That's not the Mona Lisa!: how to interpret spatial capture-recapture density surface estimates

David L. Borchers[1,*], Ian Durbach[1], Ben C. Stevenson[2], Rishika Chopara[2], Rachel Phillip[1], and Koustubh Sharma[3]

[1]Centre for Research into Ecological and Environmental Modelling, School of Mathematics and Statistics, Univeristy of St Andrews, The Observatory, St Andrews, Fife, KY16 9LZ, Scotland

[2]Department of Statistics, University of Auckland, Auckland 1010, New Zealand

[3]Snow Leopard Trust, Seattle, Washington, United States of America

[*]Corresponding author: dlb@st-andrews.ac.uk

1 Summary

1. Non-uniform denisty surfaces obtained from spatial capture-recapture (SCR) analyses are often misinterpreted and this leads to incorrect inferences about the populations under study. Spatial variation in the surface of interest is often counfused with spatial variation in the amount of information in the sample about the surface of interest. There is also often a lack of clarity about what the surface of interest really is.

2. We focus on three distinct kinds of surface: (1) the estimated activity centre density surface, (2) the estimated activity centre location surface, and (3) the estimated usage surface. The first of these estimates the intensity of the point process generating activity centres, the second estimates the realised activity centre locations, the third estimates the expected space usage. For easy visual interpretation, we use the greyscale image of the Mona Lisa as the true activity centre density surface and illustrate correct and incorrect inferences from simulated SCR surveys with this density. We also illustrate with a real SCR survey of tigers in the Nagarahole game reserve.

3. We show that treating the estimated activity centre location surface as a species distribution map or an estimate of the actiivty centre density surface results in invalid and misleading ecological inferences. This surface is highly dependent on where the detectors are placed and very different surfaces can be obtained by surveying exactly the same animals with detectors placed at different locations. A correct way to obtain a species distribuiton map or an estimate of the actiivty centre density surface is to estimate the intensity of a point process model for activity centres, which may depend on spatially-referenced covariates. Usage sufaces are obtained similarly, but include expected movement about activity centres.

4. To avoid misinterpretation, practitioners should state explicitly the kind of density surface they are estimating and should be careful to draw inferences appropriate to that kind of surface. In particular, estimated activity centre *location* surfaces should not be interpreted as if they were estimated activity centre *density* surfaces.

Keywords: Spatial capture-recapture, density surface

2 Introduction

Spatial capture-recapture (SCR) models (Efford, 2004; Borchers & Efford, 2008; Royle & Young, 2008) are now widely used to estimate animal abundance and distribution from a variety of data types, including that from camera-traps (, for example), hair snares and dung surveys (, for example), live-captures (, for example), acoustic detectors (Dawson & Efford, 2009; Kidney, Rawson, Borchers, Thomas, Marques & Stevenson, 2013; Stevenson, Borchers, Altwegg, Swift, Gillespie & Measey, 2015; Borchers, Stevenson, Kidney, Thomas & Marques, 2015; Loveridge, Kidney, S., S., Eames & Borchers, 2017, for example). These methods use the location of the detectors (traps) and the locations at which animals were detected (their spatial capture histories) to estimate animal density. The methods have two basic components: a spatial model that quantifies animal activity centre density at all points in the survey region, and an encounter model that quantifies the expected detection frequency or detection probability, given the activity centre locations and the detector locations.

SCR density estimates are often presented graphically in the form of estimated density maps, these being easy to absorb and interpret, at least on the face of it. However, there are various kinds of density map that one can produce from SCR analyses and depending on what is presented, it is easy to misinterpret the maps. The most common form of misinterpretation is treating maps that include both spatially varying uncertainty about location and spatially varying activity centre density estimates as if they were maps of activity centre density alone, but there is also a lack of clarity about whether it is activity centre density or space use density that is being presented.

Examples include Dorazio & Karanth (2017) which says that such maps effectively provide "a species distribution model, even in cases where spatial covariates of abundance are unknown or unavailable", Alexander, Gopalaswamy, Shi & Riordan (2015), which presents a map (Figure 4) that include both spatially varying uncertainty about location and spatially varying activity centre density and refers to it as the "spatial distribution of snow leopards", and Elliot & Gopalaswamy (2016), which presents the same kind of map (Figure 2) and refers to it as the "pixel-specific lion density".

Add more refs?

The problems with interpretation of such maps arises because (a) there are various kinds of "density", (b) uncertainty varies spatially and this fact must be (but is often not) taken into account when interpreting estimated density surfaces from SCR surveys, and (c) a failure to distinguish betwen activity centre density and usage density.

We start by describing different kinds of densities involved in SCR surveys, because in any discussion of density surfaces, we need to be clear about what "density" means.

2.1 Different kinds of density

It is useful to distinguish between four kinds of spatial " density", two of them dealing with activity centres, and two dealing with space usage. Conceptually, we have some point process that governs how many activity centres there are in the survey region, and where they are. Animals then use (move through and/or send a detectable signal like sound sound through) the space around their activity centres. Activity centres are governed by the point process alone; usage is governed by both the point process and the movement/propogation process about the points. With this in mind, we refer to four kinds of density as follows:

1. **The expected activity centre density** at a point is the intensity of the underlying point process that models where animals' activity centres are "on average" i.e. over many realizations of the process. It is the expected number of activity centres per unit area within a region, in the limit as the area of the region shrinks to zero. The expected activity centre density surface is the set of these densities at every point. The expected number of activity centres within some region is the volume under this surface over

the region.

2. **The realised activity centre density** is only well defined if continuous space is partitioned into what we will call cells. The realised activity centre density in a cell is the actual (as opposed to expected) number of activity centres per unit area within the cell (i.e., the number divided by the area of the cell) at the time of the survey. The realised activity centres themselves are points in space, not densities.

3. **The expected usage density** in a region is the expected number of visits per unit area of animals to the area, averaged over all possible activity centre locations, over the course of a survey (it is the expected number of visits divided by the area). The expected usage density at a point is the limit of this value as the area shrinks to zero.

4. **The realised usage density** in a region is the expected number of animal visits per unit area to the area over the duration of the survey (the expected number of visits divided by the area), *conditional* on the activity centre locations. The realised usage density at a point is the limit of this value as the area shrinks to zero.

We focus on densities 1, 2, and 4. Figure 1 shows examples of each, except that we show the realised activity centre locations rather than realised activity centre densities in sub-regions of space. Realised activity centre densities can only be plotted when space has been cut into cells; in continuous space the density is zero everywhere except at activity centre locations, where it approaches infinity.

2.2 Estimated density surfaces

If we are interested in explaining why density tends to be high in some places and lower in others, or in characterising the process that governs the distribution of activity centres, then we are primarily interested in estimating a density surface

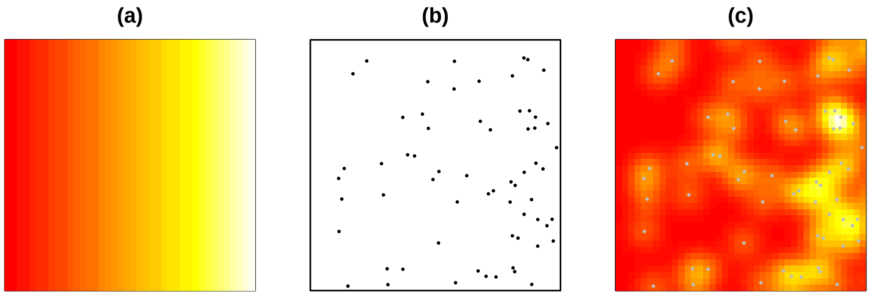

Figure 1: Examples of (a) an expected activity centre density surface, (b) a realisation of activity centres from this density surface, and (c) the associated realised usage density surface (with activity centres shown as grey dots).

like that shown in Figure 1(a). In this example, it is easting that influences this density, but in general it might be any of a wide variety of habitat or environmental covariates, some of which may be unobserved and evidenced only by spatial clustering of activity centres.

