Universidade Federal do Rio de Janeiro

Modelos Dinâmicos Bayesianos para Processos Pontuais Espaço-Temporais

Edna Afonso Reis

Rio de Janeiro

2008

Modelos Dinâmicos Bayesianos para Processos Pontuais Espaço-Temporais

Edna Afonso Reis

Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Estatística.

Orientadores: Dani Gamerman
Marina Silva Paez

Rio de Janeiro
2008

Reis, Edna Afonso

Modelos Dinâmicos Bayesianos para Processos Pontuais Espaço-Temporais – Rio de Janeiro:UFRJ/IM, 2008.

vi, 142f.: il, color.; 31cm.

Orientadores: Dani Gamerman e Marina Silva Paez

Tese (Doutorado em Estatística) – UFRJ/IM/Programa de Pós-graduação em Estatística, 2008.

Referências Bibliográficas: f. 118 – 123.

1. . 2. . 3. . I. Gamerman, Dani e Paez, Marina S. (Orient.). II. Universidade Federal do Rio de Janeiro, Instituto de Matemática. III. Título.

Modelos Dinâmicos Bayesianos para Processos Pontuais Espaço-Temporais

Edna Afonso Reis

Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Estatística.

Presidente, Prof. Dani Gamerman

IM-UFRJ

Prof.^{a} Marina Silva Paez

IM-UFRJ

Prof.^{a} Alexandra Mello Schmidt

IM-UFRJ

Prof.^{a} Nancy Lopes Garcia

IMECC-UNICAMP

Prof. Jorge Alberto Achcar

ICMC-USP

Rio de Janeiro, 08 de maio de 2008.

PARA MEU PAI,

DÁLVIO.

(*in memoriam*)

AGRADECIMENTOS

A autora expressa seus mais sinceros agradecimentos às seguintes pessoas e entidades por sua valiosa contribuição para a realização deste trabalho:

- Meus orientadores Dani e Marina, pela dedicação e paciência;
- Ramiro e Emília, pelo importante apoio na fase final;
- Minha mãe Oraida, irmãs Ilka e Tânia, amigas Esther e Romy, pelo carinho;
- Eduardo, secretário da PPG-IM, pela sua presteza e eficiência;
- Colegas e professores do PPG em Estatística da UFRJ;
- FAPERJ e CAPES, pelo suporte financeiro;
- Departamento de Estatística e Universidade Federal de Minas Gerais, pela licença concedida para realização do curso.

RESUMO

Modelos Dinâmicos Bayesianos para Processos Pontuais Espaço-Temporais

Edna Afonso Reis

Orientadores: Dani Gamerman
Marina Silva Paez

Resumo da Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Estatística.

O estudo de processos pontuais observados no espaço e no tempo tem se tornado uma importante área da Estatística Espacial. Nesta tese, é proposto um modelo espaço-temporal especificado por uma seqüência de superfícies de intensidades espaciais ligadas no tempo através de modelos dinâmicos, resultando nos denominados *processos pontuais espaciais dinâmicos*. A inferência para esses processos é feita sob a abordagem bayesiana, com utilização de métodos MCMC, como o amostrador de Gibbs e o algoritmo de Metropolis-Hastings. Os modelos e métodos de estimação propostos foram intensivamente testados em estudos simulados e aplicados em um conjunto de dados experimentais de impulsos elétricos no intestino delgado de gatos e em um conjunto de dados observacionais dos casos de doenças gastrointestinais no condado de Hampshire, no Reino Unido.

Palavras-chave: processos pontuais espaço-temporais; modelos dinâmicos; inferência bayesiana; MCMC; mapeamento de doenças.

ABSTRACT

Bayesian Dynamic Models for Space-Time Point Processes

Edna Afonso Reis

Advisors: Dani Gamerman
Marina Silva Paez

Abstract of doctoral thesis submited to the Graduate Program in Statistics of the Instituto de Matemática da Universidade Federal do Rio de Janeiro, as required to the Doctor degree in Statistics.

Point processes in time and space has gained an important role in Spatial Statistics. In this thesis, a spatio-temporal model is proposed by specifying a sequence of spatial intensity surfaces linked in time through dynamic models. This is denoted by *dynamic spatial point process*. A Bayesian inference approach was adopted and MCMC methods as Gibbs sampler and Metropolis-Hastings algorithm were used. Models and inference methods were intensively tested through simulated data. These models were applied to an experimental dataset of spikes in the small intestine of cats and to an observational dataset of cases of gastroenteric disease in the county of Hampshire, UK.

Key-words: space-time point processes; dynamic models; Bayesian inference; Monte Carlo Markov chain; disease mapping.

Lista de Figuras

2.1 Os tipos básicos de arranjos pontuais espaciais 20

2.2 Exemplo de construção de uma grade regular 25

2.3 Exemplo de construção da tesselagem de Voronoi 26

3.1 Exemplo 3.1: Mapa dos eventos gerados 36

3.2 Exemplo 3.1: Resultados de estimação dos efeitos espaciais 37

3.3 Exemplo 3.1: Histogramas das amostras a posteriori do coeficiente de regressão e dos hiperparâmetros 38

3.4 Exemplo 3.2: Mapa dos eventos gerados 38

3.5 Exemplo 3.2: Resultados de estimação dos efeitos espaciais 39

3.6 Exemplo 3.2: Histogramas das amostras a posteriori do coeficiente de regressão e dos hiperparâmetros 40

3.7 Exemplo 3.3: Especificações de prioris para σ^2 e θ. 41

3.8 Exemplo 3.3: Histogramas das amostras a posteriori de σ^2 e θ. 42

3.9 Exemplo 3.4: Mapas dos processos gaussianos e eventos gerados 43

3.10 Exemplo 3.4: Resultados de estimação dos efeitos espaciais 43

5.1 Modelo estacionário: valores gerados das log-intensidades 71

5.2 Modelo estacionário: eventos gerados 71

5.3 Modelo não-estacionário: valores gerados ds log-intensidades 72

5.4 Modelo não-estacionário: eventos gerados 72

5.5 Modelo estacionário: histogramas das amostras a posteriori dos hiperparâmetros . . 73

5.6 Modelo estacionário: inferência dos efeitos ϕ 74

5.7 Modelo estacionário: imagens dos valores reais e médias a posteriori das log-intensidades 75
5.8 Modelo não-estacionário: histogramas das amostras a posteriori dos hiperparâmetros 76
5.9 Modelo não-estacionário: inferência dos efeitos ϕ 77
5.10 Modelo não-estacionário: imagens dos valores reais e médias a posteriori das log-intensidades 78
5.11 Modelo com tendência temporal linear: mapas dos efeitos espaciais reais e eventos . 80
5.12 Modelo com tendência temporal linear: resultados de estimação dos hiperparâmetros 81
5.13 Modelo com tendência temporal linear: resultados de estimação dos efeitos espaciais 82
5.14 Modelo com tendência temporal dinâmica polinomial de primeira ordem: mapas das somas dos efeitos espaciais e temporais e da localização dos eventos gerados 84
5.15 Modelo com tendência temporal dinâmica polinomial de primeira ordem: resultados de estimação dos hiperparâmetros 85
5.16 Modelo com tendência temporal dinâmica polinomial de primeira ordem: resultados de estimação dos efeitos espaciais 86
5.17 Modelo com tendência temporal dinâmica polinomial de primeira ordem: resultados de estimação dos efeitos temporais 87
5.18 Modelo com tendência temporal dinâmica polinomial de segunda ordem: mapas das somas dos efeitos espaciais e temporais e da localização dos eventos gerados 89
5.19 Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos hiperparâmetros 90
5.20 Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos efeitos espaciais 91
5.21 Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos efeitos temporais $\mu_{[t]}$ 92
5.22 Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos efeitos temporais $\beta_{[t]}$ 92

6.1 Mapa do contorno do condado de Hampshire e eventos observados em cada ano . . 94
6.2 Totais de casos mensais nos três anos do estudo 95
6.3 Grade regular com 270 células sobreposta à região de estudo 95
6.4 Histogramas das amostra a posteriori dos hiperparâmetros 97
6.5 Mapas das médias a posteriori dos efeitos espaciais 98

6.6 Número de impulsos na grade espacial no intestino de um gato, durante 13 ondas lentas sucessivas 100
6.7 Histogramas das amostras a posteriori dos hiperparâmetros do modelo 1 105
6.8 Histogramas das amostras a posteriori dos hiperparâmetros do modelo 2 106
6.9 Histogramas das amostras a posteriori dos hiperparâmetros do modelo 3 107
6.10 Histogramas das amostras a posteriori dos hiperparâmetros do modelo 3b 108
6.11 Histogramas das amostras a posteriori dos hiperparâmetros do modelo 3c 109
6.12 Médias a posteriori e intervalos de 90% de credibilidade dos efeitos temporais . . . 110
6.13 Médias a posteriori e intervalos de credibilidade de 90% dos efeitos espaço-temporais 111
6.14 Mapas das médias a posteriori dos efeitos espaço-temporais $\phi_{[i,t]}$ dos modelos 1, 2, 3 e 3c, para $t=1,...,7$ 112
6.15 Mapas das médias a posteriori dos efeitos espaço-temporais $\phi_{[i,t]}$ dos modelos 1, 2, 3 e 3c, para $t=8,...,13$ 113
6.16 Mapas de variabilidade dos efeitos espaço-temporais $\phi_{[i,t]}$ do modelo 3, para $t=1,...,13$ 114

Lista de Tabelas

6.1 Médias a posteriori e Intervalo de Credibilidade de 90% para os hiperparâmetros. . . 104

6.2 Resultados dos critérios de seleção de modelos. 104

Sumário

Lista de Figuras **ix**

Lista de Tabelas **xii**

Capítulo 1: Introdução **1**

1.1 Introdução 1

1.2 Inferência Bayesiana 2

1.2.1 O Teorema de Bayes 3

1.2.2 Análise da Distribuição a Posteriori 3

1.2.3 Escolha da Distribuição a Priori 4

1.3 Métodos MCMC na Inferência Bayesiana 5

1.3.1 Amostrador de Gibbs 5

1.3.2 Algoritmo de Metropolis-Hastings 6

1.3.3 Avaliação da Convergência da Cadeia 7

1.4 Modelos Dinâmicos 7

1.4.1 Modelos Dinâmicos Lineares 8

1.4.2 Modelos Dinâmicos Lineares Generalizados 9

1.5 Seleção de Modelos 10

1.5.1 DIC - *Deviance Information Criterion* 11

1.5.2 EPD - *Expected Predictive Deviance* 12

1.6 Organização da Tese 13

Capítulo 2: Processos Espaciais 14

2.1 Introdução 14

2.2 Processos Espaciais 15

2.3 O Processo Gaussiano 16

2.3.1 Simulação de Dados de Processos Gaussianos 16

2.3.2 Famílias de Funções de Correlação Espaciais 17

2.4 Processos Espaciais Pontuais 18

2.4.1 Tipos de Arranjos Pontuais 19

2.4.2 Alguns Modelos para Processos Espaciais Pontuais 20

2.4.3 Simulação de Dados de Processos Espaciais Pontuais 22

2.5 Processos Pontuais Espaço-Temporais 23

2.6 Inferência via Discretização no Espaço e/ou Tempo 24

Capítulo 3: Modelos para Processos Pontuais Espaciais 27

3.1 Introdução 27

3.2 Modelo Espacial 28

3.3 Aspectos Computacionais da Inferência 29

3.3.1 Amostragem dos Efeitos Espaciais 31

3.3.2 Amostragem do Coeficiente de Regressão 33

3.3.3 Amostragem dos Parâmetros do Processo Espacial 34

3.4 Estudos de Simulação 35

3.5 Prioris de Referência 40

3.6 Efeito da Discretização no Espaço 41

Capítulo 4: Modelos para Processos Pontuais Espaço-Temporais 44

4.1 Introdução 44

4.2 Modelos Espaço-Temporais 46

4.2.1 Modelos para a Tendência Temporal 47

4.2.2 Modelos para os Efeitos Espaciais 48

4.2.3 Modelos para os Efeitos Espaço-Temporais 48

4.3 Aspectos Computacionais da Inferência 49

4.3.1 Modelo de Tendência Constante 50

4.3.2 Modelo de Tendência Determinística Linear 56
4.3.3 Modelo de Tendência Dinâmica Polinomial de Primeira Ordem 59
4.3.4 Modelo de Tendência Dinâmica Polinomial de Segunda Ordem 62
4.4 Sumário 67

Capítulo 5: Estudos de Simulação **68**
5.1 Introdução 68
5.2 Tendência Temporal Constante 69
5.3 Tendência Temporal Determinística Linear 79
5.4 Tendência Temporal Dinâmica Polinomial de Primeira Ordem 83
5.5 Tendência Temporal Dinâmica Polinomial de Segunda Ordem 87
5.6 Conclusões 88

Capítulo 6: Aplicações **93**
6.1 Introdução 93
6.2 Análise Espaço-Temporal dos Casos de Doença Gastrointestinal em Hampshire 93
6.3 Evolução Espaço-Temporal de Impulsos Elétricos no Intestino Delgado 99

Capítulo 7: Considerações Finais e Trabalhos Futuros **115**
7.1 Considerações Finais 115
7.2 Trabalhos Futuros 116
7.2.1 Eficiência Computacional do Processo de Inferência 116
7.2.2 Análise de Resíduos 117

Referências **117**

Apêndice A: **124**
A.1 O Filtro de Kalman 124
A.2 O Algoritmo FFBS 125
A.3 Algoritmo de Gamerman (1997) 126

Capítulo 1

Introdução

1.1 Introdução

Uma importante área da Estatística, conhecida como *processos pontuais espaciais*, é o estudo de processos de observação de eventos em uma dada região geográfica. Esta área tem sido estudada tanto do ponto de vista teórico, onde as propriedades probabilísticas desses processos são analisadas (Cox e Isham, 1980), quanto pelo estudo de propriedades estatísticas, onde a ênfase se dá no processo de estimação da taxa de intensidade dos eventos na região (Diggle, 2003).

Um exemplo de observação nessa área é o estudo dos locais de moradia de pessoas acometidas de uma particular doença contagiosa. Esse estudo serve para determinar possíveis padrões de distribuição geográfica do risco de contaminação. Existem vários estudos já realizados nessa área, tanto sob o ponto de vista bayesiano quanto sob o ponto de vista freqüentista, em diversos campos de aplicação, como epidemiologia (Diggle, 2000), criminologia (Liu e Brown, 2003), geologia (Ogata, 1998), dentre outros.

Uma extensão relevante do problema consiste em considerar também a dimensão temporal. Nesse caso, não só o local de ocorrência é registrado, mas também o momento. Para o exemplo de mapeamento de doenças descrito acima, esses processos têm a grande utilidade de permitir a caracterização do processo de espalhamento do risco de contaminação. Com isso, é possível estabelecer uma estratégia de controle da dispersão da doença na região, bem como implantar um sistema de alarme para detecção de novos focos ou de previsão do padrão espacial da doença em tempos futuros.

Dentro desse enfoque, uma possível estratégia é a de especificar uma seqüência de taxas de intensidades do processo no espaço ligadas através do tempo. A proposta desta tese é caracterizar de forma não-paramétrica a seqüência de taxas de intensidade do processo com vistas à definição de formas apropriadas de estimação e previsão do processo. O objetivo é a formulação de modelos levando em consideração esses aspectos e propondo formas de inferência para eles. Para isso, será tomada como ponto de partida a modelagem através de processos gaussianos usada em Gelfand *et al.* (2005) no contexto de processos espaciais contínuos, acoplada à evolução dinâmica das taxas de intensidades proposta em Paez (2004) no contexto de processos pontuais. Os modelos resultantes são chamados de *processos pontuais espaciais dinâmicos*, por terem essa estrutura de evolução das taxas de intensidade ao longo do tempo.

A inferência para esses processos é feita sob o ponto de vista bayesiano, com utilização de métodos de amostragem *Monte Carlo via Cadeias de Markov* (*MCMC*, na abreviação em inglês). Com isso, será possível obter estimativas para a seqüência de taxas de intensidade e de parâmetros que estejam presentes na sua especificação como, por exemplo, médias e variâncias de evolução temporal e medidas de correlação da dispersão espacial. Além disso, será possível especificar as distribuições preditivas para as taxas de intensidade de tempos futuros, possibilitando a previsão de futuras ocorrências de eventos do fenômeno de interesse.

Este capítulo faz uma breve revisão dos conceitos e métodos estatísticos utilizados no desenvolvimento da tese, e está organizado do seguinte modo: na próxima seção é feita uma revisão do procedimento bayesiano de inferência e na seção seguinte são descritos alguns métodos computacionais aplicados à inferência bayesiana; na Seção 1.4 é feita uma ilustração destes métodos computacionais no contexto de uma breve revisão de modelos dinâmicos; alguns critérios de seleção de modelos são apresentados na Seção 1.5; finalmente, a Seção 1.6 descreve a organização dos capítulos da tese.

1.2 Inferência Bayesiana

Nesta seção, são apresentados os conceitos básicos da inferência bayesiana necessários ao entendimento da tese. Para uma discussão ampla e detalhada sobre o tema, são recomendados os livros de Berger (1985), Bernardo e Smith (1994) e Migon e Gamerman (1999).

1.2.1 O Teorema de Bayes

No procedimento bayesiano de inferência, a informação prévia sobre o vetor de parâmetros $\boldsymbol{\theta}$, contida na distribuição a *priori* $\pi(\boldsymbol{\theta})$, é combinada com a informação dos dados $\boldsymbol{y}$, contida na função de verossimilhança $f(\boldsymbol{y}\,|\,\boldsymbol{\theta})$, resultando na distribuição a *posteriori* $\pi(\boldsymbol{\theta} \mid \boldsymbol{y})$. O teorema de Bayes é a regra desta atualização da informação sobre os parâmetros:

$$\pi(\boldsymbol{\theta}\,|\,\boldsymbol{y}) = \frac{f(\boldsymbol{y}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})}{p(\boldsymbol{y})},$$

onde

$$p(\boldsymbol{y}) = \int f(\boldsymbol{y}\,|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,d\boldsymbol{\theta}.$$

A influência relativa de cada um destes componentes, priori e verossimilhança, na informação a posteriori depende de quanto peso é dado à distribuição a priori (o quão "informativa"ela é) e do tamanho da amostra.

1.2.2 Análise da Distribuição a Posteriori

A inferência sobre os parâmetros $\boldsymbol{\theta}$ é baseada nas informações contidas na distribuição a posteriori, seja através de medidas resumo como média, variância ou percentis, ou de intervalos de probabilidade:

Definição (Intervalo de Credibilidade): C é um intervalo de credibilidade $100(1-\alpha)\%$ para um escalar θ se $\int_C \pi(\theta\,|\,\boldsymbol{y})d\theta = 1-\alpha$, com $0< \alpha <1$.

Esta definição é facilmente estendida para a situação onde $\boldsymbol{\theta}$ é um vetor e C é uma região. Para um α fixo, o intervalo C de menor amplitude é aquele que inclui os pontos de mais alta densidade a posteriori; são os chamados intervalos MDP - *máxima densidade a posteriori*.

A predição de uma observação futura $\boldsymbol{z}$, após a observação dos dados $\boldsymbol{y}$, é baseada na distribuição de $\boldsymbol{z}|\,\boldsymbol{y}$, chamada de *distribuição preditiva*, dada pela expressão

$$f(\boldsymbol{z}|\boldsymbol{y}) \;=\; \int f(\boldsymbol{z},\boldsymbol{\theta}|\,\boldsymbol{y})\,d\boldsymbol{\theta} \;=\; \int f(\boldsymbol{z}|\,\boldsymbol{\theta},\boldsymbol{y})\,\pi(\boldsymbol{\theta}|\,\boldsymbol{y})\,d\boldsymbol{\theta} \;=\; \int f(\boldsymbol{z}|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta}|\,\boldsymbol{y})\,d\boldsymbol{\theta},$$

na qual a última passagem ocorre se $\boldsymbol{z}$ e $\boldsymbol{y}$ são condicionalmente independentes dado $\boldsymbol{\theta}$.

A densidade a posteriori $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ e a distribuição preditiva $f(\boldsymbol{z}|\boldsymbol{y})$ podem ser tão complexas a ponto de não permitirem a extração analítica de informações descritivas que exijam integração. Uma maneira de contornar este problema é conduzir a inferência baseada na análise de uma amostra simulada da distribuição a posteriori. Na próxima seção, são apresentados alguns métodos bastante utilizados de obtenção de amostras da posteriori utilizando-se métodos de simulação estocástica através de cadeias de Markov.

1.2.3 Escolha da Distribuição a Priori

Migon e Gamerman (1999) apresentam diferentes formas de especificação da distribuição a priori dos parâmetros. A distribuição a priori pode ser determinada a partir de conhecimentos subjetivos ou através do uso de informações sobre o parâmetro obtidas de experimentos passados.

Um procedimento indireto é a especificação através de formas funcionais de densidades paramétricas. Os parâmetros destas formas funcionais da distribuição a priori, chamados *hiperparâmetros*, são escolhidos de modo subjetivo de acordo com informações disponíveis. Um procedimento sistemático é escolher a forma funcional da distribuição a priori de modo que as distribuições a priori e a posteriori pertençam à mesma a família de distribuições, as chamadas famílias de distribuições conjugadas:

Definição (Família de Distribuições Conjugadas): Seja $\mathcal{F} = \{f(y\,|\,\theta), \theta \in \Theta\}$ uma família de distribuições amostrais (observacionais). Uma classe $\mathcal{P}$ de distribuições é dita ser uma família conjugada com respeito a $\mathcal{F}$ se, para todo $f \in \mathcal{F}$ e $p(\theta) \in \mathcal{P}$, tem-se que $\pi(\theta\,|\,y) \in \mathcal{P}$.

As vantagens da conjugação são especialmente a facilidade da análise e a possibilidade de explorar o aspecto seqüencial do paradigma bayesiano.

Alguns analistas preferem que a influência da informação a priori na inferência seja reduzida ao mínimo, ou seja, permitem que os dados determinem a região com maior massa de probabilidade a posteriori. Este é o conceito das prioris *não-informativas* ou *de referência*, também chamadas de *vagas* ou *planas* (*flat*). Uma priori não-informativa pode ser obtida a partir de uma priori conjugada definindo-se o hiperparâmetro de escala tendendo a zero e mantendo os outros constantes. Por exemplo, uma priori Normal com média zero e variância muito alta é relativamente plana. Um parâmetro de variância pode ter distribuição a priori Gama invertida pouco informativa se seus hiperparâmetros forem escolhidos com valores suficientemente baixos.

1.3 Métodos MCMC na Inferência Bayesiana

A densidade a posteriori π pode ser muito complexa e impossível de ser amostrada diretamente. Com o uso de um método Monte Carlo via cadeias de Markov (MCMC, na abreviação em inglês) é possível gerar uma cadeia de Markov ergódica que tenha π como distribuição de equilíbrio. Assim, após a convergência da cadeia para π, os valores gerados formam uma amostra desta distribuição, que pode ser usada para cálculos de Monte Carlo.

Nesta seção serão apresentados o amostrador de Gibbs e o algoritmo de Metropolis-Hastings, utilizados na inferência bayesiana dos modelos propostos nesta tese. Uma ampla discussão destes e de outros métodos, com sua aplicação em diversos modelos, é encontrada em Gamerman e Lopes (2006).

1.3.1 Amostrador de Gibbs

Com o objetivo de obter uma amostra da distribuição à posteriori $\pi(\theta_1, ..., \theta_d \,|\, \boldsymbol{y})$, o amostrador de Gibbs (Gelfand e Smith, 1990) simula sucessivamente e repetidamente das distribuições condicionais completas de cada componente dados os demais componentes, ou seja, gera valores de θ_i de $\pi(\theta_i \mid \boldsymbol{\theta}_{-i}, \boldsymbol{y}), \quad i = 1, ..., d,$ onde $\boldsymbol{\theta}_{-i} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_d)'$. Assume-se que estas distribuições são de fácil amostragem direta.

Os passos deste esquema de amostragem são:

1. Inicialize o contador de iterações da cadeia em $j = 1$ e atribua valores iniciais $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, ..., \theta_d^{(0)})'$;

2. Obtenha um novo valor $\boldsymbol{\theta}^{(j)} = (\theta_1^{(j)}, ..., \theta_d^{(j)})'$ através da geração sucessiva de valores

$$
\begin{aligned}
\theta_1^{(j)} &\sim \pi(\theta_1 \,|\, \theta_2^{(j-1)}, ..., \theta_d^{(j-1)}, \boldsymbol{y}), \\
\theta_2^{(j)} &\sim \pi(\theta_2 \,|\, \theta_1^{(j)}, \theta_3^{(j-1)}, ..., \theta_d^{(j-1)}, \boldsymbol{y}), \\
&\vdots \\
\theta_d^{(j)} &\sim \pi(\theta_d \,|\, \theta_1^{(j)}, ..., \theta_{d-1}^{(j)}, \boldsymbol{y});
\end{aligned}
$$

3. Mude o contador de j para $j+1$ e retorne ao passo 2 até que a convergência da cadeia seja atingida.

À medida que o número de iterações cresce, a cadeia aproxima-se da sua condição de equilíbrio. Quando a convergência é atingida, o valor resultante $\boldsymbol{\theta}^{(j)}$ é uma observação de π.

Assim, na prática, a cadeia é iterada um número suficientemente grande de iterações (digamos, J) tal que se possa assmuir que a convergência foi atingida. Este é o chamado período de *burn in*. Os valores $\boldsymbol{\theta}^{(J)}, ..., \boldsymbol{\theta}^{(M)}$ são tomados como uma amostra da distribuição a posteriori de $\boldsymbol{\theta}$. Como os valores seqüenciais nesta amostra são autocorrelacionados, é usual tomar uma sub-amostra sistemática dos valores, por exemplo, a cada $k > 1$ iterações, para reduzir este efeito.

A convergência pode ser muito lenta devido à alta correlação entre os elementos de $\boldsymbol{\theta}$. Uma solução para este problema é definir subconjuntos (chamados *blocos*) dos elementos de $\boldsymbol{\theta}$ que são amostrados conjuntamente.

1.3.2 Algoritmo de Metropolis-Hastings

Novamente, o objetivo é gerar um valor de θ de uma distribuição $\pi(\theta)$. No procedimento de inferência bayesiana, esta distribuição pode ser a posteriori de θ ou algumas das distribuições condicionais completas de θ_i no amostrador de Gibbs, quando estas não são de fácil amostragem direta. A idéia do algoritmo de Metropolis-Hastings (Metropolis *et al.*, 1953; Hastings, 1970) é amostrar um valor de θ da densidade $q(x \mid y)$ (chamada de *densidade da proposta*) da qual a geração de valores é possível ou mais fácil.

Os passos deste esquema de amostragem são:

1. Inicialize o contador de iterações da cadeia em $j=1$ e atribua um valor inicial $\theta^{(0)}$;

2. Obtenha um novo valor ϕ para θ gerado da distribuição $q(\phi \mid \theta^{(j-1)})$;

3. Avalie a probabilidade de aceitação do novo valor, dada por

$$\alpha(\theta^{(j-1)}, \phi) = min\left\{ 1, \ \frac{\pi(\phi)}{\pi(\theta^{(j-1)})} \frac{q(\theta^{(j-1)} \mid \phi)}{q(\phi \mid \theta^{(j-1)})} \right\}.$$

   Se o novo valor é aceito, $\theta^{(j)} = \phi$; caso contrário, $\theta^{(j)} = \theta^{(j-1)}$;

4. Mude o contador de j para $j+1$ e retorne ao passo 2 até que a convergência da cadeia seja atingida.

Após a convergência da cadeia para sua condição de equilíbrio, digamos, na iteração J, os valores $\boldsymbol{\theta}^{(J)}, ..., \boldsymbol{\theta}^{(M)}$ constituem-se em uma amostra (correlacionada) da distribuição a posteriori de $\boldsymbol{\theta}$.

Em geral, a taxa de aceitação dos valores novos é ajustada para cerca de 50% através da definição de uma constante sintonizadora da probabilidade de aceitação do valor proposto, geralmente associada à variância da densidade da proposta q.

Assim como no amostrador de Gibbs, a amostragem dos parâmetros $\boldsymbol{\theta}$ também pode ser feita em blocos de seus elementos.

1.3.3 Avaliação da Convergência da Cadeia

A teoria de MCMC nos garante que a cadeia de Markov irá eventualmente produzir uma amostra da distribuição alvo se a cadeia é rodada por um tempo suficientemente longo. A questão de difícil resposta é saber quão longo é suficiente para garantir a convergência.

Existem métodos formais de verificação da convergência das cadeias, como o procedimento de Geweke (1992) e a estatística de Gelman e Rubin (1992), modificada por Brooks e Gelman (1998). Entretanto, nenhum destes métodos é conclusivo, fornecendo apenas indícios de convergência.

Um modo informal simples de verificação da convergência é a análise das séries temporais de várias estatísticas derivadas da cadeia de Markov, como somas, médias ou índices úteis na descrição dos dados. Considera-se que a cadeia aparentemente convergiu quando a série temporal destas estatísticas estabiliza-se.

Do mesmo modo, pode-se analisar a trajetória de pelo menos duas cadeias independentes (definidas por diferentes valores iniciais) dos próprios parâmetros e verificar se todas convergem para o mesmo ponto de estabilidade.

1.4 Modelos Dinâmicos

Modelos dinâmicos são uma ampla classe de modelos de regressão e de séries temporais nos quais os parâmetros mudam com a passagem do tempo. Eles incluem como caso particular os modelos estáticos, nos quais esta mudança temporal não existe.

Nesta seção serão apresentados os modelos dinâmicos e seus procedimentos de inferência utilizados na tese. Detalhes da modelagem, aplicações e extensa discussão do assunto podem ser encontrados no livro de West e Harrison (1997) e no recente artigo de Migon *et al.* (2005).

1.4.1 Modelos Dinâmicos Lineares

Os modelos dinâmicos lineares consistem em uma equação de regressão relacionando os parâmetros às observações e uma equação relacionando entre si os sucessivos parâmetros da regressão:

$$\begin{array}{lrcllr} \textit{Equação das observações:} & y_t & = & F_t^{'}\theta_t & + \ \epsilon_t, & \epsilon_t \sim N[0; V_t]; \\ \textit{Equação do sistema:} & \theta_t & = & G_t\theta_{t-1} & + \ \omega_t, & \omega_t \sim N[0; W_t], \end{array}$$

onde $\{y_t\}$ é uma seqüência de observações no tempo, condicionalmente independentes dados V_t e o vetor de parâmetros de estado θ_t, F_t é um vetor de variáveis explicativas e G_t é uma matriz que descreve a evolução dos parâmetros de estado. O modelo é completado com a especificação de uma priori normal para θ_1.

A seguir são apresentados dois exemplos dos chamados *modelos de tendência*.

Exemplo 1.1: O mais simples dos modelos dinâmicos é o chamado *modelo polinomial de primeira ordem*, no qual o nível da série temporal permanece localmente estável, mas varia a longos intervalos de tempo. Este modelo é descrito por:

$$\begin{array}{rcllr} y_t & = & \mu_t & + \ \epsilon_t, & \epsilon_t \sim N[0; V_t]; \\ \mu_t & = & \mu_{t-1} & + \ \omega_t, & \omega_t \sim N[0; W_t], \end{array}$$

onde μ_t é escalar. Este modelo é obtido a partir do modelo geral definindo $F_t = 1$ e $G_t = 1$.

Exemplo 1.2: O *modelo polinomial de segunda ordem* permite que haja um crescimento no nível da série com a inclusão do parâmetro escalar β_t:

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, & \epsilon_t &\sim N[0; V_t]; \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \omega_{1t}, & \omega_{1t} &\sim N[0; W_{1t}], \\ \beta_t &= \beta_{t-1} + \omega_{2t}, & \omega_{2t} &\sim N[0; W_{2t}]. \end{aligned}$$

Este modelo é obtido a partir do modelo geral tomando $F_t = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ e $G_t = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Ambos modelos serão utilizados no Capítulo 4.

Quando V_t e W_t são conhecidos, a inferência pode ser feita analiticamente e as densidades a posteriori são normais. O algoritmo *Filtro de Kalman* (Anderson e Moore, 1979) fornece as distribuições em tempo real de $\theta_t | D_t$, $\forall t$, com $D_t = \{y_1, ..., y_t\}$. Os detalhes desta inferência seqüencial são mostrados no Apêndice A.1.

Quando V_t e W_t são desconhecidos, a inferência não pode ser feita de forma analítica. Dentre as diversas alternativas existentes para se realizar uma inferência aproximada, destacam-se os procedimentos baseados em métodos MCMC (Migon *et al.*, 2005). No Apêndice A.2, é descrito o algoritmo *Forward Filtering Backward Smoothing* (FFBS), proposto por Carter e Kohn (1994) e Frühwirth-Schnatter (1994). Este é o esquema utilizado na amostragem do componente temporal nos modelos espaço-temporais propostos no Capítulo 4.

1.4.2 Modelos Dinâmicos Lineares Generalizados

West *et al.* (1985) estenderam o modelo dinâmico linear para situações nas quais as observações da série temporal pertencem à ampla *família exponencial de distribuições*. A variável aleatória Y_t tem uma distribuição pertencente à família exponencial se sua função de densidade (de probabilidade) puder ser escrita na forma

$$p(y_t \mid \eta_t, V_t) = \exp\left\{ V_t^{-1} \left[y_t \eta_t - b(\eta_t) \right] \right\} a(y_t, V_t),$$

onde η_t e V_t são parâmetros definidos de acordo com a distribuição específica; $b(\eta_t)$ e $a(y_t, V_t)$ são funções conhecidas e $\mu_t \doteq E(Y_t \mid \eta_t) = b'(\eta_t)$.

Desse modo, o *modelo dinâmico linear generalizado* é definido pelos seguintes componentes:

$$\begin{aligned} \textit{Equação das observações:} \qquad p(y_t \mid \eta_t) &\propto \exp\left\{V_t^{-1}\left[y_t\eta_t - b(\eta_t)\right]\right\}, \\ g(\mu_t) &= F_t^{'}\theta_t; \\ \textit{Equação do sistema:} \qquad \theta_t &= G_t\theta_{t-1} + \omega_t, \qquad \omega_t \sim N[0; W_t], \end{aligned}$$

onde g é uma função de ligação conhecida, contínua e monótona que projeta μ_t na reta real.

A inferência pode ser feita via métodos MCMC. Entretanto, a distribuição condicional completa dos estados $\boldsymbol{\theta} = (\theta_1, ..., \theta_T)'$ não é conhecida. Para amostrar desta distribuição, Gamerman (1997) sugere o uso de blocos dos θ_t, fazendo uma reparametrização em função dos erros $\omega_1 = \theta_1$ e $\omega_t = \theta_t - G_t\theta_{t-1}, t = 2, ..., T$. A amostragem é feita em função destes erros, evitando, assim, a lenta convergência da cadeia devido à forte correlação entre os estados θ_t. A reconstrução dos estados originais é feita facilmente através da relação $\theta_1 = \omega_1$ e $\theta_t = \sum_{l=1}^{t}\left(\prod_{k=1}^{t-l} G_{t-k+1}\right)\omega_l,\ t = 2, ..., T$.

Ravines (2006) propõe um esquema de amostragem eficiente na inferência bayesiana em modelos dinâmicos não normais e não lineares, denominado CUBS (abreviação de *Conjugate Updating Backward Sampling*). Os resultados obtidos mostram que o esquema proposto é eficiente no sentido de reduzir significativamente o tempo computacional e ser de fácil implementação.

1.5 Seleção de Modelos

A escolha entre diferentes propostas de modelos é uma etapa fundamental na análise de conjuntos de dados. Se "todos os modelos são errados, mas alguns são úteis" (Box, 1976), dentre estes modelos úteis deve-se identificar aqueles que descrevam adequadamente a informação nos dados e/ou forneçam previsões eficazes. Ainda que as ferramentas computacionais nos habilitem a ajustar modelos cada vez mais complexos, não se deve perder de vista o critério da parcimônia e a interpretabilidade do modelo.

Medir a complexidade de um modelo é mais do que contar o número de parâmetros quando se trata da comparação de modelos com efeitos fixos contra modelos que também incluem efeitos aleatórios ou ainda entre modelos não encaixados. É o caso dos modelos hierárquicos complexos nos quais o número de parâmetros não está definido claramente.

A seguir são apresentados dois conhecidos critérios de seleção de modelos - DIC e EPD, que serão utilizados neste trabalho.

1.5.1 DIC - Deviance Information Criterion

O DIC foi proposto por Spiegelhalter *et al.* (2002) como uma generalização do critério de informação de Akaike - AIC (Akaike, 1973).

Considere um modelo com um vetor de observações $\boldsymbol{y} = (y_1, ..., y_n)'$ e um vetor de parâmetros $\boldsymbol{\theta}$, cuja função de verossimilhança é denotada por $p(\boldsymbol{y}|\,\boldsymbol{\theta})$. A *deviance* do modelo é definida por $D(\boldsymbol{\theta}) = -2\log[p(\boldsymbol{y}|\,\boldsymbol{\theta})]$.

A média a posteriori da deviance, denotada por $E_{\theta|y}\left[D(\boldsymbol{\theta})\right]$, pode ser pensada como uma medida bayesiana de ajuste ou adequação do modelo. O número efetivo de parâmetros no modelo é definido como sendo a diferença entre a média a posteriori da deviance e a deviance avaliada nas médias a posteriori dos parâmetros:

$$p_D = E_{\theta|y}\left[D(\boldsymbol{\theta})\right] - D\big[E_{\theta|y}(\boldsymbol{\theta})\big]\,.$$

Quanto menor o valor de p_D, menor é a complexidade do modelo.

O DIC é então definido como a soma destes dois componentes - uma medida da bondade do ajuste e uma penalização pela complexidade do modelo:

$$DIC = E_{\theta|y}\left[D(\boldsymbol{\theta})\right] + p_D.$$

Dentre os modelos comparados, aquele com menor valor de DIC é considerado o mais adequado.

