BIOMETRICS 000, 000–000 DOI: 000
000 0000

Variance estimation for systematic designs in spatial surveys

R. M. Fewster

Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

email: r.fewster@auckland.ac.nz

SUMMARY: In spatial surveys for estimating the density of objects in a survey region, systematic designs will generally yield lower variance than random designs. However, estimating the systematic variance is well-known to be a difficult problem. Existing methods tend to overestimate the variance, so although the variance is genuinely reduced, it is over-reported, and the gain from the more efficient design is lost. The current approaches to estimating a systematic variance for spatial surveys are to approximate the systematic design by a random design, or approximate it by a stratified design. Previous work has shown that approximation by a random design can perform very poorly, while approximation by a stratified design is an improvement but can still be severely biased in some situations. We develop a new estimator based on modeling the encounter process over space. The new 'striplet' estimator has negligible bias and excellent precision in a wide range of simulation scenarios, including strip-sampling, distance-sampling, and quadrat-sampling surveys, and including populations that are highly trended or have strong aggregation of objects. We apply the new estimator to survey data for the spotted hyena *(Crocuta crocuta)* in the Serengeti National Park, Tanzania, and find that the reported coefficient of variation for estimated density is 20% using approximation by a random design, 17% using approximation by a stratified design, and 11% using the new striplet estimator. This large reduction in reported variance is verified by simulation.

KEY WORDS: Distance sampling; Encounter rate; Line transect sampling; Plot sampling; Post-stratification; Quadrat sampling; Strip sampling; Systematic sampling; Variance estimation.

1. Introduction

Systematic survey designs are popular in spatial surveys such as strip sampling, quadrat sampling, and distance sampling from lines or points. The aim of these surveys is to estimate density of animals or plants (termed 'objects') in a defined region. Systematic designs use a grid of equally spaced samplers — strips, lines, points, or quadrats — with a random start-point. They are easy to plan and implement in the field, and they generally yield lower variance than random designs in which samplers are placed randomly and independently in the survey region. This is because random designs include realizations where several samplers fall by chance into high density or low density parts of the region, whereas systematic designs ensure even coverage of the region for all realizations. In many situations, systematic designs are also more precise than stratified designs (Cochran, 1946).

The chief disadvantage of systematic designs is the difficulty of estimating the improved variance. A systematic sample is based on only one random start-point, so the samplers are not independent replicates. Wolter (1984; 1985) highlighted three common approaches to systematic variance estimation for sampling a finite population in social statistics:

1. Random estimation, ignoring the problem of non-independent samplers and using estimators derived for random designs;
2. Poststratification, approximating the systematic design by a stratified design by grouping small sets of adjacent samplers into strata, and using stratified variance estimators;
3. Modeling the process producing the finite population, for example by proposing a model for the correlation in response between adjacent members of the population.

Similar ideas are used for spatial surveys. Most analyses ignore the problem (approach 1), but there is increasing recognition that this can be misleading. Millar and Olsen (1995), Simmonds and Fryer (1996), Kingsley (2000), and D'Orazio (2003) all used poststratification (approach 2), and Fewster et al. (2009) extended this scheme to provide estimators for strip or line-transect sampling where line lengths are not equal. However, the poststratification scheme is an approximation and does not yield unbiased estimates for the variance.

The aim of this paper is to develop a new variance estimator for systematic spatial surveys. We create a model for the systematic variance, similar to approach 3 but exploiting the continuous nature of space. We show how the new variance estimator is applied to strip-sampling, line-transect distance-sampling, and quadrat or point-transect sampling surveys. We assess the estimator through a wide range of simulations, reproducing those in two recent studies in which correct variances were not always obtained (Fewster et al., 2009; Johnson, Laake, and Ver Hoef, 2010). We then apply the estimator to distance-sampling data for spotted hyenas in the Serengeti National Park, Tanzania (Durant et al., in review), and show that the new estimator can make a dramatic impact on the standard error and confidence interval width. This result is verified by further simulations. All computations are coded in the R language (R Development Core Team, 2008), and code is available from the author.

2. Strip sampling in a rectangular region

2.1 *Strip-sampling surveys*

We begin with the case of strip sampling in a rectangular region, to derive the new systematic variance estimator in a simple context. We orient the rectangular survey region on a graph with its horizontal base parallel to the x-axis, and consider systematic strip surveys consisting of k equally-spaced vertical search-strips, as in the top panels of Figure 1. The grid of strips has a random start-point. Each search-strip i has horizontal width $2w$ and vertical length $l_i = l$. All objects inside the search-strips are assumed to be detected. The number of objects detected in strip i is n_i for $i = 1, \dots, k$, and the total number of detections is $n = \sum_{i=1}^{k} n_i = k\overline{n}$. The total area surveyed is $2wL$ where $L = kl$ is the total length surveyed. The estimator for object density is $\widehat{D} = n/(2wL) = k\overline{n}/(2wL)$, a multiple of $\overline{n}$.

[Figure 1 about here.]

2.2 *Existing systematic variance estimators*

Wolter (1984; 1985) lists eight variance estimators for estimating the variance of $\overline{n}$ in a systematic survey. Wolter's context is a finite population of rk units (e.g. households), where the first or 'base' household b is randomly chosen from $b \in \{1, 2, \ldots, r\}$, so the systematic sample contains households $b, b+r, b+2r, \ldots, b+(k-1)r$ with observations $n_1, n_2, \ldots, n_k$. The variance arises from the choice of b. Fewster et al. (2009) extended Wolter's recommended estimators to systematic spatial surveys with non-rectangular survey regions, accommodating the influence of variable strip length, l_i, on the number of objects detected in the strip, n_i. Here we give the estimators $\widehat{\text{var}}(\widehat{D})$ for strip-sampling surveys on a rectangular region. We leave the extended formulas for non-rectangular regions to Fewster et al. (2009).

1. Random-line estimators ignore the systematic design and treat the samplers as independent replicates. For a rectangular region, estimator v_1 of Wolter (1984; 1985) and estimators $R1$, $R2$, and $R3$ of Fewster et al. (2009) all reduce to the simple random sampling estimator:

$$\widehat{\text{var}}_R(\widehat{D}) = \frac{k}{4w^2L^2(k-1)} \sum_{i=1}^{k} (n_i - \overline{n})^2.$$

2. Stratified estimators with non-overlapping strata approximate the systematic design by a stratified design, such that pairs of adjacent strips from the systematic survey are treated as if they were drawn from a stratified design with two units per stratum. Assuming k is even, estimator v_3 of Wolter (1984; 1985) and estimators $S1$ and $S2$ of Fewster et al. (2009) become the following estimator for a rectangular region:

$$\widehat{\text{var}}_S(\widehat{D}) = \frac{1}{4w^2L^2} \sum_{h=1}^{k/2} (n_{2h-1} - n_{2h})^2.$$

3. Stratified estimators with overlapping strata also rely on post-strata of paired strips, but allow the post-strata to overlap to improve precision. For a rectangular region, estimator v_2 of Wolter (1984; 1985) and estimators $O1$ and $O2$ of Fewster et al. (2009) become:

$$\widehat{\text{var}}_O(\widehat{D}) = \frac{k}{4w^2L^2 \times 2(k-1)} \sum_{i=1}^{k-1} (n_i - n_{i+1})^2.$$

Fewster et al. (2009) found that the random-line estimators $\widehat{\text{var}}_R$ can perform very poorly for systematic surveys, especially when object density follows strong trends in the x-direction. The stratified estimators $\widehat{\text{var}}_S$ and $\widehat{\text{var}}_O$ perform similarly to each other, much better than $\widehat{\text{var}}_R$ but still not capturing the systematic variance correctly. The boxplots in Figure 1 show why improvement is still needed: estimator R vastly overestimates the systematic variances, while estimators S and O are much better but are not reliably correct. The aim of this paper is to create a new variance estimator specifically for systematic spatial surveys, to yield further substantial improvement in both bias and precision for estimating $\text{var}(\widehat{D})$.

2.3 *Repeated-survey variance and systematic sampling frame*

In order to estimate $\text{var}(\widehat{D})$, we must specify a conceptual population of replicate surveys that create the variance in $\widehat{D}$ that we seek to estimate. We imagine repeating surveys numerous times according to some specified rules, and obtaining an estimate $\widehat{D}$ for each replicate survey. The variance of these replicate $\widehat{D}$ estimates is the 'true' $\text{var}(\widehat{D})$. The set of rules for generating conceptual replicates is called the inferential framework, because it creates the distributions from which we evaluate bias, precision, and confidence interval coverage of proposed estimators. The framework is entirely the choice of the investigator (Gregoire, 1998), but it must be clearly specified because it defines all expectations and variances. In our case, the inferential framework defines the distribution of $\widehat{D}$, and thence the true estimator variance $\text{var}(\widehat{D})$. It also defines the distribution of our estimator $\widehat{\text{var}}(\widehat{D})$, so that we can evaluate bias and precision of $\widehat{\text{var}}(\widehat{D})$ as an estimator of $\text{var}(\widehat{D})$.

We use the inferential framework specified by Fewster et al. (2009), termed the repeated-survey framework. In the repeated-survey framework, the total number of objects in the region, N, is fixed for all replicate surveys. Every survey has k search-strips with fixed systematic spacing and fixed orientation, but with a random start-point for the systematic grid. Object positions change between replicate surveys, mimicking mobile animals. For each replicate, object positions are drawn from some spatial probability density function (p.d.f.) or point process, which might include features such as clustering or trends. The same p.d.f. is used for every replicate, but we wish to derive estimators that will work well for any choice of p.d.f. Once the positions of the search-strips and the objects are determined, we obtain the data of stripwise detections $n_1, \ldots, n_k$, calculate $\widehat{D} = n/(2wL)$, and calculate our chosen estimator for $\widehat{\text{var}}(\widehat{D})$. New draws of search-strip locations and object positions are made for each replicate survey, to form the distributions of the estimators $\widehat{D}$ and $\widehat{\text{var}}(\widehat{D})$.