If we are interested only in where the activity centres are, and not in explaining why they are there, then Figure 1(b) suffices. But suppose that we observe activity centres with some error. For example, Figure 2 shows the distributions of estimated activity centre locations when the locations are estimated with bivariate normal errors with (a) small standard errors, (b) larger standard errors, and (c) standard errors increasing linearly from the centre of the plot. The estimation uncertainty "spreads" each activity centre according to a bivariate normal distribution, with greater spreading when there is greater unceraінty.

Figure 2(a) gives a reasonable visual representation of where the activity centres are. It is much more difficult to pick out individual activity centres from Figure 2(b), but it gives a reasonable representation of where the high- and low-density regions of activity centres are – much like Figure 1(a), but customised somewhat for this particular realisation of activity centre locations rather than their long-run average locations. Note, however, that these two figures are representations of exactly the same set of activity centres and that if one interprets them as plots of activity centre density, they contradict each other. Figure 2(a) says that almost all the region has low density (red in the

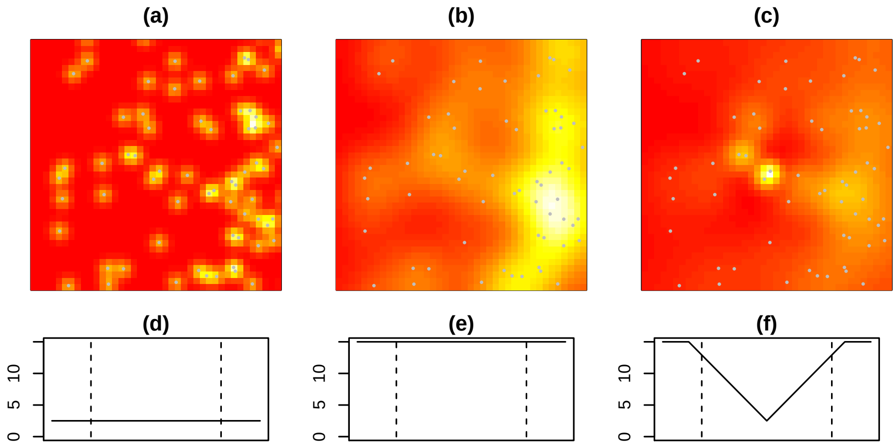

Figure 2: Examples of the density of the activity centres of Figure 1(b), when observed with bivariate normal estimation errors with standard errors (a) $\sigma = 2.5$, (b) $\sigma = 15$, and (c) $\sigma = 2.5$ at the centre of the plot, rising linearly to $\sigma = 2.5$ by the edge of the plot. True activity centres shown as grey dots. The colour scales of panels (a) to (c) are such that the highest and lowest densities in each plot is the same. Panesl (d) to (f) plot the standard errors of the observation errors against the x-axis. Vertical dashed lines show the extent of the survey region in panels (a) to (c); a buffer beyond this is included because spreading of points outside it affect the plot within the survey region.

plot) and that there are lots of small high-density regions, while Figure 2(b) says that there is much less variation in density, that there are large swathes of higher density (the yellow towards the right) and large swathes of low density towards the left. The reason that Figure 2(b) shows less variation in density is not that there is less variation in the population (there are exactly the same activity centres in both (a) and (b)), it is that we are less sure about the location of the activity centres in (b). To interpret this as less variation in activity centre density is to invite incorrect ecological inferences.

Now what about Figure 2(c)? If this is interpreted as indicating where the high and low-density regions are, it is misleading. It says that the highest density region is in the centre of the plot, and that the region with most variation in density is the central region, which is not true.

The fact that there is only small observation error in the centre of the plot and large observation error at the edges means that the activity centres near the centre are not spread much and therefore appear as higher peaks in the surface, with low regions where there are no activity centres. Near the edges of the plot, on the other hand, observation error is high and activity centres are spread a lot, which both flattens the peaks at individual activity centre locations and "fills in" the troughs where the are no activity centres. We see the same effect with the usage density maps (Figure 3), but less pronounced because the usage about the activity centres already "spreads" around points before any observation error occurs.

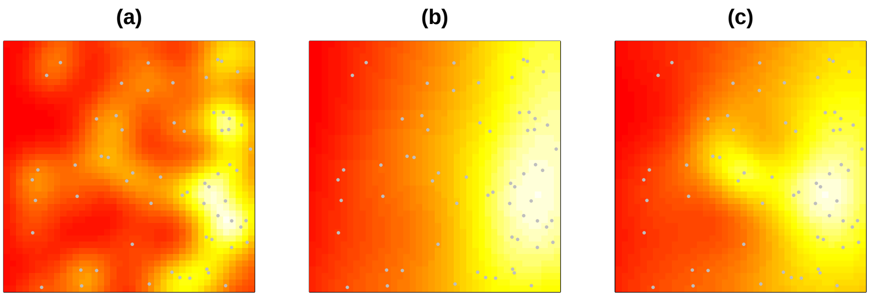

Figure 3: Examples of the usage density of Figure 1(c), when observed with bivariate normal estimation errors with standard errors (a) $\sigma = 2.5$, (b) $\sigma = 15$, and (c) $\sigma = 2.5$ at the centre of the plot, rising linearly to $\sigma = 2.5$ by the edge of the plot. True activity centres shown as grey dots. The colour scales of the three plots are such that the highest and lowest densities in each plot is the same.

It is a feature of SCR surveys that the locations of individuals farther from the detector array tend to be estimated with greater uncertainty than individuals within the array. This is illustrated in Figure 4, which shows the estimated probability density functions for two animals detected on a simulated SCR survey with a $4{\times}4$ array placed in the centre of the population shown in Figure 1(b). (The reason contours top right "avoid" the triangle is because the detection function range, estimated from the whole survey, not just the points shown, is large and if the activity centre was near the triangle, other detectors would have high probability of detecting it. The fact that they did not makes them "repel" the activity centre.)

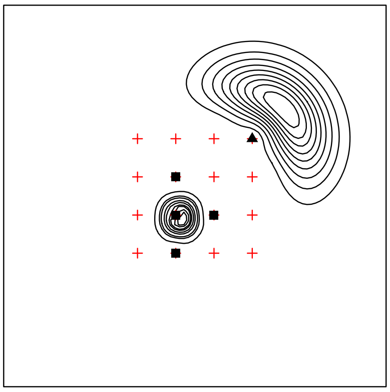

Figure 4: Estimated probability density function contours for two detections of an SCR survey of the population shown in Figure 1(b). Traps are shown as red crosses. The lower left individual was detected at traps indicated by black squares, the upper right individual only by the top right trap indicated by a black triangle.

3 SCR density estimation methods

Maximum likelihood (ML) and Bayesian SCR estimation methods are documented in a good number of papers, starting with Borchers & Efford (2008) and Royle & Young (2008), and we do not repeat the details here. Both ML and Bayesian inference are based on SCR likelihood functions that include a component specifying the activity centre density surface, which may depend on spatially-referenced covariates. (The linear density surface shown in Figure 1(a) is an example.) The density surface is typically of the form $D(\boldsymbol{s}) = \exp\left\{\beta_0 + \sum_{k=1}^{K} \beta_k x_k(\boldsymbol{s})\right\}$, where $\boldsymbol{s}$ is a point in the plane, $x_k(\boldsymbol{s})$ is the kth of K spatially-referenced covariates, evaluated at $\boldsymbol{s}$, β_0 is and intercept parameer, and β_k is the slope parameter for the kth spatially-referenced covariate. ML and Bayesian methods are able to estimate $\beta_0, \ldots, \beta_K$, and hence to estimate the density surface. We refer to this as the "estimated density surface".

