Modelling Spatial and Spatial-Temporal Data: a Bayesian approach

Robert Haining and Guangquan Li

Table of contents

Preface

Part I. Fundamentals for modelling spatial and spatial-temporal data

Chapter 1. Challenges and opportunities analysing spatial and spatial-temporal data

1.1 Introduction

1.2 Four main challenges when analysing spatial and spatial-temporal data

1.2.1 Dependency

1.2.2 Heterogeneity

1.2.3 Data sparsity

1.2.4 Uncertainty

1.2.4.1 Data uncertainty

1.2.4.2 Model (or process) uncertainty

1.2.4.3 Parameter uncertainty

1.3. Opportunities arising from modelling spatial and spatial-temporal data

1.3.1 Improving statistical precision

1.3.2 Explaining variation in space and time

1.3.2.1 Example 1: Modelling exposure-outcome relationships

1.3.2.2 Example 2: Testing a conceptual model at the small area level

1.3.2.3 Example 3: Testing for spatial spillover (local competition) effects

1.3.2.4 Example 4: Assessing the effects of an intervention

1.3.3 Investigating space-time dynamics

1.4. Spatial and spatial-temporal models: bridging between challenges and opportunities

1.4.1 Statistical thinking in analysing spatial and spatial-temporal data: the big picture

1.4.2 Bayesian thinking in a statistical analysis

1.4.3 Bayesian hierarchical models

1.4.3.1 Thinking hierarchically

1.4.3.1.1 The data model

1.4.3.1.2 The process model

1.4.3.1.3 The parameter model

1.4.3.2 Incorporating spatial and spatial-temporal dependence structures in a Bayesian hierarchical model using random effects.

1.4.3.3 Information sharing in a Bayesian hierarchical model through random effects

1.4.4 Bayesian spatial econometrics

1.5 Concluding remarks

1.6 The datasets used in the book

1.7 Exercises

Chapter 2. Concepts for modelling spatial and spatial-temporal data: an introduction to "spatial thinking"

2.1 Introduction

2.2. Mapping data and why it matters

2.3 Thinking spatially

2.3.1 Explaining spatial variation

2.3.2 Spatial interpolation and small area estimation

2.4. Thinking spatially and temporally

2.4.1 Explaining space-time variation

2.4.2 Estimating parameters for spatial-temporal units.

2.5 Concluding remarks

2.6 Exercises

2.7 Appendix: Geographic Information Systems

Chapter 3. The nature of spatial and spatial-temporal attribute data

3.1 Introduction.

3.2 Data collection processes in the social sciences

3.2.1 Natural experiments

3.2.2 Quasi-experiments

3.2.3 Non-experimental observational studies

3.3 Spatial and spatial-temporal data: properties

3.3.1 From geographical reality to the spatial database

3.3.2 Fundamental properties of spatial and spatial-temporal data.

3.3.2.1 Spatial and temporal dependence.

3.3.2.2 Spatial and temporal heterogeneity

3.3.3 Properties induced by representational choices.

3.3.4 Properties induced by measurement processes.

3.4 Concluding remarks

3.5 Exercises

Chapter 4. Specifying spatial relationships on the map: the weights matrix

4.1 Introduction

4.2 Specifying weights based on contiguity

4.3 Specifying weights based on geographical distance.

4.4 Specifying weights based on the graph structure associated with a set of points.

4.5 Specifying weights based on attribute values.

4.6 Specifying weights based on evidence about interactions.

4.7 Row standardisation.

4.8 Higher order weights matrices

4.9 Choice of *W* and statistical implications

- **4.9.1 Implications for small area estimation**
- **4.9.2 Implications for spatial econometric modelling**
- **4.9.3 Implications for estimating the effects of observable covariates on the outcome**

4.10 Estimating the *W* matrix

4.11 Concluding remarks

4.12 Exercises

4.13 Appendices

- **Appendix 4.13.1: Building a geodatabase in R**
- **Appendix 4.13.2: Constructing the W matrix and accessing data stored in a shapefile**

Chapter 5. Introduction to the Bayesian approach to regression modelling with spatial and spatial-temporal data

