What you see is not what is there: Mechanisms, models, and methods for point pattern deviations

Peter Guttorp, Janine Illian, Joel Kostensalo, Mikko Kuronen, Mari Myllymäki, Aila Särkkä, and Thordis L. Thorarinsdottir

Abstract. Many natural systems are observed as point patterns in time, space, or space and time. Examples include plant and cellular systems, animal colonies, earthquakes, and wildfires. In practice the locations of the points are not always observed correctly. However, in the point process literature, there has been relatively scant attention paid to the issue of errors in the location of points. In this paper, we discuss how the observed point pattern may deviate from the actual point pattern and review methods and models that exist to handle such deviations. The discussion is supplemented with several scientific illustrations.

Key words and phrases: ghost point, measurement error, missing data, point process, thinning.

1. INTRODUCTION

Point process models are a modeling framework for stochastic processes that consist of point occurrences in some space, typically a Euclidean space. Historically, point processes emerged from scientific problems related to, for example, astronomy, telecommunication, or life tables (Guttorp and Thorarinsdottir, 2012; Daley and Vere-Jones, 2003). More recently, point process models have been commonly used for spatial data analysis (Møller and Waagepetersen, 2007; Illian et al., 2008; Baddeley et al., 2015) with applications in such diverse

_____

Peter Guttorp is Professor II, Norwegian Computing Center, Oslo, Norway (e-mail: peter.guttorp@nr.no). Janine Illian is Professor, School of Mathematics and Statistics, University of Glasgow, Glasgow, Scotland, UK (e-mail: Janine.Illian@glasgow.ac.uk). Joel Kostensalo is Senior Scientist, Natural Research Institute Finland (Luke), Joensuu, Finland (e-mail: joel.kostensalo@luke.fi). Mikko Kuronen is Research Scientist, Natural Research Institute Finland (Luke), Helsinki, Finland (e-mail: mikko.kuronen@luke.fi). Mari Myllymäki is Senior Scientist, Natural Research Institute Finland (Luke), Helsinki, Finland (e-mail: mari.myllymaki@luke.fi). Aila Särkkä is Professor, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden (e-mail: aila@chalmers.se). Thordis L. Thorarinsdottir is Associate Professor, Department of Mathematics, University of Oslo, Oslo, Norway and Chief Research Scientist II, Norwegian Computing Center, Oslo, Norway (e-mail: thordist@math.uio.no).

fields as forestry (Stoyan and Penttinen, 2000), hydrology (Rodriguez-Iturbe et al., 1987), ecology (Law et al., 2009; Zhang et al., 2023), criminology, epidemiology, oncology (Jones-Todd et al., 2019) or seismology (Kagan and Knopoff, 1980).

At a basic level, the data are given as point patterns—a set of point coordinates describing the locations of objects or events. The data analysis generally aims to infer aspects of the joint distribution of the point locations; this paper focuses on the statistical analysis of such point patterns. Additionally, each point may be associated with a mark describing further characteristics such as size or type of object. While point pattern analysis has some commonalities with more classical statistical analysis, it is in many ways distinct.

As for most other data, observed point patterns are subject to both systematic and random observational error. The observed point pattern may thus deviate from the true, underlying point pattern of interest. However, modeling frameworks discussed in textbooks on the statistical analysis of point patterns lack explicit treatment of observational error. Rather, we have found that such discussion is sparsely scattered across diverse strands of literature, and often aimed at a specific application. The goal of this review paper is to provide a unified overview of the literature on the topic of modeling discrepant point patterns with a specific focus on identifying connections between methodological developments aimed at diverse applications. We will restrict ourselves to unmarked point patterns.

There are two main approaches to dealing with discrepant observations of a point process. As one approach, one can model the full process, i.e., the process generating both the original points and the difference between the original points and the observed ones. However, for many point process models it is difficult to write down the likelihood of the point process of interest even without taking discrepancies into account. Adding a modeling component for the discrepancies will be even more challenging, although it is sometimes feasible. As a simple example, adding i.i.d. measurement errors to the locations of a homogeneous Poisson process yields another homogeneous Poisson process of the same intensity. Another possibility is to study, analytically or by simulation, the effect of discrepancies on parameters and/or parameter estimates in a point process model. This paper provides examples of both approaches.

Since data discrepancies are closely linked to observational mechanisms, substantial space is given here to a review of relevant observational mechanisms for point processes and potential causes of discrepancies related to these. Similarly, we aim to bring together diverse applications in which discrepant observations are common.

This review is certainly not all-encompassing and the included material inadvertently somewhat reflects the authors' interests and expertise. In our literature review, we initiated searches in Google Scholar and Web of Science using the terms "point process error", "location error", "measurement error in point processes" and several variants of these. Furthermore, we searched forward through the citations of the papers we found as well as in particular scientific applications.

The remainder of this paper is organized as follows. Section 2 provides a brief introduction to point processes. Modeling frameworks that adjust for observation errors are reviewed in Section 3. The subsequent Sections 4–8 provide a review of several applications where the observation mechanisms for the point patterns cause discrepancies between actual and observed patterns. Some of these applications include unpublished material for which we show illustrative plots. We end the paper with a discussion and suggestions for future work on this topic in Section 9.

2. POINT PROCESS PRELIMINARIES

A spatial point process X is defined as a random set of locations in some spatial domain of interest, here assumed to be $\mathbb{R}^d$ for $d \in \{2, 3\}$ or a subset thereof. A point pattern $\mathbf{x} = \{x_1, \ldots, x_n\}$ is a realization of a point process observed in a bounded window W. We assume that the number $n(A)$ of points of the process in any bounded set A is finite, and also write $n(\mathbf{x})$ to denote the number of points in the observed point pattern $\mathbf{x}$. A spatial point process is called stationary if the distribution of the process is invariant under translation, while it is called isotropic if the distribution is invariant under rotation. See, e.g., Guttorp (1995) for a more detailed definition of a point process.

2.1 Summary statistics

In the analysis of point pattern data, summary statistics are widely used for exploratory analysis, for model fitting and to assess the goodness-of-fit of estimated models. There are some standard summary statistics of point patterns. The intensity $\lambda(x)$ of a point process is the expected number of points per unit area (volume). That is, letting $S_{x,\epsilon}$ be a ball of radius ϵ centered at x,

$$\lambda(x) = \lim_{\epsilon \to 0} \frac{\mathbb{E}n(S_{x,\epsilon})}{|S_{x,\epsilon}|},$$

where $|A|$ denotes the area or volume of the set A. For a stationary process the intensity is constant. A constant intensity can be estimated by

$$\hat{\lambda} = \frac{n(W)}{|W|},$$

while for a non-stationary point process, the intensity and other parameters are location dependent.

The second-order intensity measures the joint intensity at two locations and is given by

$$\lambda_2(x, y) = \lim_{\epsilon_1, \epsilon_2 \to 0} \frac{\mathbb{E}\left(n(S_{x,\epsilon_1})n(S_{y,\epsilon_2})\right)}{|S_{x,\epsilon_1}||S_{y,\epsilon_2}|},$$

assuming it exists, see, e.g., Illian et al. (2008) for more details. For a stationary isotropic point process the second-order intensity depends only on the distance $\|x - y\|$ between x and y, and we write it as $\lambda_2(\|x - y\|)$.

For stationary isotropic processes one often considers the K-function,

$$K(r) = \frac{2\pi}{\lambda^2} \int_0^r u\lambda_2(u)du.$$

This can be thought of as the expected number of additional points within distance r of an arbitrarily chosen point, normalized by the intensity. In order to estimate $K(r)$ from data, given by occurrences in a window W, we count the number of occurrence pairs of distance at most r apart. The expected value of this quantity is approximately $\lambda^2|W|K(r)$, where the approximation arises from the possibility that an event in W can be within distance r of the boundary of W, and hence have a neighbor falling outside of W. While one can correct for this edge effect, one can also simplify matters by ignoring all points that fall within distance r of the boundary. Often variants of the K-function are used, including the variance stabilizing L-function $L(r) = \sqrt{K(r)/\pi}$ and the non-cumulative pair-correlation function $g(r) = \frac{K'(r)}{2\pi r}$, which is essentially the derivative of the K-function and may be interpreted as the number of additional points *at* distance r.

Another common summary statistic is the nearest neighbor distance function G. For each point x_i in the pattern $\mathbf{x}$ we find the distance r_i to the nearest other point in $\mathbf{x}$. The empirical distribution function of the distances is a natural estimate of the theoretical nearest neighbor distribution, which of course depends on the point process model used. If instead, the distance from an arbitrary point in space to the nearest point of the process is of interest, the empirical distribution of empty space distances (or empty space statistic) can be used.

While the intensity in general does not determine the distribution of the process, the Papangelou conditional intensity provides a full characterization of the process. It is defined as

$$\lambda(u|\mathbf{x}) = \begin{cases} \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})}, & u \notin \mathbf{x} \\ \frac{f(\mathbf{x})}{f(\mathbf{x} \setminus \{u\})}, & u \in \mathbf{x} \end{cases},$$

where $f(\mathbf{x})$ is the density of the pattern $\mathbf{x}$, and is essentially the conditional probability of observing a point of the process $\mathbf{X}$ at u given the observed point pattern $\mathbf{x}$ not including u.

2.2 Point process models

The simplest stochastic model for a point process is a Poisson process, in which the counts in disjoint sets are independent Poisson random variables. If the mean of each Poisson-variable is proportional to the area of the set, the Poisson process is called homogeneous, otherwise non-homogeneous. In the homogeneous case, the conditional distribution of points, given the total number of points in the set, is that of independent uniform random variables. Therefore, a homogeneous Poisson point process is called completely spatially random. Compared to a Poisson process, a point pattern is often considered clustered or regular (although it may be neither, or it may be both at different scales). A common generalization of the non-stationary Poisson process is the doubly stochastic Poisson process (or Cox process), where the intensity function of a non-stationary Poisson process is a realization of a stochastic process. The log Gaussian Cox process, a Cox process where the logarithm of the intensity is a Gaussian process, is popular in applications since it is completely characterized by the intensity and pair correlation function of the Cox process.

The most common clustered model is the Neyman–Scott process, which consists of a primary point process of cluster centers, and for each cluster center a secondary point process of cluster points, often independently placed around the primary point. The primary process may or may not be observed.

As to regular patterns, perhaps the simplest is a Poisson process from which all pairs of points closer than a distance d are removed. Such a process is called a hard core rejection model. Other common point process models for regular patterns are Gibbs processes given by a density f with respect to a unit rate Poisson process, that is, the homogeneous Poisson process with intensity 1. For a neighborhood relation $\sim$ (describing which points are considered neighbors), the density can be written

$$f(\mathbf{x}) \propto \prod_{\mathbf{y} \subset \mathbf{x}} \phi(\mathbf{y}),$$

where the non-negative interaction function $\phi(\mathbf{y}) = 1$ if there is any pair of points (y_i, y_j) in the pattern $\mathbf{y}$ with $y_i \sim y_j$. A simple nearest neighbor interaction function is of the form $\gamma^{\mathbb{1}(\|y_i - y_j\| \leq R)}$, where $\mathbb{1}$ denotes the indicator function and $0 \leq \gamma \leq 1$ measures the strength of inhibition of points within distance R. This process is called a Strauss process. In general, it is common to consider only pairwise interactions.

The historical development of these and related point process models is discussed in Guttorp and Thorarinsdottir (2012) and technical details can be found in, e.g., Daley and Vere-Jones (2003), van Lieshout (2000), and Møller and Waagepetersen (2007).

