

Preface

Spatial point processes are mathematical models that describe the arrangement of objects that are irregularly or randomly distributed in the plane or in space. The patterns formed by the objects are analysed in many scientific disciplines; hence a great variety of objects may be considered such as atoms, molecules, biological cells, animals, plants, trees, particles, pores, or stars and galaxies. At a basic level the data simply consist of point coordinates describing the locations of the objects, but additional characteristics of the objects can also be included in the analysis. These additional characteristics may, for instance, describe the size or type of an object and are usually referred to as ‘marks’. Point process analysis is in many ways distinct from the classical statistical methodologies presented in undergraduate textbooks. However, some of the more fundamental classical statistical issues remain influential; for example, sampling, exploratory data analysis, parameter estimation, model fitting, testing of hypotheses and separating signal from noise may all form part of a point process analysis.

Point process statistics is perhaps the most developed and beautiful branch of the modern field of spatial statistics; this is perhaps because points are the most elementary of geometrical objects and lead to data structures that are particularly clear and useful. Sometimes, however, point data have to be analysed in combination with other data from variables that vary continuously in space. This requires an application of spatial statistical methods that fall outside the realm of spatial point processes.

Recent decades have seen a strong increase in the development of point process methodology, based on a profound theoretical development and driven by applications from many different fields of science. In addition to the classical fields of application such as archaeology, astronomy, particle physics and forestry, today other fields such as ecology, biology, medicine and materials science extensively apply point process methods. This development is facilitated by the advent of new and improved technologies that may be used to collect point pattern data. Whereas the first point patterns were small and collected manually, modern data sets tend to be much larger and are collected using automated methods. Ecologists have found fascinating relations in plant communities by considering plants as points marked by characteristics such as size, species or genotype. For instance,

they have shown that spatial structure determines ecological processes in the short term and ecological processes modify spatial structure in the long term. Physicists, on the other hand, have used point process methods to study physical structures, for example, packings of hard spheres or other objects, where phase transitions appear. Astronomers have analysed the spatial distribution of galaxies in the universe with particularly powerful statistical methods based on point processes.

The aim of this book is to present statistical methods that are relevant in practice to readers from all these areas. Indeed, there is no point process methodology specific to ecology or physics; the methods are universal, and ideas developed in one field of application may be of value in another. Consequently, not every example in this book is of an ecological or physical nature. Ecologists and physicists are encouraged to translate the examples into their own language. In a few cases this may be difficult. For ecologists it may be impossible to apply the idea of packings of hard spheres and for physicists cluster processes rarely are suitable models for physical phenomena. Knowledge of such structures in one discipline may eventually turn out to be equally useful in the other; for example, solutions applied to the physical problem of packings of hard spheres have the potential to be informative in spatial studies on the swarming behaviour of birds and fish or, in the planar case, on the distribution of communities of plants.

Readers are encouraged to study all the examples even if these are from outside their specific field of interest, taking into account that it is really the geometrical structure that is being analysed. A pattern originating from an entirely different area of science may well be geometrically similar to patterns formed by more familiar objects. Consider, for instance, the pattern of gold particles, which is frequently discussed in this book. A pattern with similar geometrical properties might, on a different spatial scale and with its own interpretation, also appear in quite different contexts. The results of the statistical analysis should then be translated into the terminology relevant to the reader – this might even generate surprising new ideas.

Readers from fields of applications where only planar patterns are analysed are asked for their forbearance when spheres and even the d -dimensional case are discussed, as this is sometimes necessary for the sake of brevity and elegance.

This book, which may be regarded as a successor to Stoyan and Stoyan (1994), is intended for an audience of readers with widely varying knowledge of mathematics, statistics and computer science. The authors hope that this book will prove useful to students on a variety of courses, as well as to scientists both within and outside the field of mathematical statistics, who may be interested in the underlying principles and theoretical ideas. Some readers will write their own programs for the statistical procedures, many will work with open source libraries such as `spatstat` (<http://spatstat.org>) in R (R Development Core Team, 2007; Baddeley and Turner, 2005, 2006) or with commercial software, whereas others may simply want to understand the capabilities of point process statistics or the output generated by point process software.

The authors hope that readers will enjoy the large number of examples. They are encouraged to use the data files provided in <http://www.wiley.com/go/penttinen> and to analyse these in more detail using their own software. Some of the methods are presented without examples, mainly due to lack of space, or when further explanation seems unnecessary, as for example with the numerous indices. For some methods that have been developed recently, convincing examples could not be found. These methods are nevertheless presented here in the expectation of potential future applications.

The book mainly presents mathematical-statistical facts. Proofs of these are only provided when they are considered helpful in understanding the ideas underlying the statistical methods. The mathematically inclined reader may use this book as an introductory text, as a source of examples and ideas and as a motivation for further, more detailed study of these topics in other literature such as Stoyan et al. (1995), van Lieshout (2000) and Møller and Waagepetersen (2004).

