THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Modelling and Inference for Spatio-Temporal Marked Point Processes

Ottmar Cronie

CHALMERS | UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
Göteborg, Sweden 2011

Modelling and Inference for Spatio-Temporal Marked Point Processes
Ottmar Cronie
ISBN 978-91-7385-586-0

Doktorsavhandlingar vid Chalmers Tekniska Högskola
ISSN 0346-718X
Ny serie nr 3267

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31-772 1000

Printed in Göteborg, Sweden 2011

Modelling and Inference for Spatio-Temporal Marked Point Processes

Ottmar Cronie

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis deals with inference problems related to the *growth-interaction process* (GI-process). The GI-process is a continuous time spatio-temporal point process with dynamic interacting marks (closed disks), in which the *immigration-death process* (ID-process) controls the arrivals of new marked points as well as their potential life-times. The data considered are marked point patterns sampled at fixed time points and the main area of application of the GI-process is the dynamical modelling of the trees in a forest stands.

The parameters related to the development of the marks are estimated using the least-squares (LS) approach. The death rate, which is assumed to be a function of the mark sizes, and the arrival intensity and are estimated by (approximate) maximum likelihood (ML) methods. We also propose three edge correction methods for discretely sampled (marked) spatio-temporal point processes. The edge correction methods together with the LS approach are applied to fit the GI-process to a forest stand of Scots pines.

We derive the transition probabilities of the (Markovian) ID-process, which form the likelihood function of its two parameters. We further reduce the ML-problem from two dimensions to one dimension. Given an equidistant sampling scheme and some conditions for the parameter space, we manage to prove the consistency and the asymptotic normality of the ML-estimators. The results are also evaluated numerically.

Measurements of locations and radii at breast height (rbh) made at 3 different time points of the individual trees in 10 Swedish Scots pine stands, are modelled spatio-temporally by the GI-process. A new location assignment strategy and a more flexible function for the open-growth (growth in absence of competition) are suggested in order to improve the fit. A linear relationship is found between the site productivity index (fertility) and the sizes of the trees. This relationship is exploited in the estimation of the carrying capacity parameter (theoretical upper bound for the radii). We also test the goodness-of-fit of the fitted model in terms of prediction.

By adding scaled continuous white noise to the mark growth equations, we obtain a system of stochastic differential equation (SDEs) for the mark growth. We consider the case where there is no interaction present and the mark SDEs are independent Cox-Ingersoll-Ross SDEs. Closed form expressions are available both for the transition densities and the stationary distributions. Under the assumption that the mark processes are stationary, consistency and asymptotic normality of the ML-estimators of the parameters are proved.

Keywords: Asymptotic normality, Consistency, Cox-Ingersoll-Ross process, Diffusion process, Edge correction, Goodness-of-fit, Richards growth function, Growth-interaction process, Immigration-death process, Least squares estimation, Markov process, Maximum likelihood estimation, Open-growth, Spatio-temporal marked point process, Stationarity, Stochastic differential equation, Transition density.

Acknowledgments

These last five years have been the most happy, interesting, challenging and rewarding years of my life! When entering the process of writing a PhD thesis, it is impossible to imagine what one will go through on such a journey. I must admit that, at times, I have also felt quite alone in the process, thinking that no one could relate to what I was going through emotionally. However, I now realize that I was very wrong in that respect. All the people who, in one way or another, have been a part of my life throughout this process have not only understood me, but they have supported me and cared about me (probably more than I deserved). This is my opportunity to pay my respect to all of you and thank you, from the bottom of my heart :)

To start with I would like to thank my advisor Aila Särkkä. Words truly are insufficient for express the gratitude that I feel. I would truly like to thank you for the constant support, encouragement and inspiration given to me during these past years! What probably makes me the happiest, however, is the true friendship that we have developed!

I would further like to thank my co-advisors Jun Yu and Anastassia Baxevani. Thank you for being good mentors and colleagues, but most importantly, thank you for being good friends! I have truly learnt a great deal from you!

To one of my co-authors and colleagues, Kenneth Nyström, thank you for these last couple of years. It has been nice working with you and learning from you.

Other people who I would like to thank and show gratitude for ideas and inspiration related to the writing of this thesis include: Claudia Redenbach, Eric Renshaw, Gerald van den Boogaart, Patrik Albin, Peter Guttorp and Serik Sagitov.

Also, a general thank you goes out to all the people at the department of Mathematical Sciences in Gothenburg and a few of you deserve special attention: Daniel, Dmitrii, Frank and Sofia, thank you for these years of happiness and friendship ;) Related to my work (socially or workwise), other people I would like to thank on a personal level include Alexandra, Anders, Carl, Elisabeth, Emilio, Erik, Erik, Fredrik, Hans, Hermann, Hossein, Igor, Jacques, Johan, José, Jovan, Kerstin, Lotta, Magnus, Malin, Martin, Marianne, Marie, Mattias, Michael, Tommy and Torbjörn.

To all my friends outside the department. Thank you for your love, friendship and support! Some of you deserve special attention: Adam, Alain, Alma, Anders, Andreas, Bryan, Daina, Daniel (all of you), Farhad, Felicia, Felipe, Henrik, Hilma, Joel, Jonathan, Kajsa, Laszlo, Marcel, Maria, Mauri, Mattias, Niklas, Raphael, Rikard, Sandra, Sara, Serkan and Simon.

Finally, to the people who mean the most to me. My family. I love you. Thank you for your eternal love and support! You have been with me all the way and I cannot thank you enough for everything. Bobo, thank you for your support and love!

Greetings from Dr Cronie (to be)...

List of Papers

This thesis includes the following papers.

Paper I. Cronie, O., Särkkä, A. (2011).
Some edge correction methods for marked spatio-temporal point process models. *Computational Statistics & Data Analysis* **55**, 2209-2220.

Paper II. Cronie, O., Yu, J. (2010).
The Discretely Observed Immigration-Death Process and its Maximum Likelihood Estimation. *Preprint.*

Paper III. Cronie, O., Nyström, K., Yu, J. (2011).
Spatio-Temporal Modelling of Swedish Scots Pine Stands. *Preprint.*

Paper IV. Cronie, O. (2011).
Likelihood Inference for a Stochastic Version of the Spatio-Temporal Growth-Interaction Process. *Preprint.*

Contents

Chapter 1

Introduction

In many different instances in our surrounding world we find point patterns of different kinds. Such patterns include galaxy locations, locations of earthquake epicentres, locations of cell centres and locations of trees in a forest stand. In order to help analysing point patterns the field of spatial statistics has lent a helping hand and has simultaneously also been developing through it. The field of spatial statistics incorporates a few different disciplines within the field of stochastic mathematics and in this thesis we will focus on the parts played by *stochastic geometry* (the study of random geometrical objects) and *spatial point processes* (the study of random point structures) (see e.g. [7, 10, 20, 21, 33]).

Sometimes one does not solely record the locations of the points in a point pattern but also some additional features connected to each point, such as the radii of the trees in a forest stand or the amount of seismic energy in earthquakes. This additional variable, called a *mark*, can often be quite helpful in explaining the behaviour of the point pattern in question. When focusing on the statistical analysis of these point patterns or marked point patterns, we employ *spatial point processes* or *marked spatial point processes*, respectively (see e.g. [7, 12, 20, 33, 34]). However, to a large extent, the field of spatial (marked) point processes has mainly concentrated on treating marked point patterns within a purely spatial framework. In such a setting one fully ignores that the patterns studied, in fact, almost always are results of evolutionary processes in which the changes occurring among the marks are time dependent. Such situations motivate a change of regime to an approach where one instead considers *spatio-temporal marked point processes* (see e.g. [14, 24, 36]). To fully take the evolution of these marked patterns into consideration it is reasonable to demand that the models describing them should incorporate interaction between marks during the development phases.

The application motivating the work presented in this thesis is found in forestry. Treating a forest stand which is recorded at a specific time point as a static entity, thus ignoring the temporal aspects, the literature offers a wide range of statistical tools for analysing and drawing conclusions about its inherent features, whether one includes marks or not (see e.g. [12, 20, 16, 34] to mention a few). However, here we are interested in modelling the development of a forest stand in both space and time. Figure 1.1 illustrates the type of recorded time series of marked point patterns we refer to – a data set of Swedish Scots pines recorded in 1985, 1990 and 1996 where we have scaled the radii (our marks) for more clear visualisation.

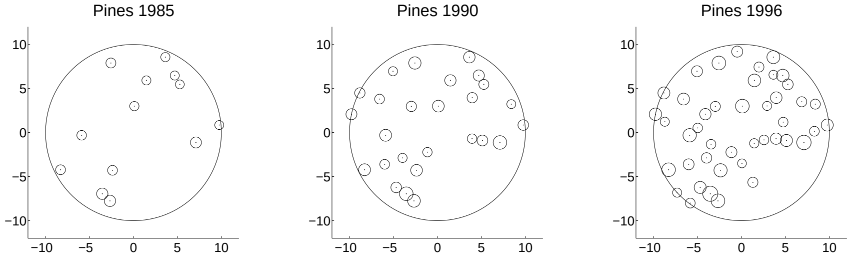

Figure 1.1: Locations and sizes (measured in metres) of Swedish Scots pines recorded in 1985 (left), 1990 (middle) and 1996 (right). The radii of the trees (marks) are scaled by a factor of 10.

A clear risk when formulating the type of spatio-temporal models we are interested in is that the models easily become too involved and we loose both transparency, interpretability and tractability (see e.g. [15]). A spatio-temporal marked point process which manages well to describe this type of spatio-temporal behaviour of a marked population is the so called *Growth-Interaction process* (GI-process) (see [28, 29, 32] or Papers I and III), which is a combination of stochastic and deterministic components (note that it has also been referred to as the *Renshaw-Särkkä growth-interaction model*). It has been used to study, among other things, the development of forest stands [32]. Since this model, in spite of being very flexible, is both tractable and easily interpreted it is quite natural to further assess its potential. In the coming chapters we will present and discuss the GI-process together with different statistical tools developed for it (and other models of this type) and further developments of it.

Since this process deals with both space and time we need to be able to fit, not only, the stochastic process controlling the arrivals and deaths of new marked points in time, but also the mechanism controlling the growth of and interaction between the marks. In the GI-process the arrivals and deaths are controlled by a so-called *immigration-death process* (ID-process) – a continuous time Markov chain.

In both Paper I and Paper II estimators for the two parameters of the ID-process are given and in Paper I we also recall how the growth and interaction parameters of the GI-process are estimated. Additionally in Paper I, for the Scots pine data in Figure 1.1, we evaluate some of the estimators of the GI-process in the context of the edge correction methods developed in Paper I. In Paper II, where the likelihood estimation of the discretely sampled ID-process is tackled, we prove that the obtained likelihood estimators are both consistent and asymptotically normally distributed. Having obtained ideas in Paper I about how to improve the fit of the GI-process, in Paper III the model is altered in the way individuals (trees) are assigned locations in the study region and in the way they grow in absence of competition (open-growth). As the main objective of Paper III is to evaluate further how the GI-process fits Scots pine data, we fit the model to a series of pine data sets of the type presented in Figure 1.1 and finally exploit a set of goodness-of-fit procedures (spatial and forestry related) to assess the fit of the predicted model. In Paper IV we modify the model by adding scaled white noise to the equations which govern the growth of the marks, and hereby the mark growth will be driven instead by a system of stochastic differential equations (SDEs). By utilizing the findings in Paper II and properties of the mark SDEs, under the restriction that there is no interaction present among the individuals, a full likelihood estimation procedure is developed in Paper IV. By putting some additional restrictions on the SDEs, consistency and asymptotic normality are proved.

Chapter 2

The process

We will here define the spatio-temporal marked point process $\Phi_M(t) = \{[X_i, M_i(t)] : i \in \Omega_t\}$, which we refer to as the *Growth-Interaction process* (GI-process) (see e.g. [32], Paper I and Paper IV). Consider a spatial study region W, which either is given by a subset of the Euclidean space $\mathbb{R}^2$ (or possibly $\mathbb{R}^3$) or by a torus (see e.g. [12]). As time $t \in [0, T) \subseteq [0, \infty)$ passes, individuals (marked points) arrive to W at random times and receive locations $X_i \in W$. Additionally, as time passes, the dynamical and interacting marks $M_i(t)$ change size until they die and leave W. In order to keep track of which individuals are alive at a given time, we define the index process $\Omega_t = \{\text{indices of individuals alive at time } t\}$. Before giving a detailed description of the GI-process, however, we first describe the immigration-death process, $\{N(t)\}_{t \geq 0}$, which controls the arrivals and deaths.

2.1 The immigration-death process

The immigration-death (ID) process, $\{N(t)\}_{t \geq 0}$, is a time-homogeneous irreducible continuous-time Markov chain (see e.g. [23]) where the possible states for which transitions $i \to j$ are possible are supplied by the state space $E = \mathbb{N} = \{0, 1, \ldots\}$. It is governed by the parameter pair $\gamma = (\alpha, \mu)$ which we here assume to take values in some compact parameter space $\Gamma \subseteq \mathbb{R}^2_+$.

