



# Methods in Ecology and Evolution

1 **Running title:** A spatial capture-recapture model for passive acoustic surveys  
2 **Number of words:** 6948  
3 **Number of tables:** 2  
4 **Number of figures:** 2  
5 **Number of references:** 42

6 **A spatial capture-recapture model to estimate**  
7 **call rate and population density from passive**  
8 **acoustic surveys**

9 **Ben C. Stevenson<sup>1\*</sup> Paul van Dam-Bates<sup>2</sup>**  
10 **Callum K. Y. Young<sup>1,3†</sup> John Measey<sup>4</sup>**

11 *1. Department of Statistics, University of Auckland, Auckland, New Zealand*  
12 *2. School of Mathematics and Statistics, University of St Andrews, St Andrews,*  
13 *United Kingdom*  
14 *3. Ministry of Business, Innovation, and Employment, Auckland, New Zealand*  
15 *4. Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch*  
16 *University, Stellenbosch, South Africa*

17 \*Corresponding author.  
18 Ben C. Stevenson  
19 ben.stevenson@auckland.ac.nz  
20 +64 9 923 8474  
21  
22 † Present address

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/2041-210X.13522](https://doi.org/10.1111/2041-210X.13522)

This article is protected by copyright. All rights reserved

Accepted Article

23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

## Summary

1. Spatial capture-recapture (SCR) models are commonly used to estimate animal population density from detections and subsequent redetections of individuals across space. In particular, acoustic SCR models deal with detections of animal vocalisations across an array of acoustic detectors. Previously published acoustic SCR methods either estimate call density (calls per unit space per unit time) rather than animal density itself, require an independently estimated call rate to estimate animal density, or discard data from all but one detected call from each individual.
2. In this manuscript, we develop a new spatial capture-recapture model that estimates both call rate and animal density from the acoustic survey alone, without requiring an independently estimated call rate. Our approach therefore alleviates the need for the additional fieldwork of physically locating and monitoring individual animals. We illustrate our method and compare it to an existing approach using a simulation study and an application to data collected on an acoustic survey of the visually cryptic Cape peninsula moss frog *Arthroleptella lightfooti*.
3. In the context of our acoustic survey, our calling animal density estimator has low bias, good precision, and confidence intervals with appropriate coverage, yielding results that are consistent with previous studies of the same species.
4. Our method can obtain accurate and precise estimates of animal density while eliminating the fieldwork burden associated with separately estimating call rate. We discuss how the development of our model's likelihood reveals a clear path to further extensions, which may incorporate features such as animal movement processes and

52 uncertain individual identification.

## 53 **1 Introduction**

54 Passive acoustic surveys are becoming increasingly popular for monitoring an-  
55 imal populations because they offer a fast, cheap, safe, and noninvasive altern-  
56 ative to relying on visual sightings or physical capture (Marques et al., 2013).  
57 Although there has been a particular focus on passive acoustic surveys for the  
58 study of cetaceans (e.g., Marques et al., 2012; Klinck et al., 2012; Harris et al.,  
59 2018), they have also been used to monitor taxa such as birds (e.g., Dawson and  
60 Efford, 2009; Buxton et al., 2013; Dent and Molles, 2016; Sebastián-González  
61 et al., 2018) and anurans (e.g., Measey et al., 2017; Benevides et al., 2019).

62 Estimating population density is a primary objective of many acoustic sur-  
63 veys. Marques et al. (2013) provide a review of statistical methods that have  
64 been developed to meet this goal. A major challenge is that not all vocalisa-  
65 tions are detected. Those produced close to recorders are easier to detect than  
66 those further away, and missed vocalisations must be accounted for in order to  
67 estimate density. A common way to overcome uncertain detection is to model  
68 the detection process with a detection function (e.g., Dawson and Efford, 2009;  
69 Marques et al., 2009, 2012; Stevenson et al., 2015; Sebastián-González et al.,  
70 2018), which describes how the probability of detection decreases with an in-  
71 creasing distance between the vocalisation and a detector.

72 Distance sampling (Buckland et al., 2001) can be used to estimate the detec-  
73 tion function, and requires observed distances between detectors and detected  
74 vocalisations. However, the locations at which detected vocalisations were pro-  
75 duced are not often observed on an acoustic survey, and so neither are the  
76 distances. Although properties of acoustic signals such as the received strength  
77 can be used to estimate distances, it is not often possible to do so with high

78 precision for all calls (e.g., see Sebastián-González et al., 2018). Obtaining dis-  
79 tances with negligible error is a fundamental assumption of distance sampling,  
80 violation of which causes biased density estimators (Marques, 2004; Borchers  
81 et al., 2010).

82 Spatial capture-recapture (SCR) models (Efford, 2004; Borchers and Efford,  
83 2008; Royle and Young, 2008; Royle et al., 2013) provide a way to estimate  
84 a detection function directly from a single survey without observing exact dis-  
85 tances between individuals and detectors. Although the first SCR methods were  
86 developed for live-trapping and area-search surveys, there has been recent de-  
87 velopment for acoustic surveys (Efford et al., 2009; Dawson and Efford, 2009;  
88 Borchers et al., 2015; Stevenson et al., 2015). Instead of requiring call locations  
89 to be observed, acoustic SCR models treat them as latent variables. Estimation  
90 of the detection function involves either integrating (for estimation by maximum  
91 likelihood) or sampling (for estimation via MCMC under a Bayesian framework)  
92 over all points at which each call may have been located. Borchers et al. (2015)  
93 developed a method to incorporate auxiliary data like times of arrival, received  
94 signal strengths, and estimated bearings into acoustic SCR models, providing  
95 additional information about call locations. Importantly, their method models  
96 measurement error in these variables, so like other SCR approaches, it does not  
97 require that the locations of the calls are perfectly resolved. So far, acoustic SCR  
98 models have been applied to populations of birds (Efford et al., 2009), whales  
99 (Marques et al., 2012), primates (Kidney et al., 2016), and frogs (Measey et al.,  
100 2017).

101 There are two primary disadvantages of the SCR methods developed by  
102 Efford et al. (2009) and Borchers et al. (2015). The first is that they estimate  
103 call density (calls produced per unit area per unit time), which confounds the  
104 two parameters that are usually of interest: animal density (individuals per

105 unit area) and call rate (calls produced per individual per unit time). Using  
106 ideas from Marques et al. (2013), Stevenson et al. (2015) showed that unbiased  
107 estimation of animal density can be calculated via the quotient of the call density  
108 estimate from SCR and an independently estimated call rate. The latter can be  
109 established by monitoring a random sample of individuals from the population  
110 separately to—but alongside—the main acoustic survey. Collecting these data  
111 can be time consuming and costly, but avoiding this exercise by applying a call  
112 rate estimated from data collected at another location or time can introduce  
113 bias. For example, call rates of ectotherms may vary with temperature (Llusia  
114 and Márquez, 2013) and between different populations of the same species (Zuk  
115 et al., 2001).

116 The second primary disadvantage is that their likelihoods are constructed  
117 under an assumption that call locations are a realisation of a Poisson point pro-  
118 cess, effectively assuming the location of one call is independent of all others.  
119 This assumption provides tractable estimates that can be computed in realistic  
120 time frames—but does not accurately describe the way call locations are distrib-  
121 uted throughout the survey area. Two calls produced by the same individual  
122 will have the same location if animals do not move, or similar locations if they  
123 do. Stevenson et al. (2015) showed that point estimates of call density (and an-  
124 imal density, if computed via an independently estimated call rate) are unbiased  
125 despite this model misspecification—but standard errors are negatively biased  
126 and confidence intervals (CIs) fail to meet their nominal coverage levels. They  
127 used a parametric bootstrap method to provide appropriate estimates of para-  
128 meter uncertainty. Because their point estimates still relied on maximising a  
129 misspecified likelihood, their method cannot make use of likelihood-based tools,  
130 for example AIC and BIC for model selection. The remaining acoustic SCR  
131 method, that of Dawson and Efford (2009), avoids both of the disadvantages

132 described above by only analysing a single call from each detected individual.  
133 However, this approach discards data from all subsequent calls, which are in-  
134 formative about the detection function.