Given spatial capture histories, ML and Bayesian methods are also able to estimate the locations of activity centres. (Locations like those shown in Figure 1(b), for example.) While activity centres are points, there is always uncertainty associated with estimating their locations, so that SCR estimates of activity centre locations are probability density functions (PDFs), not points. Estimates of these PDFs are conditional on the spatial capture histories of the individuals concerned – because the capture histories contain information on where each animal's activity centre was (see the capture histories and estimated location densities in Figure 4, for example). Details of how one obtains these estimated activity centre PDFs are contained in Section 4.3 of Borchers & Efford (2008) for ML methods and the section "Estimating derived parameters" on page 3238 of Royle, Karanth, Gopalaswamy & Kumar (2009) for Bayesian methods.

Note that we can obtain activity centre PDFs for undetected animals, because although the animals were unobserved, we know their capture histories – namely no capture at every detector. Note also that all undetected animals will have the same activity centre PDF[1] because they all have the same capture history.

Suppose that we estimate from an SCR survey that there are $\hat{N}$ animals within the survey region. If one adds up the activity centre PDFs for all n detected animals, and the $\hat{N}-n$ activity centre PDFs of the undetected animals, at all points in the survey region, one gets a surface that is in many publications (including those lised in the Introduction) interpreted as a density surface for animal locations. We refer to this as the "estimated activity centre *location* surface".

What we call an "estimated activity centre *location* surface" has been referred to as the estimated distribution, or density of *animals*. However, animals distribute themselves around their activity centres, so that activity centre density and animal density are not the same thing. Suppose for example, that we are certain that there is exactly one activity centre in a region that has surface area 1 (so that activity centre density in this region is 1). Suppose also that

[1]This is not the case if there are individual-level covariates that affect detection probability estimates, but this is a complication that we ignore here in order to present as clear and uncomplicated an exposition of they key points of this paper as we can.

the animal with activity centre in this region ranges wider than this region, and spends exactly half its time in this region. It is not certain that there is an animal in the region at any time, so that animal density will be less than 1. In this example, it would be fair to say that the *animal* density in the region is 0.5. To avoid confusion, we refer to this as the "usage density" rather than "animal density". Details of how one obtains an estimated usage denstiy surface from an estimated activity centre *location* surface are given in Appendix **??**.

Ben/Ian: Can you insert mathematical details? I think we need to do this because there are (to my knowledge) no publications that contain the details. BCS: Sure thing, I'll fill it in here. Basically it's just spreading the point around in the same way you have above for estimation uncertainty. I'm pretty sure this is linked to the idea of what some people call a 'utilisation distribution'; I'd better do a little reading about these, too.

In summary, there are three kinds of estimated surface of interest here:

- The **estimated activity centre *density* surface**: This is an estimate of the density model component of the SCR model, which governs the location of activity centres.
- The **estimated activity centre *location* surface**: This is the combined estimated activity centre densities of all animals, conditional on each animal's capture history.
- The **estimated usage surface**: This is the combined space usage density of animals, conditional on each animal's capture history.

4 Methods

We illustrate what each of the three kinds of estimated surface gives the practitioner, and what interpretations of the surfaces are valid and useful, by (a) simulating data from a density surface that has easy visual interpretation, and (b) using the Ngarahole SCR tiger survey data kindly provided by the first author of **?**.

4.1 Reproducing the Mona Lisa

For easy visual interpretation, we turned one of the most recognisable images in Western culture, the Mona Lisa, into a density surface. We created a greyscale version of a region of the original image (Figure 5, "True Density") in which greyscale values give the true density of activity centres, and lighter areas correspond to higher densities.

We then used the density surface to generate two realisations of points from the underlying process. In the first of these we generated the number of points from a single draw from a Poisson distribution with mean 7,500, resulting in 7,451 activity centres being generated, which we plot in Figure 5, "Realisation 1" as a density at 50×50 pixel resolution. This realisation has the advantage of closely reproducing the source image, and when we conduct SCR surveys with this population, it gives us an indication of the asymptotic behaviour of SCR density estimators, i.e. as sample size gets very large. We also generated a much smaller second realisation of 84 points (Figure 5, "Realisation 2"). This realisation captures the Mona Lisa only at an extremely gross level (the darkest region corresponding to the hair can be picked out if you squint at the image long enough!), but is a useful aid to understanding some properties of the estimators.

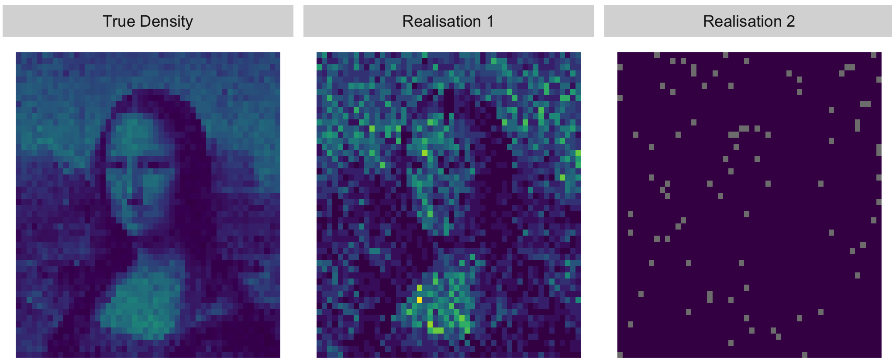

Figure 5: Input data for the Mona Lisa simulation study. A grayscale version of the Mona Lisa ("True Density") is treated as an activity centre density surface, frwom which we generated a sample of 7,451 and 84 activity centres and plotted these as densities at 50×50 pixel resolution ("Realisation 1" and "Realisation 2", respectively).

We simulated SCR surveys of the population, using a variety of detector arrays and also varying the number of detection occasions.

I think occasions are a red herring here. If these are count data, then occasion is irrelevant. What we should report is the sample size (total number of detections and number of unique individuals detected).

Different arrays and detection functions were used for the large and small populations described above. With the large number of activity centers, we used four different 3x4 arrays (Figure 6), with array centers $(19, 21)$, $(19, 33)$, $(28, 21)$, or $(28, 33)$. All arrays were spaced such that they have length 8 in the x-plane and 12 in the y-plane, and so have an average spacing of $4 = 2\sigma$. We simulated capture histories using a half-normal detection function with g_0, the probability of detection at a single detector placed at an activity centre, set to 0.5, and σ, the spatial scale parameter, set to 2. We then simulated a capture history for either one or 20 occasions on each array.

DLB: Or encounter functions?

Same comment about occasions applies - and 20 occasions is very rare, so not great example in this way.

BCS: I had a look at some R code, and I think I saw that the hazard halfnormal detection function was being used by defining λ_0 and σ. Or maybe this was just a default setting? Anyway I'm just leaving this comment here to remind me to have a look at the code later (or Ian can clarify).

When using relatively few activity centers, visual interpretation was made easier by increasing the spatial scale parameter, effectively increasing the distance animals travel from the activity centers, and also by increasing the distance between detectors. For these cases, we increased σ to 4, holding other detection function parameters at their previous values, and used a 3x4 array centered on $(18, 34)$ and with an average spacing of 8 between detectors, double that used previously. We simulated capture histories after one, three, ten, and 20 occasions.