5.1 Introduction

5.2 Introducing Bayesian analysis

- **5.2.1 Prior, likelihood and posterior: what do these terms refer to?**
- **5.2.2 Example: modelling high-intensity crime areas**

5.3 Bayesian computation

- **5.3.1 Summarizing the posterior distribution**
- **5.3.2 Integration and Monte Carlo integration**
- **5.3.3 Markov chain Monte Carlo with Gibbs sampling**
- **5.3.4 Introduction to WinBUGS**
- **5.3.5 Practical considerations when fitting models in WinBUGS**
  - **5.3.5.1 Setting the initial values**
  - **5.3.5.2 Checking convergence**
  - **5.3.5.3 Checking efficiency**

5.4 Bayesian regression models

- **5.4.1 Example I: modelling household-level income**
- **5.4.2 Example II: modelling annual burglary rates in small areas**

5.5 Bayesian model comparison and model evaluation

5.6 Prior specifications

5.6.1 When we have little prior information

5.6.2 Towards more informative priors for spatial and spatial-temporal data

5.7 Concluding remarks

5.8 Exercises

Part II Modelling spatial data

Chapter 6. Exploratory analysis of spatial data.

6.1 Introduction

6.2 Techniques for the exploratory analysis of univariate spatial data

6.2.1 Mapping

6.2.2 Checking for spatial trend

6.2.3 Checking for spatial heterogeneity in the mean

6.2.3.1 Count data

6.2.3.2 A Monte Carlo test

6.2.3.3 Continuous-valued data

6.2.4 Checking for global spatial dependence (spatial autocorrelation)

6.2.4.1 The Moran scatterplot

6.2.4.2 The global Moran's *I* statistic

6.2.4.3 Other test statistics for assessing global spatial autocorrelation

6.2.4.4 The join-count test for categorical data

6.2.4.5 The global Moran's I applied to regression residuals

6.2.5 Checking for spatial heterogeneity in the spatial dependence structure: detecting local spatial clusters

6.2.5.1 The Local Moran's *I*

6.2.5.2 The multiple testing problem when using local Moran's *I*

6.2.5.3 Kulldorff's spatial scan statistic

6.3 Exploring relationships between variables:

6.3.1 Scatterplots and the bivariate Moran scatterplot

6.3.2 Quantifying bivariate association

6.3.2.1 The Clifford-Richardson test of bivariate correlation in the presence of spatial autocorrelation

6.3.2.2 Testing for association "at a distance" and the global bivariate Moran's *I*

6.3.3 Checking for spatial heterogeneity in the outcome-covariate relationship: Geographically weighted regression (GWR)

6.4 Overdispersion and zero-inflation in spatial count data

6.4.1 Testing for overdispersion

6.4.2 Testing for zero-inflation

6.5 Concluding remarks

6.6 Exercises

6.7 Appendix: An R function to perform the zero-inflation test by van den Broek (1995).

Chapter 7 Bayesian models for spatial data I: Non-hierarchical and exchangeable hierarchical models.

7.1 Introduction

7.2 Estimating small area income: a motivating example and different modelling strategies

7.2.1 Modelling the 109 parameters non-hierarchically

7.2.2 Modelling the 109 parameters hierarchically

7.3 Modelling the Newcastle income data using non-hierarchical models

7.3.1 An identical parameter model based on Strategy 1

7.3.2 An independent parameters model based on Strategy 2

7.4 An exchangeable hierarchical model based on Strategy 3

7.4.1 The logic of information borrowing and shrinkage

7.4.2 Explaining the nature of global smoothing due to exchangeability

7.4.3 The variance partition coefficient (VPC)

7.4.4 Applying an exchangeable hierarchical model to the Newcastle income data

7.5 Concluding remarks

7.6 Exercises

7.7 Appendix: Obtaining the simulated household income data

Chapter 8 Bayesian models for spatial data II: hierarchical models with spatial dependence.

