Introduction

Robert J. Stimson

This volume is the final product from the *Australian Research Council Research Network in Spatially Integrated Social Science* (ARCRNSISS), which was funded by the Council from late 2004 to the end of 2009. The chapters in the book are based on presentations from a wide range of social scientists – mostly from Australia but also from the US, New Zealand and the UK – who were involved as instructors in the annual residential *Summer School for Graduate Students in Spatially Integrated Social Science Theory, Methods and Applications* which was attended each February by 20 to 25 graduate students. The purpose of the Summer School was for graduate students drawn from a range of disciplines in the social sciences and from across the Australian university system to receive intensive training in research taking a spatially integrated approach in the social sciences.

Over the last couple of decades there has been a growing interest in spatially integrated social science (SISS) research internationally, reflecting an increasing recognition of the importance of the roles of space and place as important consideration in analysing and modelling social and economic phenomena and in investigating human behaviour. That is reflected in part in a growing number of journals that are interdisciplinary in scope and that have an explicit focus on spatial phenomena. It is also a reflection of the degree to which much of the data used by social scientists is spatial in nature, especially data that is collected by official agencies (such as census bureaux) that is aggregate in nature and made available for spatial units that are largely *de jure*. It is also reflected in the emergence of sophisticated micro-simulation methods whereby it is possible to ‘merge’ unit record survey data (individual data) with spatially aggregated data (such as the census data) to produce simulated synthetic estimates across space of phenomena that relate to measures in those survey data collections.

Building on the explicitly spatial theoretical basis of human geography and regional science in particular, there has now emerged a powerful theoretical basis that underpins a *spatially integrated approach in social science research*. That is enhanced by researchers making use of methodologies that explicitly incorporate *space* and *place* considerations in analytical routines and models which, importantly, incorporate a consideration of variation in spatial scale and that give due recognition to the *spatial*
Introduction

aggregation–disaggregation problem and the modifiable area unit problem (MAUP) that researchers confront in analysing spatial data and in operationalizing models that use spatial data.

The chapters in this book provide a reasonable but not complete coverage of the theoretical underpinnings and the methodologies that typify research using a SISS approach. The book is intended primarily as a primer for students and budding researchers who wish to investigate social, economic and behavioural phenomena by giving explicit consideration to the roles of space and place. In the majority of the chapters as appropriate there is an emphasis on demonstrating applications of methods, tools and techniques that are used in SISS research, both long-established and relatively new approaches.

The book is organised into five parts. Part I, ‘A Spatially Integrated Social Science Approach’, has three chapters. In Chapter 1 Robert Stimson provides a discussion of the context for the evolution of a spatially integrated approach in social sciences research, pointing out that an explicitly spatial approach requires the development and use of socio-spatial theory that takes account of how space manifests itself. That has required a new meta-paradigm which has seen the development of explicit theories, methods, tools and techniques that are used in SISS analysis. The chapter provides a brief discussion of the research process and some of the issues that might be of concern in conducting SISS research. Stimson points out that the purpose of research is fundamentally about seeking explanation, and emphasizes how this is achieved through using the scientific model of investigation which includes testing theories by collecting and analysing data in operationalizing models. The chapter discusses how taking a SISS approach to conduct research in the social sciences necessitates spatializing data.

In Chapter 2 Michael Goodchild, Donald Janelle and Karl Grossner discuss how, in recent years, a ‘spatial turn’ in the social sciences has been well documented, spurred in part by: the availability of geographic information systems (GIS) and other computational tools that support a spatial perspective; the availability of geo-referenced data through Global Positioning System (GPS) and other technologies; and a recognition that a spatial perspective can lead to valuable insights and the development of new theory about social and environmental phenomena distributed over the surface and near-surface of the Earth. The authors discuss a range of views on what constitutes a spatial perspective in undertaking research, outlining a number of key fundamental spatial concepts and principles that constitute spatial knowledge. They place an emphasis on discussing geographic space, spatial processes, and understanding time–space patterns in making inferences about relational causality that enhances critical thinking in research.
In Chapter 3 Donald Janelle discusses time–space convergence, in which the connection time between places changes over an interval of time as a result of innovations in transport and communication technologies; and human extensibility, which focuses on how individual human beings and institutions exploit convergence technologies to project their presence and influence beyond their physical locales and to expand their capacities to interact at a distance from their current locations. He demonstrates how spatial reorganization is a process whereby the spatial patterns of resources and settlement systems are reconfigured to reflect differential changes in local, regional, national and global accessibility, and how a myriad of decision-makers act alone or in selective collaborations in a complex web of interdependencies. These issues represent substantial theoretical and methodological challenges for research, but they also provide useful frameworks for conducting a wide range of research issues in the social sciences, including research that undertakes ecological analysis, pattern analysis, and the investigation of spatial dynamics and spatial behaviour.