Same comment about ocaions applies.

After simulating these capture histories for these arrays, we created an estimatedactivity centre surface for each simulation. In this scenario we assumed a model with constant density and compared the resulting estimated realised activity center surface to the true population density surface (the Mona Lisa).

The next step was to introduce covariates into our density model. We generated covariates by manipulating the "Low Res" image to obtain further images using simple image processing operations like blurring and shifting. Covariates are thus all functions of the true densities but the strength of the association between the covariate and true density varies substantially. We generated four covariates: a "strong" covariate that uses a Deriche (blur) filter with a small range, effectively smoothing the image locally; a "moderate" covariate that uses the same blur filter but with a larger range, increasing the amount of smoothing; a "weak" covariate that uses the same degree of blurring as the moderate covariate but in addition shifts the image down and to the right, destroying much of the relationship between covariate and density; and a "locally strong" covariate, which uses the strong covariate values in the top right hand part of the image and the weak covariate values in the remainder of the image. These covariates are shown in the first row of plots in Figure 8. For each covariate we estimated a corresponding expected activity center density.

DLB: I thoink we should remove the shifting

We simulated capture histories and created an estimated realised animal density surface i.e. including movement, for each of these simulations and compared them to the true population density surface. All computations were done using the *secr* package in R version 3.4.3.

4.2 Results

4.2.1 Estimated realised activity centre densities with many activity centers

The same patterns hold in two dimensions under the standard wildlife survey assumptions of Poisson-distributed activity centers (with constant intensity) and detectability inversely related to distance from activity center (Figure 6 and 7). A single sampling occasion was sufficient to capture the broad features of the Mona Lisa, but only close to where detectors were located (Figure 6, first row). Away from the detectors the estimated density quickly reverted to close to the estimated mean intensity of the process. Additional sampling occasions resulted in the density of activity centers close to detectors being estimated in much greater detail, but did not affect the surface away from detectors (Figure 7, first row).

Very different relative and absolute densities were obtained depending on where traps are located, even when estimating density *in exactly the same region of the surface and where that region is close to the array* (Figure 6 and 7, second row). With a single occasion, density was always estimated to be highest nearest the corner where the trap is located (Figure 6, second row). This pattern occured because the inset region happened to occur in a region of above average density. If instead it occured in a low density region one would see the opposite pattern – low density in the corner containing a trap, increasing away from the trap. This was clearly visible when a single sampling occasion was used, because the estimated surface reverted quickly to the mean intensity. Additional sampling allowed fine detail in the density surface to be estimated close to traps, with slower reversion to mean intensity, but there was still very clear disagreement between the density surfaces returned by the different arrays (Figure 7, second row).

The densities in these plots should really be expected densities (i.e. averages over lots of simulations) - particularly for the bottom row of estimates, since you might expect substantial random variation in these from sample to sample. The top images illustrate fairly convincingly how the SCR "torch" does pretty well where it "shines" but can't see beyond, without the need to use expected values, but because you can't see pattern in the bottom ones, you are more inclined to say "So what that they are not the same - there is sampling variation." Also, we need to give sample sizes for these surveys (number detections and number of different individuals).

4.2.2 Estimated expected activity center densities with many activity centers

Introducing covariates into the density models allowed us to recover features of the Mona Lisa across the entire image, not just near where detectors were located, although good estimation of activity center locations depended heavily

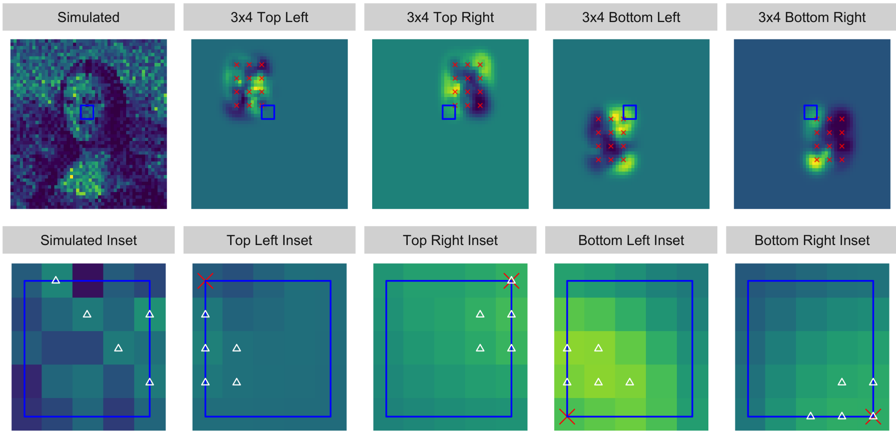

Figure 6: True activity center densities in this realisation ("Simulated") compared with realised activity center surfaces estimated using different arrays after a single sampling occasion. High density areas are indicated in yellow, low density areas in blue. Detectors are shown as red crosses. Blue squares show the same 4x4 square centered at $(25, 29)$, whose vertices are corner detectors from each of 3x4 arrays. Each plot in the second row shows an enlargement of the blue square in the plot above it. White triangles denote the five cells with the highest estimated densities in each of the second row plots.

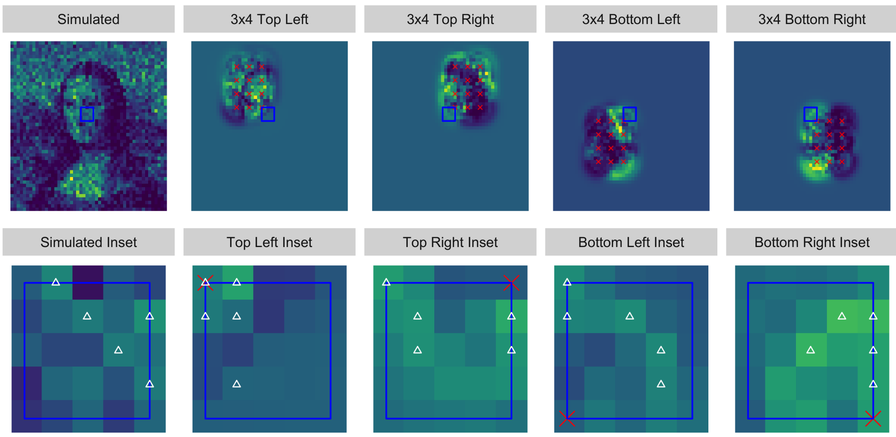

Figure 7: True activity center densities in this realisation ("Simulated") compared with activity center surfaces estimated using different arrays after 20 sampling occasions. High density areas are indicated in yellow, low density areas in blue. See the caption to Figure 6 for further annotation details.

on the availability of good covariates (Figure 8). With our "strong" covariate we recovered all of the broad features of the Mona Lisa, and many of the fine scale features such as eyes, shading of clouds, *etc.* With the "moderate" covariate we recovered broad scale features but no finer details. With a "weak" covariate, the estimated density surface essentially reverted to the mean intensity of the process across the entire region. With a "locally strong" covariate – one that is a good indicator of density in some parts of the study region but poor elsewhere – the dependency on array location was reintroduced. If the array was located where the covariate was strong, the estimated density surface was accurate in that vicinity. If the array was located where the covariate was weak, then the model estimated no relationship between covariate and density and reverted back to the mean intensity everywhere in the region (Figure 8).