8.1 Introduction

8.2 The intrinsic conditional autoregressive (ICAR) model

8.2.1 The ICAR model using a spatial weights matrix with binary entries

8.2.1.1 The WinBUGS implementation of the ICAR model

8.2.1.2 Applying the ICAR model using spatial contiguity to the Newcastle income data

8.2.1.3 Results

8.2.1.4 A summary of the properties of the ICAR model using a binary spatial weights matrix

8.2.2 The ICAR model with a general weights matrix

8.2.2.1 Expressing the ICAR model as a joint distribution and the implied restriction on W.

8.2.2.2 The sum-to-zero constraint

8.2.2.3 Applying the ICAR model using general weights to the Newcastle income data

8.2.2.4 Results

8.3 The proper CAR (pCAR) model

8.3.1 Prior choice for ρ

8.3.2 ICAR or pCAR?

8.3.3 Applying the pCAR model to the Newcastle income data

8.3.4 Results

8.4 Locally adaptive models

8.4.1 Choosing an optimal W matrix from all possible specifications

8.4.2 Modelling the elements of the W matrix

8.4.3 Applying some of the locally adaptive spatial models to a subset of the Newcastle income data

8.5 The Besag, York and Mollié (BYM) model

8.5.1 Two remarks on applying the BYM model in practice

8.5.2 Applying the BYM model to the Newcastle income data

8.6 Comparing the fits of different Bayesian spatial models

8.6.1 DIC comparison

8.6.2 Model comparison based on the quality of the MSOA-level average income estimates

8.7 Concluding remarks

8.8 Exercises

Chapter 9. Bayesian hierarchical models for spatial data: applications

9.1 Introduction

9.2 Application 1: Modelling the distribution of high intensity crime areas in a city

9.2.1 Background

9.2.2 Data and exploratory analysis

9.2.3 Methods discussed in Haining and Law (2007) to combine the PHIA and EHIA maps

9.2.4 A joint analysis of the PHIA and EHIA data using the MVCAR model

9.2.5 Results

9.2.6 Another specification of the MVCAR model and a limitation of the MVCAR approach

9.2.7 Conclusion and discussion

9.3 Application 2: Modelling the association between air pollution and stroke mortality

9.3.1 Background and data

9.3.2 Modelling

9.3.3 Interpreting the statistical results

9.3.4 Conclusion and discussion

9.4 Application 3: Modelling the village-level incidence of malaria in a small region of India

9.4.1 Background

9.4.2 Data and exploratory analysis

9.4.3 Model I: A Poisson regression model with random effects

9.4.4 Model II: A two-component Poisson mixture model

9.4.5 Model III: A two-component Poisson mixture model with zero-inflation

9.4.6 Results

9.4.7 Conclusion and model extensions

9.5 Application 4: Modelling the small area count of cases of rape in Stockholm, Sweden.

9.5.1 Background and data

9.5.2 Modelling

9.5.2.1 A "whole map" analysis using Poisson regression

9.5.2.2 A "localised" analysis using Bayesian profile regression

9.5.3 Results

9.5.3.1 "Whole map" associations for the risk factors

9.5.3.2 "Local" associations for the risk factors

9.5.4 Conclusions

9.6 Exercises

Chapter 10. Spatial econometric models

10.1 Introduction

10.2 Spatial econometric models

10.2.1 Three forms of spatial spillover

10.2.2 The spatial lag model (SLM)

10.2.2.1 Formulating the model

10.2.2.2 An example of the SLM

10.2.2.3 The reduced form of the SLM and the constraint on δ

10.2.2.4 Specification of the spatial weights matrix

10.2.2.5 Issues with model fitting and interpreting coefficients

10.2.3 The spatially lagged covariates model (SLX)

10.2.3.1 Formulating the model

10.2.3.2 An example of the SLX model

10.2.4 The spatial error model (SEM)

10.2.5 The spatial Durbin model (SDM)

10.2.5.1 Formulating the model

10.2.5.2 Relating the SDM model to the other three spatial econometric models
10.2.6 Prior specifications
