Effect of spatial overdispersion on confidence intervals for population density estimated by spatial capture–recapture

Murray G. Efford[1] and David Fletcher[2]

1. Department of Mathematics and Statistics, University of Otago, P.O. Box 56, Dunedin, New Zealand.

2. David Fletcher Consulting Limited, 67 Stornoway Street, Karitane 9471, New Zealand.

Corresponding author: Murray Efford murray.efford@otago.ac.nz

Keywords: activity centers; aggregation; Cox process; density estimation; inhomogeneous Poisson process; LGCP; overdispersion; SECR; spatial capture–recapture

Abstract

Spatially explicit capture–recapture models are used widely to estimate the density of animal populations. The population is represented by an inhomogeneous Poisson process, where each point is the activity center of an individual and density corresponds to the intensity surface. Estimates of average density are robust to unmodeled inhomogeneity, but the coverage of confidence intervals is poor when the intensity surface is stochastic. Poor coverage is due to overdispersion of the number of detected individuals n with respect to the fitted Poisson distribution. We investigated overdispersion from stochastic generating models (log-Gaussian Cox process and Thomas cluster process). Variation in a scalar measure of local density – the detection-weighted mean density – predicts overdispersion when the generating process is known. A previously proposed correction for overdispersion was successful only in limited cases: rigorous correction for spatial overdispersion requires prior knowledge of the generating process. The problem is lessened by assuming population size to be fixed, but this assumption cannot be justified for common study designs.

Introduction

Spatially explicit capture–recapture (SECR) is a method for estimating the density of an animal population by modeling observations of marked individuals at detectors in known locations. The state model is the spatial distribution of animal activity centers (AC). Activity centers are not observed directly, but are modeled as a realization of a spatial point process. Capture–recapture sampling (the observation process) provides spatial data on a sample of animals. Parameters of

the state model (specifically, the intensity surface of the point process) are estimated jointly with parameters of the spatial observation process by maximizing the marginal likelihood (Borchers & Efford, 2008) or in a Bayesian framework by data augmentation (Royle et al., 2014) or using the semi-complete data likelihood (Zhang et al., 2023). The inhomogeneous Poisson process (IHPP) is a useful general model for the intensity surface (Borchers & Efford, 2008).

The IHPP intensity surface may be modeled as a function of spatial covariates, but covariates are commonly lacking or inadequate. Fortunately, 'average' density estimated by fitting a homogeneous model is a useful population summary even when the true intensity surface is not flat. Average density estimated in this way generally shows little bias, as we confirm later by simulation, although bias may result from correlated variation in density and detection parameters (McLellan et al., 2023).

The intensity surface of the IHPP for AC may vary among realizations, i.e. be stochastic rather than constant. Stochastic-intensity IHPP are also known as Cox processes, for which there is extensive theory (e.g., Møller & Waagepetersen, 2004; Chiu et al., 2013). Possible biological reasons for stochasticity are numerous. For example, the distribution of AC may bear the imprint of stochastic demography, random aggregation of migratory prey, harvesting, local catastrophes or colonisation processes.

Animals may also aggregate in clusters that are spatially unmoored rather than tied to persistent landscape features. The mathematical analog is a cluster process. In the widely used Thomas cluster process, conceptual 'parent' locations are Poisson-distributed in the plane and each parent gives rise to a Poisson number of 'offspring' scattered about the parent according to a circular bivariate normal distribution (examples in Appendix S1). An AC distribution that results from a

cluster process cannot readily be distinguished from one that is the result of an IHPP (e.g., Diggle et al., 2013), and each realized distribution of parents corresponds to a unique IHPP intensity surface (Møller & Waagepetersen, 2004 p. 61).

Bischof et al. (2020) drew attention to poor coverage of confidence intervals for SECR estimates of average density when animals are socially aggregated in family groups or packs. Coverage improved when they applied a variance inflation factor based on the number of individuals per detector. The close connection between cluster processes and IHPP led us to investigate more generally the effect of AC overdispersion in SECR, and the utility of the variance inflation factor. Bischof et al. (2020) also addressed spatial overdispersion in the detection process when group members are detected together ('cohesion'), but we consider this only in passing.