4.2.3 Estimated realised activity centre densities with few activity centers

We observed similar patterns under the more "wildlife survey appropriate" condition in which we generated only 85 activity centers across the study region (Figure 9). In this case there is a large difference between the mean intensity surface (the Mona Lisa) and the activity center surface in this realization (85 points), and so it is not surprising that the estimated realised activity centre density surface looks nothing like the Mona Lisa (Figure 9, first row). Nevertheless, a model assuming constant density gives increasingly accurate estimates of the locations of activity centers in the vicinity of detectors as survey effort increases, but very little information is obtained elsewhere, and this does not change with survey effort (Figure 9, first row). This gives the estimated activity center surfaces a characteristic pattern – the surface becomes increasingly peaked or "spiky" around detectors as survey effort increases, but remains flat away from the array.

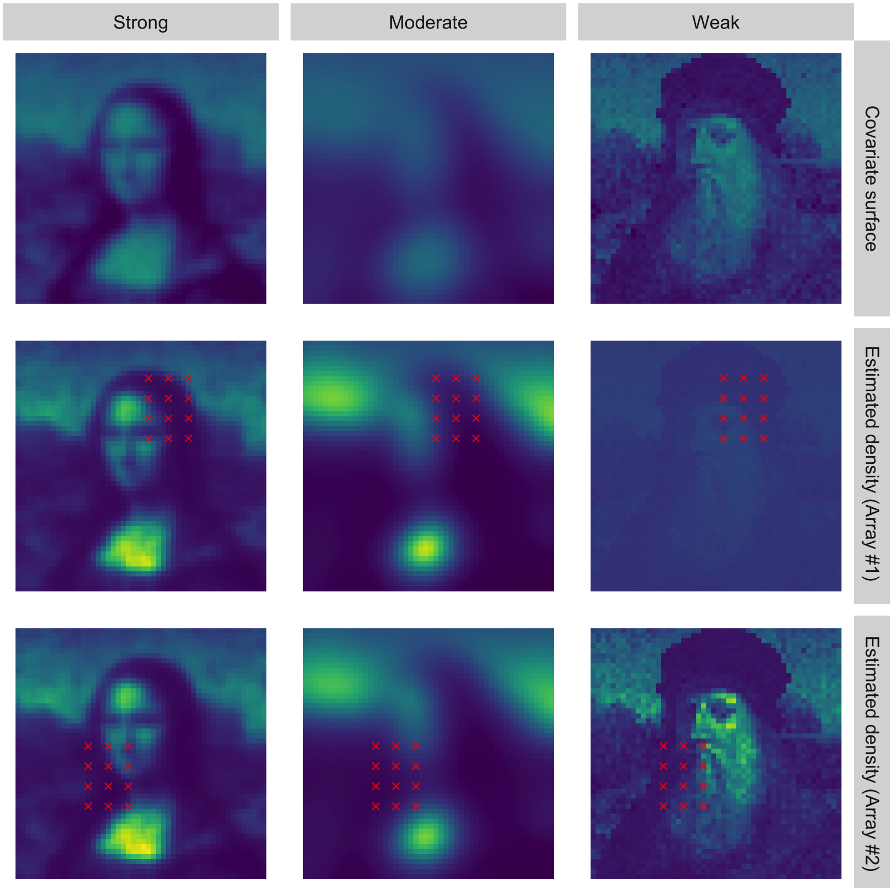

Figure 8: Expected activity center surfaces estimated using a model with density a function on one of four simulated spatially-varying covariates. Covariates are shown in the first row of plots, and were generating by manipulating the true intensity surface (Figure **??**, "Low Res") by blurring and shifting operations (see Section **??** for details). High density areas are indicated in yellow, low density areas in blue. Detectors are shown as red crosses.

4.2.4 Estimated expected animal densities with few activity centers

Any covariate model returns a surface that is some multiple of the covariate surface. Whether this is a good approximation of the true mean intensity surface depends on the strength of the covariate and sample size. With a strong covariate and sufficient sampling occasions we recovered the Mona Lisa, but with only a single occasion the direction of the relationship was incorrectly estimated, so that dark areas were predicted as light and light areas as dark (Figure 9, second row). This error was corrected by additional occasions. The same pattern occured with a moderate covariate, but the effect of the weaker covariate is clear in that we did not recover as good an approximation of the Mona Lisa (Figure 9, third row). Additional sampling would not help with this. Note that in both cases the surface we recovered is an approximation of the mean intensity surface. It does not give a good approximation to the locations of the 85 activity centers in this particular realization.

4.2.5 Estimated realised animal densities when there are few activity centers

Estimated realised animal density surfaces – those that incorporate animal movement – were smoother than the density surface of estimated realised activity centers and also less sensitive to survey effort (Figure 10). The estimated realised animal density surface adds a movement kernel that is insensitive to survey effort to a realised activity center surface that becomes more peaked as survey effort increases, so this is to be expected. Estimated realised animal density surfaces were not "just" smoothed versions of the realised activity center surfaces, however. In our example unobserved animals were estimated to be spending their time on the outskirts of the study region, far away from any detectors (Figure 10, second row), which is quite different from the homogenous surface we obtained away from detectors when looking only at activity centers (Figure 10, first row). In contrast, the realised animal density surface for captured animals *was* essentially a smoothed version of the realised activity center

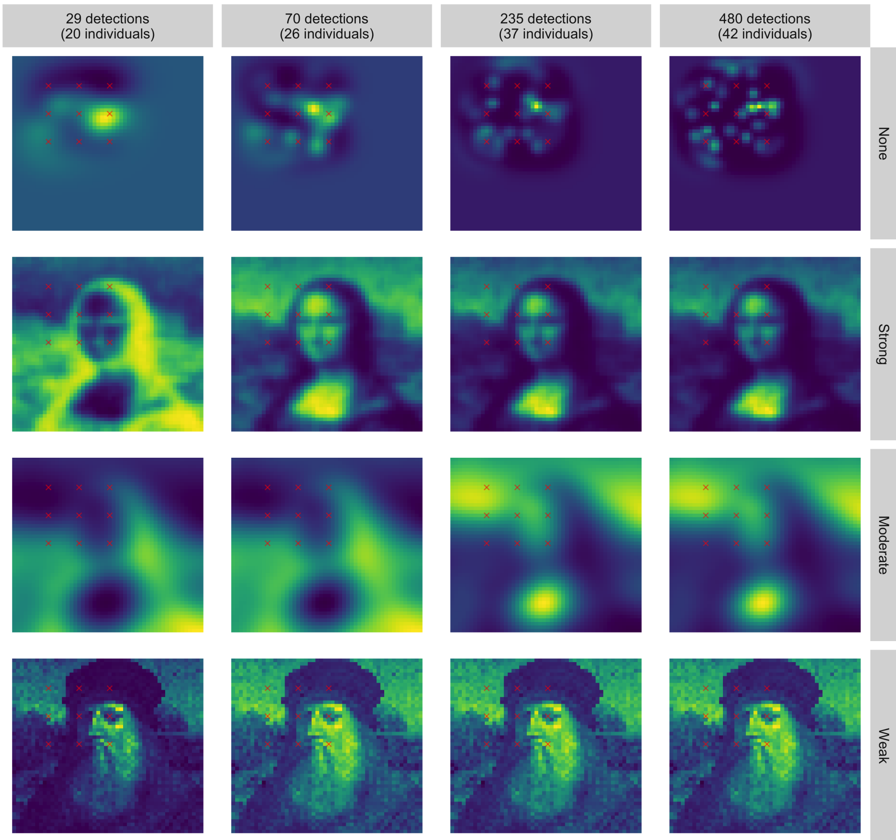

Figure 9: Estimates of realised activity center density surfaces from a constant density model (first row) and expected activity center density surfaces from a model with density depending on “strong” or “moderate” covariates (second and third rows respectively, see Section **??** for details of how covariates were simulated). 85 activity centers were generated across the entire image, drawn from a Poisson process with intensity given by the “Low Res” image in Figure **??**. High density areas are indicated in yellow, low density areas in blue. Detectors are shown as red crosses.

surface around the detector array, and so very similar in terms of the broad patterns it showed (Figure 10, third row).

