

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Yazigi, Adil; Penttinen, Antti; Ylitalo, Anna-Kaisa; Maltamo, Matti; Packalen, Petteri; Mehtätalo, Lauri

Title: Modeling Forest Tree Data Using Sequential Spatial Point Processes

Year: 2022

Version: Published version

Copyright: © 2021 The Author(s)

Rights: CC BY 4.0


Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Yazigi, A., Penttinen, A., Ylitalo, A.-K., Maltamo, M., Packalen, P., & Mehtätalo, L. (2022). Modeling Forest Tree Data Using Sequential Spatial Point Processes. *Journal of Agricultural, Biological, and Environmental Statistics*, 27(1), 88-108. <https://doi.org/10.1007/s13253-021-00470-2>



Modeling Forest Tree Data Using Sequential Spatial Point Processes

Adil YAZIGI , Antti PENTTINEN, Anna-Kaisa YLITALO, Matti MALTAMO, Petteri PACKALEN, and Lauri MEHTÄTALO

The spatial structure of a forest stand is typically modeled by spatial point process models. Motivated by aerial forest inventories and forest dynamics in general, we propose a sequential spatial approach for modeling forest data. Such an approach is better justified than a static point process model in describing the long-term dependence among the spatial location of trees in a forest and the locations of detected trees in aerial forest inventories. Tree size can be used as a surrogate for the unknown tree age when determining the order in which trees have emerged or are observed on an aerial image. Sequential spatial point processes differ from spatial point processes in that the realizations are ordered sequences of spatial locations, thus allowing us to approximate the spatial dynamics of the phenomena under study. This feature is useful in interpreting the long-term dependence and spatial history of the locations of trees. For the application, we use a forest data set collected from the Kiihtelysvaara forest region in Eastern Finland.

Key Words: Functional summary statistics; History-dependent model; Maximum likelihood; Ordered sequence; Spatial point processes.

1. INTRODUCTION

This paper discusses the modeling of sequential spatial point processes in the context of forest tree data. Our main motivation arises from issues in identifying individual trees in a forest stand inventories using airborne laser scanning (ALS) (see, e.g., [Maltamo et al. 2014](#)),

A. Yazigi (✉) School of Computing, University of Eastern Finland, PO BOX 111, 80101 Joensuu, Finland
(E-mail: adil.yazigi@uef.fi).

A. Penttinen, Department of Mathematics and Statistics, University of Jyväskylä, PL 35, 40014 Jyväskylä, Finland
(E-mail: antti.k.penttinen@jyu.fi).

A.-K. Ylitalo, Natural Resources Institute Finland (Luke), PL 2, 00791 Helsinki, Finland
(E-mail: anna-kaisa.ylitalo@luke.fi).

M. Maltamo · P. Packalen, School of Forest Sciences, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland (E-mail: matti.maltamo@uef.fi)

P. Packalen (E-mail: petteri.packalen@uef.fi).

L. Mehtätalo, Bioeconomy and Environment Unit, Natural Resources Institute Finland (Luke), Yliopistokatu 6 B, 80100 Joensuu, Finland
(E-mail: lauri.mehtatalo@luke.fi).

© 2021 The Author(s)

Journal of Agricultural, Biological, and Environmental Statistics

<https://doi.org/10.1007/s13253-021-00470-2>

where individual tree crowns are recorded by an aircraft. However, small trees often remain undetected below the crowns of bigger trees. A natural approach to such data involves ordering the detected trees according to size, then taking the trees hidden by larger tree crowns into account when estimating the number of trees in each size class (Kansanen et al. 2016; Mehtätalo 2006). A similar model could be considered for the more general modeling of forest stand dynamics and competition among trees (Pommerening and Grabarnik 2019).

Within this context, spatiotemporal point process (STPP) models could be proposed as candidates to model the intensity in question (see Cressie 1993; Daley and Vere-Jones 2003; Diggle 2013; González et al. 2016; Jensen et al. 2007). Roughly speaking, a STPP is a collection of instantaneous events, each occurring at a given spatial location x_i , with a given associated time t_i , namely $N_{st} = \{(x_i, t_i), i = 1, \dots, n\}$ for an integer $n > 0$. The temporal aspect of STPP offers a natural ordering of the points, which does not exist in the spatial dimension. In fact, the tree locations in a forest stand are fundamentally time dependent, as all the trees in a forest emerge gradually over time. It is worth noting that a one-to-one correspondence can be established between the class of the STPPs and the class of the marked point processes (MPPs) with positive quantitative marks. A MPP model can be transformed into a STPP by constructing the occurrence times based on the ordered realizations of the mark variable (Møller et al. 2016), whereas a STPP can be identified as MPP by considering event times t_i as marks (Daley and Vere-Jones 2008; Stoyan et al. 2017; Vere-Jones 2009).

In practice, the ages of trees are often unknown and hard to determine in the field. Therefore, the STPP approach does not appear at first glance to be applicable to forest tree data. Nevertheless, it is often sufficient to incorporate knowledge concerning the ordering of the trees with respect to their age. In particular, if the sample plot of interest is homogeneous in terms of growing conditions, the tree order with respect to age can be well approximated by their order with respect to size (Møller et al. 2016) given a suitable size variable, such as tree height, stem diameter or tree crown dimensions. In this case, the size distribution has to be involved in the model and the correlation between the size variable and spatial locations should be taken into account. In Møller et al. (2016), this correlation leads to the adoption of two distinct models to cover the potential dependence or independence of the size variable from spatial locations.

In this work, we employ a spatial point process model that takes the ordering of the trees into account and where the time component is disclosed as auxiliary information, namely in a *sequential spatial point process* (SSPP). Realizations of the SSPP are ordered sequences of spatial locations, which we denote by $\vec{x}_n = (x_1, \dots, x_n)$, for an integer $n > 0$, see Lieshout (2006a,b) and Lieshout and Capasso (2009). Other constructions of the SSPP model can be found in the works of Evans (1993) and Talbot et al. (2000), known as a random sequential absorption model, and Diggle et al. (1976) and Lieshout (2006b,c), where it is referred to as simple sequential inhibition. Recently, a significant expansion of the SSPP model for ordered spatial point patterns was proposed by Penttinen and Ylitalo (2016) motivated by eye movement data.

In this paper, we present applications and extensions of the model in Penttinen and Ylitalo (2016) for forest tree data. An essential property of these extensions is their *self-interaction* feature: Given the past observations x_1, \dots, x_k , the information carried by a new

point x_{k+1} is accommodated by the model, which determines the probability law. We apply these models to forest data consisting of tree locations associated with quantitative marks representing the diameters of the tree stems at breast height (DBH) and providing a natural ordering for the trees. The empirical data include three sample plots selected subjectively from a larger set of plots located in Kiihtelysvaara, Eastern Finland. In our study, we assume that all trees are generated by the same model. The self-interaction parameters for the model will be estimated via maximum likelihood, and, to evaluate the model fit, we employ several summary statistics assisted by Monte Carlo simulations. In particular, we develop new dynamic summary statistics that measure different features of the tree data, take the temporal order into account, and are meaningful for forest tree data.

2. FINITE SEQUENTIAL SPATIAL POINT PROCESS MODELS

2.1. MODEL

In a bounded window $W \subset \mathbb{R}^2$, let the observed data be a set of points (x_i, m_i) , $i = 1, \dots, n$, where n is a finite integer. Here, x_i indicates the tree location, and m_i is the corresponding mark representing the size of the tree, such as the diameter at breast height (DBH). We suppose that the following order $0 < m_n < m_{n-1} < \dots < m_2 < m_1 < \infty$ holds where m_1 and m_n are the sizes of the largest and the smallest tree, respectively.