**Part II** has two chapters on ‘Setting Up Your Research’. In Chapter 4 Robert Stimson discusses how the research process is fundamentally about generating new information to improve our understanding about social and economic phenomena, events and situations, and human behaviours, and to help explain their occurrence. Fundamentally the objective of research is to provide explanation. It is about explaining why, when, where and how those things occur, and it is thus goal-oriented. To do that we need data about those things we are investigating. The chapter provides an overview of the modes of enquiry that are typically embedded within a scientific paradigm in undertaking research, which includes the deductive and inductive approaches to reasoning, and the use of qualitative, quantitative and mixed-mode approaches. The chapter provides a discussion of the use of models and the skills needed to build and operationalize them.

A fundamental element in undertaking a research project is conducting a literature review, and how to go about doing that is discussed by Kevin O’Connor in Chapter 5. The chapter provides many useful suggestions for researchers in conducting and writing a literature review, particularly for students in undertaking research for a dissertation. It emphasizes how a literature review is a carefully constructed argument that leads to formulating a research question.

**Part III** has eight chapters that focus on ‘Data Sources, Data Collection and Information Generation’. The discussions in these chapters are more generic for social science research in general, but as the authors point out we often face specific challenges and issues that need to be addressed in using data that is spatial.

In Chapter 6 Robert Stimson discusses issues to do with using data in
research, either existing data or by generating new data. Data involves measurement, which enables manipulation using statistical and other procedures. The chapter discusses forms of data, the diverse sources of data used in research, issues arising in the explicit use of spatial data, and the issue of data quality.

A lot of research in the social sciences uses data from official census collections which provide a complete count of the population, and this is discussed by Graeme Hugo in Chapter 7. The chapter discusses how a census is the significant source of demographic, social and economic information about a nation’s population and, in addition, usually its dwellings; and it provides detailed data about population characteristics at an aggregate level but usually down to a relatively fine geospatial scale. Using the Australian census as an example, Hugo provides a discussion of the nature of census data, and its strengths and limitations for spatial analysis. Issues to do with how censuses address the concept of ‘place’ are discussed – including the hierarchy of official census geographies – as are how the de facto and de jure population of an area is determined, along with a consideration of how confidentiality is assured. The chapter provides a detailed discussion of the questions that are typically asked in a census and the type of data variables that are thus generated to give information about the official levels of census geographies. Hugo argues that the potential for using census data in spatial analysis is not always fully exploited, and it is suggested that there is perhaps a need to rethink some of the fundamental concepts on which modern censuses are based to better facilitate the analysis and understanding of contemporary spatial, demographic, social and economic change at various levels of scale.

It is common in social science research for primary data to be collected using a survey approach. In Chapter 8 Robert Stimson provides an overview of the challenges and pitfalls researchers face in conducting a survey. The chapter outlines the tasks that are involved in undertaking survey research. The discussion focuses on the choice of modes for survey data collection, issues to do with sampling – especially probability sampling – in survey design, non-probability sampling approaches, dealing with survey error and bias, the principles in designing a survey instrument to collect data (which is typically a questionnaire), the challenges that respondents face in answering survey questions, and how to deal with missing data.