10.2.7 An example: modelling cigarette sales in 46 US states
10.2.7.1 Data description, exploratory analysis and model specifications
10.2.7.2 Results
10.3 Interpreting covariate effects
10.3.1 Definitions of the direct, indirect and total effects of a covariate
10.3.2 Measuring direct and indirect effects without the SAR structure on the outcome variables
10.3.2.1 For the LM and SEM models
10.3.2.2 For the SLX model
10.3.3 Measuring direct and indirect effects when the outcome variables are modelled by the SAR structure
10.3.3.1 Understanding direct and indirect effects in the presence of spatial feedback
10.3.3.2 Calculating the direct and indirect effects in the presence of spatial feedback
10.3.3.3 Some properties of direct and indirect effects
10.3.3.4 A property (limitation) of the average direct and average indirect effects under the SLM model
10.3.3.5 Summary
10.3.4 The estimated effects from the cigarette sales data
10.4 Model fitting in WinBUGS
10.4.1 Derivation of the likelihood function
10.4.2 Simplifications to the likelihood function
10.4.3 The zeros-trick in WinBUGS
10.4.4 Calculating the covariate effects in WinBUGS
10.5 Concluding remarks
10.5.1 Other spatial econometric models and two problems of identifiability
10.5.2 Comparing the hierarchical modelling approach and the spatial econometric modelling approach: a summary
10.6 Exercises

Chapter 11 Spatial Econometric Modelling: applications
11.1 Application 1: Modelling the voting outcomes at the local authority district level in England from the 2016 EU referendum
11.1.1 Introduction
11.1.2 Data

11.1.3 Exploratory data analysis

11.1.4 Modelling using spatial econometric models

11.1.5 Results

11.1.6 Conclusion and discussion

11.2 Application 2: Modelling price competition between petrol retail outlets in a large city

11.2.1 Introduction

11.2.2 Data

11.2.3 Exploratory data analysis

11.2.4 Spatial econometric modelling and results

11.2.5 A spatial hierarchical model with t_4 likelihood

11.2.6 Conclusion and discussion

11.3 Final remarks on spatial econometric modelling of spatial data

11.4 Exercises

11.5 Appendix: Petrol price data

Part III Modelling spatial-temporal data

Chapter 12 Modelling spatial-temporal data: an introduction

12.1 Introduction

12.2 Modelling annual counts of burglary cases at the small area level: a motivating example and frameworks for modelling spatial-temporal data

12.3 Modelling small area temporal data

12.3.1 Issues to consider when modelling temporal patterns in the small area setting

12.3.1.1 Issues relating to temporal dependence

12.3.1.2 Issues relating to temporal heterogeneity and spatial heterogeneity in modelling small area temporal patterns

12.3.1.3 Issues relating to flexibility of a temporal model

12.3.2 Modelling small area temporal patterns: setting the scene

12.3.3 A linear time trend model

12.3.3.1 Model formulations

12.3.3.2 Modelling trends in the Peterborough burglary data

12.3.3.2.1 Results from fitting the linear trend model without temporal noise

12.3.3.2.2 Results from fitting the linear trend model with temporal noise

12.3.4 Random walk models

12.3.4.1 Model formulations

12.3.4.2 The RW(1) model: its formulation via the full conditionals and

its properties

12.3.4.3 WinBUGS implementation of the RW(1) model

12.3.4.4 Example: modelling burglary trends using the Peterborough data

12.3.4.5 The random walk model of order 2

12.3.5 Interrupted time series (ITS) models

12.3.5.1 Quasi-experimental designs and the purpose of ITS modelling

12.3.5.2 Model formulations

12.3.5.3 WinBUGS implementation

12.3.5.4 Results

12.4 Concluding remarks

12.5 Exercises

Appendix Three different forms for specifying the impact function, f.