From the observed data, we can write the ordered sequence of locations $\vec{x}_n = (x_1, \dots, x_n)$ with respect to the order of the m_i 's. Following Lieshout (2006a,b), the sequence \vec{x}_n is considered to be a realization of a SSPP with $\vec{x}_n \in W^n$. For $n = 0$, we have $\vec{x}_0 = \emptyset$. In order to write the distribution of the SSPP, we let (W^n, W^n) be the n -dimensional space of the ordered sequence with length $n > 0$ equipped with the Borel σ -algebra W^n and Lebesgue measure. The joint density of $\vec{x}_n = (x_1, \dots, x_n)$ with respect to Lebesgue measure is given by conditional densities of the following form:

$$f(\vec{x}_n) = f_1(x_1) \prod_{k=1}^{n-1} f_{k+1}(x_{k+1} | \vec{x}_k), \quad (1)$$

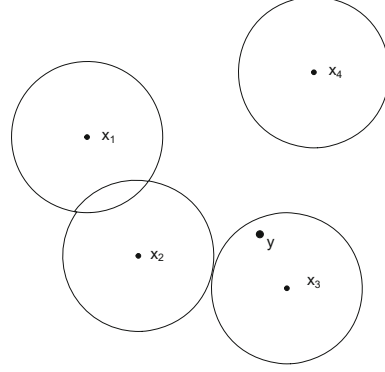
where f_1 indicates the density of the first spatial object x_1 , and $f_{k+1}(x_{k+1} | \vec{x}_k)$ is the density of a new object x_{k+1} conditional on the past sequence \vec{x}_k . Eventually, through consecutive conditioning, the distribution of a given location $x_{k+1} \in \vec{x}_n$, $1 \leq k < n$, depends on the entire past spatial history \vec{x}_k . Therefore, this enables the SSPP model to preserve the spatial information built up along the sequence, reflecting the long-term spatial dependence between the ordered locations of the trees and the spatial memory formed by this underlying dynamic. In order to give the density (1) an explicit form to model this dependence, we define the *zone of interaction* $B(x_k, r)$ (ZOI) for each tree k located at x_k , $k \in \{2, \dots, n-1\}$, driven by the radius $r > 0$ (*radius of the zone of interaction*). We introduce the *lagged clustering measure*

$$\mathcal{S}_k(y; r) = \sum_{i=1}^k \mathbf{1}_{B(x_i, r)}(y) \quad (2)$$

Figure 1. An illustration of the lagged clustering measure given a sequence

$\vec{\mathbf{x}}_4 = (x_1, x_2, x_3, x_4)$ of length 4 and a radius r . Here,

$\mathcal{S}_4(y; r) = 1$, and $\mathcal{S}_3(x_4; r) = \mathcal{S}_2(x_3; r) = \mathcal{S}_1(x_2; r) = 0$.



which counts the number of earlier balls $B(x_i, r)$, $i = 1, \dots, k$, containing the new point y (see the illustration in Fig. 1).

Based on this measure, the *self-interaction function* of the SSPP model denoted by $\pi_k(y, r)$ is represented as the reweighting probability

$$\pi_k(y, r) = \begin{cases} \theta & \text{if } \mathcal{S}_k(y; r) \geq 1, \\ 1 - \theta & \text{if } \mathcal{S}_k(y; r) = 0, \end{cases} \quad (3)$$

where $\theta \in (0, 1)$. Now, given the past $\vec{\mathbf{x}}_k$, the SSPP model at the new location y can be formulated by writing the conditional density in (1) as:

$$f_{k+1}(y | \vec{\mathbf{x}}_k) \propto \pi_k(y, r), \quad (4)$$

which, in compact form, is

$$f_{k+1}(y | \vec{\mathbf{x}}_k) \propto \theta \mathbf{1}_{\{\mathcal{S}_k(y; r) \geq 1\}}(y) + (1 - \theta) \mathbf{1}_{\{\mathcal{S}_k(y; r) = 0\}}(y).$$

The self-interaction function (3) first appeared in [Penttinen and Ylitalo \(2016\)](#) to model eye movements. Under (4), the parametric SSPP model with parameters (θ, r) is considered to be a special case since it is mainly based on the self-interaction function. Accordingly, the model is *history-dependent* and takes the full past history into account to examine the long-term dependency of y on the $\vec{\mathbf{x}}_k$ with the component $\mathcal{S}_k(y; r)$, which is controlled by the parameters (θ, r) . In fact, the parameter θ is the weight of the probability that the location y lies inside at least one of the balls $B(x_i, r)$, $i = 1, \dots, k$, formed by the sequence of the past locations $\vec{\mathbf{x}}_k$, while the weight $1 - \theta$ is when y lies outside of all the balls $B(x_i, r)$, $i = 1, \dots, k$. In other words, if p is the probability that a uniformly chosen random location falls inside the interaction zone of the previous trees, the reweighted probability is given by $\frac{p^\theta}{p^\theta + (1-p)(1-\theta)}$. Thus, the self-interaction approach is useful in the sense that it emphasizes the attraction or the inhibition exhibited by spatial locations of the trees, depending on the value of θ . In particular, if the parameter θ is close to 1, the model accepts the birth of new trees inside the zones $B(x_i, r)$ of the previous trees with high probability, which leads to clustering, whereas a lower value of θ indicates that the model favors locations

for new trees in the non-occupied areas rather than in the ZOIs of the previous trees, leading to repulsion. When $\theta = 0.5$, the model indicates no spatial interaction between trees. On the other hand, the parameter r allows the model to perceive the spatial features of the trees, such as the size of the competition area between trees when their ZOIs intersect, and the spatial coverage of the trees over the window \mathbf{W} , which is the union of all the ZOIs $B(x_i, r)$, $i = 1, \dots, k$. These features will serve as tools when developing dynamic summary statistics for sequential modeling (see Sect. 2.3).

2.2. LIKELIHOOD

We estimate the parameters involved in the model using the maximum likelihood (ML) method. This method has been applied for finite sequential point processes such as the random sequential absorption model (Lieshout 2006c), and the sequential model for eye movement (Penttinen and Ylitalo 2016). The likelihood function can be expressed as:

$$L(\vec{\mathbf{x}}_k) = g_1(x_1) \prod_{k=1}^{n-1} \alpha_k^{-1} \pi_{k+1}(x_{k+1}, r), \quad (5)$$

where $\alpha_k = \int_{\mathbf{W}} \pi_k(u, r) du$ is the normalizing constant, which depends on the sequence $\vec{\mathbf{x}}_k$, the window \mathbf{W} , and the parameters (θ, r) . The component $g_1(x_1)$ is assumed to be uniform. From (5), the likelihood is integrable since it is bounded, which makes the model well defined. We can write then log-likelihood for the model as

$$\begin{aligned} l(\theta, r) &= \log(\theta) \sum_{k=1}^{n-1} \mathbf{1}_{\{\mathcal{S}_k(x_{k+1}; r) \geq 1\}}(x_{k+1}) \\ &\quad + \log((1 - \theta)) \sum_{k=1}^{n-1} \mathbf{1}_{\{\mathcal{S}_k(x_{k+1}; r) = 0\}}(x_{k+1}) \\ &\quad - \sum_{k=1}^{n-1} \log \int_{\mathbf{W}} \pi_k(u, r) du. \end{aligned} \quad (6)$$

The normalizing integral depends on the model parameters, and it is evaluated at each step $k = 1, \dots, n - 1$. The evaluation will be based on the Riemann sum method.

2.3. MODEL EVALUATION

Assessing the goodness of fit will be conducted by using the envelope method in order to indicate the statistical variation in the summary statistic under the parametric model assumption. The envelope method allows us to compare the empirical functional summary statistics estimated from the data with the same summary statistics obtained through simulations of the fitted model based on the parametric bootstrap approach (Efron and Tibshirani 1994).

In our case, we need to employ metrics that measure various characteristics of the tree data under the sequential approach. Apparently, the usual summary functions used in the

field of spatial statistics lack order. Therefore, to describe the dynamics of the SSPP model, we establish various summary statistics as functions of order (time), and that are also justified from the forest applications point of view to measure the competition among trees for resources (e.g., light, water, nutrients). These dynamic functional summary statistics incorporate the self-interaction, area covered by the sequence (see Penttinen and Ylitalo 2016), nearest neighborhood distances, and areas of interaction between trees and are as follows:

1. *Lagged clustering statistic*: It is based on the lagged clustering measure $\mathcal{S}_k(x_{k+1}; r)$, which calculates the number of earlier tree locations (x_i 's) in $\vec{\mathbf{x}}_k = (x_1, \dots, x_k)$ within range r from the new location x_{k+1} . The variation along the time (ordered points) axis of this measure will indicate how much the number of new trees spatially closer to the earlier trees varies over the order (time).
2. *First contact distance*: We consider the distance between a given new location x_{k+1} and the past sequence $\vec{\mathbf{x}}_k$ and define it as the minimum of all the distances $\|x_{k+1} - x_i\|$, where $\|\cdot\|$ denotes the Euclidean distance and $x_i \in \vec{\mathbf{x}}_k$, i.e., $\min_{i \leq k} \|x_{k+1} - x_i\|$. This measure acts as the sequential version of the nearest neighbor distance function (see, e.g., Illian et al. 2008), but it is restricted to the past only.
3. *Proper zone statistic*: When a new location is interacting with the past, we will characterize this interaction *locally* in order to see whether it is inhibitive or attractive, and how strong it is. We will do so in terms of the interaction area for the new location. If we consider the ZOI of tree $k + 1$, i.e., $\mathbf{B}(x_{k+1}, r)$, we examine the part of this zone that does not intersect with any other zone $\mathbf{B}(x_i, r)$ formed by past locations $x_i, i = 1, \dots, k$. This shall determine a *proper zone* for x_{k+1} to be considered in competition, which is actually the complement of the overlapping area $\mathbf{B}(x_{k+1}, r) \cap \left\{ \bigcup_{i=1}^k \mathbf{B}(x_i, r) \right\}$. We quantify the proper zone statistic by measuring its normalized area with the value:

$$c_{k+1} = \frac{\left| \left\{ \mathbf{B}(x_{k+1}, r) \setminus \bigcup_{i=1}^k \mathbf{B}(x_i, r) \right\} \cap \mathbf{W} \right|}{\left| \mathbf{B}(x_{k+1}, r) \cap \mathbf{W} \right|},$$

where the symbol $|\cdot|$ denotes the area. In other words, at the ZOI of each tree, the evaluated proper zone statistic induces resource territory used by the tree without competition from larger trees, while the overlapping zone indicates the space where it competes for resources.