135 Finally, in some cases, animal distribution is of interest in addition to density,  
136 which can be estimated by SCR using an inhomogeneous Poisson process for  
137 the latent locations (Borchers and Efford, 2008; Royle et al., 2013, pp. 307–327).  
138 However, inhomogeneous processes cannot be implemented to estimate animal  
139 distribution in existing acoustic SCR methods, because they directly model call  
140 locations rather than animal locations. Applying an inhomogeneous Poisson  
141 process to call locations provides an estimated density surface that confounds  
142 spatial clustering of animals with clustering of different calls from the same  
143 animal. In other words, such an approach would not be able to distinguish  
144 between a region with high call density due to the presence of many animals,  
145 or a region with high call density due to the presence of a single animal that is  
146 particularly vocally active.

147 In this paper we develop a new SCR method to analyse detection data from  
148 acoustic surveys. Our method directly estimates a detection function, the call  
149 rate, and animal density without requiring the collection of any independent  
150 data and without discarding available data. Unlike existing methods, we directly  
151 model animal locations rather than call locations, so we can fit inhomogeneous  
152 animal density surfaces if desired. We present our method in Sections 2.1–  
153 2.3. Unlike the model of Stevenson et al. (2015), parameter estimation for  
154 our method involves maximisation of a correctly specified likelihood, so we can  
155 rely on the usual variance estimators from maximum-likelihood theory without  
156 resorting to a bootstrap procedure. We can also use other likelihood-based  
157 tools such as AIC and BIC. Like that of Dawson and Efford (2009), our method  
158 requires that individuals can be identified by their calls and that animals do not

159 move substantially during the survey.

160 We are able to generate plausible detection-to-call and call-to-individual  
161 matchings for an application of our method to an acoustic survey of the Cape  
162 peninsula moss frog *Arthroleptella lightfooti*, described in Section 2.4, and present  
163 a simulation study investigating point and variance estimator performance in  
164 Section 2.5. Our method paves the way for the development of extensions that  
165 are applicable to mobile species and/or those that cannot be identified from  
166 their calls, which we discuss in Section 4.

## 167 2 Materials and methods

### 168 2.1 DATA AND NOTATION

169 Notation used in this manuscript is summarised in Table 1. We consider a survey  
170 of duration  $z$  with  $m$  acoustic recorders at known locations, which detect calls  
171 produced by animals. Let  $\mathbf{x}_k$  be the Cartesian coordinates for the location of  
172 the  $k$ th detector. We denote our survey region  $\mathcal{S} \in \mathbb{R}^2$ , the set of all locations  
173 at which a calling individual could possibly be detected by the detector array.  
174 Like existing acoustic SCR models (Efford et al., 2009; Dawson and Efford, 2009;  
175 Borchers et al., 2015; Stevenson et al., 2015), we require that detections can be  
176 matched to calls, so we can determine when a single call is detected by more  
177 than one detector. Like Dawson and Efford (2009), we impose the additional  
178 requirement that calls can be matched to individuals, so we can determine when  
179 a single animal produces multiple detected calls.

180 SCR models require observed capture histories, each of which denotes a  
181 spatial pattern of detections (Borchers, 2012). In our case, each call has a  
182 capture history representing which microphones it was detected by. Let  $\omega_{ij} =$   
183  $(\omega_{ij1}, \dots, \omega_{ijm})$  be the capture history for the  $j$ th call produced by the  $i$ th

**Table 1** A summary of notation used throughout this manuscript.

Group	Notation	Definition
Survey design	$m$	The number of detectors.
	$\mathbf{x}_k$	The location of the $k$ th detector.
	$\mathcal{S}$	The set of all points in the survey region.
	$z$	The duration of the survey.
Latent data	$n$	The number of calling animals in $\mathcal{S}$ .
	$c_i$	The number of calls produced by the $i$ th animal during the survey.
	$\omega_{ijk}$	A detection indicator of the $j$ th call by the $i$ th animal at the $k$ th detector.
	$\boldsymbol{\omega}_{ij}$	The capture history of the $j$ th call by the $i$ th animal, where $\boldsymbol{\omega}_{ij} = (\omega_{ij1}, \dots, \omega_{ijm})$ .
	$\boldsymbol{\Omega}_i$	The capture history matrix of the $i$ th animal, where $\boldsymbol{\Omega}_i = (\boldsymbol{\omega}_{i1}, \dots, \boldsymbol{\omega}_{ic_i})$ .
	$\mathbf{T}_i$	Signal arrival times for the $i$ th animal, comprising an arrival time $t_{ijk}$ for all $\{i, j, k : \omega_{ijk} = 1\}$ .
Observed data	$n^*$	The number of animals with at least one detected call.
	$c_i^*$	The number of calls produced by the $i$ th animal that were detected by at least one detector.
	$\omega_{ijk}^*$	A detection indicator for the $j$ th detected call by the $i$ th detected animal at the $k$ th detector.
	$\boldsymbol{\omega}_{ij}^*$	The capture history of the $j$ th detected call by the $i$ th detected animal, where $\boldsymbol{\omega}_{ij}^* = (\omega_{ij1}^*, \dots, \omega_{ijm}^*)$ .
	$\boldsymbol{\Omega}_i^*$	The capture history matrix of the $i$ th detected animal, where $\boldsymbol{\Omega}_i^* = (\boldsymbol{\omega}_{i1}^*, \dots, \boldsymbol{\omega}_{ic_i^*}^*)$ .
	$\mathbf{T}_i^*$	Signal arrival times for the $i$ th detected animal, comprising an arrival time $t_{ijk}^*$ for all $\{i, j, k : \omega_{ijk}^* = 1\}$ .
Parameters	$\boldsymbol{\beta}$	A vector of coefficients characterising spatially varying calling animal density, where $\boldsymbol{\beta} = (\beta_0, \dots, \beta_Q)$ .
	$\boldsymbol{\psi}$	A vector of call detection function parameters, where $\boldsymbol{\psi} = (\lambda_0, \sigma)$ for the hazard halfnormal.
	$\mu$	The expected number of calls an animal produces per unit time.
	$\sigma_t$	The standard deviation of measurement error for recorded signal arrival times.
Functions	$D(\mathbf{s})$	Calling animal density at location $\mathbf{s}$ . Abbreviated to $D$ for homogeneous density models.
	$x_q(\mathbf{s})$	Measurement of the $q$ th spatial covariate or spline basis function at location $\mathbf{s}$ .
	$d(\mathbf{x}, \mathbf{y})$	The Euclidean distance between locations $\mathbf{x}$ and $\mathbf{y}$ .
	$g_c(d)$	The probability a call is detected at a detector at distance $d$ .
	$g(d)$	The probability at least one call from an animal is detected by a detector at distance $d$ .
	$p_c(\mathbf{s})$	The probability a call at $\mathbf{s}$ is detected by at least one detector.
	$p(\mathbf{s})$	The probability at least one call from an animal at $\mathbf{s}$ is detected by at least one detector.

184 individual, where  $\omega_{ijk} = 1$  if the call was detected by the  $k$ th detector and  
 185  $\omega_{ijk} = 0$  if not. The  $i$ th animal therefore has a capture history matrix,  $\mathbf{\Omega}_i$ ,  
 186 the rows of which are the capture histories for its individual calls. The number  
 187 of rows in  $\mathbf{\Omega}_i$  is equal to  $c_i$ , the number of calls the animal produced during  
 188 the survey, and the number of columns is equal to  $m$ . Some calls may not be  
 189 detected and have capture histories of  $\mathbf{0}_m$ , containing only zeroes. Likewise,  
 190 some animals may not produce any calls that are detected, and therefore have  
 191 capture history matrices containing only zeroes.

