Math Geosci (2017) 49:815–817
DOI 10.1007/s11004-016-9670-x

BOOK REVIEW

A. Baddeley, E. Rubak, R. Turner: Spatial Point Patterns: Methodology and Applications with R

Boca Raton, CRC Press, 2016

Edith Gabriel[1]

Received: 9 December 2016 / Accepted: 23 December 2016 / Published online: 25 January 2017

Baddeley, Rubak and Turner dedicate their book to Professor Brian D. Ripley, whose pioneering statistical methods and statistical computing for spatial point process data constitute a substantial and still highly relevant contribution to the analysis of spatial point patterns. Interest for point process statistics has exploded since Ripley's work, leading to a wide range of related scientific questions, methods and statistical programming (see, e.g., Illian et al. 2008; Chiu et al. 2013). The authors successfully take up the difficult challenge to present to statistician and nonstatistician practitioners the methodology derived from spatial point process statistics through the use of the powerful open source software `spatstat`. The interplay between rich illustrative examples and the clear and rigorous presentation of mathematical methods make the reading of this long book (almost 800 pages) amazingly clear, easy and enjoyable.

The book makes it very accessible to learn how to analyze spatial point patterns and is written with a constant attention to practical needs and cautions in use. It is organized in four parts. The first part, 'Basics', introduces the `R` software, point pattern data and their handling with `spatstat`. If it states the goals of point patterns analyses (roughly describing the structure of the pattern in terms of intensity—interpreted as the rate of occurrence of the events, potentially spatially varying—and/or interpoint interaction—capturing stochastic dependence between events), it also advices on the relevance of the point process methods and the standard assumptions to the scientific context.

The second part concerns 'Exploratory Data Analysis.' It presents nonparametric tools to investigate the intensity and details classical summary statistics used to measure dependence between points, as Ripley's K-function, the pair correlation function

✉ Edith Gabriel
edith.gabriel@univ-avignon.fr

[1] LMA EA2151, Avignon University, 84000 Avignon, France

and the functionals based on spacings between points. Focus is put on the very popular K-function, both on its description and estimation and on the related caveats in its use and interpretation in specific contexts. This function, defined for a stationary point process, corresponds to the expected number of neighbors, up to a given distance, of a typical point of the point process. Its comparison to the one obtained under a null hypothesis like complete spatial randomness indicates if a typical point has more or less neighbors than would be expected, and thus may hint at clustering or regularity, respectively, with respect to the model under the null hypothesis. Attention is paid to its common misuses and misinterpretations when the process is not stationary, as the variation in intensity may then not necessarily be related to interpoint interaction. Modified K-functions are defined under less restrictive assumptions [second-order intensity reweighted stationarity or locally rescaled second-order stationarity developed in Baddeley et al. (2000) and Hahn and Vedel Jensen (2015), respectively] to take into account this inhomogeneity. Furthermore, Hahn and Vedel Jensen (2015) provide a general method to test the hypothesis that a point pattern can be assigned one of these type of second-order stationarity as presented at the end of the book.

The third and main part of the book considers 'Statistical Inference.' It describes the usual models of point processes (Poisson, cluster, Cox and Gibbs models), their inference and validation and presents how to conduct hypothesis testing. This part reveals the strength of `spatstat`, with various models (including multi-type point patterns) and sophisticated likelihood-based methods implemented for model fitting. About one hundred pages deal with Poisson point processes, where the authors cleverly connect this model to classical statistical methods: logistic regression and maximum entropy for model fitting and residual analysis for model validation. Hypothesis testing is an important step in data analysis, covering the detection of interpoint interaction, goodness of fit and model selection. Because in spatial statistics, the distribution of the test statistics under the null hypothesis is often unknown, Monte Carlo methods are of common use. The authors detail Monte Carlo tests in their general setting and their use for tests based on simulation envelopes. They thoroughly distinguish pointwise and global envelopes, detailing the multiple testing caveat of the former. Further caveats and alternative methods are presented in Mrkvička et al. (2016), as the rank envelope tests (Myllymäki 2016), which provides global graphical envelopes together with p values.

The fourth part presents the analysis of point patterns with 'Additional Structure,' as marked point patterns, replicated patterns and point patterns on a linear network. Many spatial point process data are snapshots of spatiotemporal point processes. When the evolutionary process is of scientific interest, the temporal information can be thought of as an additional structure of the spatial point process. If statistics and models for spatiotemporal point processes are not presented in a specific chapter, they can nevertheless be found in the ones related to marked point processes, as time can sometimes be treated as a mark. It is often more relevant to consider time as an additional dimension, fundamentally different of the spatial dimensions. Most of the summary functions and models presented in this book have been extended to the spatiotemporal setting (see González et al. 2016 for a review) and are implemented in the package `stpp` (Gabriel et al. 2013).

To sum up, I highly recommend this book. It is a wonderful achievement, with a good balance between the description of statistical methods and models and their handling in practice with `spatstat`. The numerous and relevant data sets used as examples make it easier to understand results, interpretations and caveats by any scientist dealing with spatial point process data. Among the things I like in this book are the FAQ sections ending each chapter. They definitely answer questions we met at least once as users, but were sometimes afraid to ask.

References

Baddeley A, Møller J, Waagepetersen R (2000) Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. Stat Neerl 54:329–350

Chiu S, Stoyan D, Kendall W, Mecke J (2013) Stochastic geometry and its applications, 3rd edn. Wiley, New York

Gabriel E, Rowlingson B, Diggle PJ (2013) stpp: an R package for plotting, simulating and analysing Spatio-temporal point patterns. J Stat Softw 53(2):1–29

González J, Rodríguez-Cortés F, Cronie O, Mateu J (2016) Spatio-temporal point process statistics: a review. Spat Stat. doi:10.1016/j.spasta.2016.10.002

Hahn U, Vedel Jensen E (2015) Hidden second-order stationary spatial point processes. Scand J Stat 43(2):455–475

Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns. Wiley, London

Mrkvička T, Soubeyrand S, Myllymäki M, Grabarnik P, Hahn U (2016) Monte Carlo testing in spatial statistics, with applications to spatial residuals. Spat Stat 18(A):40–53

Myllymäki M, Mrkvička T, Grabarnik P, Seijo H, Hahn U (2016) Global envelope tests for spatial processes. J R Stat Soc ser B. doi:10.1111/rssb.12172