Various approaches have been used to generate non-uniform densities for the evaluation of SECR estimators and study designs. As examples of a constant-intensity IHPP, Efford et al. (2009) specified a simple east–west gradient in density and Durbach et al. (2021) generated non-uniform densities as a log-linear function of a terrain covariate from a real landscape. Stochastic-intensity IHPP were simulated by Efford and Fewster (2013) using quasi-realistic maps of habitat interspersed with non-habitat, and by Paterson et al. (2019) and Dupont et al. (2021) who specified the log-intensity surface as a Gaussian random field. Gaussian random fields underlie the log-Gaussian Cox process (LGCP) which is theoretically elegant and has wide application as a point process model (Møller et al., 1998; Diggle et al., 2013). Data from Thomas cluster processes were simulated for SECR by Efford et al. (2009) and Clark (2019), using the more general name Neyman-Scott. Bischof et al. (2020) simulated a simplified cluster distribution in which they fixed the number of clusters in a defined area, assigned each cluster a fixed number of

AC, and allowed no within-cluster dispersion (AC within a cluster shared the same location). Confidence intervals for average density depend on the reference distribution of possible AC realizations. The key distinction is between a reference distribution involving realizations of an IHPP with constant intensity surface, and a reference distribution in which the IHPP intensity surface is stochastic. Which is appropriate depends on the biology and purpose of the study and has major implications for inference, as we aim to elucidate.

Background

For specificity we focus on closed SECR models and observations at point detectors such as automatic cameras. We use $\boldsymbol{\omega}$ for the set of n observed spatial detection histories, and ϕ and θ for the parameter vectors of the state and observation models. The likelihood has the general form $L(\phi, \theta|\boldsymbol{\omega}) = \Pr(n|\phi, \theta)\Pr(\boldsymbol{\omega}|n, \phi, \theta)$ (Borchers & Efford, 2008). The intensity surface of the IHPP for AC is a function of location, $D(\mathbf{x}; \phi)$, where $\mathbf{x}$ represents the Cartesian coordinates x, y. It can make sense to fit a homogeneous density surface despite the potential for heterogeneity: we may be concerned with the average density or the total number of individuals, or any structure may be ephemeral, as in the case of clustering. We denote the estimate of density from a homogeneous model by $\hat{D}$ (without dependence on $\mathbf{x}$). Density may be estimated from a homogeneous model in two stages, first estimating θ by maximizing the second factor in the likelihood conditional on n, and then computing the Horvitz-Thompson-like estimator $\hat{D} = n/a(\hat{\theta})$ (Borchers & Efford, 2008). The effective sampling area a is a function of the detection parameters $a(\theta) = \int p_{\cdot}(\mathbf{x}; \theta)\, d\mathbf{x}$, where $p_{\cdot}(\mathbf{x}; \theta)$ models the overall probability of detection for an AC at $\mathbf{x}$ (see Appendix S5 for details) and integration is over all habitat in the plane.

We note that n is an analogue of the encounter rate in conventional distance sampling (e.g., Buckland et al., 1993). As in distance sampling, the variance of $\hat{D}$ may be approximated by the delta method as

$$\widehat{\text{var}}(\hat{D}) = \hat{D}^2 \left\{ \frac{\widehat{\text{var}}(n)}{n^2} + \frac{\widehat{\text{var}}[a(\hat{\theta})]}{a(\hat{\theta})^2} \right\}. \tag{1}$$

Maximization of the conditional likelihood provides an estimate of $\text{var}[a(\hat{\theta})]$ (Appendix S4), and this component is often small. If n is Poisson then the first term simplifies to $n/a(\hat{\theta})^2$. However, to assume wrongly that n is Poisson risks biasing estimates of $\text{var}(\hat{D})$.