In Chapter 9 Mark Western discusses the use of quantitative data in social science research, providing a general overview of quantitative analysis and an introduction to using some statistical methods to analyse data. The chapter discusses a range of issues including the nature of data, measurement scales, measurement quality, linking data and data analysis in the investigation of relationships, using samples, using statis-
tics and statistical inference in quantitative data analysis, and statistical modelling.

In Chapter 10 the focus changes to a discussion of qualitative methods in social science research, which have a long history. Authors Phillip O’Neill and Pauline McGuirk outline the rationale for using qualitative methods in research, and discuss its origins. They explicitly focus on the use of qualitative methods in a geographical context. It is pointed out how it is not productive to divide socio-spatial methods into the two neat categories ‘quantitative’ and ‘qualitative’ which is commonly done.

In Chapter 11 Andrew Beer and Debbie Faulkner focus on how to use primary data and secondary data in research, providing a discussion of the relative advantages of these approaches, and illustrating how both primary and secondary data may provide valuable insights in addressing social, economic and behavioural phenomena. They discuss how these approaches may be used to enhance understanding of the specific social phenomenon of the movement of young people out of the family home.

The imperatives and pitfalls involved in forecasting in social science research are discussed by Tony Sorensen in Chapter 12. Forecasts are statements about the future and are surrounded by uncertainty to which we often attach probabilities and error terms in making estimates, and in the social sciences those error terms are likely to be large on account of the often huge number of complexly interconnected variables in play, poor data and model specification, and the frequency of surprise events. The chapter discusses the principal ingredients in forecasts and the value of forecasts, and it provides an overview of selected forecasting techniques that are commonly used in social science research to generate potential information about the future.

In the final chapter in Part III (Chapter 13) Jacques Poot discusses the use of meta-analysis of previous empirical research findings as a research method. Meta-analysis has a long history, especially in experimental research, and it is gaining increasing use to generate information and insights on social phenomena. The chapter provides a stylized example of meta-analysis of data, describing step by step how it is conducted, and it outlines how meta-analysis can be conducted when the primary study results are summarized in a categorical way.

Part IV has 13 chapters on ‘Research Tools and Techniques and Applications’. These chapters cover a range of both long-established and relatively new methods that are used in research in the social sciences that is explicitly (or at least implicitly) spatial in context. The chapters provide examples of applications of those specific methodologies, tools and techniques that are used to analyse, model and visualize social, economic and behavioural phenomena that illustrate a SISS approach.
Using spatial data in social science research invariably involves the need to visualize data. In Chapter 14 Tung-Kai Shyy, Imran Azeezullah, Irfan Azeezullah, Robert Stimson and Alan Murray illustrate how cartographic approaches developed by geographers, and now greatly enhanced by the ready availability of GIS tools, are used in research to visualize the distributitional aspects of phenomena. Typically that involves classifying data to produce a visual pattern such as a choropleth map. The chapter discusses a number of classification approaches that have been developed to generate map displays of socio-spatial data at point and polygon levels. The choice of method for classification of spatial data influences the visual impact of the mapped data and how it might be interpreted. The chapter illustrates the importance of using benchmarked measures, such as the location quotient, and shows how sophisticated categorization routines are particularly useful for developing typologies in mapping spatial data, particularly in the development of spatial decision support tools.

The use of indexes has a long history in the social sciences, especially in economics, as does the use of spatial indexes, especially in sociology, geography and demography. In Chapter 15 Martin Watts provides an overview of the conceptual and methodological issues confronting researchers in the social sciences in investigating the distributional aspects of social phenomena, illustrating that through a discussion of the use of segregation indexes to measure social segregation, both aspatial and spatial. The challenges involved in using spatial data are paid particular attention.

Shift-share analysis is a tool that is widely used in economics, geography and regional science to analyze regional differentiation. It recognizes that all regions at one level or another are embedded in, and are constituent parts of, larger regional systems. The application of shift-share analysis has been used predominantly to analyse change over time in regional economic performance, such as employment in industry sectors. In Chapter 16 Kingsley Haynes and Jitendra Parajuli discuss and illustrate the use of the shift-share analysis as a tool to decompose a region into its fundamental building blocks. That involves decomposing a total shift in a phenomenon over time into its three additive components: a national (or other reference area) share of the total shift; a structural share of the total shift; and a residual that is the regional or differential share of the total shift. The chapter provides a discussion of the evolution of shift-share analysis illustrating how the traditional shift-share model has been extended to incorporate productivity and inter-regional trade effects.