192 We may also collect auxiliary data from each detection informative about call  
 193 location, incorporation of which can greatly improve density estimator precision  
 194 (see Borchers et al., 2015). Here we consider the collection of precise times that  
 195 acoustic signals from detected calls arrived at the detectors, where  $t_{ijk}$  provides  
 196 the time that the  $j$ th call from the  $i$ th individual arrived at the  $k$ th microphone.  
 197 Let  $\mathbf{t}_{ij}$  be a vector and  $\mathbf{T}_i$  be a matrix of arrival times associated with detections  
 198 in  $\omega_{ij}$  and  $\mathbf{\Omega}_i$ , respectively.

199 On our survey, we obtain a truncated version of the data described above,  
 200 because capture histories of calls that are not detected and capture history  
 201 matrices of animals that have no detected calls are unobservable. We denote  
 202 the truncated versions of these variables with an asterisk. We do not observe  
 203 the total number of calls produced by the  $i$ th individual,  $c_i$ , instead we observe  
 204 the total number of its calls that were detected by at least one detector,  $c_i^*$ . We  
 205 do not observe all  $c_i$  rows of its capture history matrix  $\mathbf{\Omega}_i$ , we only observe the  
 206  $c_i^*$  rows associated with its detected calls, together given by  $\mathbf{\Omega}_i^*$ . Every row of  
 207  $\mathbf{\Omega}_i^*$  is not  $\mathbf{0}_m$  and has at least one detection. We denote  $\omega_{ij}^*$  to be the capture  
 208 history of  $j$ th detected call produced by the  $i$ th animal, whereas the potentially  
 209 unobserved  $\omega_{ij}$  is its  $j$ th call, regardless of detection. We also only observe  $t_{ijk}^*$ ,  
 210  $\mathbf{t}_{ij}^*$ , and  $\mathbf{T}_i^*$ , comprising arrival times of calls that were detected. Finally, we do

211 not observe  $n$ , the total number of animals in  $\mathcal{S}$ , we only observe the number of  
 212 detected animals,  $n^*$ .

213 In the development of our likelihood below, we use  $f$  to denote a probab-  
 214 ility density function (PDF) or probability mass function (PMF), but do not  
 215 distinguish between functions for different variables, relying on their arguments  
 216 to resolve ambiguity.

## 217 2.2 THE MODEL

218 In the following sections we develop a likelihood for acoustic surveys based on  
 219 the observed data described above:  $n^*$ ,  $\mathbf{c}^* = (c_1^*, \dots, c_{n^*}^*)$ ,  $\boldsymbol{\Omega}^* = (\boldsymbol{\Omega}_1^*, \dots, \boldsymbol{\Omega}_{n^*}^*)$ ,  
 220 and  $\mathbf{T}^* = (\mathbf{T}_1^*, \dots, \mathbf{T}_{n^*}^*)$ .

### 221 2.2.1 Animal density

222 Let  $\mathbf{s}_i$  be the physical location of the  $i$ th animal during the survey. Following  
 223 Borchers and Efford (2008), we assume animals' locations are a realisation of  
 224 an inhomogeneous Poisson process over the survey region, with intensity (i.e.,  
 225 animal density)  $D(\mathbf{s})$  at location  $\mathbf{s} \in \mathcal{S}$ . Spatially varying animal density can  
 226 be modelled via a log-linear relationship

$$\log\{D(\mathbf{s})\} = \beta_0 + \sum_{q=1}^Q \beta_q x_q(\mathbf{s}), \quad (1)$$

227 where  $x_q(\mathbf{s})$  is the measurement of the  $q$ th spatial covariate measured at location  
 228  $\mathbf{s}$ , and  $\boldsymbol{\beta} = (\beta_0, \dots, \beta_Q)$  are parameters characterising the intensity function.  
 229 Alternatively,  $x_q(\mathbf{s})$  might be a basis function of a regression spline, for example  
 230 allowing a spatially smoothed density surface that varies independently of any  
 231 covariates (Borchers and Kidney, 2014). This Poisson point process model im-  
 232 plies that  $n \sim \text{Poisson}\{\int_{\mathcal{S}} D(\mathbf{s}) d\mathbf{s}\}$  (see Borchers et al., 2015), and that animals'  
 233 locations are independent of one another given the intensity surface.

234 *2.2.2 Call and individual detection probabilities*

235 The closer an individual is to a recorder, the more easily its calls are detected.  
 236 We model this detection process using a call detection function,  $g_c(d)$ , which  
 237 provides the probability that a call is detected by a recorder that is located dis-  
 238 tance  $d$  from where the call was produced. There are various parameteric forms  
 239 for detection functions available in the distance sampling and SCR literature.  
 240 Here we consider the hazard halfnormal function

$$g_c(d) = 1 - \exp \left\{ -\lambda_0 \exp \left( \frac{-d^2}{2\sigma^2} \right) \right\}, \quad (2)$$

241 because it allows perfect detection within some distance  $b$  such that  $g_c(d) \approx 1$   
 242 for  $0 < d < b$ , which is often the case for acoustic data.

243 Let  $d(\mathbf{x}, \mathbf{y})$  be the Euclidean distance between the locations  $\mathbf{x}$  and  $\mathbf{y}$ , and  
 244 so  $d(\mathbf{x}_k, \mathbf{s}_i)$  is the distance between the  $k$ th detector and the  $i$ th animal. By  
 245 assuming independence between detections of the same call at different detectors  
 246 conditional on the animal's location, the probability that a single call located  
 247 at  $\mathbf{s}$  is detected at all (i.e., by at least one detector) is the complement of the  
 248 probability of no detection by any detector,  $p_c(\mathbf{s}) = 1 - \prod_{k=1}^m [1 - g_c\{d(\mathbf{x}_k, \mathbf{s})\}]$ .

249 The number of calls produced by the  $i$ th individual has the PMF  $f(c_i)$ .  
 250 Conditional on both its location,  $\mathbf{s}_i$ , and the total number of calls it produced,  
 251  $c_i$ , the number of calls from the  $i$ th animal that are detected by at least one  
 252 detector has a binomial distribution with  $c_i$  trials and probability of success  
 253  $p_c(\mathbf{s})$ . The PMF  $f(c_i^* | c_i, \mathbf{s}_i^*)$ , for the number of detected calls from the  $i$ th  
 254 detected animal, conditional on having produced  $c_i$  total calls from location  
 255  $\mathbf{s}_i$ , is that of the zero-truncated (ZT) binomial distribution. The truncation is  
 256 because detected individuals must have produced at least one detected call.

257 Because we do not observe  $c_i$ , we require the PMF for the number of detected

258 calls from a detected animal conditional only on  $\mathbf{s}_i^*$ , which is given by

$$f(c_i^* | \mathbf{s}_i^*) = \sum_{c_i=1}^{\infty} f(c_i^* | c_i, \mathbf{s}_i^*) f(c_i). \quad (3)$$

259 A special case arises if we assume that calls are produced by a Poisson pro-  
 260 cess at the same rate for all calling individuals, which results in a closed-form  
 261 expression for the sum. In this scenario we have  $c_i \sim \text{Poisson}(\mu z)$ , where the call  
 262 rate  $\mu$  is the expected number of calls produced per unit time per individual.  
 263 This provides  $(c_i^* | \mathbf{s}_i^*) \sim \text{ZT Poisson}\{\mu z p_c(\mathbf{s}_i^*)\}$ , because the call detection pro-  
 264 cess is a thinning of the call production process by the call detection probability,  
 265 and a thinned Poisson process is itself a Poisson process. We proceed using this  
 266 Poisson assumption, but our method can incorporate alternative distributions  
 267 for  $c_i$  if desired, for example to accommodate between-individual variation in  
 268 call rates.

269 The probability that an animal located at  $\mathbf{s}$  is detected at all (i.e., at least  
 270 one of its calls is detected by at least one of the detectors) is then given by the  
 271 complement of the probability that none of its calls were detected,

$$p(\mathbf{s}) = 1 - f(c^* = 0 | \mathbf{s}). \quad (4)$$

272 Although not explicitly used in our likelihood below, the individual detection  
 273 function,  $g(d)$ , is the probability that at least one call produced by an individual  
 274 animal is detected by a particular detector at distance  $d$  from its location, and  
 275 is given by

$$g(d) = \sum_{c=0}^{\infty} [1 - \{1 - g_c(d)\}^c] f(c). \quad (5)$$

276 The area effectively sampled by the detectors is given by  $\int_S p(\mathbf{s}) d\mathbf{s}$  (Borchers,  
 277 2012).

