Spatial Statistics 26 (2018) 38–55

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

A multi-scale area-interaction model for spatio-temporal point patterns

Adina Iftimi [a,*], Marie-Colette van Lieshout [b,c], Francisco Montes [d]

[a] Department of Bioscience and Nutrition, Karolinska Institutet, Stockholm, Sweden
[b] CWI, P.O. Box 94079, NL-1090 GB Amsterdam, The Netherlands
[c] Department of Applied Mathematics, University of Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands
[d] Department of Statistics and Operations Research, University of Valencia, Doctor Moliner, 50, 46100 Burjassot-Valencia, Spain

ARTICLE INFO

Article history:
Received 26 July 2017
Accepted 5 June 2018
Available online 19 June 2018

Keywords:
Gibbs point processes
Multi-scale area-interaction model
Spatio-temporal point processes
Varicella

ABSTRACT

Models for fitting spatio-temporal point processes should incorporate spatio-temporal inhomogeneity and allow for different types of interaction between points (clustering or regularity). This paper proposes an extension of the spatial multi-scale area-interaction model to a spatio-temporal framework. This model allows for interaction between points at different spatio-temporal scales and for the inclusion of covariates. We present a simulation study and fit the new model to varicella cases registered during 2013 in Valencia, Spain.

1. Introduction

Spatio-temporal patterns are increasingly observed in many different fields, including ecology, epidemiology, seismology, astronomy and forestry. The common feature is that all observed events have two basic characteristics: the location and the time of the event. In this paper we are mainly concerned with epidemiology (Stallybrass, 1931), which studies the distribution, causes and control of diseases in a defined human population. The locations of the occurrence of cases give information on the spatial behavior of the disease, whereas the times, measured on different scales (days, weeks, years, period of times), give insights on the temporal response of the overall process. An essential point

* Corresponding author.
E-mail address: iftimi@uv.es (A. Iftimi).

https://doi.org/10.1016/j.spasta.2018.06.001

to take into consideration is that people are not uniformly distributed in space, hence information on the spatial distribution of the population at risk is crucial when analyzing spatio-temporal patterns of diseases.

Realistic models to fit epidemiological data should incorporate spatio-temporal inhomogeneity and allow for different types of dependence between points. One important class of such models is the family of Gibbs point processes, defined in terms of their probability density function (van Lieshout, 2000; Ripley, 1988, 1990), and, in particular, the sub-class of pairwise interaction processes. Well-known examples of pairwise interaction processes are the Strauss model (Kelly and Ripley, 1976; Strauss, 1975) or the hard core process, a particular case of the Strauss model where no points ever come closer to each other than a given threshold. However, pairwise interaction models are not always a suitable choice for fitting clustered patterns. A family of Markov point processes that can fit both clustered and inhibitory patterns is that of the area- or quermass-interaction models (Baddeley and van Lieshout, 1995; Kendall et al., 1999). These models are defined in terms of stochastic geometric functionals and display interactions of all orders. Methods for inference and perfect simulation are available in Dereudre et al. (2014), Häggström et al. (1999), Kendall (2000) and Møller and Helisová (2010).

Most natural processes exhibit interaction at multiple scales. The classical Gibbs processes model spatial interaction at a single scale, nevertheless *multi-scale* generalizations have been proposed in the literature (Ambler and Silverman, 2010; Gregori et al., 2003; Picard et al., 2009). In this paper we propose an extension of the spatial multi-scale area-interaction model to a spatio-temporal framework.

The outline of the paper is as follows. Section 2 provides some preliminaries in relation to notation and terminology. Section 3 gives the definition and Markov properties of our spatio-temporal multi-scale area-interaction model. Section 4 adapts simulation algorithms, such as the Metropolis–Hastings algorithm, to our context. Section 5 treats the logistic regression approach and presents a simulation study. The model is applied to a varicella data set in Section 6. Section 7 presents final remarks and a discussion of future work.

2. Preliminaries

A realization of a spatio-temporal point process X consists of a finite number $n \geq 0$ of distinct points (x_i, t_i), $i = 1, \ldots, n$, that are observed within a compact spatial domain $W_S \subset \mathbb{R}^2$ and time interval $W_T \subset \mathbb{R}$. The pattern formed by the points will be denoted by $\mathbf{x} = \{(x_i, t_i)\}_{i=1}^n$. For a mathematically rigorous account, the reader is referred to Daley and Vere-Jones (2003, 2008).

We define the Euclidean norm $\|x\| = (x_1^2 + x_2^2)^{1/2}$ and the Euclidean metric $d_{\mathbb{R}^2}(x, y) = \|x - y\|$ for $x = (x_1, x_2) \in \mathbb{R}^2$ and $y = (y_1, y_2) \in \mathbb{R}^2$. We need to treat space and time differently, thus on $\mathbb{R}^2 \times \mathbb{R}$ we consider the supremum norm $\|(x, t)\|_\infty = \max\{\|x\|, |t|\}$ and the supremum metric $d((x, t), (y, s)) = \|(x, t) - (y, s)\|_\infty = \max\{\|x - y\|, |t - s|\}$, where $(x, t), (y, s) \in \mathbb{R}^2 \times \mathbb{R}$. Note that $(\mathbb{R}^2 \times \mathbb{R}, d(\cdot, \cdot))$ as well as its restriction to $W_S \times W_T$ is a complete, separable metric space. We write $\mathcal{B}(\mathbb{R}^2 \times \mathbb{R}) = \mathcal{B}(\mathbb{R}^2) \otimes \mathcal{B}(\mathbb{R})$ for the Borel σ-algebra and ℓ for Lebesgue measure. We denote by $\oplus$ the Minkowski addition of two sets $A, B \subset \mathbb{R}^2$, defined as the set $A \oplus B = \{a + b : a \in A, b \in B\}$.

As stated in Section 1, Gibbs models form an important class of models able to fit epidemiological data exhibiting spatio-temporal inhomogeneity and interaction between points. In space, the Widom–Rowlinson *penetrable sphere model* (Widom and Rowlinson, 1970) produces clustered point patterns; the more general area-interaction model (Baddeley and van Lieshout, 1995) fits both clustered and inhibitory point patterns. In its most simple form, the area-interaction model is defined by its probability density

$$p(\mathbf{x}) = \alpha \lambda^{n(\mathbf{x})} \gamma^{-A(\mathbf{x})} \tag{1}$$

with respect to a unit rate Poisson process on W_S. Here α is the normalizing constant, $\mathbf{x}$ is a spatial point configuration in $W_S \subset \mathbb{R}^2$, $n(\mathbf{x})$ is the cardinality of $\mathbf{x}$ and $A(\mathbf{x})$ is the area of the union of discs of radius r centered at $x_i \in \mathbf{x}$ restricted to W_S. The positive scalars λ, γ and $r > 0$ are the parameters of the model. Note that, as emphasized in van Lieshout (2000), Gibbsian interaction terms can be combined to yield more complex models. Doing so, Ambler and Silverman (2010), Gregori et al. (2003) and Picard