	1.50E+01	5.03E+01
SR8	med	6	60	1.34E+03	8.83E+02	(193.13, 3083.25)	1.43E+03	-3.28E+01	3.48E+01	6.55E+175	Inf	(1106.19, 5.19E+83)	3.60E+03	3.27E+174	Inf
SR8	med	10	36	1.88E+03	5.90E+02	(955.01, 3199.69)	1.81E+03	-5.85E+00	1.90E+01	1.15E+75	3.64E+76	(1022.38, 5773.81)	2.18E+03	5.75E+73	1.15E+75
SR8	med	18	20	1.88E+03	6.24E+02	(917.11, 3458.33)	1.78E+03	-5.79E+00	2.01E+01	2.39E+03	2.21E+03	(1012.82, 5148.86)	2.10E+03	1.94E+01	7.10E+01
SR8	med	24	15	1.92E+03	5.99E+02	(926.97, 3317.25)	1.85E+03	-4.13E+00	1.91E+01	2.29E+03	9.68E+02	(1051.08, 4767.00)	2.11E+03	1.47E+01	3.20E+01
SR8	med	40	9	1.95E+03	6.72E+02	(925.28, 3551.23)	1.85E+03	-2.51E+00	2.13E+01	2.37E+03	1.35E+03	(1013.99, 5432.76)	2.09E+03	1.86E+01	4.43E+01
SR8	med	60	6	1.98E+03	6.98E+02	(966.69, 3620.12)	1.86E+03	-1.11E+00	2.21E+01	2.43E+03	1.59E+03	(1002.46, 5433.14)	2.15E+03	2.17E+01	5.20E+01

Table A.5: Table of results for all scenarios in Simulation 3. That is,  $\hat{N}_2$  and  $\hat{N}_3$  with associated percentage bias (pc.bias), standard deviation estimates (sd), 2.5 and 97.5 percentile intervals (PI), median and root mean square error (RMSE) when underlying heterogeneity is accounted for (Model=Include) or not (Model=Ignore), using the “Medium” detection function scenario, with differing effort allocated per individual (No.tri) and number of individuals (No.ind) in the trials survey.

Model	No.tri	No.ind	mean $\hat{N}_2$	95%( $\hat{N}_2$ )	median $\hat{N}_2$	pc.bias $\hat{N}_2$	rmse $\hat{N}_2$	mean $\hat{N}_3$	95%( $\hat{N}_3$ )	rmse $\hat{N}_3$	median $\hat{N}_3$	pc.bias $\hat{N}_3$
Include	6	60	4721.30	(1105,21329)	2517.74	136.06	221.56	5013.46	(1161,22572)	235.25	2671.42	150.67
Include	10	36	2197.19	(1019,4150)	2083.09	9.86	30.35	2325.65	(1070,4444)	34.01	2191.02	16.28
Include	18	20	2145.25	(1072,3754)	2021.13	7.26	23.72	2246.49	(1093,4129)	26.52	2112.75	12.32
Include	24	15	2151.91	(1101,3728)	2062.29	7.60	22.63	2241.93	(1114,4024)	24.93	2129.12	12.10
Include	40	9	2141.48	(988,3712)	2038.09	7.07	23.31	2213.02	(1016,3866)	24.97	2092.10	10.65
Include	60	6	2174.63	(1031,3914)	2051.54	8.73	23.99	2242.06	(1038,4046)	25.97	2107.60	12.10
Include	6	120	2681.54	(1084,6787)	2239.36	34.08	77.32	2826.87	(1118,6787)	82.25	2342.66	41.34
Include	10	72	2091.27	(1025,3534)	2045.04	4.56	19.79	2195.88	(1086,3754)	21.92	2133.59	9.79
Include	18	40	2031.30	(1100,3275)	1968.64	1.57	17.28	2120.96	(1128,3458)	18.99	2053.40	6.05
Include	24	30	2050.05	(1057,3352)	1992.68	2.50	18.67	2136.73	(1080,3534)	20.28	2067.47	6.84
Include	40	18	2042.20	(1003,3366)	1998.90	2.11	18.33	2122.15	(1031,3511)	19.76	2066.54	6.11
Include	60	12	2025.69	(984,3418)	1980.02	1.28	18.92	2092.76	(1000,3529)	19.94	2040.31	4.64
Include	6	180	2785.79	(1013,10156)	2206.99	39.29	88.18	2914.88	(1075,10396)	92.99	2317.74	45.74
Include	10	108	2056.51	(1088,3297)	2021.57	2.83	17.81	2159.97	(1119,3530)	19.74	2113.28	8.00
Include	18	60	2070.58	(1135,3252)	2026.26	3.53	17.79	2162.29	(1182,3353)	19.46	2116.45	8.11
Include	24	45	2042.81	(1006,3233)	2020.22	2.14	17.47	2127.61	(1056,3354)	18.83	2101.17	6.38
Include	40	27	2021.28	(1064,3189)	1965.51	1.06	16.69	2095.44	(1093,3293)	17.71	2043.94	4.77
Include	60	18	2039.94	(1088,3261)	1990.39	2.00	18.10	2114.03	(1113,3405)	19.33	2059.71	5.70
Include	6	240	2567.49	(1083,9320)	2149.25	28.37	65.03	2681.75	(1119,9320)	68.89	2246.67	34.08
Include	10	144	2038.89	(1047,3245)	2023.71	1.94	17.17	2126.32	(1076,3404)	18.55	2102.14	6.31
Include	18	80	2022.89	(1087,3271)	1999.87	1.14	17.24	2111.91	(1149,3458)	18.69	2079.23	5.59
Include	24	60	2042.61	(1071,3292)	2016.30	2.13	17.98	2127.53	(1108,3478)	19.38	2089.67	6.37
Include	40	36	2018.92	(1084,3187)	1982.37	0.95	17.09	2099.17	(1120,3333)	18.19	2056.66	4.95
Include	60	24	2037.49	(1014,3201)	2002.95	1.87	17.23	2116.57	(1061,3321)	18.36	2083.55	5.82