One way of viewing $\{N(t)\}_{t \geq 0}$ is to treat it as a special case of a birth-death process, for which the infinitesimal transition probabilities are given by

$$p_{ij}(t;\gamma) := \mathbb{P}\left(N(h+t)=j|N(h)=i\right) = \begin{cases} \lambda_i t + o(t) & \text{if } j=i+1 \\ 1-(\lambda_i+\mu_i)t+o(t) & \text{if } j=i \\ \mu_i t + o(t) & \text{if } j=i-1 \\ o(t) & \text{if } |j-i|>1, \end{cases}$$

where the birth rates are given by $\lambda_i = \alpha$, $i=0,1,\ldots$, and the death rates are given by $\mu_i = i\mu$, $i=0,1,\ldots$, (see [17], p. 268-270). Within this framework the interpretation of $\{N(t)\}_{t\geq 0}$ is the following. By letting the arrivals of new individuals to a population occur according to a Poisson process with intensity α and upon arrival assigning to all individuals independent and exponentially distributed lifetimes with mean $1/\mu$, $N(t)$ gives us the number of individuals alive at time t. Another possibility is to view it as an $M/M/\infty$ queuing system; each customer (arriving according to a Poisson process with intensity α) is being handled by its own server so that its sojourn time in the system is exponential with intensity μ and independent of all other customers.

Being a Markov process, the finite dimensional distributions of $\{N(t)\}_{t\geq 0}$ are controlled by its transition probabilities, $p_{ij}(t;\gamma)$ which are given in Paper II.

Proposition 2.1.1 (Paper II). *The transition probabilities of the ID-process are given by convolutions of Poisson densities and Binomial densities, i.e.*

$$\begin{aligned} p_{ij}(t;\gamma) &= \left(f_{Poi(\rho)} * f_{Bin(i,\mathrm{e}^{-\mu t})}\right)(j) \\ &= \sum_{k=0}^{j} f_{Poi(\rho)}(k) f_{Bin(i,\mathrm{e}^{-\mu t})}(j-k) = \sum_{k=0}^{i\wedge j} f_{Poi(\rho)}(j-k) f_{Bin(i,\mathrm{e}^{-\mu t})}(k) \\ &= \frac{\mathrm{e}^{-\frac{\alpha}{\mu}\left(1-\mathrm{e}^{-\mu t}\right)}}{j!} \sum_{k=0}^{j} \left(\frac{\alpha}{\mu}\right)^k \binom{j}{k} \frac{\mathrm{e}^{-(j-k)\mu t}}{(1-\mathrm{e}^{-\mu t})^{j-2k-i}} \frac{i!}{(i-(j-k))!} \end{aligned}$$

where $i,j \in E = \mathbb{N}$, $\gamma = (\alpha,\mu) \in \Gamma \subseteq \mathbb{R}_+^2$, $f_{Poi(\rho)}(\cdot)$ *is the Poisson density with parameter* $\rho = \frac{\alpha}{\mu}\left(1-\mathrm{e}^{-\mu t}\right)$, *and* $f_{Bin(i,\mathrm{e}^{-\mu t})}(\cdot)$ *is the Binomial density with parameters* i *and* $\mathrm{e}^{-\mu t}$. *Moreover, we have that the probability generating function (p.g.f.) of* $(N(s+t)|N(s)=i)$ *is given by*

$$G_i(s;\gamma) = \left(1+(s-1)\,\mathrm{e}^{-\mu t}\right)^i \mathrm{e}^{\rho(s-1)} \tag{2.1}$$

and

$$\begin{aligned} \mathbb{E}[N(s+t)|N(s)=i] &= i\,\mathrm{e}^{-\mu t} + \rho \qquad (2.2) \\ \mathbb{E}[N^2(s+t)|N(s)=i] &= i(i-1)\,\mathrm{e}^{-2\mu t} + (1+2\rho)i\,\mathrm{e}^{-\mu t} + \rho^2 + \rho. \end{aligned}$$

The interpretation of $p_{ij}(t;\gamma)$ is quite clear. Note that

$$\begin{aligned} f_{Poi(\rho)}(j-k) &= \mathbb{P}(j-k \text{ new arrivals during } (h,h+t)) \\ f_{Bin(i,\mathrm{e}^{-\mu t})}(k) &= \mathbb{P}(k \text{ of the } i \text{ individuals alive at time } h \text{ survive } (h,h+t)), \end{aligned}$$

so that $p_{ij}(t;\gamma)$ expresses the sum of the probabilities of all possible ways in which we can decrease i individuals to j individuals. Furthermore, when $i \leq j$, we get that $p_{ij}(t;\gamma)$ simply represents the convolution of the $Bin(i,\mathrm{e}^{-\mu t})$-density and the $Poi(\rho)$-density. One can easily show that for the marginal distributions of $\{N(t)\}_{t\geq 0}$ we have that $\mathbb{P}(N(t)=j|N(0)=0)=\mathrm{e}^{-\rho}\rho^j/j!$, i.e. $(N(t)|N(0)=0) \sim Poi\big(\frac{\alpha}{\mu}(1-\mathrm{e}^{-\mu t})\big)$, and that $(N(t)|N(0)=0) \xrightarrow{d} Poi(\alpha/\mu)$ as $t \to \infty$. Note that this invariant distribution is unique due to the positive recurrence, and it is also the same as its asymptotic distribution since every asymptotic distribution is an invariant distribution.

Proposition 2.1.2. *The ID-process is ergodic with invariant distribution* π_N *given by the Poisson distribution with mean* α/μ, *i.e.* $\pi_N(\cdot) = \mathbb{P}(Poi(\alpha/\mu) \in \cdot)$.

A further characterisation of $\{N(t)\}_{t\geq 0}$ which sometimes is useful to exploit is to consider $\{N(t)\}_{t\geq 0}$ as a Markov jump process (see Paper II).

Proposition 2.1.3 (Paper II)**.** *Let* $\gamma = (\alpha,\mu) \in \Gamma \subseteq \mathbb{R}^2_+$. $\{N(t)\}_{t\geq 0}$ *is a Markov jump process with state space* $E = \mathbb{N}$, *jump intensity function*

$$\lambda(\gamma;i) = \alpha + \mu i, \quad i \in E,$$

and transition kernel $r(\gamma;\cdot) = \{r(\gamma;i,j) : i,j \in E\}$, *where*

$$r(\gamma;i,j) = \frac{1}{\alpha+\mu i}\left(\alpha\mathbf{1}\{j=i+1\} + \mu i\mathbf{1}\{j=i-1\}\right), \quad i,j \in E.$$

2.2 The GI-process

The process $\Phi_M(t) = \{[X_i, M_i(t)] : i \in \Omega_t\}$ can be described as follows. As time elapses, the arrivals in time of new individuals to $W \subseteq \mathbb{R}^2$ and the time these individuals live in W are governed by an ID-process, $N(t)$, having arrival/birth rate $\alpha\nu(W)$ and death rate μ, where $\nu(\cdot)$ denotes volume in $\mathbb{R}^2$. Furthermore, for the $N \sim Poi(\alpha\nu(W)T)$ individuals who arrive during $[0,T)$, upon arrival at times $B_1,\dots,B_N$ they are assigned locations $X_i \in W$ (precise description given below) and initial marks $M_i(B_i) = M_i^0$, $i = 1,\dots,N$, with the latter taken either as some fixed positive value (as will be the case here), or as a value drawn from some suitable distribution ([32] considers $M_i^0 \sim Uni(0,\epsilon)$,

$\epsilon > 0$). When an individual's ($Exp(\mu)$-distributed) life time has expired at time D_i we say that its has suffered a *natural death* and we set its size to 0.

Once individual i has arrived it starts growing deterministically according to

$$M_i(t) \quad = \quad M_i^0 + \int_{B_i}^{t} dM_i(s), \quad B_i \le t \le D_i, \tag{2.3}$$

where

$$dM_i(t) \quad = \quad f(M_i(t);\theta)dt - \sum_{\substack{j\in\Omega_t \\ j\neq i}} h\left(M_i(t), M_j(t), X_i, X_j;\theta\right)dt.$$

Here $\Omega_t = \{i \in \{1,\dots,N\} : \text{individual } i \text{ is alive at time } t\}$, θ is a parameter vector, the function $f\left(M_i(t);\theta\right)$ determines the open-growth of mark i (growth in absence of competition with other (neighbouring) individuals) and $h\left(M_i(t), M_j(t), X_i, X_j;\theta\right)$ is a function handling the individual's spatial (pairwise) interaction with other individuals. We note that as a radius/mark $M_i(t)$ changes with time, also the closed disk $B_{X_i}[M_i(t)]$ with centre X_i and radius $M_i(t)$, which denotes the occupied space, will change in size.

In addition to the natural death, an individual can die *competitively* which we consider to happen as soon as $M_i(t) \le 0$, and the we set $M_i(t) = 0$ once this happens.

In Papers I and IV we assign the locations to the individuals according to $X_i \sim Uni(W)$. We note that for this choice, when we ignore the competitive deaths, at each fixed time t the locations form a spatial Poisson process with intensity $\frac{\alpha}{\mu}(1 - \mathrm{e}^{-\mu t})$, which is restricted to W. In Paper III, however, we let $X_i \sim Uni(W \setminus \bigcup_{i\in\Omega_{j\in\Omega_{B_i}}} B_{X_j}[M_j(B_i)])$, i.e. we let the location of the ith individual be uniformly distributed on the part of W which is not covered by other trees.

We note that in the case of no interaction, i.e. when $h(\cdot) = 0$, expression (2.3) turns into $dM(t)/dt = f\left(M(t);\theta\right)$, $M(0) = M_0$, which has $M(t)$ as its solution (for simplicity we here write $M(t)$ for $M_i(t)$). The literature offers a wide range of possible choices for the growth function (see e.g. [31]), and one of the models considered in this thesis is the Richards growth function (see e.g. [28, 31]), which is given by

$$\begin{aligned}
\frac{dM(t)}{dt} = f\left(M(t);\theta\right) \quad &= \quad \frac{\lambda}{\delta} M(t)\left(\left(\frac{K}{M(t)}\right)^{\delta} - 1\right), \\
M(t) \quad &= \quad K\left(1 + \left((M_0/K)^{\delta} - 1\right)\mathrm{e}^{-\lambda t}\right)^{1/\delta},
\end{aligned}$$

which is a strictly increasing growth function with *carrying capacity* (upper bound/asymptote) $K > 0$ and growth rates $\delta \neq 1$ and $\lambda > 0$. If we set $\delta = -1$, we obtain as special case the so called logistic growth function. The logistic growth function has been considered in most of the papers which deal with the GI-process (see e.g. [32] or Paper I), and the more general Richards growth function has been employed in both [28] and in Paper III for the modelling of Scots pine stands.

Just as for the individual growth function, the possible choices of spatial interaction functions are many (c.f. [22, 28, 32] for examples of interaction functions and related discussions). One example is given by (see [32])

$$h\left(M_i(t), M_j(t), X_i, X_j; \theta\right) = c\mathbf{1}\left\{B_{X_i}[rM_i(t)] \cap B_{X_j}[rM_j(t)] \neq \emptyset\right\},$$

where $\mathbf{1}\{A \in \cdot\}$ denotes the indicator function for the set A, $c \in \mathbb{R}$ is the force of interaction and $r > 0$ is the scale of interaction. Furthermore, the closed disk $B_{X_i}[rM_i(t)]$ with centre X_i and radius $rm_i(t)$ is referred to as the 'influence zone' of individual i. Since competition for resources takes place only within influence zones ([3, 37]), individuals i and j will compete only when their influence zones intersect, i.e. when $B_{X_i}[rM_i(t)] \cap B_{X_j}[rM_j(t)] \neq \emptyset$. This symmetric interaction function has the effect that small individuals have the same impact on large (neighbouring) individuals as the large individuals have on small individuals. Unless our forest stand consists of trees of similar size, this interaction function becomes unrealistic. In order to circumvent this problem we here consider instead the so called area interaction function, given by

$$h\left(M_i(t), M_j(t), X_i, X_j; \theta\right) = c\,\frac{\nu\left(B_{X_i}[rM_i(t)] \cap B_{X_j}[rM_j(t)]\right)}{\nu\left(B_{X_i}[rM_i(t)]\right)}, \tag{2.4}$$

This non-symmetric soft core interaction has the effect that large marks influence small marks more than the other way around, yet allowing the small marks to play their part. This interaction model is more realistic in tree modelling applications than symmetric interaction models (see [28, 32]). Depending on the choice of parameters, this area interaction function has the ability to generate regular as well as aggregated point patterns (despite the possible underlying uniform distribution of the locations) [27]. Note that the parameter r determines how large the range of interaction is and c mainly determines how regular the point patterns are.