278 *2.2.3 PMFs for capture histories conditional on location*

279 From Section 2.2.2, we assume that  $(\omega_{ijk} \mid \mathbf{s}_i) \sim \text{Bernoulli}[g_c\{d(\mathbf{x}_k, \mathbf{s}_i)\}]$  and  
 280 that a detection of a single call at one detector is independent of detection  
 281 at others, conditional on the call's location—these are standard acoustic SCR  
 282 assumptions. The PMF of the capture history for the  $j$ th call produced by the  
 283  $i$ th animal conditional on its location is therefore given by the following product  
 284 of Bernoulli PMFs:

$$f(\boldsymbol{\omega}_{ij} \mid \mathbf{s}_i) = \prod_{k=1}^m g_c\{d(\mathbf{x}_k, \mathbf{s}_i)\}^{\omega_{ijk}} [1 - g_c\{d(\mathbf{x}_k, \mathbf{s}_i)\}]^{1-\omega_{ijk}}. \quad (6)$$

285 However, observed capture histories of calls are zero-truncated (i.e., it is  
 286 impossible to observe  $\boldsymbol{\omega}_{ij} = \mathbf{0}_m$ ), because we only observe those relating to  
 287 detected calls. We can account for this truncation using the call detection  
 288 probabilities from Section 2.2.2, and so the PMF of the capture history of the  
 289  $j$ th detected call from the  $i$ th detected animal is

$$f(\boldsymbol{\omega}_{ij}^* \mid \mathbf{s}_i^*) = \frac{f(\boldsymbol{\omega}_{ij} = \boldsymbol{\omega}_{ij}^* \mid \mathbf{s}_i = \mathbf{s}_i^*)}{p_c(\mathbf{s}_i^*)}. \quad (7)$$

290 The PMF of a detected animals' capture history matrix, given its location and  
 291 the number of its calls that were detected, is

$$f(\boldsymbol{\Omega}_i^* \mid c_i^*, \mathbf{s}_i^*) = \prod_{j=1}^{c_i^*} f(\boldsymbol{\omega}_{ij}^* \mid \mathbf{s}_i^*). \quad (8)$$

292 *2.2.4 Individual locations*

293 In Section 2.2.3 we assumed that individuals' locations are a realisation of  
 294 an inhomogeneous Poisson process with intensity  $D(\mathbf{s})$  at location  $\mathbf{s}$ , which  
 295 implies that the PDF of a randomly selected individual's location is  $f(\mathbf{s}) =$   
 296  $D(\mathbf{s}) / \int_{\mathcal{S}} D(\mathbf{s}') d\mathbf{s}'$ . However, the location of a randomly selected detected indi-

vidual does not come from this distribution, because inhomogeneity in detected  
 animals' locations is additionally affected by spatially varying detectability: animals  
 situated closer to the detectors are more likely to have at least one of their  
 calls detected.

We proceed using the standard maximum-likelihood SCR approach (Borchers and Efford, 2008), but with an individual's detection probability specified as per our new formulation in Equation (4). Locations of detected individuals are a thinned Poisson process, with thinning via the function  $p(\mathbf{s})$ . The intensity function of the thinned process at location  $\mathbf{s}$  is  $D(\mathbf{s})p(\mathbf{s})$ , which is proportional to the PDF of a detected individual's location, providing the PDF of the  $i$ th detected individuals location:

$$\begin{aligned} f(\mathbf{s}_i^*) &\propto D(\mathbf{s}_i^*)p(\mathbf{s}_i^*) \\ &= \frac{D(\mathbf{s}_i^*)p(\mathbf{s}_i^*)}{\int_{\mathcal{S}} D(\mathbf{s})p(\mathbf{s})d\mathbf{s}}. \end{aligned} \quad (9)$$

Because locations of detected individuals are a realisation of a thinned Poisson process, we have  $n^* \sim \text{Poisson}\{\int_{\mathcal{S}} D(\mathbf{s})p(\mathbf{s})d\mathbf{s}\}$  with PDF

$$f(n^*) = \frac{\{\int_{\mathcal{S}} D(\mathbf{s})p(\mathbf{s})d\mathbf{s}\}^{n^*} \exp\{-\int_{\mathcal{S}} D(\mathbf{s})p(\mathbf{s})d\mathbf{s}\}}{n^*!}. \quad (10)$$

### 2.2.5 Incorporating auxiliary information

We use the same approach of Borchers et al. (2015) and Stevenson et al. (2015) to incorporate observed TOAs. Their method estimates independent measurement error for each detection, with standard deviation  $\sigma_t$ . This procedure acknowledges that TOAs are not observed perfectly and do not deterministically resolve call locations—even when a single call is detected by many detectors.

In Appendix S1, we describe their formulation of  $f(\mathbf{T}_i^* | \boldsymbol{\Omega}_i^*, \mathbf{c}_i^*, \mathbf{s}_i^*)$ , the PDF of the observed TOAs, conditional on the observed capture histories, detected

311 call counts, and detected animal locations.

### 312 2.2.6 The likelihood

Our likelihood function is a joint density of our observed data as a function of the model parameters,  $\theta$ , comprising the following:  $\beta$ , coefficients characterising animal density;  $\psi$ , a vector of call detection function parameters with  $\psi = (\lambda_0, \sigma)$  for the hazard halfnormal detection function;  $\mu$ , the call rate; and  $\sigma_t$ , measurement error for observed TOAs, if they have been collected. The likelihood is

$$L(\theta) = f(n^*, \mathbf{c}^*, \mathbf{\Omega}^*, \mathbf{T}^*) \quad (11)$$

Assuming independence between individuals and building in dependence on call locations as latent variables provides

$$= f(n^*) \prod_{i=1}^{n^*} \int_{\mathcal{S}} f(c_i^*, \mathbf{\Omega}_i^*, \mathbf{T}_i^*, \mathbf{s}_i^*) d\mathbf{s}_i^* \quad (12)$$

$$= f(n^*) \prod_{i=1}^{n^*} \int_{\mathcal{S}} f(\mathbf{T}_i^* | \mathbf{\Omega}_i^*, c_i^*, \mathbf{s}_i^*) f(\mathbf{\Omega}_i^* | c_i^*, \mathbf{s}_i^*) f(c_i^* | \mathbf{s}_i^*) f(\mathbf{s}_i^*) d\mathbf{s}_i^*. \quad (13)$$

313 Here,  $f(n^*)$  is provided in Equation (10), and the four functions in the product  
 314 of the integrand are provided in Appendix S1 and Equations (8), (3), and (9),  
 315 respectively.

316 Estimation can be achieved in the standard way under a maximum likeli-  
 317 hood framework. Maximising  $\log\{L(\theta)\}$  over  $\theta$  provides point estimates, while  
 318 standard errors are available by taking the square-root of the diagonal elements  
 319 of the inverse of the Hessian. The standard errors can be used to construct  
 320 Wald CIs, although likelihood-ratio CIs can also be computed.

321 2.3 SOFTWARE IMPLEMENTATION

322 A software implementation of the method described above is available on GitHub  
323 and archived on Zenodo (Stevenson et al., 2020), and is described in Appendix  
324 S2. The log-likelihood function is written in C++, which is maximised using  
325 the `nllminb()` function from the `stats` package in R (R Core Team, 2020).

326 2.4 THE *A. lightfooti* SURVEY

327 2.4.1 Survey description

328 We applied our method to data collected on two acoustic surveys of *A. lightfooti*,  
329 conducted 18 days apart on 16 May and 3 June, 2012, on Steenberg Plateau in  
330 Silvermine Nature Reserve, Table Mountain National Park, South Africa. On  
331 both occasions, six microphones were placed in identical configurations in the  
332 same seepage inhabited by calling *A. lightfooti* males. These are two of many  
333 such surveys described by Measey et al. (2017), who analysed the resulting data  
334 using the method of Stevenson et al. (2015) and used a proxy call rate to convert  
335 from call to animal density. Here we use our method to estimate both animal  
336 density and the call rate directly.