O DIC é um critério de fácil implementação em procedimentos de ajuste de modelos via MCMC. Sejam $\boldsymbol{\theta}^{(1)}, ..., \boldsymbol{\theta}^{(M)}$ uma amostra da distribuição a posteriori $p(\boldsymbol{\theta}\,|\,\boldsymbol{y})$,

$$\bar{D} = \frac{1}{M}\sum_{j=1}^{M} D\left(\boldsymbol{\theta}^{(j)}\right) \quad \text{e} \quad D(\bar{\boldsymbol{\theta}}) = D\left(\frac{1}{M}\sum_{j=1}^{M}\boldsymbol{\theta}^{(j)}\right);$$

tem-se que $DIC = 2\bar{D} - D(\bar{\boldsymbol{\theta}})$.

Em geral, tanto o componente p_D quando o DIC são valores positivos. Entretanto, a componente p_D pode ser negativa se a função de verossimilhança não for log-côncava; quando há conflito entre a distribuição a priori e a função de verossimilhança; ou ainda se a distribuição a posteriori dos parâmetros é muito assimétrica ou simétrica bimodal, de modo que a média a posteriori não seja uma boa medida de tendência central. O valor do DIC também pode ser negativo se a deviance é negativa, o que ocorre quando a densidade de probabilidade é maior que um. Entretanto, este fato não interfere no uso do critério na comparação de modelos, pois o foco está na diferença entre seus valores, não no valor do DIC propriamente.

1.5.2 EPD - Expected Predictive Deviance

Gelfand e Ghosh (1998) apresentam o EPD, um critério preditivo cujo objetivo é escolher, dentre os modelos ajustados, aquele que fornece a melhor predição de réplicas dos dados observados. A idéia é amostrar um "novo" conjunto de dados da distribuição preditiva:

$$f(y_i^N|y_i) \;=\; \int f(y_i^N|\boldsymbol{\theta})f(\boldsymbol{\theta}|y_i)d\boldsymbol{\theta},$$

onde y_i^N é visto como uma réplica (ou predição) da observação y_i. Uma vez definida uma função de discrepância $D(\boldsymbol{y}^N, \boldsymbol{y})$ entre os dados observados e preditos, o critério escolhe o modelo que minimiza a esperança a posteriori desta discrepância.

No caso de modelos normais, uma função de discrepância adequada é a soma de quadrados $D(\boldsymbol{y}^N, \boldsymbol{y})=(\boldsymbol{y}^N-\boldsymbol{y})'(\boldsymbol{y}^N-\boldsymbol{y})$, levando ao cálculo explícito do EPD por

$$EDP \;=\; \sum_{i=1}^{n} Var(Y_i^N|y_i) \;+\; \sum_{i=1}^{n} \left[E(Y_i^N|y_i) - y_i\right]^2.$$

O modelo com menor valor de EPD é escolhido como o mais adequado.

Em modelos com verossimilhança de Poisson, a função perda sugerida é a deviance usual adaptada para comparar dados reais e suas réplicas:

$$D(\boldsymbol{y}^N, \boldsymbol{y}) = -2\sum_{i=1}^{n} \left[\, y_i \log(y_i/y_i^N) - (y_i - y_i^N)\,\right].$$

Uma correção no caso de contagens baixas é dada por

$$D^*(\boldsymbol{y}^N, \boldsymbol{y}) = -2\sum_{i=1}^{n} \left[\, (y_i+0.5)\log[(y_i+0.5)/(y_i^N+0.5)] - (y_i - y_i^N)\,\right].$$

O valor de EPD é então dado pela média de $D^*(\boldsymbol{y}^N, \boldsymbol{y})$ baseada em amostradas repetidas da distribuição preditiva de $\boldsymbol{y}^N$.

O critério EPD também é de fácil implementação em algoritmos MCMC de amostragem da distribuição a posteriori dos parâmetros do modelo.

1.6 Organização da Tese

Este texto é composto de mais seis capítulos. O Capítulo 2 apresenta uma introdução aos conceitos básicos dos processos pontuais no espaço e/ou no tempo. O Capítulo 3 apresenta uma proposta de modelagem espacial. Os modelos espaço-temporais são propostos e estudados no Capítulo 4. Estudos de simulação dos modelos propostos são apresentados no Capítulo 5, enquanto sua aplicação a conjuntos de dados reais são mostrados no Capítulo 6. Finalmente, o Capítulo 7 resume as conclusões do trabalho e aponta caminhos para pesquisa futura nesta importante área da Estatística.

Capítulo 2

Processos Espaciais

2.1 Introdução

Muitos dos fenômenos estudados nas diferentes áreas do conhecimento, como saúde pública, meio-ambiente, geologia, estudos de criminalidade, dentre outras, apresentam variabilidade das observações sobre o espaço e o tempo.

Nos últimos anos tem havido um grande crescimento de técnicas e modelos estatísticos para analisar conjuntos de dados espaço-temporais. Tais dados são utilizados para detectar padrões significativos de uma variável na região, estudar sua evolução temporal, bem como fazer previsões.

Neste capítulo, primeiramente são apresentados os conceitos básicos sobre processos espaciais e os tipos de dados espaciais gerados a partir deles. Na Seção 3 é apresentado o processo gaussiano. Os processos pontuais espaciais estudados na tese são introduzidos na Seção 4. A Seção 5 introduz os processos espaço-temporais. A Seção 6 discute a necessidade de discretização do espaço para a inferência.

As definições adotadas neste texto, baseadas em Diggle (2003), não são definições formais. Será introduzida apenas a teoria básica para a compreensão do assunto. Para definições com maior rigor matemático, ver por exemplo, Cressie (1993) para processos espaciais em geral, e Daley e Vere-Jones (2003) ou Møller e Waagepetersen (2003 e 2007) para processos pontuais.

2.2 Processos Espaciais

Um processo estocástico com domínio no espaço é chamado um *processo espacial*. Um processo espacial é definido por

$$\{Z(s) : s \in D \subset \Re^2\}, \tag{2.1}$$

onde D é um conjunto de índices e $Z(s)$ é o atributo de interesse na localização s. Por simplicidade, a dimensão de D será considerada igual a 2, representando observações no plano. A natureza do conjunto D permite a definição de três principais tipos de dados espaciais, de acordo com Cressie (1993):

1. **Dados Geoestatísticos**
$Z(s)$ é uma variável aleatória observada nas localizações $s \in D$, onde D é fixo e contínuo.
Exemplos: medições do volume de chuva em estações meteorológicas de um estado, medições do nível de um poluente atmosférico em pontos de uma cidade.

2. **Dados de Área**
$Z(s)$ é uma variável aleatória observada nas localizações $s \in D$, onde D é fixo e discreto.
Exemplos: número de casos de uma doença por município de um estado, número de furtos de veículos por bairro de uma cidade.

3. **Arranjos Pontuais**
$Z(s)$ é uma variável aleatória observada nas localizações $s \in D$, onde D é um conjunto aleatório de índices.
Exemplos: as localizações de focos de incêndio em uma floresta, as residências com focos do mosquito *Aedes aegypti* em uma cidade.

Nesta tese, são estudados modelos para arranjos pontuais. Entretanto, um importante modelo para dados com D contínuo, o processo gaussiano, será utilizado na definição de componentes dos modelos propostos e, portanto, é apresentado na próxima seção.

2.3 O Processo Gaussiano

O *processo gaussiano* no plano é definido como o processo estocástico $x(\cdot)$ na região $D \in \Re^2$, com D fixo e contínuo, tal que, para $n \geqslant 1$ e localizações espaciais $s_1, \ldots, s_n$, o vetor $(x(s_1), \ldots, x(s_n))$ tem distribuição Normal multivariada com vetor de médias $\boldsymbol{m}$ e matriz de variâncias e covariâncias Σ.

As suposições usuais são:

- *estacionariedade*, que implica que $\boldsymbol{m} = \mu \underline{1}$ e $\Sigma = \sigma^2 R$, onde R é uma matriz de correlações tais que $r_{ij} = \rho(s_i - s_j; \theta)$ para uma função de correlação adequada ρ (Vide subseção a seguir.);

- *isotropia*, que implica que a função de correlação ρ_θ depende apenas da distância $\|s_i - s_j\|$ entre as localizações s_i e s_j.

A notação

$$x(\cdot) \mid \mu, \sigma^2, \theta \ \sim \ PG\left[\mu; \sigma^2; \rho(\cdot; \theta)\right]$$

será utilizada neste texto para denotar um processo gaussiano estacionário e isotrópico com média μ, variância σ^2 e função de correlação espacial ρ. A suavidade na variação espacial depende essencialmente da função de correlação espacial. Em geral, estruturas suaves podem ser obtidas com a definição, via especificação de θ, de valores altos para a correlação espacial entre localizações próximas.

Gamerman *et al.* (2007) descrevem a classe de *processos gaussianos dinâmicos*, que são obtidos como uma extensão dos processos gaussianos, quando se introduz o componente do tempo, ou como uma extensão dos modelos dinâmicos, quando a dimensão espacial é introduzida. Estes processos podem ser usados como prioris de alguns componentes de diferentes modelos espaço-temporais, como nos modelos de regressão (Gelfand *et al.*, 2005), na análise fatorial espacial dinâmica (Salazar, 2006) e nos processos pontuais espaço-temporais estudados no Capítulo 4 desta tese.

2.3.1 Simulação de Dados de Processos Gaussianos

Há vários métodos disponíveis para simulação de um campo aleatório gaussiano (Lantuéjoul, 1994). O processo gaussiano pode ser simulado usando-se métodos Monte Carlo. O domínio infinito da região de simulação é representado por uma grade $\mathcal{G}_{\mathcal{N}} = \{c_1, ..., c_N\}$, na qual cada célula c_i tem área a_i, e o processo é aproximado por seus valores da distribuição gaussiana de dimensão finita nas N células da grade. Se o processo tem intensidade e agregação moderados, as propriedades de pequena escala do campo gaussiano não são tão importantes, podendo ser adotada uma discretização mais "grosseira". O erro resultante da discretização também depende da suavidade das realizações do campo gausssiano, sendo menor quando a função de correlação espacial decresce lentamente com a distância.

A geração de processos gaussianos está implementada em linguagens de programação como o *R* (R Development Core Team, 2004) que tem disponíveis, por exemplo, as bibliotecas *RandomFields* (Schlather, 2001) e *geoR* (Ribeiro e Diggle, 2001).

2.3.2 Famílias de Funções de Correlação Espaciais

Se o processo espacial for assumido isotrópico, a função de correlação espacial $\rho(d)$ será função apenas da distância euclidiana d entre duas localizações. É desejável que esta função satisfaça as seguintes propriedades:

1. $\rho(d;\theta)$ é monótona não-crescente em d;

2. $\rho(d;\theta) \rightarrow 0$ quando $d \rightarrow \infty$;

3. Pelo menos um dos parâmetros em θ controla a taxa com que ρ decai para zero.

Há diversas famílias de funções de correlação espaciais, dentre elas as mais conhecidas e utilizadas são descritas a seguir.

Família Exponencial Potência:

$$\rho(d;\theta) = \exp\{-(d/\phi)^{\kappa}\},$$

onde $\theta = (\rho,\kappa)$, $\phi > 0$ é o parâmetro de escala e $\kappa \in (0;2]$. Quando $\kappa = 1$, tem-se o caso particular da *função de correlação exponencial*; $\kappa = 2$ corresponde à *função de correlação gaussiana*.

Família Esférica:

$$\rho(d;\rho) = \begin{cases} 1 - \frac{3}{2}(d/\phi) + \frac{1}{2}(d/\phi)^3 & , 0 \leqslant d \leqslant \phi; \\ 0 & , d > \phi, \end{cases}$$

onde $\theta = \phi > 0$ é o parâmetro de escala.

Família Matérn (Matérn, 1986):

$$\rho(d;\rho,\kappa) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)} \left(\frac{d}{\phi}\right)^{\kappa} K_{\kappa}\left(\frac{d}{\phi}\right),$$

onde $\theta = (\rho,\kappa)$, $\phi > 0$ é o parâmetro de escala e $\kappa > 0$ é o parâmetro de forma; a função $\Gamma(\cdot)$ é a função gama e κ_κ é a função modificada de Bessel do terceiro tipo de ordem κ (Abramowitz e Stegun, 1972). As funções exponencial e gaussiana também pertencem a esta família, quando $\kappa = 0.5$ e $\kappa \rightarrow \infty$, respectivamente.

2.4 Processos Espaciais Pontuais

Um *processo espacial pontual* $Z(s)$, $s \in D$, onde D é um conjunto aleatório de índices, é um processo estocástico que governa a distribuição (localização) e o número de realizações de um fenômeno nesta região do espaço. Tal processo espacial difere-se dos outros dois tipos de processos espaciais pelo fato de que o componente estocástico primário é a própria localização espacial das observações. O conjunto das localizações espaciais observadas $\boldsymbol{x} = \{x_1, ..., x_n\}$ é chamado de *arranjo pontual* e cada uma delas é usualmente chamada de *evento*, para distingui-las de pontos arbitrários no plano, denotados por s.

Os conceitos de média e covariância dos processos contínuos são definidos, para os processos pontuais, em função dos efeitos de primeira e segunda ordens. A *função de intensidade de primeira ordem* é uma medida de uniformidade e envolve o número médio de eventos por unidade de área no ponto s (Diggle, 2003):

$$\lambda(s) = \lim_{|ds| \rightarrow 0} \frac{E\left[Z(ds)\right]}{|ds|}, \tag{2.2}$$

onde $E[\ ]$ denota o valor esperado, ds é uma região infinitesimal em torno do ponto s e $|ds|$ é a área desta região. A *função de intensidade de segunda ordem* é uma medida da estrutura de dependência entre as localizações s_i e s_j (Diggle, 2003):

$$\lambda_2(s_i, s_j) = \lim_{|ds_i|,|ds_j| \to 0} \frac{E\left[Z(ds_i)Z(ds_j)\right]}{|ds_i|\,|ds_j|}.$$

Um processo pontual espacial é dito ser *fracamente estacionário* se o processo é invariante em localização, isto é,

$$\lambda(s) = \lambda \quad \forall s \in D \qquad \text{e} \qquad \lambda_2(s_i, s_j) = \lambda_2(\boldsymbol{h}) \quad \forall s_i, s_j \in D,$$

onde $\boldsymbol{h} = s_i - s_j$ é o vetor bidimensional da mudança em localização espacial do ponto s_i ao ponto s_j (Diggle, 2003). Isto equivale a dizer que o número esperado de eventos em uma localização arbitrária é constante e a dependência entre os eventos em duas localizações quaisquer depende apenas do vetor diferença $\boldsymbol{h}$ e não das localizações específicas s_i e s_j.

A função de covariância fracamente estacionária pode ser definida como *anisotrópica* ou *isotrópica*. Um processo isotrópico é invariante sob translação e rotação em um ângulo qualquer, ou seja, sua função de covariância não depende da direção de $\boldsymbol{h}$, que pode ser substituído por $h = \|s_i - s_j\|$, a distância euclidiana entre s_i e s_j.

2.4.1 Tipos de Arranjos Pontuais

Basicamente são considerados três tipos básicos de arranjos pontuais: *agregado*, *regular* e *aleatório*. A Figura 2.1 mostra uma realização simulada de cada um destes tipos.

No arranjo pontual agregado, como o próprio nome diz, os eventos aparecem formando diversos agrupamentos no espaço. É freqüentemente observado quando as sementes de planta são espalhadas nas proximidades da planta-mãe. O oposto direto da agregação é o arranjo regular, no qual os eventos não ocorrem (ou têm uma probabilidade muito baixa de ocorrer) dentro de uma certa distância uns dos outros, como, por exemplo, os centros das células biológicas. No arranjo aleatório os eventos se distribuem no espaço de maneira completamente ao acaso.

Arranjos pontuais heterogêneos podem surgir, por exemplo, da observação das posições de plantas onde a fertilidade do solo exibe uma variação espacial. Se é assumido que a fertilidade é um campo aleatório que varia espacialmente, mas que, condicional à fertilidade do solo (e possivelmente a outros fatores ambientais) as localizações das plantas são independentes, o

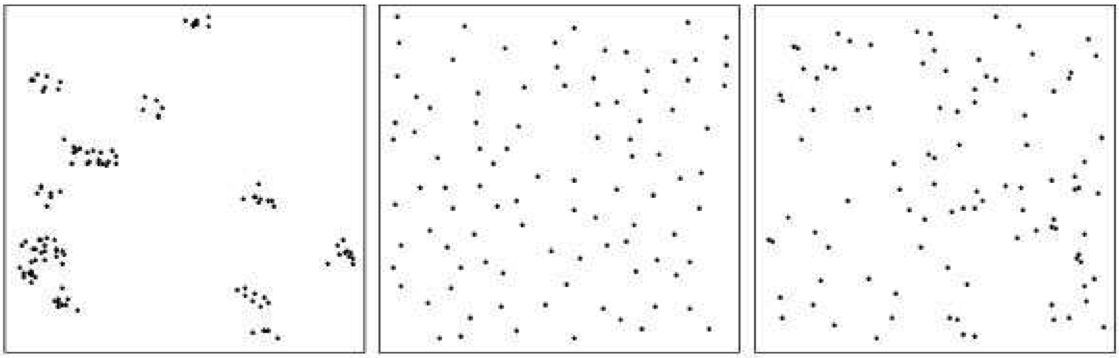


Figura 2.1: Os tipos básicos de arranjos pontuais espaciais. Da esquerda para a direita: agregado, regular e aleatório.

modelo apropriado é um processo com intensidade (de primeira ordem) variando no espaço e intensidade de segunda ordem nula. Por outro lado, se há dependência entre as plantas (como competição), é natural pensar em um processo que tenha termos de interação.

Dentre os vários tipos de processos pontuais que geram arranjos pontuais agregados, regulares ou aleatórios, são apresentados neste texto apenas os processos que serão importantes na compreensão da modelagem proposta neste trabalho. Nestes processos, não há efeito de interação entre os eventos. Assim, a eventual agregação dos eventos é atribuída unicamente à heterogeneidade na intensidade do processo.

2.4.2 Alguns Modelos para Processos Espaciais Pontuais

O mais simples dos processos espaciais pontuais é aquele em que não há efeitos de primeira nem de segunda ordens: a intensidade é constante no espaço e os eventos não interagem espacialmente. Esta situação, chamada de *aleatoriedade espacial completa*, define o processo de Poisson homogêneo.

Processo de Poisson Homogêneo

Neste processo pontual, o número de eventos N em uma região planar limitada $A \subset \Re^2$ é uma variável aleatória Poisson com média $\lambda|A|$, sendo $|A|$ a área de A; adicionalmente, condicionadas à intensidade, as contagens de eventos em regiões disjuntas são independentes. Este processo tem $\lambda(s) = \lambda$ e $\lambda_2(s_i, s_j) = \lambda^2$, ou seja, é estacionário e isotrópico.

Pela definição do modelo, a função de verossimilhança de λ não depende da localização dos eventos $\boldsymbol{x}=\{x_1,...,x_n\}$ na região A, mas apenas do número de eventos n, sendo proporcional a

$$l(\lambda;n) \;\propto\; \exp\{-\lambda|A|\}\,(\lambda|A|)^n.$$

O processo de Poisson homogêneo é útil como base de comparação, mas pouco realístico para aplicações. Ainda que não haja interação espacial entre os eventos, raramente se tem homogeneidade na intensidade. Assumindo eventos independentes, mas com a intensidade $\lambda(s)$ variando no espaço, um padrão pontual espacial pode ser modelado através do *processo de Poisson não-homogêneo*.

Processo de Poisson Não-Homogêneo

Nele, o número de eventos em uma região $A\subset\Re^2$ tem distribuição de Poisson com média $\mu(A)=\int_A \lambda(s)ds$ e, para regiões disjuntas, as contagens de eventos são independentes. Este é um processo não-estacionário, mas tem apenas efeitos de primeira ordem: a aglomeração dos eventos é resultante da heterogeneidade da intensidade, não da atração entre eventos.

A função de verossimilhança de $\lambda(\cdot)$, baseada no conjunto de eventos $\boldsymbol{x} = \{x_1,...,x_n\}$ observados na região A, é proporcional a

$$l(\lambda;\boldsymbol{x}) \;\propto\; \exp\left\{-\int_A \lambda(s)ds\right\}\prod_{z\in\boldsymbol{x}}\lambda(z).$$

Processo de Cox

Cox (1955) apresentou o *processo de Poisson duplamente estocástico*, para o qual a superfície de intensidade também é assumida ser estocástica. Assim, seja $\Lambda=\{\Lambda(s) : s \in S\}$ um campo aleatório não-negativo. Se a distribuição condicional de Z dado $\Lambda=\lambda$ é um processo de Poisson em S com função de intensidade $\lambda(s)$, então Z é um *processo de Cox dirigido por* Λ. O processo pontual resultante é estacionário e isotrópico se, e somente se, o processo Λ o é.

A decisão sobre a aleatoriedade ou não da função intensidade, ou de parte dela, depende de questões científicas do fenômeno e/ou conhecimento prévio da aplicação em particular. Quando apenas uma realização do processo pontual está disponível, não se consegue distinguir um processo de Cox de um processo de Poisson não-homogêneo.

Processo de Cox Log-Gaussiano

No processo pontual de Cox, se $log[\Lambda(\cdot)] = \Phi(\cdot)$ é um processo gaussiano, o processo pontual resultante é denominado *processo de Cox log-gaussiano* (Møller *et al.*, 1998).

Desse modo, a função de verossimilhança do processo de Cox log-gaussiano decorre diretamente da função de verossimilhança do processo de Poisson não-homogêneo, sendo dada por

$$l(\phi; \boldsymbol{x}) \;\propto\; \exp\left\{-\int_S \exp[\phi(s)]ds\right\} \prod_{z\in\boldsymbol{x}} \exp[\phi(z)],$$

na qual $\boldsymbol{x} = \{x_1, ..., x_n\}$ é a localização dos eventos observados na região S.

Esta verossimilhança não é analiticamente tratável, pois depende de um número infinito de variáveis aleatórias $\{\phi(s),\ s \in S\}$.

2.4.3 Simulação de Dados de Processos Espaciais Pontuais

A geração de conjuntos de dados simulados do processo de Poisson homogêneo e suas extensões é geralmente simples e está implementada em vários programas computacionais de análise estatística, como o *R* (R Development Core Team, 2004), que tem disponíveis as bibliotecas *Splancs* (Rowlingson e Diggle, 1993) e *Spatstat* (Baddeley e Tuner, 2005).

A geração de um arranjo pontual do processo Poisson com intensidade λ em uma região D tem dois estágios: (i) uma contagem N da distribuição de Poisson com média λ é gerada; (ii) as posições dos N eventos são determinadas pela simulação de pontos independentes e uniformes em D.

Lewis e Shedler (1979) propõem gerar um processo de Poisson não-homogêneo com função de intensidade $\lambda(x), x \in D$ através de algoritmo baseado em amostragem por rejeição. Na sua forma mais simples, este algoritmo consiste em gerar um processo de Poisson homogêneo com intensidade $\lambda_{\max} = max\{\lambda(x); x \in D\}$ e reter cada evento gerado com probabilidade $\lambda(x)/\lambda_{max}$.

O processo de Cox log-gaussiano pode ser simulado usando-se métodos Monte Carlo (Møller e Waagepetersen, 2003). Assim como sua definição, a simulação do processo de Cox log-gaussiano envolve duas etapas. Primeiramente, o campo gaussiano Φ é simulado nas N subregiões que particionam a região de estudo e, dada sua realização $\hat{\phi} = (\hat{\phi}_1, ..., \hat{\phi}_N)$, geram-se N contagens de Poisson independentes com médias $\hat{\lambda}_i = a_i \exp(\hat{\phi}_i)$, onde a_i é a área da i-ésima subregião, para $i = 1, ..., N$.

2.5 Processos Pontuais Espaço-Temporais

Um processo pontual espaço-temporal é um processo estocástico que tem como realizações pontos com coordenadas aleatórias no espaço e no tempo. Estes processos pontuais podem ser considerados como um híbrido de um componente espacial e um componente temporal (Dorai-Raj, 2001). Estendendo a definição $Z(s)$ em (2.1) para incluir o tempo, obtem-se a seguinte definição de um processo pontual espaço-temporal:

$$\{Z(s,t) : s \in D \subset \Re^2, t \in [0,T] \subset \Re\}, \tag{2.3}$$

onde D é um conjunto aleatório de índices.

Segundo Schoenberg *et al.* (2002), um processo pontual espaço-temporal Z é caracterizado unicamente pelo seu processo de intensidade condicional λ. Assim como em (2.2), a intensidade $\lambda(s,t)$ do processo na localização espacial s e no tempo t pode ser pensada como a freqüencia com a qual os eventos são esperados ocorrer em torno de uma localização (s,t) no espaço e tempo, condicionada na história a priori do processo até o tempo t, denotada por H_t. Formalmente, $\lambda(s,t)$ pode ser definida como a esperança condicional limite, como explicado a seguir. Fixe qualquer ponto (s,t) no espaço-tempo, onde $s = (s_1, s_2) \in \Re^2$. Seja $\boldsymbol{B}_\Delta$ o conjunto $(t, t+\Delta_t) \times (s_1, s_1+\Delta_{s1}) \times (s_2, s_2+\Delta_{s2})$, onde Δ é o vetor $(\Delta_t, \Delta_{s1}, \Delta_{s2})$. Então

$$\lambda(s,t) = \lim_{\Delta \to 0} E\left[Z(\boldsymbol{B}_\Delta) \,|\, H_t)\right] / |\Delta|, \tag{2.4}$$

se este limite existe.

O conjunto de dados observados deste processo é chamado de *arranjo pontual espaço-temporal*, sendo formado pelo registro $\xi = \{(x_1, t_1), ..., (x_n, t_n)\}$ das localizações espaciais x_i e respectivo tempo de ocorrência t_i dos n eventos.

Arranjos pontuais espaço-temporais são freqüentemente analisados com negligência ao componente temporal, através da investigação das propriedades de primeira e segunda ordens do processo espacial separadamente para cada período de tempo. Esta abordagem oferece uma visão limitada da evolução do padrão espacial através do tempo, pois, sem a incorporação direta de uma relação temporal entre todos os arranjos espaciais observados, muito da inferência sobre o processo pode ser perdido.

Fishman e Snyder (1976) definem e estudam uma classe geral de processos pontuais no espaço-tempo a qual chamam de *analítica*. Dorai-Raj (2001) introduz vários tipos de processos pontuais espaço-temporais juntamente com suas correspondentes definições de intensidade de primeira e segunda ordens. Ele propõe estimadores das intensidades espaço-temporais de primeira ordem usando a técnica de densidades de *kernel*.

Nas aplicações desta tese, os eventos serão analisados com a informação de espaço e de tempo. Alguns estudos agregam a informação do tempo, ou seja, analisam apenas a informação da localização espacial do eventos, como nos modelos do Capítulo 3. Outros estudos observam o tempo de ocorrência sem observar a localização espacial, ou seja, fazem a agregação no espaço. Paez e Diggle (2006), por exemplo, usam processos dinâmicos para modelar processos de Cox agregados no espaço. Gamerman (1992) apresenta um modelo dinâmico para análise estatística em processos pontuais com eventos registrados apenas no tempo e informação de covariáveis. A intensidade do processo é assumida constante em cada um dos intervalos de tempo e a inferência bayesiana é feita através de uma análise seqüencial da informação nestes intervalos sucessivos.

2.6 Inferência via Discretização no Espaço e/ou Tempo

Os modelos para processos pontuais estudados nesta tese são definidos em espaço contínuo. Entretanto, a inferência via verossimilhança é muito difícil de ser feita com espaço contínuo. Uma solução é a "discretização espacial". A região de estudo é dividida por uma partição $\mathcal{G}_{\mathcal{N}} = \{c_1, ..., c_N\}$, na qual cada célula $c_i,\ i = 1, ..., N$, tem centróide com coordenadas s_i. A variável aleatória passa a ser a contagem de eventos $Y_{[i]}$ ocorridos na i-ésima célula.

Este procedimento é adotado, por exemplo, em Møller *et al.* (1998), Brix e Møller (2001) e Beněs *et al.* (2002), nos quais o campo gaussiano é aproximado por uma *step function*, obtida via discretização da região espacial em uma grade, para que então o cálculo de sua distribuição a posteriori possa ser aproximado por um método MCMC.

A definição da partição no espaço pode ser feita de várias maneiras. Uma delas é sobrepor na região de estudo uma grade regular, como exemplificado na Figura 2.2. A região "discretizada" é constituída da união das células obtidas pela interseção da região original com a grade. Deve-se notar que as células das bordas da região original terão área menor do que a células centrais, o que deve ser incorporado no modelo.

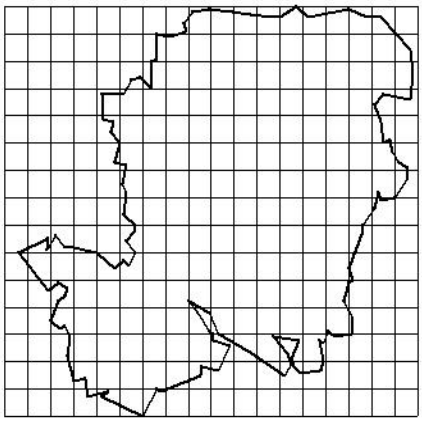


Figura 2.2: Exemplo de construção de uma grade regular sobreposta à região de estudo.

Nos arranjos pontuais espaciais com forte agregação dos eventos, o uso de uma grade regular na discretização parece ineficaz devido à criação de um grande número de células sem registro de eventos e outras, do mesmo tamanho, mas com um grande número de eventos. Uma solução seria construir uma grade com células de tamanho menor nas áreas de mais alta ocorrência de eventos, o que tornaria mais refinada a estimação da intensidade nestas regiões.

Heikkinen e Arjas (1998 e 1999), por exemplo, utilizam uma partição formada pelos polígonos de diferentes tamanhos obtidos na construção da tesselagem de Voronoi a partir dos eventos observados ou gerados especificamente para esta construção. Eles estudam modelos não-paramétricos para processos de Poisson não-homogêneos nos quais a função de intensidade é assumida constante nos polígonos.

A tesselagem de Voronoi pode ser informalmente definida do seguinte modo. Dados n pontos distintos em uma região planar S, pode-se atribuir a cada ponto s_i um polígono consistindo da parte de S que é mais próxima de s_i do que de qualquer outro dos $n-1$ pontos. Este conjunto de polígonos é chamado de tesselagem de Voronoi (ou Dirichlet).

A Figura 2.3 mostra a discretização por uma grade regular e via tesselagem de Voronoi para com um arranjo pontual fictício. Os polígonos de Voronoi são menores nas áreas com mais alta intensidade de eventos, o que certamente contribui para obtenção de um mapa de intensidades estimadas mais refinado nestas áreas. Entretanto, esta construção resulta que todas as células da discretização têm apenas um evento. A informação sobre a intensidade do processo, que na grade regular cabia à *contagem de eventos por célula*, torna-se a *área da célula polígono*.

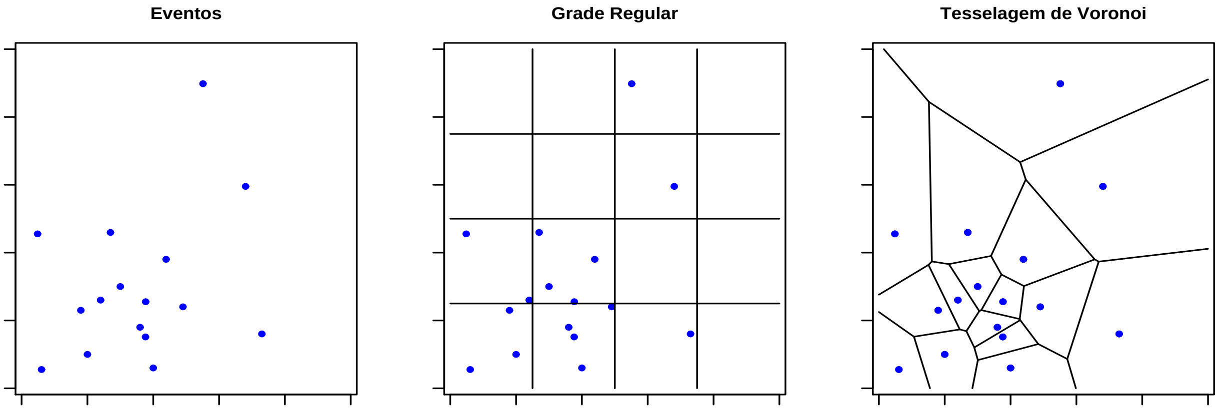


Figura 2.3: Exemplo de construção da tesselagem de Voronoi. Da esquerda para a direita: eventos observados na região, grade regular sobreposta à região, tesselagem de Voronoi sobreposta à região.

O mesmo tipo de raciocínio se aplica à discretização no tempo. O período de observação no tempo é dividido em intervalos. Registra-se, então, $Y_{[i,t]}$, o número de eventos na i-ésima célula no t-ésimo intervalo de tempo. A decisão de se criar intervalos de tempo equiespaçados ou não pode pode depender da agregação dos eventos no tempo. Se os eventos têm uma forte agregação em alguns períodos de tempo, o uso de intervalos equiespaçados pode gerar intervalos com poucos eventos observados.

Alguns estudos foram feitos para verificar o efeito da discretização no tempo ou no espaço. Paez e Diggle (2006) estudam as conseqüências de se trabalhar com diferentes níveis de discretização temporal no processo pontual de Cox agregado no espaço e concluem que, quanto maior a discretização dos dados no tempo, maior é a variância a posteriori dos parâmetros de variância e correlação temporal do processo, não afetando, entretanto, a estimação da média do processo. Considerando apenas a observação dos eventos no espaço, Waagepetersen (2004) demonstra analiticamente que as posterioris aproximadas das log-intensidades, calculadas a partir dos processos de Cox log-gaussianos discretizados, convergem para as posterioris exatas quando as áreas das células da grade tendem a zero. No Capítulo 3 é apresentado um estudo de simulação mostrando o efeito de diferentes escalas de discretização para o modelo apresentado naquele capítulo.

Capítulo 3

Modelos para Processos Pontuais Espaciais

3.1 Introdução

Neste capítulo, é estudado um modelo para a intensidade do processo pontual em eventos observados apenas no espaço, sem informação de tempo de ocorrência.

Os livros clássicos sobre processos pontuais espaciais usualmente lidam com arranjos pontuais relativamente pequenos, nos quais os métodos não paramétricos baseados em estatísticas descritivas têm um importante papel na análise. Nos últimos anos, os avanços dos métodos estatísticos computacionais, particularmente o MCMC, tiveram um grande impacto no desenvolvimento dos procedimento de inferência para processos pontuais espaciais. O foco mudou para inferência baseada na verossimilhança em modelos paramétricos, freqüentemente dependendo de covariáveis, e muitas vezes contendo também efeitos aleatórios.

Baddeley *et al.* (2006) reúne diversos estudos de casos em modelagem de processos espaciais, bem como os mais recentes avanços teóricos e metodológicos na teoria de processos pontuais no espaço. No contexto de mapeamento de doenças, são interessantes os estudos de Diggle (2000) e Richardson (2003).

O modelo estudado neste capítulo foi proposto por Beněs *et al.* (2002) no contexto de mapeamento do risco de doenças. Considera-se a situação na qual toda estrutura de dependência espacial é devida apenas à heterogeneidade espacial na intensidade do processo, e não à interação direta entre os eventos. Os exemplos de tais processos pontuais são o processo de Poisson não-homogêneo e o processo (log-gaussiano) de Cox, introduzidos no Capítulo 2.

Este capítulo está organizado do seguinte modo. Na próxima seção é apresentado o modelo espacial que servirá de base para a modelagem espaço-temporal apresentado no Capítulo 4. Na Seção 3.3 são apresentados os aspectos computacionais da inferência bayesiana implementada para o modelo. A Seção 3.4 apresentada estudos de simulação preliminares dos modelos e métodos de estimação. Nas seções finais são apresentados estudos de simulação para investigar o efeito da discretização espacial necessária para a condução da inferência e da utilização de prioris de referência. Os desenvolvimentos algébricos e estudos de simulação deste capítulo serviram como ponto de partida para a inferência nos modelos espaço-temporais do próximo capítulo.

3.2 Modelo Espacial

Seja Z um processo pontual espacial definido em $S \in \Re^2$, uma subregião no plano, para o qual considera-se o seguinte modelo hierárquico:

- No primeiro nível, Z é assumido ser um processo de Cox com função de intensidade Λ, modelada pelo produto da *intensidade da população* $\lambda_0(s)$ (ou qualquer outra variável determinística da intensidade) e da *função de risco* $\lambda(s)$:

$$\Lambda(s) = \lambda_0(s)\lambda(s), \quad \forall s \in S. \tag{3.1}$$

- No segundo nível, é proposto um modelo log-linear para a função de risco $\lambda(s)$ que incorpora informações de covariáveis espaciais $\boldsymbol{d}(s)$ e um processo Gaussiano estacionário e isotrópico $\phi(s)$:

$$log\lambda(s) = \phi(s) + \boldsymbol{\beta}'\boldsymbol{d}(s), \quad \text{com} \quad \phi(\cdot) \sim PG\left[0; \sigma^2; \rho(\cdot;\theta)\right], \tag{3.2}$$

  no qual $\boldsymbol{\beta}$ é o vetor de coeficientes de regressão desconhecidos, σ^2 é a variância do processo espacial e $\rho(h;\theta)$ é uma função de correlação espacial com parâmetro θ que depende apenas da distância h entres as localizações no espaço. Os efeitos espaciais $\phi(s)$ levam em conta a variação espacial não-explicada pelas covariáveis e a incerteza na estimação da intensidade da população.

- No terceiro nível, é escolhida uma distribuição de probabilidade a priori para os parâmetros desconhecidos do estágio anterior, denotada por

$$\pi(\boldsymbol{\beta}, \sigma^2, \theta). \tag{3.3}$$

O caso especial do modelo sem covariáveis, no qual $\boldsymbol{\beta}'\boldsymbol{d}(s) = \beta_0$, é a própria definição do modelo de Cox log-gaussiano. As covariáveis podem estar associadas diretamente ao evento observado no ponto s, como, por exemplo, características pessoais de indivíduos identificados como casos de uma doença, ou indiretamente ao próprio local de observação, como, por exemplo, uma característica ambiental.