2.2.1 The natural death rate

As previously mentioned the so called natural deaths are governed by the death process part of the ID-process. In situations where it seems plausible that the natural deaths depend on an individual's size, we may let the death rate be given by some function $\mu\eta(\cdot)$, $\mu > 0$, where $\eta(\cdot)$ is a function of the marks. This means that as time passes the $Exp(\mu\eta(M_i(t)))$-distributed remaining lifetime of an individual will change with its size. An alternative way of expressing the behaviour of the death process is to say that the conditional probability that an individual i dies naturally during $(t, t + dt)$, given $M_i(t)$, equals $\mu\eta(M_i(t))\,dt + o(dt)$. Note that if $\eta(\cdot) \equiv 1$, we retrieve the ordinary ID-process. In Paper I, we choose to evaluate the GI-process under $\eta(M_i(t)) = 1/(1 + M_i(t))$ which implies that individuals become more viable as they grow; a choice motivated by our forestry applications. In Papers II and IV, as well as in [27, 28, 29, 32] the model is chosen to have $\eta(\cdot) \equiv 1$.

2.2.2 Remarks about the competitive death

As previously mentioned, one of the possible death occurrences present in the GI-process is the competitive death. Consider the infinitesimal-size interval $(t, t + dt)$ and recall that we classify an individual as having died from competition in $(t, t + dt)$ if $M_i(t) > 0$ and $M_i(t + dt) \leq 0$. Let us call this scenario 1. Consider now an alternative approach, which we call scenario 2, where the individual suffers a competitive death if $M_i(t) > 0$ and $dM_i(t) < 0$. Now a reasonable question emerges, namely, which of the two scenarios should be used to represent competitive/interactive death for tree data. In a tree stand model one could argue that scenario 1 is a more appropriate view than scenario 2 since trees do not disappear immediately after they die. This thus indicates that they should not be removed as soon as $dM_i(t) < 0$, since dead trees occupy the ground where they have been standing some time after their deaths. Also, to some extent, dead trees inhibit the nutrient access and light absorption of other trees close to it. Furthermore, it is not reasonable that a new tree would end up very close to the centre of a one. Although a bit artificial in its nature we thus have chosen to use of scenario 1 to represent competitive deaths, just as in [32].

2.3 The stochastic GI-process

In the case where the mark equations of expression (2.3) are instead given by stochastic differential equations (SDEs), we refer to $\Phi_M(t) = \{[X_i, M_i(t)] : i \in \Omega_t\}$ as the spatio-temporal *stochastic growth-interaction* process. In the simplified case where there is no interaction between the individual, i.e. $h(\cdot) = 0$, we will simply refer to Φ_M as the spatio-temporal *stochastic growth* (SG) process. Note that in this case there is no competitive death.

We build Φ_M in two steps: The underlying marked point process $\Phi(t)$, governs the locations of the individuals on W as well their arrival times and lifetimes, and the marking process, which may be regarded as an extension of Φ, assigns the set of processes which control the growth of the marks (disks) $B_{X_i}[M_i(t)]$.

We start by describing the underlying process Φ. Consider an ID-process $N(t)$ which controls the arrival times $B_1, \ldots, B_N$, lifetimes $L_1, \ldots, L_N$ and death times $D_i = \min(B_i + L_i, T) = (B_i + L_i) \wedge T$ of the $N = \Phi(T) \sim Poi(\alpha T \nu(W))$ individuals who arrive to W at the locations $X_i \sim Uni(W)$. The process $\Phi(t)$ is given by the marked Poisson process $B_1, \ldots, B_N$ on $[0, T)$ for which the marks are given by the pairs (L_i, X_i) (note that conditional on N, $B_i \sim Uni(0, T)$). Note that here $\Omega_t = \{i \in \{1, \ldots, N\} : t \in [B_i, D_i]\}$, $\Omega_0 = \emptyset$, and $N(t) = |\Omega_t|$.

We now turn to the second part of Φ_M. Given some suitable diffusion coefficient $\sigma(x)$ and independent standard Brownian motions $W_i(t)$, $i = 1, \ldots, N$, we have that, loosely speaking, the mark radii are controlled by the system of SDEs $(dM_1(t), \ldots, dM_N(t))$, where

$$\begin{aligned} dM_i(t) &= f(M_i(t); \theta)dt - \sum_{\substack{j \in \Omega_t \\ j \neq i}} h\left(M_i(t), M_j(t), X_i, X_j; \theta\right) dt \\ &\quad + \sigma(M_i(t)) dW_i(t), \end{aligned}$$

and $M_i(t) = 0$ for $t \notin [B_i, D_i]$.

In the special case of the SG-process, which is studied in Paper IV, we let $f(x) = \lambda(1 - x/K)$, $h(\cdot) = 0$ and $\sigma(x) = \sigma\sqrt{x}$, whereby we obtain a system of independent (time-shifted) Cox-Ingersoll-Ross (CIR) processes for the growth of the marks (see e.g. [6, 13, 19]). We make this precise by letting $t \in [0, T)$ denote our *global time* and consider the ith CIR-process $\{Y_i(t)\}_{t \in [0,T)}$ where, given $Y_i(0) = M_i^0$ and the Brownian motion $W_i(t)$, the SDE generating $Y_i(t)$ is given by

$$dY_i(t) = \lambda\left(1 - Y_i(t)/K\right) dt + \sigma\sqrt{Y_i(t)} dW_i(t), \tag{2.5}$$

so that its integral form is given by

$$Y_i(t) = M_i^0 + \int_0^t \lambda \left(1 - \frac{Y_i(s)}{K}\right) ds + \int_0^t \sigma\sqrt{Y_i(s)} dW_i(s). \tag{2.6}$$

By then letting $\tau_i(t) = t - B_i$ be our ith *local time* and defining

$$M_i(t) = \begin{cases} Y_i(\tau_i(t)) & \text{for } t \in [B_i, D_i] \\ 0 & \text{for } t \notin [B_i, D_i] \end{cases}, \tag{2.7}$$

we have expressed $M_i(t)$ by means of the global time scale.

The parameters $(\lambda, K, \sigma) \in \Theta_\lambda \times \Theta_K \times \Theta_\sigma \subseteq \mathbb{R}_+^3$ in this mean-reverting SDE control different aspects of the growth: The *diffusion coefficient* σ controls the magnitude of the random individual fluctuations of the radii. The interpretation of the remaining two parameters becomes most clear by noticing that $Y_i(t)$ is a so called mean-reverting process: As $Y_i(t)$ starts to move away from its long term equilibrium K, the drift term starts pulling it back towards K and the speed at which this occurs is given by λ/K. Related to this interpretation we find that if we set $\sigma = 0$ in expression (2.5), we retrieve the GI-process (without interaction) and the size development of the disks will comprise the differential equation $dY_i(t) = \lambda\,(1 - Y_i(t)/K)\,dt$. This differential equation is often referred to as the linear growth function (see e.g. [28, 32]) and in this setting the parameter λ is referred to as the (individual) *growth rate* and recall that the upper bound K is the carrying capacity. In conclusion, Φ_M is controlled by the parameter vector $\theta = (\lambda, K, \sigma, \alpha, \mu) \in \Theta = \Theta_\lambda \times \Theta_K \times \Theta_\sigma \times \Theta_\alpha \times \Theta_\mu \subseteq \mathbb{R}_+^5$.

We note that we also may treat $\Phi_M(t)$ as a multivariate (N-dimensional) diffusion (hence a Markov process), for which all components are independent, stopped and time-shifted CIR-processes.

Regarding the initial size $M_i(B_i) = Y_i(0) = M_i^0$, a few different options are available. As previously mentioned, the choices $M_i^0 \equiv M_0 \in \mathbb{R}_+$ and $M_i^0 \sim Uni(0, \epsilon)$, $\epsilon > 0$, have already been explored (see e.g. [32] and Paper I). Here, however, we also have the further option to sample each $M_i(B_i)$ from the stationary distribution of Y_i, which turns the radius diffusion processes into strictly stationary processes.

It is also possible to represent Φ_M as a spatial entity, under the condition that $T < \infty$. By restricting a homogeneous spatial Poisson process on $\mathbb{R}^2$ with intensity αT to $W \subseteq \mathbb{R}^2$ (see e.g. [10, 12, 30, 33]), we obtain the Poisson process $\Phi' = \{X_1, \ldots, X_N\}$ with intensity measure $\Lambda(B) = \alpha T \nu(B \cap W)$, $B \in \mathcal{B}(\mathbb{R}^2)$, where $\nu(\cdot)$ denotes Lebesgue measure and $\mathcal{B}(\mathbb{R}^2)$ are the Borel sets in $\mathbb{R}^2$. By now considering $\Phi_M = \{[X_i, M_i([0,T); B_i, L_i)]\}_{i=1}^N$, which is a marked version of Φ' such that the ith mark is given by the random element $M_i([0,T); B_i, L_i) : \mathcal{X} \to V_i = \{f \in C_{[0,T)}^+ : \mathrm{supp}(f) = [B_i, D_i]\}$, where $C_{[0,T)}^+ = \{f : [0,T) \to \mathbb{R}_+ : f > 0, f \text{ continuous}\}$ and $supp(f)$ denotes the support of the function f, we have obtained a different representation of the SGI-process. Note that Φ_M is a marked spatial Poisson process for which the marks are random elements which take values in the function space $\{f \in C_{[0,T)}^+ : \mathrm{supp}(f) \subseteq [0,T)\}$, and in the case of the SGI-process these random elements are dependent.

2.3.1 Distributional properties of the SG-process

We give here some results concerning different properties of the CIR-process (they can be found in e.g. [6, 19]) and then turn to the finite dimensional distributions (fdds) of the SG-process (its likelihood function is given by the joint density).

When $2\lambda \geq \sigma^2$ the process $Y_i(t)$ stays strictly positive [6], and we note that this means that the drift of the SDE $dY_i(t)$ must be large enough, in comparison to the diffusion term, to ensure that the mean-reversion is strong enough to keep the process a.s. positive. By recalling that the individual is alive if $M_i(t) > 0$, it becomes clear that we will have to require that $2\lambda \geq \sigma^2$ so that $M_i(t) > 0$ for all $t \in [B_i, D_i]$. Moreover, since $Y_i(t)$ is a Markov process, when we require that $2\lambda \geq \sigma^2$, it is possible to derive explicit statements about the transition distributions, i.e. the distributions of the random variables $Y_i(t)|Y_i(s)$, $s \leq t$. For instance, under the hypothesis that $2\lambda \geq \sigma^2$ and $s \leq t$, the transition density of $Y_i(t)$, conditional on $Y_i(s) = y_s$, is given by the noncentral χ^2-distribution density

$$p_{Y_i}(t-s, y_t|y_s; \lambda, K, \sigma) = a\,\mathrm{e}^{-(u+v)} \left(\frac{v}{u}\right)^{q/2} I_q\left(2\sqrt{uv}\right), \tag{2.8}$$

where $a = 2\lambda/\left(\sigma^2 K\left(1-\mathrm{e}^{-(t-s)\lambda/K}\right)\right)$, $u = ay_s\,\mathrm{e}^{-(t-s)\lambda/K}$, $v = ay_t$ and $q = 2\lambda/\sigma^2 - 1$. The function $I_q(x) = \sum_{k=0}^{\infty}(x/2)^{2k+q}/k!\Gamma(k+q+1)$, $x \in \mathbb{R}$, where $\Gamma(\cdot)$ denotes the gamma function, is the modified Bessel function of the first kind of order q.

The ergodic process $Y_i(t)$ also has a stationary (invariant) distribution $\pi = \pi_{\lambda,K,\sigma}$ which is given by the Gamma distribution with shape parameter $2\lambda/\sigma^2$ and scale parameter $\sigma^2 K/2\lambda$. Hereby, the density of the stationary distribution is given by

$$\pi(x; \lambda, K, \sigma) = \frac{\left(2\lambda/\sigma^2 K\right)^{2\lambda/\sigma^2}}{\Gamma(2\lambda/\sigma^2)} x^{2\lambda/\sigma^2-1}\,\mathrm{e}^{-x(2\lambda/\sigma^2 K)}, \quad x \geq 0, \tag{2.9}$$

so that π has mean K and variance $\sigma^2 K^2/2\lambda$ and, moreover, for $s < t$, the covariance function of Y_i is given by $\mathrm{Cov}(Y_i(s), Y_i(t)) = \frac{\sigma^2 K^2}{2\lambda}\,\mathrm{e}^{-(t-s)}$. As previously mentioned $Y_i(t)$ is a Markov process and given that we start a Markov process in its stationary distribution, it is a strictly stationary process. In the case of $Y_i(t)$ this means that $Y_i(0) = M_i^0 \sim \pi$ and that its fdds are shift invariant w.r.t. time, i.e. $(Y_i(T_1), \ldots, Y_i(T_n)) =^d (Y_i(T_1+h), \ldots, Y_i(T_n+h))$ for any set of times $T_1 < \ldots < T_n$, any $h \geq 0$ and any $n \in \mathbb{N}$. Hereby the marginal/transition distributions do not change, i.e. for any (s,t), $t > s \geq 0$, $Y_i(t) \sim \pi$ and $Y_i(t)|Y_i(s) \sim \pi$.