Constant-intensity IHPP

For an intensity surface that is a deterministic function of constant spatial covariates or trend, the expected number of AC in a region A is $\text{E}[N(A)] = \int_A D(\mathbf{x};\phi)\, d\mathbf{x}$, and from the properties of an IHPP, $N(A)$ is Poisson with variance $\text{E}[N(A)]$ (e.g., Diggle, 2003 p. 67). The spatial detection process applies location-dependent thinning to the distribution of activity centers. Thinning is independent for the detector types we consider, but not for single-catch traps. The expected number of animals detected n is an integral over space of the product of density and a location-dependent detection function, i.e. $\text{E}(n) = \int_A D(\mathbf{x};\phi)\, p_{\cdot}(\mathbf{x};\theta)\, d\mathbf{x}$.

The probability of detection is assumed to approach zero at large distances from the detectors. Consider an area A that includes the detectors and all habitat within a buffer of width w. Then $\text{E}(n)$ approaches a constant as w increases, and beyond some w' further change in $\text{E}(n)$ is negligible. This is convenient for computation because numerical integration may be restricted to an arbitrary area A_w for which density estimates will be nearly unbiased given $w > w'$.

If AC are distributed as an inhomogeneous Poisson process and detections are independent, then

the locations of detected AC are also inhomogeneous Poisson, by the properties of independently thinned point processes (Chiu et al., 2013). It follows that the number of animals detected is Poisson, with variance $\mathrm{var}(n) = \mathrm{E}(n)$ (Borchers & Efford, 2008). Thus inhomogeneity of density does not in itself imply overdispersion.

Stochastic IHPP

The intensity of a Cox process (an IHPP with stochastic intensity) is a random variable. From here on we use $D(\mathbf{x})$ to mean this variable, and omit the parameter vector to lighten the notation. The specification of $D(\mathbf{x})$ is flexible and requires parameters additional to ϕ that we consider only in passing because, while they are used in simulation, they cannot easily be estimated from SECR data. For a log-Gaussian Cox process with flat expected-density surface and parameters μ and σ^2, $\mathrm{E}[D(\mathbf{x})] = \exp(\mu + \frac{\sigma^2}{2})$ regardless of the spatial scale of covariance (a third LGCP parameter). Although stochastic intensity surfaces may include a deterministic trend (e.g., Johnson et al., 2010), for simplicity we assume the process is stationary, and in particular the expected-density surface is flat, so that we can use the simpler notation $\mu_D \equiv \mathrm{E}[D(\mathbf{x})]$.

Local and global density

We show in Appendix S6 that $\mathrm{var}(n)$ for a stochastic-intensity IHPP may be written as:

$$\mathrm{var}(n) = \mathrm{E}(n) + a(\theta)^2\,\mathrm{var}(D_a), \tag{2}$$

where $\mathrm{E}(n) = a(\theta)\mathrm{E}(D_a)$ and D_a is a scalar random variable representing the realized density in the vicinity of detectors, $D_a \equiv \int D(\mathbf{x})\, p_{\cdot}(\mathbf{x};\theta)\, d\mathbf{x} / \int p_{\cdot}(\mathbf{x};\theta)\, d\mathbf{x}$. The second term in Equation (2) indicates the amount of overdispersion in n that is caused by stochasticity in the intensity of the

IHPP. The subscript a indicates that D_a is an effective density, by analogy with the effective sampling area. D_a for known θ may be computed directly from a simulated intensity surface. It is constant for repeated simulations from the same realized intensity surface $D(\mathbf{x})$, but differs among realizations. The magnitude of $\mathrm{var}(D_a)$ depends on both stochastic variation in intensity and the spatial scale of sampling, which in turn depends on the distribution of detectors, the duration of sampling, and parameter θ.

As a convenient shorthand, we use the term 'local density' for a realization of D_a and 'global density' for μ_D. As μ_D is independent of location, it can be shown that $\mathrm{E}(D_a) = \mu_D$.

Spatial overdispersion

Underestimation of the sampling variance of $\hat{D}$ results directly from overdispersion of n relative to the fitted model. We define the amount of overdispersion in n by the ratio $c_n = \mathrm{var}(n)/\mathrm{var}_0(n)$ using $\mathrm{var}_0(n)$ to denote the variance under the fitted model. We assume a Poisson model for n except where otherwise stated. Fixed population size $N(A_w)$ leads to a binomial model for n that we consider at some length in Appendix S3.