We now define the systematic sampling frame. We align the survey region such that its base extends between x-coordinates 0 and 1. The k search-strips are parallel with the y-axis. Each strip has width w on either side of its centerline. The spacing between centerlines is r, and the first centerline has x-coordinate b, so the k centerlines have x-coordinates $b, b+r, \ldots, b+(k-1)r$. Our systematic sampling frame specifies that $b \sim \text{Uniform}[w, w+r)$, so the leftmost systematic grid starts at w to accommodate the whole of the first search strip. The first strip of the rightmost grid coincides with the second strip of the leftmost grid. The kth centerline of the rightmost grid lies at $x = 1 - w$ to accommodate the whole of the last search strip. The sampling frame and repeated-survey framework are shown in the top two panels of Figure 1, highlighting the mobile objects and different grid positions.

The chosen sampling frame implies $r = (1 - 2w)/k$. It is the investigator's choice whether r is determined given k, or vice versa, or the limits of the survey region are defined to incorporate the sampling frame given r and k. This sampling frame undersamples the regions $x \in (0, w)$ and $x \in (1 - w, 1)$. Consequences of this are likely to be minor, but if necessary a wrap-around sampling frame can be used instead (Fewster and Buckland, 2004: 293).

2.4 *Components of systematic variance*

For strip-sampling on a rectangular region with $\widehat{D} = n/(2wL)$, finding $\widehat{\text{var}}(\widehat{D})$ is equivalent to finding $\widehat{\text{var}}(n)$ because L does not change between replicate surveys. The top row of Figure

1 shows the components of $\mathrm{var}(n)$ in the repeated-survey framework. The first component is the variance in n due to the grid base $b \in [w, w + r)$. If there are trends in object density across the region, as in Figure 1, the expected number n of objects encountered may change with b. The line in panel 3 shows $E(n \mid b)$ as b changes. If large-scale trends are present, there is information about them contained in the systematic sample, because the stripwise detections $n_1, \ldots, n_k$ span the width of the region.

The second component of $\mathrm{var}(n)$ is the variance for a given grid base b. This is shown by the scatter about the line in panel 3 of Figure 1, which shows n from 100 draws from the repeated-survey framework. The scatter about the line is caused by objects changing location for each replicate survey.

The law of total variance summarizes the two components of repeated-survey variance:

$$\mathrm{var}(n) = \mathrm{var}_b \left\{ E(n \mid b) \right\} + E_b \left\{ \mathrm{var}(n \mid b) \right\}. \tag{1}$$

The term $\mathrm{var}_b \left\{ E(n \mid b) \right\}$ is due to large-scale trends and describes how $E(n \mid b)$ varies as b changes, visible from the slope of the line $E(n \mid b)$ in panel 3 of Figure 1. The term $E_b \left\{ \mathrm{var}(n \mid b) \right\}$ describes the mean scatter about the line $E(n \mid b)$. We construct our estimator for $\mathrm{var}(n)$ by modeling the distribution of $[n \mid b]$.

2.5 *Striplet partition of the survey area*

Our approach to constructing a variance estimator for spatial systematic surveys is as follows. First we approximate the continuous sampling-frame $b \sim \mathrm{Uniform}[w, w + r)$ by a discrete sampling frame $b \in \{b_1, b_2, \ldots, b_B\}$ for some large B, where $b_1, \ldots, b_B$ are equally-spaced x-coordinates such that $b_1 = w$ and $b_B \leqslant w + r$: see Figure 2.

[Figure 2 about here.]

Secondly we plot the boundaries of all search-strips and centerlines in all B systematic grids based at $b_1, \ldots, b_B$. The boundaries partition the region $0 \leqslant x \leqslant 1$ into J thin parallel strips, which we term 'striplets' to distinguish them from the search-strips in the survey. These striplets will not generally be of equal width or area. They are introduced for convenience so that every search half-strip in the grids based at $b_1, \ldots, b_B$ consists of an integer number of striplets (e.g. three in Figure 2). The striplets remain invariant across survey replicates, while the changing grid location renders any particular striplet 'active' or 'inactive' for that replicate. We label the striplets $1, 2, \ldots, J$, and define the set $\mathcal{S}(b_i) \subset \{1, \ldots, J\}$ for any b_i, where $\mathcal{S}(b_i)$ is the set of 'active' striplets contained in the systematic grid based at b_i.

We next introduce a random variable $\boldsymbol{X} = (X_1, X_2, \ldots, X_J)$, where $\boldsymbol{X}$ gives the number of objects contained in striplets $1, 2, \ldots, J$ for a particular replicate survey. $\boldsymbol{X}$ describes object relocations in the repeated-survey framework. Its true distribution in the repeated-survey framework is unknown, depending on the spatial p.d.f. or point process used to generate object positions. The true repeated-survey distribution of n is created by survey replicates such that each replicate involves a single draw from the true distribution of $\boldsymbol{X}$ and a single grid base $b \sim \text{Uniform}[w, w + r)$, leading to a single value of n. The key idea of the new 'striplet' variance estimator is to approximate the repeated-survey distribution of n by modeling the distribution of $\boldsymbol{X}$ and approximating $b \sim \text{Uniform}[w, w+r)$ by $b \in \{b_1, \ldots, b_B\}$. A repeated-survey replicate is approximated by a single draw from the modeled distribution of $\boldsymbol{X}$, and a single value of $b \in \{b_1, \ldots, b_B\}$, giving the observation $n = \sum_{j \in \mathcal{S}(b)} X_j$. The modeled distribution of $\boldsymbol{X}$ is used to derive estimates for $\text{var}(n \mid b)$ and $E(n \mid b)$. We then sum over $b_1, \ldots, b_B$ using (1) to estimate $\text{var}(n)$.

2.6 *Model for $\boldsymbol{X}$*

In the repeated-survey framework, the total number of objects N is fixed, but objects are relocated between replicate surveys. A reasonable model for $\boldsymbol{X}$ is $\boldsymbol{X} \sim \text{Multinomial}(N\,;\, p_1, \ldots, p_J)$ where p_j gives the probability of a relocated object falling in striplet j. The multinomial model arises naturally if the original point process is a Poisson process such that the Poisson intensity in striplet j is proportional to p_j. The distribution of Poisson random variables, conditional on their sum being N, is multinomial.

Specifying a multinomial distribution for $\boldsymbol{X}$ means that (a) the trend function $E(n \mid b)$ can be estimated once the cell probabilities $p_1, \ldots, p_J$ are estimated; and (b) the scatter function $\text{var}(n \mid b)$ is assumed to be that of the multinomial distribution. These assumptions are found to be very robust by simulations, but there may be situations where alternative distributions for $\boldsymbol{X}$ are preferred: see Section 3.4.

The problem of constructing $\widehat{\text{var}}(\widehat{D})$ reduces to estimating the cell probabilities $p_1, \ldots, p_J$. Note that $\text{var}(\widehat{D})$ is not generated by the multinomial model for $\boldsymbol{X}$: it is generated by the repeated-survey replicates. We hope that the multinomial variance will be a good model for the repeated-survey variance, and insensitive to the choice of spatial p.d.f. for the repeated-survey replicates. However, for biological realism, we continue to evaluate our estimators against their repeated-survey distributions, not against the multinomial distribution.

We use the information from the stripwise detections $n_1, \ldots, n_k$ to estimate the cell probabilities $p_1, \ldots, p_J$. We favor a model that allows $p_1, \ldots, p_J$ to change smoothly over space. Let striplet j have area α_j, and let the x-coordinate of the midpoint be t_j. Let $\lambda(\cdot)$ be a one-dimensional intensity function, such that $\lambda(t)$ is the instantaneous intensity of objects per unit area at x-coordinate t. The striplet cell probabilities are $p_j = \alpha_j \lambda(t_j) / \sum_{m=1}^{J} \alpha_m \lambda(t_m)$. We wish to estimate $\lambda(t)$ as a smooth function of t. Suppose that the search-strips $1, \ldots, k$ have centerlines at $x = t_1^*, \ldots, t_k^*$, and areas $a_1, \ldots, a_k$, then our model is $E(n_i) = a_i \lambda(t_i^*)$. We could use an ordinary scatterplot smoother to estimate $\lambda(\cdot)$ given $n_1, \ldots, n_k$, but the offset of strip areas a_i makes it convenient to use the functionality of R package `gam` (Hastie and Tibshirani, 1990) for fitting generalized additive models (GAMs). Using `gam` with log link creates the formula $\log\{E(n_i)\} = \log(a_i) + \log\{\lambda(t_i^*)\}$, which allows a quick way of generating a smooth estimate for $\lambda(\cdot)$ while allowing for possibly different strip areas a_i. We fit the GAM with Poisson error, log link, offset term $\log(a_i)$, and smooth term `s(t, df=4)` from which $\widehat{\lambda}(t)$ can be obtained for any t . The choice of Poisson error is consistent with using a multinomial distribution for $\boldsymbol{X}$, but the estimated mean curve is fairly insensitive to the choice of error distribution. The multinomial parameter N is estimated by $\widehat{N} = A\widehat{D}$, where $A = \sum_{j=1}^{J} \alpha_j$ is the total area of the survey region.

2.7 *Striplet variance estimator*

We assemble the striplet variance estimator for $\text{var}(n)$ as follows. There are $X_1, \ldots, X_J$ objects in striplets $1, \ldots, J$, where we model $\boldsymbol{X} \sim \text{Multinomial}(N\,; p_1, \ldots, p_J)$. Given a systematic grid based at b, define $Q(b) = \sum_{j \in \mathcal{S}(b)} p_j$. Using marginal properties of the multinomial distribution, the distribution of $[n \,|\, b]$ is Binomial $(N, Q(b))$, so $E(n \,|\, b) = NQ(b)$ and $\text{var}(n \,|\, b) = NQ(b)\,\{1 - Q(b)\}$. Thus from (1):

$$
\begin{aligned}
\text{var}(n) &= \text{var}_b \{NQ(b)\} + E_b \Big[NQ(b)\,\{1 - Q(b)\} \Big] \\
&= E_b \left\{ N^2 Q(b)^2 \right\} - \left[E_b \{NQ(b)\} \right]^2 + E_b \Big[NQ(b)\,\{1 - Q(b)\} \Big] .
\end{aligned}
$$

We use sample means from the discrete set of values $b_1, \ldots, b_B$ to estimate the expectations:

$$
\widehat{\text{var}}(n) = \frac{1}{B} \sum_{b=b_1}^{b_B} \left[\widehat{N}\widehat{Q}(b) \left\{ 1 - \widehat{Q}(b) \right\} + \widehat{N}^2 \widehat{Q}(b)^2 \right] - \left\{ \frac{1}{B} \sum_{b=b_1}^{b_B} \widehat{N}\widehat{Q}(b) \right\}^2 , \qquad (2)
$$

where $\widehat{Q}(b) = \sum_{j \in \mathcal{S}(b)} \widehat{p}_j$, $\widehat{p}_j = \alpha_j \widehat{\lambda}(t_j) / \sum_{m=1}^{J} \alpha_m \widehat{\lambda}(t_m)$, and $\widehat{N} = A\widehat{D}$.

