Abstract

This thesis is about spatial statistics, which concerns the analysis of spatial data, that is, data collected at different locations in space. Examples of spaces are the d-dimensional Euclidean space $\mathbb{R}^d$ ($d \geq 2$), the sphere $\mathbb{S}^d$ ($d \geq 1$) or, if the data is collected over the time, the product space $\mathbb{S}^2 \times \mathbb{R}$. Such data is collected in many fields of science including agriculture, astronomy, biology, epidemiology, and physics. The main objective of spatial statistics is to describe and explain trends and interactions in the data. In this regard, spatial data is often studied as a realization of a general stochastic process defined over a region of interest. Considering this fact, spatial data, and so the underlying stochastic process, is classified into three classes: geostatistical data, where the stochastic process has a value at any location but only is measured at certain sites; lattice data, where the stochastic process has a value over a countable collection of spatial regions, and supplemented by a neighborhood structure; and point pattern data where the realization of the stochastic process is given by the locations of all occurrences of some event. Both geostatistical data and lattice data are modelled by random fields whilst point patterns are modelled as point process. This thesis focus on statistical models for geostatistical and point pattern data.

For both random fields and point processes, second order moment properties are used to describe spatial interaction at pair of locations. Notice that the interaction concept depends on the type of data. In fact, for random fields the interaction is between a pair of random variables at different locations and is quantified by the covariance function whilst for point processes the interaction is between pair of events and is quantified through the so-called pair correlation function or the K-function. The statistical analysis based on second order moment properties has been fairly well developed for points defined on $\mathbb{R}^d$.

Nowadays, there ia a lot of data that is collected on other spaces than $\mathbb{R}^d$. For example, global data is often collected by satellites, measuring some quantity on the Earth's surface, which, in turn, is approximated by $\mathbb{S}^2$. Moreover, if time is part of the data, then the spatial locations are on the product space between the sphere and the real line. For this type of data the statis-

tical tools developed for data observed on $\mathbb{R}^d$ are not suitable. As a matter of fact, such tools are not necessarily valid for analyze data on the sphere or the product space between the sphere and the real line. Actually, for random fields, the developed covariance function models on the Euclidean space are not necessarily valid models on the sphere and hence new statistical models are needed, whilst for point processes, the statistical models and functional summary statistics must be redefined and studied. This is the motivation of this thesis.

This thesis has three parts. Part I discusses the state of the art and review my research papers regarding random fields and point processes. For random fields, characterization theorems for covariance functions and a simulation algorithm when the spatial locations are on a $\mathbb{R}^d$, $\mathbb{S}^d$, or $\mathbb{S}^d \times \mathbb{R}$ are presented. For point processes, summary statistics and models are presented when the locations are on $\mathbb{R}^d$, $\mathbb{S}^d$, or $\mathbb{R}^d \times \mathbb{S}^k$. Part II are just copies of my six publications, where publications A–C are about random fields and publications D–F are about point processes. Publication A criticizes the Gneiting covariance function family and provides conditions such that this family yields a covariance functions with a counterintuitive statistical meaning. Publication B presents a family of covariance functions for geostatistical data collected on the sphere, capable of modeling the degree of smoothness in the mean–squared sense. Publication C shows a fast and exact algorithm to simulate random fields on $\mathbb{S}^d$ or $\mathbb{S}^d \times \mathbb{R}$. Publication D gives an extension of the K–function for point process on $\mathbb{R}^d$ to point processes on $\mathbb{R}^d \times \mathbb{S}^k$ and study the second order moment properties of point processes in the case of first and second order separability. Publication E develops an estimation procedure for the pair correlation function based on estimation equations that comes from variational equations. Paper F discusses the existence, moment properties, and estimation procedure of log-Gaussian Cox processes when the spatial locations are on the sphere. Finally, Part III shows a LASSO type estimator for multivariate log-Gaussian Cox processes with highly dimensional unknown variables.