Este modelo de efeitos fixos pode ser estendido para um modelo de efeitos aleatórios com variação no espaço, ou seja,

$$\boldsymbol{\beta}'(s)\boldsymbol{d}(s)$$

Esta idéia é explorada nos modelos de Assunção *et al.* (1999) e Gamerman *et al.* (2003) para dados de área e de Gelfand *et al.* (2003) e Paez *et al.* (2004) para dados geoestatísticos.

Seja o arranjo espacial observado $\boldsymbol{x} = (x_1, \ldots, x_n)$, com $x_i,\ i = 1, ...n$, representando as coordenadas espaciais do i-ésimo evento. A distribuição gaussiana de $\phi(\cdot)$ é vista como uma priori e a distribuição condicional de Z dados $(\phi(\cdot), \boldsymbol{\beta}, \sigma^2, \theta)$ como a verossimilhança, que, dados $\phi(\cdot)$ e $\boldsymbol{\beta}$, não depende de σ^2 e θ:

$$p(\boldsymbol{x} \mid \phi(\cdot), \boldsymbol{\beta}) \propto \exp\left\{ -\int_S \lambda_0(s) \exp[\phi(s) + \boldsymbol{\beta}'\boldsymbol{d}(s)]\, ds \right\} \prod_{z \in \boldsymbol{x}} \lambda_0(z) \exp\left[\phi(z) + \boldsymbol{\beta}'\boldsymbol{d}(z)\right]. \tag{3.4}$$

Como mencionado na Seção 2.4.2, esta verossimilhança não é analiticamente tratável, pois depende de um número infinito de variáveis aleatórias $\{\phi(s),\ s \in S\}$.

A distribuicao a posteriori de $\phi(\cdot)$ é resultante da combinação, via teorema de Bayes, de um processo Gaussiano como priori para $\phi(\cdot)$ com sua verossimilhança em (3.4).

3.3 Aspectos Computacionais da Inferência

Assim como Beněs *et al.* (2002), Møller *et al.* (1998) e Brix e Møller (2001), para viabilizar a inferência do modelo (3.1)-(3.3), é adotado o procedimento de discretização espacial do processo: a região de estudo é dividida por uma partição $\mathcal{G}_\mathcal{N} = \{c_1, ..., c_N\}$, na qual cada célula $c_i,\ i = 1, ..., N$, tem centróide com coordenadas s_i e área a_i (que incorpora também a densidade populacional). As variáveis aleatórias passam a ser as contagens de eventos $Y_{[i]},\ i = 1, ..., N$, ocorridos nas N células da partição. O modelo espacial discretizado assume que, condicionalmente à intensidade do processo nas células, estas contagens são independentes, levando a

$$\begin{aligned}
p(y_{[i]}\,|\,\lambda_{[i]}) &\propto \exp\{-a_i\lambda_{[i]}\}\cdot\lambda_{[i]}^{y_{[i]}}, & i=1,\ldots,N,\\
\log(\lambda_{[i]}) &= \boldsymbol{\beta}'\boldsymbol{d}_{[i]} + \phi_{[i]}, & i=1,\ldots,N,\\
\boldsymbol{\phi} = (\phi_{[1]}, ..., \phi_{[N]})' &\sim N\left[\underline{0}\,;\sigma^2 R_\theta\right],
\end{aligned}$$

no qual $\boldsymbol{d}_{[i]} = (d_{1[i]}, ..., d_{K[i]})'$ é o vetor de covariáveis associadas à i-ésima célula, $\boldsymbol{\beta} = (\beta_1, ..., \beta_K)'$ é o vetor de coeficientes de regressão desconhecidos, $\underline{0}$ é um vetor de comprimento N com elementos iguais a zero e $R_\theta = [R_{i,j}]_{\{i,j=1,\ldots,N\}}$ é a matriz $N\times N$ de correlações espaciais entre as células, com $R_{i,j} = \rho(\|s_i - s_j\|;\theta),\ i,j = 1, ..., N$, para uma função de correlação espacial $\rho(\cdot)$ apropriadamente escolhida.

O objetivo é inferir sobre o vetor dos efeitos espaciais $\boldsymbol{\phi}$, seus hiperparâmetros σ^2 e θ, e sobre o vetor dos coeficientes de regressão $\boldsymbol{\beta}$. A função de verossimilhança é proporcional a

$$l(\boldsymbol{\phi},\boldsymbol{\beta};\boldsymbol{y}) = \prod_{i=1}^{N} p(y_{[i]}\,|\,\phi_{[i]},\boldsymbol{\beta}) \;\propto\; \exp\left\{\sum_{i=1}^{N}\left[-a_{[i]}e^{\boldsymbol{\beta}'\boldsymbol{d}_{[i]}+\phi_{[i]}} + y_{(i)}(\boldsymbol{\beta}'\boldsymbol{d}_{[i])} + \phi_{[i]})\right]\right\}$$

e a distribuição a priori por

$$\pi(\boldsymbol{\phi},\boldsymbol{\beta},\sigma^2,\theta) \;=\; \pi(\boldsymbol{\beta})\,\pi(\boldsymbol{\phi}\,|\,\sigma^2,\theta)\,\pi(\sigma^2,\theta),$$

assumindo-se que os coeficientes de regressão $\boldsymbol{\beta}$ e os hiperparâmetros σ^2 e θ são independentes a priori. Assim, a densidade a posteriori conjunta dos efeitos aleátorios e parâmetros do modelo é

$$p(\boldsymbol{\phi},\boldsymbol{\beta},\sigma^2,\theta\,|\,\boldsymbol{y}) \;\propto\; l(\boldsymbol{\phi},\boldsymbol{\beta};\boldsymbol{y})\,\pi(\boldsymbol{\beta})\,\pi(\boldsymbol{\phi}\,|\,\sigma^2,\theta)\,\pi(\sigma^2)\,\pi(\theta).$$

Esta distribuição de densidade não pertence a uma família de distribuições conhecidas, qualquer que seja a forma funcional das prioris. Desse modo, a inferência sobre os efeitos aleatórios e demais parâmetros do modelo é feita através de uma amostra desta distribuição a posteriori, que será obtida através dos amostradores de Gibbs e de Metropolis-Hastings. Os detalhes do procedimento de amostragem são mostrados nas subseções a seguir.

A distribuição a posteriori da log-intensidade pode então ser computada usando-se métodos

MCMC. Beněs *et al.* (2002), Møller *et al.* (1998) e Brix e Møller (2001) usam o amostrador de Gibbs e o algoritmo de Metropolis-Hastings para amostrar da posteriori dos hiperparâmetros e o algoritmo de Langevin-Hastings para amostrar da posteriori do vetor de log-intensidades. Nesta tese, será usado o amostrador de Gibbs e o algoritmo de Metropolis-Hastings para amostrar os hiperparâmetros e os efeitos espaciais, como descrito a seguir.

3.3.1 Amostragem dos Efeitos Espaciais

A distribuição condicional completa do vetor de efeitos espaciais $\boldsymbol{\phi} = (\phi_{[1]}, ..., \phi_{[N]})'$ é dada por

$$p_c(\boldsymbol{\phi}\,|\,\boldsymbol{y}) \;\propto\; l(\boldsymbol{\phi},\boldsymbol{\beta};\boldsymbol{y})\,\pi(\boldsymbol{\phi}\,|\,\sigma^2,\theta) \;\propto\; \exp\left\{-\boldsymbol{A'e^{\phi}} + \boldsymbol{y'\phi} - \frac{\boldsymbol{\phi}' R_\theta^{-1}\boldsymbol{\phi}}{2\sigma^2}\right\},$$

com $\boldsymbol{A} = \left(a_1 e^{\boldsymbol{\beta d}_{[1]}}, ..., a_N e^{\boldsymbol{\beta d}_{[N]}}\right)'$ e $\boldsymbol{e^{\phi}} = \left(e^{\phi_{[1]}}, ..., e^{\phi_{[N]}}\right)'$. Esta distribuição não conjuga com a distribuição a priori Normal multivariada de $\boldsymbol{\phi}$, não pertence a uma família de densidades conhecida e tem difícil amostragem direta.

Inicialmente, foi experimentada a aplicação do esquema Metropolis-Hastings para amostragem conjunta dos elementos $\phi_{[1]}, ..., \phi_{[N]}$. Entretanto, as diversas propostas de densidades tentadas nos estudos de simulação resultaram em um número muito baixo, por vezes nulo, de valores propostos aceitos, mesmo para um grande número de iterações. Também foi experimentada a amostragem de blocos destes ϕ´s, sem sucesso mesmo para blocos pequenos, com apenas quatro elementos. Mais informações sobre estas tentativas são encontradas no Capítulo 7.

Decidiu-se, desse modo, fazer a amostragem de cada $\phi_{[i]},\ i = 1, ..., N$, individualmente. Defina $\phi_{[-i]} = (\phi_{[1]}, ..., \phi_{[i-1]}), \phi_{[i+1]}, ..., \phi_{[N]})'$, o vetor das log-intensidades excluída aquela da célula i. A distribuição a priori condicional completa de $\phi_{[i]}$ é dada por

$$\phi_{[i]}\,|\,\phi_{[-i]}, \sigma^2, \theta \;\sim\; N\left[M_i; V_i\right], \qquad i = 1, ..., N,$$
$$\text{com}\;\; M_i = B_i' H_i^{-1}\phi_{[-i]} \;\;\text{e}\;\; V_i = \sigma^2(1 - B_i' H_i^{-1} B_i),$$

onde H_i é a matriz de correlações R_θ extraídas as i-ésimas linha e coluna, B_i é o vetor formado pela i-ésima linha de R_θ sem a i-ésima coluna, para $i = 1, ..., N$.

A distribuição condicional completa da log-intensidade $\phi_{(i)}$ é dada por:

$$
\begin{aligned}
p_c(\phi_{[i]}\,|\,\phi_{[-i]},\boldsymbol{\beta},\psi,y_{[i]}) \;&\propto\; p(y_{[i]}\,|\,\phi_{[i]},\boldsymbol{\beta})\;\pi(\phi_{[i]}\,|\,\phi_{[-i]},\sigma^2,\theta), \qquad\qquad i=1,...,N,\\
&\propto\; \exp\left\{-a_i e^{\boldsymbol{\beta' d}_{[i]}+\phi_{[i]}} + y_{[i]}\phi_{[i]} - \frac{(\phi_{[i]}-M_i)^2}{2V_i}\right\}.
\end{aligned}
$$

O esquema de Metropolis-Hastings é usado para amostrar desta distribuição. Apresentamos duas propostas de densidades para amostragem: a *proposta da priori condicional* e a *proposta da posteriori de modelos lineares generalizados mistos (MLGM)*.

Na proposta da priori condicional, um novo valor $\phi^N_{[i]},\ i=1,...,N$, é amostrado da densidade a priori condicional aos valores correntes das demais log-intensidades e dos hiperparâmetros, ou seja, da distribuição Normal com média M_i e variância V_i. A probabilidade de aceitação do valor proposto é igual a $min\{1,\alpha_1(\phi_{[i]})\}$, com

$$
\alpha_1(\phi_{[i]}) = \exp\left\{-a_{[i]}e^{\beta d_{[i]}}\left(e^{\phi^N_{[i]}} - e^{\phi^V_{[i]}}\right) + \left(\phi^N_{[i]} - \phi^V_{[i]}\right) y_{[i]}\right\}, \qquad i=1,...,N.
$$

A densidade proposta da posteriori MLGM para $\phi_{[i]}$ é aquela apresentada em Gamerman (1997) e detalhada no Apêndice A.3. Nela, o novo valor $\phi^N_{[i]}$ é amostrado da distribuição Normal com média M^V_ϕ e variância V^V_ϕ, com

$$
\begin{aligned}
M^V_\phi \;&=\; V^V_\phi\left(M_i/V_i + \tilde{y}_{[i]}/\tilde{V}^V_i\right) \qquad \text{e} \quad V^V_\phi = \left(1/V_i + 1/\tilde{V}^V_i\right)^{-1}, \\
\tilde{y}^V_{[i]} \;&=\; \phi^V_{[i]} + \frac{y_{[i]} - a_i e^{\phi^V_{[i]}+\beta d_{[i]}}}{a_i e^{\phi^V_{[i]}+\beta d_{[i]}}} \qquad \text{e} \quad \tilde{V}^V_i = \left(a_i e^{\phi^V_{[i]}+\beta d_{[i]}}\right)^{-1},
\end{aligned}
\tag{3.5}
$$

e probabilidade de aceitação do valor proposto é igual a $min\{1,\alpha_2(\phi_{[i]})\}$, na qual

$$
\begin{aligned}
\alpha_2(\phi_{[i]}) \;&=\; \exp\left\{-a_i e^{\beta d_{[i]}}\left(e^{\phi^N_{[i]}} - e^{\phi^V_{[i]}}\right) + \left(\phi^N_{[i]} - \phi^V_{[i]}\right) y_{[i]}\right\} \times \\
&\times\; \exp\left\{-\frac{(\phi^V_{[i]} - M^N_\phi)^2}{2V^N_\phi} + \frac{(\phi^N_{[i]} - M^V_\phi)^2}{2V^V_\phi}\right\}\left(\frac{V^N_\phi}{V^V_\phi}\right)^{-1/2}, \qquad i=1,...,N,
\end{aligned}
$$

na qual M^N_ϕ e V^N_ϕ são dados pelas expressões em (3.5) substituindo-se $\phi^V_{[i]}$ por $\phi^N_{[i]}$.

3.3.2 Amostragem do Coeficiente de Regressão

Os coeficientes de regressão $\beta_k,\ k\!=\!1,...,K$, são assumidos independentes entre si a priori, com distribuições a priori marginais Normais:

$$\beta_k \sim N\left[a_{\beta_k}; b^2_{\beta_k}\right], \quad k = 1, ..., K.$$

Defina $\boldsymbol{\beta}_{-k}\!=\!(\beta_1, ..., \beta_{k-1}, \beta_{k+1}, ..., \beta_K)'$, ou seja, o vetor $\boldsymbol{\beta}$ dos coeficientes de regressão excluído β_k; e $\boldsymbol{d}_{-k[i]} = (d_{1[i]}, ..., d_{k-1[i]}, d_{k+1[i]}, ..., d_{K[i]})'$, o vetor das variáveis explicativas para a i-ésima célula da partição espacial.

A distribuição condicional completa de β_k é dada por

$$\begin{aligned} p_c(\beta_k \,|\, \boldsymbol{\beta}_{-k}, \boldsymbol{\phi}, \sigma^2, \theta, \boldsymbol{y}) &\propto l(\boldsymbol{\phi}, \boldsymbol{\beta}; \boldsymbol{y})\ \pi(\beta_k) \\ &\propto \exp\left\{ -\sum_{i=1}^{N} a_i e^{\boldsymbol{\beta}' \boldsymbol{d}_{[i]} + \phi_{[i]}} + \beta_k \sum_{i=1}^{N} y_{[i]} d_{k[i]} - \frac{(\beta_k - a_{\beta_k})^2}{2 b^2_{\beta_k}} \right\}, \end{aligned}$$

que não pertence a uma família de distribuições conhecida. Desse modo, a amostragem de β_k é feita através do esquema de Metropolis-Hastings: a cada iteração da cadeia do MCMC, um novo valor para β_k (denotado β_k^N) é amostrado, em função do valor da iteração anterior (denotado β_k^V), de uma distribuição proposta Normal $q(\beta_k^N \mid \beta_k^V)$ com média β_k^V e variância w_k (que também é a constante sintonizadora da taxa de aceitação dos valores propostos). Aceita-se este valor proposto com probabilidade igual a $min\{1, \alpha(\beta_k)\}$, na qual

$$\begin{aligned} \alpha(\beta_k) &= \frac{p_c(\beta_k^N \mid \boldsymbol{\phi}, \boldsymbol{y})}{p_c(\beta_k^V \mid \boldsymbol{\phi}, \boldsymbol{y})} \times \frac{q(\beta_k^V \mid \beta_k^N)}{q(\beta_k^N \mid \beta_k^V)} = \exp\{-\sum_{i=1}^{N} A_i (e^{\beta_k^N d_{k[i]}} - e^{\beta_k^V d_{k[i]}}) + \\ &+ (\beta_k^N - \beta_k^V) \sum_{i=1}^{N} y_{[i]} d_{k[i]} - \frac{(\beta_k^N - a_{\beta_k})^2 - (\beta_k^V - a_{\beta_k})^2}{2 b^2_{\beta_k}}\}, \end{aligned}$$

com $A_i = a_i e^{\phi_{[i]} + \beta_{-k[i]} d_{-k[i]}}$. O processo de amostragem continua até que se obtenha convergência da cadeia.

O procedimento de amostragem descrito acima, no qual cada coeficiente β_k é amostrado individualmente, pode ser modificado para que $\beta_1, ..., \beta_K$ sejam amostrados conjuntamente.

3.3.3 Amostragem dos Parâmetros do Processo Espacial

Os parâmetros σ^2 e θ, relacionados à distribuição dos efeitos espaciais $\boldsymbol{\phi}$, são assumidos independentes a priori com distribuições marginais gama invertida e gama, respectivamente, e denotadas por

$$\sigma^2 \sim GI[g_s; v_s] \qquad \text{e} \qquad \theta \sim G\,[g_t; v_t].$$

A parametrização da densidade $G[g;v]$ é tal que $f(x) = v^g/\Gamma(g)\ x^{g-1}e^{-vx},\ x > 0,\ g > 0,\ v > 0$, que tem esperança igual a g/v e variância igual a g/v^2. Na densidade $GI[g;v]$, a parametrização é tal que $f(x) = v^g/\Gamma(g)\ x^{-(g+1)}e^{-v/x},\ x > 0,\ g > 0,\ v > 0$, que tem esperança igual a $v/(g-1)$, se $g > 1$, e variância igual a $v^2/[(g-1)^2(g-2)]$, se $g > 2$.

O amostrador de Gibbs foi escolhido para atualização dos valores destes parâmetros individualmente. O parâmetro σ^2 possui densidade condicional completa Gama Invertida:

$$\begin{aligned} p_c(\sigma^2 \,|\, \theta, \boldsymbol{\phi}, \boldsymbol{\beta}, \boldsymbol{y}) \ &\propto\ \pi(\boldsymbol{\phi} \,|\, \sigma^2, \theta)\ \pi(\sigma^2) \\ &\propto\ (\sigma^2)^{-\left(g_s + \frac{N}{2} + 1\right)} \exp\left\{ -\frac{1}{\sigma^2} \left[v_s + \frac{\boldsymbol{\phi}' R_\theta^{-1} \boldsymbol{\phi}}{2} \right] \right\}, \quad \sigma^2 > 0, \\ \Leftrightarrow \quad \sigma^2 \mid \theta, \boldsymbol{\phi} \ &\sim\ GI\left[g_s + \frac{N}{2}\ ;\ v_s + \frac{\boldsymbol{\phi}' R_\theta^{-1} \boldsymbol{\phi}}{2} \right]. \end{aligned}$$

A cadeia σ^2 é então formada pela amostragem, a cada iteração, diretamente de sua densidade condicional completa.

A densidade condicional completa de θ, entretanto, não pertence a uma família de distribuições conhecida:

$$\begin{aligned} p_c(\theta \mid \sigma^2, \boldsymbol{\phi}, \boldsymbol{\beta}, \boldsymbol{y}) \ &\propto\ \pi(\boldsymbol{\phi} \,|\, \sigma^2, \theta)\ \pi(\theta) \\ &\propto\ \theta^{g_t - 1} |R_\theta|^{-\frac{1}{2}}\ \exp\left\{ -v_t\theta - \frac{\boldsymbol{\phi}' R_\theta^{-1} \boldsymbol{\phi}}{2\sigma^2} \right\}, \quad \theta > 0. \end{aligned}$$

Desse modo, é utilizado o esquema de amostragem de Metropolis-Hastings: a cada iteração da cadeia, um novo valor para θ (denotado θ^N) é amostrado, em função do valor da iteração anterior (denotado θ^V), da densidade proposta $q(\theta^N|\theta^V)$ log-Normal com parâmetros $log\theta^V - \frac{w_\theta}{2}$ e w_θ, de modo que seu valor esperado é $E(\theta^N \,|\, \theta^V) = \theta^V$ e seu coeficiente de variação é $CV(\theta^N \,|\, \theta^V) = (e^{w_\theta} - 1)^{1/2}$, onde w_θ é a constante sintonizadora da taxa de aceitação de novos valores. Aceita-se este valor proposto com probabilidade igual a $min\{1, \alpha(\theta)\}$, na qual

$$\alpha(\theta) \;=\; \frac{p_c(\theta^N \mid \sigma^2, \boldsymbol{\phi})}{p_c(\theta^V \mid \sigma^2, \boldsymbol{\phi})} \times \frac{q(\theta^V \mid \theta^N)}{q(\theta^N \mid \theta^V)},$$

com $q(\theta^N \mid \theta^V) = (2\pi w_\theta)^{-\frac{1}{2}} \ \exp\Big\{-\frac{1}{2w_\theta}\left(log\theta^N - log\theta^V + w_\theta\right)^2\Big\}$ e $q(\theta^V \mid \theta^N)$ definido de modo análogo trocando-se θ^N por θ^V na expressão anterior, de modo que

$$\begin{aligned}
\alpha(\theta) \;&=\; \left(\frac{|R_{\theta^N}|}{|R_{\theta^V}|}\right)^{-\frac{1}{2}} \left(\frac{\theta^N}{\theta^V}\right)^{g_t} \times \\
&\times \;\exp\Big\{-\frac{1}{2w_\theta}\left[\left(log\theta^V - log\theta^N + \frac{w_\theta}{2}\right)^2 - \left(log\theta^N - log\theta^V + \frac{w_\theta}{2}\right)^2\right]\Big\} \times \\
&\times \;\exp\Big\{-v_t(\theta^N - \theta^V) - \frac{1}{2\sigma^2}\left(\boldsymbol{\phi}' R^{-1}_{\theta^N}\boldsymbol{\phi} - \boldsymbol{\phi}' R^{-1}_{\theta^V}\boldsymbol{\phi}\right)\Big\}.
\end{aligned}$$

O processo de amostragem continua até que se obtenha convergência da cadeia.

3.4 Estudos de Simulação

Nesta seção, são apresentados os resultados de dois estudos de simulação conduzidos para a verificação da eficácia de estimação da metodologia proposta no trabalho em arranjos espaciais gerados do modelo (3.5).

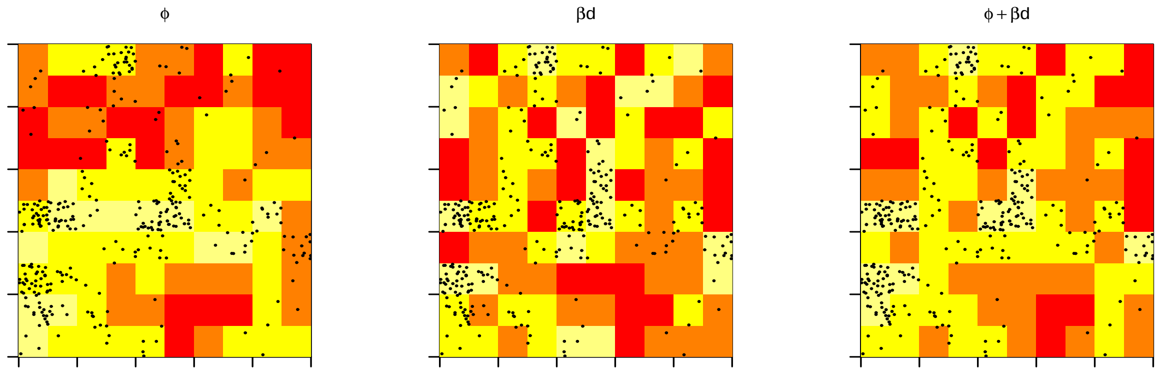
A região de simulação é o quadrado $S = [0,1] \times [0,1]$, aproximado aqui por uma grade regular com 100 células. O processo gaussiano de $\boldsymbol{\phi} = (\phi_{[1]}, ..., \phi_{[100]})'$ foi gerado com a *função de correlação espacial exponencial* $\rho(h;\theta) = exp\,(-\theta h)$ e valores escolhidos de β, σ^2 e θ em cada exemplo.

No MCMC foram geradas duas cadeias, definidas por diferentes valores iniciais, com 100 mil iterações cada. As amostras a posteriori são formadas pelos 1000 valores tomados a cada 50 das últimas 50 mil iterações.

O programa *Ox* (Doornik, 2002) foi utilizado para a codificação dos algoritmos.

Exemplo 3.1: Covariável sem estrutura espacial.

O processo de Cox log-gaussiano foi simulado com $\sigma^2 = 2$, $\theta = 4$ e $\beta = 1,5$ e valores da covariável d gerados independentemente da densidade uniforme entre 1 e 5. A Figura 3.1 mostra os eventos gerados sobrepostos às imagens de ϕ, βd e à soma destes dois termos.



$-2,6 \bullet -1 \bullet 0 \bullet 1 \bullet 2,6$ $\quad$ $1,5; \bullet 3 \bullet 4,5 \bullet 6 \bullet 7,5$ $\quad$ $-1,1 \bullet 2 \bullet 4,5 \bullet 7 \bullet 10,1$

Figura 3.1: Exemplo 3.1: Mapa dos eventos gerados sobrepostos às imagens dos valores reais dos $\phi_{[i]}$, $\beta d_{[i]}$ e $\phi_{[i]} + \beta d_{[i]}$ nas 100 células da grade.

Dois conjuntos de prioris para β, σ^2 e θ foram definidos: as chamadas *pouco informativas* (embora bem localizadas), nas quais o valor esperado e o desvio padrão são iguais ao valor real do parâmetro, o que resulta em $\beta \sim N[1,5;(1,5)^2]$, $\sigma^2 \sim GI[3;4]$ e $\theta \sim G[1;0,25]$; e as chamadas *muito informativas*, com $\beta \sim N[1,5;(0,15)^2]$, $\sigma^2 \sim GI[102;202]$ e $\theta \sim G[100;25]$, resultado da escolha da esperança igual ao valor real do parâmetro e do desvio padrão correspondente a um décimo deste valor.

A Figura 3.2 mostra os resultados da amostragem dos $\phi_{[i]}$ referentes apenas à proposta da densidade priori condicional no Metropolis-Hastings, com prioris pouco informativas, pois os resultados da proposta da posteriori MLGM e/ou prioris muito informativas são visualmente idênticos a estes. A estimação das log-intensidades $\phi_{[i]} + \beta d_{[i]}$ e dos termos $\phi_{[i]}$ isoladamente mostrou-se bastante eficaz.

Os parâmetros β, μ, σ^2 e θ também foram bem estimados, como pode ser verificado pelos histogramas de suas amostras a posteriori (Figura 3.3). Apesar da variabilidade ser muito alta nas amostras a posteriori de σ^2 e θ quando prioris pouco informativas foram escolhidas para eles, o intervalo de valores mais freqüentes em cada distribuição engloba o valor real do parâmetro.

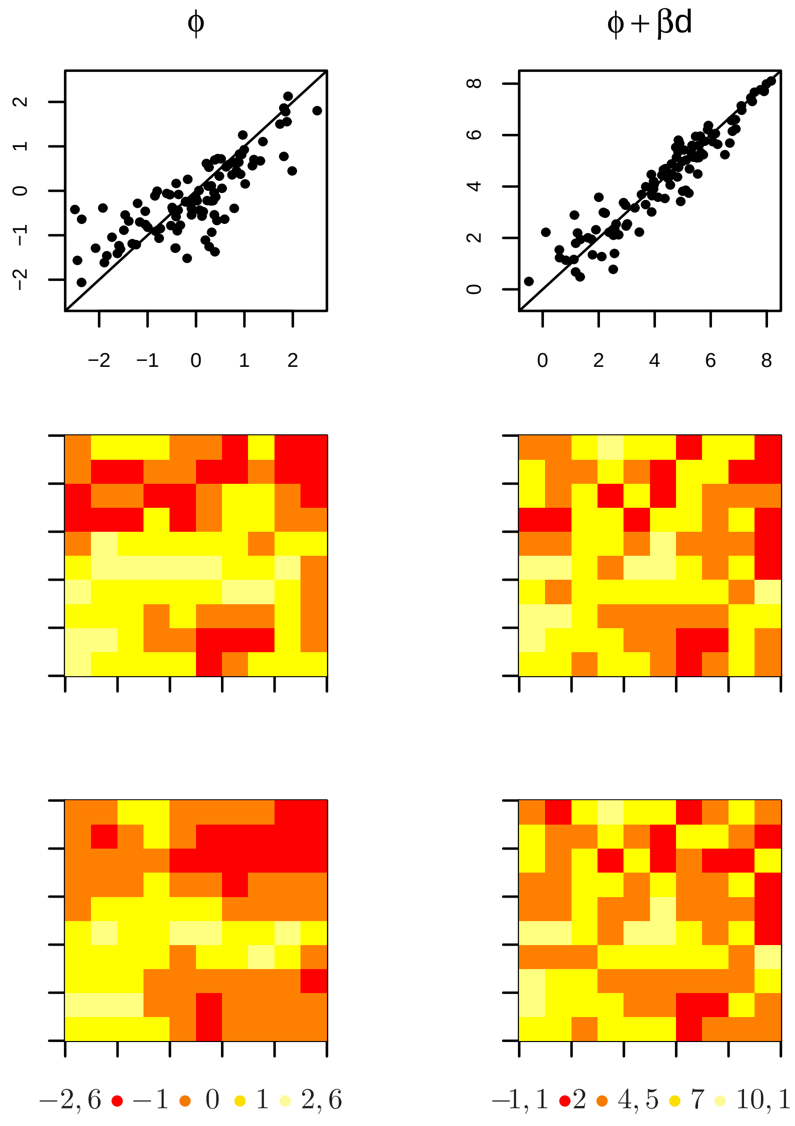


Figura 3.2: Exemplo 3.1: Resultados de estimação dos efeitos espaciais. Na primeira linha, diagramas de dispersão dos valores reais *versus* médias a posteriori dos $\phi_{[i]}$ e $\phi_{[i]} + \beta d_{[i]}$ nas 100 células da grade.

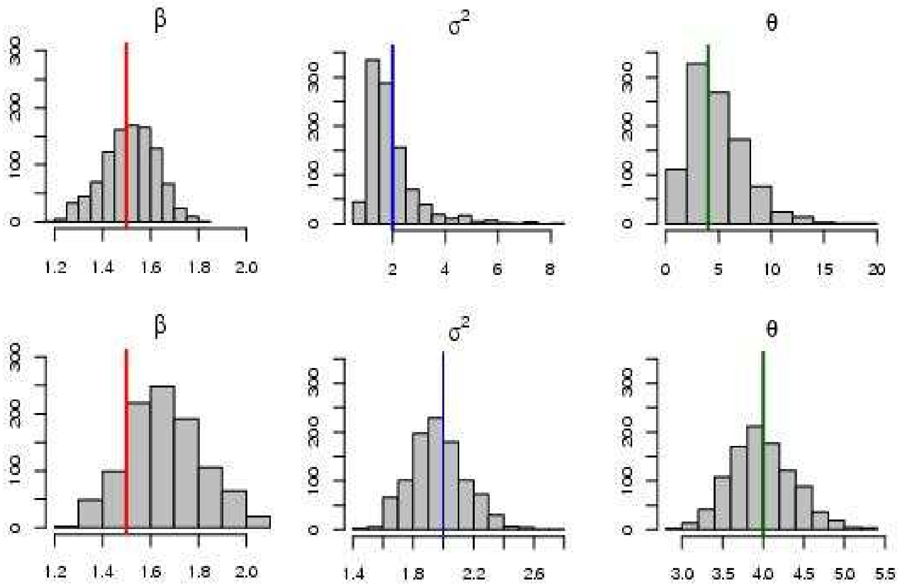


Figura 3.3: Exemplo 3.1: Histogramas das amostras a posteriori do coeficiente de regressão e dos hiperparâmetros, com prioris *pouco informativa* (primeira linha) e *muito informativa* (segunda linha). O traço vertical marca o valor real do parâmetro.

Exemplo 3.2: Covariável com forte estrutura espacial.

Na mesma situação do Exemplo 3.1, fixou-se $\sigma^2 = 2$, $\theta = 4$, $\beta = 5$ e uma covariável $d_{[i]}$ com forte estrutura espacial, como pode ser visto na Figura 3.4. Ao contrário do Exemplo 3.1, o arranjo espacial gerado reproduz mais fielmente a distribuição espacial do termo da regressão do que a distribuição espacial do processo ϕ.

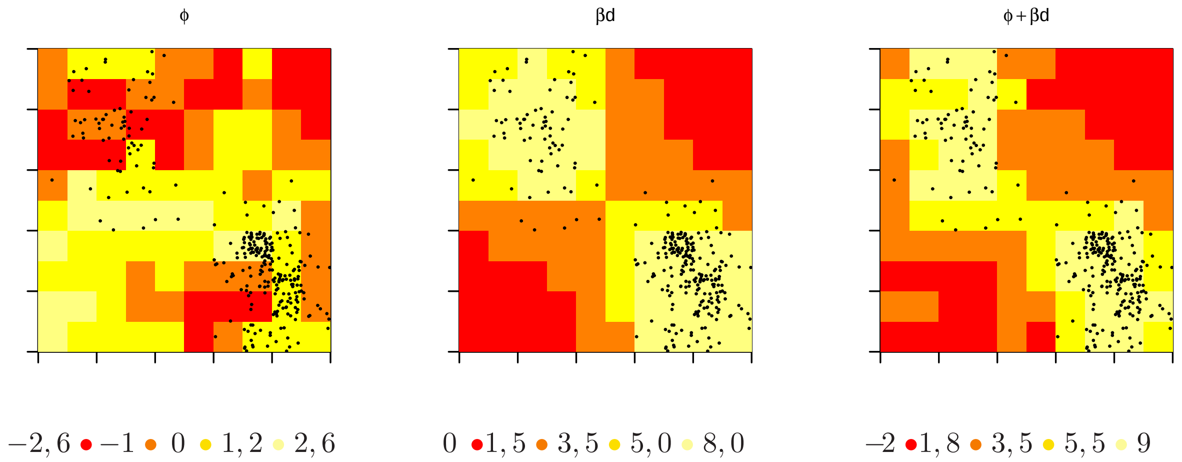


Figura 3.4: Exemplo 3.2: Mapa dos eventos gerados. Eventos gerados sobrepostos às imagens dos valores reais dos $\phi_{[i]}$, $\beta d_{[i]}$ e $\phi_{[i]} + \beta d_{[i]}$ nas 100 células da grade.

As figuras 3.5 e 3.6 mostram os resultados da amostragem dos $\phi_{[i]}$ no Metropolis-Hastings apenas da proposta da densidade priori condicional que, assim como no exemplo anterior, são idênticos aos resultados da outra proposta de amostragem. Neste exemplo, a superfície das log-intensidades $\phi_{[i]}+\beta d_{[i]}$ é melhor estimada do que a superfície dos $\phi_{[i]}$.

Na Figura 3.6, verifica-se que os parâmetros também foram bem estimados com as prioris pouco informativas escolhidas, a saber: $\beta \sim N[5;(2.5)^2]$, $\sigma^2 \sim GI[6;10]$ e $\theta \sim G[4;1]$.

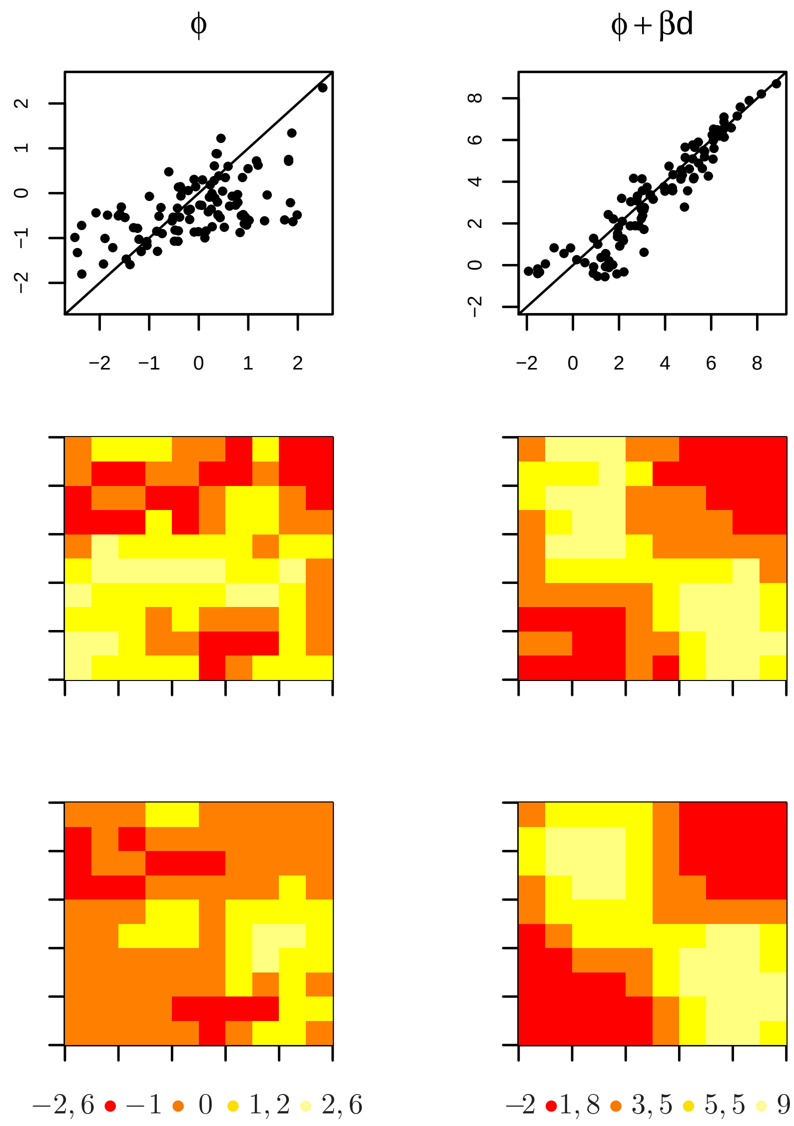


Figura 3.5: Exemplo 3.2: Resultados de estimação dos efeitos espaciais. Na primeira linha, diagramas de dispersão dos valores reais *versus* médias a posteriori dos $\phi_{[i]}$ e $\phi_{[i]}+\beta d_{[i]}$ nas 100 células da grade. Na segunda linha, imagens dos valores reais e, na terceira linha, imagens das médias a posteriori.