4. *Ball union coverage*: At each location x_k , we consider the *regionalized* form of the sequence $\vec{\mathbf{x}}_k$ by taking the union $\bigcup_{i=1}^k \mathbf{B}(x_i, r) \cap \mathbf{W}$. This union will measure the coverage and the degree of filling cast by the sequence at each time inside the window \mathbf{W} , while indicating the development of empty space in \mathbf{W} . We write it in normalized form as:

$$\frac{\left| \bigcup_{i=1}^k \mathbf{B}(x_i, r) \cap \mathbf{W} \right|}{|\mathbf{W}|}.$$

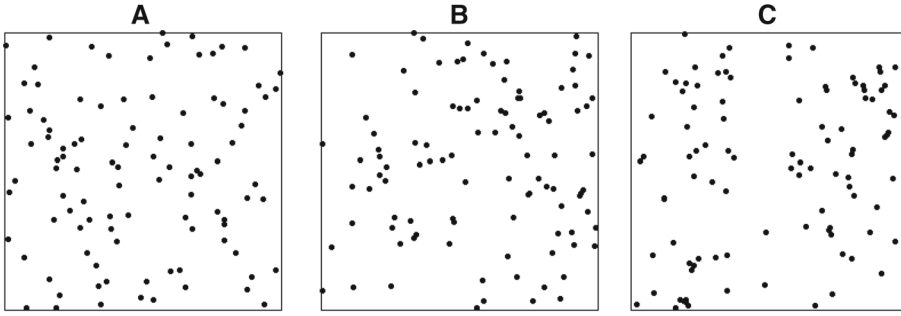


Figure 2. Simulated realizations for each of the three cases: Case A (left), case B (middle) and case C (right) .

On the other hand, the empty space statistic which is the complement of the ball union coverage is similarly given by

$$1 - \frac{\left| \bigcup_{i=1}^k B(x_i, r) \cap W \right|}{|W|}.$$

In this paper, the first three summary statistics are presented in cumulative form to express the contribution of each tree to the computed summary statistic. The ball union coverage is analytically similar to the cumulative proper zone statistic; however, its stepwise scaling is different. We turn back to this point later in the final section.

2.4. SIMULATION

Often, properties of complex models can be studied through simulations. In our case, we use simulations to observe the behavior of the model and for testing the goodness of fit through summary statistics. In this section, we examine the behavior of the SSPP model via simulations utilizing different types of self-interaction which are based on the values of the parameter θ . The simulated realizations are generated using the conditional distribution (4) with the parameters (θ, r) . Thereafter, we will employ the summary statistics listed above to see how well these statistics capture the features of the simulated data.

Each realization is a sequence \vec{x}_{100} of 100 locations in the unit square window, and for the sake of simplicity, the first two points x_1 and x_2 of the sequence \vec{x}_{100} are drawn uniformly and will be used as starting points for all realizations. The first sampled points are $x_1 = (0.20, 0.55)$ and $x_2 = (0.91, 0.81)$. The latter points of the sequence are simulated by adding locations x_k using the accept–reject method developed by von Neumann (see, e.g., Ripley 1987) such that the law of x_k given the past \vec{x}_{k-1} follows the density (4).

Since we are interested in the effect of the self-interaction characteristics on the model, the parameter r will be fixed at $r = 0.1$. We propose the following inhibitive, neutral, and attractive case values, respectively, for the parameter θ : (A) $\theta = 0.20$, (B) $\theta = 0.5$ and (C) $\theta = 0.80$.

We generate 20 realizations for each case and compute the four functional summary statistics. Simulated realizations are visualized in Fig. 2. In Fig. 3, the results related to the

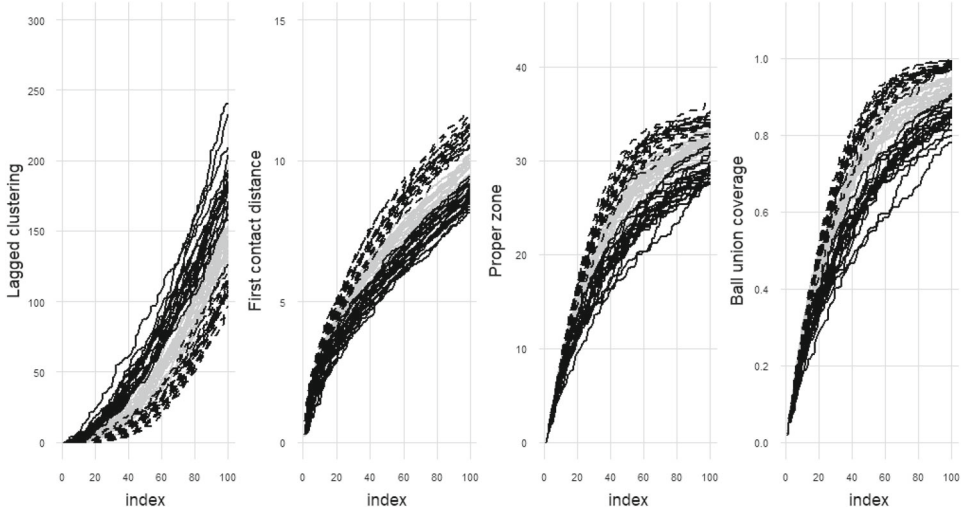


Figure 3. Summary statistics for simulated patterns of the three cases. From left to right: cumulative lagged clustering statistic, cumulative first contact distance, cumulative proper zone statistic, and the ball union coverage. Dashed lines represent the statistics for case A, gray lines represent those for case B, and dark solid lines represent those for case C .

summary statistics are given in cumulative form. The lagged clustering statistic indicates the number of earlier locations near the current location, say x_k , and the cumulative version sums all these numbers together, i.e., $\sum_{i=1}^{k-1} \mathcal{S}_i(x_{i+1}; r)$. We see that in case A, new locations avoid zones near past locations compared to case B, and the values of its cumulative first contact distance statistic are larger, thus expressing this avoidance character of the pattern compared to case C. The latter case exhibits high values of the cumulative lagged clustering statistic, indicating the new locations are favoring zones of previous locations, and its first contact distances are the smallest among these three cases, indicating a tendency toward clustering.

The ball union coverage in case A increases faster than it does for the other two cases, filling almost the whole window after just 80 points. Further, the forthcoming points indexed from 80 to 100 try to locate themselves inside the ZOIs of former points, which is indicated by a slow-up in the ball union coverage function. Accordingly, the fast coverage in case A is consistent with the increase in the proper zone gained by the points as can be seen from the proper zone statistic. On the other hand, the coverage in case C fills less than 80% of the window, expressing the clustering of the new locations in the ZOIs of the former ones. In addition, the proper zone statistic in case C shows a similar effect.

2.5. GENERALIZED SSPP MODEL

In the SSPP model above, the radius of interaction r is fixed for all $B(x_i, r), k = 1, \dots, n$. However, one might expect the interaction radius r to be a function of tree size. In this case, the parameter r is specific for each tree i , allowing the ZOI to reflect the size of the tree, i.e., larger trees have larger ZOIs and younger trees have smaller ZOIs. We express the radius

using the power equation

$$r_i = \alpha m_i^\beta \quad \alpha > 0, \quad \beta \in \mathbb{R},$$

where $m_i = d_i/2$, and d_i is the DBH of tree i (Berger et al. 2008). We refer to this model as the *generalized SSPP model* with parameters (θ, α, β) , where the SSPP model is the special case obtained by taking $\beta = 0$ and $\alpha = r$. Under this general construction, the lagged clustering measure is given by $\mathcal{S}_k(y) = \sum_{i=1}^k \mathbf{1}_{B(x_i, r_i)}(y)$. Similarly, the estimation of the parameters is ML-based, where the likelihood is given by (5) with the self-interaction function $\pi_k(y) = \mathbf{1}_{\{\mathcal{S}_k(y) \geq 1\}}(y) + (1 - \theta) \mathbf{1}_{\{\mathcal{S}_k(y) = 0\}}(y)$.