Since both the ID-process and the CIR-process are Markov processes, also the SG-process will be Markovian, and this fact is exploited in Proposition 2.3.1, where the fdds of $\Phi_M(t)$ are given. To set the framework, consider now the (sample) times $0 = T_0 < T_1 < \ldots < T_n \leq T$ and the distribution of $(\Phi_M(T_1), \ldots, \Phi_M(T_n))^T$, when we are concerned with exactly, say, $d \in \{1, \ldots, N\}$ individuals which appear at $T_1, \ldots, T_n$ (recall that N is the total number of individuals observed if we monitor the process continuously). Furthermore, provided that the joint density of $(\Phi_M(T_1), \ldots, \Phi_M(T_n))^T$ exists, when evaluated at the size-time matrix

$$\mathbf{M} = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{d1} & \cdots & m_{dn} \end{pmatrix} \in \mathbb{R}^{d\times n},$$

we will denote it by $\mathbf{p}_{T_1,\ldots,T_n}(\mathbf{M};\theta)$. It should be emphasized that the ith row of $\mathbf{M}$ represents the evaluation-sizes of the ith individual under consideration, at the respective times $T_1, \ldots, T_n$. We further also note that if $m_{ik} = 0$, we are considering the case where the ith individual is not alive at time T_k. Consequently, if a row were to contain only zeros, we would be considering an individual who is not alive at any of $T_1, \ldots, T_n$, whence that individual/row may be removed from consideration.

Proposition 2.3.1 (Paper IV). *Given $0 = T_0 < T_1 < \ldots < T_n \leq T$ and $\Phi_M(T_0)$, if we let $M_i^0 = M_0 > 0$ for all i, then the joint density of $(\Phi_M(T_1), \ldots, \Phi_M(T_n))^T$, evaluated at $\mathbf{M} \in \mathbb{R}^{d\times n}$, $d \geq 1$, is given by*

$$\begin{aligned} \mathbf{p}_{T_1,\ldots,T_n}(\mathbf{M};\theta) &= C\prod_{k=1}^{n} p_N\left(\Delta T_k, |\omega_k| \Big| |\omega_{k-1}|; \alpha\nu(W), \mu\right) \qquad (2.10) \\ &\quad \times \prod_{k=1}^{n} \prod_{i\in\omega_{k-1}\cap\omega_k} p_{Y_1}(\Delta T_k, m_{ik}|m_{i(k-1)}; \lambda, K, \sigma) \\ &\quad \times \prod_{i=1}^{d} \int_{T_{k_i-1}}^{T_{k_i}} \frac{p_{Y_1}(T_{k_i} - t, m_{i(k_i-1)}|M_0; \lambda, K, \sigma)}{T_{k_i} - T_{k_i-1}} dt, \end{aligned}$$

where $\Delta T_k = T_k - T_{k-1}$ and $\omega_k = \{i : m_{ik} > 0\}$, $k = 1, \ldots, n$, and $k_i = \min\{k : i \in \omega_k\}$, $i = 1, \ldots, d$. $C = C(\nu(W), \mathbf{M})$ is a positive constant, and the densities $p_{Y_1}(\cdot)$ and $p_N(\cdot)$ are given, respectively, by expression (2.8) and Propostion 2.1.1.

We recall that when $M_i^0 \sim \pi$, the process $Y_i(t)$ is a strictly stationary process and this will have a further impact on the joint densities in Proposition 2.3.1.

Corollary 2.3.1 (Paper IV)**.** *Given the preliminaries and notation of Proposition 2.3.1, by instead assuming that* $M_i^0 \sim \pi$*, the joint density (2.10) becomes*

$$\begin{aligned} \mathbf{p}_{T_1,\ldots,T_n}(\mathbf{M};\theta) &= C\prod_{k=1}^{n} p_N\left(\Delta T_k, |\omega_k|\middle||\omega_{k-1}|;\alpha\nu(W),\mu\right) \\ &\times \prod_{k=1}^{n}\prod_{i\in\omega_k}\pi(m_{ik};\lambda,K,\sigma). \end{aligned} \tag{2.11}$$

We may additionally require that also $N(t)$ starts in its stationary distribution π_N (see Proposition 2.1.2) so that also $N(t)$ becomes a strictly stationary process. Hereby the transition probabilities $p_N\big(\Delta T_k, |\omega_k|\big||\omega_{k-1}|;\alpha\nu(W),\mu\big)$ in (2.11) in the above corollary will be replaced by $\pi_N\left(|\omega_k|;\alpha\nu(W),\mu\right)$, which is given in Proposition 2.1.2. Note that this change will imply that $N(t) = |\Omega_t| \sim Poi(\alpha/\mu)$ for all $t \geq 0$ and under this setup, since all Y_i's are stationary, we have that $M_i(0) \sim \pi$ for all individuals $i \in \Omega_0$.

We note further that if $M_i^0 \sim \pi$, conditionally on $\Omega_0 = \emptyset$, the process $\Xi(t) = \bigcup_{i\in\Omega_t} B_{X_i}[M_i(t)]$ at each fixed time t corresponds to a Boolean model (see e.g. [33]) with *germs* $\{X_i\}_{i\in\Omega_t}$ generated from a Poisson process with intensity measure $\Lambda_t(B) = \frac{\alpha}{\mu}(1-\mathrm{e}^{-\mu t})\nu(B\cap W)$, $B \in \mathcal{B}(\mathbb{R}^2)$, and *grains* given by $\{B_{X_i}[M_i(t)]\}_{i\in\Omega_t}$, where all $M_i(t)$'s are iid $\Gamma(2\lambda/\sigma^2, \sigma^2 K/2\lambda)$-distributed. Note that this follows since Ω_t can be generated as a thinned Poisson process (see Paper II).

Chapter 3

Parameter estimation

Assume now that we sample the process at times $0 = T_0 < \ldots < T_n = T$. Then, for each $k = 1, \ldots, n$, this gives rise to a sampled marked point configuration $\mathbb{X}(T_k) = \{[\mathbf{x}_i, m_{ik}] : i \in \Omega_{T_k}\}$ (Figure 1.1 illustrates such a scenario). We start by considering the estimation of GI-process (Papers I and III), then we consider the (asymptotic) Maximum Likelihood (ML) inference for the ID-process (Paper II) and the SG-process (Paper IV), and we finally briefly discuss the edge correction methods developed in Paper I.

3.1 Estimation of the GI-process parameters

3.1.1 Estimation of the growth and interaction parameters

The following least squares approach for estimating the mark related parameters, $\theta = (\lambda, K, c, r) \in \mathbb{R}_+^2 \times \mathbb{R} \times \mathbb{R}_+$, and method for the labeling of naturally dead individuals originally was suggested in [32]. We here present it in the context of an open-growth function with two parameters λ and K and an interaction function with the parameters c and r. The procedure can easily be altered to accommodate any other open-growth and interaction functions. Let $\left\{\tilde{m}_i\left(T_{k+1}; \theta, \mathbb{X}(T_k)\right) : i \in \Omega_{T_k}\right\}$ denote the set of predictions of the actual data marks, $\left\{m_{i(k+1)} : i \in \Omega_{T_k}\right\}$, generated by equation (2.3) under the regime of θ, based on the configuration $\mathbb{X}(T_k)$ (in practice we employ the simulation algorithm presented in [32] in order to create each predicted set). If the predicted mark indicates that the individual is alive but the individual is dead in reality, this predicted individual will be treated as having died by natural causes during (T_k, T_{k+1}). The least squares estimates are then found by minimising

$$S(\theta) := \sum_{k=1}^{n-1} \sum_{i \in \Omega_{T_k}} \mathbf{1}\{i \in \Omega_{T_{k+1}}\} \left[\tilde{m}_i\left(T_{k+1}; \theta, \mathbb{X}(T_k)\right) - m_{i(k+1)}\right]^2$$

with respect to $\theta = (\lambda, K, c, r) \in \mathbb{R}_+^2 \times \mathbb{R} \times \mathbb{R}_+$, where $\mathbf{1}\{i \in \Omega_{T_{k+1}}\}$ is an indicator function being 1 if the actual data individual i is alive at time T_{k+1}.

In order to minimize $S(\theta)$ some optimization procedure is required. The approach used in [32] is to create a grid of parameter values for each of the parameters in $\theta = (\lambda, K, c, r)$ and then calculate $S(\theta)$ for all combinations of values taken from these grids. One then lets $\hat{\theta} = (\hat{\lambda}, \hat{K}, \hat{c}, \hat{r})$ be given by the combination of grid values which gives rise to the smallest value of $S(\theta)$ and either accepts $\hat{\theta}$ as one's final estimate or one creates a new, finer, grid centred around the estimated parameter values in $\hat{\theta}$ and repeats the procedure a number of times until no change in $\hat{\theta}$ takes place and the grids have all become very dense. This procedure encounters the problem that the actual optimal combination of parameters may fall outside the grids, as the grids are becoming finer, if the initial grid is not chosen correctly. Another approach which is similar in its nature to the grid search, still avoiding the aforementioned problem, is to repeatedly draw parameter values $\theta = (\lambda, K, c, r)$ where $\lambda \sim Uni(\lambda_L, \lambda_U)$, $K \sim Uni(K_L, K_U)$, $c \sim Uni(c_L, c_U)$, $r \sim Uni(r_L, r_U)$ and for each such combination calculate $S(\theta)$, choosing as final estimate the parameter combination giving rise to the smallest $S(\theta)$. This MCMC type of method, however, has the drawback that one needs to make a choice on the upper and lower bounds in the uniform distributions being drawn from. One could handle this by choosing initial intervals on which we sample while successively extending the intervals if candidates near the boundaries are the ones minimizing $S(\theta)$. Note that we do not have to bother too much about the lower bounds since most of the parameters are bounded below by 0.

Paper I adopts an MCMC-type method (see [26]) where we start by choosing initial parameter estimates, i.e. let $\lambda = \lambda_0 > 0$, $K = K_0 > 0$, $c = c_0 > 0$ and $r = r_0 > 0$, for which we calculate $S(\theta) = S(\lambda, K, c, r)$. We also define the step sizes $\delta_\lambda > 0$, $\delta_K > 0$, $\delta_r > 0$, and $\delta_c > 0$. Now, in each round we

1. randomly choose one of the parameters λ, K, r, c;
2. for our parameter of choice, say λ, let $\lambda' = \lambda + Z$, for Z drawn from $Uni(-\delta_\lambda, \delta_\lambda)$;
3. calculate $S(\theta') = S(\lambda', K, r, c)$;

4. if $S(\theta') < S(\theta)$ let $\lambda = \lambda'$, otherwise let $\lambda = \lambda$;
5. return to step 1.

We continue to run the algorithm until either $S(\theta)$ is less than some predefined minimum value or until we have not seen any decrease in $S(\theta)$ for a predefined number of consecutive runs. We let our final estimates $\hat{\theta} = (\hat{\lambda}, \hat{K}, \hat{c}, \hat{r})$ be given by the last θ obtained in the algorithm above. Note that we here utilize the information obtained in the previous step in order to stepwise get closer to the final estimate.

When minimizing $S(\theta)$, in the case of a simulated data set, it can be seen that $S(\theta)$ may not attain its minimum at the true parameter set but instead at some biased θ. This 'incorrect' shape of $S(\theta)$ is mainly due to edge effects and dependence between certain parameters. This phenomenon is illustrated in Figure 3.1. It is a plot of $S(\theta)$ as a function of only λ and K, where c and r are kept fixed at their actual values. Note that we have used the logistic growth function and the area interaction function. It is clear from the graph that $S(\theta)$ is decreasing as λ moves away from its actual value 0.2.

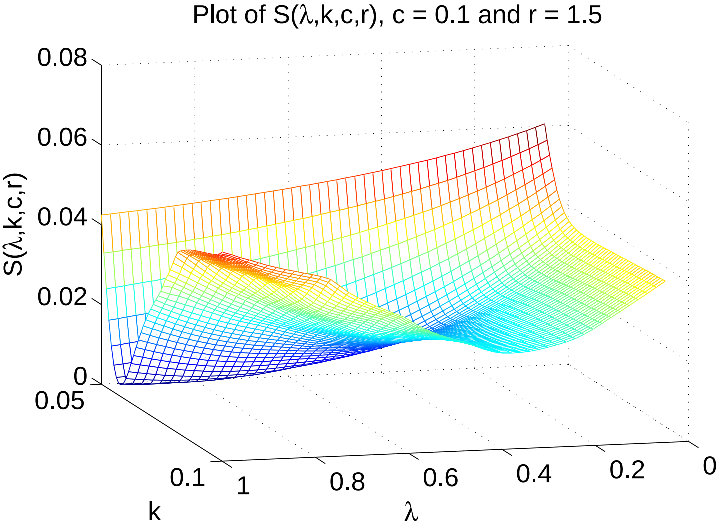

Figure 3.1: Plot of $S(\theta)$ as a function of only λ and K. c and r are kept fixed at their actual values, where $(\lambda, K, r, c) = (0.2, 0.1, 1.5, 0.1)$.