Equation (2) gives the striplet variance estimate for rectangular regions. Computation takes only about a second on a laptop computer, to set up the grid $b_1, \ldots, b_B$ and striplets $1, \ldots, J$, and fit the GAM. The estimated density variance is $\widehat{\text{var}}(\widehat{D}) = (2wL)^{-2}\, \widehat{\text{var}}(n)$.

2.8 *Simulations for strip surveys*

Figure 1 shows the striplet estimator in practice. The top row shows realizations of the repeated survey framework creating the distribution of $[n \,|\, b]$. The lower rows show two highly trended populations. In each case, a single population realization is shown, followed by the corresponding set of stripwise detections $n_1, \ldots, n_k$ for $k = 20$ and the smooth output from the fitted GAM, which is proportional to $\lambda(t)$ scaled by strip areas. This single GAM fit constructs an estimated line for $E(n \mid b)$, shown on the next panel. The line is an estimate from the single survey shown, whereas the cloud of data points shows the true distribution of $[n \mid b]$ from 10,000 simulations. The line represents the trend component $E(n \mid b)$ and of importance is its slope, rather than its height, to accurately capture $\text{var}_b \{E(n \mid b)\}$ in (1). The scatter about the line is modeled by the multinomial variance to give $E_b \{\text{var}(n \mid b)\}$. The two components are added to give the striplet estimate. Results from 10,000 repeats of this procedure are shown in the boxplots. The true repeated-survey $\text{var}(\widehat{D})$ is shown by the bold horizontal line across the boxplots. In both cases the striplet estimator $\widehat{\text{var}}(\widehat{D})$ exhibits negligible bias for $\text{var}(\widehat{D})$, and it is also much more precise than any of the other estimators.

To understand the relative importance of the two striplet variance components, we also show results from a similar procedure that omits the GAM phase and replaces $\lambda(t)$ by a constant function, such that $E(n_i) = a_i \lambda_{\text{const}} = \overline{n}$. The constant function is shown as a dashed line on the panels from the single survey replicates, and the resulting estimator is called `nogam`. This estimator ignores the information on large-scale trends. It performs poorly for Population A, because there the trend is monotonic and consequently the line $E(n \mid b)$ has large slope representing a large component of $\text{var}_b \{E(n \mid b)\}$. However, the loss in Population B is of little importance, despite the strong trend in Population B. This is because the trend is not monotonic, so the number of detections even out over the systematic grid such that the aggregate $E(n \mid b)$ changes very little over b: it is barely distinguishable from the dashed line representing a constant fit. For Population B, both the striplet and `nogam` estimators are so precise that their boxplots are barely visible. However, the inaccuracy of the `nogam` fit is sufficient to cause a slight loss in confidence interval coverage.

For simulations in Figures 1, 3 and 4, the striplet method uses a grid $b_1, \ldots, b_B$ with spacing 0.0005. The simulated data use the continuous sampling frame $b \sim \text{Uniform}[w, w{+}r)$. Confidence intervals use a log-Normal approximation described in generality in Section 4.

3. Extension to other spatial surveys

3.1 *Non-rectangular regions with perfect detection*

If the sampling region is not rectangular, the total line length L in the survey changes according to the position of the grid base b, so both n and L are random for each survey replicate. We write $L(b)$ for the line length associated with grid base b. The variance of $\widehat{D}$ is now the variance of a ratio, var $\{n/(2wL)\}$. This is easily incorporated in (1) because of the conditioning on b. We obtain $\widehat{\text{var}}(\widehat{D}) = (2w)^{-2}\,\widehat{\text{var}}(n/L)$, where

$$\widehat{\text{var}}\left(\frac{n}{L}\right) = \frac{1}{B}\sum_{b=b_1}^{b_B}\left[\frac{\widehat{N}\widehat{Q}(b)\left\{1-\widehat{Q}(b)\right\} + \widehat{N}^2\widehat{Q}(b)^2}{L(b)^2}\right] - \left\{\frac{1}{B}\sum_{b=b_1}^{b_B}\frac{\widehat{N}\widehat{Q}(b)}{L(b)}\right\}^2 . \tag{3}$$

3.2 *Line-transect distance sampling with imperfect detection*

Line-transect distance sampling (Buckland et al., 2001) is a suite of methods for estimating object density when there is imperfect detection of objects within the search-strips. The estimator $\widehat{D}$ includes a detection function, $g(x, \theta)$, which is the probability that an object in the search-strip is detected, given that it is distance x from the strip centerline. The detection parameter θ is estimated using the observed object distances from the centerline. The average detection probability for objects in the search-strip is $P_a = w^{-1}\int_0^w g(x,\theta)\,dx$. If individuals occur in groups, for example family units of animals, then the 'object' is defined as the group of individuals, and the mean group size in the population, $E(S)$, must be estimated.

The density estimator for line-transect distance sampling is

$$\widehat{D} = \frac{n}{\widehat{P}_a} \times \frac{1}{2wL} \times \widehat{E}(S) = \frac{1}{2w} \times \frac{n}{L} \times \frac{1}{\widehat{P}_a} \times \widehat{E}(S) ,$$

where the second expression factorizes $\widehat{D}$ into components of encounter rate (n/L), detection probability, and group size. The density variance is estimated using the delta method:

$$\left\{\widehat{\mathrm{cv}}(\widehat{D})\right\}^2 = \left\{\widehat{\mathrm{cv}}\left(\frac{n}{L}\right)\right\}^2 + \widehat{\mathrm{cv}}\left(\widehat{P}_a\right)^2 + \widehat{\mathrm{cv}}\left\{\widehat{E}(S)\right\}^2 , \tag{4}$$

where $\widehat{\mathrm{cv}}(\cdot)$ denotes the estimated percentage coefficient of variation (CV) of an estimator, for example $\widehat{\mathrm{cv}}(\widehat{D}) = \sqrt{\widehat{\mathrm{var}}(\widehat{D})}/\widehat{D} \times 100$.

Each set of survey data contains replicate information about distances and group sizes, so estimating the variance of $\widehat{P}_a$ and $\widehat{E}(S)$ is straightforward (Buckland et al., 2001). The striplet estimator (3) is used for the encounter rate variance, $\widehat{\mathrm{var}}(n/L)$. However, the repeated-survey replicates now include variance in which objects are detected, given the grid placement b and the object locations $\boldsymbol{X}$, so the striplet estimator must be adjusted for imperfect detection.

Given a grid base, b, and striplets $1, 2, \ldots, J$, the set $\mathcal{S}(b)$ gives the 'active' striplets in the grid. For active striplets we know the distance of the striplet midpoint to the nearest centerline in the grid. We can therefore estimate $g_j(b)$, the probability that an object located in striplet j is detected in a grid with base b, using the estimated detection function $g(\cdot, \widehat{\theta})$. Let $g_j(b) = 0$ if $j \notin \mathcal{S}(b)$. The former expression $Q(b) = \sum_{j \in \mathcal{S}(b)} p_j$ is now replaced by $Q(b) = \sum_{j=1}^{J} g_j(b) p_j$. The linewise detections $n_1, \ldots, n_k$ are thinned by the detection process, so the GAM formulation becomes $E(n_i) = P_a \times a_i \lambda(t_i^*)$ and the former offset term of $\log(a_i)$ is replaced by $\log(\widehat{P}_a) + \log(a_i)$. The estimated number of objects, $\widehat{N}$, refers to groups rather than individuals, so $\widehat{N} = n/(2wL\widehat{P}_a)$. With these three changes, the striplet variance estimator for line-transect distance sampling is given by (3) and (4).

3.3 *Two-dimensional surveys*

For two-dimensional surveys, such as quadrat or point-transect sampling, there are no variable line lengths so we use (2). Instead of dividing the region into striplets, we divide it into 'boxlets' such that the region is tesselated by a large number of tiny quadrats. This is accomplished by applying the algorithm for striplet boundaries to both the x- and the y-directions. The start-points $b = 1, \ldots, B$ now index a starting-grid in two dimensions.

Detection less than perfect, or boxlets that do not fit wholly inside a systematic grid, can be accommodated through $g_j(b)$ as in Section 3.2. We use a two-dimensional GAM with smooth term of the form `s(x, y)`, which demands adequate sampling coverage in both directions.

3.4 *Alternative distributions for* $\boldsymbol{X}$

A multinomial distribution for $\boldsymbol{X}$ arises naturally from an inhomogeneous Poisson process, so we expect it to cater for strong trends in object density. It is less clear how robust it will be to aggregation of objects. This is tested by simulations in Sections 4 and 5. The Dirichlet compound multinomial distribution might perform well if there is extreme aggregation.

If $\boldsymbol{X}$ is not modeled by a multinomial distribution, the alternative to (3) is derived using $n = \sum_{j=1}^{J} Y_j$ where Y_j is the number of objects detected in striplet j, and $[Y_j \mid X_j, b] \sim \text{Binomial}(X_j, g_j(b))$. The model for $\boldsymbol{X}$ specifies $\text{cov}(X_i, X_j)$ for all i, j, from which $[\text{cov}(Y_i, Y_j) \mid b] = g_i(b)g_j(b)\text{cov}(X_i, X_j)$. We then find $E(n \mid b)$ and $\text{var}(n \mid b)$ and combine as before.

If the chosen inferential framework requires N to change between survey replicates, this can be accommodated in the model for $\boldsymbol{X}$. Examples include $X_j \sim \text{Poisson}(\alpha_j \lambda(t_j))$ with $X_1, \ldots, X_J$ independent, or negative binomial distributions with dispersion estimated from the data, or a multivariate distribution such as the negative multinomial.

