

Preface

Interest in analysing spatial data has grown considerably in the scientific research community. This reflects the existence of well-formulated questions or hypothesis in which location plays a role, of spatial data of sufficient quality, of appropriate statistical methodology.

In writing this book I have drawn on a number of scientific and also policy-related fields to illustrate the scale of interest – actual and potential – in analysing spatial data. In seeking to provide this overview of the field I have given a prominent place to two fields of research: Geographic Information Science (GISc) and applied spatial statistics.

It is important as part of the process of understanding the results of spatial data analysis to define the relationship between geographic reality and how that reality is captured in a digital database in the form of a data matrix containing both attribute data and data on locations. The usefulness of operations on that data matrix – revising or improving an initial representation (e.g. spatial smoothing), testing hypotheses (e.g. does this map pattern contain spatial clusters of events?) or fitting models (e.g. to explain offence patterns or health outcomes in terms of socio-economic covariates) – will depend on how well the reality that is being represented has been captured in the data matrix. Awareness of this link is important and insights can be drawn from the GISc literature.

I have drawn on developments in spatial statistics which can be applied to data collected from continuous surfaces and from regions partitioned into sub-areas (e.g. a city divided into wards or enumeration districts). In covering this material I have attempted to draw out the important ideas whilst directing the reader to specialist sources and original papers. This book is not an exhaustive treatment of all areas of spatial statistics (it does not cover point processes), nor of all areas of spatial analysis (it does not include cartographic modelling).

Implementing a programme of spatial data analysis is greatly assisted if supporting software is available. Geographic information systems (GIS) software are now widely used to handle spatial data and there is a growing quantity of software some of it linked to GIS for implementing spatial statistical methods. The appendix directs the reader to some relevant software.

Readership

This book brings together techniques and models for analysing spatial data in a way that I hope is accessible to a wide readership, whilst still being of interest to the research community.

Parts of this book have been tried out on year 2 geography undergraduates at the University of Cambridge in an eight-hour lecture course that introduced them to certain areas of geographic information science and methods of spatial analysis. The parts used are chapters 1, 2, sections 3.1, 3.2.1, 3.2.3, 3.2.4(a) from chapter 3, selected sections from chapter 4 (e.g. detecting errors and outliers, areal interpolation problems), selected sections from chapter 7 (section 7.1.3, map smoothing) and some selected examples on modelling and mapping output using the normal linear regression model. In associated practicals simple methods for hot spot detection are applied (the first part of section 7.3.1(a)) together with logistic regression for modelling (along the lines of section 11.2.2(a)).

Parts of the book have been tried out on postgraduate students on a one year M.Phil. in Geographic Information Systems and Remote Sensing at Cambridge. One 16-hour course was on general methods of spatial analysis but particularly for data from continuous surfaces. In addition to some of the foundation material covered in chapters 1 to 4 there was an extended treatment of the material in section 4.4.2 with particular reference to kriging with Gaussian data (including estimation and modelling of the semi-variogram taken from chapter 10 and the references therein). A second 16-hour course dealt with exploratory spatial data analysis and spatial modelling with reference to the analysis of crime and health data. This focused on area data. The material in chapter 7 was included with an introduction provided by the conceptual frameworks described in chapter 5. The part of the course on modelling took selected material from chapter 9 and drew on examples referred to in that chapter and chapter 11.