

CDD 577.55 Ric

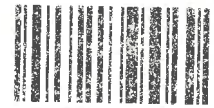
AGR 295

Spatial Data Analysis in Ecology and Agriculture Using R

Second Edition

④

070094



070094



Richard E. Plant

Departments of Plant Sciences and Biological and Agricultural Engineering
University of California, Davis



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an Informa business

Contents

Preface to the First Edition.....	xiii
Preface to the Second Edition.....	xv
Author.....	xvii
1. Working with Spatial Data.....	1
1.1 Introduction.....	1
1.2 Analysis of Spatial Data.....	5
1.2.1 Types of Spatial Data.....	5
1.2.2 The Components of Spatial Data.....	7
1.2.3 Spatial Data Models.....	7
1.2.4 Topics Covered in the Text.....	8
1.3 The Data Sets Analyzed in This Book.....	9
1.3.1 Data Set 1: Yellow-Billed Cuckoo Habitat.....	10
1.3.2 Data Set 2: Environmental Characteristics of Oak Woodlands.....	12
1.3.3 Data Set 3: Uruguayan Rice Farmers.....	13
1.3.4 Data Set 4: Factors Underlying Yield in Two Fields.....	14
1.3.5 Comparing the Data Sets.....	15
1.4 Further Reading.....	16
2. The R Programming Environment.....	19
2.1 Introduction.....	19
2.1.1 Introduction to R.....	19
2.1.2 Setting Yourself Up to Use This Book.....	20
2.2 R Basics.....	21
2.3 Programming Concepts.....	26
2.3.1 Looping and Branching.....	26
2.3.2 Functional Programming.....	28
2.4 Handling Data in R.....	30
2.4.1 Data Structures in R.....	30
2.4.2 Basic Data Input and Output.....	33
2.4.3 Spatial Data Structures.....	34
2.5 Writing Functions in R.....	43
2.6 Graphics in R.....	46
2.6.1 Traditional Graphics in R: Attribute Data.....	47
2.6.2 Traditional Graphics in R: Spatial Data.....	52
2.6.3 Trellis Graphics in R, Attribute Data.....	55
2.6.4 Trellis Graphics in R, Spatial Data.....	57
2.6.5 Using Color in R.....	59
2.7 Continuing on from Here with R.....	61
2.8 Further Reading.....	62
Exercises.....	62

3. Statistical Properties of Spatially Autocorrelated Data	67
3.1 Introduction	67
3.2 Components of a Spatial Random Process	68
3.2.1 Spatial Trends in Data	68
3.2.2 Stationarity	74
3.3 Monte Carlo Simulation	77
3.4 A Review of Hypothesis and Significance Testing	79
3.5 Modeling Spatial Autocorrelation	84
3.5.1 Monte Carlo Simulation of Time Series	84
3.5.2 Modeling Spatial Contiguity	88
3.5.3 Modeling Spatial Association in R	93
3.6 Application to Field Data	98
3.6.1 Setting Up the Data	98
3.6.2 Checking Sequence Validity	101
3.6.3 Determining Spatial Autocorrelation	102
3.7 Further Reading	103
Exercises	103
4. Measures of Spatial Autocorrelation	105
4.1 Introduction	105
4.2 Preliminary Considerations	105
4.2.1 Measurement Scale	105
4.2.2 Resampling and Randomization Assumptions	108
4.2.3 Testing the Null Hypothesis	109
4.3 Join-Count Statistics	110
4.4 Moran's I and Geary's c	114
4.5 Measures of Autocorrelation Structure	117
4.5.1 The Moran Correlogram	117
4.5.2 The Moran Scatterplot	119
4.5.3 Local Measures of Autocorrelation	121
4.5.4 Geographically Weighted Regression	124
4.6 Measuring Autocorrelation of Spatially Continuous Data	127
4.6.1 The Variogram	127
4.6.2 The Covariogram and the Correlogram	132
4.7 Further Reading	133
Exercises	133
5. Sampling and Data Collection	135
5.1 Introduction	135
5.2 Preliminary Considerations	138
5.2.1 The Artificial Population	138
5.2.2 Accuracy, Bias, Precision, and Variance	141
5.2.3 Comparison Procedures	142
5.3 Developing the Sampling Patterns	142
5.3.1 Random Sampling	142
5.3.2 Geographically Stratified Sampling	144
5.3.3 Sampling on a Regular Grid	146
5.3.4 Stratification Based on a Covariate	148
5.3.5 Cluster Sampling	153