337 2.4.2 Call and individual identification

338 The process we used to match detections to calls (i.e., to determine which detec-  
339 tions by different microphones are recordings of the same call) is described by  
340 Measey et al. (2017) and Stevenson et al. (2015). We then manually allocated a  
341 plausible matching of detected calls to individual identities using the timings of  
342 calls, plausible call locations, and received call frequencies, as described in Ap-  
343 pendix S3. We recognise that our matching procedure was a subjective exercise.  
344 We discuss this further in Section 4.1.

345 2.4.3 Data analysis

346 We analysed a 30-second subset of each recording, although the full surveys were  
347 longer. If animals are stationary, then increasing survey length does not collect  
348 additional independent data because the array continues to repeatedly detect  
349 the same individuals. Lengthening the survey beyond some limit provides a  
350 negligible improvement in terms of animal density estimator precision; the small  
351 improvement may be outweighed by data processing and computational costs.  
352 Stevenson et al. (2015) found that this limit was approached at about 25 s for  
353 their survey of *A. lightfooti*.

354 Our goal here is to illustrate the application of our method, which is able to  
355 estimate calling animal density and call rate in addition to call density, and is  
356 hereafter referred to as the ASCR-AD model. We also compare these results to  
357 those obtained using the method of Stevenson et al. (2015), which only directly  
358 estimates call density, and is hereafter referred to as the ASCR-CD model. We  
359 fitted the ASCR-CD model using the R package `ascr` (Stevenson, 2020). We  
360 fitted ASCR-AD and ASCR-CR models to the data, considering the two surveys  
361 as independent sessions.

362 Calling animal density, the call detection function, and the TOA measure-  
363 ment error were assumed to be the same for both surveys. Because the mi-  
364 crophone array spanned a small area (the largest distance between a pair of  
365 microphones was 9.22 m), it is reasonable to assume spatially homogeneous  
366 calling animal density across our survey region. Below we refer directly to the  
367 spatially constant calling animal density,  $D$ , rather than coefficients  $\beta$ . We also  
368 refer to call density,  $D_c = D\mu$ , which represents the density of calls produced  
369 per hectare per minute.

370 2.5 SIMULATION STUDY

371 We conducted a simulation study to establish the performance of the ASCR-  
372 AD animal density estimator used in Section 2.4, and make comparisons to the  
373 ASCR-CD method. In particular, our simulation study answers the following  
374 questions in the context of our *A. lightfooti* analysis: (1) does the ASCR-AD  
375 model provide an unbiased point estimator of calling animal density, and CIs  
376 with nominal coverage? (2) How do the properties of the ASCR-AD call density  
377 estimator compare to those of the ASCR-CD call density estimator?

378 We simulated 1000 data sets under the ASCR-AD model described in Sec-  
379 tions 2.2, using the detector configuration and parameter estimates from our  
380 analysis described above. We fitted both the ASCR-AD and ASCR-CD model  
381 to each data set.

382 **3 Results**

383 3.1 THE *A. lightfooti* SURVEY

384 In total, 86 calls from 11 individuals were detected on 16 May 2012, and 98  
385 calls from 14 individuals by the same array on 3 June 2012. Our ASCR-AD  
386 model estimated calling animal density and call rate separately with good pre-  
387 cision, providing estimates (with 95% CIs) of  $\hat{D} = 358.5$  (240.4, 534.6) calling  
388 individuals per hectare and  $\hat{\mu} = 18.1$  (15.5, 21.2) calls per individual per minute.  
389 Coefficients of variation (CVs) for  $\hat{D}$  and  $\hat{\mu}$  were 20.4 and 8.0%, respectively.  
390 We estimate that the area effectively sampled by each of the two surveys was  
391 350 m<sup>2</sup>. See Table 2 for estimates, standard errors, CVs and 95% CIs for all  
392 parameters provided by our ASCR-AD model, and Figure 1 for the estimated  
393 call and individual detection functions, based on estimates  $\hat{\lambda}_0$ ,  $\hat{\sigma}$  and  $\hat{\mu}$ .

394 These results are consistent with those from the ASCR-CD model fitted to

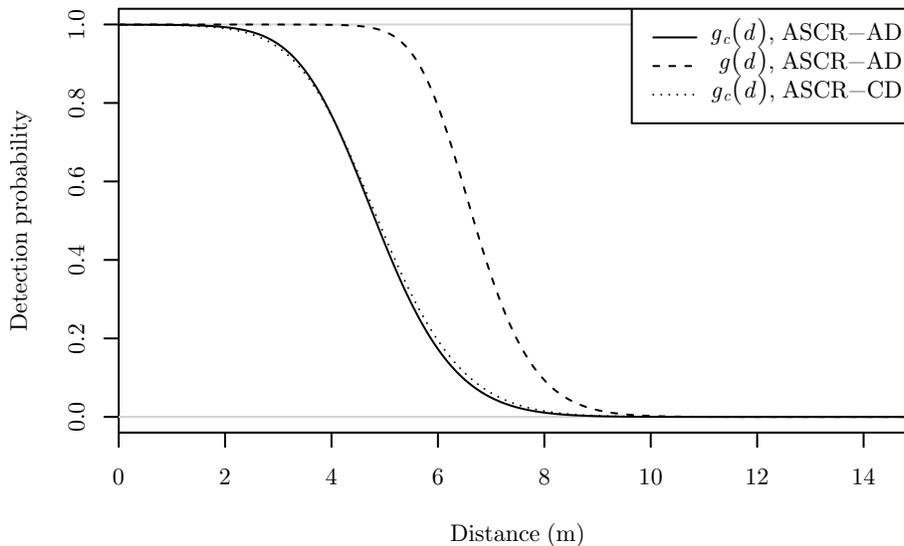
**Table 2** Parameter estimates, standard errors, CVs and CIs from the ASCR-AD model. The parameter  $D$  is reported in calling animals  $\text{ha}^{-1}$ ,  $\sigma$  in m,  $\mu$  in calls per minute, and  $\sigma_t$  in s ( $\times 10^{-4}$ ).

Parameter	Estimate	Std Error	CV (%)	95% CI
$D$	358.5	73.1	20.4	(240.4, 534.6)
$\lambda_0$	7.5	1.1	14.2	(5.7, 9.9)
$\sigma$	2.2	0.1	3.6	(2.1, 2.4)
$\mu$	18.1	1.4	8.0	(15.5, 21.2)
$\sigma_t$	10.4	0.5	4.8	(9.5, 11.4)

395 the same data, which provided a call density estimate of  $\hat{D}_c = 7470$  calls per  
 396 hectare per minute. A derived estimate of call density from our ASCR-AD  
 397 model is  $\hat{D}\hat{\mu} = 6504$  calls per hectare per minute. The two models provided  
 398 very similar estimated call detection functions (Figure 1). The ASCR-CD model  
 399 cannot compute standard errors or CIs without an independently estimated call  
 400 rate, so we are unable to directly compare standard errors, CVs, or CIs between  
 401 the two models.

### 402 3.2 SIMULATION STUDY

403 Based on our simulations, we estimate percentage bias of the call density es-  
 404 timators for our survey scenario to be  $-0.5$  and  $0.6\%$  for the ASCR-AD and  
 405 ASCR-CD models, respectively, with CVs of  $21.4$  and  $23.5\%$ . Our ASCR-AD  
 406 calling animal density estimator has an estimated bias of  $-0.6\%$ , with a CV of  
 407  $20.3\%$ . See Figure 2 for violin plots of call and animal density estimates from  
 408 the two models. In total,  $95.9\%$  of the  $95\%$  CIs for animal density from the  
 409 ASCR-AD model captured the the true parameter value.