2.3 Statistical inference

To carry out inference for a point process we need to define a likelihood (technically, the Radon–Nikodym derivative of the point process measure with respect to a unit rate homogeneous Poisson process). The log-likelihood for a non-homogeneous Poisson process with intensity $\lambda(x)$ is

$$\ln L(x_1, \ldots, x_k; \boldsymbol{\theta}) = \sum_{i=1}^{k} \ln \lambda(x_i; \boldsymbol{\theta}) - \int_W (\lambda(x); \boldsymbol{\theta}) - 1) \, dx.$$

The joint density or likelihood function for Gibbs processes includes a normalizing constant that usually cannot be computed analytically, making the likelihood inference computationally expensive. An alternative for the maximum likelihood method is the pseudo-likelihood method, where the likelihood function is replaced by a product of Papangelou conditional intensities which do not include the normalizing constant.

For a more comprehensive review of inference tools for spatial point patterns, see Møller and Waagepetersen (2007), Illian et al. (2008) and Baddeley et al. (2015).

3. MODELS FOR DEVIANT POINT PATTERNS

Observed point patterns may deviate from the actual point pattern in a multitude of ways. There may be *recording errors* in location recordings. For example, the coordinates of points may have been accidentally switched in the field records. A variant is rounding of data due to changes in precision of measurement devices. As an illustration, early Icelandic volcanic eruptions are known only to the

nearest year, or locations of trees are recorded only up to a certain accuracy dependent on the measurement device used. More generally, points can be displaced due to *measurement error*. For example, the location of an earthquake is determined using inversion of arrival times of primary waves (P-waves) and secondary waves (S-waves) at several seismometers. The location process itself includes uncertainty, as does each arrival time used in the inversion.

Points may be lost by *censoring*, for example due to location measurement error pushing the point outside the observation area, but also by points being located too close together to be observed as separate points. Conversely, a *ghost pattern* in the form of points that do not belong to the point pattern of interest may appear. An example might be a pattern formed by landmines, where many mine-like objects are located, and the real land mines must be identified. A second example with similar issues is automatic detection of animals in remote sensing images or through drone surveys, e.g., detection of seal pups on ice floes or orangutan nests in a rainforest (Milne et al., 2021).

Thinning occurs when a process different from that generating the true pattern operates on the pattern by removing points. This can, e.g., occur in transect sampling, where the probability of observing a point can be related to the distance from the transect. Alternatively, a smooth *deformation* of the point pattern may happen. This can, for example, result from optical or instrumental distortion of observations from space.

Below, we review modeling frameworks that have been proposed to account for these various deviations in the observations.

3.1 A contamination model for point process data

A fairly general model for *contamination* of point processes is a combination of thinning, ghost pattern, and measurement error. Following Assunção and Guttorp (1999) we denote the actual point process (or more precisely, its counting measure) by X, an independent superposed point process by Y', and a deletion process by X' taking on values 0 or 1, corresponding to deletion or retention, respectively. Writing the contaminated process Y we have

$$Y(B) = \int_B X'(x) X^*(dx) + Y'(B).$$

Here X^* corresponds to the point locations of X perturbed by measurement error.

Alternative models have been proposed for more specific contaminations. For a hard-core process X, Redenbach et al. (2009) suggest a mechanism for creating outliers violating the hard-core condition by picking a point in the true pattern at random and adding another point (outlier) in the vicinity of this point.

In presence-only and citizen science data, false absences are accounted for by detection probabilities as discussed in Section 3.4 below. To account for false positives, on the other hand, Royle and Link (2006) introduce a mixture model that estimates the probability of detecting an absent species. Here, the estimated probability can be inflated by detection heterogeneity, and a reliable estimation of false positives requires a subset of the observations to be unambiguously true detections, see Altwegg and Nichols (2019) and references therein.

3.2 Likelihood inference for deviant observations conditional on unobserved truth

Lund and Rudemo (2000) develop a likelihood inference approach for the Assunção–Guttorp modeling framework specialized to the case where Y' is a Poisson process and the deletion process consists of independent deletions. Specifically, they consider a pair of finite point processes, the true process $X = \{X_i : i \in I\}$ and the process of imperfect observation $Y = \{Y_i : i \in J\}$ of X, where $I = \{1, \ldots, n\}$ and $J = \{1, \ldots, m\}$. Let $g_{gh}(\cdot)$ be the density for ghost points, $p(\cdot)$ the retention probability, and $g(\cdot|x_i)$ the displacement density. Then, the conditional likelihood of Y given X with respect to the unit rate Poisson process becomes

$$L(\mathbf{y}|\mathbf{x}) = \sum_{\substack{I_1 \subseteq I \\ J_1 \subseteq J \\ |I_1| = |J_1|}} \sum_{\pi} L_1 L_2 L_3,$$

where the inner sum is over all bijections $\pi : I_1 \to J_1$ for two finite sets I_1 and J_1 with the same cardinality (denoted by $|\cdot|$). Furthermore,

$$L_1 = \prod_{i \in I_1} p(x_i) g(y_{\pi(i)}|x_i)$$

handles displacement of points that are retained with probability $p(\cdot)$,

$$L_2 = \prod_{i \in I \setminus I_1} \left(p(x_i) \int_{\mathbb{R}^d \setminus W} g(y|x_i)\, dy + 1 - p(x_i) \right)$$

handles cencoring of points that are retained with probability $p(\cdot)$, and

$$L_3 = \exp\left(|W| - \int_W g_{gh}(y)\, dy \right) \prod_{j \in J \setminus J_1} g_{gh}(y_j)$$

handles placing of ghost points.

Lund and Rudemo (2000) use this framework to estimate tree-top locations of Norway spruce (*Picea abies*) from an aerial photograph having the candidate tree locations determined either by kernel smoothing (Dralle and Rudemo, 1996, 1997) or template matching (Larsen and Rudemo, 1998) as the starting point. Lund et al. (1999)

suggest a Bayesian analysis in a similar set-up (without censoring). Three Gibbs point process models, namely Poisson process, Strauss process, and a process with a logistic pairwise interaction function, are considered as prior distributions for the unobserved true point pattern. Given the prior distribution for X and the model parameters, a Metropolis–Hastings sampler is constructed to sample from the posterior distribution of X given Y. In a follow-up paper (Lund and Thönnes, 2004), the Metropolis–Hastings algorithm is replaced by perfect simulation.

In an application to geocoded spatial data, Fanshawe and Diggle (2011) propose a general model-based solution to correct for positional errors in various types of spatial data. Specifically, they derive the likelihood function for a linear Gaussian geostatistical model incorporating positional errors. However, the inference method is not computationally feasible for even moderately large data sets. In a follow-up work, Fronterrè et al. (2018) propose a more computationally efficient composite likelihood approach for the case where the positional errors are due to geomasking.

3.3 Perturbed summary statistics

If the locations are measured with error, the observed locations can be regarded as random perturbations of the underlying correct locations. Summary statistics, such as Ripley's K-function, are not the same for the perturbed and the correct processes since typically, perturbation makes the process more Poisson-like. Diggle (1993) derives the relationship between the perturbed K-function and the original K-function when the (radially symmetric) disturbance model is known. Let g denote the convolution of two independent perturbations. Then the K-function of the perturbed pattern becomes

$$K(r) = \frac{2\pi}{\lambda} \int_0^\infty R(r, u) u \lambda_2(u) du,$$

where

$$R(r, x) = \int_{\|u\| \le r} g(x - u) du$$

is the probability that a point at x gets perturbed to some point in the region $\|u\| \le r$. We see that the result of the perturbation is to smear out the original K-function.

The impact of measurement error on the nearest neighbor distribution function G and the empty space statistic F is studied and applied to tropical forest data in Bar-Hen et al. (2013). The unobserved true point process is assumed to be stationary and the perturbation model as in Chakraborty and Gelfand (2010). The summary statistics for the true process X and the distribution of the noise ϵ are estimated from the observed process Y by using the influence of simulated perturbations on the observed point process.

Using a Bayesian framework, Lund et al. (1999) suggest posterior estimates for the summary functions K, G, and F of a true, unknown point pattern when the observed pattern is exposed to thinning, displacement, or a ghost pattern. They further provide sampling variation in each case. Kuronen et al. (2021) estimate parameters of an actual process based on summary statistics of functions robust with respect to errors.

3.4 Adjusted intensity estimates

Contamination in various forms may also be included in the modeling and estimation of the intensity. One widely used approach to estimating the intensity of a point process model is based on kernel tools. If these are used in the context of a contaminated point process, deconvolution methods need to be applied. Diggle (1985) suggests a kernel estimator of the point process intensity for temporal non-stationary point processes observed with *location errors*, subsequently extended to the spatial setting by Cucala (2008). In the simplest case of no edge correction, the intensity estimate (with $\mathcal{F}$ denoting the Fourier transform and b the bandwidth) is

$$\lambda_b^*(x) = \mathcal{F}^{-1}\big(\mathcal{F}(\hat{\lambda}_b)/\mathcal{F}(g)\big)(x).$$

Here, g is the density of the location errors and

$$\hat{\lambda}_b(u) = \int \frac{1}{b^2} k\Big(\frac{u - x}{b}\Big) dX(x)$$

is a kernel estimate of the intensity of point process X using kernel k.

Chakraborty and Gelfand (2010) consider the case where the observed point pattern is a result of random perturbation of the true point locations. They assume that the true unobserved point pattern is a realization of a Cox process, where the intensity is modeled by a scaled Gaussian mixture distribution. The perturbation model is $y_i = x_i + \epsilon_i$, where the noise variables ϵ_i are i.i.d. random variables. Two cases are considered: the perturbation may cause a true point moving outside the observation window and being missed or a point outside the window moving inside the window and being falsely counted as a point within the window. Given the observed noisy data, the true intensity function is estimated in a hierarchical Bayesian framework, where the unknown point pattern is modeled by a Cox process and the true number of points is one of the model parameters. The modeling framework is applied to estimate the intensity of ecological field data in the Cape Floristic region in South Africa.

Another scenario concerns point processes where some of the points that are present have not been detected during data collection. This is common in what is referred to as distance sampling. Distance sampling is used, in particularly, in animal ecology for gathering data on, e.g., wildlife populations where the area of interest is typically too big to be sampled exhaustively in its entirety

(Buckland et al., 2001, 2015). The most common examples for this are line transect sampling and point transect sampling. In line transect sampling, an observer moves along a linear transect observing the locations of visible objects along the way as well as their distance from the observer, while in point transect sampling, the observer stands in a location and observes animals within a certain radius, e.g., by listening for birds (Buckland, 2006). Similarly, one of the approaches for deriving stand level characteristics from terrestrial laser scanning (TLS) data in forest inventory applications is based on distance sampling methodology (Ducey and Astrup, 2013).

The effect that objects closer to the observer are more likely to be observed than those further away is included in the modeling framework using a detection probability, decreasing in distance from the observer, as a component of the density. That is, under the assumption that the distribution of all points is uniform with respect to distance from the transect the model is given by a thinned point process and hence an adjusted intensity with unknown thinning probability that may evolve in time and space (Hedley and Buckland, 2004; Renner et al., 2015; Yuan et al., 2017). The uniformity assumption is typically justified due to the small size of the transects relative to the survey area of interest.

Traditional distance sampling methodology has discretised space and modeled the data as Poisson counts, but more recently, point process approaches that model in continuous space have been developed (Stoyan, 1982; Waagepetersen and Schweder, 2006). Chandler et al. (2011) suggest a hierarchical model for inferring the density of so-called unmarked populations–that is, data where the identity of an individual is not known–subject to temporary emigration and imperfect detection. They assume the true process is a homogeneous Poisson point process. Then, the observed count $n(\mathbf{x}_{it})$ within a plot i at time t depends on a detection probability p and N_{it}, the true number of individuals present in plot i at time t. N_{it} again is a random number that depends on the total number of individuals that could possibly be detected in plot i over the duration of sampling and the level of temporary emigration at time t. The model is developed for general sampling designs and applied to repeated removal sampling data collected on Chestnut-sided Warbles in the White Mountain National Forest in the United States.