5.4	Methods for Variogram Estimation.....	154
5.5	Estimating the Sample Size	157
5.6	Sampling for Thematic Mapping.....	158
5.7	Design-Based and Model-Based Sampling	159
5.8	Further Reading	164
	Exercises	164
6.	Preparing Spatial Data for Analysis.....	167
6.1	Introduction	167
6.2	Quality of Attribute Data.....	168
6.2.1	Dealing with Outliers and Contaminants	168
6.2.2	Quality of Ecological Survey Data	170
6.2.3	Quality of Automatically Recorded Data.....	170
6.3	Spatial Interpolation Procedures	175
6.3.1	Inverse Weighted Distance Interpolation.....	175
6.3.2	Kriging Interpolation	180
6.3.3	Cokriging Interpolation.....	183
6.4	Spatial Rectification and Alignment of Data	188
6.4.1	Definitions of Scale Related Processes.....	188
6.4.2	Change of Coverage.....	190
6.4.3	Change of Support.....	193
6.5	Further Reading	196
	Exercises	197
7.	Preliminary Exploration of Spatial Data.....	199
7.1	Introduction	199
7.2	Data Set 1.....	201
7.3	Data Set 2.....	214
7.4	Data Set 3.....	229
7.5	Data Set 4.....	240
7.6	Further Reading	251
	Exercises	251
8.	Data Exploration Using Non-Spatial Methods: The Linear Model.....	255
8.1	Introduction	255
8.2	Multiple Linear Regression	255
8.2.1	The Many Perils of Model Selection	255
8.2.2	Multicollinearity, Added Variable Plots, and Partial Residual Plots	261
8.2.3	A Cautious Approach Model Selection as an Exploratory Tool.....	269
8.3	Building a Multiple Regression Model for Field 4.1.....	270
8.4	Generalized Linear Models	281
8.4.1	Introduction to Generalized Linear Models.....	281
8.4.2	Multiple Logistic Regression Model for Data Set 2.....	288
8.4.3	Logistic Regression Model of Count Data for Data Set 1.....	295
8.4.4	Analysis of the Counts of Data Set 1: Zero-Inflated Poisson Data	299
8.5	Further Reading	303
	Exercises	304

9. Data Exploration Using Non-Spatial Methods: Nonparametric Methods	307
9.1 Introduction	307
9.2 The Generalized Additive Model	307
9.3 Classification and Regression Trees (a.k.a. Recursive Partitioning)	317
9.3.1 Introduction to the Method	317
9.3.2 The Mathematics of Recursive Partitioning	320
9.3.3 Exploratory Analysis of Data Set 2 with Regression Trees	321
9.3.4 Exploratory Analysis of Data Set 3 with Recursive Partitioning	328
9.3.5 Exploratory Analysis of Field 4.1 with Recursive Partitioning	334
9.4 Random Forest	339
9.4.1 Introduction to Random Forest	339
9.4.2 Application to Data Set 2	342
9.5 Further Reading	345
Exercises	345
10. Variance Estimation, the Effective Sample Size, and the Bootstrap	347
10.1 Introduction	347
10.2 Bootstrap Estimation of the Standard Error	351
10.3 Bootstrapping Time Series Data	355
10.3.1 The Problem with Correlated Data	355
10.3.2 The Block Bootstrap	357
10.3.3 The Parametric Bootstrap	360
10.4 Bootstrapping Spatial Data	362
10.4.1 The Spatial Block Bootstrap	362
10.4.2 The Parametric Spatial Bootstrap	366
10.4.3 Power of the Tests	368
10.5 Application to the EM38 Data	368
10.6 Further Reading	371
Exercises	372
11. Measures of Bivariate Association between Two Spatial Variables	373
11.1 Introduction	373
11.2 Estimating and Testing the Correlation Coefficient	376
11.2.1 The Correlation Coefficient	376
11.2.2 The Clifford et al. (1989) Correction	378
11.2.3 The Bootstrap Variance Estimate	381
11.2.4 Application to the Example Problem	383
11.3 Contingency Tables	386
11.3.1 Large Sample Size Contingency Tables	386
11.3.2 Small Sample Size Contingency Tables	393
11.4 The Mantel and Partial Mantel Statistics	398
11.4.1 The Mantel Statistic	398
11.4.2 The Partial Mantel Test	401
11.5 The Modifiable Areal Unit Problem and the Ecological Fallacy	404
11.5.1 The Modifiable Areal Unit Problem	404
11.5.2 The Ecological Fallacy	408
11.6 Further Reading	410
Exercises	410

