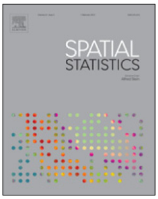
Spatial Statistics 20 (2017) 191–205

Contents lists available at ScienceDirect

Spatial Statistics

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



On parameter estimation for doubly inhomogeneous cluster point processes

Tomáš Mrkvička [a,*], Samuel Soubeyrand [b]

[a] Department of Applied Mathematics and Informatics, Faculty of Economics, University of South Bohemia, Studentská 13, 37005 České Budějovice, Czech Republic

[b] BioSP, INRA, 84914, Avignon, France

ARTICLE INFO

Article history:
Received 13 October 2016
Accepted 24 March 2017
Available online 13 April 2017

Keywords:
Bayesian estimation
Birth–Death–Move algorithm
Group dispersal model
Independent dispersal model
MCMC algorithm

ABSTRACT

Nowadays, spatial inhomogeneity and clustering are two important features frequently observed in point patterns. These features often reveal heterogeneity of processes/factors involved in the point pattern formation and interaction determining the relative locations of points. Thus, inhomogeneous cluster point processes can be viewed as flexible and relevant models for describing point patterns observed in biology, forestry and economics for example. In this article, we consider cluster point processes with double inhomogeneity in which locations of cluster centers are drawn under an inhomogeneous parametric intensity function and the distribution of clusters is spatially inhomogeneous and depends on a given parametric function. We propose a Bayesian estimation procedure based on an MCMC algorithm to simultaneously estimate inhomogeneity parameters, cluster parameters and cluster centers.

This modeling and estimation framework was applied to a toy case study dealing with the small-scale dispersal of spores of a fungal pathogen infecting plants.

---

* Corresponding author.
E-mail address: mrkvicka.toma@gmail.com (T. Mrkvička).

http://dx.doi.org/10.1016/j.spasta.2017.03.005

1. Introduction

Cluster point processes (Illian et al., 2008; Stoyan et al., 1995) form an important group of models often used in biology, forestry and economics, for example, to describe aggregated patterns of weeds, seedlings or firms (Arbia et al., 2012; Brix and Chadœuf, 2002; Stoyan and Penttinen, 2000). These models are most often built as doubly stochastic processes: firstly, cluster centers are randomly drawn under a given spatial point process; secondly, daughter points are randomly spread around cluster centers. Considering this typical construction, a cluster point process is usually specified by the distribution of cluster centers, the distribution of the number of points per cluster (thereafter called cluster size), and the distribution of daughter points around their cluster center (thereafter called cluster spread).

In various fields, improvements of technology have enabled scientists to collect large-scale point data sets. In general, mechanisms involved at large scale generate spatial heterogeneity in data that must be included in point process models built to analyze such data. Usually, second order intensity reweighted stationarity (Baddeley et al., 2000) are used to represent inhomogeneous processes because this is the most tractable type of inhomogeneity. Mechanically, it can be realized by thinning a homogeneous pattern by a thinning function, which is applied on every point independently. This type of inhomogeneity was assumed for example in the analysis of tropical forest data sets (Waagepetersen and Guan, 2009), where a two-step parameter estimation method was proposed. This method provides estimates of the inhomogeneity parameters in the first step, assuming no interaction between points, and it provides estimates of the cluster parameters in the second step, conditionally on the inhomogeneity estimated in the first step. Mrkvička et al. (2014) contrastively incorporated spatial inhomogeneity in cluster centers and applied the two-step estimation method for estimating parameters. Both types of inhomogeneities previously mentioned allow variation in the number of observed points, but they do not allow variation in cluster spread. The latter variation was allowed in locally scaled models introduced by Hahn et al. (2003), where the cluster spread decreases with increasing intensity of points. Mrkvička (2014) reviewed all these types of inhomogeneity in cluster point processes and proposed an approach to select the type of inhomogeneity contained in observed point patterns. Mrkvička et al. (2014) showed that the Bayesian method based on an MCMC algorithm is, among the competing methods, the most precise one (with respect to the mean square error) for estimating parameters of cluster process with inhomogeneous cluster centers. Moreover, this method can be easily adapted to other kinds of inhomogeneous models. Due to these reasons, the Bayesian estimation method based on an MCMC algorithm is applied here for the case of doubly inhomogeneous cluster point processes. We have also studied a moment method but obtained unsatisfactory results certainly due to the complexity of the model.

In this paper, we consider doubly inhomogeneous cluster point processes where cluster centers are drawn under an inhomogeneous parametric intensity function and the distribution of clusters is inhomogeneous in space and depends on a given parametric function. The inhomogeneity of the distribution of clusters can affect the spatial distribution of daughter points around their group center and/or the number of points per cluster, i.e. the cluster spread and size. A particular case of such doubly inhomogeneous cluster point process was built in Soubeyrand et al. (2011) to represent the dispersal of groups of particles (e.g. pollen, seeds or spores) from a point source (see also Allard and Soubeyrand, 2012, in this journal). In this context, the observed point pattern is formed by the deposit locations of particles and, in this pattern, groups of deposited particles form clusters. Intuitively, group dispersal can occur when several particles are (i) released from a source of particles because of a single wind gust, (ii) transported in the air into a varying volume of air, and (iii) deposited over a more or less spatially limited area. This heuristic vision of group dispersal leads to consider two inhomogeneities:

1. The intensity of cluster centers may vary in space depending on the strength and direction of wind gusts, the mass of particles and their shape (generally, in dispersal studies, the intensity of deposited particles decreases with distance from the source; Austerlitz et al., 2004; Tufto et al., 1997);
2. The spread of clusters increases with travelled distance (particles from the same group travel in an expanding air volume because of air turbulences accumulating with travelled distance).