More generally, Schmidt et al. (2022) provide a recent overview and a decision support for practitioners to select an appropriate inference approach based on the type of thinning probabilities present in the data collection design when estimating population sizes based on imperfect detection. For a probability-based design it is possible to calculate the inclusion probability p_s that a given individual is included in the sample. The detection process of individuals within the sample unit can be divided into p_p, the probability that the individual is present within the sample unit, p_a, the probability that the individual is available (above ground, above water, vocalizing, etc.) for observation given that it is present, and p_d, the probability that an individual is detected given that it is present and available. Commonly, objects close to the observer are more likely to be observed than those further away.

Intensity estimates may also be adjusted for errors by including relevant covariate information in the model for the intensity. For example, to account for study site selection bias, Renner et al. (2015) include distance from main roads and distance from urban areas as predictors for modeling the density of *Eucalyptus sparsifolia* trees based on presence records from incidental sightings of the species. Furthermore, intensity is sometimes estimated based on coarsened data or on several data sources having different spatial resolutions. For example, (Huang et al., 2014) and (Chang et al., 2015) investigate risk factors for endometrial cancer based on a case-control study, where complete addresses are available, and a surveillance system, where addresses and covariate information are known only at zip code level.

3.5 Adjusting the intensity with a detection probability

Here, we describe further details of the modeling setting where the intensity estimate is adjusted with a detection probability. Specifically, for a distance sampling procedure at a fixed point in time, denote the true spatial process by X, and the process of imperfect observation of X by Y, as in Section 3.2 above. Assume the observer is moving at a constant speed v along a line transect and denote the perpendicular distance from the line transect by $z(x)$. Then,

$$\mathbb{P}(\text{object at } x \text{ detected}|x \in X) = 1 - \exp\left(-\frac{h(z(x))}{v}\right)$$
$$= g(z(x), v),$$

where h is an aggregated detection hazard along a path and g is the aggregated detection function. The intensity of the observed process Y is then given by

$$\lambda^*(x) = \mathbb{P}(\text{object at } x \text{ detected}|x \in X)\lambda(x),$$

where λ is the intensity of the true process X.

Hayes and Buckland (1983) derive the commonly-used hazard-rate model $h(z(x)) = -(z(x)/\sigma)^{-b}$ for $b, \sigma > 0$, while the detection function is widely given by the half-normal density $g(z(x)) = \exp(-z(x)^2/(2\sigma^2))$ with $\sigma > 0$, or a generalization thereof. Only the latter results in a log-linear probability model and is thus more computationally convenient for inference under a point process modeling framework. Yuan et al. (2017) consider such an approach where the unknown detection function is modeled and its parameters estimated jointly with the underlying point process, i.e. as a thinning of a point process
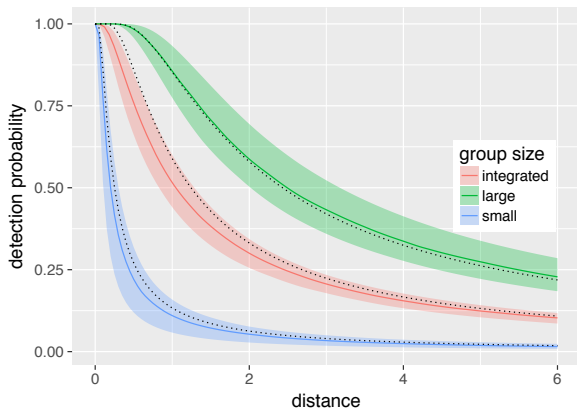
FIG 1. *Estimated detection functions for different group sizes for a thinned marked point process resulting from distance sampling.*

accordingly, in particular a log Gaussian Cox process. The intensity is obtained by multiplying the intensity of the log Gaussian Cox process by the detection probability given above.

In a more general scenario, detection can also depend on properties of the objects represented by the points, e.g. the size of an animal or animal group, in other words on the mark in a marked point process. For a mark m observed in location x with distribution $p(m|x)$, the intensity becomes

$$\lambda^*(x, m) = \mathbb{P}(\text{object at } x \text{ detected}|x \in X, m)p(m|x)\lambda(x).$$

For illustration, we consider an example of the pattern formed by the locations of groups of pantropical dolphins (*Stenella attenuata*), with varying sizes of groups to which a model was fitted. Here detection depends on the mark *groupsize*. Figure 1 shows the estimated half-normal detection functions. Larger groups of animals are easier to detect at larger distances than those at close distances. The model was fitted using the software package *inlabru*, which allows the fitting of models using integrated nested Laplace approximations (INLA) without the requirement of a linear relationship between the parameters and the response (Bachl et al., 2019).

4. RECORDING AND MEASUREMENT ERRORS

We now move from a general discussion of modeling frameworks to specific applications, starting with a discussion of various causes of recording and measurement errors.

4.1 Microscopy data

In our context, microscopy data refers to point patterns that are derived from analyzing objects in microscopy images. Here, we give three examples of microscopy data that can be analyzed by using point processes: confocal microscopy images of epidermal nerve fibers, scanning transmission electron micrograph (STEM) images

of silica gel, and photoactivated localization microscopy (PALM) images of proteins in biological cells. Depending on the microscopy technique used, different types of complications may occur.

In the nerve fiber example, the points where the nerve trees enter the epidermis (outermost layer of the skin), and the points where the nerve fibers terminate, have been analyzed as point patterns in several studies in 2D and 3D, see, e.g., Olsbo et al. (2013), Andersson et al. (2016) and Konstantinou and Särkkä (2021). The nerve fibers can be followed from one xy-image to another but the accuracy of the point locations depends on the chosen resolution, which is typically lower in the z-direction than in the xy-plane. In the data mentioned here, the resolution in xy is $0.83 \times 0.83\ \mu m^2$ compared to $2\ \mu m$ in z. So far, this difference has not been taken into account in the analysis of such nerve fiber images.

In 2D STEM images of silica, the intensity response is increasing with increasing mass thickness of silica. Therefore, 3D silica nanoparticle structures can be reconstructed from the 2D images if the mass thickness-intensity function is known or can be estimated. Assuming that the intensity response is monotonically increasing with increasing mass thickness, Nordin et al. (2014) estimate the functional form by using a maximum likelihood approach. Which functional form, e.g., linear or power law, is chosen and how the estimated number of silica nanoparticles are located in the z-direction will affect the resulting 3D point pattern.

Photoactivated localization microscopy (PALM) is a powerful imaging technique for characterization of protein organization in biological cells. Due to the stochastic blinking of the fluorescent probes and some camera discretization effects, a cluster of observations is detected instead of a single protein. Jensen et al. (2022) suggest a model, the independent blinking cluster point process, for such data and a method to recover the ground truth based on images observed with the blinking artifact.

4.2 Particle counters

Particle contagion is important to identify in machine oil, in clean rooms such as operating theaters in hospitals, and in air quality, among many such applications. The most common type of errors are counting and identification errors.

Rutherford et al. (1910) had developed a method for counting alpha particles in radioactive emissions using an electrometer and photographic film. However, it was tedious to use, and instead they use a scintillation method invented by Regener (1909), where a zinc oxide surface gives off a flash when hit by a particle. The scintillations on a small area of the zinc oxide surface were viewed in a microscope, with time reserved for resting the eyes, and recorded on a paper string by hitting an electric switch.

Clearly, there is a possibility of missing a scintillation, and of observing a spurious one. In addition, the time of the record is likely offset by a random period of time between seeing the scintillation and hitting the switch. For a homogeneous Poisson process this is of no consequence, but when the process is non-homogeneous or non-Poisson, there are different consequences. Modern radiation counters are based on the Geiger-Müller tube, which has a brief dead time after registering a particle.

Martens (1968) discusses measurement errors in devices that count (and size) air pollution particles using light scattering. He lists, in addition to sizing errors, three kinds of counting errors, namely instrumental , statistical, and imaginary errors. The term imaginary errors refers to non-repeatability of measurements. The statistical errors are largely due to incorrect particle sizing, while the instrumental errors also can occur, e.g., when two particles are interpreted as one, or there is inhomogeneity of the viewing volume illumination.

Modern tools for particle counting use laser beams, either measuring particle shadows or laser light scattering. The errors are similar: coincidences where two particles are lined up with the laser beam which only sees one shadow, particles that are too small to cast an interpretable shadow, and the same types of light scattering problems as when using incoherent light. There are a variety of other methods, many mentioned in Zhang et al. (2009).

4.3 Earthquake locations

The location of the epicenter of an earthquake is determined from translating the temporal distance between P- and S-waves into a distance, and then triangulating the distance from all the seismometers recording the earthquake. Errors in timing, which can be due to lack of or imprecise coordination with a timing source (Hable et al., 2018), then translate to location errors through the triangulation process. More generally, the arrival times can be written as path integrals through the unknown (or at best partially known) structure of earth. The most common approach is iterated linearization (e.g., Menke, 1989). It can, however, be unreliable or unstable in the presence of outliers (Husen et al., 2003).

The ETAS (Epidemic Type Aftershock Sequence) model, proposed by Ogata (1998) and extended by Ogata and Zhuang (2006), is a spatio-temporal point process model for earthquakes. To investigate the effect of measurement error, Schneider (2021) estimates the parameters of an ETAS model to a catalog in the Pacific Northwest, for both the catalog locations as well as jittered locations of the seismic events according to the uncertainties given in the catalog. Figure 2 shows boxplots of original estimates (obtained by MCMC) and nine jittered estimates. We see that the parameters μ (background rate), c (temporal offset of aftershocks) and p (temporal decay rate of aftershocks) are similar to the estimates from unjittered data, while the estimates of K (aftershock productivity) are biased downwards for the jittered data, while estimates for d (spatial offset of aftershocks) and q (spatial decay rate) are biased upwards, while those for q also are substantially more uncertain.

4.4 Geocoded data

Spatial epidemiology is a subfield of epidemiology focused on the study of the spatial distribution of health outcomes. The data analysis is commonly performed by using a spatial point process model to analyze geocoded location data, that is, text-based description of a location—such as an address or the name of a place—transformed to geographic coordinates. Data of this type similarly arises in, e.g., microeconomics and demographic analyses.

Geocoded data have two different kinds of potential error mechanisms (Fronterrè et al., 2018). Imprecision may be introduced in the geocoding due to incorrect placement of an address along a street segment or the same geocoded location may be assigned to multiple addresses based on a post-code system. Automatic geocoding may also fail for a substantial proportion of the available data and these failures can be spatially clustered (Zimmerman et al., 2010). Alternatively, it may not be possible to report precise measurement locations due to issues of confidentiality. Random or deterministic perturbation of the locations is then applied in a process known as geomasking (Armstrong et al., 1999). See Jacquez (2012) for an overview of the impacts of geocoding positional error in health analysis.