In Chapter 17 William Mitchell focuses on the implications of using spatial data in econometric analysis which predominantly uses regression models. The chapter discusses the specific field of spatial econometrics, showing how we encounter the spatial autocorrelation problem when
modelling spatial data based on arbitrarily demarcated spatial units. Ignoring this spatial dependence between neighbouring regions (spatial units for which data is available) leads to biased regression results. The chapter takes the famous data for the city of Columbus in the US used by regional scientist Luc Anselin to illustrate how spatial econometric techniques that have been developed by regional scientists as part of the new field of spatial statistics enable us to address the spatial autocorrelation problem through the use of spatially weighted regression methods by making appropriate weight adjustments to the data matrix.

It is common in a SISS approach for a researcher to seek to develop typologies out of the analysis of spatial data. In Chapter 18 Roger Stough discusses some of approaches that are used to identify spatial clusters illustrated in the context of investigating aspects of regional economies as an example to show how spatial pattern data may be investigated in the social sciences. Methodologies to identify and measure industry clusters and visualization tools to represent them are discussed. The chapter discusses how it is important to differentiate between absolute and relative concentration in a cluster, and the importance of investigating change over time – the dynamics of clusters. Stough also discusses the structural properties of clusters and how techniques such as spider diagrams may be used to measure cluster performance. The life cycle of clusters is also addressed.

Research into some social and behavioural phenomena in the social sciences involves the analysis and modelling of interactions or flows in a spatial framework, where the data used is in the form of a data matrix measuring interactions and flows between places, such as movements of people between origins and destinations for inter-regional migrations. In Chapter 19 Martin Bell and Dominic Brown focus on this phenomenon, referring to the extensive literature in geography, demography and regional science on analysing spatial interactions and flows. The chapter discusses such models in the context of investigating inter-regional migration, demonstrating the application of a series of statistical measures and mapping routines to visualize the flow patterns.

Approaches to visualizing flows data is considered further by Jonathan Corcoran and Chris Brunsdon in Chapter 20. Their focus is on the application of circular statistics – a specialized branch of statistics that has been used in a number of scientific fields but rarely in the social sciences – as a tool to visualize and analyse directional data by exploring its cyclical properties. The chapter discusses the principles of circular statistics and introduces a series of circular measures and graphical devices that can be applied to explore cyclic data, illustrating their applications to analyse patterns in hoax calls of urban fires and the patterns in journey-to-work data for a city.
The nature of networks is of interest to many social science researchers. In Chapter 21 psychologists Galina Daraganova and Philippa Pattison provide an introduction to the concepts and methods of social network analysis, and illustrate their application in the social sciences. The primary focus of these methods is the analysis of connections among a group of social actors. The chapter discusses graph theoretic concepts and methods for exploring cohesive subgroups within networks. The chapter addresses the structural properties of actors, and discusses the statistical analysis of graphs in the investigation of networks, referring to a number of widely used software packages.

Social scientists are making greater use of path analysis which models the effects of independent variables on dependent variables via intervening or mediating or moderating variables, investigating the pathways by which variables affect each other. These models, which are also referred to as structural equation models, are discussed by Rod McCrea in Chapter 22. Path analysis is particularly useful in analysing unit record data collected through survey research whereby the objective is to explain how a behavioural phenomenon – as measured by the dependent variable – might be the outcome of independent or exogenous variables which predict the intervention and/or might be influenced by mediating variable(s) which, in turn, might predict the dependent or endogenous variable. By adding additional variables and paths, models can quickly become much more complex to describe, so they are invariably represented diagrammatically to assist with explication. In the context of a SISS approach, the chapter provides examples of where using path analysis models – both regression and more complex models using measured variables and using latent variables – have been used to investigate the links between objective dimensions (including spatial objective measures relating to an urban environment) and subjective evaluations of the urban environment (including levels of satisfaction with urban domains) in the context of quality of urban life research. Path analysis is a particularly useful analytical tool that may be used to test hypotheses.