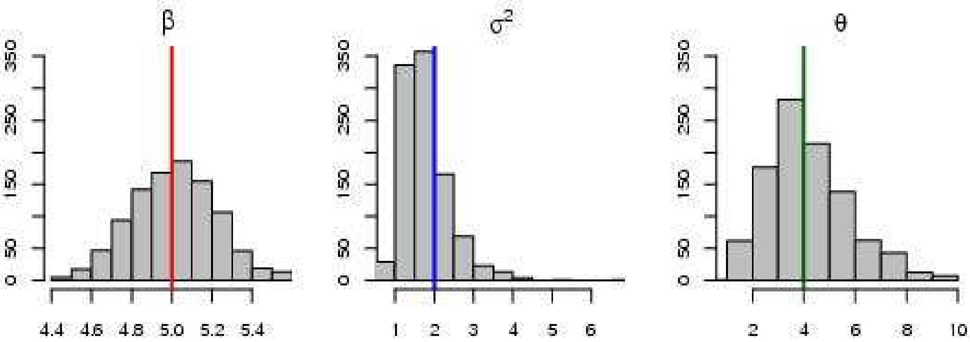


Figura 3.6: Exemplo 3.2: Histogramas das amostras a posteriori do coeficiente de regressão e dos hiperparâmetros. O traço vertical marca o valor real do parâmetro.

3.5 Prioris de Referência

Berger *et al.* (2001) propõem uma priori de referência para os parâmetros de processos espaciais em espaço contínuo modelados como processos gaussianos com função de média descrita por um modelo linear e função covariância descrita por uma função de correlação espacial (por exemplo, exponencial, esférica, etc.) com poucos parâmetros desconhecidos.

O modelo espacial para processos pontuais descrito em (3.1-3.3) tem sua função de log-intensidade descrita como um processo gaussiano desta natureza. Desse modo, a priori de referência dos hiperparâmetros deste processo no modelo discretizado $\boldsymbol{\phi} \sim N\,[X\boldsymbol{\beta};\sigma^2 R_\theta]$ poderia ser aproximada pela priori de Berger *et al.* (2001) por

$$\begin{aligned}
\pi(\beta,\sigma^2,\theta) \;&\propto\; \frac{\pi(\theta)}{\sigma^2}, \quad \text{para} \quad (\beta,\sigma^2,\theta) \;\in\; \Re^p \times (0,\infty), \times (0,\infty) \qquad (3.6)\\
\text{com} \quad \pi(\theta) \;&\propto\; \left\{ tr(W_\theta^2) - \frac{1}{n-p}\,[tr(W_\theta)]^2 \right\}^{1/2},\\
\text{na qual} \quad W_\theta^R \;&=\; S_\theta^R\, R_\theta^{-1}\, P_\theta^R, \qquad P_\theta^R \;=\; I - X(X'R_\theta^{-1}X)^{-1}X'R_\theta^{-1}
\end{aligned}$$

e $S_\theta^R = (\delta/\delta\theta)R_\theta$ denotando a matriz obtida diferenciando-se R_θ elemento por elemento.

Exemplo 3.3: O modelo espacial com $X\beta\!=\!\underline{1}\,\mu$ e função de correlação espacial exponencial, com $\mu\!=\!1$, $\sigma^2\!=\!1$ e $\theta\!=\!0,8$ foi simulado na região quadrada $[0,1]\!\times\![0,1]$ dividida em uma grade regular com 100 células. Tem-se que $R_\theta[i,j] = e^{-\theta\, d_{ij}}$ e $S_\theta^R[i,j] = -d_{ij}\; e^{-\theta\, d_{ij}}$, sendo d_{ij} a distância entre os centróides s_i e s_j das células i e j da grade.

A Figura 3.7 mostra as densidades das prioris de referência marginais de σ^2 e θ com duas especificações Gama Inversa e Gama, respectivamente. Para cada uma das especificações de prioris foram simuladas cadeias de tamanho 150 mil. As amostras finais das posterioris dos parâmetros são formadas por 1000 valores tomados a um intervalo de 50 iterações após 100 mil iterações. Os histogramas destas amostras a posteriori são mostrados na Figura 3.8. Pode-se concluir que, neste exemplo, a forma da distribuição a posteriori e o intervalo de valores de maior densidade não se modifica significativamente entre as especificações a priori.

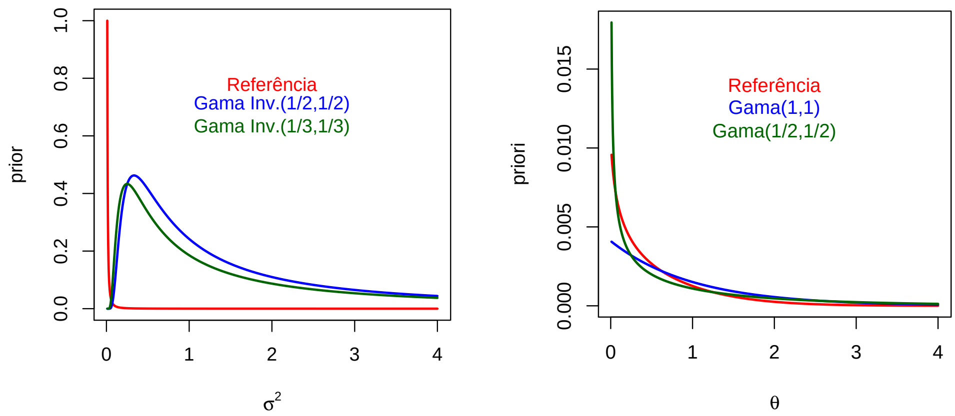


Figura 3.7: Exemplo 3.3: Especificações de prioris para σ^2 e θ.

3.6 Efeito da Discretização no Espaço

Em algumas situações pode ser necessário substituir a discretização espacial do processo, feita originalmente por uma grade com grande número de células, por outra grade com um número menor de células. Os valores gerados ou estimados da intensidade do processo para a grade maior precisam, assim, ser reproduzidos para o mapa com número menor de células. Neste caso, adotamos a média dos valores das log-intensidades ϕ nas células da grade maior que compõem cada célula da grade menor como uma estimativa para o valor de ϕ naquela nova célula. Ou seja, o valor aproximado para a intensidade $\lambda_k = log(\phi_k)$ na k-ésima célula resultante da união de n_k células da grade anterior é dado por

$$\tilde{\lambda}_k = log(\tilde{\phi}_k) = log\left(\frac{\sum_{i=1}^{n_k}\phi_{[i]}}{n_k}\right). \qquad (3.7)$$

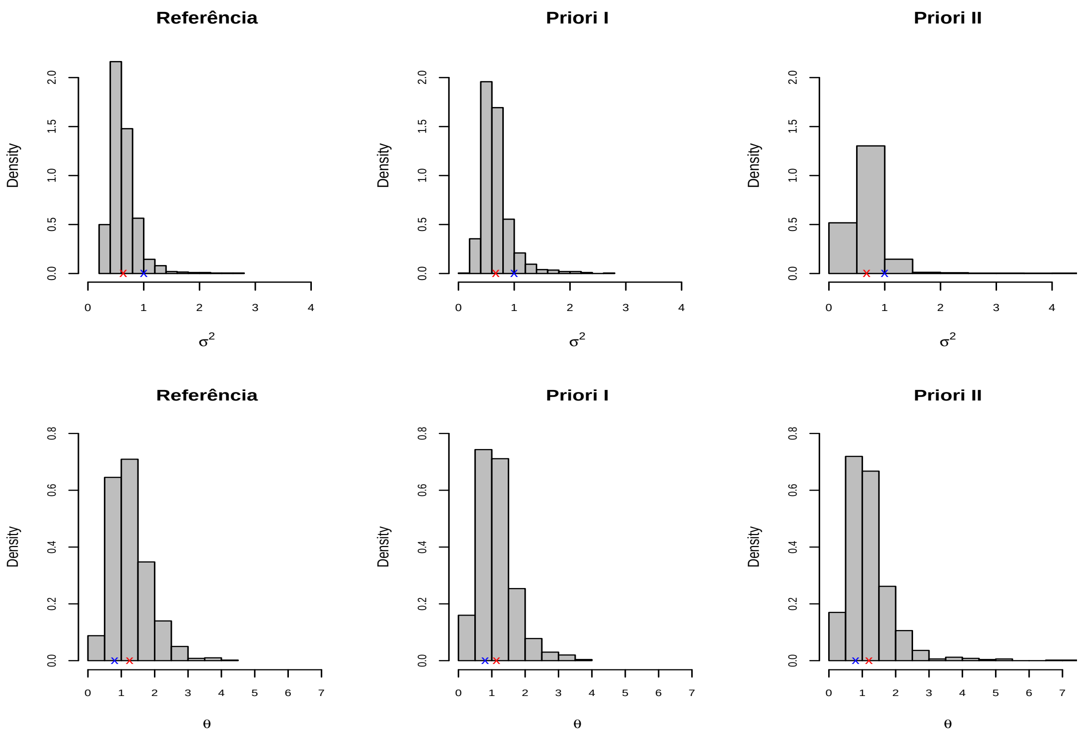


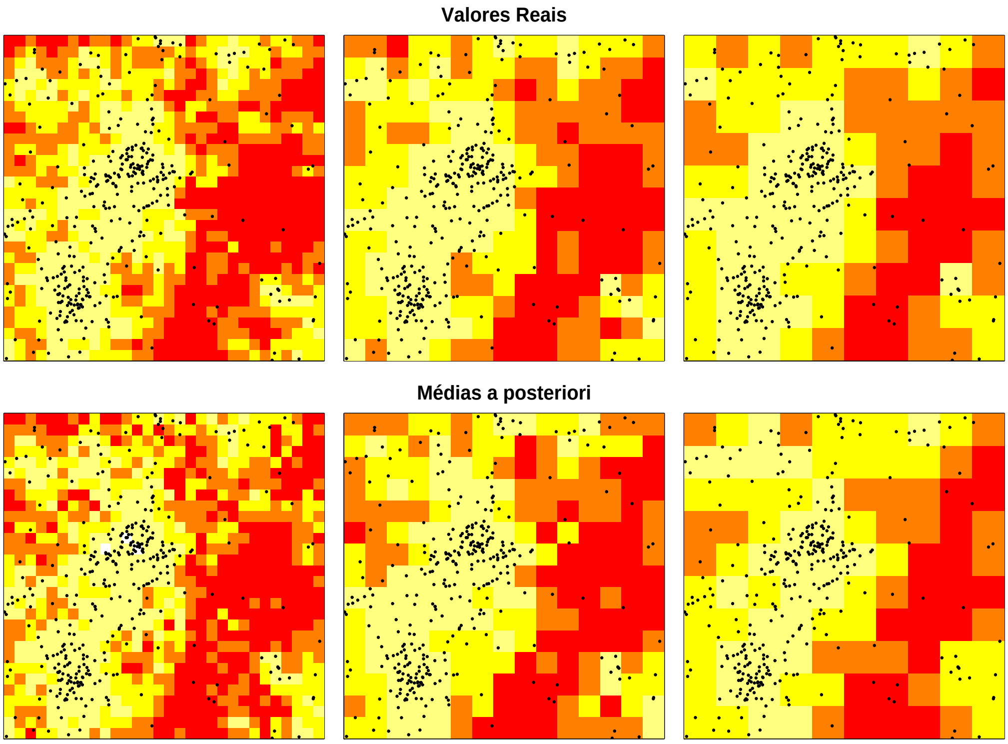
Figura 3.8: Exemplo 3.3: Histogramas das amostras da posteriori de σ^2 e θ geradas a partir das três especificações de prioris da Figura 3.7.

Com o objetivo de estudar o efeito da discretização espacial na estimação da intensidade do processo espacial latente, é descrito a seguir um estudo de simulação no qual a inferência sobre o processo espacial é feita em diferentes níveis de agregação de um processo pontual gerado a partir do nível mais refinado.

Exemplo 3.4: Um estudo de simulação foi realizado para ilustrar a aplicação da aproximação proposta em (3.7). O processo gaussiano ϕ foi gerado com $\beta = 5$ (intercepto, sem covariáveis), $\sigma^2 = 2$, $\theta = 8$ e função de correlação espacial exponencial na região quadrada $[0, 1] \times [0, 1]$ dividida por uma grade regular com 900 células.

A primeira linha de mapas da Figura 3.9 mostra estes valores gerados e os valores aproximados para as grades menores com 225 e 100 células, resultantes do agrupamento das células da grade original em grupos de quatro e nove células adjacentes, respectivamente. A aplicação de (3.7) preservou as principais características espaciais da grade original nas grades menores.

A abordagem bayesiana de inferência descrita na Seção 3.3 foi conduzida com prioris $\sigma^2 \sim GI[1, 1]$, $\theta \sim G[1, 1]$ e $\pi(\beta) \propto 1$. Foram geradas cadeias MCMC com 150 mil iterações. As amostras das posterioris são formadas por 1000 valores tomados a cada 50 das 50 mil iterações finais. As médias da distribuição a posteriori mostraram-se bastante próximas dos



$0,4 \bullet 3,8 \bullet 4,8 \bullet 5,8 \bullet 8,8 \circ 9,0.$

Figura 3.9: Exemplo 3.4: Mapas dos processos gaussianos e eventos gerados. Na primeira linha, valores reais do processo gaussiano ϕ gerados na grade com 900 células (coluna da esquerda) e aproximados para as grades com 225 células (coluna do meio) e 100 células (coluna da direita). Na segunda linha, médias das distribuições a posteriori para as respectivas grades. Os eventos gerados estão sobrepostos em todos os mapas.

valores reais (Figura 3.10) e reproduziram o padrão espacial da log-intensidade em todos os níveis de discretização (Figura 3.9).

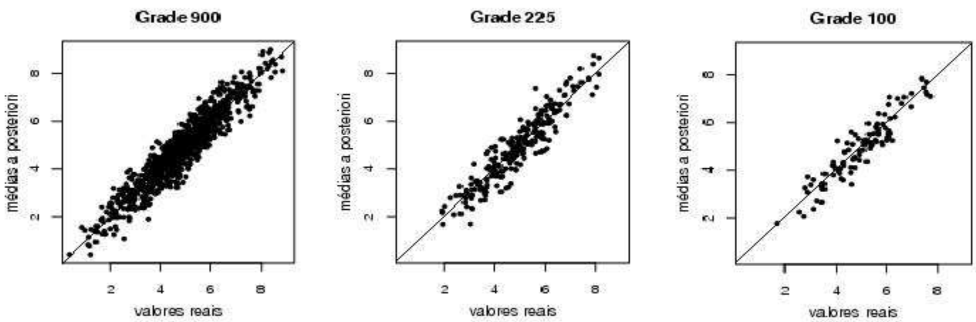


Figura 3.10: Exemplo 3.4: Resultados de estimação dos efeitos espaciais. Diagramas de dispersão dos valores reais do campo gaussiano ϕ do exemplo 1 e de suas estimativas dadas pelas médias a posteriori nas grades com 900, 225 e 100 células espaciais, respectivamente.

Capítulo 4

Modelos para Processos Pontuais Espaço-Temporais

4.1 Introdução

Um arranjo pontual espacial formado pelas localizações de eventos em uma região do $\Re^2$ é, freqüentemente, o resultado de um processo dinâmico que ocorre no tempo tanto quanto no espaço. Por exemplo, o processo pontual das localizações de certo tipo de árvore em uma floresta evolue no tempo à medida que novas árvores nascem e árvores velhas morrem.

Um processo pontual espaço-temporal poderia ser obtido a partir de qualquer processo pontual no espaço $\Re^2$ tratando o tempo como um eixo espacial adicional (ou seja, em $\Re^3$), ou ainda, tratar a dimensão temporal como a marca de um processo pontual espacial marcado. Entretanto, esta abordagem falha ao não explorar a natureza unidirecional do tempo, uma característica não encontrada no domínio espacial, no qual as dependências ocorrem em todas as direções.

Recentemente, foram propostos alguns procedimentos para análise de arranjos pontuais espaço-temporais, cada um motivado por uma aplicação particular. Uma importante distinção na prática está entre arranjos pontuais espaciais para os quais os eventos ocorrem continuamente no tempo, e aqueles para os quais a escala do tempo é genuinamente discreta ou está discretizada pelo registro agregado dos eventos em períodos de tempo. A modelagem em tempo discreto é exemplificada em Diggle *et al.* (2005a) pelo estudo dos registros anuais da distribuição espacial dos casos de tuberculose bovina em Cornwall. A modelagem em tempo contínuo é exemplificada pelo modelo de Diggle *et al.* (2005b) para os casos de doença gastrointestinal em Hampshire, no Reino Unido.

Brix e Diggle (2001) descrevem uma classe flexível de processos pontuais espaço-temporais baseada em modelos de Cox log-gaussianos. No contexto de mapeamento de doenças, a intensidade do processo no espaço e tempo é definida pela equação $\lambda(s,t)=\rho(s)\pi(s,t)$, na qual $\rho(s)$ é um processo determinístico descrevendo a variação espacial da população e $\pi(s,t)$ é a função de risco, definida por um processo espaço-temporal de Ornstein-Uhlenbeck, descrito no tempo através de equações diferenciais estocásticas. A inferência, uma tarefa difícil neste contexto, foi feita com estimadores de momentos. Nessa modelagem, tanto o espaço como o tempo são definidos como contínuos, mas tratados de forma discretizada na inferência.

Um modo intuitivamente natural de especificar um modelo espaço-temporal para um processo pontual é através de sua intensidade condicional em cada localização e tempo dada a história do processo até este tempo.

No contexto de processos espaciais contínuos, Gelfand *et al.* (2005) fazem a modelagem de fenômenos espaço-temporais através de processos gaussianos, onde o espaço é visto como contínuo (dados geoestatísticos) e o tempo é tomado como discreto. A idéia é enxergar os dados como uma série temporal de processos espaciais, adaptando o esquema de modelos dinâmicos a um modelo espaço-temporal univariado com coeficientes variando espacialmente. Assim, a variável resposta $y(s,t)$, observada na localização espacial $s\in S=\{s_1,...,s_{N_s}\}$ e no tempo $t\in T=\{t_1,...,t_{N_t}\}$, é modelada por covariáveis cujos coeficientes variam no espaço e no tempo segundo o modelo

$$\begin{aligned} y(s,t) &= \boldsymbol{x}(s,t)'\boldsymbol{\gamma}(s,t)+\epsilon(s,t), \qquad \epsilon(s,t)\sim N[0;\sigma^2_\epsilon], \ \text{independentes} \\ \boldsymbol{\gamma}(s,t) &= \boldsymbol{\beta}_t+\boldsymbol{\beta}(s,t). \end{aligned}$$

A evolução dos estados é descrita por

$$\begin{aligned} \boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1}+\boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t\sim N[0;\Sigma_\eta] \ \text{independentes} \\ \boldsymbol{\beta}(s,t) &= \boldsymbol{\beta}(s,t\text{–}1)+\boldsymbol{w}(s,t), \qquad \boldsymbol{w}(s,t)\sim PG \ \text{multivariados independentes.} \end{aligned}$$

A modelagem da variação espacial, incorporada no modelo a partir dos erros $w(s,t)$, é baseada na hipótese de isotropia.

Em processos pontuais, se o objetivo é analisar unicamente a variação temporal de processos contínuos, pode-se trabalhar com a agregação da intensidade do processo no espaço. Paez (2004) propõe um processo de Cox contínuo no tempo, cuja intensidade a cada período de tempo t é definida por $\Lambda(t)$, representando a intensidade média do processo no espaço para t fixo. Uma forma de definir $\Lambda(t)$ é pelo produto da intensidade populacional $\rho(t)$, supostamente conhecida, e a função de risco $\pi(t)$. A autora propõe um modelo log-gaussiano para a função de risco $\pi(t)$, incorporando um processo autoregressivo no tempo, $\gamma(t)$, e covariáveis $\boldsymbol{x}(t)$ que tratam a variação temporal não explicada por $\rho(t)$, tal que $\pi(t)=\exp\{\gamma(t)+\boldsymbol{\beta x}'(t)\}$. Supõe-se que a correlação entre $\gamma(t_i)$ e $\gamma(t_j)$, para t_i e t_j períodos de tempo tais que $t_i<t_j$, depende da distância temporal (t_j-t_i).

A proposta neste capítulo é agregar as abordagens de Brix e Diggle (2001) e Paez (2004) para processos pontuais com a abordagem de Gelfand *et al.* (2005) para processos contínuos. Na próxima seção, este modelo espaço-temporal é especificado em detalhes. Na seção seguinte são apresentados os aspectos computacionais da inferência em alguns casos especiais. Estudos com dados simulados destes modelos são mostrados no próximo capítulo e aplicações a conjuntos de dados reais são apresentadas no Capítulo 6.

4.2 Modelos Espaço-Temporais

O modelo espacial apresentado no Capítulo 3 pode ser estendido para incluir a dimensão do tempo. Seja o processo pontual espaço-temporal $Z(s,t)$, para o qual s são as coordenadas espaciais em uma região $S\subset\Re^2$ e $t\in[0,T]$ é o instante de tempo.

Assume-se que $Z(s,t)$, para cada t fixo, é um processo de Cox com função de intensidade $\lambda(s,t)$ modelada por

$$log\,[\,\lambda(s,t)\,] \;=\; \mu(t) + \zeta(s) + \phi(s,t), \tag{4.1}$$

onde $\mu(t)$ é a *tendência temporal*, comum a todos os pontos no espaço, $\zeta(s)$ é o *efeito puramente espacial*, comum a todos os instantes de tempo, e $\phi(s,t)$ são os *efeitos espaço-temporais*, específicos de cada ponto e tempo. Cada um destes efeitos pode ser decomposto em um componente determinístico e outro estocástico. O modelo é então completado com a especificação de distribuições a priori para os parâmetros dos componentes destes três efeitos.

4.2.1 Modelos para a Tendência Temporal

A tendência temporal $\mu(t)$ do modelo (4.1) pode ser modelada livremente, com a combinação de componentes determinísticos e estocásticos, dependência em covariáveis que tomam diferentes valores ao longo do tempo.

O modelo determinístico é representado por:

$$\mu(t) = \boldsymbol{f}(t)'\boldsymbol{\beta}, \tag{4.2}$$

onde $\boldsymbol{\beta}$ é o vetor de coeficientes de regressão e $F'(t)$ é um vetor de covariáveis medidas no tempo ou o próprio tempo, como no seguinte caso particular:

$$\mu(t) \; = \; \beta_0 \; + \; \beta_1 t. \tag{4.3}$$

Outro caso particular importante, que merece destaque, é aquele em que o nível da intensidade do processo não muda com tempo, ou seja, a tendência temporal é constante:

$$\mu(t) \; = \; \mu. \tag{4.4}$$

No modelo estocástico, a tendência temporal pode, por exemplo, ter evolução dinâmica:

$$\begin{aligned} \mu(t) \; &= \; F_t' \, \boldsymbol{\beta}(t), \\ \boldsymbol{\beta}(t) \; &= \; G_t \, \boldsymbol{\beta}(t{-}1) \; + \; \boldsymbol{\upsilon}(t), \qquad \boldsymbol{\upsilon}(t) \sim N\left[\underline{0}\,; \Omega_t\right], \end{aligned} \tag{4.5}$$

onde $\boldsymbol{\beta}(t)$ é o vetor de estados no tempo t, F_t e G_t são matrizes conhecidas e $\underline{0}$ é um vetor com elementos iguais a zero. Neste caso, assume-se que a observação dos eventos foi feita a tempo discreto ou discretizado. Um exemplo desta especificação é o *modelo dinâmico polinomial de primeira ordem*

$$\mu_t = \mu_{t-1} + \upsilon_t, \qquad \upsilon_t \sim N\left[0; \omega^2\right], \; t{=}2, ..., T, \qquad \mu_1 \sim N\left[\mu_0; \tau_0^2\right], \tag{4.6}$$

que contém o modelo (4.4) como caso particular ao se tomar $\mu_0 = \mu$ e fazendo $\omega^2 \rightarrow \infty$ e $\tau_0^2 \rightarrow \infty$.

Outro exemplo da especificação em (4.5) é o *modelo dinâmico polinomial de segunda* *ordem*

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \upsilon_t, \quad \upsilon_t \sim N\left[0; \omega_1^2\right], \; t=2, ..., T, \quad \mu_1 \sim N\left[\mu_0; \tau_0^2\right], \tag{4.7}$$

$$\beta_t = \beta_{t-1} + \nu_t, \quad \nu_t \sim N\left[0; \omega_2^2\right], \; t=2, ..., T, \quad \beta_1 \sim N\left[\beta_0; \kappa_0^2\right]. \tag{4.8}$$

4.2.2 Modelos para os Efeitos Espaciais

Assim como a tendência temporal, os efeitos puramente espaciais $\zeta(s)$ do modelo (4.1) podem ser escritos pela soma de componentes determinísticos ($\zeta_d(s)$) e estocásticos ($\zeta_e(s)$).

A parte deteminística destes efeitos espaciais pode, por exemplo, envolver covariáveis definidas em cada localização espacial, mas que não variam no tempo:

$$\zeta_d(s) = \boldsymbol{x}(s)'\boldsymbol{\alpha},$$

e a parte estocástica pode ser definida, por exemplo, por um processo gaussiano na região de estudo:

$$\zeta_e(\cdot) \sim PG[0; \gamma^2; \rho_\zeta(\cdot; \kappa)].$$

4.2.3 Modelos para os Efeitos Espaço-Temporais

Os efeitos espaço-temporais $\phi(s,t)$ do modelo (4.1) são usualmente descritos como processos espaciais gaussianos independentes no tempo.

Seguindo o trabalho de Paez (2004), a proposta nesta tese é modelar $\phi(\cdot, t), \; t=1, ..., T$, como processos gaussianos (estacionários e isotrópicos no espaço) autorregressivos e *estacionários no tempo*

$$\phi(s,t) \; = \; \eta\, \phi(s, t{-}1) + \omega(s,t), \qquad \omega(\cdot, t) \sim PG\left[0; (1{-}\eta^2)\sigma^2; \rho_\phi(\cdot; \theta)\right], \tag{4.9}$$

onde $0 < \eta < 1$ é o parâmetro de correlação temporal e $\phi(\cdot, 1) \sim PG\left[0; \sigma^2; \rho_\phi(\cdot; \theta)\right]$; ou *não-estacionários no tempo*

$$\phi(s,t) \; = \; \phi(s, t{-}1) + \omega(s,t), \qquad \omega(\cdot, t) \sim PG\left[0; \sigma^2; \rho_\omega(\cdot; \theta)\right], \tag{4.10}$$

com $\phi(\cdot,1)\sim PG\,[0;\tau^2;\rho_\phi(\cdot;\gamma)]$.

As equações de evolução em (4.9) e (4.10) se aplicam ao caso usual de agrupamento dos arranjos espaciais em intervalos de tempos equiespaçados. Adaptações nestas equações e nas expressões para a estimação dos parâmetros podem ser feitas para que elas também se apliquem ao caso mais geral de tempos não-esquiespaçados.

4.3 Aspectos Computacionais da Inferência

A inferência bayesiana via métodos MCMC nos modelos espaço-temporais é feita através da discretização do espaço em N células e a T intervalos de tempo discretos e equiespaçados.

Assume-se que as contagem de eventos $Y_{[i,t]}$ na i-ésima célula (com área unitária) e no t-ésimo intervalo de tempo (de comprimento unitário) são, condicional à intensidade do processo $\lambda_{[i,t]}$, independentes entre as células espaciais e intervalos de tempo, e que

$$\begin{aligned} p(y_{[i,t]}\,|\,\lambda_{[i,t]}) &\propto e^{-\lambda_{[i,t]}}\,\lambda_{[i,t]}^{y_{[i,t]}}, \qquad i=1,...,N, \quad t=1,...,T, \\ \log\left(\lambda_{[i,t]}\right) &= \mu_{[t]} + \phi_{[i,t]}. \end{aligned} \tag{4.11}$$

Nesta seção são mostrados os detalhes de cálculos das distribuições condicionais completas dos seguintes casos particulares dos modelos para a tendência temporal $\mu_{[t]}$:

- Dois modelos determinísticos

$$\begin{aligned} \text{Constante:} \quad \mu_{[t]} &= \mu, & t=1,...,T; \\ \text{Linear no tempo:} \quad \mu_{[t]} &= \beta_0+\beta_1 t, & t=1,...,T; \end{aligned}$$

- Dois modelos estocásticos com evolução temporal dinâmica polinomial

$$\begin{aligned} \text{1a ordem:} \quad \mu_{[t]} &= \mu_{[t-1]}+\upsilon_{[t]}, & \upsilon_{[t]}\sim N\left[0;\omega^2\right],\ t=2,...,T, \quad & \mu_{[1]}\sim N\left[\mu_0;\tau_0^2\right]; \\ \text{2a ordem:} \quad \mu_{[t]} &= \mu_{[t-1]}+\beta_{[t-1]}+\upsilon_{1[t]}, & \upsilon_{1[t]}\sim N\left[0;\omega_1^2\right],\ t=2,...,T, \quad & \mu_{[1]}\sim N\left[\mu_0;\tau_0^2\right], \\ \beta_{[t]} &= \beta_{[t-1]}+\upsilon_{2[t]}, & \upsilon_{2[t]}\sim N\left[0;\omega_2^2\right],\ t=2,...,T, \quad & \beta_{[1]}\sim N\left[\beta_0;\kappa_0^2\right]. \end{aligned}$$

Estes modelos são combinados com o modelo de efeitos espaciais $\phi_{[i,t]}$ autorregressivos e estacionários no tempo, ou seja,

$$\phi_{[\cdot,t]} = \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}, \text{ com } \epsilon_{[\cdot,t]} \sim N\left[\underline{0}\,;(1-\eta^2)\sigma^2 R_\theta\right],\ t=2,...,T, \text{ e } \phi_{[\cdot,1]} \sim N\left[\underline{0}\,;\sigma^2 R_\theta\right],$$

onde $\underline{0}$ é um vetor de comprimento N com elementos iguais a zero e $R_\theta = [R_{i,j}]_{\{i,j=1,...,N\}}$, com $R_{i,j} = \rho(\|s_i - s_j\|;\theta)$, é a matriz $N \times N$ de correlações espaciais entre as células, para uma função de correlação espacial isotrópica $\rho(\cdot;\theta)$ apropriadamente escolhida.

4.3.1 Modelo de Tendência Constante

Assume-se que, no modelo (4.11),

$$\begin{aligned}
\mu_{[t]} &= \mu, & t=1,...,T; \qquad (4.12)\\
\phi_{[\cdot,t]} &= \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}, \quad \epsilon_{[\cdot,t]} \sim N\left[\underline{0}\,;(1-\eta^2)\sigma^2 R_\theta\right], & t=2,...,T, \qquad (4.13)\\
\phi_{[\cdot,1]} &\sim N\left[\underline{0}\,;\sigma^2 R_\theta\right].
\end{aligned}$$

Para facilitar a inferência, o modelo é reparametrizado definindo-se $\varphi_{[i,t]} \doteq \log(\lambda_{[i,t]})$, ou, equivalentemente,

$$\varphi_{[\cdot,t]} = \underline{1}\,\mu + \phi_{[\cdot,t]}, \qquad t=1,...,T; \tag{4.14}$$

Aplicando na equação (4.13) a substituição dada na equação (4.14), verifica-se que o modelo reparametrizado é um modelo dinâmico linear autorregressivo,

$$\begin{aligned}
\varphi_{[\cdot,t]} &= \eta\varphi_{[\cdot,t-1]} + (1-\eta)\mu\underline{1} + \epsilon_{[\cdot,t]}, \qquad (4.15)\\
&\quad \text{com} \quad \epsilon_{[\cdot,t]} \sim N\left[\underline{0}\,;(1-\eta^2)\sigma^2 R_\theta\right], \text{ independentes } \forall\, t = 2,...,T,\\
&\quad \text{e} \quad \varphi_{[\cdot,1]} \sim N\left[\underline{1}\mu\,;\sigma^2 R_\theta\right].
\end{aligned}$$

Note que a densidade a priori condicional completa de $\boldsymbol{\varphi}$ é

$$\begin{aligned}
\pi\left(\boldsymbol{\varphi}\,|\,\mu,\eta,\sigma^2,\theta\right) &= \pi\left(\varphi_{[\cdot,1]}\,|\,\mu,\eta,\sigma^2,\theta\right)\cdot\prod_{t=2}^{T}\pi\left(\varphi_{[\cdot,t]}\,|\,\varphi_{[\cdot,1]},...,\varphi_{[\cdot,t-1]},\mu,\eta,\sigma^2,\theta\right)\\
&= \pi\left(\varphi_{[\cdot,1]}\,|\,\mu,\sigma^2,\theta\right)\cdot\prod_{t=2}^{T}\pi\left(\varphi_{[\cdot,t]}\,|\,\varphi_{[\cdot,t-1]},\mu,\eta,\sigma^2,\theta\right), \qquad (4.16)
\end{aligned}$$

onde as distribuições deste produtório são as Normais do modelo (4.15).

Os componentes deste modelo são a média μ, os efeitos $\boldsymbol{\varphi} = (\varphi'_{[\cdot,1]}, ..., \varphi'_{[\cdot,T]})'$, e seus hiperparâmetros η, σ^2 e θ, que têm distribuição a posteriori proporcional a

$$p\left(\boldsymbol{\varphi}, \mu, \eta, \sigma^2, \theta \,|\, \boldsymbol{y}\right) \;\propto\; p(\boldsymbol{y} \,|\, \boldsymbol{\varphi}) \cdot \pi(\boldsymbol{\varphi} \,|\, \mu, \eta, \sigma^2, \theta) \cdot \pi(\mu) \cdot \pi(\eta, \sigma^2, \theta). \tag{4.17}$$

Para fazer inferência sobre (4.17), serão utilizados métodos MCMC de amostragem. Para isso, as condicionais completas de cada parâmetro ou bloco de parâmetros precisa ser calculada. Neste modelo, todos os efeitos/parâmetros serão amostrados individualmente de sua condicional completa, diretamente ou através de Metropolis-Hastings. Os cálculos são mostrados a seguir.