12. The Mixed Model	413
12.1 Introduction	413
12.2 Basic Properties of the Mixed Model	417
12.3 Application to Data Set 3.....	419
12.4 Incorporating Spatial Autocorrelation.....	422
12.5 Generalized Least Squares	429
12.6 Spatial Logistic Regression.....	431
12.6.1 Upscaling Data Set 2 in the Coast Range	431
12.6.2 The Incorporation of Spatial Autocorrelation.....	436
12.7 Further Reading	443
Exercises	444
13. Regression Models for Spatially Autocorrelated Data	445
13.1 Introduction	445
13.2 Detecting Spatial Autocorrelation in a Regression Model.....	450
13.3 Models for Spatial Processes	452
13.3.1 The Spatial Lag Model.....	452
13.3.2 The Spatial Error Model.....	454
13.4 Determining the Appropriate Regression Model	455
13.4.1 Formulation of the Problem.....	455
13.4.2 The Lagrange Multiplier Test	456
13.5 Fitting the Spatial Lag and Spatial Error Models.....	458
13.6 The Conditional Autoregressive Model.....	460
13.7 Application of Simultaneous Autoregressive and Conditional Autoregressive Models to Field Data	462
13.7.1 Fitting the Data	462
13.7.2 Comparison of the Mixed Model and Spatial Autoregression.....	465
13.8 Further Reading	466
Exercises	466
14. Bayesian Analysis of Spatially Autocorrelated Data	467
14.1 Introduction	467
14.2 Markov Chain Monte Carlo Methods	471
14.3 Introduction to WinBUGS.....	478
14.3.1 WinBUGS Basics.....	478
14.3.2 WinBUGS Diagnostics.....	481
14.3.3 Introduction to R2WinBUGS	483
14.3.4 Generalized Linear Models in WinBUGS.....	490
14.4 Hierarchical Models	492
14.5 Incorporation of Spatial Effects.....	498
14.5.1 Spatial Effects in the Linear Model	498
14.5.2 Application to Data Set 3.....	501
14.5.3 The spBayes Package	505
14.6 Comparison of the Methods.....	509
14.7 Further Reading	510
Exercises	511

15. Analysis of Spatiotemporal Data	513
15.1 Introduction	513
15.2 Spatiotemporal Data Interpolation	513
15.2.1 Representing Spatiotemporal Data	513
15.2.2 The Spatiotemporal Variogram	518
15.2.3 Interpolating Spatiotemporal Data	523
15.3 Spatiotemporal Process Models	525
15.3.1 Models for Dispersing Populations	525
15.3.2 A Process Model for the Yield Data	526
15.4 Finite State and Time Models	529
15.4.1 Determining Finite State and Time Models Using Clustering	529
15.4.2 Factors Underlying Finite State and Time Models	538
15.5 Bayesian Spatiotemporal Analysis	543
15.5.1 Introduction to Bayesian Updating	543
15.5.2 Application of Bayesian Updating to Data Set 3	546
15.6 Further Reading	550
Exercises	551
16. Analysis of Data from Controlled Experiments	553
16.1 Introduction	553
16.2 Classical Analysis of Variance	554
16.3 The Comparison of Methods	559
16.3.1 The Comparison Statistics	559
16.3.2 The Papadakis Nearest-Neighbor Method	561
16.3.3 The Trend Method	562
16.3.4 The "Correlated Errors" Method	563
16.3.5 Published Comparisons of the Methods	565
16.4 Pseudoreplicated Data and the Effective Sample Size	566
16.4.1 Pseudoreplicated Comparisons	566
16.4.2 Calculation of the Effective Sample Size	567
16.4.3 Application to Field Data	569
16.5 Further Reading	571
Exercises	572
17. Assembling Conclusions	573
17.1 Introduction	573
17.2 Data Set 1	573
17.3 Data Set 2	578
17.4 Data Set 3	583
17.5 Data Set 4	586
17.6 Conclusions	590
Appendix A: Review of Mathematical Concepts	593
Appendix B: The Data Sets	619
Appendix C: An R Thesaurus	627
References	635
Index	657