How to integrate unit record survey-based data and spatial objective data is attracting more attention in SISS research. In recent times there has been substantial effort put into developing more sophisticated methods to do so. This is discussed in Chapters 23 through 25.

In Chapter 23 Prem Chhetri and Robert Stimson demonstrate how a series of methods and tools derived from GIS technologies may be used to achieve that integration, once again illustrating that in the context of modelling aspects of quality of urban life by way of example. Specifically the chapter shows how tools such as the weighted average operator may be used to derive composite indexes, and how the ordered average operator
may be used for aggregating multiple indicators to form and map overall scores representing subjective evaluations of environmental settings. The chapter also illustrates how GIS routines such as the neighbourhood operation routine and statistical decomposition tools such as principal components analysis may be used to derive generalized spatial representation of subjective evaluations of an environmental setting such as residential areas in a city. Finally, the chapter shows how GPS technology may be used to enhance the investigation of human spatial behaviour in a space–time context when linked to a survey research approach to data collection.

A further application of GIS technology in social science research is illustrated in Chapter 24. Using the collective efficacy framework that sociologists have developed to investigate variations in crime and communities, Tung-Kai Shyy, Lorraine Mazerolle, Kate Riseley and Robert Stimson demonstrate the development of a web-based GIS and the application of an online tool to support the analysis of survey data which collected information on community variations in crime. The chapter describes how a spatial database may be built to allow the mapping and identification of spatial patterns and spatial clusters and show the relationships between different layers of information such as crime and collective efficacy using survey data.

One of the important methodological developments in SISS research is the use of micro-simulation tools to create synthetic data by linking large-scale unit record survey data with small-area aggregated spatial data such as that available in census data collections. This is discussed by Ann Harding and Robert Tanton in Chapter 25. The authors describe a programme of research at the National Centre for Social and Economic Modelling (NATSEM) in Australia that has continued for more than a decade. The methodologies result in a set of relatively well-established procedures for reweighting sample survey micro-data to census small-area targets, with the resultant creation of a synthetic micro-database of individuals and households for each small area. Such an approach has wide application to inform public policy by simulating the potential impacts of changes in social and economic policy, such as the likely impact of a change in income tax or welfare payments, on different individuals or households.

In Chapter 26 David Rohde and Jonathan Corcoran discuss graphical models and Bayesian networks as a spatial analytical tool in which conditional probability is used to investigate causal relationships. They apply this to a study investigating the causal relationships between the installation of smoke alarms in residential dwellings and fire-related deaths.

Finally, Part V has two chapters on ‘Producing Research Output’. In Chapter 27 Brian Head, who has worked in both academe and public
Introduction

policy agencies, addresses how quality social science research can be undertaken and communicated in ways that are useful for government policy-makers and other end-users so as to improve impacts but without undermining reliability, objectivity and independence. He also discusses applied research that is commissioned under research contracts. There is a discussion of the expressed and implicit needs of government decision-makers for certain types of information about the nature of social problems, the effectiveness of programmes, and the cost or benefit of various alternative options. Head points to gaps in knowledge about the extent to which research findings are actually utilized by potential end-users, and emphasizes that there is widespread agreement that research findings are most influential when they are well targeted and well communicated. Achieving that is enhanced if researchers and policy-makers try to learn more about each other’s business, and to better manage their mutual expectations in a joint quest for improved policy-relevant understanding. The chapter discusses how to go about choosing policy-significant and policy-relevant issues for research in the social sciences, and what it is about research findings that policy-makers might appreciate and value in providing an evidence base to assist policy formulation and implementation.

In Chapter 28 Rebecca Wickes and Tara McGee – who participated in one of the ARCRNSISS Summer Schools as doctoral (PhD) candidates – discuss their experience in successfully undertaking research for their doctoral thesis and the lessons they have to pass on, including going about choosing an advisor, getting started, developing networks, managing time, writing up the research to complete the dissertation, and understanding what examiners want.