Amostragem de φ:

Como no modelo espacial do Capítulo 3, a amostragem das log-intensidades é feita individualmente para cada célula i e tempo t. Deste modo, definindo $\varphi_{[-i,-t]}$ como o vetor $\boldsymbol{\varphi}$ excluído o efeito $\varphi_{[i,t]}$, a distribuição condicional completa da log-intensidade $\varphi_{[i,t]}$ é proporcional a

$$\begin{aligned} p_c(\varphi_{[i,t]}| \varphi_{[-i,-t]}, \mu, \eta, \sigma^2, \theta, \boldsymbol{y}) \;&\propto\; p(\boldsymbol{y}| \boldsymbol{\varphi})\, \pi(\boldsymbol{\varphi}| \mu, \eta, \sigma^2, \theta) \;\propto\; \prod_{t=1}^{T}\prod_{i=1}^{N} p(y_{[i,t]}| \varphi_{[i,t]}) \\ &\times\; \pi\left(\varphi_{[\cdot,1]}| \mu, \sigma^2, \theta\right) \prod_{t=2}^{T} \pi\left(\varphi_{[\cdot,t]}| \varphi_{[\cdot,t-1]}, \mu, \eta, \sigma^2, \theta\right) \\ &\propto\; p(y_{[i,t]}| \varphi_{[i,t]})\, \pi\left(\varphi_{[\cdot,t]}| \varphi_{[\cdot,t-1]}, \mu, \eta, \sigma^2, \theta\right) \\ &\times\; \pi\left(\varphi_{[\cdot,t+1]}| \varphi_{[\cdot,t]}, \mu, \eta, \sigma^2, \theta\right). \end{aligned} \tag{4.18}$$

Definindo $\varphi_{[-i,t]} = (\varphi_{[1,t]}, ..., \varphi_{[i-1,t]}, \varphi_{[i+1,t]}, ..., \varphi_{[N,t]})'$ o vetor das log-intensidades no tempo t excluída aquela da célula i, tem-se que

$$\begin{aligned} p_c(\varphi_{[i,1]} \,|\, \varphi_{[-i,-1]}, \mu, \eta, \sigma^2, \theta, \boldsymbol{y}) \;&\propto\; p(y_{[i,1]} \,|\, \varphi_{[i,1]})\; \pi(\varphi_{[i,1]} \,|\, \varphi_{[-i,1]}, \mu, \eta, \sigma^2, \theta) \pi(\varphi_{[i,1]} \,|\, \mu, \eta, \sigma^2, \theta) \\ &\propto\; \exp\left\{ -e^{\varphi_{[i,1]}} + y_{[i,1]}\varphi_{[i,1]} - \frac{(\varphi_{[i,1]} - K_i)^2}{2Q_i} \right\}; \end{aligned}$$

$$
\begin{aligned}
p_c(\varphi_{[i,t]}\,|\,\varphi_{[-i,-t]},\mu,\eta,\sigma^2,\theta,\boldsymbol{y}) \;&\propto\; p(y_{[i,t]}\,|\,\varphi_{[i,t]})\;\pi(\varphi_{[i,t]}\,|\,\varphi_{[-i,t]},\varphi_{[\cdot,t-1]},\mu,\eta,\sigma^2,\theta)\\
&\times\;\pi(\varphi_{[i,t]}\,|\,\varphi_{[-i,t]},\varphi_{[\cdot,t+1]},\mu,\eta,\sigma^2,\theta)\\
&\propto\;\exp\left\{-e^{\varphi_{[i,t]}}+y_{[i,t]}\varphi_{[i,t]}-\frac{(\varphi_{[i,t]}-L_{it})^2}{2P_i}\right\},\quad t=2,...,T-1;
\end{aligned}
$$

$$
\begin{aligned}
p_c(\varphi_{[i,T]}\,|\,\varphi_{[-i,-T]},\mu,\eta,\sigma^2,\theta,\boldsymbol{y}) \;&\propto\; p(y_{[i,T]}\,|\,\varphi_{[i,T]})\;\pi(\varphi_{[i,T]}\,|\,\varphi_{[-i,T]},\phi_{[\cdot,T-1]},\mu,\eta,\sigma^2,\theta)\\
&\propto\;\exp\left\{-e^{\varphi_{[i,T]}}+y_{[i,T]}\varphi_{[i,T]}-\frac{(\varphi_{[i,T]}-E_{iT})^2}{2W_i}\right\};
\end{aligned}
$$

onde H_i é a matriz de correlações R_θ extraídas as i-ésimas linha e coluna, B_i é o vetor formado pela i-ésima linha de R_θ sem a i-ésima coluna, e

$$
\begin{aligned}
M_i \;&=\; \mu+B_iH_i^{-1}(\varphi_{[-i,1]}-\mu\underline{1}) \qquad \text{e} \qquad V_i=\sigma^2(1-B_iH_i^{-1}B_i'),\\
E_{it} \;&=\; \eta\varphi_{[i,t-1]}+(1-\eta)\mu\underline{1}+B_iH_i^{-1}\left[\;(\varphi_{[-i,t]}-\mu\underline{1})-\eta(\varphi_{[-i,t-1]}-\mu\underline{1})\;\right] \quad \text{e} \quad W_i=(1-\eta^2)V_i,\\
F_{it} \;&=\; \left\{\varphi_{[i,t+1]}-(1-\eta)\mu\underline{1}+B_iH_i^{-1}\left[\;\eta(\varphi_{[-i,t]}-\mu\underline{1})-\;(\varphi_{[-i,t+1]}-\mu\underline{1})\;\right]\right\}\eta^{-1} \quad \text{e} \quad Z_i=\eta^{-2}W_i,\\
K_i \;&=\; Q_i(M_iV_i^{-1}+F_{i1}Z_i^{-1}) \quad \text{e} \quad Q_i=(\,V_i^{-1}+Z_i^{-1})^{-1},\\
L_{it} \;&=\; P_i(E_{it}W_i^{-1}+F_{it}Z_i^{-1}) \quad \text{e} \quad P_i=(W_i^{-1}+Z_i^{-1})^{-1}.
\end{aligned}
$$

As duas propostas de densidades para amostragem da posteriori Metropolis-Hastings são as mesmas do modelo espacial. Na **proposta da priori condicional**, um novo valor $\varphi^N_{(i,t)}$ é amostrado da densidade a priori condicional aos valores correntes das demais log-intensidades e dos hiperparâmetros, ou seja, da distribuição

$$
\varphi_{[i,t]}\mid\varphi_{[-i,-t]},\mu,\eta,\sigma^2,\theta\;\sim\;\begin{cases} N[\,K_i;\,Q_i\,], & t=1,\\ N[\,L_{it};\,P_i\,], & t=2,...,T-1,\\ N[E_{iT};W_i], & t=T,\end{cases} \tag{4.19}
$$

com probabilidade de aceitação do valor proposto igual a $min\{1,\alpha_1(\varphi_{(i,t)})\}$, na qual

$$
\alpha_1(\varphi_{[i,t]})=\exp\Big\{-\Big(e^{\varphi^N_{[i,t]}}-e^{\varphi^V_{[i,t]}}\Big)+\big(\varphi^N_{[i,t]}-\varphi^V_{[i,t]}\big)\,y_{[i,t]}\Big\}.
$$

Na **proposta da posteriori MLGM**, o novo valor $\varphi^N_{[i,t]}$ é amostrado da densidade Normal com média M_{φ}^{V} e variância V_{φ}^{V} tais que

$$
\begin{aligned}
M_{\varphi}^{V} &= V_{\varphi}^{V}\left(m_{\varphi}/v_{\varphi}+\tilde{y}_{[i,t]}/\tilde{V}_{[i,t]}^{V}\right) \qquad \text{e} \qquad V_{\varphi}^{V}=\left(1/v_{\varphi}+1/\tilde{V}_{[i,t]}^{V}\right)^{-1}, \\
\tilde{y}_{[i,t]}^{V} &= \varphi_{[i,t]}^{V}+\frac{y_{[i,t]}-e^{\varphi_{[i,t]}^{V}}}{e^{\varphi_{[i,t]}^{V}}} \qquad \text{e} \qquad \tilde{V}_{[i,t]}^{V}=e^{-\varphi_{[i,t]}^{V}},
\end{aligned}
\tag{4.20}
$$

m_{φ} e v_{φ} são, respectivamente, a média e variância da distribuição a priori condicional de $\varphi_{[i,t]}$, dadas na equação (4.19). A probabilidade de aceitação do valor proposto é igual $min\{1,\alpha_2(\varphi_{(i,t)})\}$, na qual

$$
\begin{aligned}
\alpha_2(\varphi_{[i,t]}) &= exp\left\{-\left(e^{\varphi_{[i,t]}^{N}}-e^{\varphi_{[i,t]}^{V}}\right)+\left(\varphi_{[i,t]}^{N}-\varphi_{[i,t]}^{V}\right)y_{[i,t]})\right\}\times \\
&\times \; exp\left\{-\frac{(\varphi_{[i,t]}^{V}-M_{\phi}^{N})^2}{2V_{\varphi}^{N}}+\frac{(\varphi_{[i,t]}^{N}-M_{\varphi}^{V})^2}{2V_{\varphi}^{V}}\right\}\left(\frac{V_{\varphi}^{N}}{V_{\varphi}^{V}}\right)^{-1/2},
\end{aligned}
$$

na qual M_{φ}^{N} e V_{φ}^{N} são dados pelas expressões em (4.20) substituindo-se $\varphi_{[i,t]}^{V}$ por $\varphi_{[i,t]}^{N}$.

Amostragem de μ:

De acordo com (4.17), a distribuição condicional completa de μ é dada por

$$
p_c\big(\mu\,|\,\boldsymbol{\varphi},\eta,\sigma^2,\theta,\boldsymbol{y}\big)\propto\pi\big(\boldsymbol{\varphi}\,|\,\mu,\eta,\sigma^2,\theta\big)\,\pi(\mu)\,,
$$

ou seja, não depende das contagens de eventos $\boldsymbol{y}$. Deste modo, a inferência sobre μ pode ser conduzida enxergando-se as equações (4.15) como o modelo de regressão

$$
\boldsymbol{\delta}_{\varphi} \;=\; X\mu \;+\; \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon}\sim N\,[\,\underline{0}\,;W]\,,
$$

onde $\boldsymbol{\delta}_{\varphi}=(\delta'_{\varphi_1},...,\delta'_{\varphi_T})'$, com

$$
\begin{aligned}
\delta_{\varphi_1} &= \varphi_{[\cdot,1]}, \\
\delta_{\varphi_t} &= \varphi_{[\cdot,t]}-\eta\,\varphi_{[\cdot,\,t-1]}, \qquad t=2,...,T,
\end{aligned}
$$

$$X = \begin{pmatrix} \underline{1} \\ (1-\eta)\,\underline{1} \\ \ldots \\ (1-\eta)\,\underline{1} \end{pmatrix}_{NT\times 1} \quad \text{e} \quad W = diag\big(\sigma^2, (1-\eta)\sigma^2, \ldots, (1-\eta)\sigma^2\big) \otimes R_\theta,$$

onde $diag(\sigma^2, (1-\eta)\sigma^2, \ldots, (1-\eta)\sigma^2)$ é uma matriz diagonal de dimensão T.

Com a escolha da distribuição a priori $\mu \sim N[a_\mu; b_\mu^2]$, a distribuição condicional completa de μ é dada por

$$\begin{aligned} p_c\big(\mu \,|\, \boldsymbol{\delta}_\varphi, \eta, \sigma^2, \theta\big) &\propto p\left(\boldsymbol{\delta}_\varphi \,|\, \mu, \eta, \sigma^2, \theta\right)\; \pi(\mu) \\ &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\delta}_\varphi - X\mu)' W^{-1} (\boldsymbol{\delta}_\varphi - X\mu)\right\} \exp\left\{-\frac{1}{2}(\mu - a_\mu)^2 (b_\mu^2)^{-1}\right\} \\ &\propto \exp\left\{-\frac{(\mu - M_\mu)^2}{2C_\mu}\right\}, \end{aligned}$$

onde

$$\begin{aligned} M_\mu &= C_\mu \left(X' W^{-1} \delta_\varphi + a_\mu b_\mu^{-2}\right) \\ &= C_\mu \left\{\sigma^{-2} \underline{1}' R_\theta^{-1} \left[\varphi_{[\cdot,1]} + (1-\eta)(1-\eta^2)^{-1} \sum_{t=2}^{T} (\varphi_{[\cdot,t]} - \eta\, \varphi_{[\cdot,\,t-1]})\right] + a_\mu b_\mu^{-2}\right\} \\ C_\mu &= \left(X' W^{-1} X + b_\mu^{-2}\right)^{-1} \\ &= \left\{\sigma^{-2} \underline{1}' R_\theta^{-1} \underline{1}\, [1 + (T-1)(1-\eta)^2 (1-\eta^2)^{-1}] + b_\mu^{-2}\right\}^{-1}. \end{aligned}$$

Ou seja, $\mu \mid (\boldsymbol{\varphi}, \eta, \sigma^2, \theta, \boldsymbol{y}) \sim N[M_\mu; C_\mu]$. Desse modo, a amostragem de μ pode ser feita diretamente de sua distribuição condicional completa.

Amostragem de η, σ^2 e θ:

Os hiperparâmetros η, σ^2 e θ são assumidos independentes a priori, ou seja, a densidade a priori $\pi(\eta, \sigma^2, \theta)$ é o produto das densidades a priori marginais, escolhidas tais que

$$\begin{aligned} \eta &\sim U[a_\eta; b_\eta], \quad 0 \le a_\eta < b_\eta \le 1, \\ \sigma^2 &\sim GI[g_s; v_s] \;\text{ e} \\ \theta &\sim G[g_t; v_t]. \end{aligned}$$

Em cada iteração da cadeia MCMC, uma vez amostrados $\boldsymbol{\varphi}$ e $\boldsymbol{\mu}$, os valores de $\boldsymbol{\phi}$ são recuperados através da equação $\phi_{[i,t]} = \varphi_{[i,t]} - \mu$. Defina:

$$\begin{aligned} \delta_1 &= \phi_{[\cdot,1]}, \\ \delta_t &= \phi_{[\cdot,t]} - \eta\, \phi_{[\cdot,\, t-1]}, \qquad t = 2, ..., T. \end{aligned}$$

A distribuição condicional completa da variância σ^2 é dada por

$$\begin{aligned} p_c(\sigma^2 \,|\, \eta, \mu, \theta, \boldsymbol{\phi}, \boldsymbol{y}) &\propto \pi(\boldsymbol{\phi} \,|\, \eta, \sigma^2, \theta)\; \pi(\sigma^2) \\ &\propto (\sigma^2)^{-\left(g_s + \frac{TN}{2} + 1\right)} \exp\left\{ -\frac{1}{\sigma^2} \left[v_s + \sum_{t=1}^{T} \delta_t \right] \right\}, \quad \sigma^2 > 0, \end{aligned}$$

ou seja, $\sigma^2 \mid (\eta, \mu, \theta, \boldsymbol{\phi}, \boldsymbol{y}) \sim GI\left[g_s + \dfrac{TN}{2}\,;\, v_s + \dfrac{1}{2}\left(\delta_1 + (1-\eta^2)^{-1} \displaystyle\sum_{t=2}^{T} \delta_t \right) \right]$,

sendo portanto, de fácil amostragem direta.

O parâmetro η é amostrado via Metropolis-Hastings, pois sua distribuição condicional completa,

$$\begin{aligned} p_c(\eta \,|\, \boldsymbol{\phi}, \mu, \sigma^2, \theta, \boldsymbol{y}) &\propto \pi(\boldsymbol{\phi} \,|\, \eta, \sigma^2, \theta)\; \pi(\eta) \\ &\propto (1-\eta^2)^{-\frac{N(T-1)}{2}} \exp\left\{ -(1-\eta^2)^{-1} \sum_{t=2}^{T} \frac{\delta_t}{2\sigma^2} \right\} I_{[a_\eta, b_\eta]}(\eta), \end{aligned}$$

é de difícil amostragem direta. Dessa forma, a cada iteração da cadeia, um novo valor para η (denotado η^N) é amostrado da densidade proposta $U[0;1]$ e aceito com probabilidade igual a $min\{1, \alpha(\eta)\}$, na qual

$$\alpha(\eta) = \left[\frac{1 - (\eta^N)^2}{1 - (\eta^V)^2} \right]^{-\frac{N(T-1)}{2}} \exp\left\{ -\frac{1}{2} \sum_{t=2}^{T} \left[\frac{\delta_t^N}{1 - (\eta^N)^2} - \frac{\delta_t^V}{1 - (\eta^V)^2} \right] \right\}.$$

A distribuição condicional completa de θ também é de difícil amostragem direta:

$$\begin{aligned} p_c(\theta \,|\, \eta, \mu, \sigma^2, \boldsymbol{\phi}, \boldsymbol{y}) &\propto \pi(\boldsymbol{\phi} \,|\, \eta, \sigma^2, \theta)\; \pi(\theta) \\ &\propto \theta^{g_t - 1} |R_\theta|^{-\frac{T}{2}} \exp\left\{ -v_t \theta - \frac{1}{2\sigma^2} \left(\delta_1 + (1 - \eta^2)^{-1} \sum_{t=2}^{T} \delta_t \right) \right\}, \quad \theta > 0. \end{aligned}$$

Desse modo, no algoritmo de Metropolis-Hastings, um novo valor θ (denotado θ^N) é amostrado, em função do valor da iteração anterior (θ^V), da densidade proposta log-normal com parâmetros

$log\theta^V - \frac{w_\theta}{2}$ e w_θ, de modo que seu valor esperado é $E(\theta^N | \theta^V) = \theta^V$ e seu coeficiente de variação é $CV(\theta^N | \theta^V) = (e^{w_\theta} - 1)^{1/2}$, sendo w_θ a constante sintonizadora da taxa de aceitação de novos. Aceita-se este valor proposto com probabilidade igual a $min\{1, \alpha(\theta)\}$, na qual

$$\begin{aligned}
\alpha(\theta) &= \left(\frac{|R_{\theta^N}|}{|R_{\theta^V}|}\right)^{-\frac{1}{2}} \left(\frac{\theta^N}{\theta^V}\right)^{g_t} exp\left\{-v_t(\theta^N - \theta^V) - \frac{1}{2\sigma^2}\left(\delta_1^N - \delta_1^V\right)\right\} \times \\
&\times \ exp\left\{-\frac{1}{2w_\theta}\left[\left(log\theta^V - log\theta^N + \frac{w_\theta}{2}\right)^2 - \left(log\theta^N - log\theta^V + \frac{w_\theta}{2}\right)^2\right]\right\}.
\end{aligned}$$

4.3.2 Modelo de Tendência Determinística Linear

Assume-se que, no modelo (4.11),

$$\begin{aligned}
\mu_{[t]} &= \beta_0 + \beta_1 \cdot t, && t = 1, ..., T; && (4.21)\\
\phi_{[\cdot,t]} &= \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}, \quad \epsilon_{[\cdot,t]} \sim N\left[\underline{0}\,; (1-\eta^2)\sigma^2 R_\theta\right], && t = 2, ..., T, && (4.22)\\
\phi_{[\cdot,1]} &\sim N\left[\underline{0}\,; \sigma^2 R_\theta\right].
\end{aligned}$$

A reparametrização também é feita definindo-se $\varphi_{[i,t]} \doteq \log(\lambda_{[i,t]})$, ou, equivalentemente,

$$\varphi_{[\cdot,t]} = \underline{1}\,\mu_{[t]} + \phi_{[\cdot,t]}, \qquad t = 1, ..., T. \tag{4.23}$$

Aplicando na equação (4.22) a substituição dada na equação (4.23), verifica-se que o modelo reparametrizado é um modelo dinâmico linear autorregressivo,

$$\begin{aligned}
\varphi_{[\cdot,t]} &= \eta\varphi_{[\cdot,t-1]} + [(1-\eta)\beta_0 + (t - \eta(t-1))\beta_1]\,\underline{1} + \epsilon_{[\cdot,t]}, && (4.24)\\
&\text{com} \quad \epsilon_{[\cdot,t]} \sim N\left[\underline{0}\,; (1-\eta^2)\sigma^2 R_\theta\right], \ \text{independentes} \ \forall\, t = 2, ..., T,\\
&\text{e} \qquad \varphi_{[\cdot,1]} \sim N\left[(\beta_0 + \beta_1)\,\underline{1}\,, \sigma^2 R_\theta\right].
\end{aligned}$$

Os componentes deste modelo são os efeitos $\boldsymbol{\varphi} = (\varphi'_{[\cdot,1]}, ..., \varphi'_{[\cdot,T]})'$, os parâmetros $\boldsymbol{\beta} = (\beta_0, \beta_1)'$ e os hiperparâmetros η, σ^2 e θ, que têm distribuição a posteriori proporcional a

$$p\left(\boldsymbol{\varphi}, \boldsymbol{\beta}, \eta, \sigma^2, \theta \,|\, \boldsymbol{y}\right) \propto p(\boldsymbol{y} \,|\, \boldsymbol{\varphi}) \cdot \pi(\boldsymbol{\varphi} \,|\, \boldsymbol{\beta}, \eta, \sigma^2, \theta) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\eta, \sigma^2, \theta). \tag{4.25}$$

Para fazer inferência sobre (4.25), serão utilizados métodos MCMC de amostragem. Para isso, as condicionais completas de cada parâmetro ou bloco de parâmetros precisam ser calculadas. Neste modelo, todos os efeitos e parâmetros serão amostrados individualmente de sua condicional completa, diretamente ou através de Metropolis-Hastings. Os cálculos são mostrados a seguir.

Amostragem de $\boldsymbol{\beta}=(\beta_0,\beta_1)'$:

De acordo com (4.25), a distribuição condicional completa de $\boldsymbol{\beta}$ é dada por

$$p_c\left(\boldsymbol{\beta}\,|\,\boldsymbol{\varphi},\eta,\sigma^2,\theta,\boldsymbol{y}\right) \;\propto\; \pi(\boldsymbol{\varphi}\,|\,\boldsymbol{\beta},\eta,\sigma^2,\theta)\cdot\pi(\boldsymbol{\beta}),$$

ou seja, não depende das contagens de eventos $\boldsymbol{y}$. Deste modo, a inferência sobre $\boldsymbol{\beta}$ pode ser conduzida enxergando-se as equações (4.24) como o modelo de regressão

$$\boldsymbol{\delta}_\varphi \;=\; X\boldsymbol{\beta} \;+\; \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon}\sim N\left[\,\underline{0}\,;W\right],$$

onde $\boldsymbol{\delta}_\varphi=(\delta'_{\varphi_1},...,\delta'_{\varphi_T})'$, com

$$\begin{aligned}
\delta_{\varphi_1} &= \varphi_{[\cdot,1]},\\
\delta_{\varphi_t} &= \varphi_{[\cdot,t]}-\eta\,\varphi_{[\cdot,t-1]}, \qquad t=2,...,T,
\end{aligned}$$

$$X=\begin{pmatrix}
\underline{1} & \underline{1}\\
(1-\eta)\,\underline{1} & (2-\eta)\,\underline{1}\\
... & ...\\
(1-\eta)\,\underline{1} & (T-(T-1)\eta)\,\underline{1}
\end{pmatrix}_{NT\times 2} \quad \text{e} \quad W=diag\big(\sigma^2,(1-\eta)\sigma^2,...,(1-\eta)\sigma^2\big)\otimes R_\theta,$$

onde $diag(\sigma^2,(1-\eta)\sigma^2,...,(1-\eta)\sigma^2)$ é uma matriz diagonal de dimensão T.

Com a escolha da distribuição a priori $\boldsymbol{\beta}\sim N[\boldsymbol{b}_0;\mathcal{B}_0]$, a distribuição condicional completa de $\boldsymbol{\beta}$ é dada por

$$\begin{aligned}
p_c\left(\boldsymbol{\beta}\,|\,\boldsymbol{\delta}_\varphi,\eta,\sigma^2,\theta\right) &\propto\; p\left(\boldsymbol{\delta}_\varphi\,|\,\boldsymbol{\beta},\eta,\sigma^2,\theta\right)\;\pi(\boldsymbol{\beta})\\
&\propto\; \exp\left\{-\frac{1}{2}(\boldsymbol{\delta}_\varphi-X\boldsymbol{\beta})'W^{-1}(\boldsymbol{\delta}_\varphi-X\boldsymbol{\beta})\right\}\;\exp\left\{-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{b}_0)'\mathcal{B}_0^{-1}(\boldsymbol{\beta}-\boldsymbol{b}_0)\right\}\\
&\propto\; \exp\left\{-\frac{1}{2}(\boldsymbol{\beta}-\mathcal{A}_1)'\mathcal{B}_1^{-1}(\boldsymbol{\beta}-\mathcal{A}_1)\right),
\end{aligned}$$

com $\mathcal{A}_1 = \mathcal{B}_1 \left(X' W^{-1} \delta_\varphi + \mathcal{B}_0^{-1} \boldsymbol{b}_0 \right)$ e $\mathcal{B}_1 = \left(X' W^{-1} X + \mathcal{B}_0^{-1} \right)^{-1}$. Ou seja, $\boldsymbol{\beta} \mid \boldsymbol{\varphi}, \eta, \sigma^2, \theta \sim N\left[\mathcal{A}_1; \mathcal{B}_1\right]$. Desse modo, a amostra de $\boldsymbol{\beta}$ pode ser feita diretamente de sua distribuição condicional completa.

Amostragem de φ:

Como no modelo de tendência constante, os efeitos $\varphi_{[i,t]}$ são amostrados individualmente. Verifica-se na equação (4.24) que o vetor $\varphi_{[\cdot,t]}$ depende apenas de $\boldsymbol{\varphi}$ através de $\varphi_{[\cdot,\,t-1]}$ e $\varphi_{[\cdot,\,t+1]}$. Assim, a densidade condicional completa de $\varphi_{[i,t]}$ é dada por

$$
\begin{aligned}
p_c\big(\varphi_{[i,t]} \,|\, \boldsymbol{\varphi}_{[-i,-t]}, \boldsymbol{\beta}, \eta, \sigma^2, \theta, \boldsymbol{y}\big) &\propto p(\boldsymbol{y} \,|\, \boldsymbol{\varphi}) \cdot \pi(\varphi \,|\, \boldsymbol{\beta}, \eta, \sigma^2, \theta) \\
&\propto p(y_{[i,t]} \,|\, \varphi_{[i,t]}) \cdot \pi\big(\varphi_{[i,t]} \,|\, \varphi_{[-i,\,t]}, \varphi_{[\cdot,\,t-1]}, \boldsymbol{\beta}, \eta, \sigma^2, \theta\big) \\
&\times \qquad\qquad\qquad \pi\big(\varphi_{[i,t]} \,|\, \varphi_{[-i,\,t]}, \varphi_{[\cdot,\,t+1]}, \boldsymbol{\beta}, \eta, \sigma^2, \theta\big) \,.
\end{aligned}
$$

Desse modo, a amostragem de cada $\varphi_{[i,t]}$ via Metropolis-Hastings é idêntica à de $\phi_{[i,t]}$ do modelo de tendência constante, a menos das seguintes alterações nas médias

$$
\begin{aligned}
M_i &= \mu_{[1]} + B_i H_i^{-1} \left(\varphi_{[-i,1]} - \mu_{[1]} \underline{1} \right), \\
E_{it} &= \eta \varphi_{[i,t-1]} + (\mu_{[t]} - \eta \mu_{[t-1]}) \underline{1} + B_i H_i^{-1} \left[\, \varphi_{[-i,t]} - \eta \varphi_{[-i,t-1]} - (\mu_{[t]} - \eta \mu_{[t-1]}) \underline{1} \,\right], \\
F_{it} &= \eta^{-1} \left\{ \left[\, \varphi_{[i,t+1]} - (\mu_{[t]} - \eta \mu_{[t-1]}) \underline{1} \right] + B_i H_i^{-1} \left[\, \eta \varphi_{[-i,t]} - \varphi_{[-i,t+1]} + (\mu_{[t]} - \eta \mu_{[t-1]}) \underline{1} \,\right] \right\}.
\end{aligned}
$$

Amostragem de η, σ^2 e θ:

Em cada iteração da cadeia MCMC, uma vez amostrados $\boldsymbol{\varphi}$ e $\boldsymbol{\mu}$, os valores de $\boldsymbol{\phi}$ são recuperados e a amostragem de η, σ^2 e θ é então feita como no modelo de tendência constante.

4.3.3 Modelo de Tendência Dinâmica Polinomial de Primeira Ordem

Assume-se que, no modelo (4.11),

$$\begin{aligned}
\mu_{[t]} &= \mu_{[t-1]} + \upsilon_{[t]}, & \upsilon_{[t]} &\sim N\left[0;\omega^2\right], & t&=2,...,T, \qquad (4.26)\\
\mu_{[1]} &\sim N\left[\mu_0;\tau_0^2\right]; \\
\phi_{[\cdot,t]} &= \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}, & \epsilon_{[\cdot,t]} &\sim N\left[\underline{0}\,;(1-\eta^2)\sigma^2 R_\theta\right], & t&=2,...,T, \qquad (4.27)\\
\phi_{[\cdot,1]} &\sim N\left[\underline{0}\,;\sigma^2 R_\theta\right].
\end{aligned}$$

A reparametrização deste modelo é a mesma do modelo anterior: define-se $\varphi_{[i,t]} \doteq \log(\lambda_{[i,t]})$, ou, equivalentemente,

$$\varphi_{[\cdot,t]} = \underline{1}\,\mu_{[t]} + \phi_{[\cdot,t]}, \qquad t=1,...,T. \tag{4.28}$$

Aplicando na equação (4.27) a substituição dada na equação (4.29), verifica-se que o modelo reparametrizado é um modelo dinâmico linear autorregressivo,

$$\begin{aligned}
\varphi_{[\cdot,t]} &= \eta\varphi_{[\cdot,t-1]} + \underline{1}\,(\mu_{[t]} - \eta\,\mu_{[t-1]}) + \epsilon_{[\cdot,t]}, & \epsilon_{[\cdot,t]} &\sim N\left[\underline{0}\,;(1-\eta^2)\sigma^2 R_\theta\right], \qquad (4.29)\\
\mu_{[t]} &= \mu_{[t-1]} + \upsilon_{[t]}, & \upsilon_{[t]} &\sim N\left[\underline{0}\,;\omega^2\right],\\
\mu_{[1]} &\sim N\left[\mu_0;\tau_0^2\right] \text{ e } \varphi_{[\cdot,1]}\,|\,\mu_{[1]} \sim N\left[\underline{1}\,\mu_{[1]};\sigma^2 R_\theta\right].
\end{aligned}$$

Os componentes deste modelo são os efeitos $\boldsymbol{\varphi} = (\varphi'_{[\cdot,1]}, ..., \varphi'_{[\cdot,T]})'$, $\boldsymbol{\mu} = (\mu_{[1]}, ..., \mu_{[T]})'$ e os hiperparâmetros η, σ^2, θ e ω^2, que têm distribuição a posteriori proporcional a

$$p\left(\boldsymbol{\varphi}, \boldsymbol{\mu}, \eta, \sigma^2, \theta, \omega^2\,|\,\boldsymbol{y}\right) \propto p(\boldsymbol{y}\,|\,\boldsymbol{\varphi}) \cdot \pi(\boldsymbol{\varphi}\,|\,\boldsymbol{\mu}, \eta, \sigma^2, \theta) \cdot \pi(\boldsymbol{\mu}\,|\,\omega^2) \cdot \pi(\omega^2) \cdot \pi(\eta, \sigma^2, \theta). \tag{4.30}$$

Para fazer inferência sobre (4.30), serão utilizados métodos MCMC de amostragem. Neste modelo, os efeitos temporais $\boldsymbol{\mu}$ são amostrados conjuntamente através de uma adaptação do algoritmo FFBS. Os efeitos espaço-temporais $\boldsymbol{\varphi}$ e demais parâmetros são amostrados individualmente de sua condicional completa, como nos modelos anteriores. Os cálculos são mostrados a seguir.

<u>**Amostragem de $\boldsymbol{\mu}$**:</u>

Dentro da amostragem no MCMC, o modelo (4.29) pode ser visto com um modelo dinâmico linear no qual as "observações" são os efeitos $\varphi_{[i,t]}$ e os "parâmetros de estado" são os efeitos temporais $\mu_{[t]}$. Desse modo, uma adaptação do algoritmo FFBS, descrita a seguir, é usada para amostrar $\mu_{[t]},\ t=1,...,T$.

<u>Passo FF: Filtro de Kalman</u>

Definindo-se $D_t = \{\varphi_{[\cdot,1]},...,\varphi_{[\cdot,t]}\}$, este passo faz a passagem de $p(\mu_{[t-1]} \mid D_{t-1})$ para $p(\mu_{[t]} \mid D_t)$, sucessivamente para $t = 1,...,T$.

De (4.26), $\mu_{[t]} \mid \mu_{[t-1]}, D_{t-1} \sim N\big[\mu_{[t-1]};\omega^2\big]$ e, definindo $\mu_{[t-1]} \mid D_{t-1} \sim N[m_{t-1};c_{t-1}]$ tem-se

$$\begin{bmatrix}\mu_{[t]}\\ \mu_{[t-1]}\end{bmatrix} \mid D_{t-1} \sim N\left[\begin{bmatrix}m_{t-1}\\ m_{t-1}\end{bmatrix};\begin{bmatrix}c_{t-1}+\omega^2 & c_{t-1}\\ c_{t-1} & c_{t-1}\end{bmatrix}\right].$$

E por (4.29), $\varphi_{[\cdot,t]} \mid \mu_{[t]},\mu_{[t-1]},D_{t-1} \sim N\big[\,\underline{1}(\mu_{[t]}-\eta\mu_{[t-1]}) + \eta\,\varphi_{[\cdot,t-1]}\,;\,(1-\eta^2)\sigma^2 R_\theta\big]$. Então

$$\begin{bmatrix}\varphi_{[\cdot,t]}\\ \mu_{[t]}\\ \mu_{[t-1]}\end{bmatrix} \mid D_{t-1} \sim N\left[\begin{bmatrix}\underline{1}(1-\eta)m_{t-1}+\eta\varphi_{[\cdot,t-1]}\\ m_{t-1}\\ m_{t-1}\end{bmatrix};\begin{bmatrix}\Upsilon_{t-1} & \chi_{t-1} & (1-\eta)c_{t-1}\underline{1}\\ \chi'_{t-1} & c_{t-1}+\omega^2 & c_{t-1}\\ (1-\eta)c_{t-1}\underline{1}' & c_{t-1} & c_{t-1}\end{bmatrix}\right], \quad (4.31)$$

com $\Upsilon_{t-1}=[(1-2\eta+\eta^2)c_{t-1}+\omega^2]\underline{1}\,\underline{1}'+(1-\eta)\sigma^2 R_\theta$ e $\chi_{t-1}=[(1-\eta)c_{t-1}+\omega^2]\underline{1}$.

Assim,

$$\mu_{[t]} \mid \varphi_{[\cdot,t]},D_{t-1} = \mu_{[t]} \mid D_t \sim N\left[m_t;c_t\right], \quad t=2,...,T, \quad (4.32)$$
$$m_t = m_{t-1} + \chi'_{t-1}\Upsilon^{-1}_{t-1}\,[\varphi_{[\cdot,t]}-\eta\varphi_{[\cdot,t-1]}-\underline{1}(1-\eta)m_{t-1}],$$
$$c_t = c_{t-1}+\omega^2 - \chi'_{t-1}\Upsilon^{-1}_{t-1}\,\chi_{t-1};$$
$$\mu_{[1]} \mid \varphi_{[\cdot,1]},D_0 \equiv \mu_{[1]} \mid D_1 \sim N\left[m_1;c_1\right], \quad (4.33)$$
$$m_1 = \mu_0 + \underline{1}'\tau_0^2(\sigma^2 R_\theta + \underline{11}'\tau_0^2)^{-1}(\varphi_{[\cdot,1]} - \underline{1}\,\mu_0)$$
$$c_1 = \tau_0^2 - \underline{1}'\tau_0^2(\sigma^2 R_\theta + \underline{11}'\tau_0^2)^{-1}\underline{1}\tau_0^2.$$

Passo BS: Suavização Retropectiva

Para este passo, é necessária a obtenção da distribuição de $\mu_{[t]},\ t=1,..,T$, condicional à $\mu_{[t+1]},...,\mu_{[T]}$ e $D_T=\{\varphi_{[\cdot,1]},...,\varphi_{[\cdot,T]}\}$. Mas note que

$$
\begin{aligned}
p\big(\mu_{[t]}|\,\mu_{[t+1]},...,\mu_{[T]},D_T\big) &\propto\ p\big(\mu_{[t]},\mu_{[t+1]},...,\mu_{[T]},\varphi_{[\cdot,t+1]},...,\varphi_{[\cdot,T]},D_t\big)\\
&\propto\ p\big(\varphi_{[\cdot,t+1]},...,\varphi_{[\cdot,T]},\mu_{[t+1]},...,\mu_{[T]}|\,\mu_{[t]},D_t\big)\,p\big(\mu_{[t]}|\,D_t\big)\\
&\propto\ p\big(\varphi_{[\cdot,t+1]},...,\varphi_{[\cdot,T]}|\,\mu_{[t+1]},...,\mu_{[T]},\mu_{[t]},D_t\big)\\
&\times\ p\big(\mu_{[t+1]},...,\mu_{[T]}|\,\mu_{[t]},D_t\big)\,p\big(\mu_{[t]}|\,D_t\big)\\
&\propto\ p\big(\varphi_{[\cdot,t+1]}|\,\mu_{[t]},\mu_{[t+1]},D_t\big)\,p\big(\mu_{[t+1]}|\,\mu_{[t]},D_t\big)\,p\big(\mu_{[t]}|\,D_t\big)\,, \qquad \text{por (4.29)}\\
&\propto\ p\big(\varphi_{[\cdot,t+1]},\mu_{[t+1]},\mu_{[t]}|\,D_t\big)\\
&\propto\ p\big(\mu_{[t]}|\,\mu_{[t+1]},\varphi_{[\cdot,t+1]},D_t\big)\\
&\propto\ p\big(\mu_{[t]}|\,\mu_{[t+1]},D_{t+1}\big)\,. \qquad (4.34)
\end{aligned}
$$

Quase todas as passagens acima são triviais, baseadas na aplicação direta da regra de multiplicação $p(x,y,z)=p(x\,|\,y,z)p(y\,|\,z)p(z)$, para qualquer coleção de quantidades aleatórias x, y e z.

Passo BS.1: Obtenção de $p(\mu_{[t]}\,|\,\mu_{[t+1]},D_{t+1})$, para $t=1,...,T-1$.

Substituindo t por $t+1$ na distribuição conjunta (4.45) obtida no Filtro de Kalman, obtem-se

$$\mu_{[t]}\,|\,\mu_{[t+1]},\varphi_{[\cdot,t+1]},D_t \ \equiv\ \mu_{[t]}\,|\,\mu_{[t+1]},D_{t+1}\ \sim\ N\,[n_t;v_t]\,, \qquad (4.35)$$

com $\quad n_t=m_t+A_tB_t^{-1}C_t\quad$ e $\quad v_t=c_t-A_tB_t^{-1}A_t'$,

$$A_t=\Big[\ \underline{1}'(1-\eta)c_t \quad c_t\Big],\ B_t=\begin{bmatrix}\Upsilon_t & \chi_t\\ \chi_t' & c_t+\omega^2\end{bmatrix} \text{ e } C_t=\begin{bmatrix}\varphi_{[\cdot,t+1]}-\underline{1}(1-\eta)m_t-\eta\varphi_{[\cdot,t]}\\ \mu_{[t+1]}-m_t\end{bmatrix}.$$

Passo BS.2: Obtenção da amostra de $\boldsymbol{\mu}$.

(a) Amostre $\mu_{[T]}\,|\,D_T$ da $N\,[m_T;c_T]$ em (4.46) e faça $t=T-1$;

(b) Amostre $\mu_{[t]}\,|\,\mu_{t+1},D_{t+1}$ da $N\,[n_t;v_t]$ em (4.48);

(c) Decresça t para $t-1$ e retorne ao passo (b) até $t=1$.

Amostragem de ω^2:

Assumindo a priori que $\omega^2 \sim GI[g_\omega; v_\omega]$, a distribuição condicional completa de ω^2 (que depende apenas de $\boldsymbol{\mu}$) é dada por:

$$
\begin{aligned}
p_c\left(\omega^2 \mid \boldsymbol{\mu}\right) &\propto \pi\left(\boldsymbol{\mu} \mid \omega^2, \mu_0, \tau_0^2\right) \cdot \pi\left(\omega^2\right) \;\;\propto\;\; \prod_{t=2}^{T} \pi\left(\mu_{[t]} \mid \mu_{[t-1]}, \omega^2\right) \cdot \pi\left(\omega^2\right) \\
&\propto \prod_{t=2}^{T}\left[(\omega^2)^{-\frac{1}{2}}\, exp\left\{-\frac{1}{\omega^2}\frac{(\mu_{[t]}-\mu_{[t-1]})^2}{2}\right\}\right] \cdot \left[(\omega^2)^{-(g_\omega+1)}\, exp\left\{-\frac{v_\omega}{\omega^2}\right\}\right] \\
&\propto (\omega^2)^{-\frac{T-1}{2}-(g_\omega+1)}\, exp\left\{-\frac{1}{\omega^2}\left[v_\omega + \frac{1}{2}\sum_{t=2}^{T}(\mu_{[t]}-\mu_{[t-1]})^2\right]\right\} \\
\Rightarrow \quad \omega^2 \mid \boldsymbol{\mu} &\sim GI\left[\, g_\omega + \frac{T-1}{2}\;;\; v_\omega + \frac{1}{2}\sum_{t=2}^{T}(\mu_{[t]}-\mu_{[t-1]})^2 \,\right].
\end{aligned}
$$

Neste caso, a amostragem de ω^2 é feita diretamente da sua condicional completa.

Amostragem de $\boldsymbol{\varphi}$, η, σ^2 e θ:

O procedimento é idêntico àquele do modelo de tendência linear.