More reliable confidence intervals might in principle be obtained by expanding the fitted model to include variance due to the generating process, presumably at the expense of estimating one or more extra parameters. However, the generating process is usually unknown, and we would stand little chance of fitting it with data from a single realization even if the process were known.

A model-free approach to overdispersion has entered capture–recapture studies (Burnham et al. 1987, pp. 243–246; Lebreton et al., 1992; Anderson et al., 1994; White & Cooch, 2017) from generalized linear modeling (Wedderburn, 1974). An empirical estimate of the overdispersion

ratio, based on a goodness-of-fit statistic, may be used to adjust variances and achieve near-nominal coverage of confidence intervals. The focus for non-spatial capture–recapture is on non-independence and heterogeneity among animals, for which the observed and expected frequencies of capture histories provide the required empirical index (the quantity $\hat{c}$). For spatial overdispersion we require a specifically spatial measure.

Johnson et al. (2010) constructed a measure of spatial overdispersion from the observed and expected numbers of detections on distance-sampling transects, but coverage of adjusted confidence intervals for density remained poor in some scenarios. Bischof et al. (2020) formed a similar measure from the observed (n_k) and expected ($\mathrm{E}(n_k)$) numbers of SECR detections per detector k:

$$\hat{c} = \frac{1}{1+\bar{s}} \frac{1}{K-\nu} \sum_{k=1}^{K} \frac{(n_k - \mathrm{E}(n_k))^2}{\mathrm{E}(n_k)}, \tag{3}$$

where K is the number of detectors, ν is the number of estimated parameters, and $\bar{s} = \frac{1}{K} \sum_{k=1}^{K} [n_k - \mathrm{E}(n_k)]/\mathrm{E}(n_k)$ is a correction for sparse data due to Fletcher (2012). The estimator assumes each n_k is drawn from a detector-specific Poisson distribution. Bischof et al. (2020) applied $\hat{c}$ as a variance inflation factor to simulated data from clustered AC, with mixed results. We extend their simulations to AC from Cox processes and cast light on discrepancies.

Simulations

For data generated from one realization of a stochastic-intensity IHPP or cluster process $\hat{D}$ is an estimate of local density, and estimating $\mathrm{var}(n)$ assuming a Poisson distribution for n should provide intervals with good coverage of the corresponding realization of D_a. However, such an estimate of $\mathrm{var}(n)$ lacks information on the component $\mathrm{var}(D_a)$ in Equation (2) and we predict the

resulting intervals will show poor coverage of μ_D. We used simulation to test these predictions. The sampling design consisted of a 12×12 array of binary proximity detectors (*sensu* Efford et al., 2009) operated for 5 occasions; detector spacing was 2σ where σ was the spatial scale parameter of a hazard halfnormal detection function on each occasion (i.e. $\lambda(d) = \lambda_0 \exp[d^2/(2\sigma^2)]$) where $\lambda_0 = 0.5$ and $\lambda(d)$ is the hazard of detection for an AC at distance d from a particular detector. AC with $\mathrm{E}(N)$ set at 256 were distributed within an area extending $w = 4\sigma$ beyond the detectors. Area A_w was discretized as cells of width 0.5σ. $\mathrm{E}(n) \approx 180.7$ and $a/|A_w| \approx 0.7058$ for all scenarios, where $|A_w|$ is the area of A_w. A homogeneous-density SECR model was fitted to each dataset by maximizing the full likelihood (Borchers & Efford, 2008). Relative bias and confidence interval coverage were computed with respect to both local and global definitions of true density, both with and without adjustment for overdispersion when $\hat{c} > 1$. Further details are given in Appendix S1. For IHPP simulations, the log-intensity surface was modeled as a Gaussian random field with exponential covariance. Simulations used six levels of variance $V \in \{0, 0.125, 0.25, 0.5, 0.75, 1.0\}$ and constant spatial scale 10σ. We also simulated clustered AC from a Thomas process with a scale of 2σ and six levels of the expected number of offspring per parent $\mu \in \{1, 2, 4, 8, 16, 32\}$. D_a was computed from each realized distribution of parents. The intensity of the parent process was μ_D/μ.