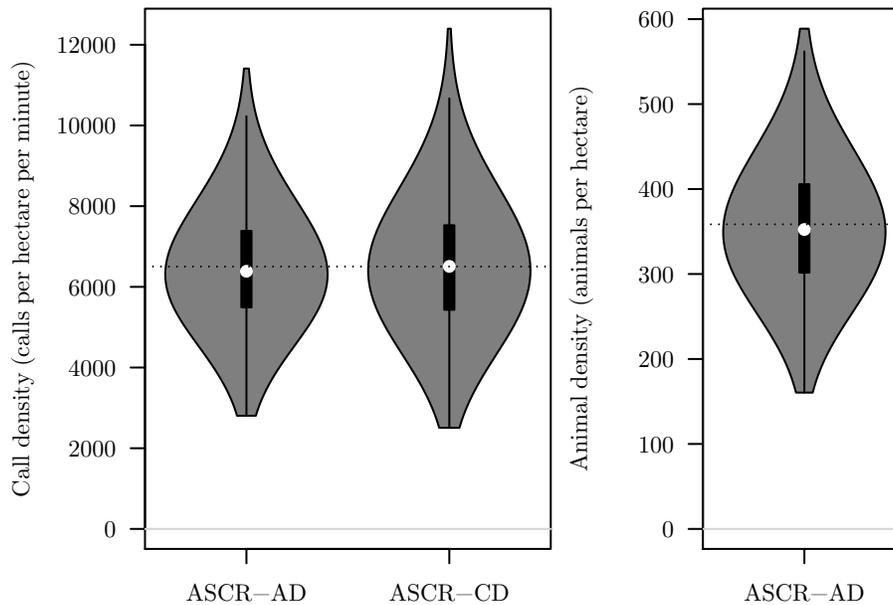


**Fig. 1** Call and individual detection functions ( $g_c(d)$  and  $g(d)$ , respectively) estimated from the acoustic survey data by the ASCR-AD model, and the call detection function estimated by the ASCR-CD model. The ASCR-CD model does not estimate an individual detection function.

#### 410 **4 Discussion**

411 We have described a novel method to estimate calling animal density from acoustic  
 412 tic surveys. Importantly, we disentangle call density and animal abundance,  
 413 allowing separate estimates of animal density and call rates from the acoustic  
 414 survey alone—which is not possible using acoustic SCR models in the existing  
 415 literature.

416 Parameter estimates from our application are consistent with previous stud-  
 417 ies of *A. lightfooti*. Based on seven surveys conducted at the same site, Measey  
 418 et al. (2017) estimated similar call densities to us, and call detection functions  
 419 that also dropped to zero by a distance of 10 m (see our estimate in Figure  
 420 1). Stevenson et al. (2015) physically located and monitored calls from 8 indi-  
 421 viduals, which produced 16.2 calls individual per minute, on average, consistent



**Fig. 2** Violin plots of call density (left) and animal density (right) estimates from our simulation study. Horizontal dotted lines indicate the true density values used to simulate the data.

422 with our estimate (with 95% CI) of 18.1 (15.5, 21.2).

423 Our simulation study reveals that our ASCR-AD model provides an animal  
 424 density estimator with negligible bias and an associated CI with near-nominal  
 425 coverage. For comparison, if we were to treat the objective function maximised  
 426 by the ASCR-CD model as the correct likelihood function, then only 55.0% of  
 427 the 95% CIs for call density capture the true parameter value. This replicates the  
 428 finding of Stevenson et al. (2015) that likelihood-based CIs from the ASCR-CD  
 429 model do not reach their nominal coverage levels due to the likelihood failing to  
 430 acknowledge dependence between capture histories of calls produced by the same  
 431 individual. Moreover, our simulations demonstrate that our model produces  
 432 more precise estimates of call density than the ASCR-CD model, presumably  
 433 because we make use of observed animal identities, and because our model is  
 434 fitted by maximising a correctly specified likelihood.

435 In the remainder of this section, we discuss extensions to our method that  
436 are now possible based on likelihood described in Section 2, and the application  
437 of our method more generally to survey scenarios other than what we have  
438 considered here.

#### 439 4.1 INDIVIDUAL IDENTIFICATION

440 Although we could allocate a plausible set of individual identities to detected  
441 calls from *A. lightfooti* (Section 2.4), we may not have allocated identities per-  
442 fectly. Nevertheless, we anticipate accurate identification from acoustic signals  
443 will become more common in the future with the development of methods to  
444 distinguish animal identities from properties of vocalisations, similarly to cur-  
445 rent efforts in identifying individual animals from photographs. Our application  
446 illustrates the utility of our model when applied to such data sets.

447 Developing models that account for uncertain animal identities is a research  
448 topic of particular recent focus (e.g., see Augustine et al., 2018, 2019; Borchers  
449 et al., in press; Chandler and Royle, 2013; Stevenson et al., 2019), but, at  
450 present, existing methods in the literature are not directly applicable to the type  
451 of acoustic survey we have considered here. Sampling from possible matchings  
452 of detections to individuals is one way to deal with uncertain animal identit-  
453 ies in capture-recapture models (Augustine et al., 2018, 2019), while another is  
454 maximising a likelihood computed via a sum over all such allocations (Borchers  
455 et al., in press). Both introduce considerable computational complexity to model  
456 fitting. Either approach involves a likelihood conditional on a proposed match-  
457 ing of detections to individuals, and this likelihood is what we have provided in  
458 Section 2.2.6. Future methods developed specifically for passive acoustic data  
459 could use properties of detected calls, such as their dominant frequencies, to  
460 inform plausible matchings to individuals.

461 The ASCR-CD model of Stevenson et al. (2015) does not require individual

462 identification and performed comparably to our ASCR-AD model in estimating  
463 call density, despite using a misspecified likelihood. Although requiring inde-  
464 pendently collected call rate data to estimate animal density and produce stand-  
465 ard errors and CIs, this approach remains a computationally efficient alternative  
466 to modelling uncertain identities via sampling or summing over possible alloc-  
467 ations if call rates can be obtained. However, collecting call rate data from  
468 *A. lightfooti* individuals presents a substantial fieldwork burden due to their  
469 cryptic nature. In this scenario, methods that estimate animal density directly  
470 can alleviate this burden, either by dealing directly with observed identities (as  
471 ours does here) or by accounting for identification uncertainty.

#### 472 4.2 INCORPORATING ANIMAL MOVEMENT

473 We assumed that animals are stationary, and so all calls produced by the same  
474 individual have the same location. While our method may be seen as an iter-  
475 ation in the progression towards the estimation of calling animal density in all  
476 situations, it is already applicable to a wide range of uses. In many inverteb-  
477 rates, for example stridulation in Orthoptera (crickets, katydids, grasshoppers)  
478 and cicadas, males call from stationary positions and females travel towards the  
479 calling male (Greenfield, 1997; Fonseca, 2014). Calling vertebrate ectotherms,  
480 like reptiles and amphibians, have similar advertising systems. Volent calling en-  
481 dotherms, birds and bats, often call from roosting congregations where they can  
482 be recorded vocalising in relatively stationary positions (e.g., see Furmankiewicz  
483 et al., 2011).

484 Nevertheless, a possible extension is to allow for mobile animals, so that not  
485 all  $c_i$  calls produced by the  $i$ th animal have location  $s_i$ . Instead, each call could  
486 have a different location, where the  $j$ th call from the  $i$ th animal has location  
487  $s_{ij}$ . Calls from the same animal still cannot be considered independent, because  
488 they will have similar locations if they are produced close together in time.

489 A movement model would be required to specify  $f(\mathcal{S}_i)$ , where  $\mathcal{S}_i$  is a matrix  
490 comprising all  $c_i$  locations of the  $i$ th animal's calls. Movement models for un-  
491 observed activity centres have been incorporated into SCR before (e.g., Ergon  
492 and Gardner, 2014; Royle et al., 2016; Glennie et al., 2019), but only for surveys  
493 where activity centres move between discrete sampling occasions. Accounting  
494 for movement within a single acoustic survey is more complicated because a  
495 continuous-time model is required instead. Locations of an individual are never  
496 directly observed; two possible methods to deal with the latent movement of  
497 each individual are (1) to integrate over possible animal movement trajec-  
498 tories to compute a likelihood to maximise, or (2) sample from these trajectories  
499 within an MCMC scheme under a Bayesian framework. Both provide a sub-  
500 stantial computational challenge.