3. APPLICATION

3.1. DATA

To demonstrate the performance of the SSPP models constructed in Sect. 2, we apply them to forest stand data. Our data were collected from the Kiihtelysvaara region, which is situated in northern Carelia, Eastern Finland. The dominant tree species is the Scots pine (*Pinus sylvestris* L.), and all trees with DBHs exceeding 4cm or heights exceeding 4 m were mapped from a total of 79 sample plots, located randomly within selected forest stands such that each plot was entirely within the same forest stand. The data were originally collected to study the performance of different aerial forest inventory methods, see Packalen et al. (2013) for details. Three plots exhibiting clearly different spatial patterns for tree locations were subjectively selected for use in this paper, namely Plots I, II, and III.

The datasets of the three plots consist of tree locations and their corresponding DBHs. Plot I contains 118 trees with minimum and maximum DBHs of 2.30 cm and 41.95 cm, respectively, with a mean value of 12.11 cm. In Plot II, there are 160 trees with DBHs ranging from 2.30 cm to 21.60 cm and a mean value of 9.24 cm. Plot III has 101 trees with a minimum diameter of 3.35 cm, maximum diameter of 40.45 cm, and a mean value of 14.66. Figure 4 (first row) shows the patterns describing the spatial structures of these sample plots at bounded window sizes of $30 \times 30 \text{ m}^2$ (Plot I), $25 \times 25 \text{ m}^2$ (Plot II), and $25 \times 25 \text{ m}^2$ (Plot III). At this point, we will ignore any edge effects, and all variations in the models are restricted to the bounded windows. However, we will discuss the edge-effects at the end of this section.

As preliminary stage of the study, an explanatory analysis will be carried out using the second-order structures of the *static* point patterns presented in Fig. 4 (second and third rows). Here, we adopt the centered L-function and the G-function to check complete spatial randomness (CSR). More details on the L- and G-functions can be found in Illian et al. (2008), Møller and Waagepetersen (2004). Employing these summary statistics will indicate whether their static nature can help to reveal the dynamic introduced by the sequential approach. To test the CSR hypothesis, a global envelope test is performed using extreme rank length (ERL) ordering (Myllymäki et al. 2017; Mrkvíčka et al. 2018). The test is conducted using 2499 simulations of the CSR on the interval $[0, 6]$ (in m) for r , and the evaluation is

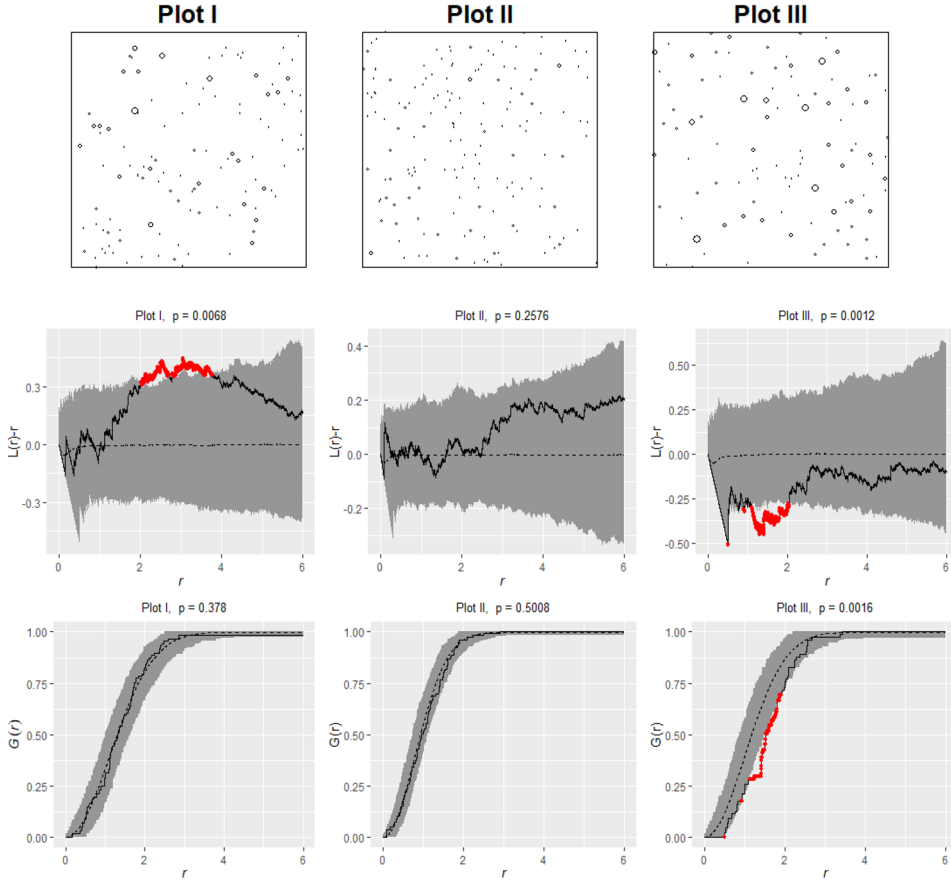


Figure 4. First row: Positions of trees in Plots I, II, and III in Kiihtelysvaara (Eastern Finland) in windows of size $30 \times 30 \text{ m}^2$ (Plot I), $25 \times 25 \text{ m}^2$ (Plot II), and $25 \times 25 \text{ m}^2$ (Plot III), where the radius of a disc is $2 \times \text{DBH}$. Global envelope tests for the summary $L(r) - r$ (second row) and for the G function (third row) of Plot I (left), Plot II (middle), and Plot III (right) are based on 2499 simulation of CSR. The gray areas show the 95% global envelopes. The solid black lines are the data functions observed in the sample plots, and the dashed lines represent the (estimated) theoretical expectation.

represented graphically in Fig. 4 together with their estimated p -values (second and third rows). The R package `spatstat` was used to produce the plots (Baddeley et al. 2015).

For Plot I, the estimated centered L-function exhibits a clustering character for the scale 2.19–3.31 m. However, the nearest neighbor function G does not reject the CSR property. In Plot II, it is clear that the test did not reject the CSR property for either the L- or G-function, while for Plot III, the CSR is rejected, and the plot shows a regular character at a scale of 0.89–2.04 m.

3.2. MODEL FITTING

To fit the sequential models to the observed data, we first write the ordered sequences $\vec{x}_n = (x_1, \dots, x_n)$ of the observed locations in each plot. The adopted order is based on the DBH, where x_1 is the location of the largest tree, and x_n the location of the smallest tree

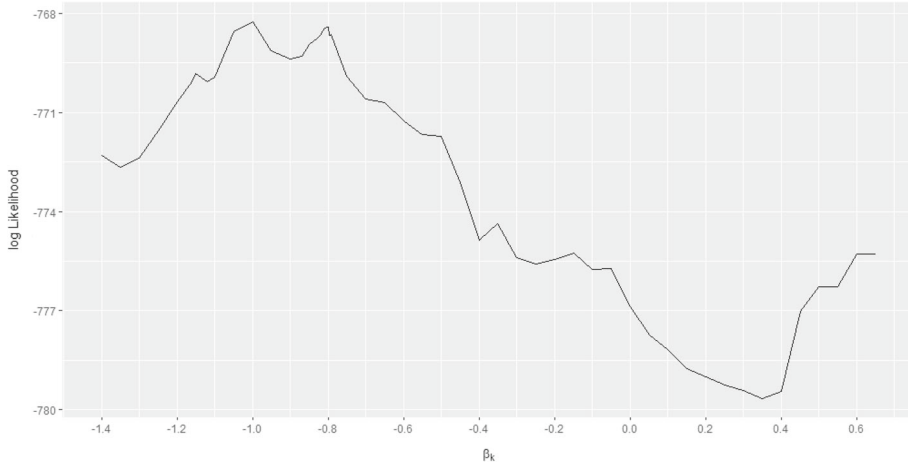


Figure 5. The graph of the log-likelihood $l(\hat{\theta}_k, \hat{\alpha}_k, \beta_k)$ as a function of β_k .

in the plot. In the first model, where the ZOIs are similar for all trees, trees are allowed to grow inside the stems of the other trees as the value of the radius r approaches 0. This is problematic since the stems cannot overlap. To adjust the model for this problem, we set a small hard-core distance around each tree so that no new trees are allowed to grow there. For each plot, this hard-core distance, which we denote by r_{\min} , is equal to the radius of the largest tree in the plot, i.e., $r_{\min} = \text{DBH}_{\max}/2$, where DBH_{\max} is the DBH of the largest tree. The hard-core distances are (in m) $r_{\min} = 0.21$, $r_{\min} = 0.11$, $r_{\min} = 0.20$ for Plot I, Plot II, and Plot III, respectively.