4. Simulation studies

We evaluate the striplet estimator using extensive simulations with line-transect distance sampling, intended to provide a stringent test of the estimator's efficacy. The simulation scenarios are sometimes extreme compared with likely scenarios for field surveys. However, the distance-sampling software Distance (Thomas et al., 2010) has thousands of registered users, whose field surveys are likely to include some extraordinary cases.

We compare the striplet method against the estimators detailed in Fewster et al. (2009) and outlined in Section 2.2. Briefly, estimators $R2$ and $R3$ are random-line estimators. Estimator $R2$ is preferred over $R3$, being more robust to spatial trends. Estimator $R3$ is the default in versions 5 and below of the Distance software, while estimator $R2$ is the default in versions 6 and above. Estimators $S1$ and $S2$, and $O1$ and $O2$, give different ways of implementing poststratification with non-overlapping strata and overlapping strata respectively. They all perform similarly in practice, although $O1$ and $O2$ tend to be a little more precise.

We focus on the following challenging survey characteristics for estimating $\text{var}(n/L)$.

— Strong trends in object density in the x-direction across the survey region.

— High variability in line length $L(b)$ as b changes, especially when low $L(b)$ is associated with high encounter rates and vice versa. This can happen in practice when line lengths are dictated by some geographical feature that also impacts density, such as fiords in marine surveys. Fewster et al. (2009) focused on this and the previous scenario.

— Object aggregation, in which objects tend to be attracted to other objects. Johnson et al. (2010) treated this case from a model-based perspective. They used an overdispersion correction factor, but this did not yield the correct variances when aggregation was strong.

— High coverage of the survey area. Systematic designs have large gains in efficiency relative to the other designs when the covered area $2wL$ is a large fraction of the total area, because inefficiencies caused by overlapping strips in the random and stratified designs become substantial. Hence this scenario is challenging for estimators that approximate the systematic design by random designs ($R2$ and $R3$) or stratified designs ($S1$ to $O2$).

— Small number of lines, k. Estimators $R2$ to $O2$ are based on the variance of linewise encounter rates, so their performance is severely impeded if k is small. Buckland et al. (2001: 232) advise that a minimum of $k = 10$ lines should be used.

Confidence intervals for all simulations are calculated by assuming a log-Normal distribution for $\widehat{D}$, justified asymptotically by Fewster and Jupp (2009). The 95% confidence interval is $(\widehat{D}/C,\ \widehat{D} \times C)$, where $C = \exp\left[z_\alpha \sqrt{\log\left\{1 + \widehat{\text{var}}(\widehat{D})/\widehat{D}^2\right\}}\right]$, and where z_α is the upper 0.025 point of the Normal(0, 1) distribution. We use this expression for the striplet estimator. For estimators $R2$ to $O2$, we replace z_α by a Student t_{df} quantile as in Fewster et al. (2009), with d.f. given by the Satterthwaite approximation in Buckland et al. (2001).

4.1 *Distance sampling with strong trends and disparate line lengths*

We redo the simulations of Fewster et al. (2009) on a triangle-shaped region with strong trends in object density. The triangle shape creates considerable variability in total line length, $L(b)$, as b changes. Object positions are simulated using beta distributions on both the x and the y coordinates. The simulations include scenarios where encounter rate is highly correlated with line length. Distance sampling data are simulated using a half-normal detection function, and the detection function is estimated from the data in each simulation.

Figure 3 shows simulation results for $\widehat{\mathrm{se}}(\widehat{D})$ from four spatial p.d.f.s corresponding to populations 1, 2, 3, and 6 from Fewster et al. (2009). The boxplots show estimator performance for two difficult scenarios: the first with a small number of lines ($k = 10$), and the second with high coverage of the survey area (coverage = 80%). The striplet estimator is the only one to exhibit negligible bias in all of the survey scenarios, and it is by far the most precise of the estimators. The gain in precision is even more marked when examining $\widehat{\mathrm{se}}(n/L)$ instead of $\widehat{\mathrm{se}}(\widehat{D})$, because the overall $\widehat{\mathrm{cv}}(\widehat{D})$ in (4) is roughly half composed of an invariant contribution from estimating the detection function, $\widehat{\mathrm{cv}}(\widehat{P}_a)$.

[Figure 3 about here.]

The results shown are characteristic of a large number of similar simulations. The performance of the striplet estimator is always similar to that in Figure 3. Confidence interval coverage is close to 95%, the estimator displays no discernible bias, and relative to the other estimators the mean confidence interval width is reduced while the precision of $\widehat{\mathrm{se}}(n/L)$ is greatly improved. The other estimators rarely underestimate variance, but sometimes have confidence interval coverage close to 100%, and are much less precise. A whimsical observation is that the striplet variance estimator often outperforms the sample variance estimator constructed by literally repeating the survey 100 times and finding the sample var(n/L). The striplet estimator is often more precise, and is obtained at 1% of the cost.

The striplet estimator (3) relies upon the estimated $g(\cdot, \widehat{\theta})$ to gain $\widehat{Q}(b)$, which will induce some dependence between components of $\widehat{\mathrm{cv}}(\widehat{D})$ in (4). In these simulations, the maximum correlation between components was 0.02, so the dependence appears to be minimal.

4.2 *Distance sampling with overdispersed Poisson point process*

Johnson et al. (2010) describe a model-based approach to distance sampling and implement it in the R package `DSpat`, which includes functions for generating intensity maps and simulating object locations using a Cox process. Points are generated as a log-Gaussian Cox process, so points follow a Poisson process for which the log-intensity is itself a Gaussian random field with possible spatial correlation. When there is no spatial correlation in the Gaussian random field, the process is an inhomogeneous Poisson process. When there is spatial correlation, the points are aggregated and the linewise detections are overdispersed relative to the inhomogeneous Poisson process. Johnson et al. (2010) did not obtain the correct variances in the overdispersed case. We reproduce their simulations with the striplet estimator, but with the important difference that we keep N constant between survey replicates in the repeated-survey framework, whereas Johnson et al. allowed N to vary in accordance with the point process. This variance in N might explain their difficulty in producing the correct variances, because the single sample contains little information about var(N) when the point process is overdispersed. We also keep the covariate map fixed between replicates, in accordance with a fixed habitat.

[Figure 4 about here.]

We reproduced all the simulations of Johnson et al. (2010) in the repeated-survey framework. The survey region is rectangular, so $S1 = S2$, $O1 = O2$, and $R2 = R3$. Figure 4 shows results from simulations modeled on the worst of the results in Figure 2 of Johnson et al.

(2010). The other simulations all gave similar results. The striplet estimator has negligible bias and good confidence interval coverage, is by far the most precise, and yields mean reported %CVs lower than any other estimator. For the inhomogeneous Poisson process without overdispersion (top row of Figure 4), the true variances from the repeated-survey framework are very similar to those from Johnson et al.'s (2010) model-based framework (J. L. Laake, pers. comm.), even though they were obtained from different conceptual replicate schemes. Johnson et al. (2010) obtained the correct variances in this case. The spatial aggregation in the overdispersed process can be seen in the second row of Figure 4. Overdispersion has negligible impact on the true repeated-survey variances in these cases, which can be seen by comparing the bold horizontal lines in the two rows of Figure 4. However, overdispersion has a detrimental impact on all estimators except the striplet estimator.

Because the striplet estimator is underpinned by a multinomial spatial relocation model, it should be possible to observe poor performance by violating this model with highly aggregated populations. The striplet estimator is robust to all scenarios examined by Johnson et al. (2010), which they termed high overdispersion. However, it eventually underestimates variance when subjected to a combination of extreme aggregation, high coverage, large sample sizes, and small numbers of transect lines. For example, with overdispersed point processes similar to those in Figure 5 (top row), confidence interval coverage fell to 80% with $k = 20$, $E(n) = 400$, and coverage 50%. The same settings with $k = 40$ restored confidence interval coverage to 94%. The non-striplet estimators each gave approximately 100% coverage in both cases. The array of influences is complex, and general guidelines for robustness to aggregation are difficult to provide. We recommend that simulations are used to check robustness to anticipated levels of aggregation for a particular survey region and design. A different model for $\boldsymbol{X}$ might be needed in some applications.

4.3 *Quadrat sampling with overdispersed Poisson point process*

We implemented the 'boxlet' method of Section 3.3 for estimating $\text{var}(\widehat{D})$ in quadrat-sampling surveys, and applied it to the populations simulated in Section 4.2. Detection was assumed perfect within a quadrat. The striplet estimator was compared against the following estimators: (i) $P1$ from Fewster et al. (2009), which is gained from the sample variance of quadrat counts and ignores the systematic design; (ii) $P4$ from Fewster et al. (2009), which uses the scheme of D'Orazio (2003) to correct for a two-dimensional systematic design; and (iii) the estimator of Millar and Olsen (1995), which is a quadrat sampling equivalent of $O1$ with four quadrats per post-stratum.

Results for quadrat sampling were generally similar to those shown in Figures 3 and 4, with the striplet estimator outperforming all the other estimators in terms of bias and (to a lesser extent) precision. However, if two-dimensional grids were used with only five lines of quadrats in either direction, the striplet estimator tended to underestimate variance for overdispersed point processes. The non-striplet estimators also trended downwards in this situation. The number of lines or rows appears to be more important than the sampling coverage: for example, switching from grid dimensions 5×10 to 10×10 while maintaining the same coverage restored the good performance of the striplet estimator in every case examined, while switching from 5×10 to 5×20, 5×40, or 5×80 did not bring any improvement. Simulations were conducted with coverage of 4% and 25%, with N ranging from 300 to 10 000, and using the overdispersion settings in Johnson et al. (2010).

5. Spotted hyena survey, Serengeti plains, Tanzania

We apply the striplet method to line-transect distance-sampling surveys of spotted hyena *(Crocuta crocuta)* on a 2428 km^2 region in the Serengeti plains of Tanzania (Durant et al., in review). The survey area is shown in Figure 5. The data were collected as part of two surveys of carnivores, the first during the wet season in May 2005, and the second during the dry season in October 2005. For each survey, $k = 29$ parallel north-south transects were placed at 2 km intervals. Sightings correspond to detected hyena groups which ranged in size from 1 to 10 individuals (mean 1.9). Hyena groups were observed to a distance of $w = 0.5$ km.