Simulations confirmed the lack of bias in density estimates from a homogeneous fitted model even when AC were highly clumped ($V > 0.25$; details in Appendix S1). Confidence intervals also achieved the nominal coverage of local density, but failed to cover the global density of AC from Cox or cluster processes (Fig. 1 and Appendix S1). The term for $\widehat{\mathrm{var}}(n)$ in Equation (1) contributed > 98% of $\widehat{\mathrm{var}}(\hat{D})$ across all simulation scenarios.

When we applied the correction of Bischof et al. (2020) the results were not encouraging. Interval coverage invariably increased (Fig. 1), but the adjustment fell well short of providing nominal coverage for all clumped scenarios. Inflating the variance resulted in intervals for local density that were unnecessarily wide, with coverage approaching 100% (Fig. 1).

Overdispersion in n

We now focus on $\mathrm{var}(n)$ which drives the poor coverage of intervals for $\hat{D}$ from stochastic intensity processes, including cluster processes (e.g., Fig. 1). Coverage could be restored with a variance inflation factor incorporating the overdispersion in n. From Equation (2), for Poisson $n|D_a$ we have $c_n = 1 + a(\theta)\mathrm{var}(D_a)/\mathrm{E}(D_a)$. This allows us to predict c_n from moments of the distribution of D_a. Simulation is at present the only practical and general method for computing these moments (mean and variance), but we need to simulate only the intensity-generating phase of a Cox process, i.e. random fields for LGCP, and the distribution of parents for the Thomas process. To evaluate $\hat{c}$ as an estimate of c_n we conducted further simulations of the Cox process for the distribution of AC followed by spatial sampling to obtain n, but without the time-consuming step of fitting an SECR model. Details and extended results are in Appendix S2. The results in Fig. 2 confirm that the empirical measure of overdispersion $\hat{c}$ underestimates the overdispersion of n. For the Thomas process with negligible within-cluster dispersion (scale ≈ 0) the empirical measure differed only by a near-constant factor of ≈ 0.65, but all other scenarios were worse.

Overdispersion in detection

Our focus has been on the AC point process and its effect on the overdispersion of n and the coverage of confidence intervals for density. Overdispersion in the detection process may also reduce the coverage of confidence intervals, and we briefly consider how this might arise. One possibility is that stochasticity in the intensity surface may lead not only to overdispersion of n, but also to overdispersion in estimates of detection parameters, summarized as $a(\hat{\theta})$ in Equation (1). We conducted simulations to test for an effect of stochastic AC generating processes on the coverage of confidence intervals for $a(\hat{\theta})$, but found none (Appendix S4).

More interesting is the reverse possibility that non-independence in the detection process generates or amplifies overdispersion in n. We found evidence for a small effect of cohesion (synchronous detection of group members) on n (Appendix S4).

Discussion

The Poisson point process of SECR may be strongly inhomogeneous without causing significant bias in estimates of sampling variance from homogeneous models. This is expected because the number of detected individuals n remains Poisson under repeated sampling from an inhomogeneous intensity surface i.e. n is not overdispersed. However, stochastic variation in the intensity surface (vis. a Cox process) invariably results in overdispersion of n and poor coverage of confidence intervals with respect to the global density (Fig. 1). The problem is less extreme for models fitted assuming a fixed population size in some area A around the detectors (Appendix S3), but this arbitrary assumption cannot be justified for common study designs.

Whether users of SECR should be concerned about spatial overdispersion depends on their understanding of the population and the scope of inference. It will often be satisfactory to limit inference to the surveyed population, implicitly conditioning on the prevailing IHPP intensity surface, and then no overdispersion arises. However, inference across multiple realizations of a stochastic IHPP intensity surface is clearly compromised. In our scheme, a varying intensity surface that is fully determined by known covariates is not stochastic when these are included in the model (e.g., Diggle, 2003 p. 70), and no overdispersion is expected. Scenarios that combine deterministic and stochastic effects are of course possible (e.g., Johnson et al., 2010).