When determining the MLE of (θ, r) for the SSPP model and of (θ, α, β) for the generalized SSPP model, we perform the optimization routines implemented in the R package `nloptr` (Johnson 2010). Here, we employ the deterministic-search algorithm `Direct-L` (see Gablonsky and Kelley 2001) to obtain the global optimum, then adjust it for greater accuracy using the Nelder–Mead simplex algorithm (Richardson and Kuester 1973). To re-check the results, we maximize the likelihood functions over a grid of parameters values and achieve the same results. For the generalized SSPP model, we adopted a profile likelihood approach: First, we use `nloptr` to find the maximum likelihood estimates $\hat{\theta}_k$ and $\hat{\alpha}_k$ for a range of fixed values of β_k . Second, we select the $(\hat{\theta}_k, \hat{\alpha}_k, \hat{\beta}_k)$ with maximum log-likelihood as the MLE $(\hat{\theta}, \hat{\alpha}, \hat{\beta})$. Figure 5 illustrates this approach for Plot I, where the maximum of $l(\hat{\theta}_k, \hat{\alpha}_k, \beta_k)$ is at $\beta_k = -1$.

The MLEs for the Plots I, II, and III are summarized in Table 1 for the SSPP model and in Table 2 for the generalized SSPP model. The 95% confidence intervals for the parameter estimates are constructed using the parametric bootstrap method (see Efron and Tibshirani 1994, p. 53), where the standard error is computed from 100 bootstrap replications. For each replication, we simulated a realization of the fitted model, and then, we fit the model to this simulated realization to obtain the replicated parameters. The point estimates of parameter r for the basic SSPP model range between 1.31 m (Plot III) and 2.60 m (Plot I). The estimate of θ is 0.65 for Plot I, indicating tendency to cluster new trees in the neighborhood of previous trees. For the other two plots, the estimates are 0.37 and 0.22, indicating that new

Table 1. Parameter estimates for the SSPP model for Plots I, II and III, together with the 95% bootstrap CIs the values of the log-likelihood at maximum, and the AIC values

	Plot I	Plot II	Plot III
$\hat{\theta}$	0.66	0.37	0.24
\hat{r}	2.6	1.31	1.80
CI for $\hat{\theta}$	(0.65, 0.67)	(0.35, 0.38)	(0.23, 0.25)
CI for \hat{r}	(2.54, 2.66)	(1.12, 1.49)	(1.78, 1.82)
$l(\hat{\theta}, \hat{r})$	-776.73	-1009.88	-621.65
AIC	1557.46	2023.76	1247.30

Table 2. Parameter estimates of the generalized SSPP model for Plots I, IIs and III, their associated 95% bootstrap CIs, the values of the log-likelihood at maximum, and the AIC values

	Plot I	Plot II	Plot III
$\hat{\theta}$	0.79	0.30	0.20
$\hat{\alpha}$	0.18	17.04	1.20
$\hat{\beta}$	-1.00	0.99	-0.15
CI for $\hat{\theta}$	(0.76, 0.82)	(0.27, 0.33)	(0.19, 0.21)
CI for $\hat{\alpha}$	(0.08, 0.28)	(16.11, 17.97)	(1.05, 1.36)
CI for $\hat{\beta}$	(-1.04, -0.95)	(0.92, 1.05)	(-0.19, -0.12)
$l(\hat{\theta}, \hat{\alpha}, \hat{\beta})$	-768.24	-1003.34	-618.92
AIC	1542.48	2012.68	1243.84

tree locations tend to avoid those of existing trees. None of the 95% confidence intervals for θ include the value 0.5. The estimates of θ are also quite similar also in the generalized SSPP model.

Next, we evaluate the models by computing the four summary statistics introduced earlier in Sect. 2.3. The evaluation is based on the 95% global envelope test (Myllymäki et al. 2017; Mrkvička et al. 2018). Figures 6 and 7 show the 95% envelopes obtained from 2499 simulated realizations of the fitted SSPP model and the fitted generalized model, respectively, together with their associated p -values. When simulating, we use the observed values of the first two locations as starting points for the simulated sequence in order to reduce unexpected variations in the simulated realizations. Based on the p -values and the fact that the 95% envelopes cover summary statistics of the observed data, we conclude that both SSPP models fit well, indicating agreements between the data and the fitted models. Comparing the model fit using the evaluated likelihoods at the ML estimates and the AIC values based on them, the generalized SSPP model provides a better fit than the reduced SSPP model. The negative values of $\hat{\beta}$ for Plot I and Plot III indicate that the larger trees have smaller ZOIs compared to the smaller trees. For Plot II, $\hat{\beta} \approx 1$ and $\hat{\alpha}$ is quite high, indicating increasing of ZOI as a function of tree size. The parameter estimates imply that the interaction radii for Plot II range within (0.41, 3.74), where the first value is the interaction radius of the smallest tree and the second value is that of the largest tree. For Plot I and Plot III, these ranges are (0.43, 7.83) and (1.37, 2.00), respectively, where the first value is associated with the

MODELING FOREST TREE DATA USING SEQUENTIAL SPATIAL POINT PROCESSES

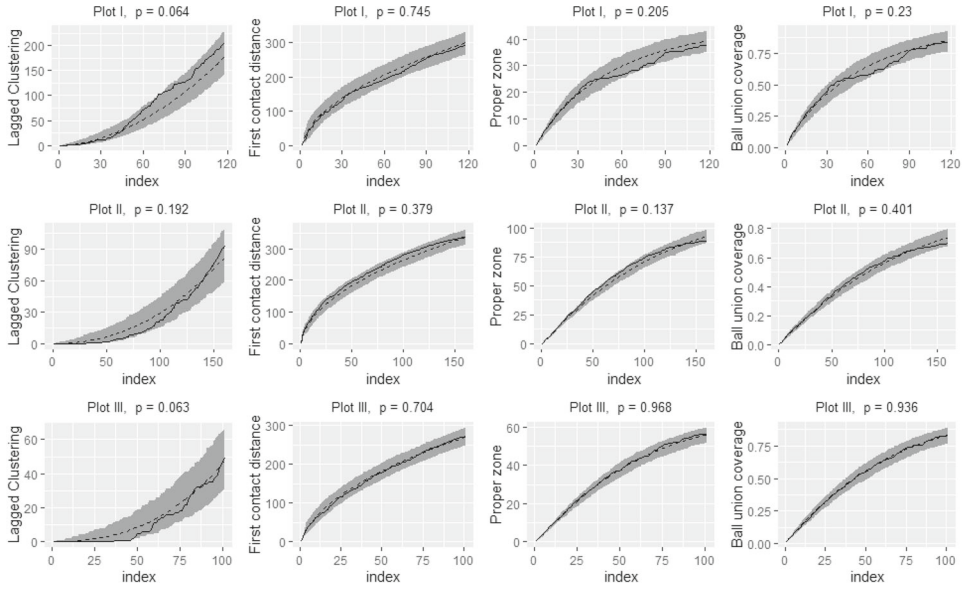


Figure 6. The SSPP model: summary statistics for the observed locations (solid lines), together with the 95% global envelopes (shaded areas) calculated from 2499 simulations of the fitted model .

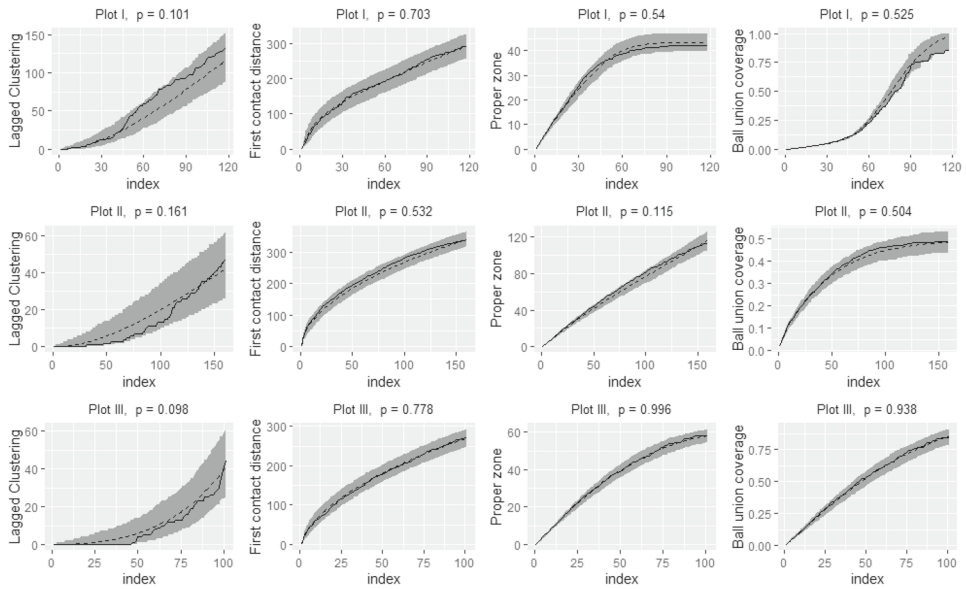


Figure 7. The generalized SSPP model: summary statistics for the observed locations (solid lines), together with the 95% global envelopes (shaded areas) calculated from 2499 simulations of the fitted model .

largest tree and the second value with the smallest tree in the plot. We will give a possible explanation for these results in the discussion.