#### 501 4.3 ACCOUNTING FOR BETWEEN-INDIVIDUAL AND BETWEEN-CALL HET- 502 EROGENEITY

503 Unmodelled heterogeneity in detection probabilities introduces bias to density  
504 estimates provided by capture-recapture models (Otis et al., 1978). Acoustic  
505 SCR methods explicitly model one specific mechanism that induces heterogen-  
506 eity: the call detection function accounts for the fact that calls produced closer  
507 to detectors are more likely to be detected than those produced further away.  
508 However, additional sources of heterogeneity may exist that cause calls to have  
509 different detection functions. For example, some individuals may consistently  
510 call louder than others making their calls more easily detected, and two calls  
511 from the same animal may not be equally detectable due to fluctuations in  
512 source signal strength or background noise.

513 Methods that do not acknowledge which calls were produced by which animal  
514 (e.g., Efford et al., 2009; Borchers et al., 2015; Stevenson et al., 2015) cannot  
515 separate heterogeneity between individuals from heterogeneity between calls.

516 However, our model could be extended by including hierarchical individual-  
517 level and call-level random effects in the detection function to account for both  
518 between-individual and between-call sources of heterogeneity.

#### 519 4.4 DISTRIBUTION OF CALL FREQUENCIES

520 In our application and simulation study, we assumed a Poisson distribution for  
521 the number of calls produced by each individual, consistent with calling animals  
522 producing calls according to a Poisson process at the same underlying rate. We  
523 did so to provide simpler expressions for  $f(c_i^* | \mathbf{s}_i^*)$  and  $g(d)$ , given our models  
524 were fitted primarily for illustrative purposes.

525 Nevertheless, if desired, our method can accommodate either underdisper-  
526 sion or overdispersion relative to the Poisson distribution by using an appropri-  
527 ate alternative for  $f(c_i)$ . This decision comes at the expense of added compu-  
528 tational complexity due to the infinite sum in Equation (3).

#### 529 4.5 FURTHER APPLICATIONS

530 We applied our model to data collected on two passive acoustic surveys of *A.*  
531 *lightfooti* conducted in the same seepage. Other studies have established that  
532 SCR models like ours can be used more generally for other taxa, in more complex  
533 habitats, and at larger spatiotemporal scales.

534 Kidney et al. (2016) and Marques et al. (2012) have shown that acoustic  
535 SCR models can be applied to species that are detectable at much larger spatial  
536 scales, although larger spacings between adjacent detectors are required; one  
537 recommendation is a spacing of two times  $\sigma$  (Efford and Fewster, 2013). A  
538 practical consequence of a large spacing is an inability to connect all detect-  
539 ors to a recorder with a single clock, precluding the collection of TOAs that  
540 are comparable between detectors. Our method can be applied without TOAs

541 simply by omitting the PDF  $f(\mathbf{T}_i^* | \boldsymbol{\Omega}_i^*, c_i^*, \mathbf{s}_i^*)$  from the likelihood (Equation  
542 (13)), although this affects estimator precision (Borchers et al., 2015).

543       Following Borchers and Efford (2008), our method accommodates estimation  
544 of inhomogeneous density for studies in more complex habitats. Moreover, SCR  
545 models that involve a homogeneous density assumption provide reliable estim-  
546 ates of abundance and average density even for complex habitats, for example  
547 with small patches of suitable habitat interspersed amongst large regions with  
548 no individuals (Efford and Fewster, 2013).

549       Cluster survey designs, which involve deploying small clusters of detectors  
550 at various locations across the landscape, are potentially useful for studies mon-  
551 itoring a large, complex region. Dawson and Efford (2009) used a cluster design  
552 for their acoustic survey of ovenbirds *Seiurus aurocapilla*. Clark (2019) con-  
553 ducted a comprehensive simulation study investigating the performance of SCR  
554 density estimators applied to data from cluster designs.

#### 555       4.6 CONCLUDING REMARKS

556 Although our method requires individual identification of calling animals and  
557 does not allow for animals relocating within the survey region, it provides a  
558 correctly specified likelihood under its assumptions. By doing so, we obtain  
559 point and variance estimates using standard maximum-likelihood theory, and  
560 can use other likelihood-based tools like information-theoretic model selection  
561 criteria. Moreover, our model can fit an inhomogeneous density surface, and  
562 does not require supplementary data on call rate at the expense of fieldwork  
563 effort to identify call density from animal density.

564       We have provided clear direction towards new methods that relax our as-  
565 sumptions and data requirements. The potential models we described through-  
566 out Sections 4.1–4.3 introduce new latent components over and above the animal  
567 locations we dealt with in constructing our likelihood, and remain topics of fu-

568 ture research.

#### 569 ACKNOWLEDGEMENTS

570 This work was funded by an EPSRC/NECR PhD grant (No. EP/1000917/1),  
571 by the EPSRC through a Doctoral Fellowship Prize, and by the Royal Society  
572 of New Zealand through Marsden Grant 19-UOA-211. JM thanks the DSI-NRF  
573 Centre of Excellence for Invasion Biology.

574 Data collection was made possible due to access facilitated by CapeNature  
575 (Permit No. AAA007-00084-0035) and South African National Parks (Table  
576 Mountain National Park). Funding for the frog survey was received from the  
577 National Geographic Society/Waite Grants Program (No. W184-11). We thank  
578 Tanya Scott, Res Altwegg, and David Borchers for assisting with recordings.

579 We are grateful to the associate editor, an anonymous reviewer, and Eliza  
580 Grames for constructive comments that improved this paper.

#### 581 DATA AVAILABILITY

582 Data collected on the acoustic surveys of *A. lightfooti*, along with code to  
583 analyse these data and conduct the simulation study, are available in a GitHub  
584 repository (<https://github.com/b-steve/scr-cuerate>) and are archived on  
585 Zenodo (Stevenson et al., 2020).

#### 586 AUTHORS' CONTRIBUTIONS

587 BCS conceived the ideas for this study. BCS and CKYY developed the statisti-  
588 cal methodology. BCS, PvDB, and CKYY wrote the software and conducted  
589 the simulation study. JM collected the data. BCS and PvDB analysed the data.  
590 BCS led the writing of the manuscript. All authors contributed to revision of  
591 the original draft.

592 **References**

- 593 Augustine BC, Royle JA, Kelly MJ, Satter CB, Alonso RS, Boydston EE, Crooks KR.  
594 2018. Spatial capture-recapture with partial identity: An application to camera  
595 traps. *The Annals of Applied Statistics* **12**: 67–95.
- 596 Augustine BC, Royle JA, Murphy SM, Chandler RB, Cox JJ, Kelly MJ. 2019. Spa-  
597 tial capture-recapture for categorically marked populations with an application to  
598 genetic capture-recapture. *Ecosphere* **10**: e02627.
- 599 Benevides FL, Mauntz WJ, Jacobsen CJ, Hara AH. 2019. Estimating density of calling  
600 male *Eleutherodactylus coqui* in Hawaii from audio recordings of the nighttime frog  
601 chorus. *Bioacoustics* **28**: 101–114.
- 602 Borchers DL. 2012. A non-technical overview of spatially explicit capture-recapture  
603 models. *Journal of Ornithology* **152**: 435–444.
- 604 Borchers DL, Efford MG. 2008. Spatially explicit maximum likelihood methods for  
605 capture-recapture studies. *Biometrics* **64**: 377–385.
- 606 Borchers DL, Kidney D. 2014. Flexible density surface estimation for spatially explicit  
607 capture-recapture surveys. CREEM Technical Report 2014-1.
- 608 Borchers DL, Marques TA, Gunnlaugsson T, Jupp P. 2010. Estimating distance  
609 sampling detection functions when distances are measured with errors. *Journal*  
610 *of Agricultural, Biological, and Environmental Statistics* **15**: 346–361.
- 611 Borchers DL, Nightingale P, Stevenson BC, Fewster RM. in press. A latent capture  
612 history model for digital aerial surveys. *Biometrics* .
- 613 Borchers DL, Stevenson BC, Kidney D, Thomas L, Marques TA. 2015. A unifying  
614 model for capture-recapture and distance sampling surveys of wildlife populations.  
615 *Journal of the American Statistical Association* **110**: 195–204.
- 616 Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L. 2001.  
617 *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*.  
618 Oxford: Oxford University Press.
- 619 Buxton RT, Major HL, Jones IL, Williams JC. 2013. Examining patterns in nocturnal  
620 seabird activity and recovery across the Western Aleutian Islands, Alaska, using  
621 automated acoustic recording. *The Auk* **130**: 331–341.
- 622 Chandler RB, Royle JA. 2013. Spatially explicit models for inference about density  
623 in unmarked or partially marked populations. *The Annals of Applied Statistics* **7**:  
624 936–954.
- 625 Clark JD. 2019. Comparing clustered sampling designs for spatially explicit estimation  
626 of population density. *Population Ecology* **61**: 93–101.
- 627 Dawson DK, Efford MG. 2009. Bird population density estimated from acoustic sig-  
628 nals. *Journal of Applied Ecology* **46**: 1201–1209.
- 629 Dent JM, Molles LE. 2016. Call-based identification as a potential tool for monitoring  
630 great spotted kiwi. *Emu—Austral Ornithology* **116**: 315–322.

