
Contents

List of Tables	xiii
Preface	xv
Author	xvii
I Background	1
1 Introduction	3
1.1 Datasets	5
2 Bayesian Inference and Modeling	19
2.1 Likelihood Models	19
2.1.1 Spatial Correlation	20
2.2 Prior Distributions	22
2.2.1 Propriety	22
2.2.2 Noninformative Priors	22
2.3 Posterior Distributions	24
2.3.1 Conjugacy	25
2.3.2 Prior Choice	25
2.4 Predictive Distributions	26
2.4.1 Poisson–Gamma Example	26
2.5 Bayesian Hierarchical Modeling	27
2.6 Hierarchical Models	27
2.7 Posterior Inference	28
2.7.1 A Bernoulli and Binomial Example	30
2.8 Exercises	34
3 Computational Issues	35
3.1 Posterior Sampling	35
3.2 Markov Chain Monte Carlo Methods	36
3.3 Metropolis and Metropolis–Hastings Algorithms	37
3.3.1 Metropolis Updates	38
3.3.2 Metropolis–Hastings Updates	38
3.3.3 Gibbs Updates	38
3.3.4 M–H versus Gibbs Algorithms	39
3.3.5 Special Methods	40

3.3.6	Convergence	40
3.3.7	Subsampling and Thinning	45
3.4	Perfect Sampling	47
3.5	Posterior and Likelihood Approximations	48
3.5.1	Pseudolikelihood and Other Forms	48
3.5.2	Asymptotic Approximations	50
3.6	Exercises	53
4	Residuals and Goodness-of-Fit	55
4.1	Model GOF Measures	55
4.1.1	Deviance Information Criterion	56
4.1.2	Posterior Predictive Loss	57
4.2	General Residuals	59
4.3	Bayesian Residuals	61
4.4	Predictive Residuals and the Bootstrap	62
4.4.1	Conditional Predictive Ordinates	63
4.5	Interpretation of Residuals in a Bayesian Setting	64
4.6	Exceedence Probabilities	64
4.7	Exercises	67
II	Themes	71
5	Disease Map Reconstruction and Relative Risk Estimation	73
5.1	Introduction to Case Event and Count Likelihoods	73
5.1.1	Poisson Process Model	73
5.1.2	Conditional Logistic Model	75
5.1.3	Binomial Model for Count Data	76
5.1.4	Poisson Model for Count Data	76
5.2	Specification of the Predictor in Case Event and Count Models	77
5.2.1	Bayesian Linear Model	79
5.3	Simple Case and Count Data Models with Uncorrelated Random Effects	80
5.3.1	Gamma and Beta Models	82
5.3.2	Log-Normal/Logistic-Normal Models	84
5.4	Correlated Heterogeneity Models	84
5.4.1	Conditional Autoregressive (CAR) Models	86
5.4.2	Fully-Specified Covariance Models	90
5.5	Convolution Models	91
5.6	Model Comparison and Goodness-of-Fit Diagnostics	92
5.6.1	Residual Spatial Autocorrelation	94
5.7	Alternative Risk Models	96
5.7.1	Autologistic Models	96
5.7.2	Spline-Based Models	101
5.7.3	Zip Regression Models	102
5.7.4	Ordered and Unordered Multicategory Data	107

5.7.5	Latent Structure Models	108
5.8	Edge Effects	111
5.8.1	Edge Weighting Schemes and MCMC Methods	113
5.8.2	Discussion and Extension to Space–Time	115
5.9	Exercises	116
5.9.1	Maximum Likelihood	116
5.9.2	Poisson–Gamma Model: Posterior and Predictive Inference	117
5.9.3	Poisson–Gamma Model: Empirical Bayes	117
6	Disease Cluster Detection	119
6.1	Cluster Definitions	119
6.1.1	Hot Spot Clustering	121
6.1.2	Clusters as Objects or Groupings	121
6.1.3	Clusters Defined as Residuals	121
6.2	Cluster Detection using Residuals	122
6.2.1	Case Event Data	122
6.2.2	Count Data	126
6.3	Cluster Detection Using Posterior Measures	130
6.4	Cluster Models	133
6.4.1	Case Event Data	133
6.4.2	Count Data	143
6.4.3	Markov Connected Component Field (MCCF) Models	148
6.5	Edge Detection and Wombling	149
7	Ecological Analysis	151
7.1	General Case of Regression	151
7.2	Biases and Misclassification Error	158
7.2.1	Ecological Biases	158
7.3	Putative Hazard Models	165
7.3.1	Case Event Data	166
7.3.2	Aggregated Count Data	172
7.3.3	Spatiotemporal Effects	176
8	Multiple Scale Analysis	185
8.1	Modifiable Areal Unit Problem (MAUP)	185
8.1.1	Scaling Up	185
8.1.2	Scaling Down	187
8.1.3	Multiscale Analysis	187
8.2	Misaligned Data Problem(MIDP)	190
8.2.1	Predictor Misalignment	191
8.2.2	Outcome Misalignment	198
8.2.3	Misalignment and Edge Effects	200