[Figure 5 about here.]

The spotted hyena data are interesting because the behavior of the animals changes between seasons. In the wet season, hyenas move from throughout the Serengeti to congregate in the short-grass plains in the east of the study area, where there is abundant food (Hofer and East, 1993). The study area holds large numbers of these non-territorial 'commuters'. In the dry season, the animals commute away from the study area, density is much lower, and the animals remaining in the study area are largely territorial residents.

These behavioral characteristics are reflected in the transect data shown in the top panels of Figure 6. The plots show the number of sightings on each transect, together with their fitted values from the GAM with 4 d.f. described in Section 2.6. In the wet season, the encounter rate is high and the data are clearly overdispersed relative to the Poisson GAM. The dispersion parameter from the fitted GAM is 3.8. This suggests there is aggregation or clustering of groups, perhaps due to the clustering of prey herds (Hofer and East, 1993). In the dry season the data do not appear overdispersed, and the dispersion parameter is 0.7. This is consistent with territorial behavior at low density.

Hyena density is estimated using `Distance 5.0` (Thomas et al., 2010). The wet and dry season surveys had respectively $n_{\text{wet}} = 186$ and $n_{\text{dry}} = 53$ detected groups, each over $L = 1199$ km of transects. The estimated densities of individuals are $\widehat{D}_{\text{wet}} = 1.00$ and $\widehat{D}_{\text{dry}} = 0.25$ km^{-2}. The detection functions are selected by AIC to be half-normal with two cosine adjustments in each case, yielding average within-strip detection probabilities of $\widehat{P}_{a,\text{wet}} = 0.26$ (CV=7.04%) and $\widehat{P}_{a,\text{dry}} = 0.27$ (CV=14.3%). The estimated mean hyena group sizes using size-bias regression are $\widehat{E}(S)_{\text{wet}} = 1.7$ (CV=5.04%) and $\widehat{E}(S)_{\text{dry}} = 1.5$ (CV=7.73%). The overall CV of $\widehat{D}$ for either survey is given by (4), with $\widehat{\mathrm{cv}}(n/L)$ gained from the striplet estimator or estimators $R2$ to $O2$. This reduces to $\widehat{\mathrm{cv}}(\widehat{D})_{\mathrm{wet}} \simeq \sqrt{\{\widehat{\mathrm{cv}}(n/L)^2_{\mathrm{wet}} + 75\}}$ and $\widehat{\mathrm{cv}}(\widehat{D})_{\mathrm{dry}} \simeq \sqrt{\left\{\widehat{\mathrm{cv}}(n/L)^2_{\mathrm{dry}} + 264\right\}}$.

For the wet season, the choice of estimator for $\widehat{\mathrm{cv}}(n/L)$ has a dramatic effect on the estimate. The random-line estimators $R2$ and $R3$ yield 17.0% and 17.6% for $\widehat{\mathrm{cv}}(n/L)$. Each of the stratified estimators $S1$, $S2$, $O1$, and $O2$ give estimated CVs of about 14.5%. The striplet estimator gives an estimated CV of only 6.9%. The standard error estimated by the striplet method is therefore less than half of that estimated by any other method. We verify this surprising result by simulation below. The overall $\widehat{\mathrm{cv}}(\widehat{D})_{\mathrm{wet}}$ is 20% using $R3$, as reported by `Distance 5.0`; 17% using each of the stratified estimators $S1$ to $O2$, and 11% using the striplet estimator. The estimated number of individuals, $\widehat{N}_{\mathrm{ind}}$, is 2430, with confidence interval width over 1900 for $R2$ and $R3$, about 1700 using each of $S1$ to $O2$, and 1060 for the striplet estimator. The choice of encounter rate variance estimator therefore has a profound impact on all reported measures of uncertainty.

For the dry season, the choice of estimator is less important. The random-line estimators $R2$ and $R3$ give a $\widehat{\mathrm{cv}}(n/L)$ of 16% and 15%, the stratified estimators $S1$ and $S2$ give 10%, and $O1$ and $O2$ give 13%. The striplet estimator also gives 13%. Estimates of $\widehat{\mathrm{cv}}(\widehat{D})_{\mathrm{dry}}$ range from 19.3% ($O1$ and $O2$) to 22.7% ($R2$), with the striplet estimator in the middle with 20.7%. Finally, $\widehat{N}_{\mathrm{ind}} = 619$ with confidence interval width ranging from about 480 ($S1$ and $S2$), through 510 (striplet), to 570 ($R2$).

The striplet estimator is not sensitive to the degrees of freedom used in the underlying GAM. As the d.f. is changed from 1 to 26, the estimated encounter rate CV varies only between 6.8% and 7.1% for the wet season, and between 12.8% and 12.9% for the dry season. The striplet method was implemented using $B = 69$ grid start-points and $J = 2006$ striplets, so that adjacent systematic grids were displaced from each other by about 30m.

To verify that these results are credible, we use the `DSpat` package to simulate populations on the Serengeti region with characteristics as close as possible to those in the observed data. Example simulations are shown in Figure 5. Each simulated point corresponds to a group of hyenas. Points are generated by the log-Gaussian Cox process described by Johnson et al. (2010). The study area is divided into two habitats: long-grass plains and short-grass plains. The total number of points is fixed at the estimated number of groups in the real data, namely 1428 for the wet season and 400 for the dry season. The relative intensities of the point process in the two habitats are chosen such that the mean numbers of detections per habitat are approximately the same as the observed numbers in the real data.

For the wet season, spatial correlation in the Gaussian random field underlying the Cox process is chosen so that the mean dispersion parameter from GAMs fitted to the simulated data is 3.8, equal to that in the real data. The resulting strong level of aggregation is seen in the two realizations in Figure 5. The aggregations change location between simulations, which could mimic hyenas following herds of prey which aggregate and disperse. For the dry season, the dispersion parameter of 0.7 is well within the distribution of dispersion parameters generated when the Gaussian random field is uncorrelated. Therefore no correlation is used, and the dry-season simulations reproduce an inhomogeneous Poisson process without aggregation. Distance-sampling data are simulated for both surveys with mean detection probability matching the estimated values from the real surveys. The detection function is estimated from the simulated detections in each simulation.

[Figure 6 about here.]

The simulation results for $\widehat{\text{cv}}(n/L)$ are shown in the bottom panels of Figure 6. All results from the real surveys, shown as horizontal lines marked with arrows on the plots, fall within the distributions expected from the simulations. In particular, the simulations validate the surprisingly large gain in CV by using the striplet method with the wet-season data. For both seasons, both the real-data striplet result and the mean of the striplet results for simulated data are very close to the correct $\text{cv}(n/L)$ gained from the simulated sample of n/L values. The striplet estimator is by far the best in terms of both bias and precision, especially for the wet season, for which the real-data gains over the stratified estimators are close to the scenario average. The differences between estimators for the real dry-season data are within sampling variability. Simulation results for $\widehat{\text{se}}(n/L)$, $\widehat{\text{se}}(\widehat{D})$, and $\widehat{\text{cv}}(\widehat{D})$ closely mirror those for $\widehat{\text{cv}}(n/L)$ shown in Figure 6.

Repeating the wet-season simulations with different levels of spatial correlation in the Gaussian random field gives results similar to those in Section 4.2. As the aggregation level is varied from none, through moderate, to the high level shown in Figure 6, the correct encounter rate CV calculated from the simulations changes only from 6.9% to 7.5%. The striplet mean CV mirrors these slight changes. However, the performance of all other estimators from $R2$ to $O2$ changes substantially. The four stratified estimators $S1$ to $O2$ are almost unbiased when there is no aggregation. As aggregation gets stronger, the estimator performance deteriorates, with most of the estimates being too high.

6. Discussion

We have assessed the striplet variance estimator for systematic designs in a wide range of standard and extreme spatial situations, and its performance has been uniformly good except for some cases of extreme aggregation and small numbers of lines. In bias and precision it consistently outperforms all other estimators from Fewster et al. (2009). The estimator is very robust to overdispersion for one-dimensional transect designs, although it is less robust for two-dimensional quadrat sampling. Extension to other spatial designs such as zigzag schemes (Strindberg and Buckland, 2004) is the topic of future work.

In this paper we have used the repeated-survey framework with constant N. Changing the framework to incorporate variable N might be desirable if the survey area has a permeable boundary so that N changes as animals move in and out of the area, but it introduces problems of model misspecification. The overdispersed process considered by Johnson et al. (2010) yields very high variance in N: e.g. $\text{var}(N) \simeq 10E(N)$ when $E(N) = 1000$. Such a situation might occur if the survey area is a subregion of a larger region, with free movement across the boundary and heavy clustering. However, in this situation it would be unwise to restrict surveys to the subregion, unless it is believed that the survey data from the subregion somehow contain information about the variance of N. The repeated-survey framework with constant N restricts variance estimation to a simpler scenario, but one that is well-defined and for which we now have a robust estimator. It extends the traditional design framework by incorporating variance due to animal movement between survey replicates.

In accordance with the simulation results in Section 4.2, the wet-season spotted hyena survey did not present difficulties for the striplet estimator, despite being strongly overdispersed. It did, however, create substantial difficulties for all other estimators, resulting in an important gain in reported precision when using the striplet estimator. In fact, performance of the striplet estimator is quite similar to performance of a naive Poisson-based estimator $\widehat{\text{var}}(n/L) = n/L^2$ for the Serengeti region. However, the naive estimator performs very poorly for the triangle-based simulations, so it cannot be recommended in practice. Variability in $L(b)$ and highly trended populations present problems for the naive Poisson estimator, whereas aggregated populations present problems for estimators $R2$ to $O2$ and the model-based procedure of Johnson et al. (2010). The striplet estimator performs well in all of these situations, unless it is subjected to extreme aggregation coupled with few transect lines.

Acknowledgements

I am indebted to Meggan Craft and Sarah Durant for providing the spotted hyena data, and to numerous staff who conducted the surveys. The surveys were funded by Frankfurt Zoological Society and the Wildlife Conservation Society, and were supported by Serengeti

National Park Authority and workers from seven Serengeti carnivore projects. Thanks to Tanzania National Parks, the Ngorongoro Conservation Area Authority, Tanzania Wildlife Research Institute, and to Len Thomas for suggesting the data. Many thanks to the associate editor and two referees whose constructive comments greatly improved the manuscript.