4.5 Target tracking/Telemetry data

An alternative to distance sampling in animal ecology discussed in Section 3.4 is to place a global positioning system (GPS) tracker on the moving object(s) of interest. GPS locations, for example in wildlife tracking (Johnson and Kuhn, 2013), have measurement error, and often these errors are ignored. However, there has been some work on looking at the error structure of GPS observations, for example by placing location trackers in different ecological areas. Williams et al. (2012) find that measurement error is higher, by about a factor of two, in areas of dense canopies. Among the error sources are the number of satellites finding the GPS, the temporal resolution of GPS (particularly when following a moving particle, such as an animal), and errors in the positioning or clock of the receiving satellites (Langley, 1997). Ranacher et al. (2016) shows that the autocorrelation in GPS locations lead to systematic overestimation of the distance traveled by the object. While many applications to telemetry data don't apply point process approaches, Arce Guillen et al. (2023) use a log-Gaussian Cox process model to account for the autocorrelation of subsequent observations in animal tracking data.
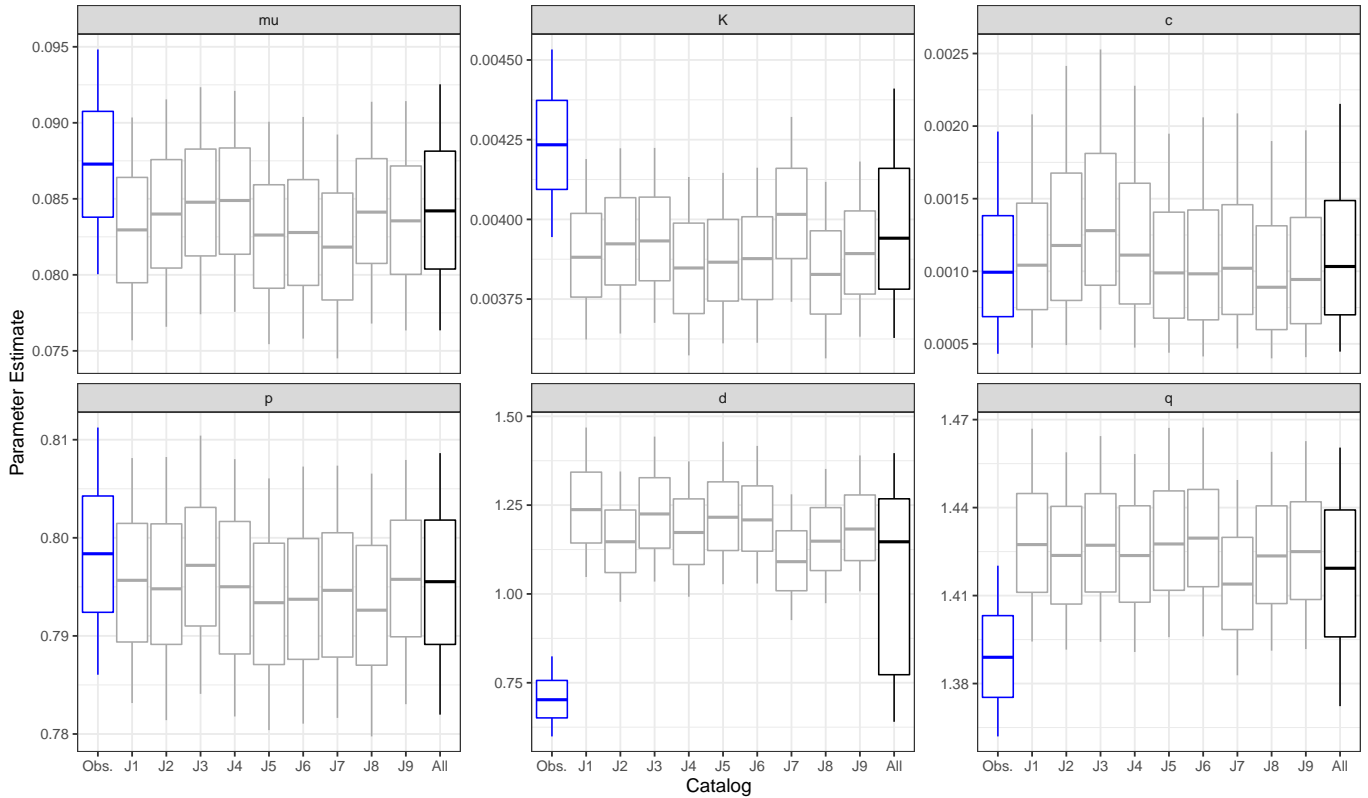
FIG 2. *Box plots summarizing posterior densities of ETAS parameters for the reported baseline catalog (blue), as well as nine jittered versions of this catalog (grey), and for the grand posterior encompassing the reported and 99 jittered catalogs (black). From Schneider (2021).*

Glennie et al. (2021) combine animal movement data with distance sampling data in order to account for animal movement independent of the observer, using a spatial hidden Markov model to approximate the underlying spatial process.

5. GHOST PATTERNS

In the applications discussed below, errors in the observation mechanism cause the introduction of an additional ghost pattern in the data, rather than errors in the actual pattern of interest.

5.1 Mine fields

Land mines kill some 20,000 people worldwide each year. The problem of identifying the location of land mines from satellite, airplane or drone pictures has been considered in the point process literature at least since Dasgupta and Raftery (1998) and Byers and Raftery (1998). Typically, land mines are located among a clutter of rocks, metal junk, and other features that can be misidentified as mines. Holmes et al. (1995) describes a multi-spectral satellite imaging approach to finding land mines, while other methods include infrared thermography (Thành et al. (2009)), 3D voxel radar (Brockner et al. (2018)), and polarimetric camera (Connor et al. (2008)).

A point process approach to identification of land mines is based on viewing the observations as a superposition of two point processes: mines and mine-like non-mines. Cressie and Lawson (2000) assume that the two processes are cluster processes with the cluster centers being independent Strauss processes with different parameters, and the cluster points distributed around the cluster centers using a Matérn radial distance function. A hierarchical approach using Markov chain Monte Carlo (MCMC) is taken to estimate the model parameters.

5.2 Spatial pattern of air bubbles in polar ice

Polar ice has information on the climate of the past. However, to be able to interpret the ice core records, one has to know how old the ice is. One way to determine the age of ice at different depths is by estimating the deformation of ice based on the air bubbles in the ice samples.

Redenbach et al. (2015) analyze ice samples drilled in Antarctica. The data consist of locations of air bubbles extracted from computer tomographic (CT) images of ice samples imaged inside a cold room at $-15°C$. In addition to the "real" air bubbles, the ice samples contain some relaxation (extra) bubbles that appear when the ice core is pulled out from the drilling hole. Such bubbles do not give

any information on the motion (deformation) of the ice but disturb the directional analysis necessary to gain some information on the deformation. Therefore, each bubble is classified either as real bubble or noise so that the noise bubbles can be removed prior the deformation analysis.

The set-up is very similar to the mine detection problem. However, in this case, it is not necessary that every bubble is classified correctly. It is more important that the spatial distribution of the true bubbles can be reconstructed. For example, in a close pair of bubbles most likely one of them is real and the other one noise and it does not matter which one is classified as real and which one as noise.

Redenbach et al. (2015) suggest the following model for the bubble configuration: The real air bubble pattern is regular and modeled by a Strauss process Y_1 with the intensity related parameter $\beta > 0$, interaction range R, and interaction strength γ. The noise bubbles are modeled as a Poisson process Y_0 with intensity λ_0. Then, the complete observed point process Y is a superposition of Y_1 and Y_0. The parameters $(\lambda_0, \beta, \gamma, R)$ as well as the classification of each point (real or noise) were estimated in a Bayesian framework. In Redenbach et al. (2015), an MCMC approach is constructed, and a computationally less demanding method based on variational Bayes approximation is suggested in Rajala et al. (2016).

Ideally, the points would be divided into two groups, points with high posterior probability (real bubbles) and points with low posterior probability (noise). However, in addition to these two groups, a group of points having intermediate posterior probabilities is found. Typically, the points in this group are R-close pairs of points. Some suggestion how to classify the points in this group are suggested in Redenbach et al. (2015) and Rajala et al. (2016). When the model parameters are known the spatial pattern of the true bubbles is estimated quite well. If the parameters need to be estimated simultaneously with the classification, isolated noise bubbles are typically classified as real bubbles, and as a result, the intensity λ_0 of the noise bubbles is underestimated and the intensity related parameter β of the true bubble process overestimated (too many real bubbles and too few noise bubbles).

6. THINNING PROCESSES

In many applications, locations of individuals in space and time are recorded, but for practical reasons a specific observation process has been chosen, which can be seen as a thinning operation on the point pattern. The thinning can appear systematic or at random, as discussed below.

6.1 Area sub-sampling in animal ecology

In some cases, the point pattern has only been partially observed as there are some areas where sampling cannot take place, e.g., for accessibility issues. In other cases the area of interest especially in animal ecology and conservation is simply too big—e.g., an entire ocean—to observe the pattern everywhere. As a result, the point pattern has only been observed in one or several sub-plots or transects, which have been extensively sampled. This may have been done through what is referred to as plot sampling or transect sampling using areal-video-surveys or drones. It is however of interest to predict into the entire area surrounding the sampled subareas by assuming that some covariates of interest are known within the whole area, with information on the structure of the point pattern and the relationship between the point pattern and the covariates taken from smaller subareas. In these scenarios, the detection probabilities are 1 in the surveyed areas and 0 in the areas that have not been surveyed. In the modeling approach, areas with no points in the surveyed areas and areas with points in the unsurveyed areas have to be distinguished (Williamson et al., 2022).

6.2 Distance sampling in ecology

As mentioned in Section 3.4, distance sampling is a common method for gathering data in animal ecology, where detection probabilities depend on the distance from the observer. In one of the earlier works based on marked point processes, each object (animal) is described by a disc, where the center of the disc is the location of the object and the radius of the disc is the associated sighting distance (Högmander, 1991). An animal is observed if and only if the transect hits the disc connected to it.

In some terrains, a subset of the objects might not be visible, e.g., when observing ocean mammals from a water vessel, or due to temporary emigration. This effect is called availability bias and its estimation commonly requires additional information from independent data (Laake and Borchers, 2004). Animal movement can cause substantial bias in density estimation from line transect or point surveys. Movement in response to the observer's presence is usually accounted for in the data collection procedure rather than the data modeling procedure (Borchers et al., 1998; Palka and Hammond, 2001).

In some cases, detection probabilities vary due to the individuals' behavior or their properties. For example, smaller individuals (or groups of individuals) are more easily missed than larger individuals. In the context of the remote sensing of forests (see Section 7), smaller trees are more likely to be missed than larger trees.

Distance sampling techniques have also been used to estimate the number of trees per hectare and basal area from a terrestrial laser scan (Ducey and Astrup, 2013; Astrup et al., 2014). In order to correct for the non-detection of trees, which also here is a serious issue, Astrup et al. (2014) assume that the detectability decreases with increasing distance from the location of the scanner.

6.3 Citizen science

For presence-only and citizen science data, spatial patterning in presence locations may result from the behavior of observers rather than the behavior of the study species, e.g., due to selection bias in study sites based on accessibility, false absences and false detections (Warton et al., 2013; Altwegg and Nichols, 2019; Sicacha-Parada et al., 2021a). More generally, biases related to specific data collection mechanisms are mitigated by merging data of different types (Dorazio, 2014; Altwegg and Nichols, 2019; Bowler et al., 2019; Renner et al., 2019; Miller et al., 2021; Watson et al., 2021; Martino et al., 2021).

7. MULTITYPE DEVIANCE

Several of the types of discrepancies mentioned above can occur when point patterns of trees are detected using remote sensing such as airborne laser scanning (ALS), aerial photographs (see Section 3.2) or hyperspectral data. Depending on how the data are collected and processed, some portion of trees may be missed, e.g., due to large trees blocking the visibility of smaller trees or due to a cluster of trees being erroneously observed as a single large tree. In addition, there may be displacement error related to measurement error or as a consequence of image geometry and lighting conditions. Furthermore, some number of ghost trees can be generated, e.g., by double counting or mistaking branches for tree tops.

7.1 Airborne laser scanning

In ALS, laser pulses are emitted from an instrument installed in an airplane and directed to a field plot (Vosselman and Maas, 2010). As a result, a 3D point cloud representing the return locations of the laser pulses is obtained. Individual tree detection (ITD) methods, often based on finding local maxima, are then used to locate tree tops (e.g., Roussel et al., 2020; Roussel and Auty, 2022). Bayesian methods utilizing certain tree canopy shapes have also been developed (e.g., Lahivaara et al., 2014). However, no matter how sophisticated the ITD method, errors do occur. Therefore, available ITD methods are not able to reconstruct the spatial structure of the forest (Packalen et al., 2013). A few approaches have been developed to correct for undetected trees in ITD methods (Kansanen et al., 2016; Mehtätalo et al., 2022; Kostensalo et al., 2023). Alternatively, it has been proposed to predict the spatial structure of the forest at stand level (Pippuri et al., 2012; Häbel et al., 2021).

