

1      That's not the Mona Lisa! How to interpret  
2      spatial capture-recapture density surface  
3      estimates

4      David L. Borchers<sup>1,\*</sup>, Ian Durbach<sup>1</sup>, Rishika Chopara<sup>2</sup>, Ben C.  
5      Stevenson<sup>2</sup>, Rachel Phillip<sup>1</sup>, and Koustubh Sharma<sup>3</sup>

6      <sup>1</sup>Centre for Research into Ecological and Environmental Modelling,  
7      School of Mathematics and Statistics, Univeristy of St Andrews,  
8      The Observatory, St Andrews, Fife, KY16 9LZ, Scotland

9      <sup>2</sup>Department of Statistics, University of Auckland, Auckland 1010,  
10      New Zealand

11      <sup>3</sup>Snow Leopard Trust, Seattle, Washington, United States of  
12      America

13      \*Corresponding author: [dlb@st-andrews.ac.uk](mailto:dlb@st-andrews.ac.uk)

14      **S1    Simulation study for Bayesian models**

15      Results presented in Section 4 were generated by fitting maximum-likelihood  
16      SCR models to simulated data. In this appendix we reproduce results from  
17      Section 4 using Bayesian models fitted via MCMC to demonstrate that our  
18      results are not simply a consequence of adopting a classical approach. In Section  
19      S1.1 we describe our Bayesian models, in Section S1.2 we present results of our  
20      simulation study, and in Section S1.3 we discuss similarities and differences

21 between these results and those presented in the main manuscript based on  
 22 maximum-likelihood models.

### 23 **S1.1 Model fitting**

24 We fitted Bayesian versions of the maximum-likelihood models presented in  
 25 Section 4 to each data set. Again, we used models with constant density to  
 26 estimate realised AC and realised usage surfaces, and a model with inhomoge-  
 27 neous density characterised by a log-linear relationship with a spatial covariate  
 28 to estimate expected AC density surfaces.

29 We fitted our models in NIMBLE (insert reference) using data augmentation  
 30 (REFERENCE: Tanner and Wong, 1987), which has become the prevailing  
 31 way to fit SCR models under a Bayesian framework. This approach involves  
 32 sampling a superpopulation of  $M$  activity centres, including those of the  $n$   
 33 animals detected on the SCR survey. We have an indicator variable  $z_i$  for the  
 34  $i$ th animal, denoting whether the  $i$ th animal in the augmented population is  
 35 ‘exists’ in a given MCMC iteration. Rather than directly estimating  $N$ , the  
 36 population size, we estimate the data augmentation parameter,  $\psi$ , the expected  
 37 proportion of the animals in the superpopulation for which the indicator is equal  
 38 to 1. For each MCMC iteration we obtain a sample from the posterior of  $N$  using  
 39  $\sum_{i=1}^M z_i$ . A sample from the posterior for animal density can be obtained by  
 40 dividing by the area of the survey region. Further details on data augmentation  
 41 can be found in Kery AND Schaub, 2012 (REFERENCE).

We used the following uninformative priors for the detection function param-  
 eters, specifying a prior for  $\log\{1/(2\sigma^2)\}$  rather than  $\sigma$  directly, as per (ROYLE  
 TEXTBOOK):

$$\begin{aligned}\lambda_0 &\sim \text{Gamma}(0.001, 0.001) \\ \log\left(\frac{1}{2\sigma^2}\right) &\sim \text{Uniform}(-10, 10)\end{aligned}$$

For the constant density model, the activity centres were given a uniform

prior distribution over the survey region and the data augmentation parameter was given a uniform prior from 0 to 1. For the inhomogeneous density model, animal density at location  $\mathbf{x}$  is given by  $D(\mathbf{x}) = \exp\{\beta_0 + \beta_1 y(\mathbf{x})\}$ , where  $y(\mathbf{x})$  is a measurement of a covariate at location  $\mathbf{x}$ . We used the following uninformative priors for the coefficients  $\beta_0$  and  $\beta_1$ :

$$\beta_0 \sim \text{Uniform}(-10, 10)$$

$$\beta_1 \sim \text{Uniform}(-10, 10)$$

42 When we fit each constant density model, we ran () MCMC iterations, where  
 43 we set M to be equal to (). We also used a thinning value of (), an adaptation  
 44 interval of () and () burn-in iterations.