Clustering of AC for behavioral reasons is more problematic. Although clustering may be treated mathematically as a special case of a Cox process, īt does not make biological sense to condition on the IHPP intensity surface from one realized set of 'parent' locations. The scope of inference will usually encompass different cluster locations. The sampling variance of density is therefore underestimated for populations with dynamic clusters, as recognised by Bischof et al. (2020). Our simulation results extend those of Bischof et al. (2020) to scenarios with random cluster size and non-zero within-cluster dispersion. These factors reduce the magnitude of actual overdispersion (points are shifted to the left on the x-axis of Fig. 2b).

The variance inflation factor $\hat{c}$ of Bischof et al. (2020) uses internal evidence from one realization of a Cox process to predict overdispersion of n across many realizations. This has proved to be inadequate, both empirically and from the simple logic that the generating process cannot be inferred from a static pattern. When AC come from an IHPP with fixed intensity rather than a Cox process, variance adjustment using $\hat{c}$ results in greater than nominal coverage and is therefore counterproductive.

We defined the effective density D_a as a scalar-valued function, an integral of the IHPP intensity surface weighted by location-specific detection probability. For a Cox process D_a is a random variable. We see two uses for D_a. The first is a general insight: when the intensity surface is stochastic, interval coverage remains good in simulations when each density estimate is compared to D_a for the intensity surface specific to the realization. This is because n is not overdispersed for the IHPP in each realization. Secondly, the value of $\mathrm{var}(D_a)$ can be determined by simulation given a model for the Cox process and values for the detection parameters. Externally acquired knowledge of the generating process is thereby transformed into a usable estimate of c_n. This is plausible in the case of clustering, for which cluster size μ and within-cluster dispersion have biological interpretations, although field estimation will be difficult. Software is described in Appendix S2.

We followed Bischof et al. (2020) in inflating the entire sampling variance of $\hat{D}$ by $\hat{c}$. It is more logical to multiply only the first variance component in Equation (1) by factor c_n, which reduces the effect on $\mathrm{var}(\hat{D})$. Overdispersion of AC slightly increased uncertainty regarding the detection process ($\mathrm{var}[a(\hat{\theta})]$), but coverage of confidence intervals for $a(\hat{\theta})$ was not impaired. If a suitable variance inflation factor is found for overdispersion due to non-independence and heterogeneity in the detection process then it should logically be applied to the second factor in Equation (1). Complete cohesion of detection within spatial clusters is a special case in which the variance components are affected equally and inflation of $\mathrm{var}(\hat{D})$ with the empirical measure $\hat{c}$ is justified. However, complete within-group cohesion seems unlikely if group members are dispersed in space. Cohesion causes overdispersion in detection itself (e.g., Appendix S4), which will be quantitatively important in studies where $\mathrm{var}[a(\hat{\theta})]$ is large.

Data availability

Simulation results are archived in Zenodo at https://doi.org/10.5281/zenodo.14873669 (Efford 2025). Efford (2025) also includes an R script to generate the text figures and a vignette detailing the simulations (summaries for each scenario are in an appendix). Other code required to perform the simulations is publicly available as the R package 'overdispsim' at https://github.com/MurrayEfford/overdispsim, archived in Zenodo at https://doi.org/10.5281/zenodo.15199523.

Acknowledgments

We thank Tilman Davies and Adrian Baddeley for help with simulation of the log-Gaussian Cox process.

References

Anderson, D. R., K. P. Burnham, and G. C. White. 1994. AIC model selection in overdispersed capture–recapture data. *Ecology* 75:1780–1793.

Bischof, R., P. Dupont, C. Milleret, J. Chipperfield, and J. A. Royle. 2020. Consequences of ignoring group association in spatial capture–recapture analysis. *Wildlife Biology* wlb.00649.

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

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock, K. H. 1987. Design and analysis methods for fish survival experiments based on release–recapture. American Fisheries Society Monograph 5.

Chiu, S. N., D. Stoyan, W. S. Kendall, and J. Mecke. 2013. Stochastic geometry and its applications. Third edition. Wiley, Chichester, UK.