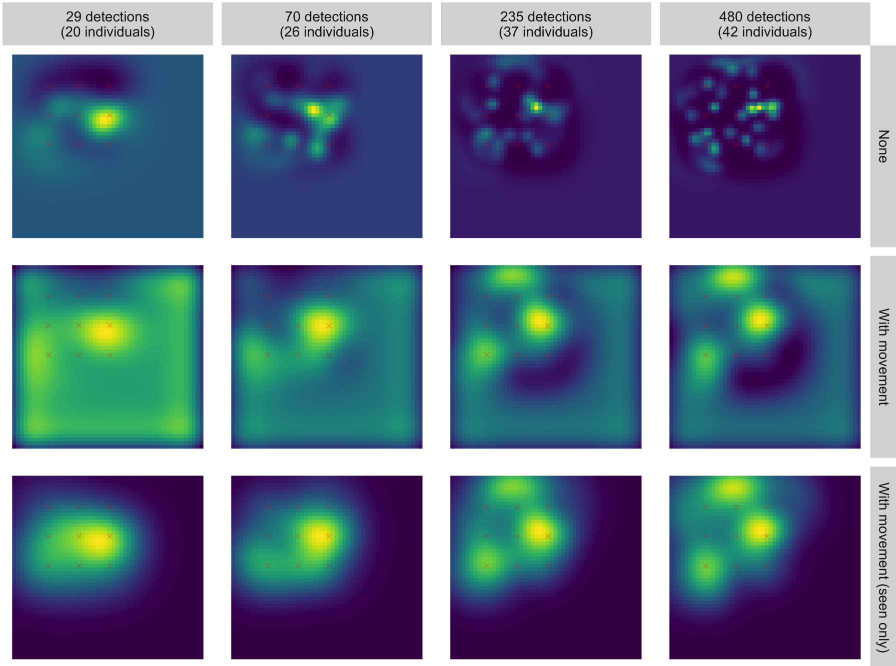

Figure 10: Estimates of realised activity center density surfaces from a constant density model (first row) and realised animal density surfaces incorporating animal movement for both observed and unobserved animals (second row) and for observed animals only (third row). High density areas are indicated in yellow, low density areas in blue. Detectors are shown as red crosses.

5 Camera-trap survey of tigers in Nagarahole, India

5.1 Materials and methods

We reanalysed data obtained from a camera trap survey of tigers *Panthera tigris* living in and around the Nagarahole Tiger Reserve of Karnataka, India, as reported in ?. A full description of the survey can be found in the original reference. The original survey used a spatial array of 162 motion-activated camera traps, these being placed at 2–3 km intervals throughout the area (Figure

11, "All traps").

We reanalyzed this data in a likelihood-based framework, first with a model assuming constant density and with three different trap arrays. The first array was the same one employed in the original study. The second was a subset of traps that excluded a large number traps in the interior of the study region, thus leaving a substantial part of the study area unsurveyed (Figure 11, "Subset #1"). The third used another subset of traps that excluded eight detectors from each of two interior areas of the survey area in which the original survey showed the density of activity centers to be particularly high (Figure 11, "Subset #2").

say how many and what proportion

We then fitted a number of covariate models in which density was assumed to depend on longitude and latitude. We fitted a variety of linear and smooth functions for each of longitude and latitude; the model selected by the AIC was one including a linear effect of latitude only, and we report results from this model only. Finally, we generated realised animal densities for a constant density model with all traps.

5.2 Results

The same broad patterns were visible in our reanalysis of the Nagarahole tiger survey (Figures 11 to 13).

5.2.1 Estimated realised activity center densities

The full array of traps used in the original Nagarahole study clearly showed three areas of high activity center density in the interior of the study region, along $E \approx 625$ and $N = 1324, 1330, 1336$ respectively (Figure 11, "All traps, no cov.").

When we reran the survey on a subset of traps that excludes traps in the interior of the study region, high density areas in the interior of the region were replaced by a flat surface indicating a homogenous low density, and the three high density regions described above were not detected (Figure 11, "Subset #1, no cov."). We also observed some regions where estimated density *increased* after the removal of the interior traps (see the easternmost detectors in Figure 11, "Subset #1, no cov."). This occurs when animals have their activity centers near to, but outside, the area circumscribed by an array – estimated activity centers then tend to be pulled towards the traps that they are closest to.

With the second subset of traps, which exclude eight detectors from each of two high density interior areas, the constant density model still recognized that activity centers are located in these areas, but the estimated locations of these activity centers showed a clear shift from what was found in the original survey (Figure 11, "Subset #2, no cov."). The estimated location of the northernmost of the two activity centers moved to the south east, while the other activity center moved to the south.

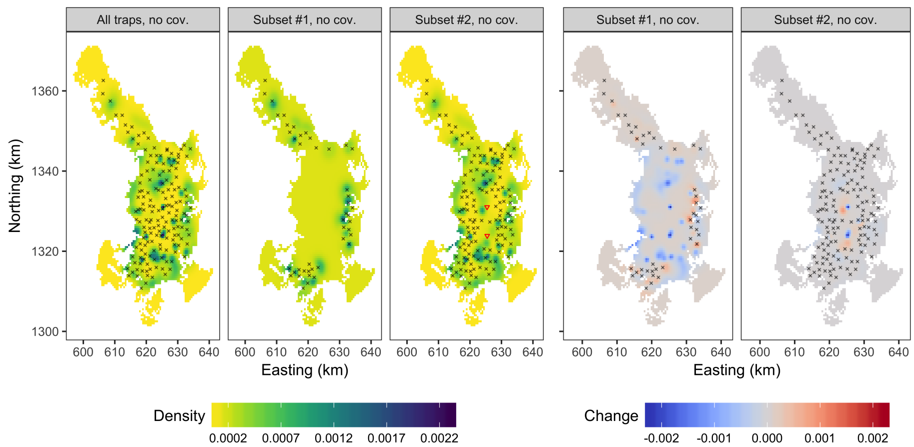

Figure 11: Estimated realised activity center densities of tigers in Nagarahole Tiger Sanctuary, India, obtained using different camera trap arrays. Plots (a), (b), and (c) show estimated densities; plots (d) and (e) show differences between the estimated densities obtained using using trap subset #1 and #2 and those obtained using all traps. Detectors are shown as black crosses.

5.2.2 Estimated expected activity center densities

The model with the lowest AIC was one expressing mean intensity as a linear function of latitude. The estimated density surface obtained from this model showed the estimated mean intensity increasing southwards across the region, with mean intensity in the extreme south roughly four times that in the extreme north (Figure 12, "All traps, northing"). Estimates of expected activity center density were much less variable than estimates of realised activity center density, and were also less sensitive to changes in the array of traps, provided that the array provided sufficient coverage of the covariate space to estimate the covariate relationship (Figure 12, "Subset #1, northing" and 'Subset #2, northing").