Note that, for instance, two different sets of c and r may result in similar interactions, due to the form of (2.4). In order to control the estimation routine, so that this risk of bias is reduced, the approach of Paper I is to find good starting values, $(\lambda_0, K_0, c_0, r_0)$, (as opposed to arbitrarily chosen ones) and to choose sensible step sizes, δ_λ, δ_K, δ_c, δ_r. Further details about this fine-tuning of the optimization can be found in [8].

3.1.2 Estimation of the death and arrival rates

We here give the form of the ML-estimators used for the estimation of α and μ in Paper I. The estimator $\hat{\mu}$ takes the form of the natural death rate function $\eta(\cdot)$ into consideration, and the α-estimator partially compensates for the unobserved individuals who arrive and die during the same sample interval, (T_k, T_{k-1}). Note that in Paper I and Paper III, the function $\eta(M_i(t)) = 1/(1 + M_i(t))$ is used.

Denote by $L_1, \ldots, L_{n_T}$ the random lifetimes of the n_T individuals who have died from natural causes by time T_n, given some natural death rate function $\mu\eta(M_i(t))$ (recall that we label an individual i as naturally dead once the predicted mark $\tilde{m}_i\left(T_{j+1}; \theta, m_i(T_j)\right) > 0$ while the actual data individual is alive at T_{j+1}, during the calculation of $S(\theta)$). Furthermore, let $t^0_{i(L_1)}, \ldots, t^0_{i(L_{n_T})}$ denote the birth times of the individuals having these lifetimes. Also let $T_{j,i(L_k)}$ be the last sample time at which individual $i(L_k)$ was observed alive and let $\tilde{m}_{i(L_k)}(T_{j,i(L_k)})$ denote the prediction of its mark at $T_{j,i(L_k)}$. Furthermore, under the same natural death rate regime, let $S_1, \ldots, S_{m_T}$ denote the m_T random lifetimes of the individuals who are still alive at time T_n and $m_{i(S_l)}(T_n)$ the size of each such individual at the final sample time. The (approximate) ML-estimator of the death rate, μ, is given by

$$\begin{aligned} \hat{\mu} \quad = \quad & n_T \Bigg/ \Bigg(\sum_{k=1}^{n_T} \eta\left(\tilde{m}_{i(L_k)}\left(T_{j,i(L_k)}\right)\right) \left(T_{j,i(L_k)} - t^0_{i(L_k)}\right) \\ & + \sum_{l=1}^{m_T} \eta(m_{i(S_l)}(T_n)) \left(T_n - t^0_{i(S_l)}\right) \Bigg). \end{aligned} \tag{3.1}$$

Note that the process is observed only at the sampled time points $0 = T_0 < T_1 < \ldots < T_n = T$ so that the actual birth times (and death times) of the individuals remain unknown. Conditioned on the number of individuals arriving during $(T_{j-1}, T_j]$ the arrival times of the individuals will be uniformly distributed on (T_{j-1}, T_j) (see e.g. [23]). Thus, when estimating μ, for each interval $(T_{j-1}, T_j]$ we simulate as many $Uni(T_{j-1}, T_j)$-distributed birth times as there are observed newcomers and these are in turn assigned to all individuals observed for the first time at T_j. The question regarding which arrival time to assign to which individual is solved by giving the first arrival time to the individual who is the largest at time T_j, the second arrival time to the individual which is the second largest at time T_j and so forth. This will have the consequence that the lifetimes will be random. By repeating this procedure a suitable number of times, each time simulating new random birth times, we could generate a set of estimates of μ which are used to estimate a standard error for $\hat{\mu}$. In the case of $\eta(\cdot) \equiv 1$, expression (3.1) reduces to the estimator found in [32].

Let $N_{T_j} = \left| \bigcup_{j=1}^{n} \Omega_{T_j} \right|$, $j = 1, \ldots, n$, denote the number of individuals observed at sample times up to T_j. [32] proposes a simple estimator for the arrival intensity α. However, this estimator underestimates α since it does not take into account the unobserved individuals who arrive and die within the same sample interval (see [32]). In order to (partially) compensate for these unobserved individuals who arrive and die in the same sample interval, (T_k, T_{k-1}), when estimating α we use the following estimator (see [8] or the Appendix for its derivation);

$$\hat{\alpha} = \frac{N_{T_n}}{T_n \nu(W)} + \frac{1}{T_n \nu(W)} \sum_{j=1}^{n} \left\lfloor N_{T_n} \frac{\Delta T_{j-1}}{T_n} \left(1 - \mathrm{e}^{-\hat{\mu}\eta(m_i^0)\Delta T_{j-1}}\right) \right\rfloor, \tag{3.2}$$

where $\lfloor x \rfloor$ denotes the integer part of x, $\Delta T_{j-1} = T_j - T_{j-1}$ and $\hat{\mu}$ is the estimate of μ found previously. Note that the first term in expression (3.2) is the estimator found in [32].

3.2 Estimation in the ID-process

We will here look at the estimation of (α, μ) when the ID-process, $\{N(t)\}_{t \geq 0}$, is considered as its own entity. The results presented in this section can be found in Paper II.

Assume now that we sample $\{N(t)\}_{t \geq 0}$ as $N_1, \ldots, N_n$ at the respective times $0 = T_0 < T_1 < \ldots < T_n$. Since the likelihood function for $\gamma = (\alpha, \mu) \in \Gamma$, $L_n(\gamma)$, is given by the joint density of the distribution of $(N(T_1), \ldots, N(T_n))$, by the Markov property of $N(t)$ it can be factorised into a product of transition probabilities, i.e. $L_n(\gamma) = \mathbb{P}(N(T_1) = N_1) \prod_{k=2}^{n} p_{N_{k-1}N_k}(t; \gamma)$. By assumption we condition on $N(T_0) = 0$, so that the log-likelihood will be given by

$$l_n(\gamma) = \sum_{k=1}^{n} \log p_{N_{k-1}N_k}(\Delta T_{k-1}; \gamma), \tag{3.3}$$

where $\Delta T_{k-1} = T_k - T_{k-1}$. In the case of equidistant sampling, i.e. $\Delta T_{k-1} = t$ for each $k = 1, \dots, n$, the log-likelihood takes the form

$$l_n(\gamma) = \sum_{i,j \in E} N_n(i,j) \log p_{ij}(t;\gamma), \tag{3.4}$$

where $N_n(i,j) = \sum_{k=1}^{n} \mathbf{1}\{(N_{k-1}, N_k) = (i,j)\}$.

Hereby, for each of the sampling schemes, the likelihood estimator of $\gamma = (\alpha, \mu) \in \Gamma$ (obtained by replacing N_k by $N(T_k)$, $k = 0, 1, \dots$, in the expressions (3.3) and (3.4)) will be defined as

$$(\hat{\alpha}_n, \hat{\mu}_n) = \hat{\gamma}_n = \arg\max_{\gamma \in \Gamma} l_n(\gamma). \tag{3.5}$$

3.2.1 The ML-estimators

The ML-estimator for $\gamma = (\alpha, \mu)$ is given by solving the system of equations

$$\begin{cases} \frac{\partial}{\partial \alpha} l_n(\gamma) &= \sum_{i,j \in E} N_n(i,j) \frac{\partial}{\partial \alpha} \log p_{ij}(t;\gamma) = 0 \\ \frac{\partial}{\partial \mu} l_n(\gamma) &= \sum_{i,j \in E} N_n(i,j) \frac{\partial}{\partial \mu} \log p_{ij}(t;\gamma) = 0. \end{cases}$$

As no closed form solution can be found by solving theses likelihood equations, numerical methods have to be employed in order to get ML-estimates. What is possible, however, is to express the estimator of α as a function of both the sample and the parameter μ, hence reducing the maximisation to a one dimensional problem.

Proposition 3.2.1. *The ML-estimator, $\hat{\gamma}_n = (\hat{\alpha}_n, \hat{\mu}_n)$, is found by maximising $l_n(\hat{\alpha}_n(\mu), \mu)$ over $\Gamma_2 \subseteq \mathbb{R}_+$ (the projection of Γ onto the μ-axis), i.e.*

$$\begin{aligned} \hat{\mu}_n &= \arg\max_{\mu \in \Gamma_2} l_n(\hat{\alpha}(\mu), \mu) \\ \hat{\alpha}_n &= \hat{\alpha}_n(\hat{\mu}_n), \end{aligned}$$

where

$$\begin{aligned} \hat{\alpha}_n(\mu) &:= \frac{\mu/(1 - \mathrm{e}^{-\mu t})}{2\left(\frac{1-\mathrm{e}^{-\mu t}}{\mu t} - \mathrm{e}^{-\mu t}\right) - 1} \frac{1}{n} \sum_{i,j \in E} N_n(i,j)(j - i\,\mathrm{e}^{-\mu t}) \\ &= \frac{\mu}{2\left(\frac{1-\mathrm{e}^{-\mu t}}{\mu t} - \mathrm{e}^{-\mu t}\right) - 1} \frac{1}{n} \left(\frac{\mathrm{e}^{-\mu t} N_n - N_0}{1 - \mathrm{e}^{-\mu t}} + \sum_{k=0}^{n} N_k \right). \end{aligned}$$

3.2.2 Asymptotic properties of the ML-estimators

Assume now that we sample $N(t)$ at the times $T_n = nt$, $n \in \mathbb{N}$, $t > 0$ (equidistant sampling). The following two results show that the ML-estimator (3.5) is strongly consistent (Proposition 3.2.2) and asymptotically Gaussian (Proposition 3.2.3). We denote by $\gamma_0 = (\alpha_0, \mu_0) \in \Gamma$ the true parameter pair of the ID-process. These results can be found in Paper II. For further discussions on ML-estimation in Markov processes and asymptotic properties thereof, see e.g. [2, 4, 11, 18, 35].

Proposition 3.2.2. *Let* Γ *be any compact subset of* $\mathbb{R}^2_+$*. Then the maximum likelihood estimator for the ID-process satisfies*

$$(\hat{\alpha}_n, \hat{\mu}_n) \xrightarrow{a.s.} (\alpha_0, \mu_0)$$

as $n \to \infty$*, where* $(\alpha_0, \mu_0) \in \Gamma$ *is the true parameter pair.*

Proposition 3.2.3. *Let* Γ *be any compact subset of* $\mathbb{R}^2_+$*. Furthermore, assume that* $(\log(\alpha_0 + \mu_0) - \log(\alpha_0))/\mu_0 \geq 2t$*. Then, as* $n \to \infty$*,* $\sqrt{n}\,((\hat{\alpha}_n, \hat{\mu}_n) - (\alpha_0, \mu_0))$ *converges in distribution to the two-dimensional zero-mean Gaussian distribution with covariance matrix,* $I(\gamma_0)^{-1}$*, given by*

$$I(\gamma_0)^{-1} = \frac{\mu_0}{t\left(\left(1 + \mathrm{e}^{-\mu_0 t}\right)\rho_0(\Xi - 1) - 1\right)} \tag{3.6}$$
$$\times \begin{pmatrix} \frac{\rho_0\left(2\tau_0 - \mu_0 t\left(1 - \mathrm{e}^{-\mu_0 t}\right)\right) + \frac{\rho_0^2}{\mu_0 t}(\Xi - 1)(\tau_0 - \mu_0 t)^2}{(1 - \mathrm{e}^{-\mu_0 t})^2} & 1 + \frac{\rho_0}{\mu_0 t}(\Xi - 1)(\tau_0 - \mu_0 t) \\ 1 + \frac{\rho_0}{\mu_0 t}(\Xi - 1)(\tau_0 - \mu_0 t) & \frac{1}{\mu_0 t}(\Xi - 1)\left(1 - \mathrm{e}^{-\mu_0 t}\right)^2 \end{pmatrix},$$

where $\Xi = \sum_{i,j \in E} \frac{\left(p_{i(j-1)}(t;\gamma_0)\right)^2}{p_{ij}(t;\gamma_0)} \pi_{\gamma_0}(i)$*,* $\tau_0 = 1 - \mathrm{e}^{-\mu_0 t} - \mu_0 t\, \mathrm{e}^{-\mu_0 t}$ *and* $\rho_0 = \frac{\alpha_0}{\mu_0}(1 - \mathrm{e}^{-\mu_0 t})$*. Here* $\pi_{\gamma_0}(\cdot) = \mathbb{P}(Poi(\alpha_0/\mu_0) \in \cdot)$ *is the invariant distribution of the ID-process.*

3.3 Maximum likelihood inference in the SG-process

Conditionally on $\Phi_M(T_0) = \Phi_M(0)$, assume now that we sample the SG-process $\Phi_M(t)$ as $\phi_1, \ldots, \phi_n$ at the sample times $T_1, \ldots, T_n$ on the compact region W. Here $\phi_k = \left(\mathbf{1}_{\omega_k}(1) m_{1k}, \ldots, \mathbf{1}_{\omega_k}(N) m_{dk}\right)^T$, $\omega_k = \{\text{indices of individuals present at time } T_k\}$, $k = 1, \ldots, n$, and $d = |\bigcup_{k=1}^n \omega_k|$. Now, based on this sampling scheme we want to find the Maximum Likelihood (ML) estimate of the parameter vector $\theta = (\lambda, K, \sigma, \alpha, \mu) \in \Theta$.