45 When fitting each inhomogeneous density model, we ran () MCMC itera-  
 46 tions, and used a value of () for M. The thinning value was (), with an adaptation  
 47 interval of () and () burn-in iterations.

## 48 S1.2 Results

## 49 S1.3 Discussion

# 50 S2 Estimation of realised usage density

51 Estimation of realised usage density is a similar process for both maximum  
 52 likelihood and Bayesian approaches: we sum usage densities for each individual  
 53 animal, each of which is calculated by convolving the estimated PDF of its  
 54 activity centre with an individual usage distribution.

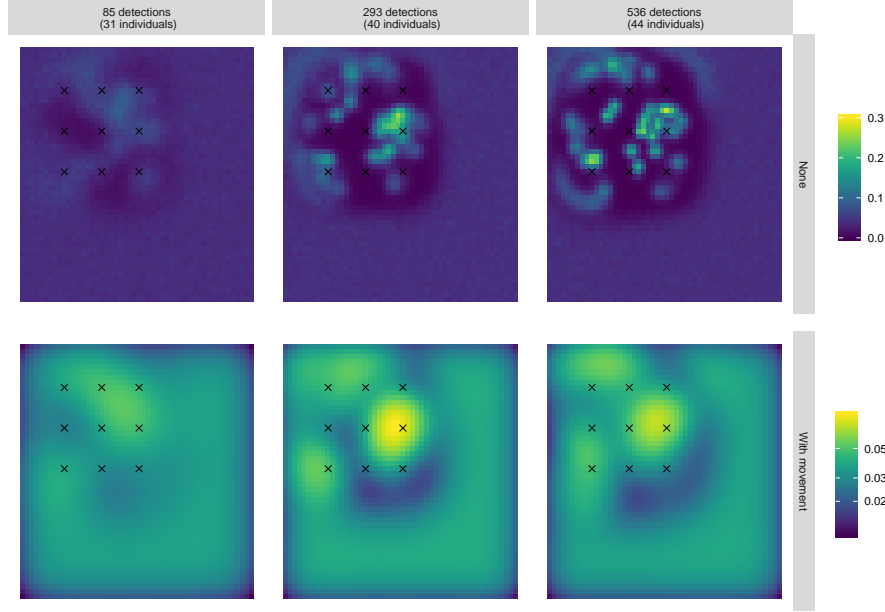


Figure 1: Figure 9

## S2.1 The maximum likelihood approach

For maximum likelihood, the estimated usage density for the  $i$ th animal, with capture history  $\omega_i$ , is given by

$$f_{s|\omega}(s | \omega_i; \hat{\theta}) = \int f_{x|\omega}(x | \omega_i; \hat{\theta}) f_{s|x}(s | x; \hat{\theta}) dx, \quad (1)$$

where

- $\hat{\theta}$  is a vector containing the maximum likelihood estimates of the encounter function parameters;
- $f_{s|\omega}(s | \omega_i; \hat{\theta})$  is the estimated usage distribution, providing the probability density of finding an individual with capture history  $\omega_i$  at location  $s$  at a randomly selected point in time;
- $f_{x|\omega}(x | \omega_i; \hat{\theta})$  is the estimated PDF of the activity centre of an individual with capture history  $\omega_i$  (see Section 3); and
- $f_{s|x}(s | x; \hat{\theta})$  is the estimated usage distribution of the individual con-

ditional on the activity centre, providing the probability density of the individual being at location  $\mathbf{s}$  given that its activity centre is at  $\mathbf{x}$ .

Estimated usage density at location  $\mathbf{s}$  is then given by  $\hat{D}_u(\mathbf{s}) = \sum_i f_{\mathbf{s}|\boldsymbol{\omega}}(\mathbf{s} | \boldsymbol{\omega}_i; \hat{\boldsymbol{\theta}})$ , noting that the sum is over individuals that were not detected, with capture histories  $(0, \dots, 0)$ , along with those that were.

It's unclear to me whether we directly estimate a detection function or an encounter function in our models. Below I assume the reader will know what an encounter function is, but it might need to be explained more explicitly. It's important that we construct the individual usage distribution using an encounter function rather than a detection function, because the rate at which an animal visits a location is proportional to the encounter function, but not to the detection function.