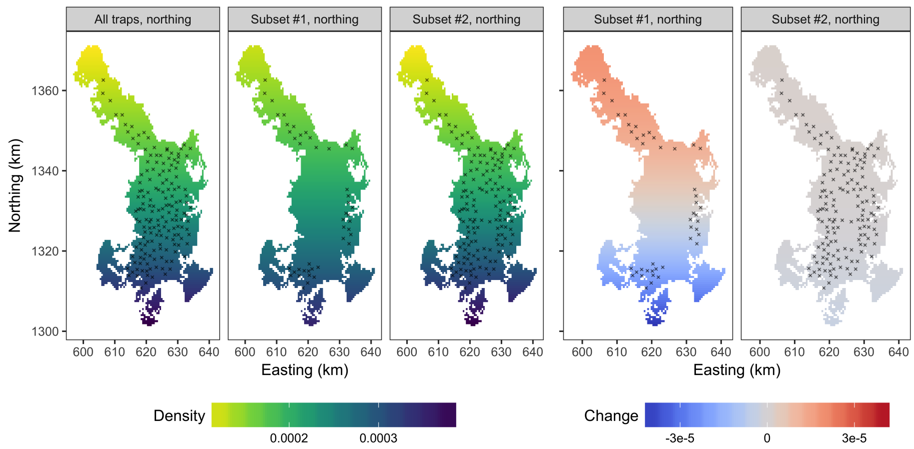

Figure 12: Estimated expected activity center density of tigers in Nagarahole Tiger Sanctuary, India, obtained using different camera trap arrays. Plots (a), (b), and (c) show estimated intensities (expected activity center densities); plots (d) and (e) show differences between the estimated intensities obtained using using trap subset #1 and #2 and those obtained using all traps. Detectors are shown as black crosses.

5.2.3 Estimated realised animal densities

The estimated realised animal density surface differed markedly from the realised activity center density surface, with these differences neatly illustrating the different purposes of the two surfaces (Figure 13). Activity center densities were highest in those cells where sufficient information had been gathered to precisely identify where a single tiger's activity center was. Adding movement to the surface had the effect of dispersing each area of high (activity center) density across a much wider area, the extent of which depended on the estimated range of movement. The estimated spatial scale parameter for the fitted half-normal detection function we used was $\sigma = 1.85$km, so that animals can move a substantial distance from their activity centers, relative to the size of the study area. As a result, animal density was highest in areas in which there were several activity centers in relatively close proximity to one another, even if the location of these activity centers was less precisely known than other activity centers. This occured in areas near the southern and south-western borders of the reserve, as well as in a central location near $N = 1340$ (Figure 13). In contrast, animal density was low in areas that contained only a single activity center, even if the location of the activity center was precisely known (for example at $N = 1330$, $E = 624$).

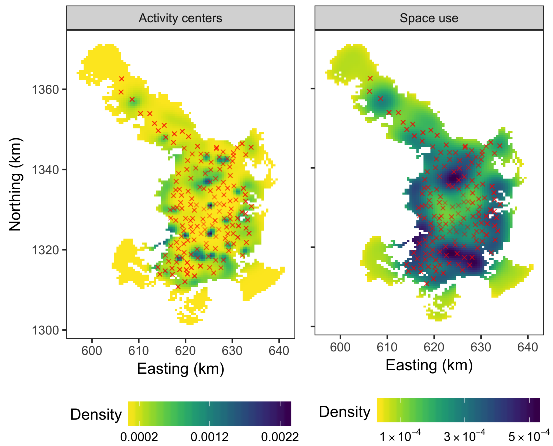

Figure 13: Estimated (a) realised activity center density surfaces from a constant density model and (b) realised animal density surfaces for tigers in Nagarahole Tiger Sanctuary, India. High density areas are indicated in blue, low density areas in yellow. Detectors are shown as red crosses.

6 Discussion

The realised activity centre density obtained from an SCR model cannot be interpreted as a species distribution model. Species distribution models predict where species are likely to occur by correlating environmental covariates with species occurrence or species density. Their rationale is to find favourable habitats and predict that animals will be found in similar habitats across the study region. A species distribution model will tend to place higher densities at locations where environmental covariates are most favourable, and spatial variation in the density surface will depend mostly on how environmental covariates change across space.

In contrast, the density on a realised activity center surface is often placed in spikes where the model is most certain that an activity center is located. The shape and location of these spikes depends on where traps are located and also on survey effort. Different arrays produce quite different results for the same populations

DLB: I removed these words "and these results can be improbable, in the sense that high density spikes can occur at locations in unfavourable habitats, if there happen to be activity centers at these locations at the time of the survey." as this seems a weak argument to me.

A useful metaphor here is of SCR as a torch shining a light onto the true activity center density surface – what you see depends on where you shine the torch (trap locations) and how brightly you shine it (survey effort). If you interpret the darkness outside of the beam to mean that there is little or nothing outside the beam, you fundamentally misunderstand the nature of torches and will draw fundamentally incorrect conclusions.

DLB: Extended the metaphor with this last sentence, which I think works well?

Realised activity center surfaces tend to be flat away from where traps are located. It is important to understand that this flatness reflects a lack of knowledge about the density surface away from traps, and not that the density surface is falt awah from traps. This point is clearly stated in standard SCR texts[2] but is misinterpreted whenever researchers explicitly or implicitly treat realised activity center densities as maps of the spatial distribution of activity centers across the study area (unless the study area is very densely surveyed), as is

[2] "As we move away from 'where the data live' (away from the trap array) we see that the density approaches the mean density. This is a property of the estimator as long as the detection function decreases sufficiently rapidly as a function of distance. Relatedly, it is also a property of statistical smoothers such as splines, kernel smoothers, and regression smoothers—predictions tend toward the global mean as the influence of data diminishes" (p165-166 Royle, Chandler, Sollman & Gardner, 2013)

done in Alexander *et al.* (2015). Another way to see that flatness away from traps reflects uncertainty rather than homogenous density is to plot lower and upper percentiles at each pixel, rather than just the posterior mean – the differences between these percentiles would be large away from traps and small close to traps. It seems that this is rarely done, or at least reported in the literature; a practice that would be worth changing.

DLB: Add other examples?

More importantly for people actually conducting SCR surveys is that the realised density estimates obtained close to traps (and even inside the trap array) *also* depend on where traps are located. The inset plots of Figure 6 and 7 show the same region in space, and this region lies within a 2.5σ range of all trap arrays, where one would expect to be making inferences about activity centers. We obtained very different density surfaces in this area depending on where traps were located. If one was using SCR to identify areas of high density e.g. for conservation purposes, or to locate animals, different areas would be identified depending on which array was used.

When considering realised activity centres, SCR models answer the question "where is an animal with *this* spatial capture history likely to have its activity center?" The answer is always contingent on where traps are located. Changing the locations of detectors also changes the expected capture history, and thus the answer to the question of where the activity center is located. This occurs regardless of whether one works in a Bayesian or frequentist framework. Precisely the same is true of the estimated realised activity center density surface, which simply sums estimated activity centers across animals. In this case the question being addressed is "where are the animals with *these spatial capture histories* likely to have *their* activity centers?" The dependency on trap location applies to activity centers estimated for detected animals and for those that were not detected. In the latter case we have limited information and our estimates thus become just "nowhere near where traps are located".

None of this precludes realised activity center surfaces from being useful sources of information, but they do need to be interpreted with care. For practical purposes this means always interpreting them with the caveat that they depend on where traps are located. Realised activity centre densities do not give proper answers to questions like "where are the high- and low-density regions?" because the highest and lowest points of the surface will always be at or near traps; not because these are high- or low-density regions of space, but because this is where the capture histories make us most certain that animals are, or are not, present. They also cannot answer questions like "are animals clustered in space?" or "is animal density heterogeneous?" because the realised density surface will always exhibit variability, even if animal densities are truly a realisation of a homogeneous Poisson point process.