Diggle, P. J. 2003. Statistical analysis of spatial point patterns. Second edition. Hodder Arnold, London, UK.

Diggle, P. J., P. Moraga, B. Rowlingson, and B. M. Taylor. 2013. Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. *Statistical Science* 28:542–563.

Efford, M. G. 2025. Simulations of overdispersed activity centres in spatially explicit capture–recapture. [Data sets and R code]. *Zenodo*. https://doi.org/10.5281/zenodo.14873669.

Efford, M. G., D. L. Borchers, and A. E. Byrom. 2009. Density estimation by spatially explicit capture–recapture: likelihood-based methods. In *Modeling demographic processes in marked populations*, D. L. Thomson, E. G. Cooch, and M. J. Conroy (eds), 255–269. Springer, New York, USA.

Efford, M. G., and R. M. Fewster. 2013. Estimating population size by spatially explicit capture–recapture. *Oikos* 122:918–928.

Fletcher, D. J. 2012. Estimating overdispersion when fitting a generalized linear model to sparse data. *Biometrika* 99:230–237.

Johnson, D. S., J. L. Laake, and J. M. Ver Hoef. 2010. A model-based approach for making ecological inference from distance sampling data. *Biometrics* 126:310–318.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals - a unified approach with case-studies. *Ecological Monographs* 62:67–118.

McLellan, B. A., E. Howe, R. R. Marrotte, and J. M. Northrup. 2023. Accounting for heterogeneous density and detectability in spatially explicit capture–recapture studies of carnivores. *Ecosphere* 14:e4669.

Møller, J., A. R. Syversveen, and R. P. Waagepetersen. 1998. Log Gaussian Cox processes. *Scandinavian Journal of Statistics* 25:451–482.

Møller, J., and R. P. Waagepetersen. 2004. *Statistical inference and simulation for spatial point processes*. Chapman & Hall/CRC, Boca Raton, Florida, USA

Royle, J. A., R. B. Chandler, R. Sollmann, and B. Gardner. 2014. Spatial capture–recapture. Academic Press, Waltham, Massachusetts, USA.

Wedderburn, R. W. M. 1974. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* 61:439–47.

White, G. C., and E. G. Cooch. 2017. Population abundance estimation with heterogeneous encounter probabilities using numerical integration. *Journal of Wildlife Management* 81:322–336.

Zhang, W., J. D. Chipperfield, J. B. Illian, P. Dupont, C. Milleret, P. de Valpine, and R. Bischof.

2023. A flexible and efficient Bayesian implementation of point process models for spatial capture–recapture data. *Ecology* 104 e3887.

Figure captions

Fig. 1. Coverage of 95% confidence intervals for $\hat{D}$ as function of clumping when AC followed (a) log-Gaussian Cox process with variance V, and (b) Thomas cluster process with expected AC per cluster μ. Estimates related to either the global density (○) or the local, detection-weighted density (●). Intervals were unadjusted (solid line) or adjusted with the variance inflation factor of Bischof et al. (2020) (dashed line). 1000 simulations per scenario; 95% confidence intervals.

Fig. 2. Effect of stochastic AC generating processes on the relationship between the overdispersion of n and an empirical measure of overdispersion $\hat{c}$. (a) Log-Gaussian Cox process with variance V and three levels of scale, and (b) Thomas (Neyman-Scott) distributions with various parameter levels. Note log scale on both axes. 10000 simulations per scenario.