We can now define the self-interaction function for each plot in terms of the classes of radius r using the hard-core distance r_{\min} and the estimates $\hat{\theta}$ and \hat{r} for the SSPP model.

We denote this function as π . For the generalized SSPP model, the self-interaction function denoted by π_i is specified for each tree i with DBH d_i . It is expressed using the estimates $\hat{\theta}$ and $\hat{\alpha} (d_i/2)^{\hat{\beta}}$. We have

$$\pi = \begin{cases} 0 & \text{on } [0, r_{\min}), \\ \hat{\theta} & \text{on } [r_{\min}, \hat{r}), \\ 1 - \hat{\theta} & \text{on } [\hat{r}, r_{\max}), \end{cases} \quad \pi_i = \begin{cases} 0 & \text{on } [0, d_i/2), \\ \hat{\theta} & \text{on } [d_i/2, \hat{\alpha}(d_i/2)^{\hat{\beta}}), \\ 1 - \hat{\theta} & \text{on } [\hat{\alpha}(d_i/2)^{\hat{\beta}}, r_{\max}), \end{cases}$$

where r_{\max} stands for the maximum value of r in each sample plot and is simply the diameter of the window. For convenience, r_{\max} is set to 42.43 m, 35.36 m, and 35.36 m for Plot I, Plot II, and Plot III, respectively.

3.3. IS THE DYNAMIC APPROACH NEEDED?

Even though the dynamic nature of a forest stand is well supported for the application, one might question whether the self-interactive feature of the model provides a significant improvement over a static model. This subsection explores this issue from two points of view: (i) through permuting the order of tree locations in terms of size, and (ii) through a dynamic nearest neighbor statistic.

Permutations: The dynamics presented in this paper are based on the ordering of the locations, the effect of which can be explored by randomly permuting the order of observed locations. We note that the number of all the possible orders of the sequence is $n!$, which is the total number of permutation on the set $\{1, \dots, n\}$. Since the dynamic of the sequential model is provided by the self-interaction component in the model, $\mathcal{S}_{i-1}(x_i; r)$, a different order will result in a different self-interaction function. Consequently, if a functional summary statistic based on the original order is very extreme compared with the statistics based on the permuted orders, we can conclude that the order really matters. To analyze this effect, we compared the lagged clustering statistic of the observed locations under the observed order, with the computed lagged clustering statistic based on permuted orders. In Fig. 8, we perform the 95% global envelope test based on the lagged clustering statistic of 2499 permutations of the observed sequence in each plot. When calculating the statistic, we use the fitted value of the radius based on the observed order, \hat{r} , for all the permuted sequences in the test. In Fig. 8, the test rejects the *exchangeability* of the order for all sample plots, which means that the observed order is distinguishable.

The lagged clustering statistic is sensitive to the re-ordering of the locations and shows a clear effect when a different order is adopted. However, this effect cannot be seen through arbitrarily chosen summary statistics, in general. In particular, the other summary statistics used in this paper did not indicate as clear of an importance for the order as the lagged clustering statistic.

Nearest neighbor: With the exception of Plot III, the parameter estimates $\hat{\theta}$ for both models are not consistent with the summary statistics estimated in Fig. 4. In Plot III, these static summary statistics indicate a regular pattern for which the sequential models provide a low value of $\hat{\theta}$. However, the centered L-function and the G-function fail to reveal the inhibition character of Plot II, whereas the SSPP models indicate a low value of $\hat{\theta}$. On the

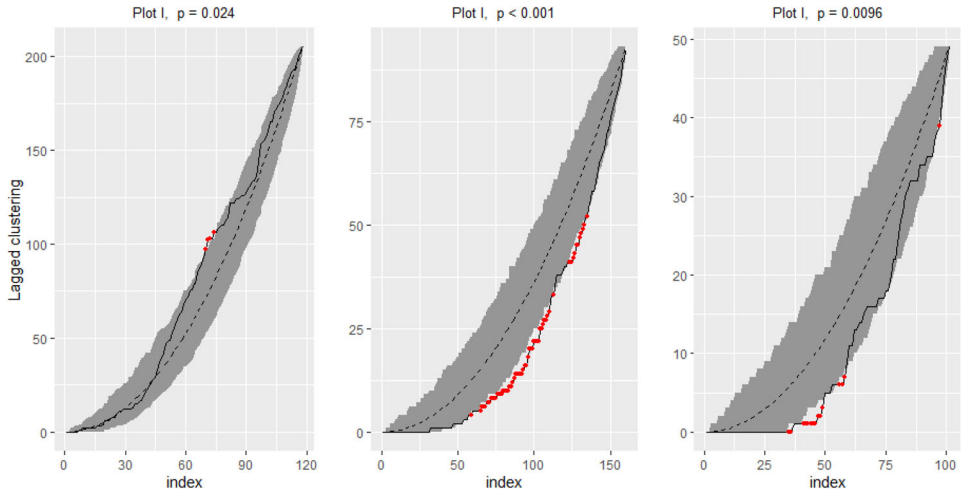


Figure 8. The 95 % global envelope tests based on the lagged clustering statistic for 2499 permutations of the observed sequence, where the solid black line indicates the lagged statistic for the observed order .

other hand, for Plot I, the high value of $\hat{\theta}$ indicating spatial clustering has not been reflected well by the estimated G-function. These problems have arisen due to the static nature of the L- and the G-functions. Essentially, it is possible to have dependence in terms of time or size ordering in the sequence, when the static or marginal (over time) point pattern looks like a Poisson process.

To incorporate order into the G -statistic, we use a dynamic (unnormalized) alternative to the G-function that is expressed in terms of order. The new developed statistic is based on the first contact distance statistic introduced in Sect. 2.3. First, we fix a distance r . Then, denoting the first contact distance of a point x_i by $D_i = \min_{j < i} \|x_i - x_j\|$, we define

$$\bar{D}_i(r) = \begin{cases} 1 & \text{if } D_i \leq r, \\ 0 & \text{elsewhere,} \end{cases}$$

which is the probability that the first contact distance of a point x_i with respect to its past \vec{x}_{i-1} lies within the range of r . Clearly, the quantity \bar{D}_i depends on the past only. Figure 9 shows the results of the 95 % global envelope tests based on this (unnormalized) dynamic nearest-neighbor statistic. The optimal value of the fixed distance r is taken as \hat{r} , the fitted value for the SSPP model. We compute the new dynamic statistic of the observed sequence of locations against 2499 simulations of the null model, which, in our case, is the sequential model with $\theta = 0.5$, i.e., the non-interaction model. In the simulations, we generate the same number of points as there are in the observed plots, and we use the fitted value of r . According to the graphs and the corresponding p -values, the dynamic nearest neighbor function rejects the null hypothesis ($\theta = 0.5$) since the computed statistics for the observed locations in the plots are not completely inside the envelopes. Moreover, the computed dynamic statistic for Plot I is above the envelope revealing clustering, while it is below the envelope for Plots II and III, which indicates spatial inhibition.

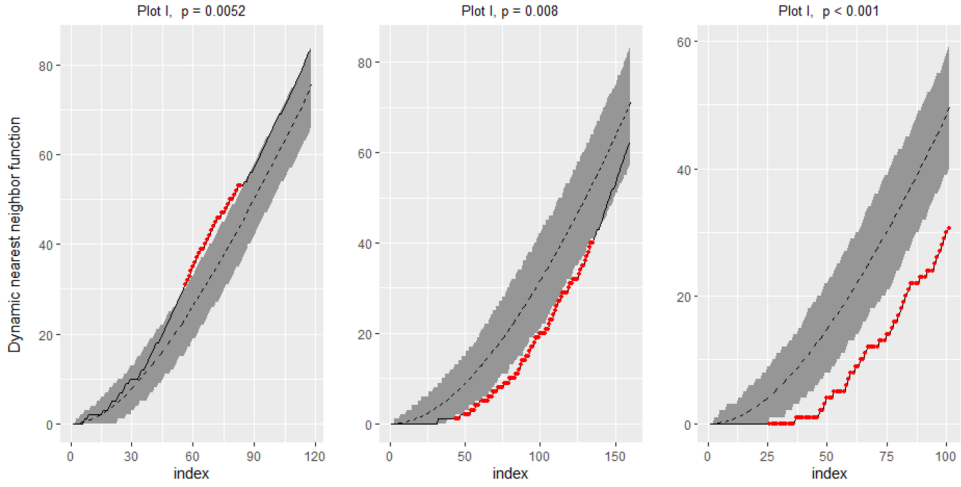


Figure 9. Plot I (left), Plot II (middle) and Plot III (right). The figure shows the results of the 95 % global envelope tests based on the cumulative $\bar{d}_i(\hat{r})$ statistic generated by 2499 simulations of the null model (no interaction), where the solid black lines represent the computed cumulative value of $\bar{d}_i(\hat{r})$ of the ordered sequences for the plots .