When estimating the location of a given activity center, the bias caused by trap locations is lowest if the activity center occurs near the center of a dense array of traps, and is highest if traps are all on one side of the activity center or if detections are only made at a single trap. Thus bias can be reduced by using a design that makes it likely that all activity centers in the study region are surrounded by a network of traps. This will be unachievable for most wildlife surveys, as it requires a large number of traps covering an area beyond the study region, and ideally placed at random *[[[note: I say 'beyond the region' so that activity centers at the borders are also in the center of some array, but not sure this is correct – ????]]]* . In summary, it is incorrect to interpret the realised activity center density surface as if it indicated where animals currently have their activity centers.

DLB: The paragraph above seems both redundant (we have already made the point about misinterpretation of activity density surfaces quite well I think), and unclear in that I don't know what "bias" is being referred to in it. I suggest we remove the paragraph.

There is a way of using SCR so that it can be interpreted as a species distribution model – by modelling the mean intensity of the underlying process as a function of enviromental covariates. Covariates allow one to see beyond the spatial extent of the array (see Figure 8), provided that the relationship between covariate and response is a good one, and that traps cover a sufficient range of covariate values to estimate that relationship. The resulting surfaces are no longer tied to one particular realisation of the Poisson process. Rather, they show the (estimated) mean intensity of the underlying process assumed to generate activity centers; in other words, the estimated expected density of activity centers across all realisations. Expected densities will be highest where environmental covariates are most favourable (such as further south in Figure 12). They answer the questions "Where are the high- and low-density regions?" and "What spatial variables are good predictors of the high- and low-density regions?" in a way that is consistent with how this question is answered by species distribution models. They predict where we would expect to see activity centers, if we were able to observe multiple independent populations distribute themselves across the study region.

Using covariate models, and associated model-based inference, is not without issues – there is a danger of extrapolating the density surface beyond the range of covariates around the traps, and the relationship with density and covariate is assumed to be the same everywhere as it is around the traps. The extent to which the expected activity center surface predicts where animals have their activity center *in this realization* depends on the strength of the covariate relationship and on the number of activity centers, each of which is an independent draw from the underlying process. In the Nagarahole survey, for example, there is a relatively weak northing covariate and a relatively small number of activity centers, and the estimated expected activity center density surface provides very little information about the location of current activity centers.

The concept of an activity center is central to SCR models, but for many applications of SCR it may be more appropriate to consider a distribution of space use, taking into account all locations where an animal may have been present, rather than a distribution over activity center locations only. The detection function or encounter function estimated as part of an SCR model provides information about how far from its activity center an animal may move.

This can be easily integrated with the estimated realised activity center density to give an estimated realised *animal* density surface. The resulting surface effectively smooths the realised activity center density surface, with the amount of smoothing determined by the distances that animals move, as given by the detection function. As it is based on activity center locations, the animal density surface also depends on where traps are located and on survey effort. However, it depends less heavily on these factors than the activity center surface because the detection function does not depend on them. In particular, the realised animal density surface quickly stops becoming increasingly "peaked" as more survey occasions are added.

Ultimately, the appropriate density surface to use depends on the aims of the researcher. We have argued that the estimated realised activity center density surface should not be used as a species distribution model, because of the strong dependence on trap location and survey effort. But if the goal is to identify the activity centers of *some* animals currently in the study region (and it does not matter which ones) then it may well be an efficient way of locating these, particularly at the center of the array. If the goal is to actually *find* an animal in the study region, then it is less important where animals have their activity centers and more important to know where they spend their time, and the realised animal density surface is most useful. If the goal is to estimate where animals (not just the ones in the current realisation) are likely to have activity centers, then this is a species distribution question and the expected activity center surface, with intensity a function of covariates, should be used.

BCS: Should this say "realised **usage** density instead?

7 Conclusions

This paper demonstrates that the summed posterior distribution of estimated ranges across animals obtained from SCR – what we call the realised activity center density surface – cannot be used as a species distribution model. We illustrated this point in a number of ways, first with a binomial point process, then by using the Mona Lisa to simulate a Poisson point process, and finally using data from a real-world camera trap survey. All these examples returned the consistent message that realised activity center density surfaces differ depending on trap location. This dependency is most obvious at large spatial scales, where moving a trap array is like "shining a torch" on a particular part of the study area, but is also present within the region in and around the trap array itself. Our main messages are:

1. Realised activity center density surfaces cannot be interpreted as SDMs. This is both because the surface makes inferences about one realisation of a spatial point process, whereas SDMs make inferences about the long run average of the process; and because the surface depends systematically on where traps are located.

2. The realised activity center density surface typically shows highest peaks and deepest troughs close to the center of arrays, defaulting to close to the mean of the underlying process away from the array. A flat density away from traps reflects a lack of knowledge, and not constant density. We should expect some areas away from traps to show substantial deviations from the process mean – it is just that we do not know which areas.

3. An SCR model that models mean activity center density as a function of environmental covariates can be interpreted as a SDM. Here the key difference is that the surface obtained from the covariate model – what we call an expected activity center surface – is a statement about the mean intensity of the underlying process, and is independent of array location provided that the environmental covariate space has been sufficiently sampled.

4. Realised activity center densities can be extended into realised animal densities by the addition of animal movement. This is done by distributing the probability mass associated with each possible location of a particular activity center across the entire region in which, conditional of that loca-

tion being the true one, an animal might be detected. The extent of this region is given by the estimated detection function parameters.

References

Alexander, J.S., Gopalaswamy, A.M., Shi, K. & Riordan, P. (2015) Face value: towards robust estimates of snow leopard densities. *PlosOne*.

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

Borchers, D.L., Stevenson, B.C., Kidney, D., Thomas, L. & Marques, T.A. (2015) A unifying model for capture–recapture and distance sampling surveys of wildlife populations. *Journal of the American Statistical Association*, **110**, 195–204.

Dawson, D. & Efford, M.G. (2009) Bird population density estimated from acoustic signals. *Ecology*, **46**, 1201–1209.

Dorazio, R.M. & Karanth, K.U. (2017) A hierarchical model for estimating the spatial distribution and abundance of animals detected by continuous-time recorders. *PlosOne*, **12**.

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

Elliot, N.B. & Gopalaswamy, A.M. (2016) Toward accurate and precise estimates of lion density. *Conservation Biology*, **31**, 934–943.

Kidney, D., Rawson, B., Borchers, D., Thomas, L., Marques, T. & Stevenson, B. (2013) An efficient acoustic density estimation method with human detectors, applied to gibbons in cambodia. *Ecology*, p. submitted.

Loveridge, R., Kidney, D., S., T., S., E., Eames, J.C. & Borchers, D.L. (2017) First systematic survey of green peafowl pavo muticus in northeastern cambodia reveals a population stronghold and preference for disappearing riverine habitat. *CambodianJournal of Natural History*, p. 157–167.

Royle, J., Chandler, R., Sollman, R. & Gardner, B. (2013) *Spatial capture-recapture*. Academic Press, Boston.

Royle, J., Karanth, K., Gopalaswamy, A. & Kumar, N. (2009) Bayesian inference in camera-trapping studies for a class of spatial capture-recapture models. *Ecology*, **90**, 3233–3244.

Royle, J. & Young, K. (2008) A hierarchical model for spatial capture-recapture data. *Ecology*, **89**, 2281–2289.

Stevenson, B.C., Borchers, D.L., Altwegg, R., Swift, R.J., Gillespie, D.M. & Measey, G.J. (2015) A general framework for animal density estimation from acoustic detections across a fixed microphone array. *Methods in Ecology and Evolution*, **6**, 38–48.