The likelihood function of the parameters of the SG-process, $\mathcal{L}_n(\theta)$, is given by the joint density of $(\Phi_M(T_1), \dots, \Phi_M(T_n))$, evaluated at $(\phi_1, \dots, \phi_n)$ and treated as a function of $\theta \in \Theta$. Therefore, depending on whether we choose $M_i(0)$ to be fixed or drawn from the stationary distribution, we end up evaluating either expression (2.10) or expression (2.11) when we evaluate $\mathcal{L}_n(\theta)$.

3.3.1 ML-estimation: $M_i^0 = M_0 \in \mathbb{R}_+$

When we let all $Y_i(0) = M_i^0 = M_0 \in \mathbb{R}_+$ be given by the same fixed value, from expression (2.10) we obtain

$$\mathcal{L}_n(\theta) = C\mathcal{L}_{1,n}(\theta)\mathcal{L}_{2,n}(\theta)\mathcal{L}_{3,n}(\theta) \propto \mathcal{L}_{1,n}(\theta)\mathcal{L}_{2,n}(\theta)\mathcal{L}_{3,n}(\theta),$$

where, for $k_i = \min\{k : i \in \omega_k\}$,

$$\begin{aligned}
\mathcal{L}_{1,n}(\theta) &= \prod_{k=1}^{n} \prod_{i \in \omega_{k-1} \cap \omega_k} p_{Y_1}(\Delta T_k, m_{ik} | m_{i(k-1)}; \lambda, K, \sigma) \\
\mathcal{L}_{2,n}(\theta) &= \prod_{i \in \bigcup_{k=1}^{n} \omega_k} \frac{1}{\Delta T_{k_i}} \int_0^{\Delta T_{k_i}} p_{Y_1}(t, m_{i(k_i-1)} | M_0; \lambda, K, \sigma) dt \\
\mathcal{L}_{3,n}(\theta) &= \prod_{k=1}^{n} p_N \left(\Delta T_k, |\omega_k| \Big| |\omega_{k-1}|; \alpha\nu(W), \mu \right).
\end{aligned}$$

The (rescaled) log-likelihood is given by

$$\begin{aligned}
l_n(\theta) &= \log\left(C^{-1}\mathcal{L}_n(\theta)\right) = \log \mathcal{L}_{1,n}(\theta) + \log \mathcal{L}_{2,n}(\theta) + \log \mathcal{L}_{3,n}(\theta) \\
&=: l_{1,n}(\theta) + l_{2,n}(\theta) + l_{3,n}(\theta),
\end{aligned}$$

and the ML-estimator of $\theta \in \Theta$, based on $(\Phi_M(T_1), \dots, \Phi_M(T_n))$, will be given by

$$\begin{aligned}
\widetilde{\theta}_n &:= \widetilde{\theta}_n \left(\Phi_M(T_1), \dots, \Phi_M(T_n)\right) \qquad (3.7) \\
&= \arg\max_{\theta \in \Theta} l_n(\theta; \Phi_M(T_1), \dots, \Phi_M(T_n)) \\
&= \arg\max_{\theta \in \Theta} (l_{1,n}(\theta) + l_{2,n}(\theta) + l_{3,n}(\theta)) \\
&= \widetilde{\theta}_{1,n} + \widetilde{\theta}_{2,n} \\
&= \arg\max_{\theta \in \Theta_\lambda \times \Theta_K \times \Theta_\sigma \times \{0\}^2} \{l_{1,n}(\theta) + l_{2,n}(\theta)\} + \arg\max_{\theta \in \{0\}^3 \times \Theta_\alpha \times \Theta_\mu} l_{3,n}(\theta),
\end{aligned}$$

whereby we may estimate the parameters of the ID-process and the parameters related to the mark growth separately. Moreover, since there is no closed form expression available for the ML-estimator $(\widetilde{\alpha}_n, \widetilde{\mu}_n)$ of the ID-process (see [9]), there is also no closed form for $\widetilde{\theta}_n$ in (3.7). Hence, in modelling situations one has to rely on numerical methods to find $\widetilde{\theta}_n$.

3.3.2 ML-estimation: $M_i^0 \sim \pi$

Under the assumption that we start the diffusions in their stationary distributions, $M_i^0 \sim \pi$, from expression (2.11) we obtain the likelihood function

$$\mathcal{L}_n(\theta) = C\mathcal{L}_{1,n}(\theta)\mathcal{L}_{2,n}(\theta) \propto \mathcal{L}_{1,n}(\theta)\mathcal{L}_{2,n}(\theta)$$

and the (rescaled) log-likelihood

$$\begin{aligned} l_n(\theta) &= \log\left(C^{-1}\mathcal{L}_n(\theta)\right) = \log \mathcal{L}_{1,n}(\theta) + \log \mathcal{L}_{2,n}(\theta) \\ &=: l_{1,n}(\theta) + l_{2,n}(\theta), \end{aligned}$$

where

$$\begin{aligned} l_{1,n}(\theta) &= \log\left(\prod_{k=1}^{n}\prod_{i\in\omega_k} \pi(m_{ik};\lambda,K,\sigma)\right) = \sum_{k=1}^{n}\sum_{i\in\omega_k} \log \pi(m_{ik};\lambda,K,\sigma) \\ l_{2,n}(\theta) &= \log\left(\prod_{k=1}^{n} p_N\left(\Delta T_k, |\omega_k| \Big| |\omega_{k-1}|; \alpha\nu(W), \mu\right)\right) \\ &= \sum_{k=1}^{n} \log p_N\left(\Delta T_k, |\omega_k| \Big| |\omega_{k-1}|; \alpha\nu(W), \mu\right). \end{aligned}$$

Here, just as in the fixed initial value case of Section 3.3.1, we deal with the separate estimators

$$\hat{\theta}_n = \hat{\theta}_{1,n} + \hat{\theta}_{2,n} = \underset{\theta\in\Theta_\lambda\times\Theta_K\times\Theta_\sigma\times\{0\}^2}{\arg\max} l_{1,n}(\theta) + \underset{\theta\in\{0\}^3\times\Theta_\alpha\times\Theta_\mu}{\arg\max} l_{2,n}(\theta) \quad (3.8)$$

and, similarly, there is no closed form expression available for $\hat{\theta}_n$.

3.3.3 Asymptotic inference under stationarity

When dealing with asymptotic spatial statistics, there are different types of asymptotics which may be considered. In the case of the SG-process, within the framework of so called increasing domain asymptotics (see e.g. [39]), there essentially are two different ways to increase the total number of individuals considered, and consequently also the number of transitions taking place between pairs of consecutive sample times T_{k-1} and T_k; Either increase the number of sample points or increase the size of W. We here consider the approach where we increase the number of sample times, i.e. we apply the equidistant sampling scheme $T_k = k\Delta$, $k = 1, \dots, n$, $\Delta > 0$, where $T = T_n = n\Delta$. We will denote by $\theta_0 = (\lambda_0, K_0, \sigma_0, \alpha_0, \mu_0) \in \Theta$ the true parameter vector value which generates Φ_M, and we assume that Θ is a subset of $\mathbb{R}^5_+$ such that

$$\Theta \cap \{(\lambda, K, \sigma, \alpha, \mu) \in \mathbb{R}^5_+ : 2\lambda < \sigma^2\} = \emptyset. \tag{3.9}$$

Recall that this is required to keep the $Y_i(t)$'s positive.

Theorem 3.3.1 (Consistency). *Let Θ be a compact subset of $\mathbb{R}^5_+$ such that (3.9) holds. Then, for $\theta_0 \in \Theta$, the estimator $\hat{\theta}_n$ in expression (3.8) is strongly consistent, i.e. as $n \to \infty$,*

$$\hat{\theta}_n \xrightarrow{a.s.} \theta_0.$$

Now, by putting some additional restrictions on the parameters we may also prove the following theorem.

Theorem 3.3.2 (Asymptotic normality). *Let θ_0 be in the interior of Θ, where Θ is a compact subset of $\mathbb{R}^5_+$ such that (3.9) holds. Require further that θ_0 and $\Delta > 0$ are such that $(\log(\alpha_0 + \mu_0) - \log(\alpha_0))/\mu_0 \geq 2\Delta$.*

Assume that λ_0 is known, so that $\hat{\theta}_n = (\hat{K}_n, \hat{\sigma}_n, \hat{\alpha}_n, \hat{\mu}_n)$ is the ML-estimator of $\theta_0 = (K_0, \sigma_0, \alpha_0, \mu_0)$. Then, as $n \to \infty$, we obtain

$$\sqrt{n}\big(\hat{\theta}_n - \theta_0\big) \xrightarrow{d} \mathbf{Y} \sim N\left(\mathbf{0}_{4\times1}, \begin{bmatrix} \frac{\mu_0}{\alpha_0}\frac{K_0^2\sigma_0^2}{2\lambda_0} & 0 & \mathbf{0}_{1\times2} \\ 0 & \frac{\mu_0}{\alpha_0}\frac{\sigma_0^4}{8\lambda_0 C(\theta_0)} & \mathbf{0}_{1\times2} \\ \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} & I_N(\theta_0)^{-1} \end{bmatrix}\right),$$

where $C(\theta) = \frac{2\lambda}{\sigma^2}\psi'\left(\frac{2\lambda}{\sigma^2}\right) - 1 > 0$, $\psi(x) = \Gamma'(x)/\Gamma(x)$, $\mathbf{0}_{i\times j}$ denotes the $i \times j$ zero matrix and the 2×2 matrix $I_N(\theta_0)^{-1}$, which can be found in expression (3.6), is the covariance matrix related to the ID-process.

Similarly, when σ_0 is known, we estimate $\theta_0 = (\lambda_0, K_0, \alpha_0, \mu_0)$ by $\hat{\theta}_n = (\hat{\lambda}_n, \hat{K}_n, \hat{\alpha}_n, \hat{\mu}_n)$ and, as $n \to \infty$, we obtain

$$\sqrt{n}\big(\hat{\theta}_n - \theta_0\big) \xrightarrow{d} \mathbf{Y} \sim N\left(\mathbf{0}_{4\times1}, \begin{bmatrix} \frac{\mu_0}{\alpha_0}\frac{\lambda_0\sigma_0^2}{2C(\theta_0)} & 0 & \mathbf{0}_{1\times2} \\ 0 & \frac{\mu_0}{\alpha_0}\frac{K_0^2\sigma_0^2}{2\lambda_0} & \mathbf{0}_{1\times2} \\ \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} & I_N(\theta_0)^{-1} \end{bmatrix}\right).$$

3.4 Spatio-temporal edge correction

When sampling real data, $\{\mathbb{X}(T_k)\}_{k=1}^{n}$, one usually considers all individuals within some region A (in Figure 1.1 circular) which is part of some larger region W. The individuals in A interact with each other but simultaneously also with the individuals present outside A, i.e. the individuals in $B = W \setminus A$. So, if one were to estimate some statistics and/or model parameters in a situation where the interaction among (neighbouring) individuals plays a role, by only taking into consideration the individuals in A the estimators may generate biased estimates since the interaction between the individuals in A and those in B would be neglected. The effects of the absence of the information regarding this interaction are commonly referred to as *edge effects*. The risk that the edge effects generate biases rapidly increases when one deals with small quantities of data in A, as is the case with our tree data set introduced in Figure 1.1. Hence, some type of correction method is needed (see e.g. [12, 20, 38]).

We here give the idea behind the edge correction methods proposed in Paper I. One starts by finding initial (possibly biased) estimates of the model parameters, $\hat{\theta}_*$, based on the original data set (region A). Then, under the regime of $\hat{\theta}_*$, we wish to find the expected model behaviour when restricted to region B (possibly conditioned on the actual data in A), $\mathbb{E}_{\hat{\theta}_*}[\Phi_M[0,T]|_B]$. By doing so we wish to establish the expected interaction between the individuals in B and the individuals in region A. With $\mathbb{E}_{\hat{\theta}_*}[\Phi_M[0,T]|_B]$ at hand we now re-estimate the model parameters from the actual data (region A), however, this time allowing for $\mathbb{E}_{\hat{\theta}_*}[\Phi_M[0,T]|_B]$ to interact with the actual data during the estimation. Once these new estimates have been obtained, we let them replace $\hat{\theta}_*$ and repeat the above procedure again. By continuing in this fashion we have an iterative procedure which we stop once it has fulfilled a given predefined convergence criterion.

The three edge correction methods presented in Paper I. We refer to as them as *The simple correction method*, *The rotated surrounding correction method* and *The influenced growth correction method*, and they are all explained for the GI-process but they may be applied to other spatial and spatio-temporal (marked) point processes as well. In the algorithms presented in Paper I the large rectangular window W will be wrapped onto a torus when we generate the individuals in the outer region, B, (see e.g. [12, 25, 29, 38]).