Mehtätalo (2006) proposes a Horvitz–Thompson-like estimator to correct for trees missed in airborne laser scanning analyses. The methodology was further developed in Kansanen et al. (2016, 2021); Mehtätalo et al. (2022). In this approach, detected trees are weighted reflecting their detectability utilizing stochastic geometry and canopy size. The core idea is that large trees are always detected and thus are given weight 1. Smaller trees on the other hand are less likely to be detected and thus they have a weight larger than one, since the detected small trees are only some fraction of the actual number of small trees in the forest. Then one can estimate, e.g., the stand density with the equation

$$\widehat{n} = \sum_{i \in detected} \frac{1}{p_i},$$

where p_i is the detection probability for each detected tree.

Kostensalo et al. (2023) reconstruct tree patterns from ALS data. Optimized ITD approaches detect only the largest, dominant trees (and trees in sparse plots) and miss approximately 40% of the trees. In addition, there are about 10% false discoveries. Clustering is also problematic for ITD, as clustered trees are often detected as a single tree.

Once the optimized ITD is applied for plots with no ground measurements, a model is used to predict the number of missing trees as well as the number of false discoveries. The placement of the missing trees is based on a resampling-type approach, where a detected tree is selected at random, and a simulated compensating tree is based at a random angle from the detected tree at a distance sampled from the nearest neighbor distribution of the ground measurements. Finally, a proportion of the smallest detected trees are removed as false discoveries. The models for the number of missing trees and false discoveries is based on the plots with ground measurements available.

Application to one plot in the Mikkeli area in Southern Finland is presented in Fig. 3. The model has been trained with 163 forest plots and the plot is one of the cross-validation test plots. The L-functions with isotropic edge correction of the three point patterns of Fig. 3 are presented in Fig. 4, for the proposed (new) method with a 95% global envelope (Myllymäki et al., 2017; Myllymäki and Mrkvička, 2023) constructed from 1000 repeated simulations of the undetected trees. While the L-function of the ITD pattern is systematically too regular, the L-function of the proposed method is in excellent agreement with the L-function based on ground measurements. Thus, while the simulated trees might not be placed exactly where the missing trees are located, the created point pattern captures well the clustering and regularity of the forest plot in addition to the intensity.

7.2 Single terrestrial laser scanning

Terrestrial laser scanning (TLS) is based on a similar principle as ALS, but the laser pulses are emitted from a device that has a fixed location in the field (Vosselman

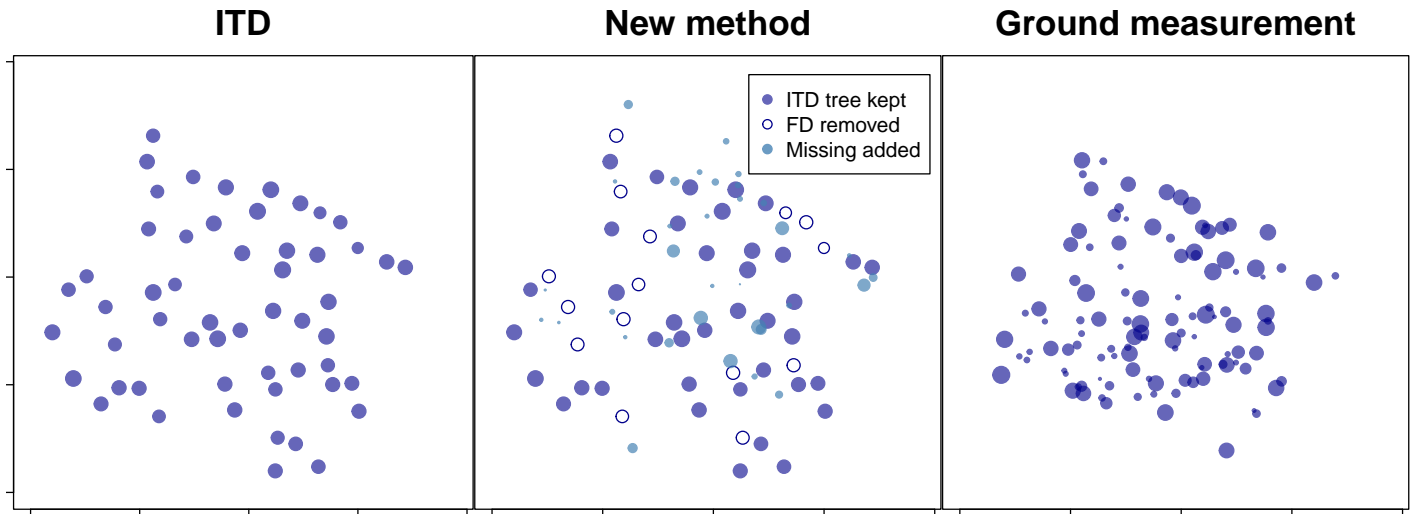**ITD**   **New method**   **Ground measurement**

FIG 3. *Left: The pattern of trees detected from ALS data by individual tree detection (ITD). Middle: Tree pattern created using the correction method suggested in* Kostensalo et al. (2023). *Right: Ground measurements of trees in an area of ca. 400 m^2 in Mikkeli, Finland.*
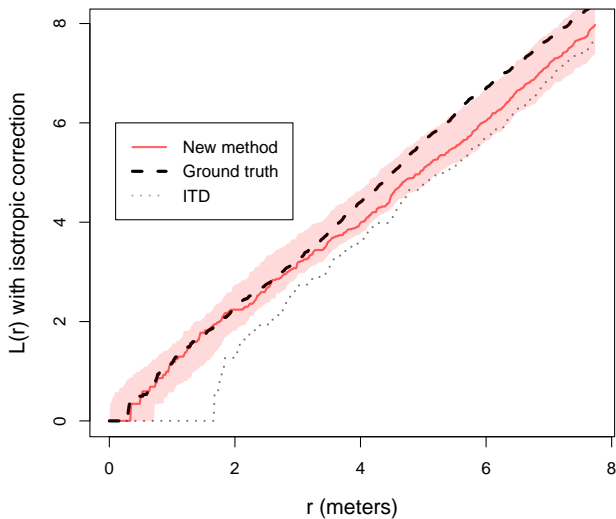

FIG 4. *L-function of the three point patterns displayed in Fig. 3. The solid red line of the new method corresponds to the particular simulation round shown in the central panel of Fig. 3 and the shaded area is the 95% global envelope estimated from 1000 repeated simulations.*

and Maas, 2010). If the forest is scanned from a single location, some of the trees are covered by the ones closer to the scanner and are not detected or are only partially detected. Multiple scans can be used to remove or almost remove the problem of occlusion of trees in a fixed area, but still the extraction of the locations of trees and tree characteristics remains challenging (Pitkänen et al., 2019). Typically, TLS data have been used to estimate only first-order characteristics of tree data such as density and total basal area (Kuronen et al., 2019; Kansanen et al., 2021; Ducey and Astrup, 2013) as well as to derive characteristics of trees (Pitkänen et al., 2021).

Kuronen et al. (2019) construct estimators for forest characteristics such as tree density or basal area based on the visible area of the scanner. They assume that the tree trunks and the cross-sections of tree trunks are perfect circles (no understory vegetation or tree branches affecting the visibility). To estimate the total basal area, the area of each tree is multiplied by a detection function which is positive only for the detected trees. In addition to the location and size of the tree, the detection function depends on the entire point pattern of the trees since the other trees affect whether a tree is detected. When the tree pattern is a realization of a Poisson process, the estimator works quite well but the choice of detection function affects the results. The best results were obtained with the detector, where a tree is detected if any part of it is detected. If the tree pattern is regular, the basal area is estimated larger than it is and this positive bias increases with the degree of regularity. If the tree pattern is clustered, the bias tends to be negative and its absolute value increases with the degree of clustering. Kansanen et al. (2021) propose an alternative estimator using sampling theory instead of point processes. Their estimator is slightly more accurate and has an analytic formula for variance.

8. OUTLIERS

An outlier in a spatial point pattern is a point (or points) that is different from the other points in some way. For example, it can be an isolated point or a pair of points that are much closer together than any other pair of points. We include a discussion of outliers, as these are commonly caused by contamination rather than the true data generating mechanism.

8.1 Outlier detection

Different methods for goodness-of-fit checking of a particular point process model and finding outliers have been presented in the literature. Stoyan and Grabarnik (1991) introduce exponential energy marks that can be used to detect outliers. The points x_i, $i = 1, \ldots, n$ of a point process X are marked by $m_i = 1/\lambda(x_i | \mathbf{x} \backslash \{x_i\})$, where $\lambda(\cdot | \cdot)$ is the Papangelou conditional intensity. Extreme values of the estimated mark may indicate outliers (with respect to the underlying model). A deviance residual for heterogeneous Poisson processes is defined by Lawson (1993). The resulting local intensity estimate can be used to find outlying points. Using the Papangelou conditional intensity, Baddeley et al. (2005) introduce residuals for spatial point processes that correspond to the usual residuals for Poisson log-linear regression. The residuals can be used to find outliers, for example isolated points or points with close neighbors. Baddeley et al. (2013) define diagnostics for detecting points with unusual prediction values or points having extraordinarily large influence on the model parameters. We note that the model-based residuals are also used for local model diagnostics (e.g. D'Angelo et al., 2022; Myllymäki et al., 2021). In addition to these model-based approaches, nonparametric functions such as distance to the nearest neighbor can be used to find outliers whether they are isolated points or close pairs of points. Other local statistics such as 'individual' L and pair-correlation functions (Stoyan and Stoyan, 1994, p. 290) or product density LISA functions (Anselin, 1995; Cressie and Collins, 2001) computed at the point locations can also be used to detect outliers.

8.2 Outliers due to imperfect image segmentation

Kuronen et al. (2021) regard sweat gland patterns extracted from a video as realizations of spatial point processes. In the patterns observed in the end of the videos, there are some unexpected close pairs of points (sweat glands). A closer inspection reveals that some of the sweat spots are detected as two nearby spots due to the several image analysis steps that are needed to extract the point patterns from the videos. By following some of the videos frame by frame, they conclude that some nearby pairs of spots are probably produced by a single gland.

An example of such an erroneous pair of sweat glands is given in Figure 5. To avoid the time consuming manual segmentation of the videos, (Kuronen et al., 2021) regard the falsely detected extra points as noise and introduce two ways to handle the noise. In the first approach, the observed pattern is modeled as a realization of a sequential point process, namely a mixture of a soft-core model for the true sweat gland locations and a uniform distribution for the noise points. In the second approach, the underlying unobserved true sweat gland pattern is generated first by using a disturbed simple sequential inhibition process
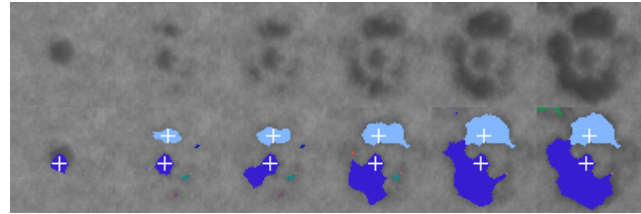

FIG 5. *Six frames of one sweat gland of a video (top row) and the result of the image segmentation (bottom row). The colored blobs indicate sweat produced by different identified sweat glands with locations indicated by pluses.*

which is then randomly thinned to obtain the observed pattern. The parameters are estimated using an approximate Bayesian computation algorithm (Beaumont, 2010) applied to summary statistics that are robust to the noise points.