References

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L. and Thomas, L. (2001). *Introduction to Distance Sampling.* Oxford: Oxford University Press.

Cochran, W. G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *Annals of Mathematical Statistics* **17,** 164–177.

D'Orazio, M. (2003). Estimating the variance of the sample mean in 2-dimensional systematic sampling. *Journal of Agricultural, Biological, and Environmental Statistics* **8,** 280–295.

Durant, S. M., Craft, M., Hilborn, R., Bashir, S., Hando, J., and Thomas, L. (in review). Long term trends in carnivore abundance using distance sampling in Serengeti National Park, Tanzania.

Fewster, R. M., Buckland, S. T., Burnham, K. P., Borchers, D. L., Jupp, P. E., Laake, J. L., and Thomas, L. (2009). Estimating the encounter rate variance in distance sampling. *Biometrics* **65,** 225–236.

Fewster, R. M. and Jupp, P. E. (2009). Inference on population size in binomial detectability models. *Biometrika* **96,** 805–820.

Fewster, R. M. and Buckland, S. T. (2004). Assessment of distance sampling estimators. In *Advanced Distance Sampling,* S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers, and L. Thomas (eds), pp. 281-306. Oxford: Oxford University Press.

Gregoire, T. G. (1998). Design-based and model-based inference in survey sampling: appreciating the difference. *Canadian Journal of Forest Research* **28,** 1429–1447.

Hastie T. J. and Tibshirani, R. J. (1990). *Generalized additive models.* London: Chapman and Hall.

Hofer H. and East M. L. (1993). The commuting system of Serengeti spotted hyaenas: how a predator copes with migratory prey. II. Intrusion pressure and commuters' space use. *Animal Behaviour* **46,** 559–574.

Johnson, D. S., Laake, J. L., and Ver Hoef, J. M. (2010). A model-based approach for making ecological inference from distance sampling data. *Biometrics* **66,** 310–318.

Kingsley, M. C. S. (2000). Numbers and distribution of beluga whales, *Delphinapterus leucas*, in James Bay, eastern Hudson Bay, and Ungava Bay in Canada during the summer of 1993. *Fishery Bulletin* **98,** 736–747.

Millar, R. B. and Olsen, D. (1995). Abundance of large toheroa (*Paphies ventricosa* Gray) at Oreti Beach, 1971–90, estimated from two-dimensional systematic samples. *New Zealand Journal of Marine and Freshwater Research* **29,** 93–99.

R Development Core Team. (2008). R: *A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3900051-070.

Simmonds, E. J. and Fryer, R. J. (1996). Which are better, random or systematic acoustic surveys? A simulation using North Sea herring as an example. *ICES Journal of Marine Science* **53,** 39–50.

Strindberg, S. and Buckland, S. T. (2004). Zigzag survey designs in line transect sampling. *Journal of Agricultural, Biological and Environmental Statistics* **9,** 443–461.

Thomas, L., Buckland, S. T., Rexstad, E. A., Laake, J. L., Strindberg, S., Hedley, S. L., Bishop, J. R. B., Marques, T. A., and Burnham, K. P. (2010). Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* **47,** 5–14.

Wolter, K. M. (1984). An investigation of some estimators of variance for systematic sampling. *Journal of the American Statistical Association* **79,** 781–790.

Wolter, K. M. (1985). *Introduction to Variance Estimation.* New York: Springer-Verlag.

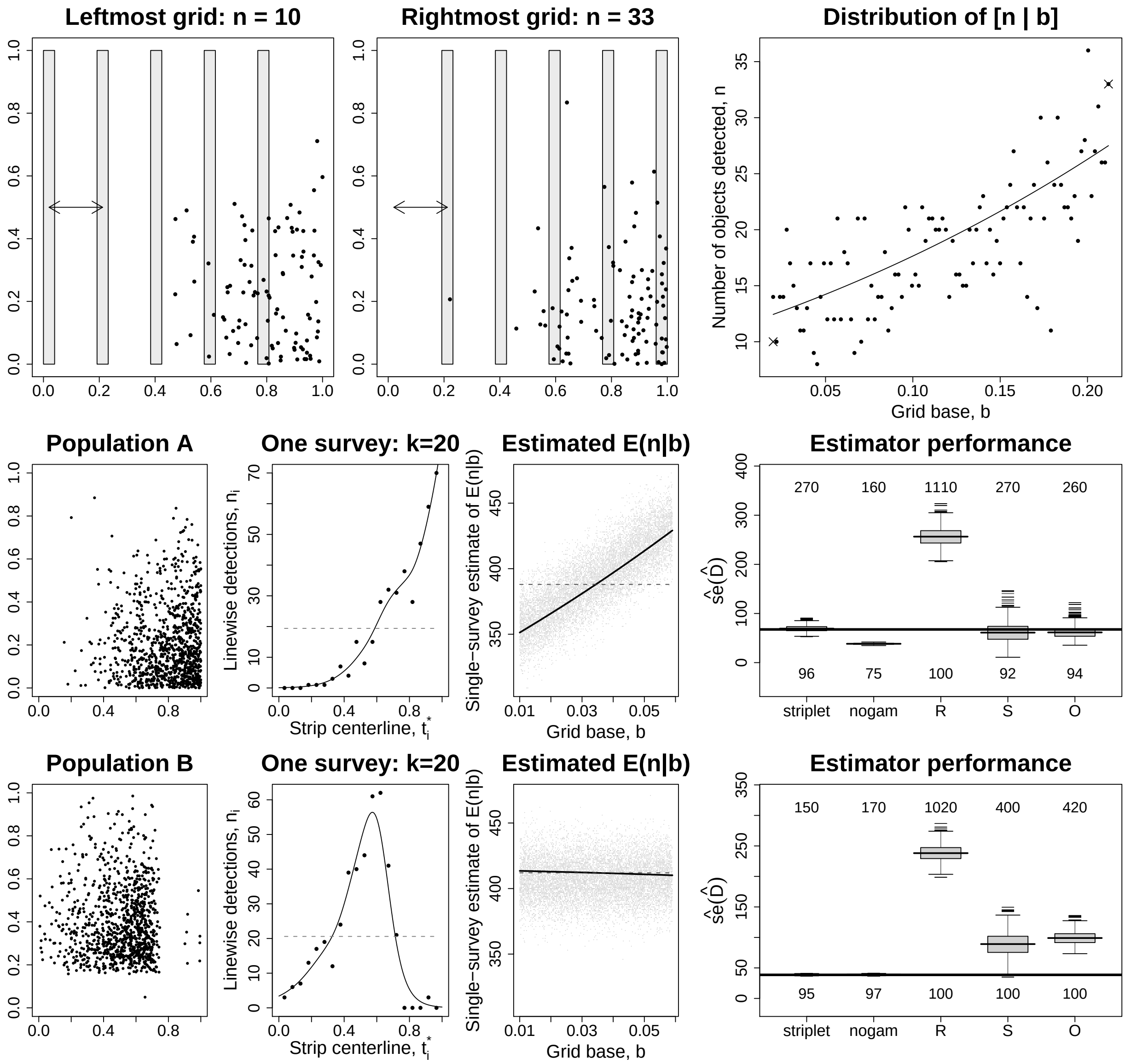

Figure 1. Illustration of the repeated-survey framework and striplet variance estimator. The top row shows two realizations of the repeated-survey framework with $k = 5$ lines, $w = 0.02$ and $N = 100$. Object positions and grid location change for each realization. Detection is perfect within the shaded strips. Arrows mark the allowable values for the grid base, b, from 0.02 to 0.212. The distribution of total detections, n, is shown as a function of grid base $b \in [0.02, 0.212]$ in the third panel, with the two featured realizations marked by crosses and the true mean for this population shown by the line. The lower two rows show the components of the striplet estimator for two highly trended populations, using $k = 20$, $w = 0.01$, and $D = N = 1000$. Panel 2 on each row shows the GAM fit to stripwise detections n_i from a single survey. This GAM is used to construct the estimated $E(n \,|\, b)$ as shown in the bold lines on panel 3. Dashed lines across panels 2 and 3 show a constant-mean fit, termed `nogam`. Also shown as clouds on panel 3 are the true distributions of $[n \,|\, b]$ from 10,000 simulations. Panel 4 shows estimator performance for $\widehat{\text{se}}(\widehat{D})$, with the true $\text{se}(\widehat{D})$ from 10,000 simulations given as a bold line across the plot. Numbers above and below the boxplots show respectively mean width and coverage of nominal 95% confidence intervals. For Population A, $\text{var}_b \{E(n \,|\, b)\}$ is high (seen by the sharp slope in panel 3) and the `nogam` estimator is poor. For Population B, $\text{var}_b \{E(n \,|\, b)\}$ is negligible (flat in panel 3) and both `striplet` and `nogam` are extremely close to the true $\text{se}(\widehat{D})$.