Chapter 4

Future work and extensions

There are some remarks which may be addressed about the general development of the GI-process. Note that the GI-process here is presented for a single species. However, it can easily be extended to include the scenario where interaction takes place also between different species, living and interacting within the same study region. This extension is made by letting each species be governed by both its unique open-growth function and interaction function, and the latter can be different within and between species. Hereby the amount an individual is affected by its neighbours not only depends on its distance to the neighbours and the neighbours' sizes, but also on the species of the neighbours.

An improvement of Paper II that possibly can be made is to improve the invertibility condition given in Proposition 3.2.3 in Chapter 3 so that asymptotic normality holds for all $(\alpha_0, \mu_0) \in \Theta$. Furthermore, in order to become more realistic in applications, $N(t)$ could be extended by letting the arrival intensity, α, and the death rate, μ, be non-constant functions of time, or in themselves Markov chains (in the latter case $N(t)$ thus becomes a hidden Markov model). Results similar to the ones found in Paper II could be established and the type of modelling done in Paper I could be developed.

Regarding the development of Paper IV, we could employ some other positive diffusion for the growth of the marks. It should be noted that the linear growth function, which is the drift function in the CIR-process, is a special case of the Richards growth function (see e.g. [28, 31]). Hence, a further possibility would be to use the Richards growth function, or one of its other special cases, as drift in the mark-SDEs $dY_i(t)$.

A further modification which may be made is to change the diffusion term $\sigma(Y_i(t);\theta) = \sigma\sqrt{Y_i(t)}$ into any other diffusion term which keeps $Y_i(t)$ positive, e.g. $\sigma(Y_i(t);\theta) = \sigma Y_i(t)^\gamma$, $\gamma > 0$, which is the diffusion coefficient found in the CKLS-model (see e.g. [5]). Note that when applying these changes, we would typically not have known closed form expressions for the transition densities, $p_{Y_i}(t, y_1|y_0;\theta)$. The transition densities are know only for a few special cases, including the CIR-process. Therefore, we have to use different approximated/pseudo likelihood methods for the estimation of the parameters (see [19] for a good general overview).

Our final goal is to ML-estimate all parameters of the full SGI-process, i.e. to include also the spatial interaction function $h(\cdot)$ in the SDEs. Note that on a compact space-time domain this amounts to considering a multivariate diffusion. Here the lack of closed form expression for the transition densities remains and, just as for the previous adjustments suggested, the estimation requires that we employ approximated/pseudo likelihood methods. For instance, [1] suggests an approach where the transition densities of multivariate diffusions may be approximated by series expansions based on hermite polynomials. Note further that within this setting, in order to reduce edge effects (absence of individuals outside the boundary of W), it would be sensible to choose W to be a torus. Furthermore, instead of using the edge correction methods in Paper I, an altered version could be considered which is more in the lines of an EM-algorithm. This would also allow us to study convergence properties of the edge corrected ML-estimators of the SGI-process from a theoretical perspective.

Thus far we have introduced only natural deaths in the SG(I)-process. It should be possible also to introduce competitive deaths as well, however, this would entail a slightly different formulation of the diffusions $M_i(t)$, $i = 1, \dots, N$. By defining the death-time of individual i to be (the stopping-time) $\zeta_i = \inf\{t > B_i : M_i(t) = 0\} \wedge D_i$, it follows that if M_i reaches the absorbing state $M_i(t) = 0$ for some $t \in (B_i, D_i)$, where $D_i = B_i + L_i$, it stays 0 and we say that it has suffered a competitive death. Furthermore, if it does not die from competition during (B_i, D_i) it will still die at time D_i, i.e. at its natural death time. As soon as $t > \zeta_i$ the interaction between $M_i(t)$ and the other marks will terminate, hence we remove individual i from consideration.

Paper I motivated the study conducted in Paper III. Although the modifications of the GI-process made in Paper III improved the fit to the Scots pine data type considered, some improvements can still be made. It was seen that spatial characteristics and basal area of simulated predictions were similar to the ones of the data. However, it was also seen that the empirical diameter distributions of the data trees and the simulated predicted trees differed, and we further also saw that the predicted and observed number of alive individuals differed substantially. The former seems to be the result of the chosen growth/interaction functions, whereas the latter seems to be the result of having only very few (3) sample time points to estimate the ID-process. Hence, although the GI-process described the data quite well in most aspects, there are still improvements to be made. As a first step one should try to fit the model to data sets with more observed time points. A further step which could be made is to evaluate other kinds of growth and interaction functions to see if the fit is improved. It is also our belief that once we have developed the ML-estimation scheme for the SGI-process (with interaction), a better approach to fitting the process to data will be available. We further hope that this could improve the modelling of the considered (pine) stands.

Chapter 5

Summary of Papers

Paper I: Some edge correction methods for marked spatio-temporal point process models

In this paper we consider the GI-process/RS-model where the death rate of the underlying immigration-death process depends on each individual's mark size, as opposed to the approach used in [32] where the death rate was constant.

We then discuss the estimation of the parameters when the process is sampled discretely in time. To improve the estimation of the growth and interaction parameters, three edge correction methods for (marked) spatio-temporal point processes are proposed. They are all based on the idea of placing an approximate expected behaviour of the process at hand outside the study region. We then let these simulated realizations outside the study region interact with the data during the estimation. We estimate this expected behaviour by simulating realizations of the process, under a parameter choice based on some non-edge corrected initial estimates, and for each such realization we generate new estimates which we average over to get our final estimates. By rerunning the whole procedure and using our edge corrected estimates to generate the surrounding realizations, we have created an iterative procedure which we stop once some given stopping criterion is fulfilled. Furthermore, we discuss three different approaches to run this type of edge correction and we present each of them in the context of the RS-model. When we numerically evaluate our edge corrected estimation procedures for the RS-model we see that we manage to reduce the bias substantially, compared to when no edge correction is applied.

We finally evaluate the performance of the edge corrected estimation of the model by fitting it to a Scots pine data set.

Paper II: The Discretely Observed Immigration-Death Process and its Maximum Likelihood Estimation

In this paper we consider the immigration-death process, $N(t)$, and specifically we treat the ML-estimation of the parameter pair governing it, $\theta = (\alpha, \mu) \in \Theta \subseteq \mathbb{R}_+^2$, when Θ is compact and $N(t)$ is sampled discretely in time; $0 = T_0 < T_1 < \ldots < T_n$, $N(T_0) = 0$.

In order to find the likelihood structure of this continuous time Markov chain we derive its transition probabilities, and further, we manage to reduce the likelihood maximisation from a two dimensional problem to a one dimensional problem, where we maximise the likelihood, $L(\alpha, \mu) = L(\hat{\alpha}_n(\mu), \mu)$, over the projection of Θ onto the μ-axis.

Furthermore, by considering $N(t)$ as a Markov jump process we have shown that, under an equidistant sampling scheme, $T_k = kt$, $t > 0$, $k = 1, \ldots, n$, the sequence of ML-estimators, $\hat{\theta}_n(N(T_1), \ldots N(T_n))$, is consistent and asymptotically Gaussian. The asymptotic normality requires the Fisher information matrix invertability condition $(\log(\alpha_0 + \mu_0) - \log(\alpha_0))/\mu_0 \geq 2t$, where (α_0, μ_0) is the underlying parameter pair. These results are further verified through simulations. In the simulations we see that the estimates approach the actual parameters and also that the empirical distribution of the estimates show strong indications of Gaussianity, even when the invertability condition is not fulfilled.

Paper III: Spatio-Temporal Modelling of Swedish Scots Pine Stands

Motivated by the study in Paper I, we here alter the GI-process to obtain a better fit to Scots pine stands. Specifically, we propose a new (hard core) arrival strategy and a different open-growth function. The space-time data set considered here consists of ten Scots pine plots and each plot is measured at three occasion. Since the trees in the plots are young, it is hard to estimate the carrying capacity correctly, so we estimate the carrying capacity (upper bound for the size of a tree) from a separate set of data which contains older pines. A linear relationship is found between the carrying capacity and the so called site productivity index, which is a fertility measure. The arrival and death intensities are estimated by means of the estimators in Paper II and the growth and interaction parameters are estimated using the edge corrected least squares approach of Paper I. The fit of the model is tested by comparing estimated summary statistics of the data and the simulated fitted model. Two of the summary statistics used deal with the spatial structures of the plots and the remaining summary statistics, which are common in forestry, mainly focus on the sizes of the trees. In comparison to Paper I, we here improve the fit of the model.

Paper IV: Likelihood Inference for a Stochastic Version of the Spatio-Temporal Growth-Interaction Process

In this paper we define a version of the GI-process (see e.g. [32]) – the SGI-process – which is given by the GI-process with mark sizes driven by stochastic differential equations (SDEs) instead of ordinary differential equations. Furthermore, we concentrate on a special case, the SG-process, where there is no interaction between the marks. The specific SDEs used here are iid CIR-processes. Since its building blocks, the immigration-death process and the CIR-process, are Markov processes also the SG-process is a Markov process. By exploiting this property we have derived its transition densities, which take closed forms since its building blocks have closed form transition densities. The transition densities are then in turn used to write down a full likelihood function based on the process sampled according to a discrete sample scheme. Furthermore, given that we start the CIR-process in its stationary distribution, it becomes strictly stationary. Under this condition and some additional conditions on the parameters, by using an equidistant sampling scheme, we prove consistency and asymptotic normality of the ML-estimators. We finally evaluate both types of estimators numerically, i.e. we compare those derived under the stationarity assumption to those derived when stationarity is not assumed. As a study of robustness against the stationarity assumption, we compare the biases obtained for both types of estimators when applied to discretely sampled realizations of the process, which have been generated in the non-stationary setting. We conclude that the non-stationary estimator generally performs quite well whereas the stationary estimator works well only for certain parts of the parameter space.

Appendix A

Appendix

In this Appendix we recall the derivation of the estimator for the arrival intensity $\hat{\alpha}$ which was derived in [8]. It is the estimator used in Paper I but its derivation was excluded in Paper I.

When constructing our α-estimator we wish to somehow compensate for the unobserved individuals who arrive and die during the same interval (T_{j-1}, T_j), $j = 1, \ldots, n$.

For each $j = 1, \ldots, n$, let N_{T_j} be the number of individuals observed at sample times up until T_j, i.e. $N_{T_j} = \left| \bigcup_{i=1}^{j} \Omega_{T_i} \right|$, where Ω_t consists of the indices of the individuals alive at t and $|A|$ denotes the cardinality of a set A. Recall that $\eta(\cdot)$ is the function which controls the death rate (see Paper I). Further, let $B(t) \geq 0$ denote the number of arrivals to W by time t. Instead of considering $\Delta B(T_{j-1}) = \Delta N_{T_{j-1}}$, where $\Delta B(T_{j-1}) = B(T_j) - B(T_{j-1})$ and $\Delta N_{T_{j-1}} = N_{T_j} - N_{T_{j-1}}$, and let our likelihood be based on these independent $Poi(T_j - T_{j-1})$-distributed increments, as was done in [32], we here consider

$$\begin{aligned} \Delta B(T_{j-1}) &= \Delta N_{T_{j-1}} \\ &+ \underbrace{\mathbb{E}\left[\sum_{k=1}^{\Delta B(T_{j-1})} \mathbf{1}\left\{ \text{Individual } k \text{ dies in } (T_{j-1}, T_j) \right\} \right]}_{I}, \end{aligned} \tag{A.1}$$

where $\mathbf{1}\{\cdot\}$ is an indicator function. In other words, we add to the observed increments the expected number of individuals arriving and dying during (T_{j-1}, T_j).

Let $L_k^{\Delta T_{j-1}}$ denote the lifetime of individual $k \in \{1, \ldots, \Delta B(T_{j-1})\}$ in (A.1) and denote by $t_i^0 \sim Uni(T_{j-1}, T_j)$ its arrival-time (the jumps of a Poisson process occurring in given time interval are uniformly distributed on that interval [23]). Recall that m_k^0 is its (deterministic) initial size. By the memoryless-property of the exponential distribution and by Fubini's theorem the expectation in expression (A.1) can be written as

$$\begin{aligned} I &= \mathbb{E}\left[\sum_{k=1}^{\Delta B(T_{j-1})} \mathbf{1}\left\{T_{j-1} < t_k^0 + L_k^{\Delta T_{j-1}} < T_j\right\}\right] \qquad \text{(A.2)} \\ &= \mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{\Delta B(T_{j-1})} \mathbf{1}\left\{T_{j-1} < t_k^0 + L_k^{\Delta T_{j-1}} < T_j\right\} \middle| \Delta B(T_{j-1})\right]\right] \\ &= \mathbb{E}\left[\sum_{k=1}^{\Delta B(T_{j-1})} \frac{1}{\Delta T_{j-1}} \int_{T_{j-1}}^{T_j} \mathbb{E}\left[\mathbf{1}\left\{T_{j-1} < x_k + L_k^{\Delta T_{j-1}} < T_j\right\}\right] dx_k\right] \\ &= \mathbb{E}\left[\sum_{k=1}^{\Delta B(T_{j-1})} \frac{1}{\Delta T_{j-1}} \int_{T_{j-1}}^{T_j} P\left(L_k^{\Delta T_{j-1}} < T_j - T_{j-1}\right) dx_k\right] \\ &\approx \mathbb{E}\left[\sum_{k=1}^{\Delta B(T_{j-1})} \left(1 - \mathrm{e}^{-\mu\eta(m_i^0)\Delta T_{j-1}}\right)\right] \\ &= \alpha\nu(W)\Delta T_{j-1}\left(1 - \mathrm{e}^{-\mu\eta(m_i^0)\Delta T_{j-1}}\right). \end{aligned}$$

Since the actual μ is unknown we will replace it by its estimate, $\hat{\mu}$, found in expression 3.1. Furthermore, this expression also contains α, the parameter we want to estimate. We deal with this by replacing α by an initial estimate, namely, $\hat{\alpha}_0 = N_{T_n}/(T_n\nu(W))$, which is given in [32].