To complement the analysis in (Kuronen et al., 2021), we further consider one of the sweat gland patterns (see Figure 6, left) as an example of a pattern with outlying points, or outlying point pairs. For each point, we compute the distance to its nearest neighbor and as a result identify eight extraordinarily close pairs of points (the first three bars of the histogram in Figure 7). It turns out that these outlier pairs are incorrectly identified sweat gland locations as shown in Figure 5. To see how these outlier pairs affect the analysis of the point pattern, we randomly remove one point in each close pair (red discs of Figure 6, left) and estimate the pair correlation function for the original and modified patterns. The behavior at small distances is clearly different in the patterns with and without the outliers (Figure 6, right).

Then, we investigate how the outlier pairs affect the parameter estimates of a model. Motivated by Kuronen et al. (2021), we use a softcore Gibbs point process model with the density function

$$f(\mathbf{x}) \propto \beta^{n(\mathbf{x})} \exp\left(-\sum_{i<j} \left(\frac{\sigma}{\|x_i - x_j\|}\right)^{2/\kappa}\right),$$

where $\beta > 0$ is an intensity related parameter, and $\sigma > 0$ and $\kappa \in (0, 1)$ interaction parameters. Large values of σ and small values of κ indicate strong interaction. The model is fitted to the original and modified patterns using spatstat (Baddeley et al., 2015). Especially, the interaction parameter σ is affected by the erroneous points (see Table 1).

8.3 Outlying clusters and patterns

In addition to outlying points or pairs of points, there can be an outlying cluster in a point pattern. Kuronen et al. (2022) study patterns of Norway spruce seedlings. Figure 8 shows one of the seedling patterns. There is tendency of seedlings to form small clusters. However, there is one extraordinarily large cluster in the middle of the plot. Such
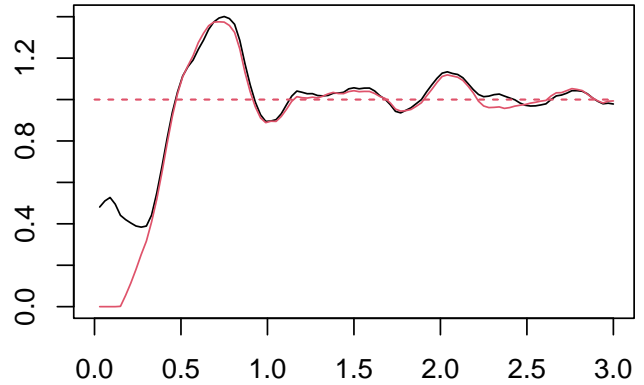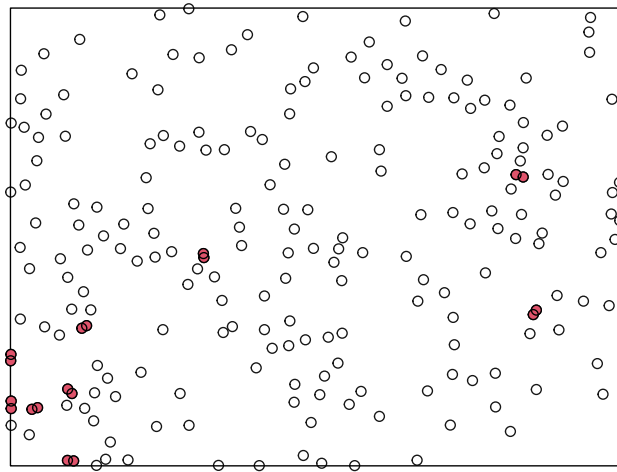
FIG 6. *Left: The detected sweat gland locations in a window of size 17.5×13 mm^2. The close pairs of gland locations identified by using the nearest neighbor distances are shown as red discs. Right: The pair-correlation function of the original pattern (black) and the pattern, where one of the points in each outlier pair has been removed (red).*
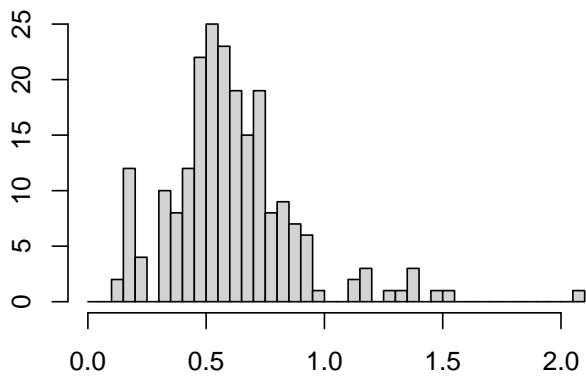

FIG 7. *Histogram of the nearest neighbor distances of the original pattern in Figure 6.*

TABLE 1

Pseudo-likelihood estimates of the parameters (95% confidence intervals) for softcore Gibbs process with all points and with outliers removed. $\kappa = 0.2$ was fixed.

	β	σ	κ
All points	0.95 (0.83, 1.09)	0.10 (0, 0.13)	0.2
Outliers removed	1.11 (0.95, 1.29)	0.30 (0.23, 0.32)	0.2

FIG 8. *Norway spruce seedlings in a forest stand in Finland.*

clusters are probably not errors (like in PALM data in Section 4.1) but real, although rare, observations. Therefore, the possibility of having such outlying clusters should be included in the model. One possibility would be to model the seedling data by a cluster model where the expected cluster size varies. In the simplest case, the expected number of points in a cluster can take two values, small and large, the large one having a small probability to occur. In general, some inform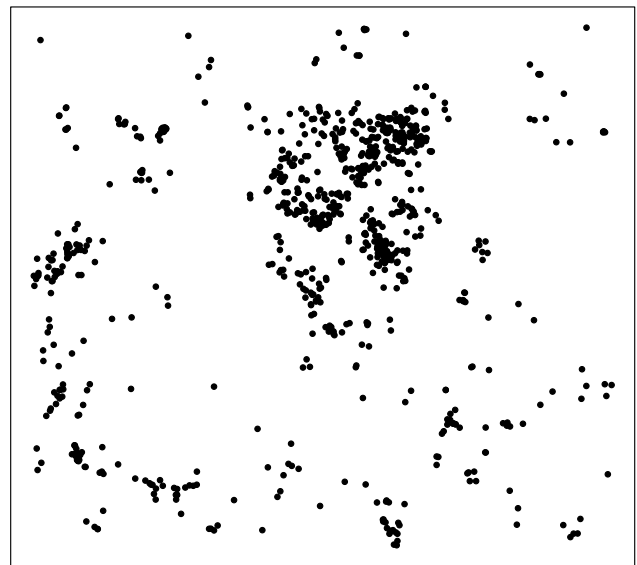ation on the underlying, e.g., soil properties would be essential for determining where such clusters are likely to appear. Construction of such models is left for future work.

Slightly related to outlying clusters, Coeurjolly (2017) introduce a median based estimation approach to estimate the intensity of stationary point processes which is robust to outlying areas in the data. An outlying area can be an area without any or with very few points or an area with many points (an outlying cluster) compared to the remaining area.

Finally, if repetitions are available, one of the patterns can be different from the others and considered an outlying pattern. Kuronen et al. (2021) analyze 15 sweat gland
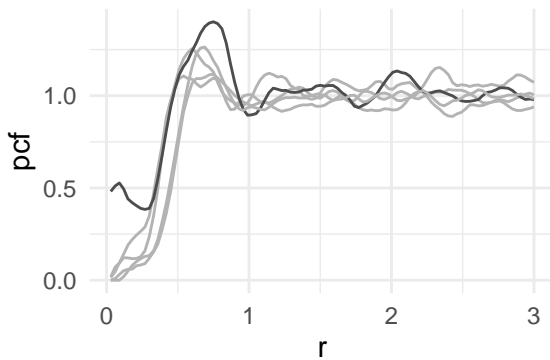
FIG 9. *Pair-correlation functions of activated sweat gland patterns from five healthy subjects.*

videos of which five are from healthy subjects. The estimated pair-correlation functions for the healthy patterns are shown in Figure 9. One of the patterns (black), also shown in Figure 6, shows a different behavior than the other patterns with more points within a short distance from each other. As pointed out earlier, some close pairs of points in this outlying pattern are incorrectly recorded as two sweat gland locations and by removing such erroneous points, the point pattern is no longer an outlying pattern, see Figure 9 and Figure 6 (right). However, if no errors are found, the possibility of outlying patterns should be included in the modeling approach.

9. DISCUSSION AND CONCLUSIONS

In this paper, we have reviewed observation mechanism that may lead to observed spatial point patterns that deviate from the true underlying point pattern of interest, as well as available statistical inference approaches to handle such deviations. The deviations include such diverse things as location errors, censoring, thinning, deformation and ghost patterns or points. Thus, a diverse set of models have been developed to handle these deviations and they are found scattered in various strands of literature, often published in application-specific journals.

Much of the work reviewed in this paper is relatively old, and the computational approaches used may no longer be available. Indeed, it is our experience that unless software is made publicly available, new methodology is very unlikely to be implemented by others. Software for some of the analyses discussed above are available in a GitHub repository at https://github.com/PointProcess/PointPatternDeviations. There are many software packages dealing with point pattern analysis, but few, if any, explicitly allow errors in the locations of points. It would be beneficial if package developers would include simple approaches, such as second order estimates for points with i.i.d. location errors.

Data types with deviant observation mechanisms are, on the other hand, constantly evolving. For example, the availability of image data has increased rapidly in recent years due to advances in satellite and drone technology, to the point where the large data quantities require the use of automated detection algorithms (Eckerstorfer et al., 2016). In some applications, see e.g. Eckerstorfer et al. (2016) and Salberg (2015), the algorithms have a high detection while they also tend to suffer from high false detection rates, in effect generating ghost points, which must be accounted for in the subsequent data analysis. An appropriate modeling of the detection error can provide valuable input for improving the detection algorithms, calling for a direct synergy between the machine learning detection techniques and the statistical models.

In this paper, we have restricted ourselves to errors in point locations, that is, in unmarked point patterns. Additional discrepancies may occur in the marks attached to the points, if relevant. For example, in the multitype case the points can be incorrectly classified, and in the case of real valued marks, the marks can be measured with error. However, we consider a comprehensive discussion of mark discrepancies outside of the scope of the current paper.

Similarly, we have not provided a comprehensive discussion of every aspect of the chosen topics, especially when the discussion requires leaving the spatial setting. For a thinned point process, for example, the thinning probability may be changing over space. This may, e.g., be due to varying sampling effort or spatially varying underreporting. A point pattern may have a gap where an unobserved process has thinned the points. A temporal example is given by Lee and Brillinger (1979), who studied historical Chinese earthquakes, where the probability of an earthquake being recorded depends on which imperial dynasty rules, and the value the dynasty sets on recording historical events. Their approach is parametric, while Guttorp and Thompson (1990) estimated the observation probability by smoothing the counting process. Both papers were analyzing the autointensity function, a common second order parameter function for temporal point processes.

Varying thinning probabilities are common in citizen science data, and can sometimes be modeled using covariates such as survey data or data on other species, e.g., Sicacha-Parada et al. (2021b). In the case where a point pattern has gaps due to a not completely known process, one possibility is to assume two overlapping random fields, one with (almost) zero intensity which operates where the empty areas are; this would also work if the deletion process is not deleting all the points, for example using a variant of the approach in Jones-Todd et al. (2018).

In conclusion, we have found that there is currently no systematic way in the literature to deal with point pattern deviations. This review illustrates many of the aspects that

would need to be included in such an approach, while others, including approaches to handle erroneous detection in automatic detection algorithms, have yet to be developed. We are thus still far from being able to produce an all-encompassing approach with the topic offering some interesting open research questions.