- 631 Efford MG. 2004. Density estimation in live-trapping studies. *Oikos* **106**: 598–610.
- 632 Efford MG, Dawson DK, Borchers DL. 2009. Population density estimated from loc-  
633 ations of individuals on a passive detector array. *Ecology* **90**: 2676–2682.
- 634 Efford MG, Fewster RM. 2013. Estimating population size by spatially explicit  
635 capture-recapture. *Oikos* **122**: 918–928.
- 636 Ergon T, Gardner B. 2014. Separating mortality and emigration: Modelling space use,  
637 dispersal and survival with robust-design spatial capture-recapture data. *Methods*  
638 *in Ecology and Evolution* **5**: 1327–1336.
- 639 Fonseca PJ. 2014. Cicada acoustic communication. In Hedwig B (ed.) *Insect Hearing*  
640 *and Acoustic Communication*. Berlin, Heidelberg: Springer, 101–121.
- 641 Furmankiewicz J, Ruczyński I, Urban R, Jones G. 2011. Social calls provide tree-  
642 dwelling bats with information about the locations of conspecifics at roosts. *Ethology*  
643 **117**: 480–489.
- 644 Glennie R, Borchers DL, Murchie M, Harmsen BJ, Foster RJ. 2019. Open population  
645 maximum likelihood spatial capture-recapture. *Biometrics* **75**: 1345–1355.
- 646 Greenfield MD. 1997. Acoustic communication in orthoptera. In Gangwere SK, Mur-  
647 alirangan MC, Muralirangan M (eds.) *The Bionomics of Grasshoppers, Katydid,*  
648 *and Their Kin*. Wallingford, UK: CAB International, 197–230.
- 649 Harris DV, Miksis-Olds JL, Vernon JA, Thomas L. 2018. Fin whale density and  
650 distribution estimation using acoustic bearings derived from sparse arrays. *Journal*  
651 *of the Acoustical Society of America* **143**: 2980–2993.
- 652 Kidney D, Rawson B, Borchers DL, Stevenson BC, Marques TA, Thomas L. 2016.  
653 An efficient acoustic density estimation method with human detectors applied to  
654 gibbons in Cambodia. *PLoS ONE* **11**: e0155066.
- 655 Klinck H, Mellinger DK, Klinck K, Bogue NM, Luby JC, Jump WA, Shilling GB,  
656 Litchendorf T, Wood AS, Schorr GS, Baird RW. 2012. Near-real-time acoustic  
657 monitoring of beaked whales and other cetaceans using a Seaglider<sup>TM</sup>. *PLoS ONE*  
658 **7**: e361282012.
- 659 Llusia D, Márquez R. 2013. Calling behaviour under climate change: Geographical  
660 and seasonal variation of calling temperatures in ectotherms. *Global Change Biology*  
661 **19**: 2655–2674.
- 662 Marques TA. 2004. Predicting and correcting bias caused by measurement error in  
663 line transect sampling using multiplicative error models. *Biometrics* **60**: 757–763.
- 664 Marques TA, Thomas L, Martin SW, Mellinger DK, Jarvis S, Morrissey RP, Ciminello  
665 C, DiMarzio N. 2012. Spatially explicit capture-recapture methods to estimate  
666 minke whale density from data collected at bottom-mounted hydrophones. *Journal*  
667 *of Ornithology* **152**: 445–455.
- 668 Marques TA, Thomas L, Martin SW, Mellinger DK, Ward JA, Morettie DJ, Harris  
669 D, Tyack PL. 2013. Estimating animal population density using passive acoustics.  
670 *Biological Reviews* **88**: 287–309.

- 671 Marques TA, Thomas L, Ward J, DiMarzio N, Tyack PL. 2009. Estimating cetacean  
672 population density using fixed passive acoustic sensors: An example with Blainville's  
673 beaked whales. *The Journal of the Acoustical Society of America* **125**: 1982–1994.
- 674 Measey GJ, Stevenson BC, Scott T, Altwegg R, Borchers DL. 2017. Counting chirps:  
675 Acoustic monitoring of cryptic frogs. *Journal of Applied Ecology* **54**: 894–902.
- 676 Otis DL, Burnham KP, White GC, Anderson DR. 1978. Statistical inference from  
677 capture data on closed animal populations. *Wildlife Monographs* **62**: 3–135.
- 678 R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R  
679 Foundation for Statistical Computing, Vienna, Austria.  
680 URL <http://www.R-project.org/>
- 681 Royle JA, Chandler RB, Sollmann R, Gardner B. 2013. *Spatial Capture-Recapture*.  
682 Amsterdam: Academic Press.
- 683 Royle JA, Fuller AK, Sutherland C. 2016. Spatial capture-recapture models allowing  
684 Markovian transience or dispersal. *Population Ecology* **58**: 53–62.
- 685 Royle JA, Young KV. 2008. A hierarchical model for spatial capture-recapture data.  
686 *Ecology* **89**: 2281–2289.
- 687 Sebastián-González E, Camp RJ, Tanimoto AM, de Oliveira PM, Lima BB, Marques  
688 TA, Hart PJ. 2018. Density estimation of sound-producing terrestrial animals using  
689 single automatic acoustic recorders and distance sampling. *Avian Conservation and*  
690 *Ecology* **13**: 7.
- 691 Stevenson BC. 2020. *ascr: Acoustic spatial capture-recapture in AD Model Builder*. R  
692 package version 2.2.2.  
693 URL <https://github.com/b-steve/ascr>
- 694 Stevenson BC, Borchers DL, Altwegg R, Swift RJ, Gillespie DM, Measey GJ. 2015. A  
695 general framework for animal density estimation from acoustic detections across a  
696 fixed microphone array. *Methods in Ecology and Evolution* **6**: 38–48.
- 697 Stevenson BC, Borchers DL, Fewster RM. 2019. Cluster capture-recapture to account  
698 for identification uncertainty on aerial surveys of animal populations. *Biometrics*  
699 **75**: 326–336.
- 700 Stevenson BC, van Dam-Bates P, Young CKY, Measey J. 2020. Code and data to  
701 accompany the paper “A spatial capture-recapture model to estimate call rate and  
702 population density from passive acoustic surveys”. *Zenodo* .  
703 URL <https://doi.org/10.5281/zenodo.4083577>
- 704 Zuk M, Rotenberry JT, Simmons LW. 2001. Geographical variation in calling song of  
705 the field cricket *Teleogryllus oceanicus*: The importance of spatial scale. *Journal of*  
706 *Evolutionary Biology* **14**: 731–741.