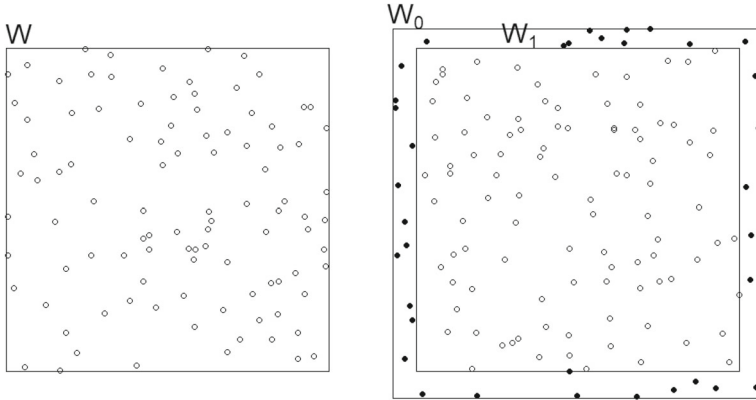


Figure 10. Edge effect cases: (Left) The first case represents isolated effects. (Right) The second case involves possible edge-effects. Here, the black dots indicate the trees in the buffer zone .

3.4. EDGE EFFECTS UNDER THE SEQUENTIAL APPROACH

In spatial point pattern analysis, edge corrections play a role in both the estimation of summary statistics and the likelihood inference. The most popular edge-correction approaches are Horwitz–Thompson type weighting and the use of a reduced window together with conditioning, see, e.g., Illian et al. (2008). For sequential point process data, the weighting method is intractable but the second method is applicable when the observation window is large enough.

The sizes of the sample plots in our study, namely Plot I, Plot II, and Plot III, are small. Our analyses ignored trees outside the observed windows of these plots. So, in order to explore the possible edge effects on the estimation under sequential modeling, we check the variations in the estimated parameters for the model in two distinct situations: without

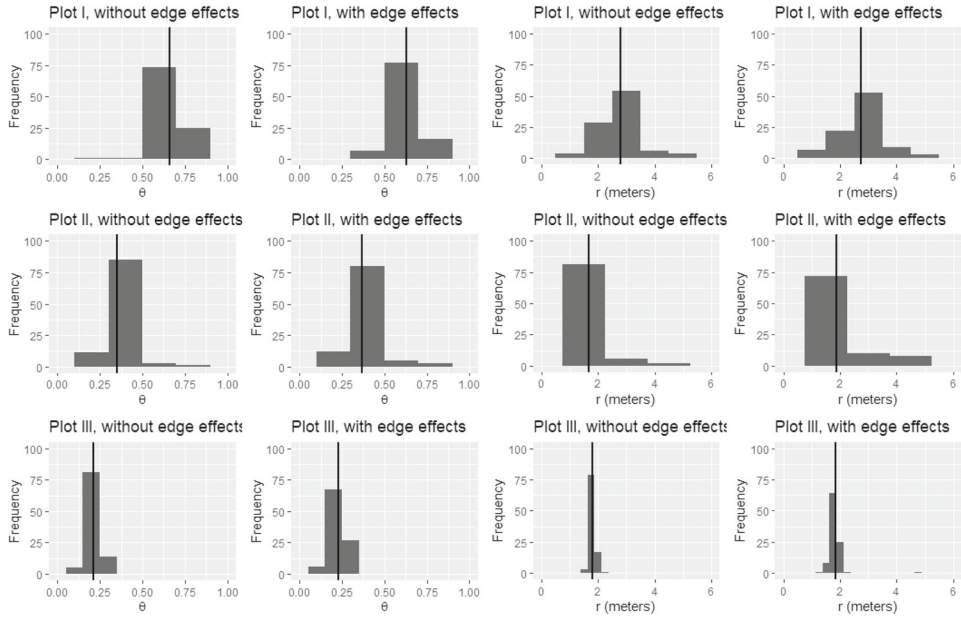


Figure 11. Histograms of the parameter estimates θ and r for 100 realizations simulated under the fitted model I for Plot I, Plot II and Plot III. The dark solid lines indicate the mean values .

edge effects and with edge effects (Fig. 10). For the first case, in which the effects of the trees outside the observation window \mathbf{W} are ignored, we simulate realizations from the fitted model inside the observation window \mathbf{W} , which is of the same size and contains the same number of trees as the sample plot. In the second case, effects from the trees in a buffer zone are taken into account in the estimation. The realizations are simulated from the same fitted model, but in a larger window \mathbf{W}_0 contains a buffer zone around the window \mathbf{W}_1 which is of the same size as \mathbf{W} . In these simulations, we condition on the number of trees inside \mathbf{W}_1 such that it is equal to the number of trees in \mathbf{W} . The analysis is done only for the SSPP model, since the generalized model would require additional assumptions on the sizes of the trees that lie in the buffer zone. The estimate \hat{r} obtained via the SSPP model will be used as the buffer zone width. We simulate 100 realizations of the fitted model in both cases and estimate the parameters of the model for each realization in \mathbf{W} and \mathbf{W}_1 so that the variations in the estimates can be compared. Figure 11 shows histograms of the parameter estimates $\hat{\theta}$ and \hat{r} for these 100 realizations simulated under the fitted SSPP model for Plots I–III. The histograms of the parameter estimates for the two situations are similar, indicating that the buffers have only minor effect on the parameter estimates.

In the case of size-determined ordering, we hypothesize that the boundary effect is small in the early part of the sequence, which consists of large trees, although the points in the reduced observation window are near the boundary. Also, small trees outside the reduced observation window do not affect the estimators excessively. Hence, the boundary effect is, in our opinion, less severe than would be the case in (static) point pattern analysis. Our simulation example supports this hypothesis.

4. DISCUSSION

The mechanism of the sequential model introduced in this paper for forest tree data represents a powerful modeling tool since it offers a causal description through successive conditioning. It reveals the information contained in the history of the ordered spatial locations, which is interpreted in terms of future-past dependence in a dynamic form. Although we adopt a history-dependent model which uses the full history to model the dependence, other ranges of history can be also of interest, such as restricting the history at a given location x_k , up to the last m points x_{k-m}, \dots, x_{k-1} , $m < k$. This type of history range shall define the so-called m -memory point processes, (Snyder and Miller 1991), where $m = 1$ stands for the Markovian case. The authors intend to investigate these types of history under the sequential approach in future work. The length of history should, of course, also depend on the size of the sample plot.

Evaluating the model requires constructing summary statistics that measure the essential properties of the studied phenomena. For example, the summary statistics used in Penttinen and Ylitalo (2016), which are based on the distances between successive points in eye movements, might be applicable for animal movement data, but are not very informative and relevant in the context of forest tree data. Thus, we proposed statistics that measure useful properties of forest tree data and take into account ordering based on the sizes of the trees (hence, the age). In particular, the lagged clustering and first contact distance statistics determine the distances of new trees compared to the earlier trees. For the proper zone and ball union statistics, the disc of radius r around each tree can be thought as an interaction zone or approximate crown projection to the ground. Therefore, the proper zone statistic measures the contribution of the current tree of the total area occupied by the trees up until this points of time, and the ball union coverage statistic measures the total occupied area up until the current instance. Moreover, we developed a dynamic alternative statistic to the nearest neighbor distance distribution function, which lacked efficacy when interpreting the data in the sample plots. The Monte Carlo simulation showed that these statistics are affected by the changes in the applied model, and their application in assessing the forest data proved that they are useful also in practice. Since the functions are in terms of time/order, they allow us to catch the variations in the feature in question.