4.3.4 Modelo de Tendência Dinâmica Polinomial de Segunda Ordem

Assume-se que, no modelo (4.11),

$$
\begin{aligned}
\mu_{[t]} &= \mu_{[t-1]} + \beta_{[t-1]} + \upsilon_{1[t]}, \quad \upsilon_{1[t]} \sim N\left[0; \omega_1^2\right], \;\; t=2,...,T, \quad \mu_{[1]} \sim N\left[\mu_0; \tau_0^2\right], \quad (4.36)\\
\beta_{[t]} &= \beta_{[t-1]} + \upsilon_{2[t]}, \qquad \upsilon_{2[t]} \sim N\left[0; \omega_2^2\right], \;\; t=2,...,T, \quad \beta_{[1]} \sim N\left[\beta_0; \kappa_0^2\right], \quad (4.37)\\
\phi_{[\cdot,t]} &= \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}, \qquad \epsilon_{[\cdot,t]} \sim N\left[\underline{0}\,; (1-\eta^2)\sigma^2 R_\theta\right], \;\; t=2,...,T, \quad (4.38)\\
\phi_{[\cdot,1]} &\sim N\left[\underline{0}\,; \sigma^2 R_\theta\right].
\end{aligned}
$$

Definindo $\theta_t = (\mu_{[t]}, \beta_{[t]})'$,as equações (4.36) e (4.37) podem ser reescritas como

$$\theta_t = G\theta_{t-1} + \upsilon_{[t]}, \qquad \upsilon_{[t]} \sim N[\underline{0}\,; W] \quad e \quad \theta_{[1]} \sim N[M_0; T_0],$$

com

$$G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \; \upsilon_{[t]} = \begin{pmatrix} \upsilon_{1[t]} \\ \upsilon_{2[t]} \end{pmatrix}, \; W = \begin{pmatrix} \omega_1^2 & 0 \\ 0 & \omega_2^2 \end{pmatrix}, \; M_0 = \begin{pmatrix} \mu_0 \\ \beta_0 \end{pmatrix} \;\; e \;\; T_0 = \begin{pmatrix} \tau_0^2 & 0 \\ 0 & \kappa_0^2 \end{pmatrix}.$$

Novamente, definindo $\varphi_{[\cdot,t]} = \underline{1}\,\mu_{[t]} + \phi_{[\cdot,t]}$ e $g\!=\!(1\ 0)$, o modelo em (4.36) e (4.37) pode ser escrito como

$$\varphi_{[\cdot,t]} = \underline{1}\,g\,\theta_t \quad + \phi_{[\cdot,t]} \tag{4.39}$$

$$\theta_t = G\theta_{t-1} \quad + \upsilon_{[t]} \tag{4.40}$$

Substituindo as expressões (4.38) e (4.40) na equação (4.39), tem-se

$$\begin{aligned}
\varphi_{[\cdot,t]} &= \underline{1}\,g\,(G\theta_{t-1} + \upsilon_{[t]}) + (\eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}), \qquad\qquad t\!=\!2,...,T,\\
&= \underline{1}\,(\mu_{[t-1]} + \beta_{[t-1]} + \upsilon_{1[t]}) + \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}\\
&= \underline{1}\,(\eta\mu_{[t-1]} + (1\!-\!\eta)\mu_{[t-1]} + \beta_{[t-1]} + \upsilon_{1[t]}) + \eta\phi_{[\cdot,t-1]} + \epsilon_{[\cdot,t]}\\
&= \eta(\underline{1}\,\mu_{[t-1]} + \phi_{[\cdot,t-1]}) + \underline{1}\left[(\mu_{[t-1]} + \beta_{[t-1]} + \upsilon_{1[t]}) - \eta\mu_{[t-1]}\right] + \epsilon_{[\cdot,t]}\\
&= \eta\varphi_{[\cdot,t-1]} + \underline{1}(\mu_{[t]} - \eta\mu_{[t-1]}) + \epsilon_{[\cdot,t]},\\
\text{e} \quad \varphi_{[\cdot,1]} &= \underline{1}\,g\,\theta_1 + \phi_{[\cdot,1]} = \underline{1}\,\mu_{[1]} + \epsilon_{[\cdot,1]}.
\end{aligned}$$

Assim, o modelo reparametrizado é o modelo dinâmico linear autorregressivo

$$\begin{aligned}
\varphi_{[\cdot,t]} \mid \boldsymbol{\theta}, \boldsymbol{\psi}_1, D_{t-1} &\sim N\left[\eta\varphi_{[\cdot,t-1]} + (\mu_{[t]} - \eta\mu_{[t-1]})\underline{1}\,;(1-\eta^2)\sigma^2 R_\theta\right]\\
\varphi_{[\cdot,1]} \mid \boldsymbol{\theta}, \boldsymbol{\psi}_1, D_0 &\sim N\left[\mu_{[1]}\underline{1}\,;\,\sigma^2 R_\theta\right]
\end{aligned} \tag{4.41}$$

com $\boldsymbol{\theta} = (\theta'_{[1]},...,\theta'_{[T]})'$, $D_{t-1} = \{\varphi_{[\cdot,1]},...,\varphi_{[\cdot,t-1]}\}$ e $\boldsymbol{\psi}_1 = (\eta,\sigma^2,\theta)'$. Deve-se notar que, na distribuição condicional em (4.41), a dependência em $(\theta'_{[1]},...,\theta'_{[T]})'$ se resume apenas ao conhecimento de $\mu_{[t]}$ e $\mu_{[t-1]}$, como no modelo de primeira ordem.

Os componentes deste modelo são os efeitos $\boldsymbol{\varphi} = (\varphi'_{[\cdot,1]},...,\varphi'_{[\cdot,T]})'$, $\boldsymbol{\mu} = (\mu_{[1]},...,\mu_{[T]})'$, $\boldsymbol{\beta} = (\beta_{[1]},...,\beta_{[T]})'$ e os hiperparâmetros η, σ^2, θ, ω_1^2 e ω_2^2, que têm distribuição a posteriori proporcional a

$$\begin{aligned}
p\left(\boldsymbol{\varphi},\boldsymbol{\mu},\boldsymbol{\beta},\eta,\sigma^2,\theta,\omega_1^2,\omega_2^2\,|\,\boldsymbol{y}\right) \;\propto\; & p(\boldsymbol{y}|\,\boldsymbol{\varphi})\cdot\pi(\boldsymbol{\varphi}|\,\boldsymbol{\mu},\boldsymbol{\beta},\eta,\sigma^2,\theta)\cdot\pi(\boldsymbol{\mu},\boldsymbol{\beta}|\,\omega_1^2,\omega_2^2)\\
& \times\; \pi(\omega_1^2)\cdot\pi(\omega_2^2)\cdot\pi(\eta,\sigma^2,\theta).
\end{aligned} \tag{4.42}$$

Para fazer inferência sobre (4.42), serão utilizados métodos MCMC de amostragem. Neste modelo, os efeitos temporais $(\boldsymbol{\mu},\boldsymbol{\beta})$ são amostrados conjuntamente através de uma adaptação do algoritmo FFBS. Os efeitos espaço-temporais $\boldsymbol{\varphi}$ e demais parâmetros são amostrados individualmente de sua condicional completa, como nos modelos anteriores. Os cálculos são mostrados a seguir.

Amostragem de $\theta_t = (\mu_t, \beta_t)$

Dentro da amostragem no MCMC, o modelo (4.41) pode ser visto com um modelo dinâmico linear no qual as "observações" são os efeitos $\varphi_{[i,t]}$ e os "parâmetros de estado" são os efeitos temporais $\theta_{[t]}$. Do mesmo modo que no modelo de primeira ordem, uma adaptação do algoritmo FFBS, descrita a seguir, é usada para amostrar $\theta_{[t]},\ t=1,...,T$.

Primeiramente, deve-se escrever $\varphi_{[\cdot,t]}\,|\,\mu_{[t]},\mu_{[t-1]},D_{t-1}$ em função da $\theta_{[t]}$ e $\theta_{[t-1]}$:

$$(\mu_{[t]}-\eta\mu_{[t-1]}) = (1\ 0)\begin{pmatrix}\mu_{[t]}\\ \beta_{[t]}\end{pmatrix} - \eta\,(1\ 0)\begin{pmatrix}\mu_{[t-1]}\\ \beta_{[t-1]}\end{pmatrix} = (1\ 0)\,\theta_{[t]} - \eta\,(1\ 0)\,\theta_{[t-1]} = g\,\theta_{[t]} - \eta\,g\,\theta_{[t-1]}$$

e, assim, obter que

$$\varphi_{[\cdot,t]} \mid \theta_{[t]},\theta_{[t-1]},\boldsymbol{\psi}_1,D_{t-1} \;\sim\; N\left[\eta\varphi_{[\cdot,t-1]}+[g\,\theta_{[t]} - \eta\,g\,\theta_{[t-1]}]\underline{1}\,;(1-\eta^2)\sigma^2R_\theta)\right]. \qquad (4.43)$$

Passo FF: Filtro de Kalman

Definindo-se $D_t = \{\varphi_{[\cdot,1]},...,\varphi_{[\cdot,t]}\}$, este passo faz a passagem de $p(\theta_{[t-1]} \mid D_{t-1})$ para $p(\theta_{[t]}\,|\,D_t)$ sucessivamente para $t = 1,...,T$.

Para $t = 1$:

De (4.39), $\theta_{[1]} \sim N[M_0;T_0]$ e, de (4.41), $\varphi_{[\cdot,1]} \mid \theta_{[1]} \sim N\left[\underline{1}\,g\theta_{[1]};\sigma^2R_\theta\right]$. Desse modo, tem-se

$$\begin{bmatrix}\varphi_{[\cdot,1]}\\ \theta_{[1]}\end{bmatrix} \sim N\left[\begin{bmatrix}\underline{1}\,gM_0\\ M_0\end{bmatrix};\begin{bmatrix}\underline{1}\,gT_0g'\underline{1}'+\sigma^2R_\theta & \underline{1}\,gT_0\\ T_0g'\underline{1}' & T_0\end{bmatrix}\right],$$

e, assim,

$$\begin{aligned}\theta_{[1]} \mid \varphi_{[\cdot,1]},D_0 = \theta_{[1]}\,|\,D_1 &\sim N\left[M_1;C_1\right], \qquad (4.44)\\ M_1 &= M_0 + T_0g'\underline{1}'(\underline{1}\,gT_0g'\underline{1}' + \sigma^2R_\theta)^{-1}(\varphi_{[\cdot,1]} - \underline{1}\,gM_0),\\ C_1 &= T_0 - T_0g'\underline{1}'(\underline{1}\,gT_0g'\underline{1}' + \sigma^2R_\theta)^{-1}\underline{1}\,gT_0.\end{aligned}$$

Para $t=2,...,T$:

De (4.39), $\theta_{[t]}\,|\,\theta_{[t-1]},D_{t-1}\sim N\left[G\theta_{[t-1]};W\right]$ e, definindo $\theta_{[t-1]}\,|\,D_{t-1}\sim N[M_{t-1};C_{t-1})]$ tem-se

$$\begin{bmatrix}\theta_{[t]}\\ \theta_{[t-1]}\end{bmatrix}|\,D_{t-1}\sim N\left[\begin{bmatrix}GM_{t-1}\\ M_{t-1}\end{bmatrix};\begin{bmatrix}GC_{t-1}G'+W & GC_{t-1}\\ C_{t-1}G' & C_{t-1}\end{bmatrix}\right].$$

E por (4.43), $\varphi_{[\cdot,t]}\mid\theta_{[t]},\theta_{[t-1]},D_{t-1}\sim N\left[\eta\varphi_{[\cdot,t-1]}+(g\,\theta_{[t]}-\eta\,g\,\theta_{[t-1]})\underline{1}\,;(1-\eta^2)\sigma^2R_\theta\right]$. Assim,

$$\begin{bmatrix}\varphi_{[\cdot,t]}\\ \theta_{[t]}\\ \theta_{[t-1]}\end{bmatrix}|\,D_{t-1}\sim N\left[\begin{bmatrix}\underline{1}\,(g\,GM_{t-1}-\eta gM_{t-1})+\eta\varphi_{[\cdot,t-1]}\\ GM_{t-1}\\ M_{t-1}\end{bmatrix};\begin{bmatrix}\Upsilon_{t-1} & \chi_{t-1} & \gamma_{t-1}\\ \chi'_{t-1} & GC_{t-1}G'+W & GC_{t-1}\\ \gamma'_{t-1} & C_{t-1}G' & C_{t-1}\end{bmatrix}\right], \quad (4.45)$$

com
$$\Upsilon_{t-1}=\underline{1}\,[g\,(GC_{t-1}G'+W+\eta^2C_{t-1}-2\eta GC_{t-1})]\,\underline{1}'+(1-\eta)\sigma^2R_\theta,$$
$$\chi_{t-1}=\underline{1}\,[g\,(GC_{t-1}G'+W-\eta GC_{t-1})],$$
$$\gamma_{t-1}=\underline{1}\,[g\,(GC_{t-1}-\eta\,C_{t-1})].$$

Assim,

$$\theta_{[t]}\mid\varphi_{[\cdot,t]},D_{t-1}\;\equiv\;\theta_{[t]}\,|\,D_t\;\sim\;N\left[M_t;C_t\right],\quad t=2,...,T, \quad (4.46)$$
$$M_t=GM_{t-1}+\chi'_{t-1}\Upsilon^{-1}_{t-1}\,[\varphi_{[\cdot,t]}-\eta\varphi_{[\cdot,t-1]}-\underline{1}(g\,G\,M_{t-1}-\eta\,g\,M_{t-1})].$$
$$C_t=GC_{t-1}G'+GWG'-\chi'_{t-1}\Upsilon^{-1}_{t-1}\,\chi_{t-1}.$$

Passo BS: Suavização Retropectiva

Assim como no modelo de primeira ordem (equação 4.34), também neste modelo obtém-se

$$\begin{aligned}\pi\big(\theta_{[t]}\,|\,\theta_{[t+1]},...,\theta_{[T]},D_T\big)&\propto\pi\big(\theta_{[t]},\theta_{[t+1]},...,\theta_{[T]},\varphi_{[\cdot,t+1]},...,\varphi_{[\cdot,T]},D_t\big)\\ &\propto\pi\big(\varphi_{[\cdot,t+1]},...,\varphi_{[\cdot,T]},\theta_{[t+1]},...,\theta_{[T]}\,|\,\theta_{[t]},D_t\big)\,\pi\big(\theta_{[t]}\,|\,D_t\big)\\ &\propto\pi\big(\varphi_{[\cdot,t+1]},...,\varphi_{[\cdot,T]}\,|\,\theta_{[t+1]},...,\theta_{[T]},\theta_{[t]},D_t\big)\\ &\times\pi\big(\theta_{[t+1]},...,\theta_{[T]}\,|\,\theta_{[t]},D_t\big)\,\pi\big(\theta_{[t]}\,|\,D_t\big)\\ &\propto\pi\big(\varphi_{[\cdot,t+1]}\,|\,\theta_{[t]},\theta_{[t+1]},D_t\big)\,\pi\big(\theta_{[t+1]}\,|\,\theta_{[t]},D_t\big)\,\pi\big(\theta_{[t]}\,|\,D_t\big),\quad\text{por (4.41)}\\ &\propto\pi\big(\varphi_{[\cdot,t+1]},\theta_{[t+1]},\theta_{[t]}\,|\,D_t\big)\propto\pi\big(\theta_{[t]}\,|\,\theta_{[t+1]},\varphi_{[\cdot,t+1]},D_t\big)\\ &\propto\pi\big(\theta_{[t]}\,|\,\theta_{[t+1]},D_{t+1}\big)\,. \qquad (4.47)\end{aligned}$$

Quase todas as passagens acima são triviais, baseadas na aplicação direta da regra de multiplicação $p(x,y,z)=p(x\,|\,y,z)p(y\,|\,z)p(z)$, para qualquer coleção de quantidades aleatórias x, y e z.

Passo BS.1: Obtenção de $\pi(\theta_{[t]}\,|\,\theta_{[t+1]},D_{t+1})$, para $t=1,...,T-1$.

Substituindo t por $t+1$ na distribuição conjunta (4.45) obtida no Filtro de Kalman, obtem-se

$$\theta_{[t]}\,|\,\theta_{[t+1]},\varphi_{[\cdot,t+1]},D_t \;\equiv\; \theta_{[t]}\,|\,\theta_{[t+1]},D_{t+1} \;\sim\; N\left[N_t;V_t\right], \tag{4.48}$$

com $\quad N_t = M_t + A_t B_t^{-1} E_t \quad$ e $\quad V_t = C_t - A_t B_t^{-1} A_t'$,

$$A_t = \begin{bmatrix}\gamma_t' & C_t G'\end{bmatrix},\; B_t = \begin{bmatrix}\Upsilon_t & \chi_t \\ \chi_t' & G C_t G' + W\end{bmatrix},\; E_t = \begin{bmatrix}\varphi_{[\cdot,t+1]} - \eta\varphi_{[\cdot,t]} - \underline{1}(g\,G\,M_t - \eta\,g\,M_t) \\ \theta_{[t+1]} - G M_{t+1}\end{bmatrix}.$$

Passo BS.2: Obtenção da amostra de $\theta_1,...,\theta_T$.

(a) Amostre $\theta_{[T]}\,|\,D_T$ da $N\left[M_T;C_T\right]$ em (4.46) e faça $t=T-1$;

(b) Amostre $\theta_{[t]}\,|\,\theta_{t+1},D_{t+1}$ da $N\left[N_t;V_t\right]$ em (4.48);

(c) Decresça t para $t-1$ e retorne ao passo (b) até $t=1$.

Amostragem de ω_1^2 e ω_2^2:

Assumindo a priori que $\omega_1^2 \sim GI[g_{\omega 1};v_{\omega 1}]$, a distribuição condicional completa de ω_1^2 (que depende apenas de $\boldsymbol{\mu}$ e $\boldsymbol{\beta}$) é dada por:

$$\begin{aligned}
p_c\left(\omega_1^2 \mid \boldsymbol{\mu},\boldsymbol{\beta}\right) &\propto \pi\left(\boldsymbol{\mu} \mid \boldsymbol{\beta},\omega_1^2\right)\cdot\pi\left(\omega_1^2\right) \;\propto\; \prod_{t=2}^{T}\pi\left(\mu_{[t]} \mid \mu_{[t-1]},\beta_{[t-1]},\omega_1^2\right)\cdot\pi\left(\omega_1^2\right) \\
&\propto \prod_{t=2}^{T}\left[(\omega_1^2)^{-\frac{1}{2}}\,exp\left\{-\frac{1}{\omega_1^2}\frac{(\mu_{[t]}-\mu_{[t-1]}-\beta_{[t-1]})^2}{2}\right\}\right]\cdot\left[(\omega_1^2)^{-(g_{\omega 1}+1)}\,exp\left\{-\frac{v_{\omega 1}}{\omega_1^2}\right\}\right] \\
&\propto (\omega_1^2)^{-\frac{T-1}{2}-(g_{\omega 1}+1)}\,exp\left\{-\frac{1}{\omega_1^2}\left[v_{\omega 1}+\frac{1}{2}\sum_{t=2}^{T}(\mu_{[t]}-\mu_{[t-1]}-\beta_{[t-1]})^2\right]\right\}
\end{aligned}$$

$$\Rightarrow\quad \omega_1^2 \mid \boldsymbol{\mu},\boldsymbol{\beta} \;\sim\; GI\left[g_{\omega 1}+\frac{T-1}{2}\;;\;v_{\omega 1}+\frac{1}{2}\sum_{t=2}^{T}(\mu_{[t]}-\mu_{[t-1]}-\beta_{[t-1]})^2\right].$$

E, assumindo a priori que $\omega_2^2 \sim GI[g_{\omega 2};v_{\omega 2}]$, a distribuição condicional completa de ω_2^2 (que depende apenas de $\boldsymbol{\beta}$) é dada por:

$$
\begin{aligned}
p_c\left(\omega_2^2 \mid \boldsymbol{\beta}\right) &\propto \pi\left(\boldsymbol{\beta} \mid \omega_2^2\right) \cdot \pi\left(\omega_2^2\right) \propto \prod_{t=2}^{T} \pi\left(\beta_{[t]} \mid \beta_{[t-1]}, \omega_2^2\right) \cdot \pi\left(\omega_2^2\right) \\
&\propto \prod_{t=2}^{T}\left[(\omega_2^2)^{-\frac{1}{2}}\, exp\left\{-\frac{1}{\omega_2^2}\frac{(\beta_{[t]}-\beta_{[t-1]})^2}{2}\right\}\right] \cdot \left[(\omega_2^2)^{-(g_{\omega 2}+1)}\, exp\left\{-\frac{v_{\omega 2}}{\omega_2^2}\right\}\right] \\
&\propto (\omega_2^2)^{-\frac{T-1}{2}-(g_{\omega 2}+1)}\, exp\left\{-\frac{1}{\omega_2^2}\left[v_{\omega 2}+\frac{1}{2}\sum_{t=2}^{T}(\beta_{[t]}-\beta_{[t-1]})^2\right]\right\} \\
\Rightarrow \quad \omega_2^2 \mid \boldsymbol{\beta} &\sim GI\left[\, g_{\omega 2}+\frac{T-1}{2}\ ;\ v_{\omega 2}+\frac{1}{2}\sum_{t=2}^{T}(\beta_{[t]}-\beta_{[t-1]})^2\,\right].
\end{aligned}
$$

Amostragem de $\boldsymbol{\varphi}$, η, σ^2 e θ:

O procedimento é idêntico àquele do modelo de tendência linear (e dinâmico de primeira ordem).

4.4 Sumário

Neste capítulo, foi proposto um modelo log-linear para a intensidade de processos pontuais espaço-temporais com decomposição de componentes em efeitos puramente temporais, efeitos puramente espaciais e efeitos variando no espaço e no tempo. Cada um destes três tipos de efeitos pode ser modelado por componentes determinísticos e/ou estocásticos.

Os cálculos do procedimento de bayesiano de inferência via MCMC foram apresentados em detalhes para vários casos particulares deste modelo. O próximo capítulo mostra estudos com dados simulados destes modelos particulares.

Capítulo 5

Estudos de Simulação

5.1 Introdução

Este capítulo é dedicado à apresentação de estudos de simulação conduzidos com os modelos propostos no capítulo anterior. Nestes estudos, observações dos processos pontuais são geradas, em uma região fictícia, de acordo com escolhas arbitrárias dos parâmetros do modelo.

O principal objetivo destes estudos é verificar a eficiência dos metódos de estimação, valendo-se do fato de que os valores reais dos parâmetros são conhecidos. Adicionalmente, pode-se medir a velocidade de convergência das cadeias MCMC e o tempo de processamento dos algoritmos.

Nestes estudos, serão simulados conjuntos de dados do modelo (4.1), proposto no capítulo anterior, para o logaritmo da intensidade $\lambda(s,t)$ do processo na localização espacial s e no tempo t:

$$log\left[\lambda(s,t)\right] \;=\; \mu(t) + \zeta(s) + \phi(s,t). \tag{5.1}$$

Cada uma das seções deste capítulo é dedicada a um dos seguintes modelos para a tendência temporal $\mu(t)$: constante (modelo (4.4)), determinística linear (modelo (4.3)), estocástica dinâmica polinomial de primeira (modelo (4.6)) e segunda ordens (modelo (4.7-4.8)). Em todos os casos, os efeitos espaço-temporais $\phi(s,t)$ são modelados por processos gaussianos autorregressivos (modelo (4.9)). Como na Seção 4.3, nestes estudos simulados os efeitos puramente espaciais $\zeta(s)$ são especificados como nulos, embora seja muito fácil a incorporação de efeitos espaciais não-nulos nos modelos daquela seção.

5.2 Tendência Temporal Constante

Nesta seção, assume-se que, no modelo (5.1), a tendência temporal constante $\mu(t) = \mu$, $\forall t$ e efeitos espaço-temporais $\phi(s,t)$ modelados como processos gaussianos autoregressivos *estacionários no tempo*

$$\phi(s,t) \; = \; \eta\,\phi(s,t\text{–}1) + \omega(s,t), \qquad \omega(\cdot,t) \sim PG\left[0; (1\text{–}\eta^2)\sigma^2; \rho_\phi(\cdot;\theta)\right], \tag{5.2}$$

com $0 < \eta < 1$ e $\phi(\cdot,1) \sim PG\,[0;\sigma^2;\rho(\cdot;\theta)]$, ou *não-estacionários no tempo*

$$\phi(s,t) \; = \; \phi(s,t\text{–}1) + \omega(s,t), \qquad \omega(\cdot,t) \sim PG\left[0; \sigma^2; \rho_\omega(\cdot;\theta)\right], \tag{5.3}$$

com $\phi(\cdot,1) \sim PG\,[0;\tau^2;\rho_\phi(\cdot;\gamma)]$.

A região de simulação é uma região quadrada dividida por uma grade regular com $N = 100$ células (10×10) com áreas unitárias, em $T = 10$ períodos de tempos equiespaçados de comprimento unitário.

O primeiro conjunto de dados foi gerado com tendência temporal constante $\mu = 1$ e sucessivos processos gaussianos $\phi_{[\cdot,t]}$, estacionários no tempo (modelo (5.2)), gerados com $\eta = 0,67$, $\sigma^2 = 1$, $\theta = 0,8$ e função de correlação espacial exponencial $\rho_\phi(h) = e^{-h\theta}$. A Figura 5.1 mostra os valores gerados das log-intensidades $(\phi_{[i,t]} + \mu)$ nas células ao longo do tempo. Na Figura 5.2 é mostrada a seqüência de arranjos pontuais espaciais nos intervalos de tempo, totalizando 1987 eventos.

Foram escolhidas distribuições a priori *vagas* para μ e η, a saber, $\pi(\mu) \propto 1$ e $\eta \sim U[0;1]$. Para σ^2 e θ foram definidos dois conjuntos de prioris: as chamadas *pouco informativas*, nas quais o valor esperado e o desvio padrão são iguais ao valor real do parâmetro, o que resulta em $\sigma^2 \sim GI[3;2]$ e $\theta \sim G[1;1,25]$; e as chamadas *muito informativas*, com $\sigma^2 \sim GI[102;101]$ e $\theta \sim G[100;125]$, resultado da escolha da esperança igual ao valor real do parâmetro e do desvio padrão correspondente a um décimo deste valor.

O segundo conjunto de dados foi gerado com tendência temporal constante $\mu = 0$ e sucessivos processos gaussianos $\phi_{[\cdot,t]}$, não-estacionários no tempo (modelo (5.3)). O processo inicial $\phi_{[\cdot,1]}$ foi gerado com os valores $\sigma^2 = 1$, $\theta = 0,8$ e função de correlação espacial exponencial $\rho_\phi(h) = e^{-h\theta}$; para $t = 2, ..., 10$, os sucessivos processos $\phi_{[\cdot,t]}$ foram gerados com $\tau^2 = 0,9$, $\gamma = 0,6$ e função de correlação espacial exponencial $\rho_\omega(h) = e^{-h\gamma}$. A Figura 5.3 mostra os valores gerados das log-intensidades ($\phi_{[i,t]} + \mu$) nas células ao longo do tempo e a Figura 5.4 mostra a seqüência de arranjos pontuais espaciais nos intervalos de tempo, totalizando 17873 eventos. A comparação das Figuras 5.1 e 5.3 mostra a característica não-estacionária do segundo modelo, evidenciada no aumento da variabilidade dos valores dos efeitos ϕ ao longo do tempo. Além disso, o número de eventos gerados ao longo do tempo manteve-se aproximadamente constante no modelo estacionário (Figura 5.2), mas aumentou bastante no modelo não-estacionário (Figura 5.4).

Do mesmo modo que no primeiro conjunto de dados, para o segundo modelo foi escolhida a priori $\pi(\mu) \propto 1$ e dois conjuntos de prioris para as variância σ^2 e parâmetro de correlação espacial θ: pouco informativas, com $\sigma^2 \sim GI[3;2]$, $\theta \sim G[1;1,25]$, $\tau^2 \sim GI[3;1,8]$ e $\gamma \sim G[1;1,67]$; e muito informativas, com $\sigma^2 \sim GI[102;101]$, $\theta \sim G[100;125]$, $\tau^2 \sim GI[102;90,9]$ e $\gamma \sim G[100;166,7]$.

Para cada um dos quatro modelos definidos (modelo estacionário com prioris pouco e muito informativas; modelo não-estacionário com prioris pouco e muito informativas) foram geradas duas cadeias de tamanho 50 mil, resultantes de dois diferentes conjuntos de valores iniciais para os parâmetros. As duas amostras a posteriori de cada parâmetro são compostas de 1000 valores tomados a cada 25 da segunda metade das respectivas cadeias.

Foram utilizadas as duas propostas de densidades de amostragem dos $\phi_{[i,t]}$ no algoritmo de Metropolis-Hastings definidas no Capítulo 3: *proposta da priori* e *proposta MLGM*.

Os resultados da inferência com as amostras da proposta da priori são mostrados nas figuras 5.5 a 5.7 para o modelo estacionário, e nas figuras 5.8 a 5.10 para o modelo não-estacionário. Os resultados da proposta MLGM são visualmente idênticos a estes mostrados e, por isso, foram suprimidos.

As estimativas tomadas das amostras a posteriori para os efeitos ϕ tiveram uma boa concordância com os valores reais, tornando-se mais precisas à medida em que a intensidade aumenta. As médias a posteriori parecem reproduzir bem o padrão espacial destes efeitos. Também as estimativas dos hiperparâmetros ficaram muito próximas a seus valores reais usados na geração dos dados, mesmo para as prioris consideradas pouco informativas. A diferença entre as duas especificações de prioris surgiu apenas na maior varibilidade dos valores nas amostras a posteriori relacionadas à especificação de prioris pouco informativas.

Estes resultados sugerem que os métodos de estimação propostos são adequados para dados dos processos pontuais espaço-temporais estudados.

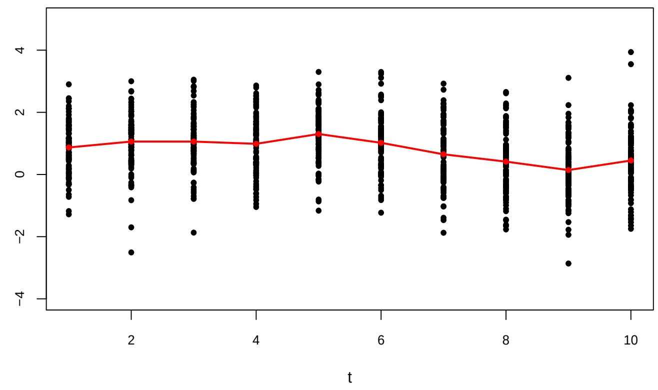


Figura 5.1: Modelo estacionário: valores gerados das log-intensidades $(\phi_{[i,t]}+\mu)$ nas 100 células da grade em 10 intervalos de tempo. A linha traçada é a média dos valores em cada tempo.

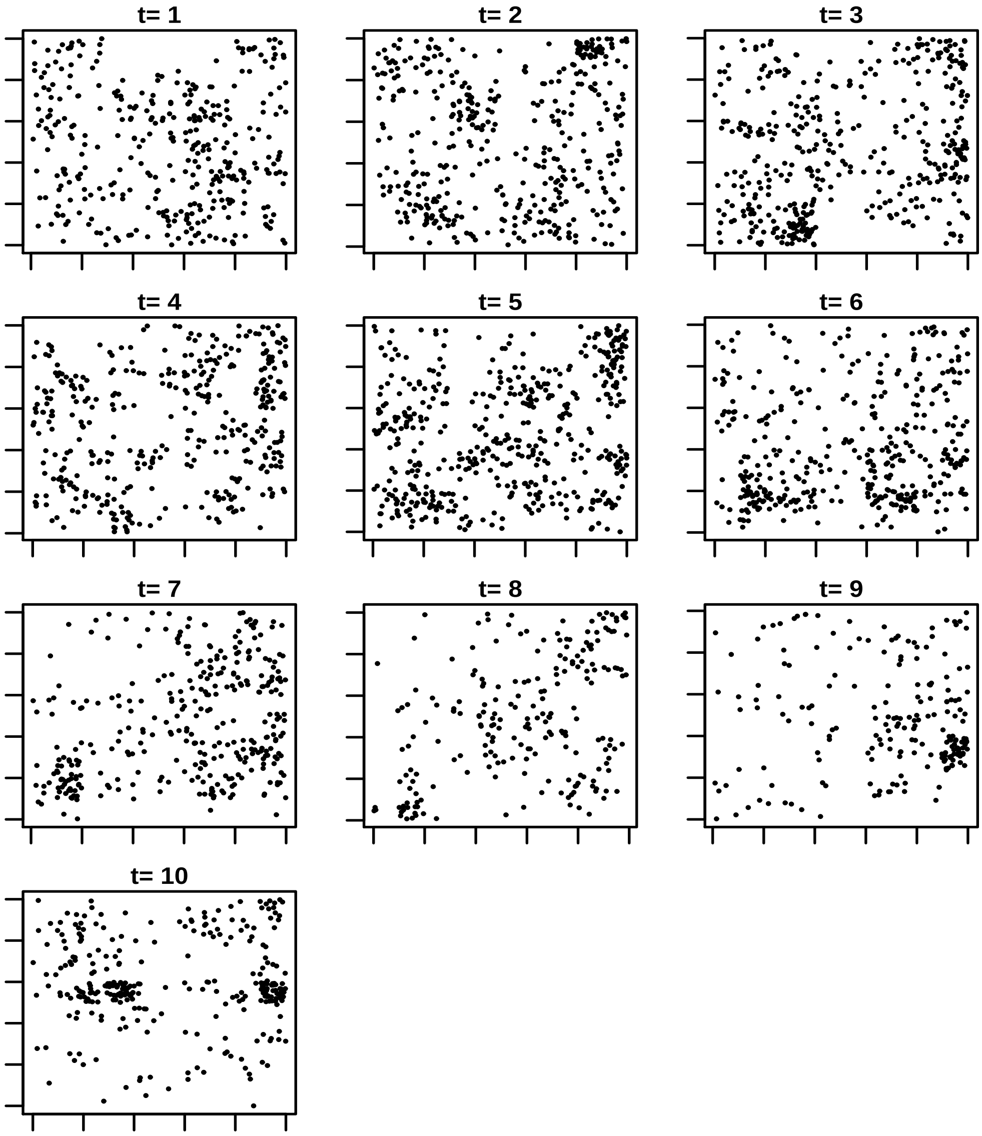


Figura 5.2: Modelo estacionário: eventos gerados nos 10 intervalos de tempo.

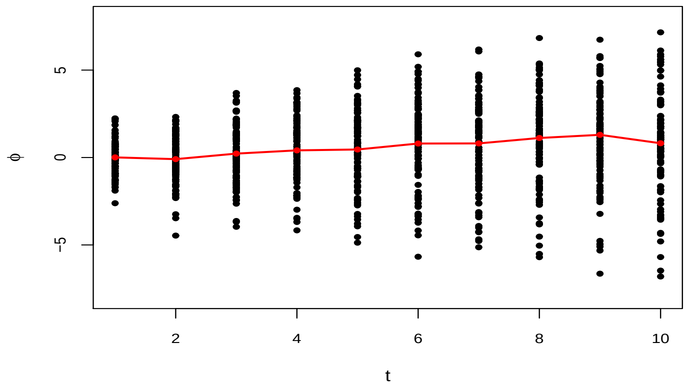


Figura 5.3: Modelo não-estacionário: valores gerados das log-intensidades ($\phi_{[i,t]}+\mu$) nas 100 células da grade em 10 tempos. A linha traçada é a média dos valores em cada tempo.

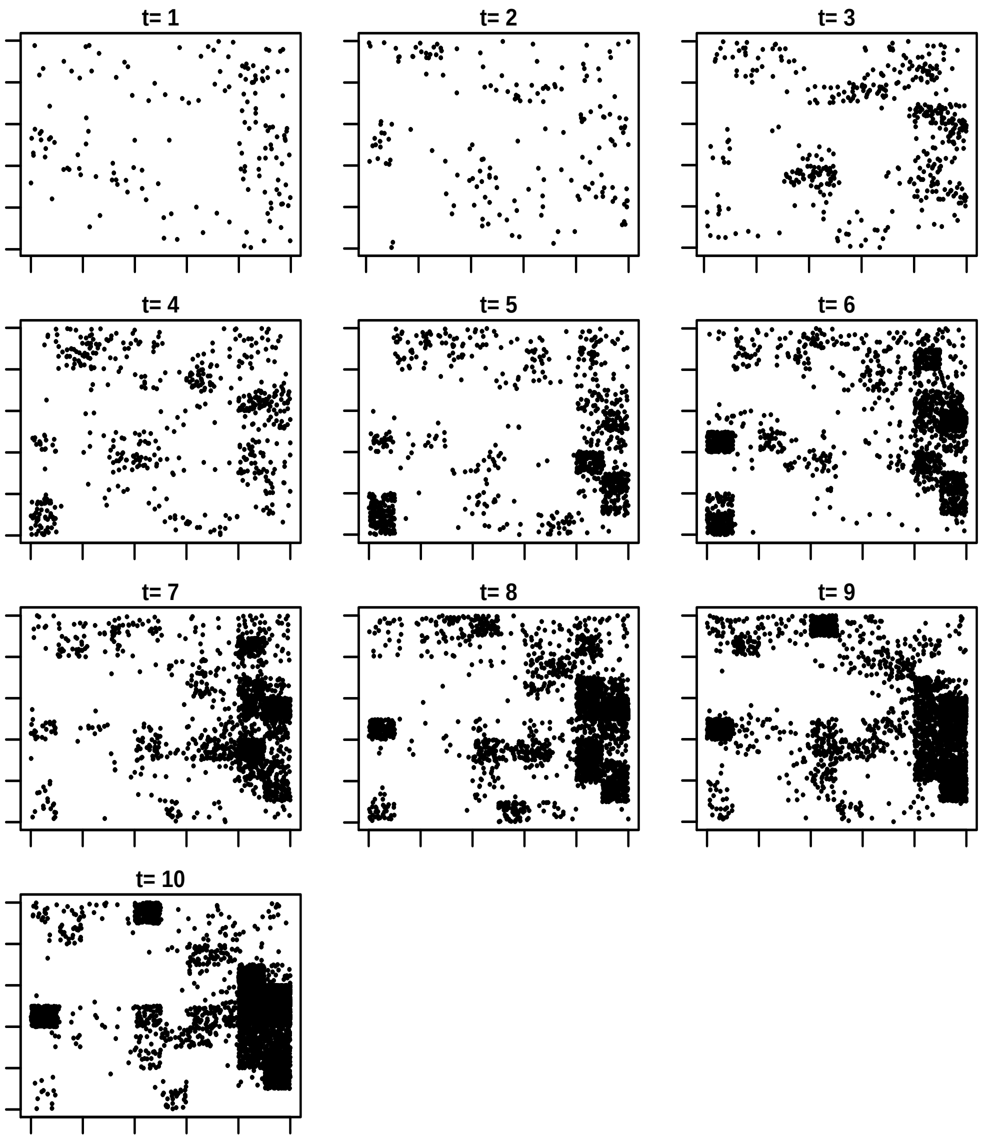


Figura 5.4: Modelo não-estacionário: eventos gerados nos 10 intervalos de tempo.