Model	No.tri	No.ind	mean $\hat{N}_2$	95%( $\hat{N}_2$ )	median $\hat{N}_2$	pc.bias $\hat{N}_2$	rmse $\hat{N}_2$	mean $\hat{N}_3$	95%( $\hat{N}_3$ )	rmse $\hat{N}_3$	median $\hat{N}_3$	pc.bias $\hat{N}_3$
Include	6	60	4138.80	(970,22649)	2172.97	106.94	202.16	4790.94	(1117,27327)	240.22	2513.07	139.55
Include	10	36	1926.23	(982,3213)	1852.13	-3.69	19.58	2210.61	(1042,3998)	25.21	2091.38	10.53
Include	18	20	1861.06	(942,3086)	1808.82	-6.95	17.90	2122.67	(1075,3647)	21.60	2030.25	6.13
Include	24	15	1883.09	(910,3340)	1790.55	-5.85	20.14	2147.98	(1020,3979)	24.20	2011.39	7.40
Include	40	9	1895.71	(970,3097)	1861.21	-5.21	17.68	2146.36	(1085,3579)	21.19	2083.51	7.32
Include	60	6	1861.07	(922,3122)	1828.71	-6.95	18.00	2110.00	(1018,3656)	21.21	2042.09	5.50
Include	6	120	3376.35	(1019,15964)	2102.91	68.82	132.51	3841.16	(1113,18804)	156.14	2408.43	92.06
Include	10	72	1861.17	(997,2943)	1821.28	-6.94	16.35	2122.90	(1124,3635)	19.69	2057.66	6.15
Include	18	40	1837.20	(982,2959)	1783.94	-8.14	16.50	2088.74	(1129,3410)	18.58	2027.25	4.44
Include	24	30	1851.55	(1026,2890)	1806.94	-7.42	15.54	2107.24	(1127,3385)	17.86	2042.88	5.36
Include	40	18	1811.50	(975,2906)	1766.42	-9.43	16.64	2063.67	(1107,3361)	18.42	2009.24	3.18
Include	60	12	1819.48	(994,2896)	1791.85	-9.03	16.18	2076.43	(1122,3363)	18.38	2019.72	3.82
Include	6	180	2910.77	(1050,14265)	2026.35	45.54	101.58	3293.62	(1186,16478)	118.65	2268.44	64.68
Include	10	108	1838.38	(1030,2956)	1810.21	-8.08	15.83	2087.77	(1101,3368)	17.77	2054.92	4.39
Include	18	60	1833.70	(1075,2778)	1802.39	-8.31	15.09	2094.25	(1210,3212)	17.02	2048.91	4.71
Include	24	45	1823.66	(940,2898)	1777.21	-8.82	16.48	2078.56	(1077,3375)	18.33	2037.35	3.93
Include	40	27	1818.59	(936,2848)	1788.99	-9.07	16.12	2069.16	(1063,3272)	17.62	2042.81	3.46
Include	60	18	1828.61	(968,2819)	1798.27	-8.57	16.01	2089.18	(1113,3251)	17.81	2039.85	4.46
Include	6	240	2651.19	(1046,12777)	1961.28	32.56	88.57	2997.83	(1144,14290)	103.07	2197.53	49.89
Include	10	144	1854.96	(1010,2935)	1824.07	-7.25	15.89	2105.07	(1149,3352)	17.84	2069.63	5.25
Include	18	80	1844.56	(1001,2735)	1826.75	-7.77	14.91	2091.64	(1142,3090)	16.59	2088.20	4.58
Include	24	60	1819.84	(1010,2691)	1790.57	-9.01	14.92	2075.60	(1131,3151)	16.47	2043.42	3.78
Include	40	36	1830.37	(1044,2776)	1808.92	-8.48	15.17	2085.76	(1183,3205)	16.73	2072.80	4.29
Include	60	24	1805.61	(941,2813)	1790.79	-9.72	15.79	2060.07	(1075,3179)	16.77	2036.91	3.00