9	Multivariate Disease Analysis	201
9.1	Notation for Multivariate Analysis	201
9.1.1	Case Event Data	201
9.1.2	Count Data	202
9.2	Two Diseases	202
9.2.1	Case Event Data	202
9.2.2	Count Data	204
9.2.3	Georgia County Level Example (3 Diseases)	206
9.3	Multiple Diseases	207
9.3.1	Case Event Data	209
9.3.2	Count Data	216
9.3.3	Multivariate Spatial Correlation and MCAR Models	219
9.3.4	Georgia Chronic Ambulatory Care-Sensitive Example	222
10	Spatial Survival and Longitudinal Analysis	227
10.1	General Issues	227
10.2	Spatial Survival Analysis	228
10.2.1	Endpoint Distributions	228
10.2.2	Censoring	229
10.2.3	Random Effect Specification	230
10.2.4	General Hazard Model	232
10.2.5	Cox Model	232
10.2.6	Extensions	233
10.3	Spatial Longitudinal Analysis	234
10.3.1	General Model	237
10.3.2	Seizure Data Example	237
10.3.3	Missing Data	241
10.4	Extensions to Repeated Events	243
10.4.1	Simple Repeated Events	243
10.4.2	More Complex Repeated Events	244
10.4.3	Fixed Time Periods	247
11	Spatiotemporal Disease Mapping	255
11.1	Case Event Data	255
11.2	Count Data	257
11.2.1	Georgia Low Birth Weight Example	262
11.3	Alternative Models	266
11.3.1	Autologistic Models	266
11.3.2	Latent Structure ST Models	268
11.4	Infectious Diseases	271
11.4.1	Case Event Data	272
11.4.2	Count Data	273
11.4.3	Special Case: Veterinary Disease Mapping	276

A Basic R and WinBUGS	283
A.1 Basic R Usage	283
A.1.1 Data	283
A.1.2 Graphics	284
A.2 Use of R in Bayesian Modeling	287
A.3 WinBUGS	290
A.3.1 Simulation	291
A.3.2 Model Code	291
A.4 R2WinBUGS Function	298
A.5 BRugs	302
A.6 Maps on R and GeoBUGS	305
B Selected WinBUGS Code	307
B.1 Code for the Convolution Model (Chapter 5)	307
B.2 Code for Spatial Spline Model (Chapter 5)	308
B.3 Code for the Spatial Autologistic Model (Chapter 6)	308
B.4 Code for Logistic Spatial Case Control Model (Chapter 6)	309
B.5 Code for PP Residual Model (Chapter 6)	309
B.5.1 Same Model with Uncorrelated Random Effect	310
B.6 Code for the Logistic Spatial Case-Control Model (Chapter 6)	310
B.7 Code for Poisson Residual Clustering Example (Chapter 6)	312
B.8 Code for the Proper CAR Model (Chapter 7)	312
B.9 Code for the Multiscale Model for PH and County Level Data (Chapter 8)	313
B.10 Code for the Shared Component Model for Georgia Asthma and COPD (Chapter 9)	314
B.11 Code for the Seizure Example with Spatial Effect (Chapter 10)	315
B.12 Code for the Knorr-Held Model for Space-Time Relative Risk Estimation (Chapter 11)	316
B.13 Code for the Space-Time Autologistic Model (Chapter 11)	316
C R Code for Thematic Mapping	319
References	321
Index	339

List of Tables

5.1	Comparison of convolution and uncorrelated heterogeneity models for the Georgia oral cancer dataset	93
5.2	DICs for three models: autologistic with no random effects; autologistic with UH component; convolution model with UH and CH	98
7.1	Model fitting results for a variety of models for the South Carolina congenital mortality data, (See text for details).	154
7.2	Results for a variety of models fitted to 1988 respiratory cancer incident counts for counties of Ohio	175
7.3	Ohio respiratory cancer (1979–1988): putative source model fits	181
8.1	Goodness of fit results for separate and joint models for Georgia oral cancer PH-county level data	190
8.2	Model fit results for the point misalignment example	197
9.1	Model comparisons for the three disease examples: joint, common, and shared	207
9.2	Correlation of spatial random effects under the MCAR model for the Georgia chronic disease example: upper triangle: correlation of the spatially structured effects; lower triangle: correlation of the relative risks	225
10.1	Comparison of four models for the seizure data: basic and Models 1–3	239
10.2	Posterior average estimates of the model parameters under the four different seizure models	240
10.3	Parameter estimates for the basic model compared to the best fitting model (Model 1)	240
11.1	Space–time models for the Georgia oral cancer dataset; models are explained in text.	264
11.2	Autologistic space–time models: models 1–4; convolution model (Model 5)	267