Chapter 13 Exploratory analysis of spatial-temporal data

13.1 Introduction

13.2 Patterns of spatial-temporal data

13.3 Visualizing spatial-temporal data

13.4 Tests of space-time interaction

13.4.1 The Knox test

13.4.1.1 An instructive example of the Knox test and different methods to derive a p-value.

13.4.1.2 Applying the Knox test to the malaria data

13.4.2 Kulldorff's space-time scan statistic

13.4.2.1 Application: the simulated small area COPD mortality data

13.4.3 Assessing space-time interaction in the form of varying local time trend patterns

13.4.3.1 Exploratory analysis of the local trends in the Peterborough burglary data

13.4.3.2 Exploratory analysis of the local time trends in the England COPD mortality data

13.5 Concluding remarks

13.6 Exercises

Chapter 14 Bayesian hierarchical models for spatial-temporal data I: space-time separable models

14.1 Introduction

14.2 Estimating small area burglary rates over time: setting the scene

14.3 The space-time separable modelling framework

14.3.1 Model formulations

14.3.2 Do we combine the space and time components additively or multiplicatively?
14.3.3 Analysing the Peterborough burglary data using a space-time separable model
14.3.4 Results
14.4 Concluding remarks
14.5 Exercises

Chapter 15 Bayesian hierarchical models for spatial-temporal data II: space-time inseparable models
15.1 Introduction
15.2 From space-time separability to space-time inseparability: the big picture
15.3 Type I space-time interaction
15.3.1 Example: a space-time model with Type I space-time interaction
15.3.2 WinBUGS implementation
15.4 Type II space-time interaction
15.4.1 Example: two space-time models with Type II space-time interaction
15.4.2 WinBUGS implementation
15.5 Type III space-time interaction
15.5.1 Example: a space-time model with Type III space-time interaction
15.5.2 WinBUGS implementation
15.6 Results from analysing the Peterborough burglary data: Part I
15.7 Type IV space-time interaction
15.7.1 Strategy 1: extending Type II to Type IV
15.7.2 Strategy 2: extending Type III to Type IV
15.7.2.1 Examples of strategy 2
15.7.3 Strategy 3: Clayton's rule
15.7.3.1 Structure matrices and Gaussian Markov random fields
15.7.3.2 Taking the Kronecker product
15.7.3.3 Exploring the induced space-time dependence structure via the full conditionals
15.7.4 Summary on Type IV space-time interaction
15.8 Concluding remarks
15.9 Exercises

Chapter 16 Modelling spatial-temporal data: applications
16.1 Introduction
16.2 Application 1: evaluating a targeted crime reduction intervention
16.2.1 Background and data
16.2.2 Constructing different control groups

16.2.3 Evaluation using ITS

16.2.4 WinBUGS implementation

16.2.5 Results

16.2.6 Some remarks

16.3 Application 2: assessing the stability of risk in space and time

16.3.1 Studying the temporal dynamics of crime hotspots and coldspots: background, data and the modelling idea.

16.3.2 Model formulations

16.3.3 Classification of areas

16.3.4 Model implementation and area classification

16.3.5 Interpreting the statistical results

16.4 Application 3: detecting unusual local time patterns in small area data

16.4.1 Small area disease surveillance: background and modelling idea

16.4.2 Model formulation

16.4.3 Detecting unusual areas with a control of the false discovery rate

16.4.4 Fitting BaySTDetect in WinBUGS

16.4.5 A simulated dataset to illustrate the use of BaySTDetect

16.4.6 Results from the simulated dataset

16.4.7 General results from Li et al. (2012) and an extension of BaySTDetect

16.5 Application 4: Investigating the presence of spatial-temporal spillover effects on village-level malaria risk in Kalaburagi, Karnataka, India

16.5.1 Background and study objective

16.5.2 Data

16.5.3 Modelling

16.5.4 Results

16.5.5 Concluding remarks

16.6 Conclusions

16.7 Exercises

Part IV Directions in spatial and spatial-temporal data analysis

Chapter 17 Modelling spatial and spatial-temporal data: future agendas?

17.1 Topic 1: Modelling multiple related outcomes over space and time

17.2 Topic 2: Joint modelling of georeferenced longitudinal and time-to-event data

17.3 Topic 3: Multiscale modelling

17.4 Topic 4: Using survey data for small area estimation

17.5 Topic 5: Combining data at both aggregate and individual levels to improve ecological inference

17.6 Topic 6: Geostatistical modelling

17.6.1 Spatial dependence

17.6.2 Mapping to reduce visual bias

17.6.3 Modelling scale effects

17.7 Topic 7: Modelling count data in spatial econometrics

17.8 Topic 8: Computation