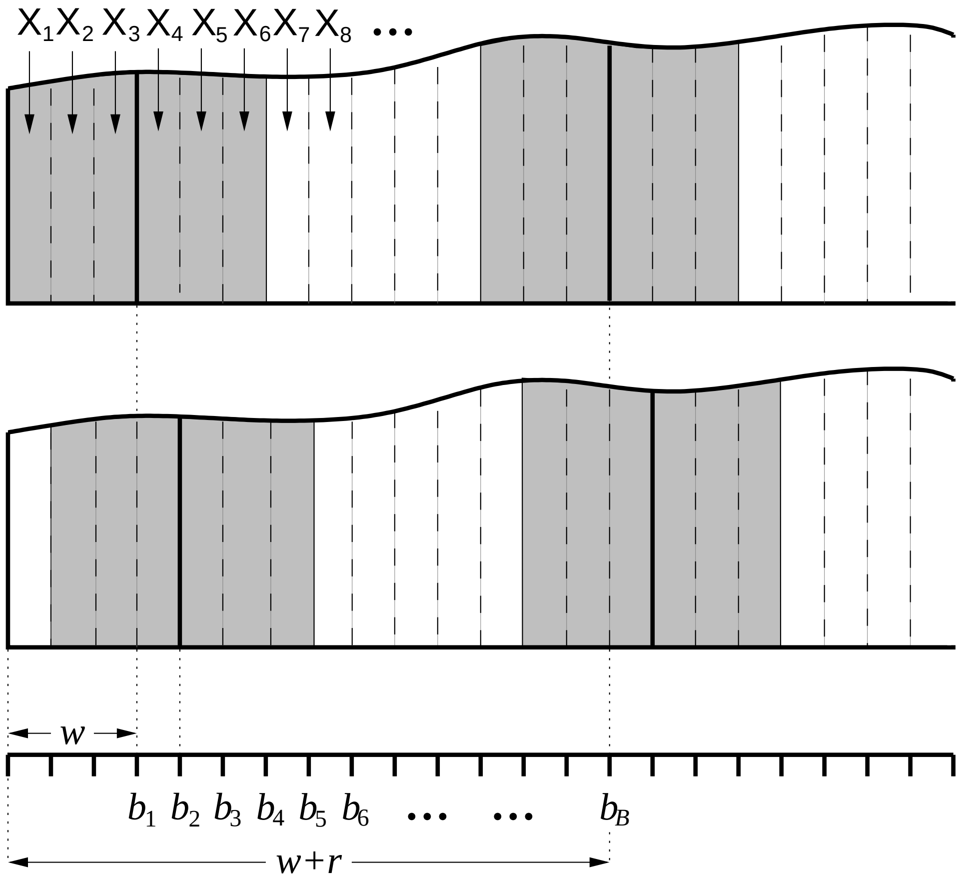

Figure 2. Striplet concept. Search strips in the systematic design are marked by shading. Each search strip is covered by many striplets, with striplet boundaries marked by dashed lines. Transect lines are marked in bold and have spacing r. The uniform sampling frame $b \sim \text{Uniform}[w, w + r)$ is approximated by a discrete sampling frame in which the possible start values for the systematic grids are $b_1 = w, b_2, \ldots, b_B$, where $b_B \leqslant w + r$. The two rows show the first and second grids, with $b = b_1$ and $b = b_2$ respectively. The random variables $X_1, X_2, \ldots, X_J$ denote the number of objects available for detection in each striplet for a particular replicate of the repeated-survey framework.

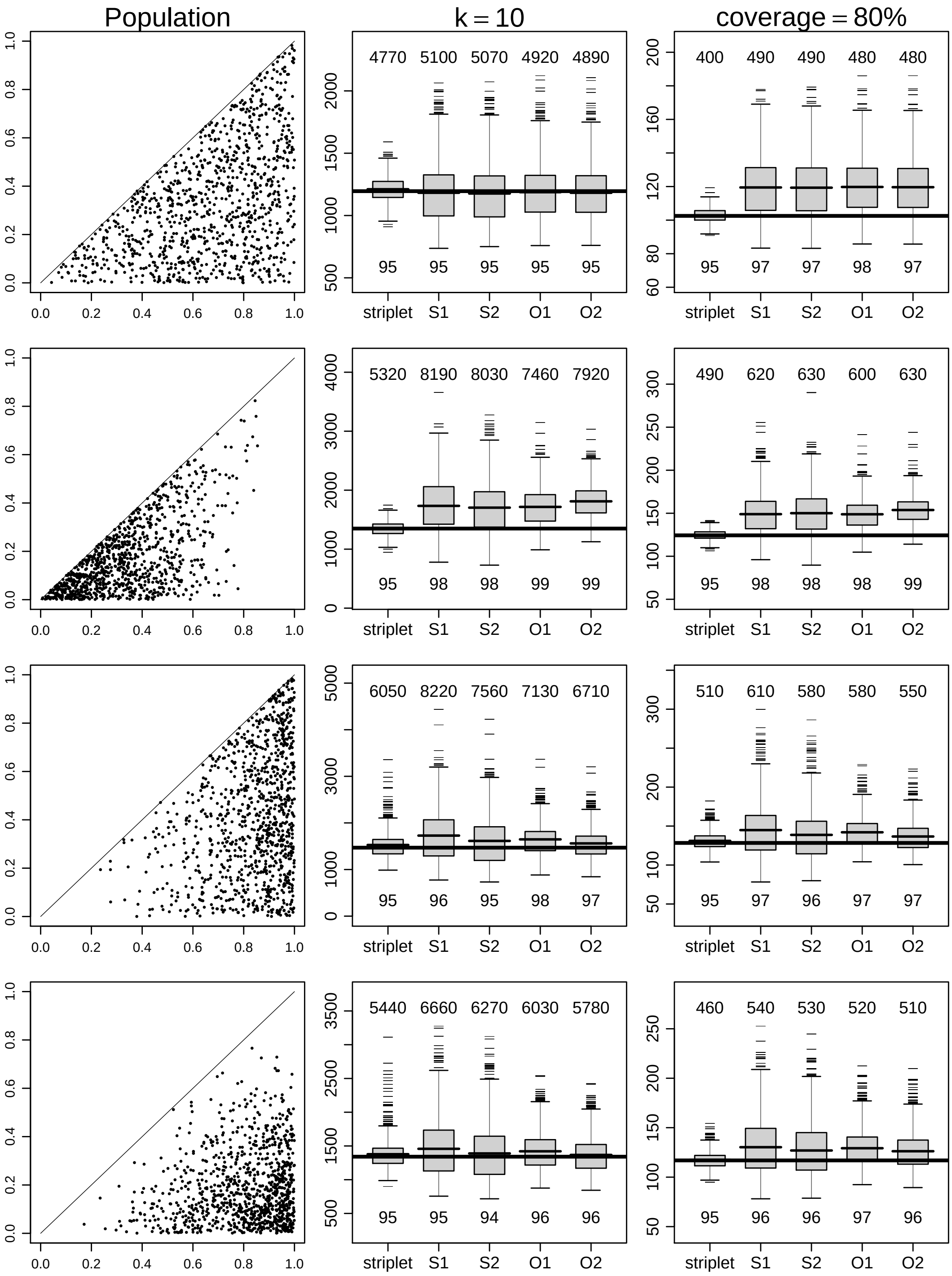

Figure 3. Triangle populations 1, 2, 3, and 6 from Fewster et al. (2009), and distributions of $\widehat{\mathrm{se}}(\widehat{D})$. Each boxplot shows $\widehat{\mathrm{se}}(\widehat{D})$ from 1000 simulations. Thin horizontal lines on the boxes show the estimator means; numbers above and below the boxplots show mean confidence interval width and coverage respectively. The bold line across the plot shows the correct $\mathrm{se}(\widehat{D})$, each from 10 000 simulations. Boxes are drawn between the upper and lower quartiles, and whiskers extend to the last observation within 1.5 times the interquartile range from the quartiles. The two simulation scenarios are: (i) $k = 10$, coverage = 4%, $E(n)$=120, $D =$ 10 000, $w = 0.002$, $P_a = 0.6$; (ii) coverage = 80%, $k = 20$, $E(n) = 480$, $D = 2000$, $w = 0.02$, $P_a = 0.6$. The first column shows population realizations with true density $D = 2000$.