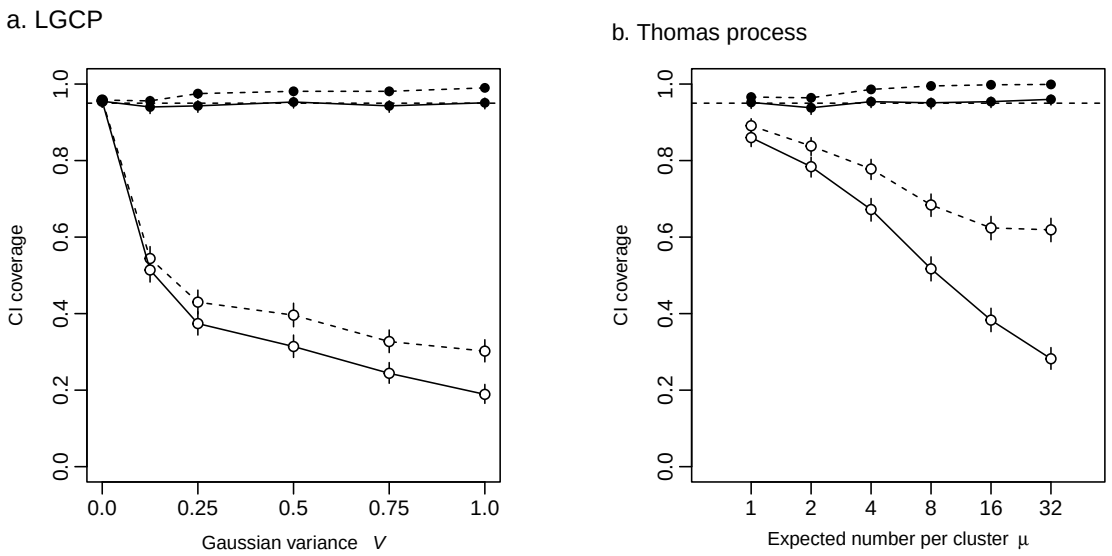
a. LGCP
CI coverage
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.25
0.5
0.75
1.0
Gaussian variance V
b. Thomas process
CI coverage
0.0
0.2
0.4
0.6
0.8
1.0
1
2
4
8
16
32
Expected number per cluster μ

Figure 1

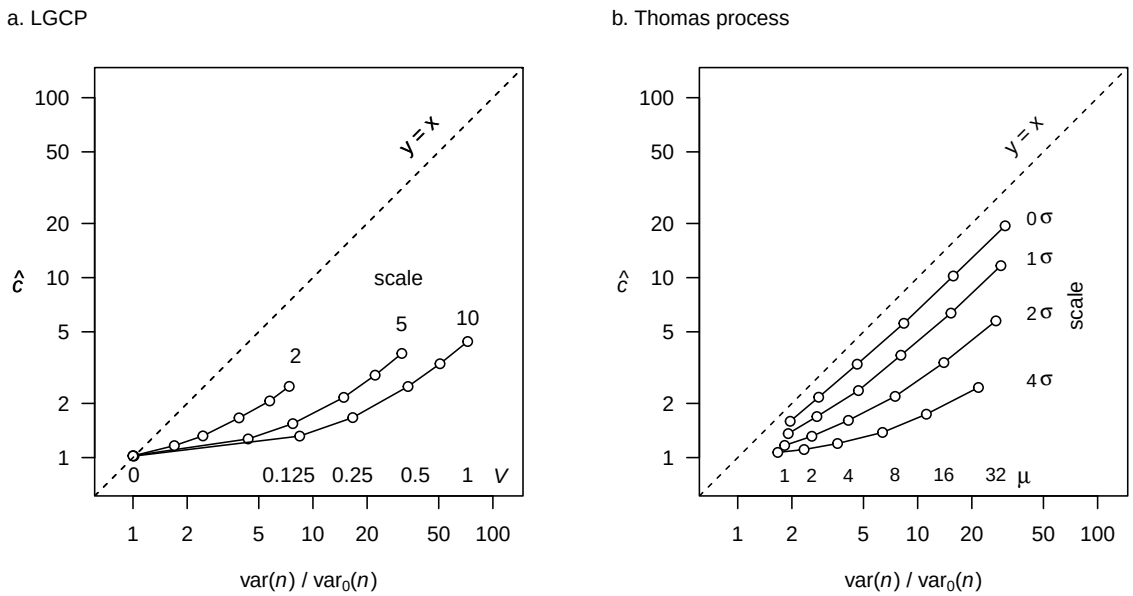
a. LGCP
b. Thomas process
y = x
scale
2
5
10
0
0.125
0.25
0.5
1
V
0 σ
1 σ
2 σ
4 σ
1
2
4
8
16
32
μ
var(n) / var0(n)

Figure 2