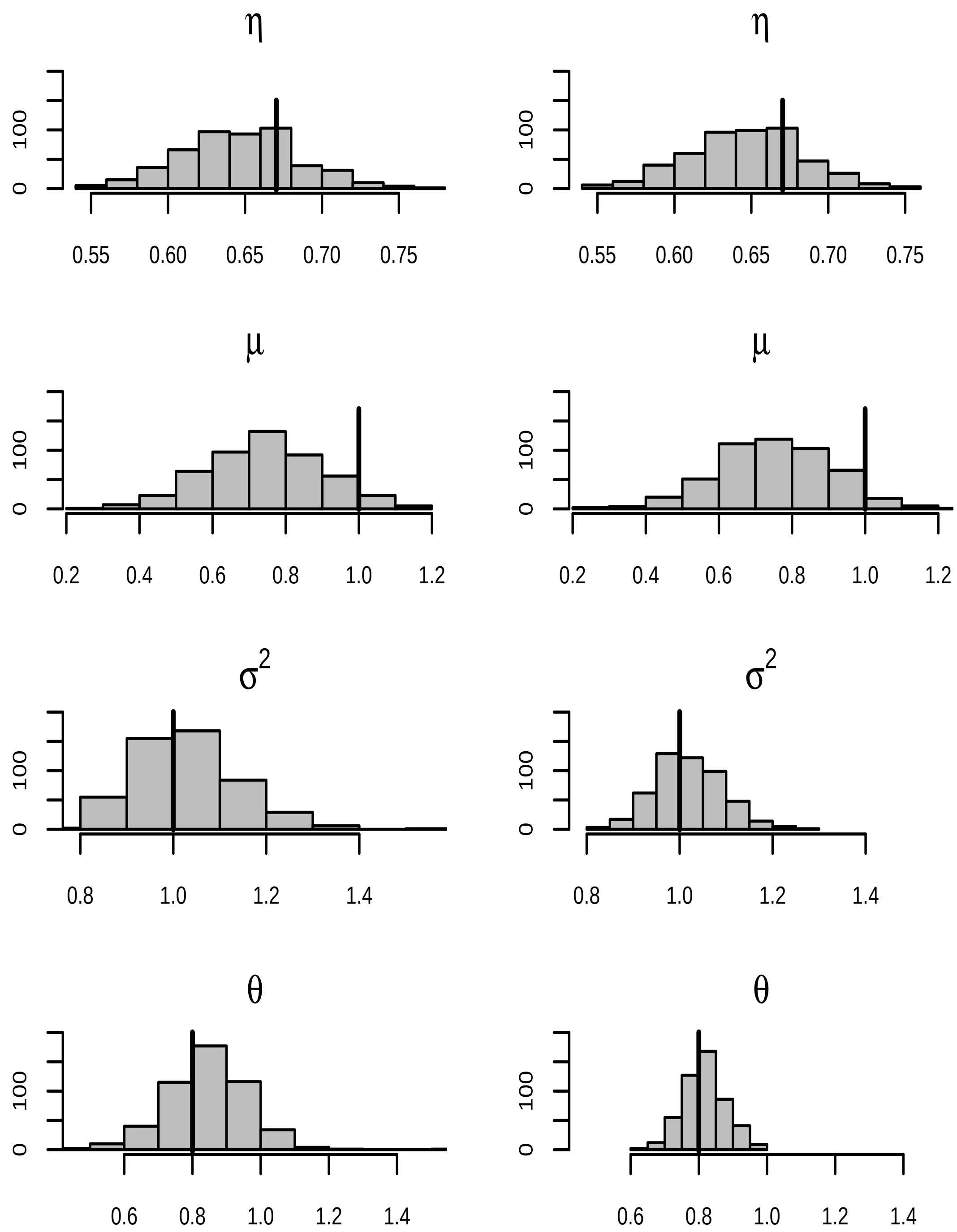


Figura 5.5: Modelo estacionário: histogramas das amostras a posteriori dos hiperparâmetros. Os gráficos na primeira coluna são relativos às prioris pouco informativas $\sigma^2 \sim GI[3;2]$ e $\theta \sim G[1;1,25]$. Os gráficos na segunda coluna são relativos às prioris muito informativas $\sigma^2 \sim GI[102;101]$ e $\theta \sim G[100;125]$. Em ambos os casos foram escolhidas a prioris $\pi(\mu) \propto 1$ e $\eta \sim U[0;1]$. O traço vertical marca o valor real do parâmetro.

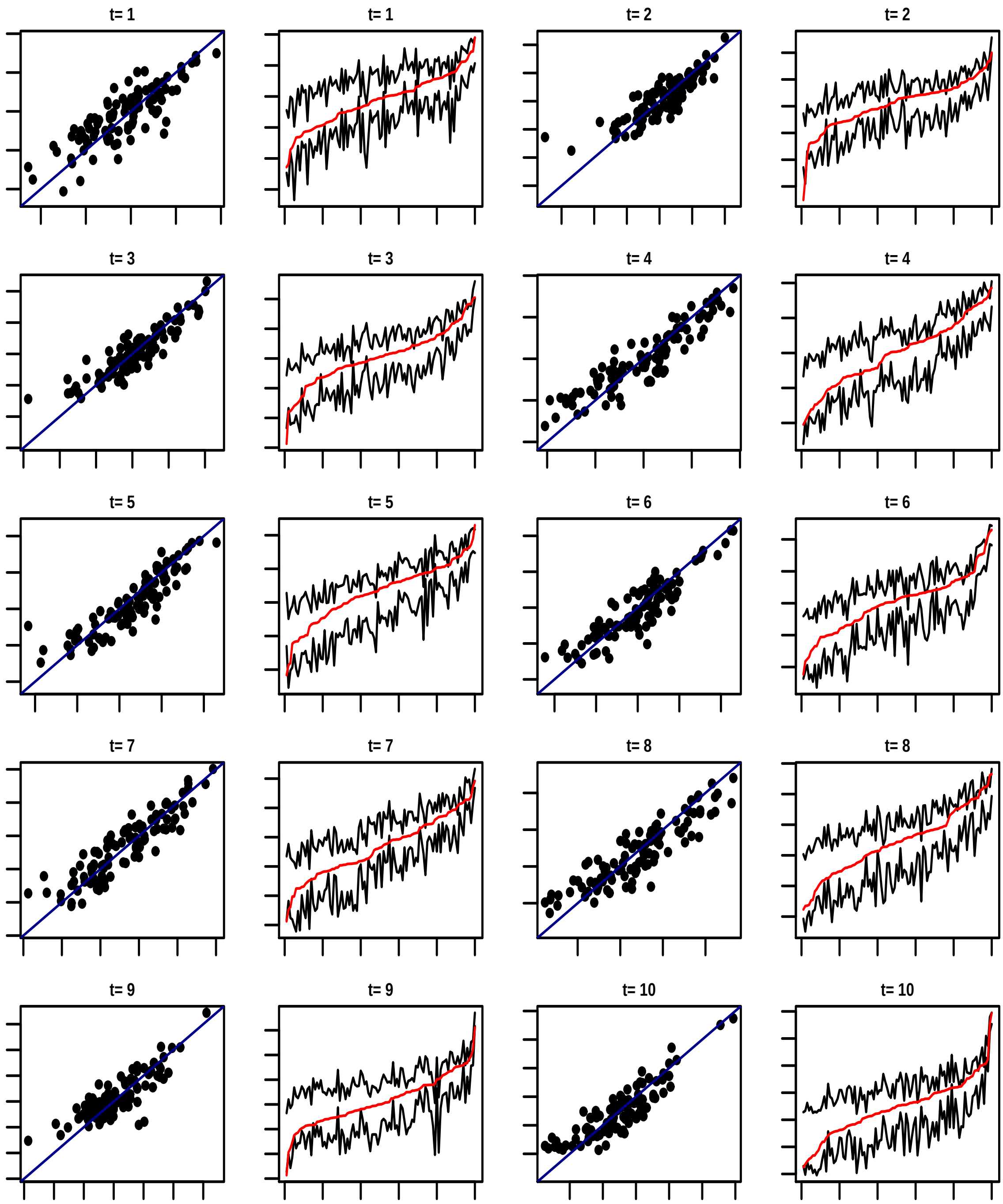


Figura 5.6: Modelo estacionário: inferência dos efeitos ϕ nas 100 células nos 10 tempos. A primeira e a terceira colunas mostram as médias a posteriori *versus* os valores reais; a segunda e quarta colunas mostram os valores reais em ordem crescente (em vermelho) e seus respectivos intevalos de 90% de credibilidade. A leitura da seqüência dos tempos é feita da esquerda para direita, de cima para baixo.

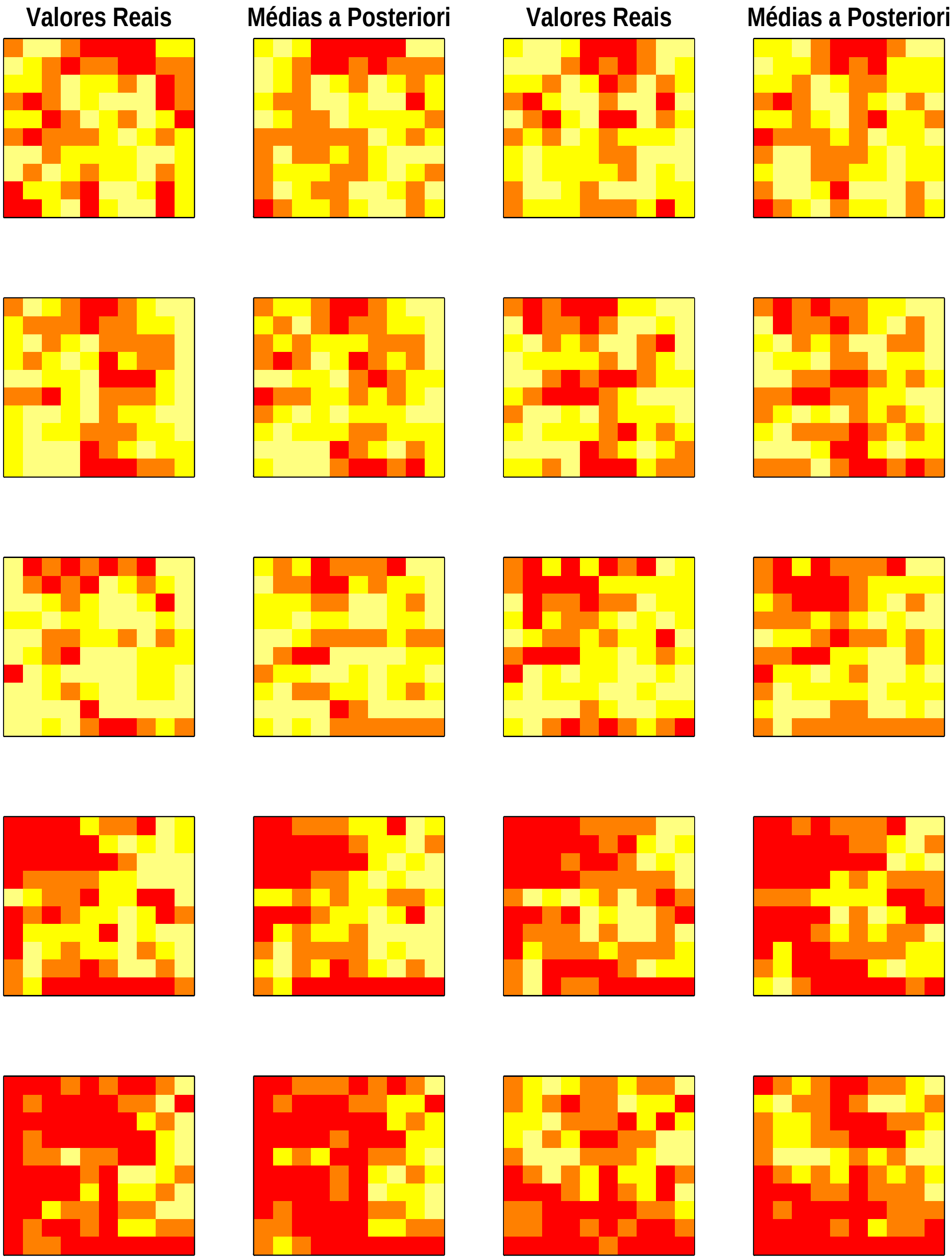


Figura 5.7: Modelo estacionário: imagens dos valores reais e médias a posteriori das log-intensidades $(\phi_{[i,t]} + \mu)$ nas 100 células nos 10 tempos. Os valores são crescentes de vermelho a amarelo-claro. A leitura da seqüência dos tempos é da esquerda para direita, de cima para baixo.

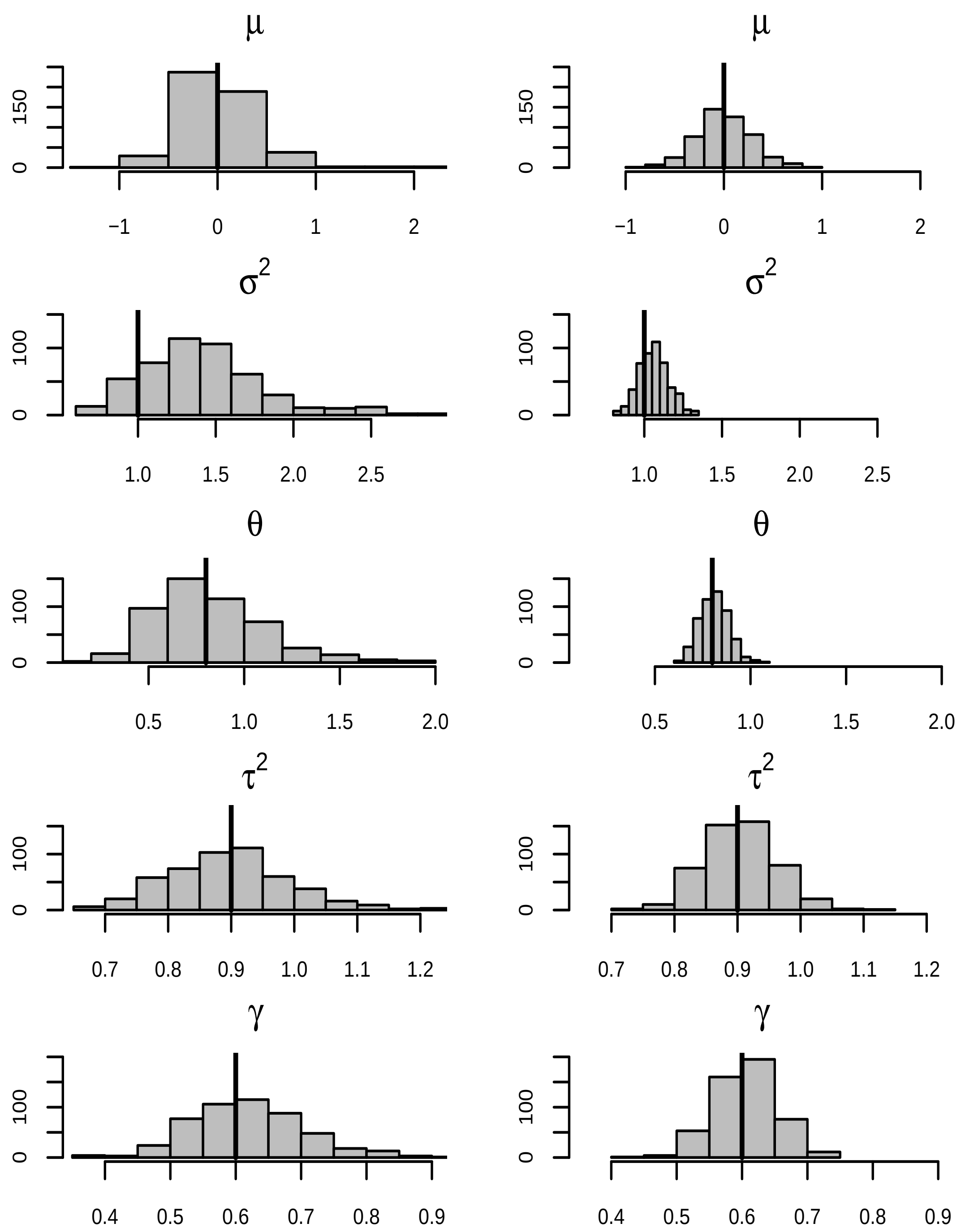


Figura 5.8: Modelo não-estacionário: histogramas das amostras a posteriori dos hiperparâmetros. Os gráficos na primeira coluna são relativos às prioris pouco informativas $\sigma^2 \sim GI[3;2]$, $\theta \sim G[1;1,25]$, $\tau^2 \sim GI[3;1,8]$ e $\gamma \sim G[1;1,67]$. Os gráficos na segunda coluna são relativos às prioris muito informativas $\sigma^2 \sim GI[102;101]$, $\theta \sim G[100;125]$, $\tau^2 \sim GI[102;90,9]$ e $\gamma \sim G[100;166,7]$. Em ambos os casos foram escolhidas a prioris $\pi(\mu) \propto 1$ e $\eta \sim U[0;1]$. O traço vertical marca o valor real do parâmetro.

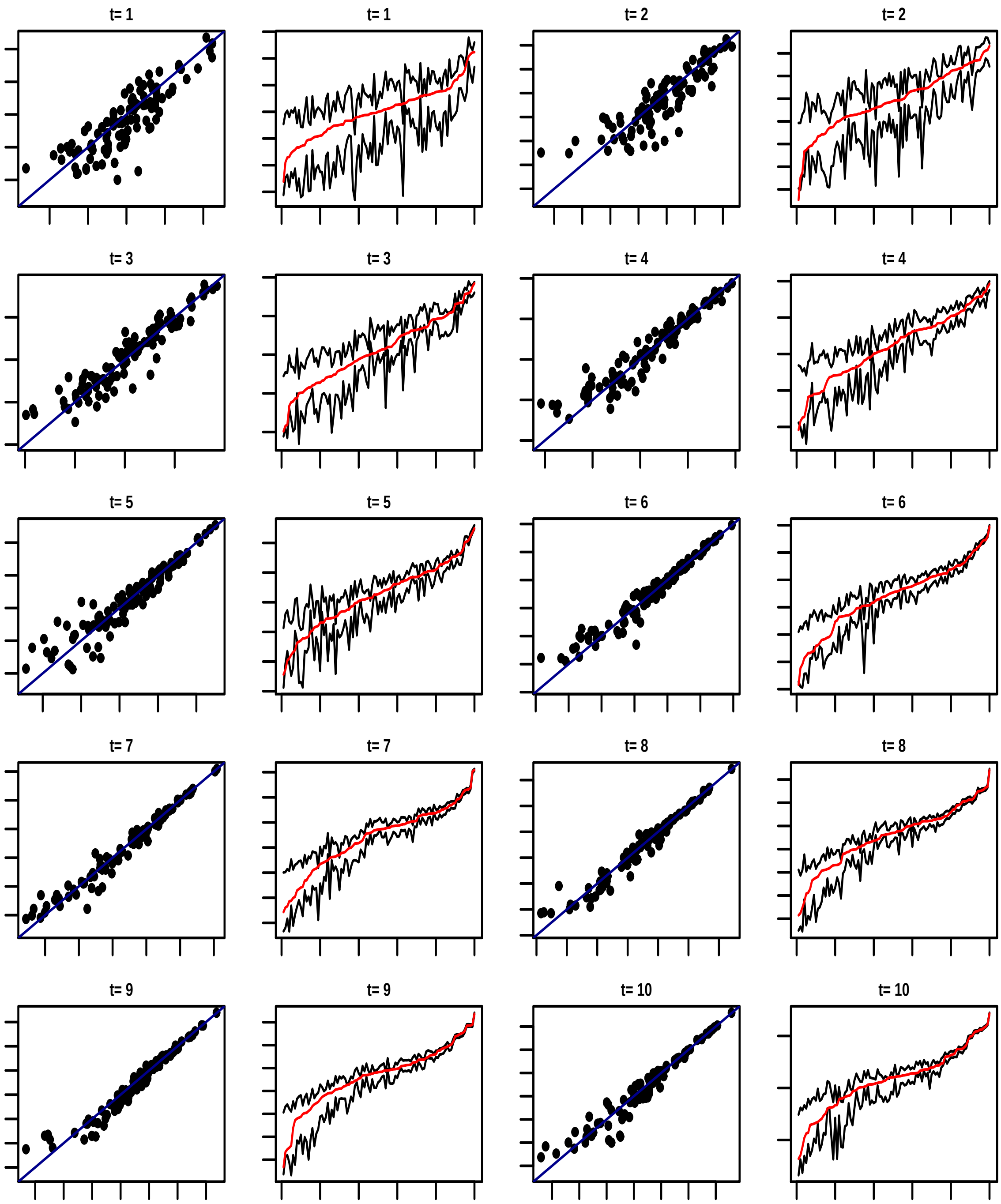


Figura 5.9: Modelo não-estacionário: inferência dos efeitos ϕ nas 100 células nos 10 tempos. A primeira e a terceira colunas mostram as médias a posteriori *versus* os valores reais; a segunda e quarta colunas mostram os valores reais em ordem crescente (em vermelho) e seus respectivos intevalos de 90% de credibilidade. A leitura da seqüência dos tempos é feita da esquerda para direita, de cima para baixo.

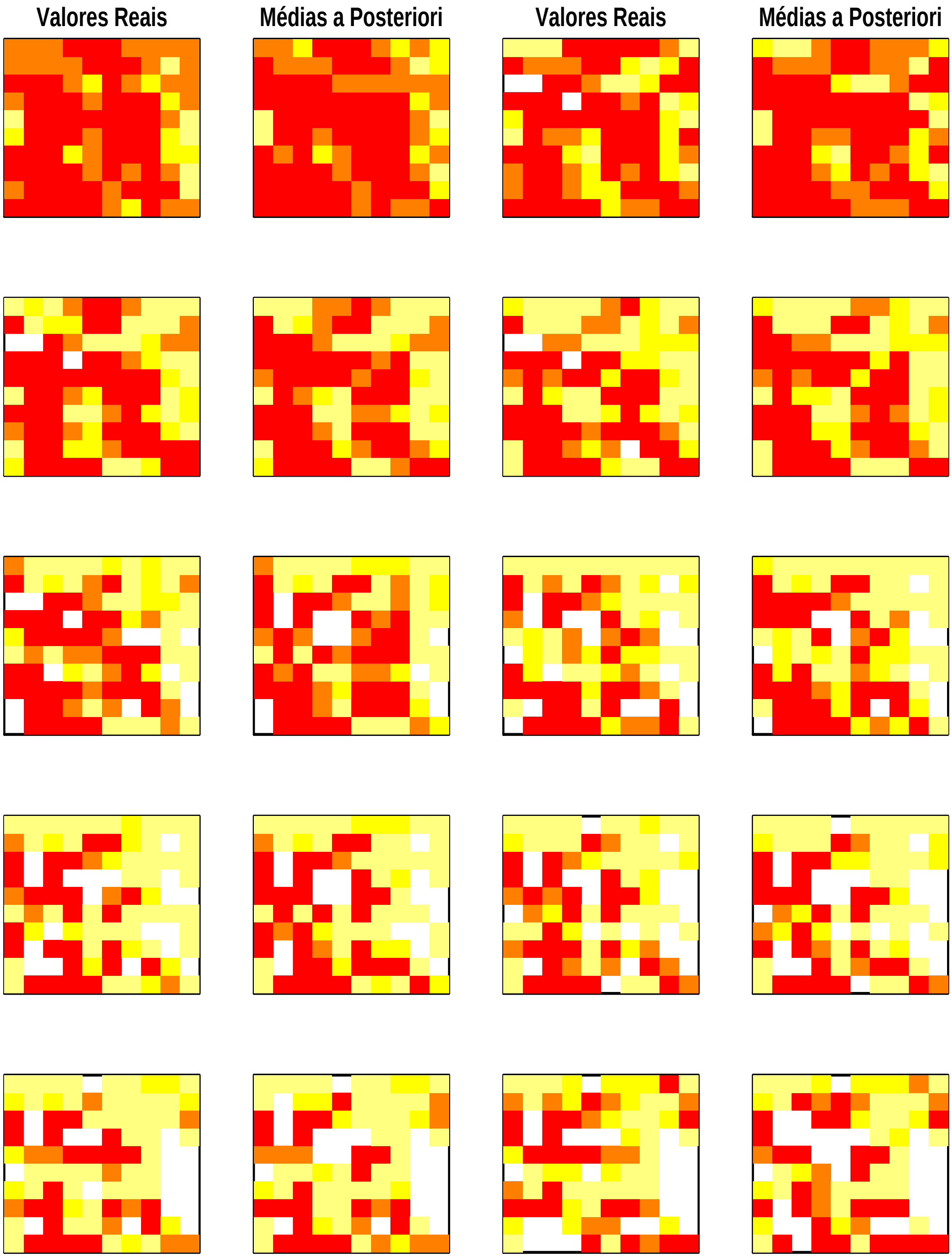


Figura 5.10: Modelo não-estacionário: imagens dos valores reais e médias a posteriori das log-intensidades $(\phi_{[i,t]} + \mu)$ nas 100 células nos 10 tempos. Os valores são crescentes de vermelho a branco. A leitura da seqüência dos tempos é da esquerda para direita, de cima para baixo.

5.3 Tendência Temporal Determinística Linear

Nesta seção, assume-se que, no modelo (5.1), a tendência temporal linear no tempo

$$\mu(t)=\beta_0+\beta_1\cdot t,\ \forall t,$$

e efeitos espaço-temporais $\phi(s,t)$ modelados como processos gaussianos autoregressivos estacionários no tempo

$$\phi(s,t) \ = \ \eta\,\phi(s,t\text{–}1) + \omega(s,t), \quad \omega(\cdot,t)\sim PG\left[0;(1\text{–}\eta^2)\sigma^2;\rho(\cdot;\theta)\right]$$

com $0<\eta<1$ e $\phi(\cdot,1)\sim PG\,[0;\sigma^2;\rho(\cdot;\theta)]$.

A região espacial da simulação é um quadrado dividido em uma grade regular com $N=400$ células (20×20), cada uma com área unitária. A janela de observação no tempo é formada por $T=20$ intervalos de tempo equiespaçados.

Os valores escolhidos para os parâmetros foram $\beta_0=-2$, $\beta_1=0,15$, $\sigma^2=0,1$, $\theta=0,2$ e $\eta=0,8$. A Figura 5.11 mostra os mapas com os valores dos efeitos espaciais $\phi_{[i,t]}$ e a localização dos eventos gerados.

As prioris pouco informativas escolhidas foram: $\sigma^2\sim GI[1;1]$, $\theta\sim G[1;1]$, $\eta\sim U[0;1]$ e $\boldsymbol{\beta}=(\beta_0,\beta_1)'\sim N[\,\underline{0}\,;\,0{,}1\boldsymbol{I}\,]$. Foram geradas duas cadeias de tamanho 100 mil, uma para cada conjunto de valores iniciais dos parâmetros. As duas amostras a posteriori de cada parâmetro são compostas de 1000 valores tomados a cada 50 da segunda metade das respectivas cadeias.

A Figura 5.12 mostra os resultados para os parâmetros σ^2, θ, η, β_0 e β_1. A inspeção visual do traço das cadeias (lado esquerdo da figura) para estes parâmetros não mostra sinais de que as cadeias não tenham convergido. Verifica-se, nos histogramas das amostras a posteriori (lado direito da figura), que os resultados da estimação destes parâmetros foram bastante satisfatórios.

Os resultados de estimação dos efeitos espaciais $\phi_{[i,t]}$ são mostrados na Figura 5.13. A comparação das estimativas dadas pelos médias a posteriori mostra que a estimação também foi satisfatória para estes efeitos. Verifica-se que a variação das estimativas é maior para os efeitos mais baixos, o que era esperado, dado que estes efeitos estão relacionados à células com menor número de eventos.

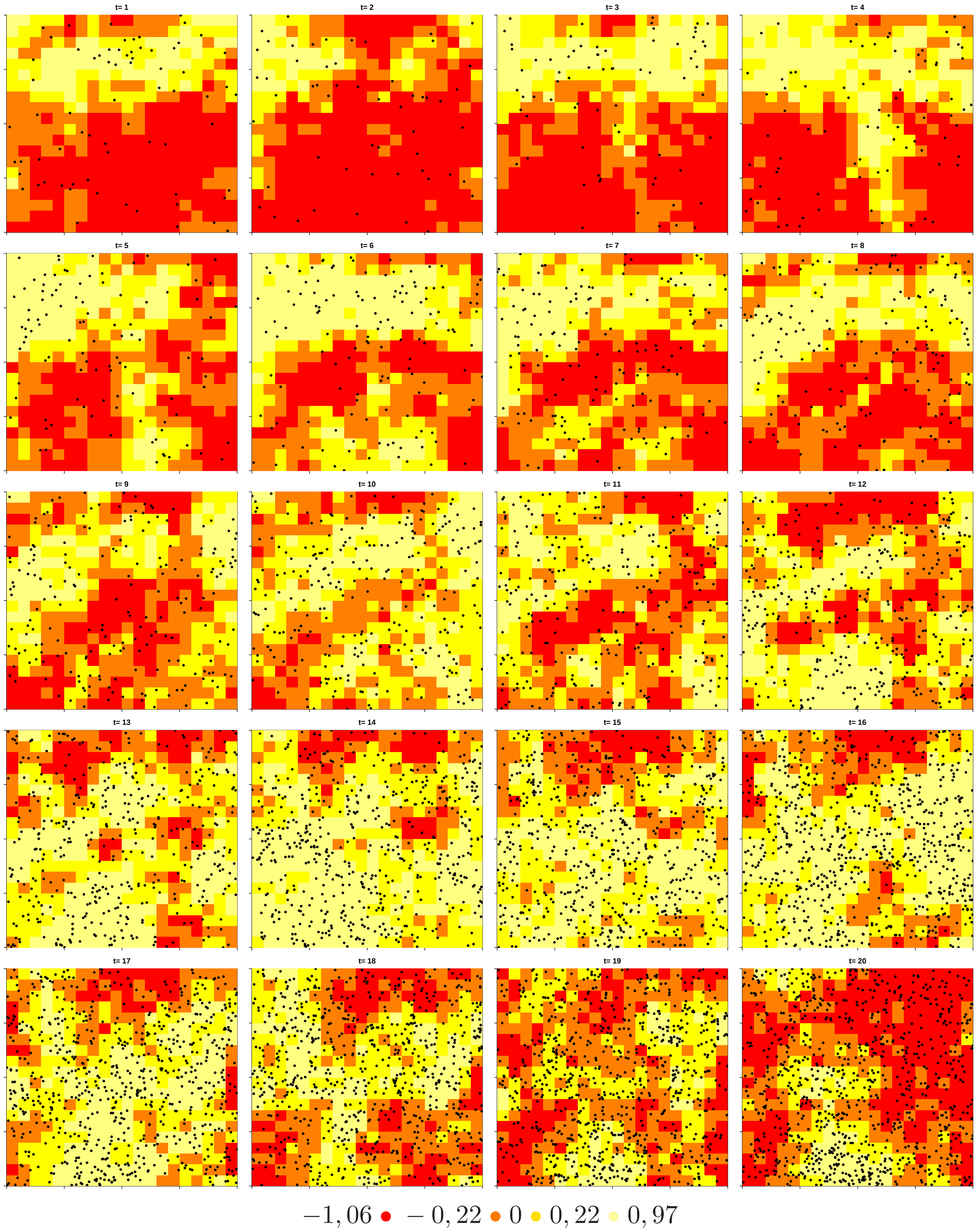


Figura 5.11: Modelo com tendência temporal linear: mapas dos efeitos espaciais reais $\phi_{[\cdot,t]}$, para $t=1,...,20$, e localização dos eventos gerados.

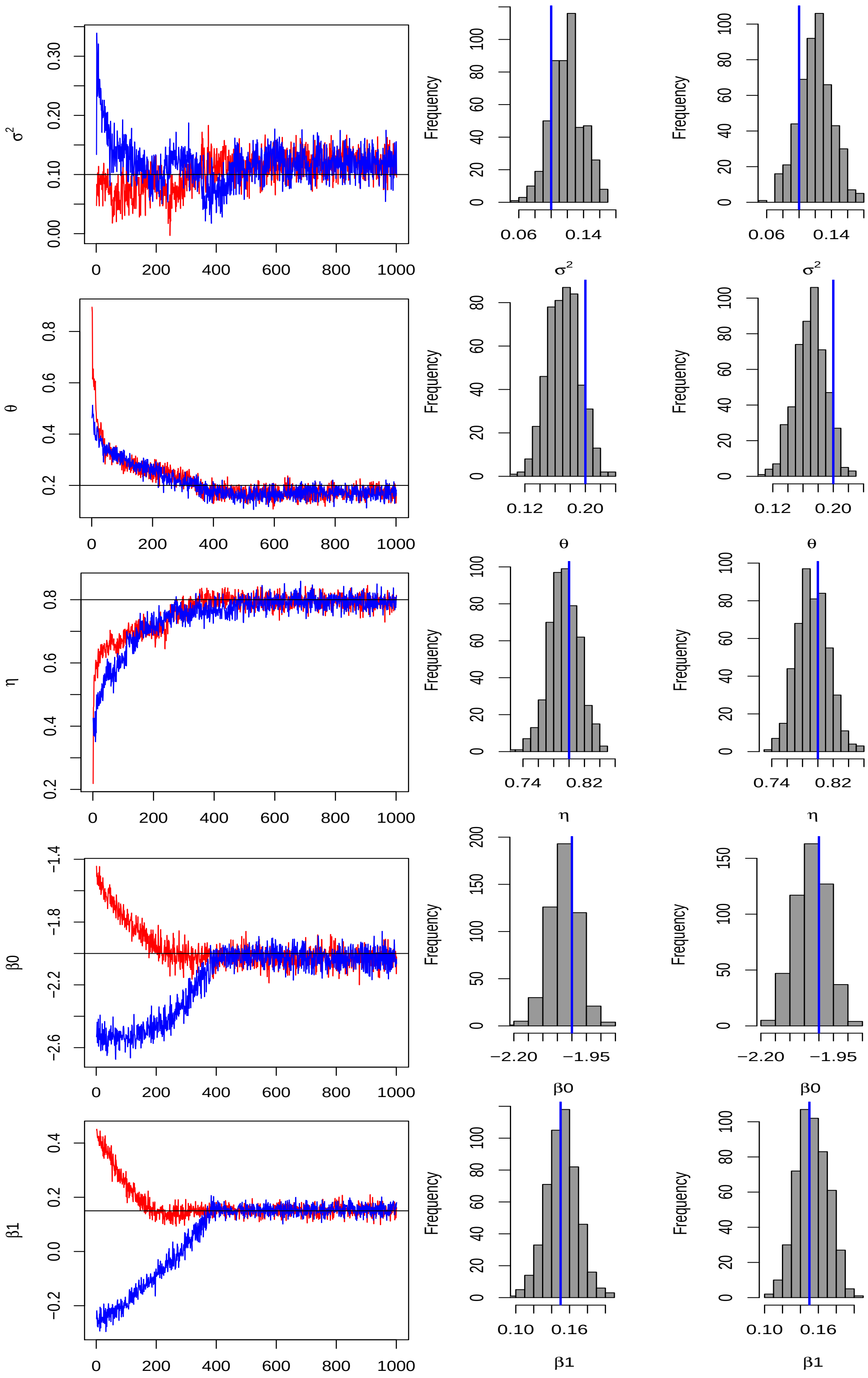


Figura 5.12: Modelo com tendência temporal linear: resultados de estimação dos hiperparâmetros σ^2, θ, η, β_0 e β_1. Lado esquerdo: traços da duas cadeias geradas, mostradas a cada 100 iterações. Lado direito: histogramas das duas amostras a posteriori.