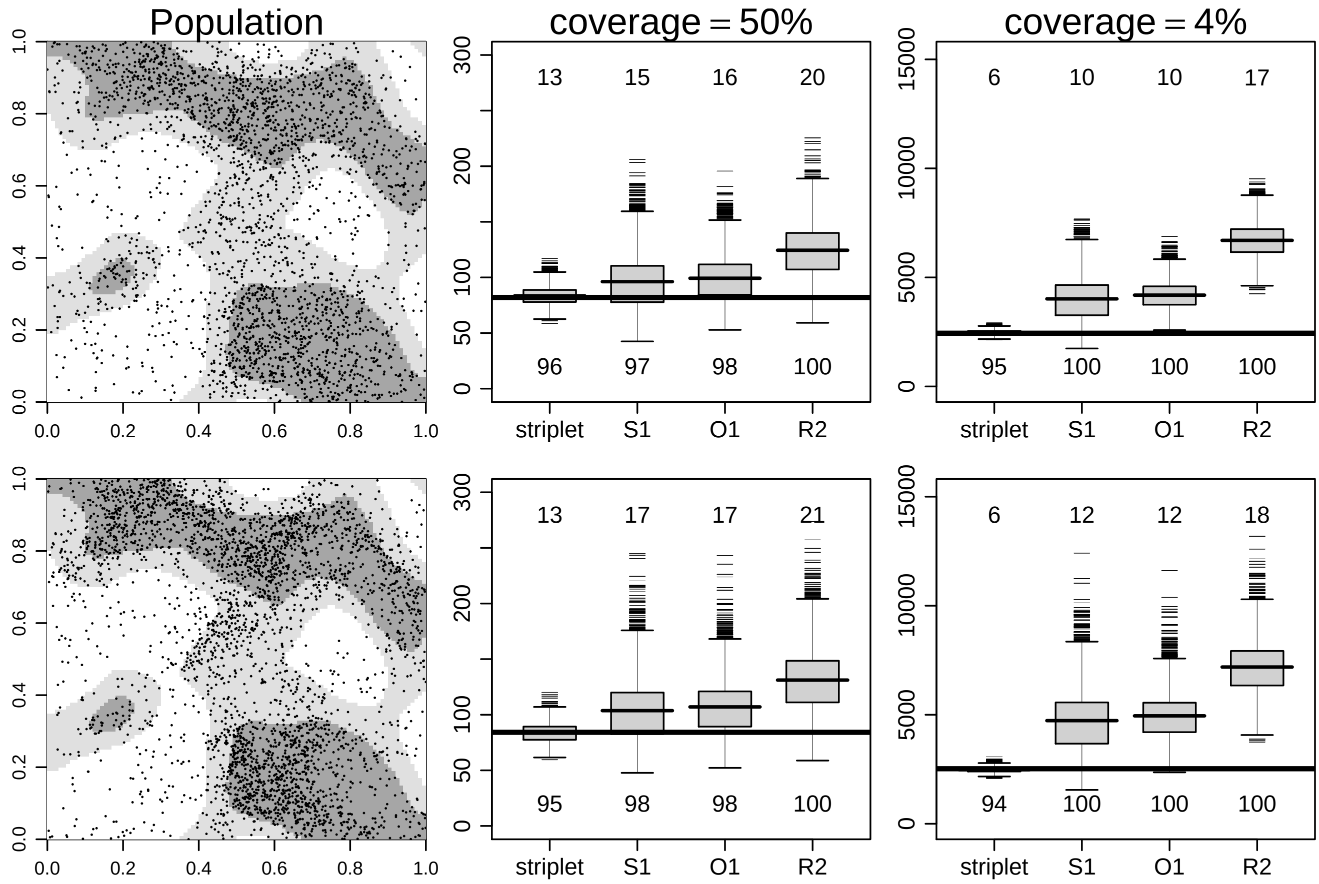

Figure 4. Simulated populations similar to those in Johnson et al. (2010), and the resulting distributions of $\widehat{\text{se}}(\widehat{D})$ using 3000 simulations for each scenario. The numbers above the boxplots are the mean %CVs reported from each estimator. The numbers below the boxplots are confidence interval coverages for nominal 95% confidence intervals. The first row shows an inhomogeneous Poisson process without overdispersion and associated results, and the second row shows results with high overdispersion. The first column shows example point processes using $N = 2000$ points. Shading indicates habitat type, with darker shading associated with higher object density. The two simulation scenarios are: (i) coverage = 50%, $E(n) = 75$, $k = 10$, $D = N = 600$, $w = 0.025$, $P_a = 0.25$; (ii) coverage = 4%, $E(n) = 400$, $k = 10$, $D = N = 40\,000$, $w = 0.002$, $P_a = 0.25$.

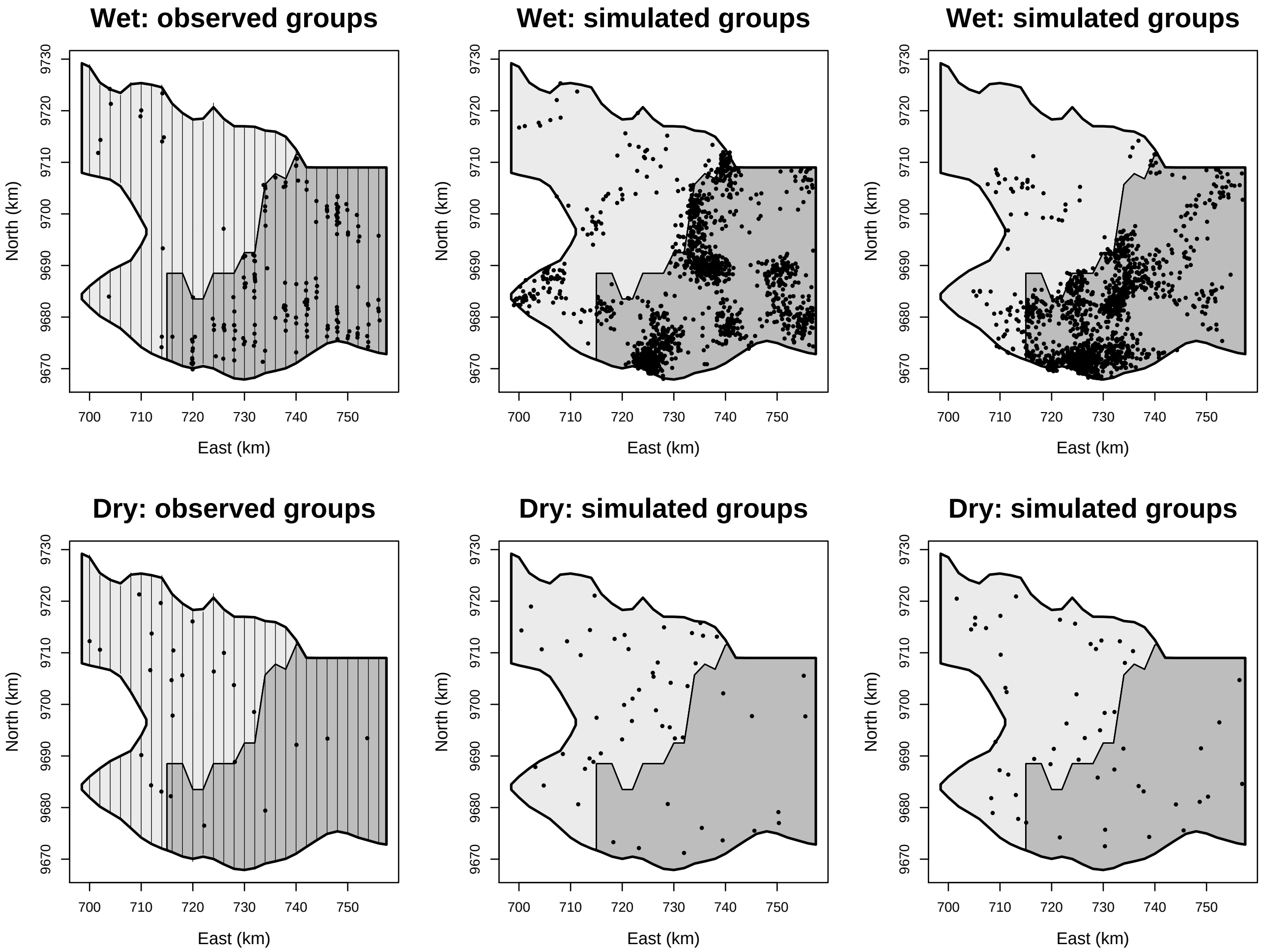

Figure 5. Serengeti survey region. Light shading in the west denotes long-grass plains, and dark shading in the east denotes short-grass plains. The rows denote wet and dry seasons. The first panels show locations of transects and observed hyena groups. East-west coordinates are observed, while north-south coordinates are generated at random from the selected clustering model. The second and third panels show two random realizations of groups as used in the simulation study. The wet-season simulations are heavily clustered, while the dry-season simulations are not clustered.

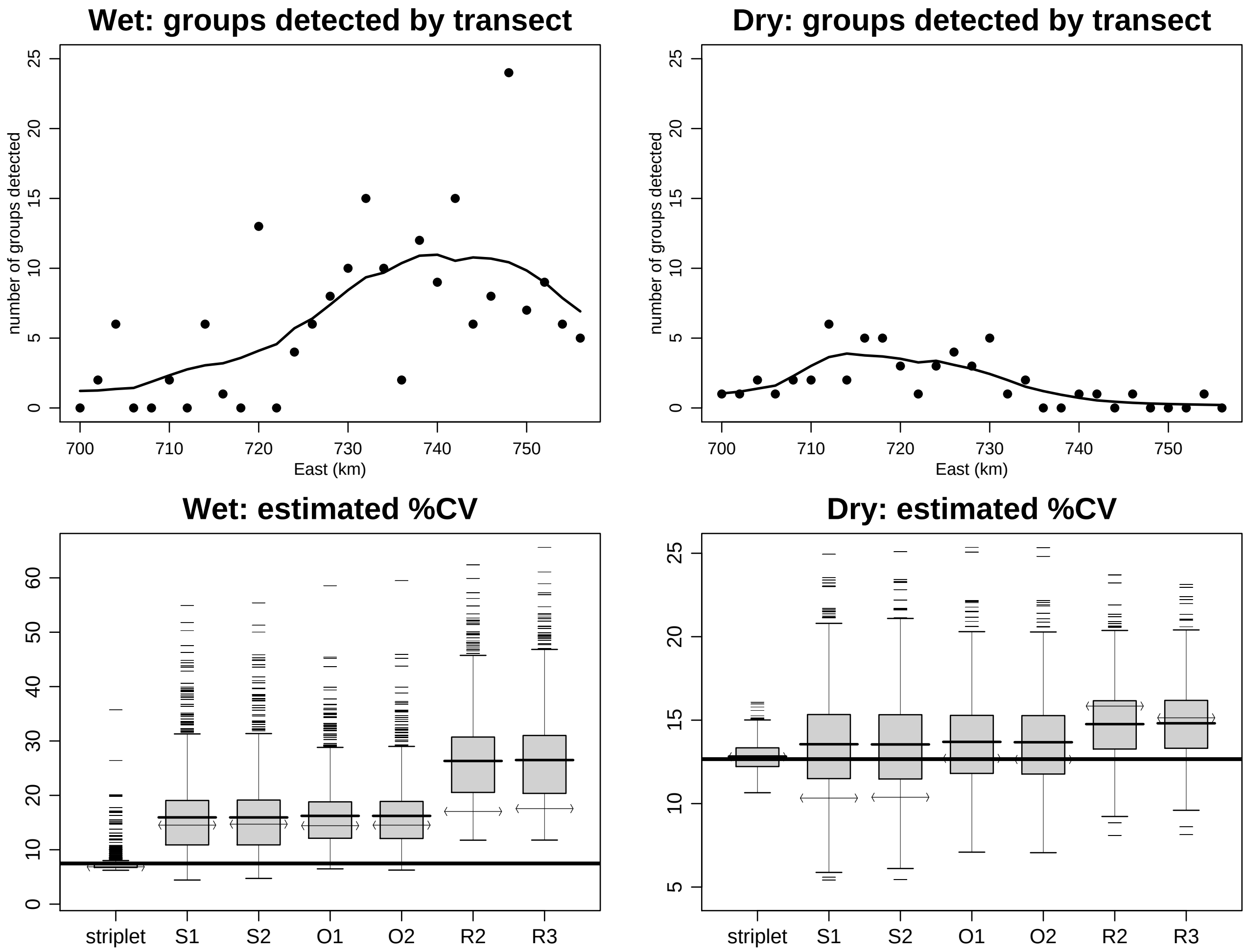

Figure 6. Model fit and simulations for the Serengeti data. The top panels show the observed number of detections on each transect, n_i for $i = 1, \dots, 29$. The solid lines show the fitted values $E\{n(t)\}$ from a GAM with 4 d.f. They are not completely smooth because the area of the surveyed strips changes non-smoothly between transects. The bottom panels show the distributions of 1000 estimated %CVs for encounter rate based on the simulations described in the text. Boxplot details are the same as before. The thin lines marked with arrows across each boxplot give the estimated %CV from the real data. The bold horizontal lines crossing the whole plot give the correct %CV as obtained from the 1000 simulations.