ACKNOWLEDGEMENTS

REFERENCES

Altwegg, R. and Nichols, J. D. (2019). Occupancy models for citizen-science data. *Methods in Ecology and Evolution*, 10(1):8–21.

Andersson, C., Guttorp, P., and Särkkä, A. (2016). Discovering early diabetic neuropathy from epidermal nerve fiber patterns. *Statistics in Medicine*, 35:4427–4442.

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2):93–115. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1538-4632.1995.tb00338.x.

Arce Guillen, R., Lindgren, F., Muff, S., Glass, T. W., Breed, G. A., and Schlaegel, U. E. (2023). Accounting for unobserved spatial variation in step selection analyses of animal movement via spatial random effects. *bioRxiv*, pages 2023–01.

Armstrong, M. P., Rushton, G., and Zimmerman, D. L. (1999). Geographically masking health data to preserve confidentiality. *Statistics in medicine*, 18(5):497–525.

Assunção, R. and Guttorp, P. (1999). Robustness in point processes. *Annals of the Institute of Statistical Mathematics*, 51:657–678.

Astrup, R., Ducey, M. J., Granhus, A., Ritter, T., and von Lüpke, N. (2014). Approaches for estimating stand-level volume using terrestrial laser scanning in a single-scan mode. *Canadian Journal of Forest Research*, 44(6):666–676.

Bachl, F. E., Lindgren, F., Borchers, D. L., and Illian, J. B. (2019). inlabru: an r package for bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution*, 10(6):760–766.

Baddeley, A., Chang, Y.-M., and Song, Y. (2013). Leverage and influence diagnostics for spatial point processes. *Scandinavian Journal of Statistics*, 40(1):86–104.

Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: methodology and applications with R*. CRC press.

Baddeley, A., Turner, R., Møller, J., and Hazelton, M. (2005). Residual analysis for spatial point processes. *Journal of the Royal Statistical Society B*, 67(5):617–666.

Bar-Hen, A., Chadoeuf, J., Dessard, H., and Monestiez, P. (2013). Estimating second order characteristics of point processes with known independent noise. *Statistics and Computing*, 23:297–309.

Beaumont, M. (2010). Approximate bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41:379–406.

Borchers, D. L., Zucchini, W., and Fewster, R. M. (1998). Mark-recapture models for line transect surveys. *Biometrics*, pages 1207–1220.

Bowler, D. E., Nilsen, E. B., Bischof, R., O'Hara, R. B., Yu, T. T., Oo, T., Aung, M., and Linnell, J. D. (2019). Integrating data from different survey types for population monitoring of an endangered species: the case of the Elås deer. *Scientific reports*, 9(1):1–14.

Brockner, B., Williams, K., Luke, R., Sheen, D., Dowdy, J., Anderson, D., and Veal, C. (2018). Convolutional neural network based side attack explosive hazard detection in three dimensional voxel radar. In *Detection and sensing of mines, explosive objects and obscured targets XXIII*, page 51.

Buckland, S. T. (2006). Point-transect surveys for songbirds: Robust methodologies. *The Auk*, 123(2):345–357.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2001). *Introduction to distance sampling: estimating abundance of biological populations*. Oxford (United Kingdom) Oxford Univ. Press.

Buckland, S. T., Rexstad, E. A., Marques, T. A., and Oedekoven, C. S. (2015). *Distance sampling: methods and applications*, volume 431. Springer.

Byers, S. and Raftery, A. (1998). Nearest-neighbor clutter removal for estimating features in spatial point processes. *Journal of the American Statistical Association*, 93:577–584.

Chakraborty, A. and Gelfand, A. (2010). Analyzing spatial point patterns subject to measurement error. *Bayesian Analysis*, 5(1):97–122.

Chandler, R. B., Royle, J. A., and King, D. I. (2011). Inference about density and temporary emigration in unmarked populations. *Ecology*, 92(7):1429–1435.

Chang, X., Waagepetersen, R., Yu, H., Ma, X., Holford, T. R., Wang, R., and Guan, Y. (2015). Disease risk estimation by combining case-control data with aggregated information on the population at risk. *Biometrics*, 71(1):114–121.

Coeurjolly, J.-F. (2017). Median-based estimation of the intensity of a spatial point process. *Annals of the Institute of Statistical Mathematics*, 69:303–331.

Connor, B., Carrie, I., Craig, R., and Parsons, J. (2008). Discriminative imaging using LWIr polarimetric imaging. In *Proceedings of the SPIE Security and Defence Conference*.

Cressie, N. and Collins, L. B. (2001). Analysis of spatial point patterns using bundles of product density LISA functions. *Journal of Agricultural, Biological, and Environmental Statistics*, 6(1):118.

Cressie, N. and Lawson, A. (2000). Hierarchical probability models and Bayesian analysis of mine locations. *Advances in Applied Probability*, 32:315–330.

Cucala, L. (2008). Intensity estimation for spatial point processes observed with noise. *Scand. J. Statist.*, pages 322–334.

Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes Volume I: elementary theory and methods*. Springer.

Dasgupta, A. and Raftery, A. (1998). Detecting features in spatial point processes with clutter via model-based clustering. *Journal of the American Statistical Association*, 93:294–302.

Diggle, P. (1985). A kernel method for smoothing point process data. *J. Roy. Statist. Soc. Ser. C*, 34:138–147.

Diggle, P. (1993). Point process modelling in environmental epidemiology. In Barnett, V. and Turkman, F., editors, *Statistics for the environment*, pages 89–110. Wiley.

Dorazio, R. M. (2014). Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. *Global Ecology and Biogeography*, 23(12):1472–1484.

Dralle, K. and Rudemo, M. (1996). Stem number estimation by kernel smoothing of aerial photos. *Canadian Journal of Forest Research*, 26:1228–1236.

Dralle, K. and Rudemo, M. (1997). Automatic estimation of individual tree positions from aerial photos. *Canadian Journal of Forest Research*, 27:1728–1736.

Ducey, M. J. and Astrup, R. (2013). Adjusting for nondetection in forest inventories derived from terrestrial laser scanning. *Canadian Journal of Remote Sensing*, 39(5):410–425.

D'Angelo, N., Siino, M., D'Alessandro, A., and Adelfio, G. (2022). Local spatial log-Gaussian Cox processes for seismic data. *AStA Advances in Statistical Analysis*, 106(4):633–671.

Eckerstorfer, M., Bühler, Y., Frauenfelder, R., and Malnes, E. (2016). Remote sensing of snow avalanches: Recent advances, potential, and limitations. *Cold Regions Science and Technology*, 121:126–140.

Fanshawe, T. and Diggle, P. (2011). Spatial prediction in the presence of positional error. *Environmetrics*, 22(2):109–122.

Fronterrè, C., Giorgi, E., and Diggle, P. (2018). Geostatistical inference in the presence of geomasking: a composite-likelihood approach. *Spatial Statistics*, 28:319–330.

Glennie, R., Buckland, S. T., Langrock, R., Gerrodette, T., Ballance, L., Chivers, S., and Scott, M. (2021). Incorporating animal movement into distance sampling. *Journal of the American Statistical Association*, 116(533):107–115.

Guttorp, P. (1995). *Stochastic modeling of scientific data*. Chapman & Hall.

Guttorp, P. and Thompson, M. L. (1990). A note on point process analysis of Chinese earthquake data. *Pure Appl. Geophys.*, 133:541–546.

Guttorp, P. and Thorarinsdottir, T. L. (2012). What happened to discrete chaos, the Quenouille process, and the sharp Markov property? Some history of stochastic point processes. *International Statistical Review*, 80(2):253–268.

Häbel, H., Balazs, A., and Myllymäki, M. (2021). Spatial analysis of airborne laser scanning point clouds for predicting forest structure. *Mathematical & Computational Forestry & Natural Resource Sciences*, 13(1):15–28.

Hable, S., Sigloch, K., Barruol, G., Stähler, S., and Hadziioannou, C. (2018). Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations. *Geophysical Journal International*, 214(3):2014–2034.

Hayes, R. and Buckland, S. (1983). Radial-distance models for the line-transect method. *Biometrics*, pages 29–42.

Hedley, S. L. and Buckland, S. T. (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics*, 9(2):181–199.

Högmander, H. (1991). A random field approach to transect counts of wildlife populations. *Biometrical journal*, 33:1013–1023.

Holmes, Q., Schwartz, C., Seldin, J., Wright, J., and Witter, L. (1995). Adaptive multispectral CFAR detection of land mines. In Dubey, A. C., Cindrich, I., Ralston, J. M., and Rigano, K. A., editors, *Detection Technologies for Mines and Minelike Targets*, volume 2496 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 421–432.

Huang, H., Ma, X., Waagepetersen, R., Holford, T. R., Wang, R., Risch, H., Mueller, L., and Guan, Y. (2014). A new estimation approach for combining epidemiological data from multiple sources. *Journal of the American Statistical Association*, 109(505):11–23. PMID: 24683281.

Husen, S., Kissling, E., Deichmann, N., Wiemer, S., Giardini, D., and Baer, M. (2003). Probabilistic earthquake location in complex three-dimensional velocity models: Application to Switzerland. *J.Geophys. Res.*, 108.

Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*. Wiley.

Jacquez, G. M. (2012). A research agenda: does geocoding positional error matter in health GIS studies? *Spatial and spatio-temporal epidemiology*, 3(1):7–16.

Jensen, L. G., Williamson, D., and Hahn, U. (2022). Semiparametric point process modeling of blinking artifacts in PALM. *The Annals of Applied Statistics*, 16:1500–1523.

Johnson, D.S.and Hooten, M. and Kuhn, C. (2013). Estimating animal resource selection from telemetry data using point process models. *Journal of Animal Ecology*, 82:1155–64.

Jones-Todd, C. M., Caie, P., Illian, J. B., Stevenson, B. C., Savage, A., Harrison, D. J., and Bown, J. L. (2019). Identifying prognostic structural features in tissue sections of colon cancer patients using point pattern analysis. *Statistics in medicine*, 38(8):1421–1441.

Jones-Todd, C. M., Swallow, B., B., I. J., and Toms, M. (2018). A spatiotemporal multispecies model of a semicontinuous response. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 67:705–722.

Kagan, Y. Y. and Knopoff, L. (1980). Spatial distribution of earthquakes: The two point correlation function. *Geophys. J. R. Astron. Soc.*, 62:303–320.

Kansanen, K., Packalen, P., Maltamo, M., and Mehtätalo, L. (2021). Horvitz-Thompson-like estimation with distance-based detection probabilities for circular plot sampling of forests. *Biometrics*, 77(2):715–728.

Kansanen, K., Vauhkonen, J., Lähivaara, T., and Mehtátalo, L. (2016). Stand density estimators based on individual tree detection and stochastic geometry. *Canadian Journal of Forest Research*, 46(11):1359–1366.

Konstantinou, K. and Särkkä, A. (2021). Spatial modeling of epidermal nerve fiber patterns. *Statistics in Medicine*, 40:6479–6500.

Kostensalo, J., Mehtátalo, L., Tuominen, S., Packalen, P., and Myllymäki, M. (2023). Recreating structurally realistic tree maps with airborne laser scanning and ground measurements. *Remote Sensing of Environment*, 298:113782.

Kuronen, M., Henttonen, H., and Myllymäki, M. (2019). Correcting for nondetection in estimating forest characteristics from single-scan terrestrial laser measurements. *Canadian Journal of Forest Research*, 46:96–103.

Kuronen, M., Myllymäki, M., Loavenbruck, A., and Särkkä, A. (2021). Point process models for sweat gland activation observed with noise. *Statistics in Medicine*, 40:2055–2072.

Kuronen, M., Särkkä, A., Vihola, M., and Myllymäki, M. (2022). Hierarchical log Gaussian Cox process for regeneration in uneven-aged forests. *Environmental and Ecological Statistics*, 40(29):185–205.