The authors have tried to present many different methods developed in different fields of point process statistics that merit communication to a broader audience. This leads to an extensive presentation of non-parametric statistical methods. But it turned out to be impossible to present all the existing knowledge of point process statistics. In general, this book focuses on traditional and proven methods, which are preferred over mathematically complex methods and, therefore, such developing areas as spatio-temporal and Bayesian point pattern modelling are only briefly discussed.

This book is not intended to be read from cover to cover. Of course the chapters and sections have a logical order and the book can be read in this way. But the reader is perhaps more likely to regard some of the material as less important at a first reading. Hence, initially some sections may be ignored to provide a general understanding of the methods. These are marked by *. The reader is encouraged to jump from one chapter to another, and this is facilitated by a comprehensive index and a notation index. In particular, Chapter 3 may be ignored at a first reading unless a reader is specifically interested in finite point processes. Some basic knowledge of the ideas in Chapter 4 is helpful for an understanding of Chapter 3. Before reading Section 6.6 the reader should have read Section 3.6.

Chapter 1 presents fundamentals of the underlying theory, motivating examples, sampling methods and historical remarks. Chapter 2 studies a particular fundamental model, the Poisson process and, closely related to this, tests of the hypothesis of complete spatial randomness. Chapter 3 considers finite point processes – processes that exist only within a bounded window, which influences the distribution of the points. The important particular case of finite Gibbs processes (or Markov point processes) is discussed in much detail. The pivotal Chapter 4 presents the statistical theory for stationary point processes. It is this theory that many scientists refer to as ‘point process statistics’, as it comprises the K -function and second-order methods in general. Towards the end of the chapter some methods for clearly inhomogeneous patterns are also discussed. Chapter 5 discusses an analogous theory for marked point processes and presents a wealth of statistical characteristics

and methods. Chapter 6 introduces a suite of stationary point process models, after initially discussing general principles of model building. It considers classical models such as Cox processes, cluster processes, hard-core processes and stationary Gibbs processes. Additionally, spatial-temporal processes are considered, an area that is still in its infancy and currently undergoing rapid development. Furthermore, statistical methods are presented which enable the analysis of correlations between point processes and random fields or fibre processes. Finally, Chapter 7 presents important approaches to parameter estimation for point process models, and explains various, typically simulation-based tests for point process models.

All the methods are illustrated by many examples with data from various areas of application. These examples aim to show the reader the wide range and potential applications of the statistical methodology. They are listed for readers' convenience on pp. xvii–xix.

This book differs from others on point process statistics not only in its application-oriented approach but also in some technical aspects. Densities play a central role, in particular the pair correlation function, since these functions are easier to interpret than cumulative functions such as Ripley's K -function. Furthermore, many non-parametric methods and many new characteristics for marked point processes are presented. However, some of the more recent Markov chain Monte Carlo methods are discussed in less detail.

It is a pleasure for the authors to thank all those friends and colleagues who helped by providing data, suggesting references, allowing us to view unpublished manuscripts, and reading and commenting on preliminary drafts of sections or chapters. In alphabetical order these include: A. J. Baddeley, F. Ballani, S. Barot, A. Bellmann, U. Berger, D. Burslem, R. Capobianco, O. Davies, P. Diggle, D.J. Daley, A. Elsnér, C. Geiss, U. Hahn, J. Heikkinen, L. Heinrich, H. Hildenbrandt, A. Holroyd, V. Isham, U. Jansen, A. Järvinen, S. Kärkkäinen, G. Last, M.N.M. van Lieshout, K. Lochmann, K. Mäkisara, V. J. Martínez, T. Mattfeldt, J. Mecke, A. Mikhail, M. Myllymäki, M.-A. Moravie, M. Nummelin, Y. Ogata, J. Ohser, J. Pfänder, A. Pommerening, T. Rajala, E. Renshaw, B. D. Ripley, A. Särkkä, K. Schladitz, M. Sonntag, S. Soubeyrand, U. Tanaka, E. Tomppo, S. Torquato, M. Vihola, A. Wade, K. Wälder, A. Weiss, R. van de Weygaert and S. Wolf. We would also like to thank the students of the point process seminar at the University of Jyväskylä for helpful discussions. T. Rajala undertook the hard job of producing the many figures for this book.

In the summer of 2006 D.J. Daley showed a draft of Chapter 15 of Daley and Vere-Jones (2008) to one of the authors along with Vere-Jones (2008), which aided the writing of Section 6.10 on space–time processes. The book also benefited from fruitful discussions with U. Hahn on the statistics of non-stationary point processes. V. J. Martínez supported the authors in writing the text in Section 1.3.4. A. J. Baddeley informed the authors about the calculation of set-geometrical quantities. And M. N. M. van Lieshout discussed with the authors questions of the theory of finite point processes, leading to an improved presentation of Chapter 3.

This work was partially supported by the Academy of Finland (projects 111156 and 208284).

The authors are also grateful for the assistance of M. Robakowski, who did a large part of the technical work. Finally, they wish to express their gratitude to K. Sharples of John Wiley & Sons, Ltd for her continuous support, to Richard Leigh for his careful work and constructive suggestions as copy editor.

The authors

November 2007

