SPRINGER BRIEFS IN REGIONAL SCIENCE

Manfred M. Fischer · Jinfeng Wang

Spatial Data Analysis

Models, Methods and Techniques

Springer

Preface

The centrality of space and location has always been taken as granted in geography and regional science. But recent attention to the spatial dimension of phenomena has also increased in the mainstream of the social sciences, and increasingly also in the natural sciences like ecology. A growing number of social scientists have taken up the use of new methodologies and technologies (such as geographic information systems, global positioning systems, remote sensing, spatial statistics and spatial econometrics) in the empirical work. In addition, increased attention is paid to location and spatial interaction in theoretical frameworks.

In broad terms, one might define spatial analysis as the quantitative analysis of spatial phenomena that are located in geographical space (Bailey and Gatrell 1995). It would be too ambitious to cover such a broad field in one textbook (see Fischer and Getis 2010 for an accounting of the diversity of the field). We thus decided to limit the scope to that important subset of spatial analysis which is known as *spatial data analysis*. In doing so, we are concerned with the situation where observational data are available on some process operating in geographic space, and consider models, methods and techniques to describe or explain the behaviour of this process and its possible relationship to other spatial phenomena. By defining spatial data analysis in this way we place the book in the area of statistical description and modelling of spatial data, and restrict ourselves to a particular set of methods. In doing so, we exclude some important quantitative methods such as, for example, various forms of network analysis and location-allocation analysis that would be included under the more general heading of spatial analysis.

Whether or not spatial data analysis is a separate academic field, the fact remains that in the last twenty years spatial data analysis has become an important by-product of the interest in and the need to understand spatial data. By spatial data we mean data which relate to observations with a spatial reference where spatial reference may be explicit, as in a postal address or a grid reference, or implicit, as a pixel in remote sensing.

The past decades have generated a number of excellent texts on the subject (see, for example, Cliff and Ord 1981; Upton and Fingleton 1985; Anselin 1988b;

Griffith 1988; Ripley 1988; Cressie 1993; Haining 1990, 2003; Bailey and Gatrell 1995; LeSage and Pace 2009). Most of these are addressed to the researcher. This text book is directed at introducing spatial data analysis to the graduate student, from a "data-driven" rather than a "theory-led" perspective. With this overall objective in mind, we have not attempted to discuss exhaustively the whole area of spatial data analysis, but have restricted the discussion to the analysis of two major types of spatial data: area data defined as data associated with a fixed set of areas or zones covering the study area, and spatial interaction (or origin–destination flow) data defined as measurements each of which associated with a link or pair of locations representing points or areas in geographical space.

We have restricted ourselves to a subset of models, methods and techniques which we believe to be relatively accessible and useful for analysing these types of spatial data. The topics discussed in this book include a mixture of both informal/exploratory methods and techniques on the one side, and formal statistical modelling, parameter estimation and hypothesis testing on the other.

The book is divided into two parts. Each of these parts is as self-contained as possible. The first, Part I, considers the analysis of *area data*. The areas may form a regular lattice, as with remotely sensed images, or be a set of irregularly shaped areas, such as administrative districts. The second part, Part II, shifts attention to the analysis of spatial interaction data which are related to pairs of points or areas. Such data—called origin–destination flow or spatial interaction data—are relevant in studies of transport planning, population migration, journey-to-work, shopping behaviour, freight flows, and even the transformation of information and knowledge.

We do not consider spatiotemporal data, but assume that the data are purely spatial, either having been aggregated over time or referring to fixed points in time. Issues of measurement, storage and retrieval of spatial data are important, but outside the scope of this textbook. GISystems offer software tools that facilitate—through georeferencing—the integration of spatial and non-spatial, qualitative and quantitative data in a database that can be managed under one system environment (see Longley et al. 2001 for a discussion). In keeping the text in a manageable number of pages we assume our reader to have a moderate level of general background in statistics and mathematics.

We acknowledge the generous general support provided by the Institute for Economic Geography and GIScience, Vienna University of Economics and Business. We have benefitted greatly from the technical assistance Thomas Seyffertitz (Institute for Economic Geography and GIScience) provided. His expertise in handling several word processing systems, formatting, and indexing, together with his care and attention to detail, helped immeasurably. Last but not at least we thank the editor of the book series, Dr. Henk Folmer, for his valuable comments on an earlier version of the manuscript.

Vienna, Beijing, May 2011

Manfred M. Fischer
Jinfeng Wang

Contents