Here we constructed the individual usage distribution under the assumption that the density of an individual being at location  $\mathbf{s}$  given its activity centre is at  $\mathbf{x}$  is proportional to the encounter function  $h\{d(\mathbf{s}, \mathbf{x}); \hat{\boldsymbol{\theta}}\}$ , where  $d(\mathbf{s}, \mathbf{x})$  is the Euclidean distance between  $\mathbf{s}$  and  $\mathbf{x}$ , and so

$$f_{\mathbf{s}|\mathbf{x}}(\mathbf{s} | \mathbf{x}; \hat{\boldsymbol{\theta}}) = \frac{h\{d(\mathbf{s}, \mathbf{x}); \hat{\boldsymbol{\theta}}\}}{\int h\{d(\mathbf{s}', \mathbf{x}); \hat{\boldsymbol{\theta}}\} d\mathbf{s}'}, \quad (2)$$

where the denominator is a normalising constant.

## S2.2 The Bayesian approach

Bayesian models fitted via MCMC can directly sample activity centres of detected individuals, and also of undetected individuals using data augmentation (ROYLE TEXTBOOK REFERENCE), thus obtaining samples from  $f_{\mathbf{x}|\boldsymbol{\omega}}(\mathbf{x} | \boldsymbol{\omega})$  for each individual. We can use these samples directly to obtain the following approximation of the  $i$ th individual's usage distribution:

$$f_{\mathbf{s}|\boldsymbol{\omega}}(\mathbf{s} | \boldsymbol{\omega}_i) \approx \frac{1}{M} \sum_{j=1}^M f_{\mathbf{s}|\mathbf{x}}(\mathbf{s} | \mathbf{x}_{(j)}, \boldsymbol{\theta}_{(j)}), \quad (3)$$

where  $\mathbf{x}_{(j)}$  and  $\boldsymbol{\theta}_{(j)}$  are the activity centre and a vector of encounter function parameters that were sampled on the  $j$ th of  $M$  total MCMC iterations, respectively. The estimated usage distribution is therefore not conditional on one particular set of estimated parameter values, but instead considers the range of values across the posterior distribution of  $\boldsymbol{\theta}$ .

Err, we used  $M$  earlier for superpopulation size. What should this be?

### S2.3 Discussion

I'm not sure that this is the best place for the discussion below, but leaving it here for now.

We constructed individual usage distributions using the encounter function from our SCR model, but this may not always be appropriate. For example, if individuals cannot fully explore their home range within the duration of the survey, then we would not expect the spatial range of the detection function to match the extent of an animal's usage distribution.

Even for longer surveys, it may not be sensible to relate the range of the encounter function to the size of the region used by an individual even for longer surveys, so care should be taken when this practice is used. For example, Tenan, Pedrini, Bragalanti, Groff & Sutherland (2017) found that the spatial scale of the encounter rate function for brown bears (*Ursus arctos*) estimated using SCR was not consistent with spatial usage parameters estimated from other data sources, although Popescu, de Valpine & Sweitzer (2014) did not detect any such inconsistency for a population of fishers (*Pekania pennanti*). If alternative data sources are available (e.g., telemetry, or opportunistic data such as hair or scat samples) they may be incorporated for improved estimation of individual usage distributions (Tenan *et al.*, 2017).

Our method also assumes that home ranges are circular, however their shapes are likely to be modified by variables relating to population and landscape connectivity (see Drake, Lambin & Sutherland, in press, for a review).

## 110 **References**

- 111 Drake, J., Lambin, X. & Sutherland, C. (in press) The value of considering  
112 demographic contributions to connectivity: a review. *Ecography*.
- 113 Popescu, V.D., de Valpine, P. & Sweitzer, R.A. (2014) Testing the consistency  
114 of wildlife data types before combining them: the case of camera traps and  
115 telemetry. *Ecology and Evolution*, **4**, 933–943.
- 116 Tenan, S., Pedrini, P., Bragalanti, N., Groff, C. & Sutherland, C. (2017) Data  
117 integration for inference about spatial processes: A model-based approach to  
118 test and account for data inconsistency. *PlosOne*, **12**, e0185588.