In order for expression (A.1) to be treated as an actual Poisson process increment it needs to be integer valued, hence

$$\Delta B(T_{j-1}) = \Delta N_{T_{j-1}} + \left\lfloor N_{T_n} \frac{\Delta T_{j-1}}{T_n} \left(1 - \mathrm{e}^{-\hat{\mu}\eta(m_i^0)\Delta T_{j-1}}\right)\right\rfloor, \qquad \text{(A.3)}$$

where $\lfloor x \rfloor$ denotes the integer part of x. For convenience we will denote the right hand side of (A.3) by $J(\Delta T_{j-1}, \Delta N_{T_{j-1}}, \hat{\mu}, N_{T_n})$. We end up with the likelihood function

$$\begin{aligned} L(\alpha) &= \prod_{j=1}^{n} \mathbb{P}\left(\Delta B(T_{j-1}) = J(\Delta T_{j-1}, \Delta N_{T_{j-1}}, \hat{\mu}, N_{T_n})\right) \qquad \text{(A.4)} \\ &= \prod_{j=1}^{n} \frac{\mathrm{e}^{-\alpha\nu(W)\Delta T_{j-1}}(\alpha\nu(W)\Delta T_{j-1})^{J(\Delta T_{j-1}, \Delta N_{T_{j-1}}, \hat{\mu}, N_{T_n})}}{J(\Delta T_{j-1}, \Delta N_{T_{j-1}}, \hat{\mu}, N_{T_n})!} \end{aligned}$$

and by evaluating $\mathrm{d}l(\alpha)/\mathrm{d}\alpha = 0$ we finally arrive at the estimator

$$\hat{\alpha} = \underbrace{\frac{N_{T_n}}{T_n\nu(W)}}_{=\hat{\alpha}_0} + \frac{1}{T_n\nu(W)} \sum_{j=1}^{n} \left\lfloor N_{T_n} \frac{\Delta T_{j-1}}{T_n} \left(1 - \mathrm{e}^{-\hat{\mu}\eta(m_i^0)\Delta T_{j-1}}\right) \right\rfloor . \qquad \text{(A.5)}$$

Since $\hat{\mu} > 0$, $\Delta T_{j-1} > 0$ and $\eta(x) > 0$, for all $x > 0$, and since $f(x) = 1 - \mathrm{e}^{-x}$ is strictly increasing and bounded below by 0 and above by 1, for $x > 0$, it is clear that $\hat{\alpha}$ is increasing with $\hat{\mu}$ and

$$\hat{\alpha}_0 = \lim_{\hat{\mu}\to 0} \hat{\alpha}|_{\hat{\mu}} < \hat{\alpha} < \lim_{\hat{\mu}\to\infty} \hat{\alpha}|_{\hat{\mu}} = \hat{\alpha}_0 + \frac{1}{T_n\nu(W)} \sum_{j=1}^{n} \left\lfloor N_{T_n} \frac{\Delta T_{j-1}}{T_n} \right\rfloor .$$

For a random variable $Z = X + Y$ it holds that $\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$. Let now $X = \hat{\alpha}_0$ and let Y be the sum in expression (A.5). Since X and Y are positively correlated (both contain N_{T_n}) and since $\mathrm{Var}(Y) \geq 0$ it is clear that $\mathrm{Var}(\hat{\alpha}) > \mathrm{Var}(\hat{\alpha}_0)$ for all $\hat{\mu} > 0$. This implies that the trade off for using $\hat{\alpha}$ instead of $\hat{\alpha}_0$ is a higher standard error. Furthermore, as $\hat{\alpha}$ is increasing with $\hat{\mu}$, so is $\mathrm{Var}(\hat{\alpha})$.

Table A.1 gives us the estimated means and standard errors (s.e.) of $\hat{\alpha}$ (and $\hat{\alpha}_0$) for a few values of $\hat{\mu}$, based on 30 realizations simulated on $W = \{\mathbf{y}\mathbb{R}^2 : \|\mathbf{y}\| \leq 10\}$ from the parameters $\alpha = 0.007$, $\mu = 0.02$, $\lambda = 0.08$, $K = 0.1$, $c = 2$, and $r = 2$ (model set-up as in Paper I).

In estimations of μ based on simulated realizations it has been observed that there seems to be no indication of over-estimation of μ. As one can see in Table A.1, on average $\hat{\alpha}_0$ under-estimates α more than $\hat{\alpha}$ does when $\hat{\mu} \leq \mu$, in the above scenario indicating that $\hat{\alpha}$ is preferred to $\hat{\alpha}_0$. Note also the smaller standard error of $\hat{\alpha}_0$.

$\alpha = 0.007$	Est. mean	Est. s.e.	Est. bias (%)
$\hat{\alpha}_0 = \lim_{\hat{\mu}\to 0} \hat{\alpha}\|_{\hat{\mu}}$	0.0060	0.0008	-0.00099 (-14%)
$\hat{\alpha}$ ($\hat{\mu} = 0.0002$)	0.0060	0.0008	-0.00099 (-14%)
$\hat{\alpha}$ ($\hat{\mu} = 0.002$)	0.0061	0.0009	-0.00089 (-13%)
$\hat{\alpha}$ ($\hat{\mu} = 0.02$)	0.0074	0.0011	0.00044 (6%)
$\hat{\alpha}$ ($\hat{\mu} = 0.1$)	0.0102	0.0014	0.00320 (46%)
$\hat{\alpha}$ ($\hat{\mu} = 0.2$)	0.0111	0.0016	0.00411 (59%)
$\hat{\alpha}$ ($\hat{\mu} = 5$)	0.0119	0.0017	0.00489 (70%)
$\lim_{\hat{\mu}\to\infty} \hat{\alpha}\|_{\hat{\mu}}$	0.0119	0.0017	0.00489 (70%)

Table A.1: Estimated means, standard errors (s.e.), and biases of $\hat{\alpha}$ (and $\hat{\alpha}_0$), based on 30 realizations simulated on $W = \{\mathbf{y}\mathbb{R}^2 : \|\mathbf{y}\| \leq 10\}$ from the parameters $\alpha = 0.007$, $\mu = 0.02$, $\lambda = 0.08$, $K = 0.1$, $c = 2$, and $r = 2$ (model set-up as in Paper I).

Bibliography

[1] Aït-Sahalia, Y. Closed-Form Likelihood Expansions for Multivariate Diffusions. *The Annals of Statistics*, 36:906–937, 2008.

[2] Basawa, I.V., Prakasa Rao, B.L.S. *Statistical Inference for Stochastic Processes.* Academic press, 1980.

[3] Berger, U., Hildenbrandt, H. A new approach to spatially explicit modelling of forest dynamics: spacing, ageing and neighbourhood competition of mangrove trees. *Ecological Modelling*, 132:287–302, 2000.

[4] P. Billingsley. *Statistical Inference for Markov Processes.* The University of Chicago Press, 1961.

[5] Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B. An empirical investigation of alternative models of the short-term interest rate. *Journal of Finance*, 47:1209–1227, 1992.

[6] Cox, J.C., Ingersoll, J.E., Ross, S.A. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407, 1985.

[7] Cressie, N.A.C. *Statistics for Spatial Data.* Wiley Series in Probability and Statistics, revised edition, 1993.

[8] O. Cronie. Some edge correction methods for marked spatio-temporal point process models. Technical report, 2009. Preprint available at `www.math.chalmers.se/Math/Research/Preprints/2010/9.pdf`.

[9] Cronie, O., Yu, J. Maximum likelihood estimation in a discretely observed immigration-death process. Technical Report 2010:1, Swedish University of Agricultural Sciences - Centre of Biostochastics, February 2010.

[10] Daley, D.J., Vere-Jones, D. *An Introduction to the Theory of Point Processes.* Springer Series in Statistics, second edition, 2008.

[11] Dehay, D., Yao, J.-F. On likelihood estimation for discretely observed Markov jump processes. *Australian & New Zealand Journal of Statistics*, 49:93–107, 2007.

[12] P. Diggle. *Statistical Analysis of Spatial Point Patterns.* Oxford university press, second edition, 2001.

[13] Feller, W. Two Singular Diffusion Problems. *Annals of Mathematics*, 54:173–182, 1951.

[14] Finkenstadt, B., Held, L., Isham, V. *Statistical Methods for Spatio-Temporal Systems.* Chapman & Hall/CRC, 2006.

[15] Goreaud, F., Loreau, M., Millier, C. Spatial structure and the survival of an inferior competitor: a theoretical model of neighborhood competition in plants. *Ecological Modelling*, 158:1–19, 2002.

[16] Grabarnik, P., Särkkä, A. Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions. *Ecological Modelling*, 220:1232–1240, 2009.

[17] Grimmett, G., Stirzaker, D. *Probability and Random Processes.* Oxford university press, third edition, 2001.

[18] P. Guttorp. *Statistical Inference for Branching Processes.* Wiley Series in Probability and Mathematical Statistics, 1991.

[19] Iacus, S.M. *Simulation and Inference for Stochastic Differential Equations: with r examples.* Springer, 2008.

[20] Illian J., Penttinen A., Stoyan H., Stoyan D. *Statistical Analysis and Modelling of Spatial Point Patterns.* Wiley-Interscience, 2008.

[21] Möller, J., Waagepetersen, R.P. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman & Hall/CRC Press, 2004.

[22] Nord-Larsen, T. Modeling individual-tree growth from data with highly irregular measurement intervals. *Forest Science*, 52:198–208, 2006.

[23] J.R. Norris. *Markov Chains.* Cambridge series in statistical and probabilistic mathematics, 1997.

[24] S.L. Rathbun. Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes. *Journal of Statistical Planning and Inference*, 51:55–74, 1996.

[25] E. Renshaw. The linear spatial-temporal interaction process and its relation to $1/\omega$-noise. *J. Roy. Statist. Soc. B*, 56:75–91, 1994.

[26] E. Renshaw. Applying the saddlepoint approximation to bivariate stochastic processes. *Mathematical Biosciences*, 168:57–75, 2000.

[27] Renshaw, E., Comas, C. Space-time generation of high intensity patterns using growth-interaction processes. *Statistics and Computing*, 19:423–437, 2009.

[28] Renshaw, E., Comas, C., Mateu, J. Analysis of forest thinning strategies through the development of space-time growth-interaction simulation models. *Stochastic Environmental Research and Risk Assessment*, 23:275–288, 2009.

[29] Renshaw, E., Särkkä, A. Gibbs point processes for studying the development of spatial-temporal stochastic processes. *Computational Statistics & Data Analysis*, 36:85–105, 2001.

[30] Schneider, R., Weil, W. *Stochastic and Integral Geometry.* Springer, 2008.

[31] Seber, G., Wild, C. *Nonlinear Regression.* Wiley series in probability and mathematical statistics, 1989.

[32] Särkkä, A., Renshaw, E. The analysis of marked point patterns evolving through space and time. *Computational Statistics & Data Analysis*, 51:1698–1718, 2006.

[33] Stoyan, D., Kendall, W., Mecke, J. *Stochastic Geometry and its Applications.* John Wiley & sons, second edition, 1995.

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

[35] van der Vaart, A.W. *Asymptotic Statistics.* Cambridge series in statistical and probabilistic mathematics, 1998.

[36] Verre-Jones, D. Some models and procedures for space-time point processes. *Environmental and Ecological Statistics*, 16:173–195, 2009.

[37] Weiner, J., Damgaard, C. Size-asymmetric competition and size-asymmetric growth in a spatially explicit zone-of-influence model of plant competition. *Ecological Research*, 21:707–712, 2006.

[38] Williams, M.S., Williams, M.T., Mowrer, H.T. A boundary reconstruction method for circular fixed-area plots in environmental survey. *Journal of agricultural, biological, and environmental statistics*, 6:479–494, 2001.

[39] Zhang, H., Zimmerman, D. Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*, 92:921–936, 2005.