Laake, J. and Borchers, D. (2004). *Methods for incomplete detection at distance zero*, pages 108–189. Oxford University Press.

Lahivaara, T., Seppanen, A., Kaipio, J. P., Vauhkonen, J., Korhonen, L., Tokola, T., and Maltamo, M. (2014). Bayesian approach to tree detection based on airborne laser scanning data. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5):2690–2699.

Langley, R. (1997). The GPS error budget. *GPS World*, 8:51–56.

Larsen, M. and Rudemo, M. (1998). Optimizing templates for finding trees in aerial photographs. *Pattern Recognition Letters*, 19:1153–1162.

Law, R., Illian, J., Burslem, D. F., Gratzer, G., Gunatilleke, C., and Gunatilleke, I. (2009). Ecological information from spatial patterns of plants: insights from point process theory. *Journal of Ecology*, 97(4):616–628.

Lawson, A. (1993). A deviance residual for heterogeneous spatial Poisson processes. *Biometrics*, 49:889–897.

Lee, W. H. K. and Brillinger, D. R. (1979). On Chinese earthquake history–an attempt to model an incomplete data set by point process analysis. *Pure Appl. Geophys.*, 117:1229–1257.

Lund, J., Penttinen, A., and Rudemo, M. (1999). Bayesian analysis of spatial point patterns from noisy observations. *In the Ph.D. thesis Statistical inference and perfect simulation for point processes observed with noise by J Lund, Department of Mathematics and Physics, Royal Veterinary and Agricultural University, Denmark.*

Lund, J. and Rudemo, M. (2000). Models for point processes observed with noise. *Biometrika*, 87(2):235–249.

Lund, J. and Thönnes, E. (2004). Perfect simulation and inference for point processes given noisy observations. *Computational statistics*, 19:317–336.

Martens, A. (1968). Errors in measurement and counting of particles using light scattering. *Journal of the Air Pollution Control Association*, 18:661–663.

Martino, S., Pace, D. S., Moro, S., Casoli, E., Ventura, D., Frachea, A., Silvestri, M., Arcangeli, A., Giacomini, G., Ardizzone, G., et al. (2021). Integration of presence-only data from several sources: a case study on dolphins' spatial distribution. *Ecography*, 44(10):1533–1543.

Mehtätalo, L. (2006). Eliminating the effect of overlapping crowns from aerial inventory estimates. *Canadian Journal of Forest Research*, 36(7):1649–1660.

Mehtätalo, L., Yazigi, A., Kansanen, K., Packalen, P., Lähivaara, T., Maltamo, M., Myllymäki, M., and Penttinen, A. (2022). Estimation of forest stand characteristics using individual tree detection, stochastic geometry and a sequential spatial point process model. *International Journal of Applied Earth Observation and Geoinformation*, 112:102920.

Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press.

Miller, D. L., Fifield, D., Wakefield, E., and Sigourney, D. B. (2021). Extending density surface models to include multiple and double-observer survey data. *PeerJ*, 9:e12113.

Milne, S., Martin, J. G., Reynolds, G., Vairappan, C. S., Slade, E. M., Brodie, J. F., Wich, S. A., Williamson, N., and Burslem, D. F. (2021). Drivers of bornean orangutan distribution across a multiple-use tropical landscape. *Remote Sensing*, 13(3):458.

Møller, J. and Waagepetersen, R. P. (2007). Modern statistics for spatial point processes. *Scandinavian Journal of Statistics*, 34(4):643–684.

Myllymäki, M. and Mrkvička, T. (2023). GET: Global envelopes in R. arXiv:1911.06583 [stat.ME].

Myllymäki, M., Mrkvička, T., Seijo, H., Grabarnik, P., and Hahn, U. (2017). Global envelope tests for spatial processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(2):381–404.

Myllymäki, M., Kuronen, M., and Mrkvička, T. (2021). Testing global and local dependence of point patterns on covariates in parametric models. *Spatial Statistics*, 42:100436.

Nordin, M., Abrahamsson, C., Hamngren Blomqvist, C., Röding, M., Olsson, E., Nydén, M., and Rudemo, M. (2014). Estimation of mass thickness response of embedded aggregated silica nanospheres from high angle annular dark-field scanning transmission electron micrographs. *Journal of Microscopy*, 253:166–170.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50:379–402.

Ogata, Y. and Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectonophysics*, 413:13–23.

Olsbo, V., Myllymäki, M., Waller, L. A., and Särkkä, A. (2013). Development and evaluation of spatial point process models for epidermal nerve fibers. *Mathematical Biosciences*, 243(2):178–189.

Packalen, P., Vauhkonen, J., Kallio, E., Peuhkurinen, J., Pitkänen, J., Pippuri, I., Strunk, J., and Maltamo, M. (2013). Predicting the spatial pattern of trees by airborne laser scanning. *International Journal of Remote Sensing*, 34(14):5154–5165.

Palka, D. and Hammond, P. (2001). Accounting for responsive movement in line transect estimates of abundance. *Canadian journal of fisheries and aquatic sciences*, 58(4):777–787.

Pippuri, I., Kallio, E., Maltamo, M., Peltola, H., and Packalén, P. (2012). Exploring horizontal area-based metrics to discriminate the spatial pattern of trees and need for first thinning using airborne laser scanning. *Forestry*.

Pitkänen, T. P., Raumonen, P., and Kangas, A. (2019). Measuring stem diameters with TLS in boreal forests by complementary fitting procedure. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147:294–306.

Pitkänen, T. P., Raumonen, P., Liang, X., Lehtomäki, M., and Kangas, A. (2021). Improving TLS-based stem volume estimates by field measurements. *Computers and Electronics in Agriculture*, 180:105882.

Rajala, T., Redenbach, C., Särkkä, A., and Sormani, M. (2016). Variational Bayes approach for classification of points in superpositions of point processes. *Spatial Statistics*, 15:85–99.

Ranacher, P., Brunauer, R., Trutschnig, W., Van der Spek, S., and Reich, S. (2016). Why GPS makes distances bigger than they are. *International Journal of Geographical Information Science*, 30(2):316–333.

Redenbach, C., Särkkä, A., Freitag, J., and Schladitz, K. (2009). Anisotropy analysis of pressed point processes. *Advances in Statistical Analysis*, 93(3):237–261.

Redenbach, C., Särkkä, A., and Sormani, M. (2015). Classification of points in superpositions of Strauss and Poisson processes. *Spatial Statistics*, 12:81–95.

Regener, E. (1909). Über Zählung der α-Theilchen durch die Szintillation und über die Größe der elektrische Elementarquantums. *Sitz. Ber. d. K. Preuss. Akad*, pages 948–965.

Renner, I. W., Elith, J., Baddeley, A., Fithian, W., Hastie, T., Phillips, S. J., Popovic, G., and Warton, D. I. (2015). Point process models for presence-only analysis. *Methods in Ecology and Evolution*, 6(4):366–379.

Renner, I. W., Louvrier, J., and Gimenez, O. (2019). Combining multiple data sources in species distribution models while accounting for spatial dependence and overfitting with combined penalized likelihood maximization. *Methods in Ecology and Evolution*, 10(12):2118–2128.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V. (1987). Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond. A*, 410:269–288.

Roussel, J.-R. and Auty, D. (2022). *lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications*. R package version 4.0.1.

Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R., Meador, A. S., Bourdon, J.-F., de Boissieu, F., and Achim, A. (2020). lidR: An R package for analysis of airborne laser scanning (ALS) data. *Remote Sensing of Environment*, 251:112061.

Royle, J. A. and Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, 87(4):835–841.

Rutherford, E., Geiger, H., and Bateman, H. (1910). LXXVI. The probability variations in the distribution of α-particles. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 20:698–707.

Salberg, A.-B. (2015). Detection of seals in remote sensing images using features extracted from deep convolutional neural networks. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1893–1896. IEEE.

Schmidt, J. H., Thompson, W. L., Wilson, T. L., and Reynolds, J. H. (2022). Distance sampling surveys: using components of detection and total error to select among approaches. *Wildlife Monographs*, 210(1):e1070.

Schneider, M. (2021). *Improving Uncertainty Quantification and Visualization for Spatiotemporal Earthquake Rate Models for the Pacific Northwest*. PhD thesis, University of Washington.

Sicacha-Parada, J., Steinsland, I., Cretois, B., and Borgelt, J. (2021a). Accounting for spatial varying sampling effort due to accessibility in citizen science data: A case study of moose in Norway. *Spatial Statistics*, 42:100446.

Sicacha-Parada, J., Steinsland, I., Cretois, B., and Borgelt, J. (2021b). Accounting for spatial varying sampling effort due to accessibility in citizen science data: A case study of moose in Norway. *Spatial Statistics*, 42:100446.

Stoyan, D. (1982). A remark on the line transect method. *Biometrical Journal*, 24(2):191–195.

Stoyan, D. and Grabarnik, P. (1991). Second-order characteristics for stochastic structures connected with Gibbs point processes. *Mathematische Nachrichten*, 151:95–100.

Stoyan, D. and Penttinen, A. (2000). Recent applications of point process methods in forestry statistics. *Statistical Science*, 15:61–78.

Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes, and Point Fields : Methods of Geometrical Statistics*. Wiley, Chichester, New York.

Thành, N., Hào, D., and Sahli, H. (2009). Infrared thermography for land mine detection. In Hammoud, R., editor, *Augmented Vision Perception in Infrared: Algorithms and Applied Systems*, pages 3–36. Springer.

van Lieshout, M. (2000). *Markov point processes and their applications*. Imperial College Press.

Vosselman, G. and Maas, H.-G. (2010). *Airborne and terrestrial laser scanning*. Whittles Publishing.

Waagepetersen, R. and Schweder, T. (2006). Likelihood-based inference for clustered line transect data. *Journal of Agricultural, Biological, and Environmental Statistics*, 11:264–279.

Warton, D. I., Renner, I. W., and Ramp, D. (2013). Model-based control of observer bias for the analysis of presence-only data in ecology. *PloS One*, 8(11):e79168.

Watson, J., Joy, R., Tollit, D., Thornton, S. J., and Auger-Méthé, M. (2021). Estimating animal utilization distributions from multiple data types: A joint spatiotemporal point process framework. *The Annals of Applied Statistics*, 15(4):1872–1896.

Williams, D. M., Dechen Quinn, A., and Porter, W. F. (2012). Impact of habitat-specific GPS positional error on detection of movement scales by first-passage time analysis. *PloS One*, 7(11):e48439.

Williamson, L. D., Scott, B. E., Laxton, M., Illian, J. B., Todd, V. L. G., Miller, P. I., and Brookes, K. L. (2022). Comparing distribution of harbour porpoise using generalized additive models and hierarchical Bayesian models with integrated nested Laplace approximation. *Ecological Modelling*, 470:110011.

Yuan, Y., Bachl, F. E., Lindgren, F., Borchers, D. L., Illian, J. B., Buckland, S. T., Rue, H., and Gerrodette, T. (2017). Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. *The Annals of Applied Statistics*, 11(4):2270–2297.

Zhang, H., Chon, C., Pan, X., and Li, D. (2009). Methods for counting particles in microfluidic applications. *Microfluid Nanofluid*, 7:739–749.

Zhang, W., Chipperfield, J. D., Illian, J. B., Dupont, P., Milleret, C., de Valpine, P., and Bischof, R. (2023). A flexible and efficient b ayesian implementation of point process models for spatial capture–recapture data. *Ecology*, 104(1):e3887.

Zimmerman, D. L., Li, J., and Fang, X. (2010). Spatial autocorrelation among automated geocoding errors and its effects on testing for disease clustering. *Statistics in medicine*, 29(9):1025–1036.