Our model was motivated primarily by forest inventories based on airborne laser scanning (ALS), where a lidar device carried by aircraft takes repeated canopy height measurements of the forested area below. The resulting set of echoes (points) are used to detect individual trees from the point cloud. Nevertheless, only trees with crowns not covered by the crowns of larger trees can be detected from the point cloud. Therefore, the intensity of the spatial point process (stand density, trees per ha) is underestimated, and the mean tree size is overestimated (Kansanen et al. 2016). However, the hidden trees can be taken into account in the estimation of stand density by modeling the detectability of trees mechanistically as a function of tree size to adjust the observed trees by the estimated detectability using a Horvitz–Thompson-like estimator (Kansanen et al. 2016, 2019). Nevertheless, the method implicitly assumes that the density of small trees in the parts of the forest that are not covered by the crowns of larger trees is similar to that in the parts that are covered. Using the notations of the current paper, this assumption is equivalent to when $\theta = 0.5$. This may

not be realistic because larger trees have initially affected the birth of new trees in the past. The model proposed in the current paper could be used to estimate the relative intensity of the small trees within the influence zones of larger trees compared to non-influenced areas. The application would be especially straightforward if the parameter r could be directly taken as the (observed or predicted) crown radius of the tree in question, and parameter θ could be predicted with sufficient accuracy based on the remotely sensed forest data and stand management records. This situation will be investigated in a follow-up paper. The step function-type interaction kernel will be especially useful in this application, and the results of this study indicate that it allows a sufficient simplification of reality. For other uses of the model, a smooth kernel might be better justified.

We presented two models: the SSPP, where the interaction radius is constant, and a generalized model, where the radius of the ZOI is a function of tree size based on a power model. In initial analyses, we also considered other functions, but they did not perform better than the power function. The dependence of the radius on tree size is realistic because large trees also have large crowns and root systems; thus, they affect neighboring trees more strongly than small trees do. However, their locations were determined a long time ago when they were smaller. Therefore, their current sizes can affect the locations of neighboring trees only by killing them off via competition, i.e., by thinning the process. In this study, the generalized model showed a better performance than the simple model, in which the interaction between trees is expressed in terms of the size of the trees. Initially, we expected for the generalized model that $\beta > 0$. Instead, the MLE of β was negative for Plot I and Plot III. This indicates that the interaction radius of large trees is smaller than for small trees. The explanation might be that the largest trees have lost their branches in the bottom part of the stem, and consequently do not have much effect on the growing conditions of the smallest trees. On the other hand, the medium-sized trees may still have branches at lower level, which have an effect on the smallest trees. These explanations actually suggest even more flexible generalization of the model, fitting of which would require a more extensive data set.

Although the SSPP model was able to read patterns with different aggregations, the concept of the sequential model presented in this paper differs from other spatial point processes in the sense that it is interpreting the structures of observed patterns beyond the general regular-clustered definition. It is reading their aggregation within a certain specific range given by the parameter r . In contrast, this parameter could be fixed using the approach in [Penttinen and Ylitalo \(2016\)](#), or, more generally, a collection of suitable r values representing the range of the competition strength could be used. The former would be an advantage in laser scanning application, where the crown radius of an observed tree could be used as a fixed r . Then, the only parameter to be estimated would be θ .

The order of the spatial dimension used in the paper emphasizes the dynamics of the forest stand, as we believe that the larger trees appeared earlier than the small ones. Hence, our objective was to switch from the static point pattern to a dynamic pattern emphasizing the great importance of the spatial history, and revealing the long-term dependence among the ordered sequence of tree locations. However, other orders can be used when applying the sequential spatial point processes depending on the context and observed phenomena.

For example, in forest data, alternative orders could be developed with respect to tree height or tree age, if known.

ACKNOWLEDGEMENTS

We would like to thank the two anonymous reviewers for their valuable comments. This research was financially supported by the Academy of Finland research Project 310073 and the flagship program (UNITE, Decision Number 337655).

Funding Open access funding provided by University of Eastern Finland (UEF) including Kuopio University Hospital.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

[Received June 2020. Revised August 2021. Accepted August 2021.]

REFERENCES

- Baddeley A, Rubak E, Turner R (2015) Spatial point patterns: methodology and applications with R. Chapman & Hall/CRC, Spatial point patterns: methodology and applications with R
- Berger U, Hildenbrandt H, Grimm V (2008) Towards a standard for the individual-based modeling of plant populations: self-thinning and the field-of-neighborhood approach. *Nat Resour Model* 15:39–54
- Cressie NAC (1993) Statistics for spatial data. Wiley, New York
- Daley DJ, Vere-Jones D (2003) An introduction to the theory of point processes. Volume I: elementary theory and methods. Springer, New York
- Daley DJ, Vere-Jones D (2008) An introduction to the theory of point processes. Volume II: general theory and structure. Springer, New York
- Diggle PJ (2013) Statistical analysis of spatial and spatio-temporal point patterns. Chapman & Hall/CRC, New York
- Diggle PJ, Besag J, Gleaves JT (1976) Statistical analysis of spatial point patterns by means of distance methods. In: *Biometrics*, pp 65–667
- Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. Chapman and Hall/CRC Press, New York
- Evans JV (1993) Random and cooperative sequential adsorption. *Rev Mod Phys* 65:1281–1330
- Gablonsky J, Kelley C (2001) A locally-biased form of the DIRECT algorithm. *J Glob Optim* 21:27–37
- González JA, Rodríguez-Cortés FJ, Cronie O, Mateu J (2016) Spatio-temporal point process statistics: a review. *Spatial Stat* 18:505–544
- Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modeling of spatial point patterns. Wiley, Chichester
- Jensen EBV, Jónsdóttir KY, Schmiegel J, Barndorff-Nielsen OE (2007) Spatio-temporal modelling with a view to biological growth. In: *Statistical methods for spatio-temporal systems, Monographs on statistics and applied probability*, pp 47–75

- Johnson SG (2010). The NLOpt nonlinear-optimization package. <http://github.com/stevengj/nlopt>
- Kansanen K, Vauhkonen J, Lähivaara T, Mehtätalo L (2016) Stand density estimators based on individual tree detection and stochastic geometry. *Can J For Res* 46:1359–1366
- Kansanen K, Vauhkonen J, Lähivaara T, Seppänen A, Maltamo M, Mehtätalo L (2019) Estimating forest density and structure using Bayesian individual tree detection, stochastic geometry, and distribution matching. *ISPRS J Photogram Remote Sens* 152:66–78
- Lieshout M (2006a) Campbell and moment measures for finite sequential spatial processes. *Proc Prague Stoch* 48:215–224
- Lieshout M (2006b) Markovianity in space and time. *IMS Lect Notes Monogr Ser* 48:154–168
- Lieshout M (2006c) Maximum likelihood estimation for random sequential adsorption. *Adv Appl Probab (SGSA)* 38:889–898
- Lieshout M, Capasso V (2009) Sequential spatial processes for image analysis. In: *Electronic transactions on numerical analysis*
- Maltamo M, Naesset E, Vauhkonen J (2014) *Forestry applications of airborne laser scanning concepts and case studies*. Springer, Berlin
- Mehtätalo L (2006) Eliminating the effect of overlapping crowns from aerial inventory estimates. *Can J For Res* 36:1649–1660
- Møller J, Waagepetersen RP (2004) *Statistical inference and simulation for spatial point processes*. Chapman & Hall/CRC, New York
- Møller J, Ghorbani M, Rubak E (2016) Mechanistic spatio-temporal point process models for marked point processes, with a view to forest stand data. *Biometrics* 72:687–696
- Mrkvička T, Myllymäki M, Jilek M, Hahn U (2018) A one-way ANOVA test for functional data with graphical interpretation. [ArXiv:1612.03608](https://arxiv.org/abs/1612.03608)
- Myllymäki M, Mrkvička, T., Grabarnik, P., Seijo, H., and Hahn, U. (2017) Global envelope tests for spatial processes. *J R Stat Soc Ser B (Stat Methodol)* 79:381–404
- Packalen P, Vauhkonen J, Kallio E, Peuhkurinen J, Pitkänen J, Pippuri I, Strunk J, Maltamo M (2013) Predicting the spatial pattern of trees by airborne laser scanning. *Int J Remote Sens* 34(14):5154–5165
- Penttinen A, Ylitalo A-K (2016) Deducing self-interaction in eye movement data using sequential spatial point processes. *Spatial Stat* 17:1–21
- Pommerening A, Grabarnik P (2019) *Individual-based methods in forest ecology and management*, 1st edn. Springer, Cham
- Richardson JA, Kuester JL (1973) The complex method for constrained optimization. *Commun ACM* 16:487–489
- Ripley BD (1987) *Stochastic simulation*. Wiley, New York
- Snyder DL, Miller MI (1991) *Random point processes in time and space*. Springer, Berlin
- Stoyan D, Rodriguez-Cortes FJ, Mateu J, Gille W (2017) Mark variograms for spatio-temporal point processes. *Spatial Stat* 20:125–147
- Talbot J, Tarjus G, Van Tassel PR, Viot P (2000) From car parking to protein adsorption: an overview of sequential adsorption processes. *Colloids Surf A Physicochem Eng Aspects* 165:28–324
- Vere-Jones D (2009) Some models and procedures for space-time point processes. *Environ Ecol Stat* 16:173–195