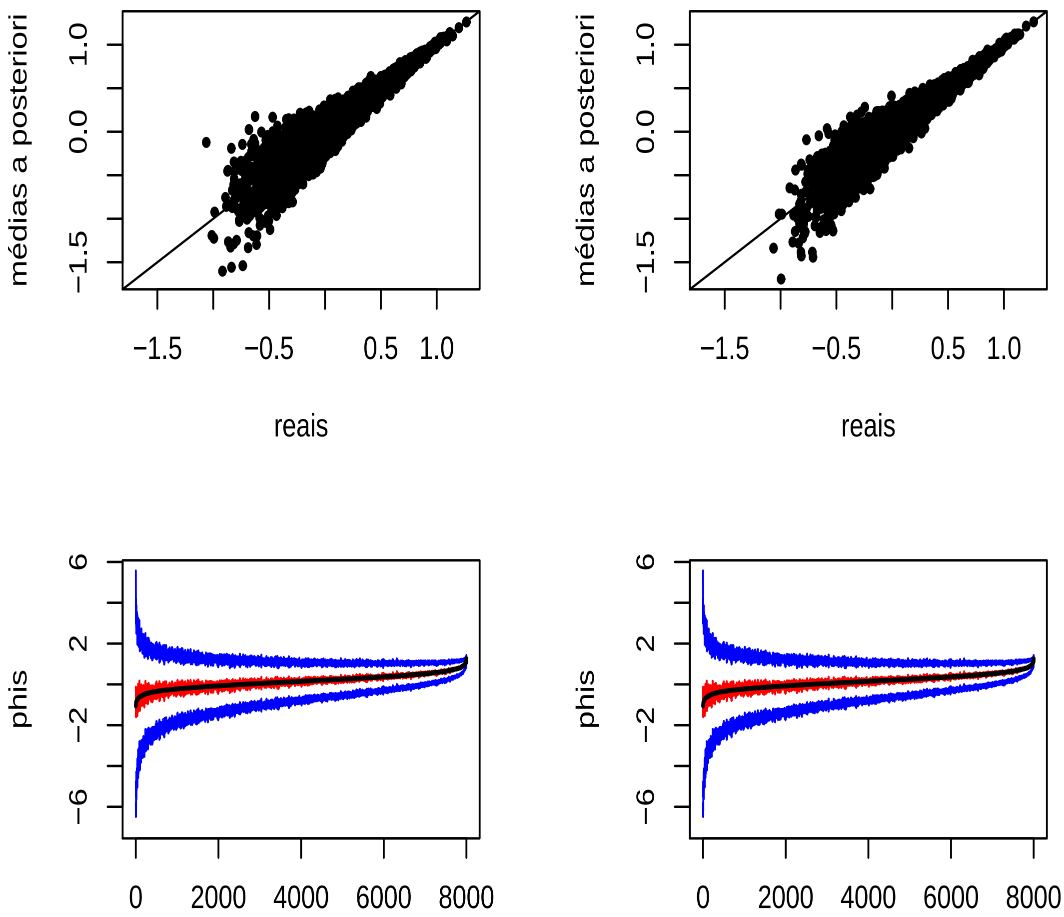


Figura 5.13: Modelo com tendência temporal linear: resultados de estimação dos efeitos espaciais $\phi_{[i,t]}$ nas duas amostras geradas. Lado superior: diagramas de dispersão dos valores reais versus as médias a posteriori. Lado inferior: médias a posteriori (em vermelho), intervalos de 90% de credibilidade (em azul) e valores reais (em preto). Os valores no eixo das ordenadas estão dispostos em ordem crescente dos valores reais.

5.4 Tendência Temporal Dinâmica Polinomial de Primeira Ordem

Nesta seção, assume-se que, no modelo (5.1), a tendência temporal estocástica dinâmica de primeira ordem

$$\mu_{[t]} = \mu_{[t-1]} + \upsilon_{[t]}, \qquad \upsilon_{[t]} \sim N\left[0; \omega^2\right],\ t=2,...,T, \quad \mu_{[1]} \sim N\left[\mu_0; \tau_0^2\right],$$

e efeitos espaço-temporais $\phi(s,t)$ modelados como processos gaussianos autoregressivos estacionários no tempo

$$\phi(s,t) = \eta\, \phi(s,t\text{–}1) + \omega(s,t), \quad \omega(\cdot,t) \sim PG\left[0; (1\text{–}\eta^2)\sigma^2; \rho(\cdot;\theta)\right]$$

com $0<\eta<1$ e $\phi(\cdot,1) \sim PG\left[0; \sigma^2; \rho(\cdot;\theta)\right]$.

A região espacial e janela temporal da simulação são as mesmas do modelo de tendência linear ($N=400$ células e $T=20$ tempos).

Os valores escolhidos para os parâmetros foram $\sigma^2=0,1$, $\theta=0,2$, $\eta=0,8$ e $\omega^2=0,01$. A Figura 5.14 mostra as somas dos efeitos espaço-temporais $\phi_{[i,t]}$ e temporais $\mu_{[t]}$ reais e localização dos eventos gerados em cada intervalo de tempo.

As prioris pouco informativas escolhidas foram $\sigma^2 \sim GI[1;1]$, $\theta \sim G[1;1]$, $\eta \sim U[0;1]$ e $\omega^2 \sim GI[1;1]$. Foram geradas duas cadeias de tamanho 100 mil, uma para cada valor inicial das quantidades a serem estimadas. As duas amostras a posteriori de cada parâmetro são compostas de 1000 valores tomados a cada 50 da segunda metade das respectivas cadeias.

A Figura 5.15 mostra os resultados para os parâmetros σ^2, θ, η e ω^2. A inspeção visual do traço das cadeias (lado esquerdo da figura) para estes parâmetros não mostra sinais de que as cadeias não tenham convergido. Verifica-se, nos histogramas das amostras a posteriori (lado direito da figura), que os resultados da estimação destes parâmetros foram bastante satisfatórios.

Os resultados de estimação dos efeitos espaciais $\phi_{[i,t]}$ são mostrados na Figura 5.16. A comparação das estimativas dadas pelos médias a posteriori mostra que a estimação também foi satisfatória para estes efeitos. Verifica-se novamente que a variação das estimativas é maior para os efeitos mais baixos. Da mesma forma, a Figura 5.17 mostra que os efeitos puramente temporais $\mu_{[t]}$ foram muito bem estimados.

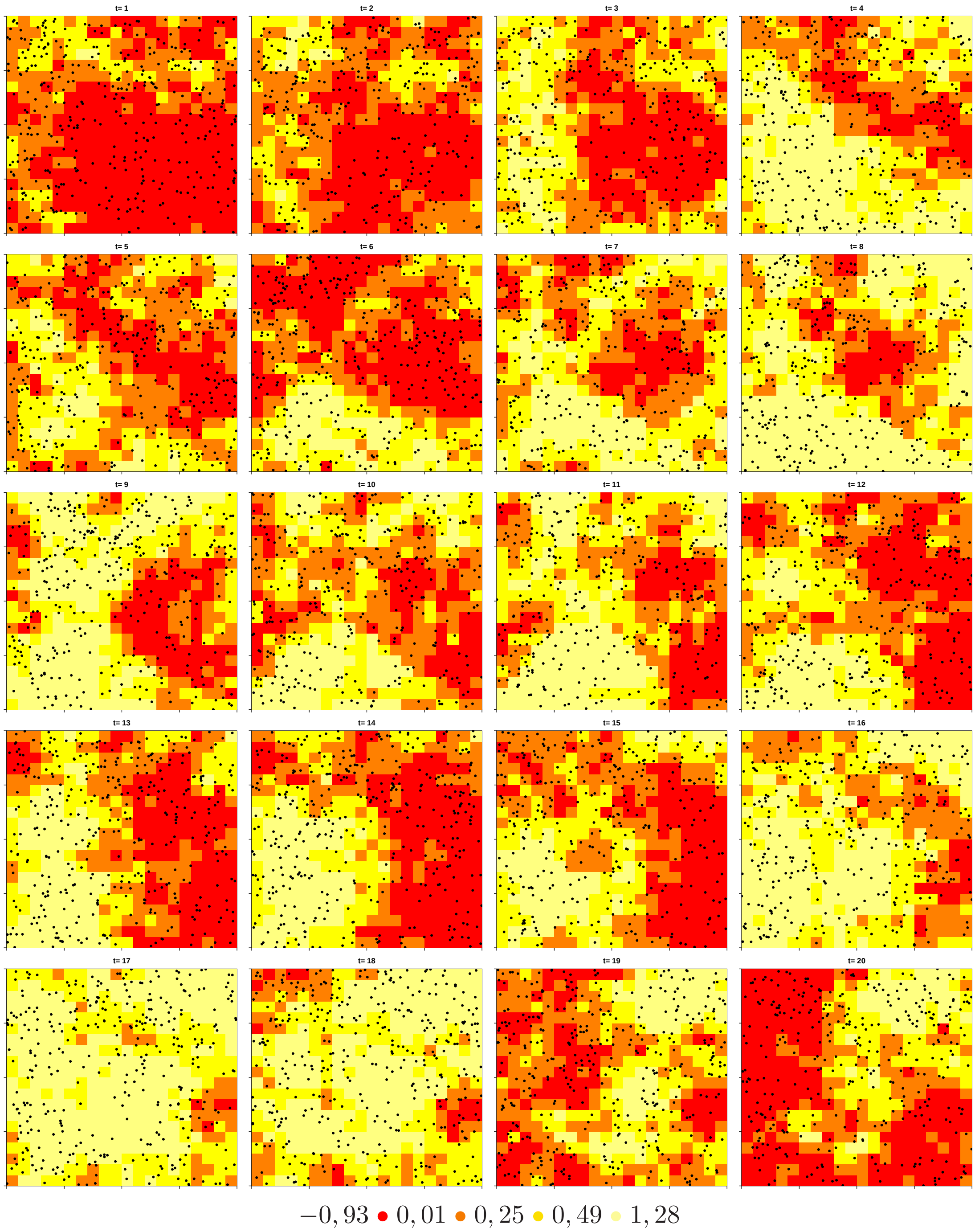


Figura 5.14: Modelo com tendência temporal dinâmica polinomial de primeira ordem: mapas das somas dos efeitos espaciais $\phi_{[\cdot,t]}$ e temporais $\mu_{[t]}$ reais, para $t = 1, ..., 20$, e localização dos eventos gerados.

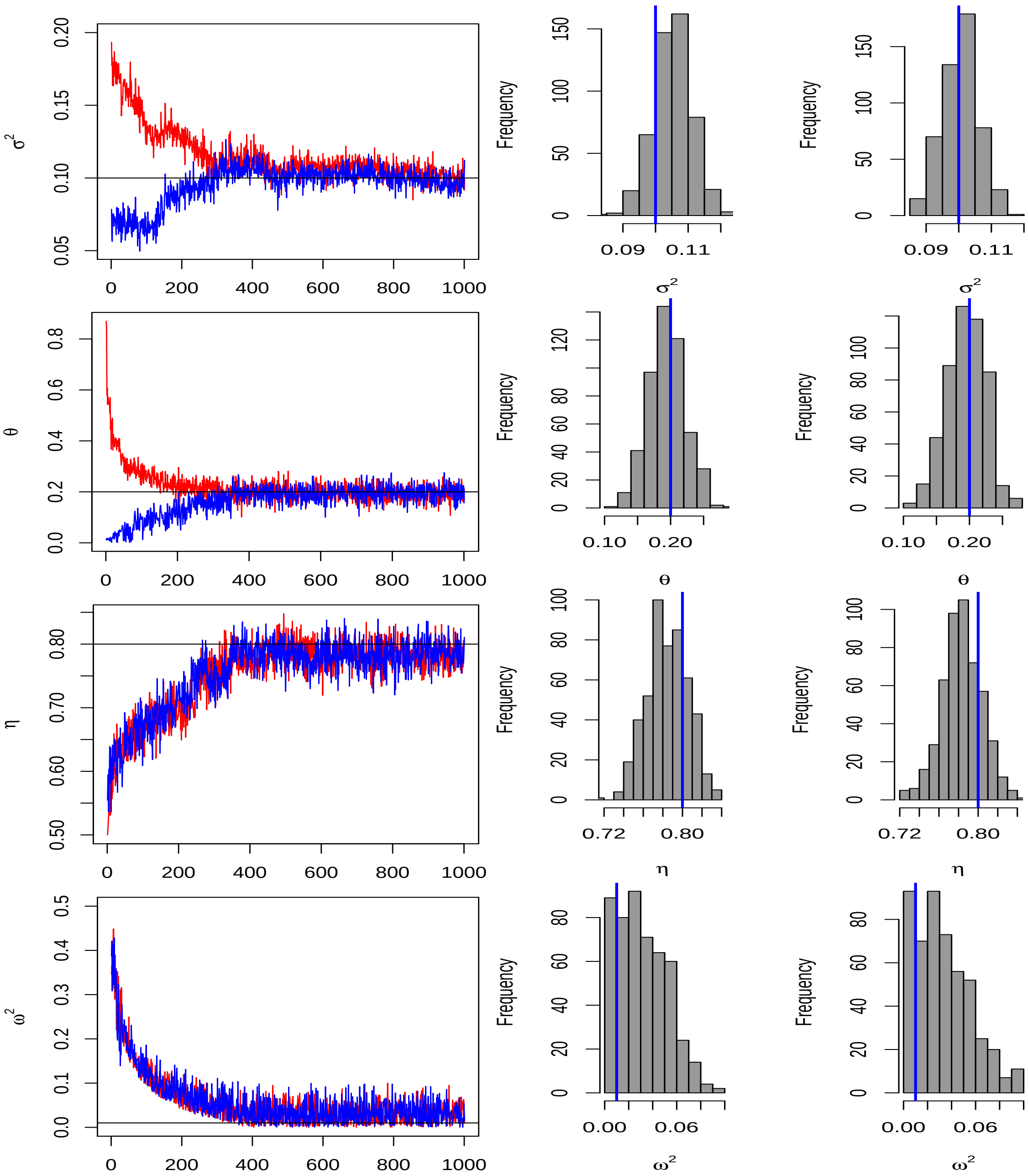


Figura 5.15: Modelo com tendência temporal dinâmica polinomial de primeira ordem: resultados de estimação dos hiperparâmetros σ^2, θ, η e ω^2. Lado esquerdo: traços da duas cadeias geradas, mostradas a cada 100 iterações. Lado direito: histogramas das duas amostras a posteriori.

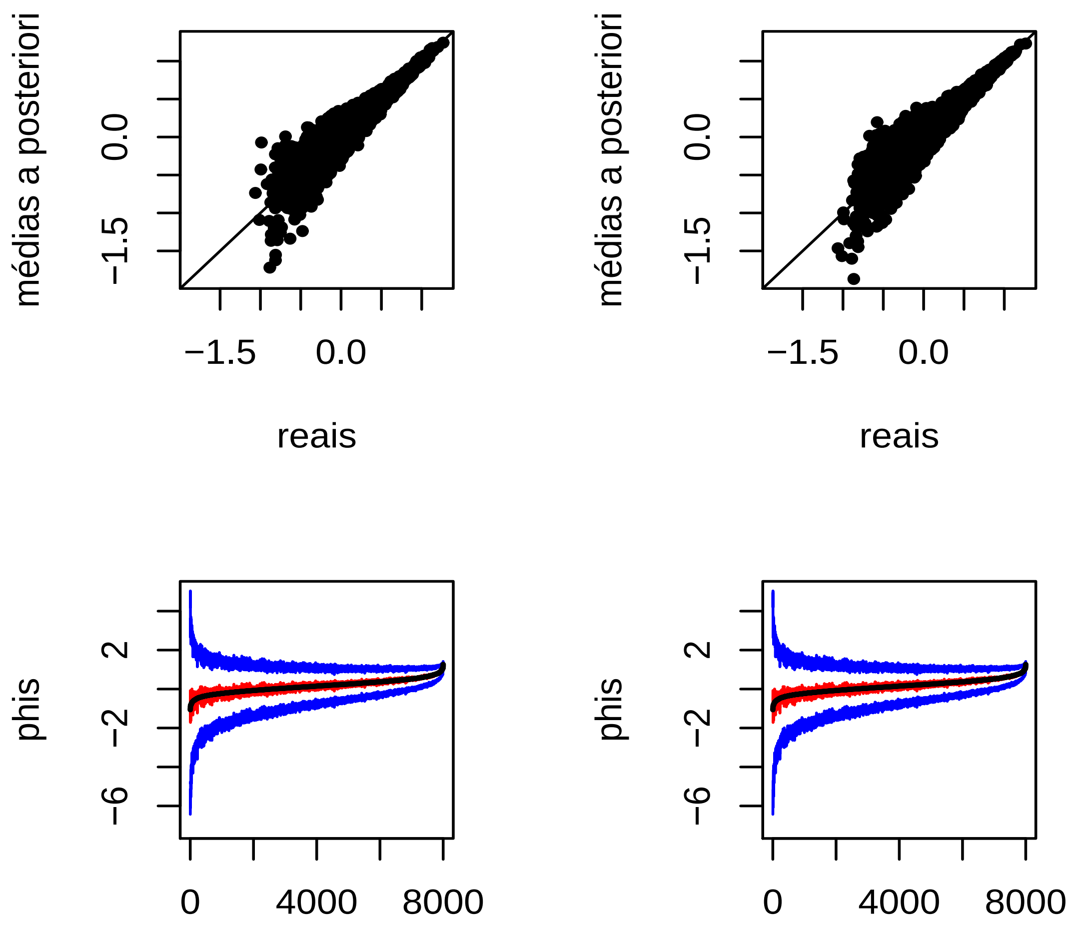


Figura 5.16: Modelo com tendência temporal dinâmica polinomial de primeira ordem: resultados de estimação dos efeitos espaciais $\phi_{[i,t]}$ nas duas amostras geradas. Lado superior: diagramas de dispersão dos valores reais versus as médias a posteriori. Lado inferior: valores reais (em preto), médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul). Os valores no eixo das ordenadas estão dispostos em ordem crescente dos valores reais.

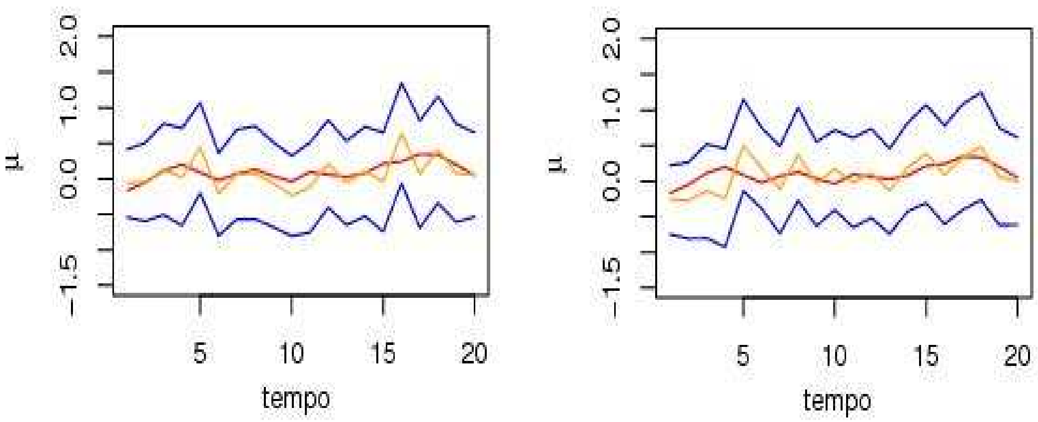


Figura 5.17: Modelo com tendência temporal dinâmica polinomial de primeira ordem: resultados de estimação dos efeitos temporais $\mu_{[t]}$ nas duas amostras geradas: valores reais (em alaranjado), médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul).

5.5 Tendência Temporal Dinâmica Polinomial de Segunda Ordem

Nesta seção, assume-se que, no modelo (5.1), a tendência temporal estocástica dinâmica

$$\begin{aligned}
\mu_{[t]} &= \mu_{[t-1]} + \beta_{[t-1]} + \upsilon_{1[t]}, \quad && \upsilon_{1[t]} \sim N\left[0; \omega_1^2\right], \ t=2,...,T, \quad && \mu_{[1]} \sim N\left[\mu_0; \tau_0^2\right], \\
\beta_{[t]} &= \beta_{[t-1]} + \upsilon_{2[t]}, && \upsilon_{2[t]} \sim N\left[0; \omega_2^2\right], \ t=2,...,T, && \beta_{[1]} \sim N\left[\beta_0; \kappa_0^2\right].
\end{aligned}$$

e efeitos espaço-temporais $\phi(s,t)$ modelados como processos gaussianos autoregressivos estacionários no tempo

$$\phi(s,t) \ = \ \eta\, \phi(s,t{-}1) + \omega(s,t), \quad \omega(\cdot,t) \sim PG\left[0; (1{-}\eta^2)\sigma^2; \rho(\cdot;\theta)\right]$$

com $0<\eta<1$ e $\phi(\cdot,1) \sim PG\left[0; \sigma^2; \rho(\cdot;\theta)\right]$.

A região espacial e janela temporal da simulação são as mesmas dos dois modelos anteriores ($N=400$ células e $T=20$ tempos).

Os valores escolhidos para os parâmetros foram $\sigma^2 = 0,1$, $\theta = 0,2$, $\eta = 0,8$, $\omega_1^2 = 0,01$ e $\omega_2^2 = 0,0025$. A Figura 5.18 mostra as somas dos efeitos espaço-temporais $\phi_{[i,t]}$ e temporais $\mu_{[t]}$ reais e localização dos eventos gerados em cada intervalo de tempo.

As prioris pouco informativas utilizadas foram $\sigma^2 \sim GI[1;1]$, $\theta \sim G[1;1]$, $\eta \sim U[0;1]$, $\omega_1^2 \sim GI[0,1;0,1]$ e $\omega_2^2 \sim GI[0,1;0,1]$. Foram geradas duas cadeias de tamanho 100 mil, uma para cada valor inicial das quantidades a serem estimadas. As duas amostras a posteriori de cada parâmetro são compostas de 1000 valores tomados a cada 50 da segunda metade das respectivas cadeias.

A Figura 5.19 mostra os resultados para os parâmetros σ^2, θ, η, ω_1^2 e ω_2^2. A inspeção visual do traço das cadeias (lado esquerdo da figura) para estes parâmetros não mostra sinais de que as cadeias não tenham convergido. Verifica-se, nos histogramas das amostras a posteriori (lado direito da figura), que os resultados da estimação destes parâmetros foram bastante satisfatórios.

Os resultados de estimação dos efeitos espaciais $\phi_{[i,t]}$ são mostrados na Figura 5.20. A comparação das estimativas dadas pelos médias a posteriori mostra que a estimação também foi satisfatória para estes efeitos. Verifica-se que a variação das estimativas é maior para os efeitos mais baixos, o que era esperado, dado que estes efeitos estão relacionados à células com menor número de eventos. Da mesma forma, a Figura 5.21 e 5.22 mostram que os efeitos puramente temporais $\mu_{[t]}$ e $\beta_{[t]}$ foram bem estimados.

5.6 Conclusões

Os resultados dos estudos simulados mostraram que os modelos podem ser bem reconhecidos pelos dados, com um boa concordância das estimativas com os valores reais.

Embora existam algumas técnicas de verificação da convergência das cadeias, optou-se por verificá-la pela observação do traços de duas cadeias independentes em alguns parâmetros. Não foi possível armazenar todas as cadeias de todos efeitos (seriam, por exemplo, 8045 cadeias no útimo modelo). A convergência das cadeias é atingida em um número relativamente baixo de iterações, tendo em vista o grande número de efeitos e parâmetros a serem estimados.

Os algoritmos foram codificados no programa Ox (Doornik, 2002) e rodaram em computador doméstico (processador AMD Athlon XP 2200, 1.8Ghz, 1.0 GB RAM). O tempo de processamento de cada cadeia é muito grande, cerca de 100 horas para os casos com as grades de 400 células e 20 intervalos de tempo.

Um estudo mais amplo teria diferentes combinações de parâmetros, prioris e com réplicas em cada um destas combinações, mas o tempo dispendido seria muito grande para o prazo disponível.

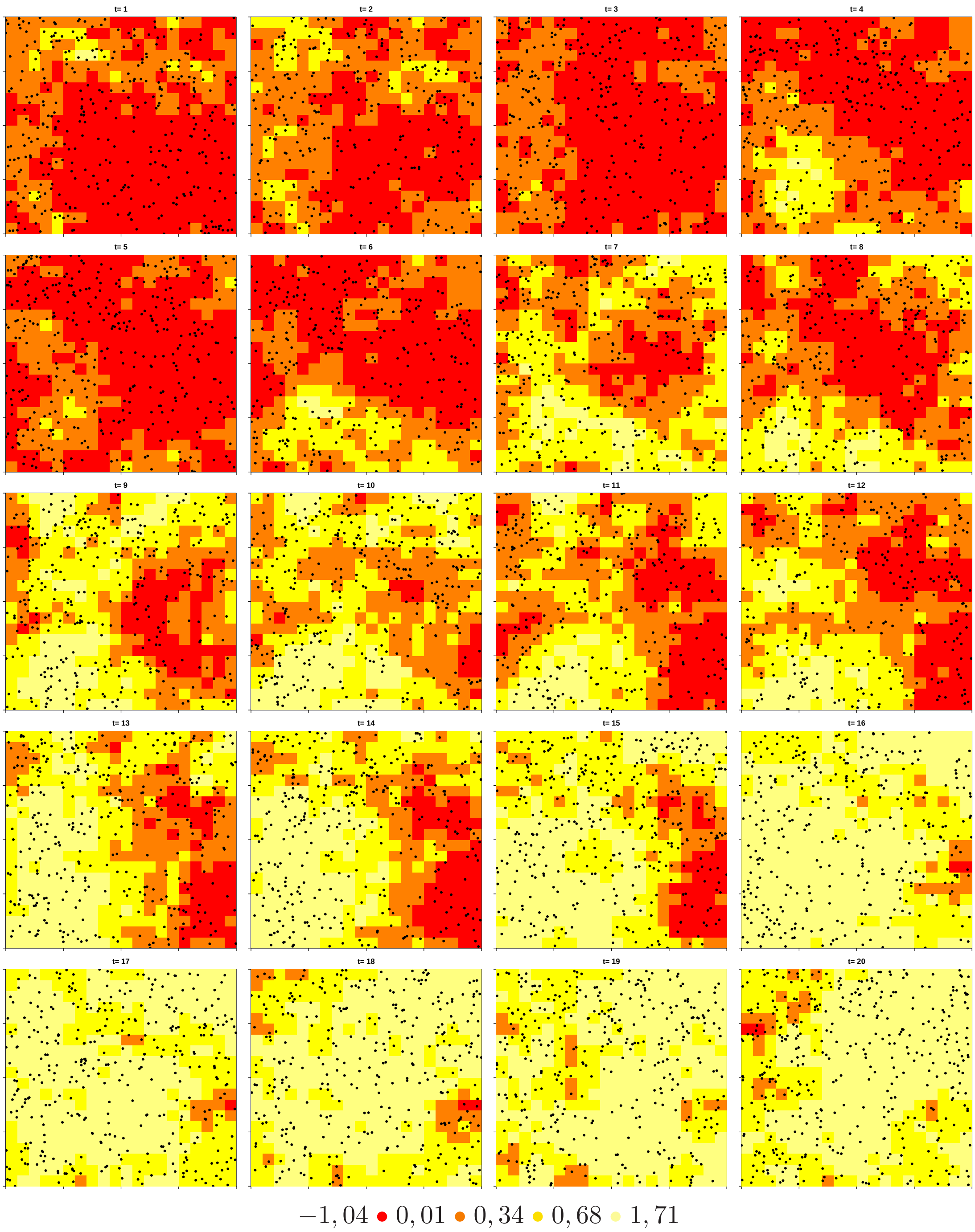


Figura 5.18: Modelo com tendência temporal dinâmica polinomial de segunda ordem: mapas das somas dos efeitos espaciais $\phi_{[\cdot,t]}$ e temporais $\mu_{[t]}$ reais, para $t = 1, ..., 20$, e localização dos eventos gerados.

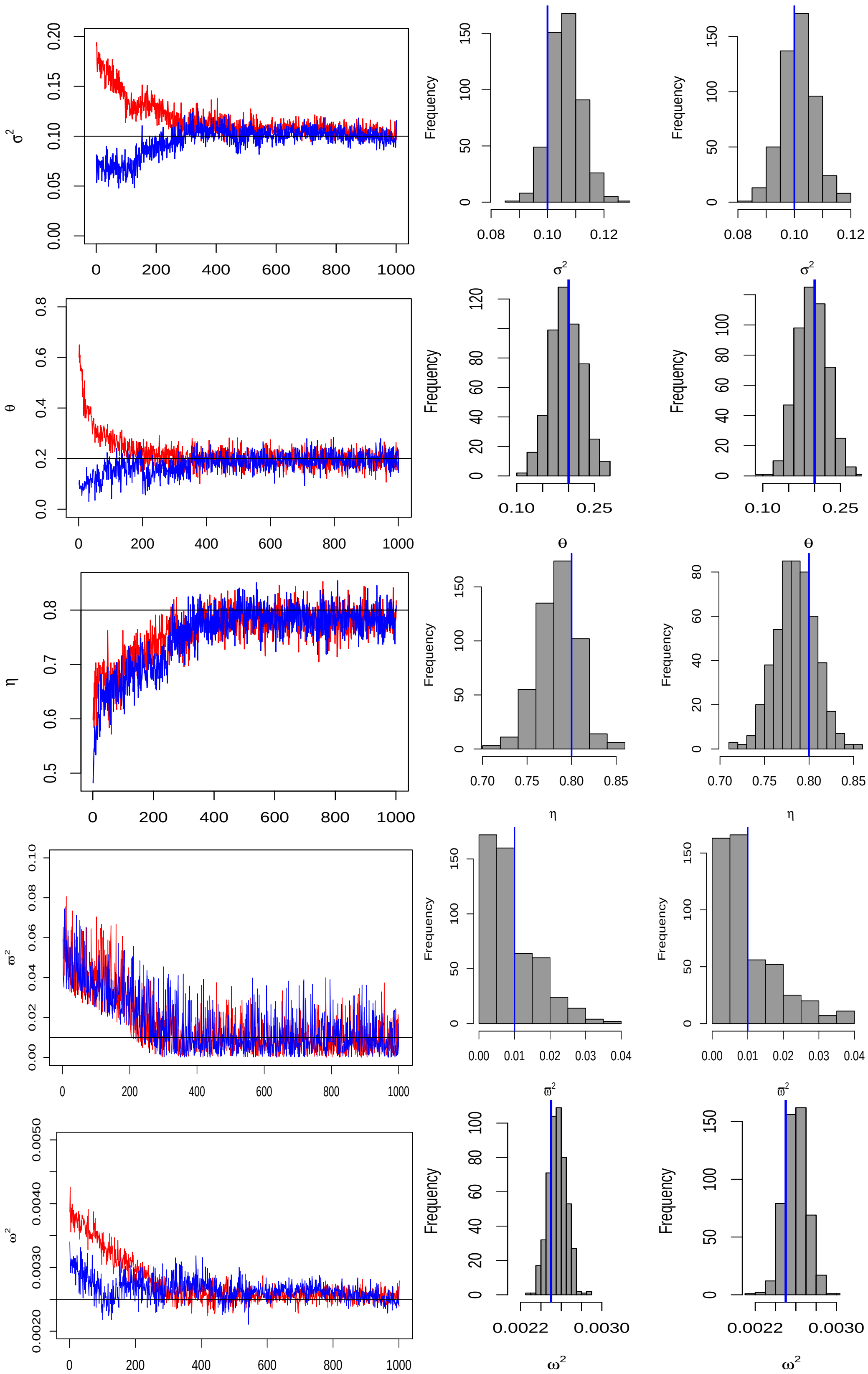


Figura 5.19: Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos hiperparâmetros σ^2, θ, η, ω_1^2 e ω_2^2. Lado esquerdo: traços da duas cadeias geradas, mostradas a cada 100 iterações. Lado direito: histogramas das duas amostras a posteriori.

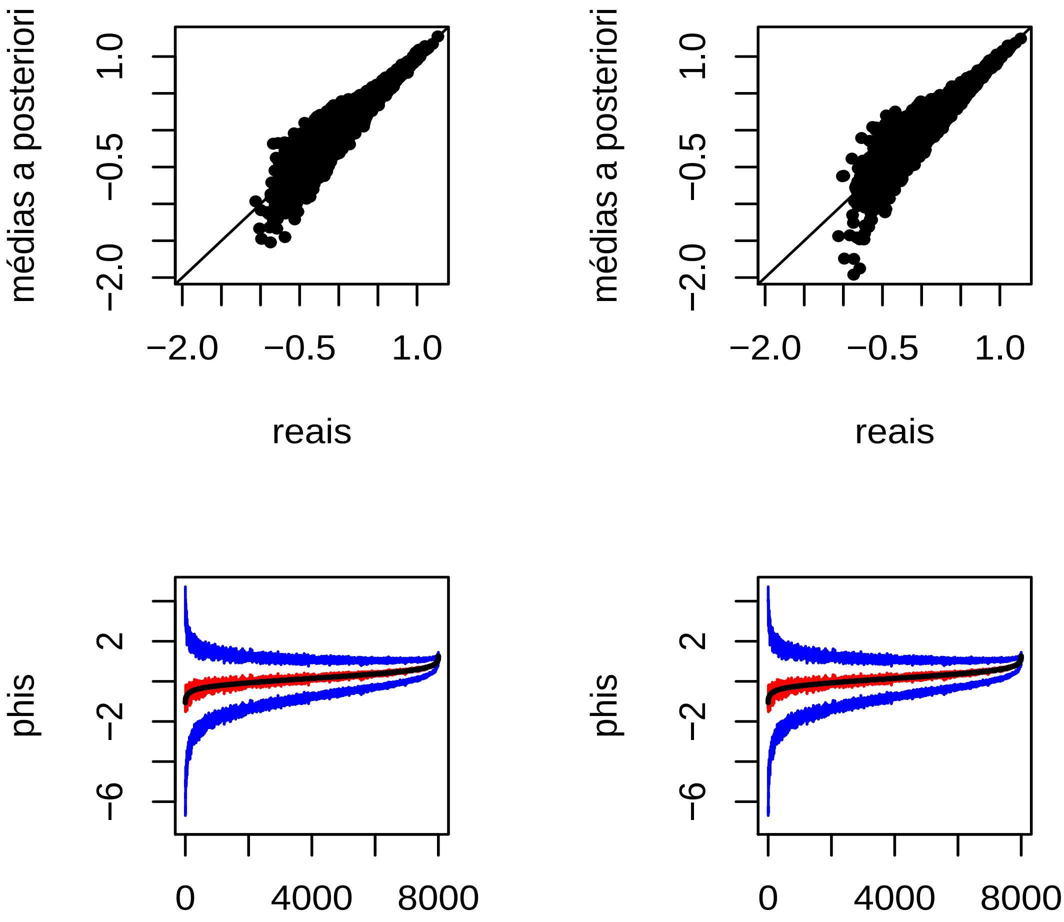


Figura 5.20: Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos efeitos espaciais $\phi_{[i,t]}$ nas duas amostras geradas. Lado superior: diagramas de dispersão dos valores reais versus as médias a posteriori. Lado inferior: valores reais (em preto), médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul). Os valores no eixo das ordenadas estão dispostos em ordem crescente dos valores reais.

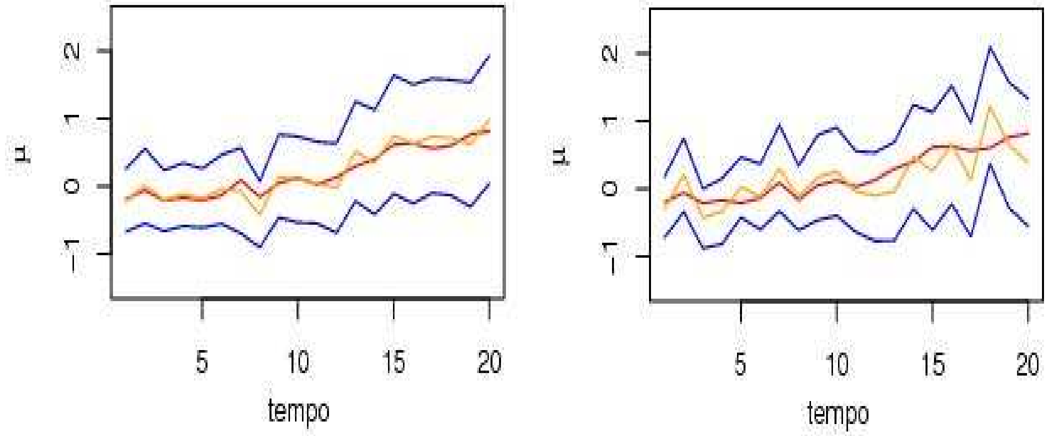


Figura 5.21: Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos efeitos temporais $\mu_{[t]}$ nas duas amostras geradas: valores reais (em alaranjado), médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul).

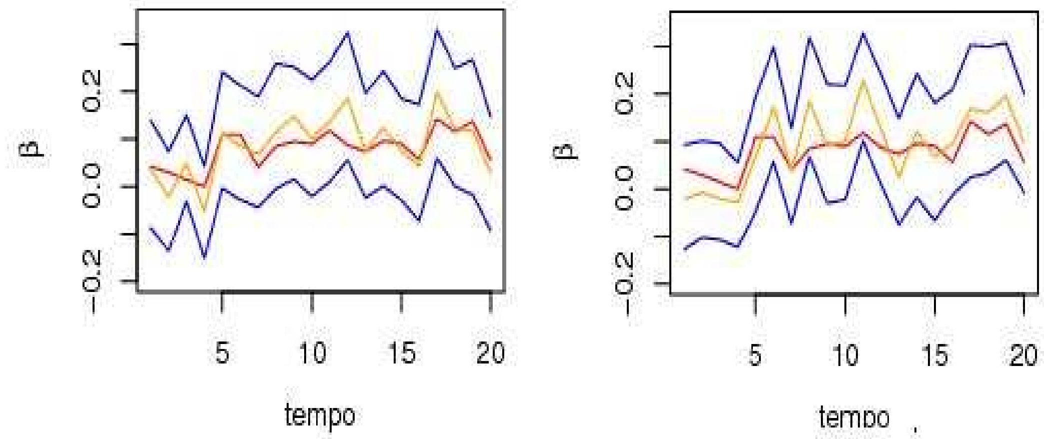


Figura 5.22: Modelo com tendência temporal dinâmica polinomial de segunda ordem: resultados de estimação dos efeitos temporais $\beta_{[t]}$ nas duas amostras geradas: valores reais (em alaranjado), médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul).

Capítulo 6

Aplicações

6.1 Introdução

Neste capítulo são apresentadas duas aplicações dos modelos propostos no Capítulo 4 a conjuntos de dados reais analisados na literatura. O primeiro conjunto de dados foi analisado por Diggle *et al.* (2005b) para vigilância epidemiológica em tempo real dos casos de doença gastrointestinal em Hampshire, no Reino Unido. A segunda aplicação consiste na análise de dados de neuro-gastroenterologia, o ramo da medicina que estuda o funcionamento dos neurônios no intestino. Os dados foram analisados originalmente em Faes *et al.* (2006).

6.2 Análise Espaço-Temporal dos Casos de Doença Gastrointestinal em Hampshire

O projeto AEGISS (*Ascertainment and Enhancement of Gastrointestinal Infection Surveillance and Statistics*) vem sendo desenvolvido na Grã-Bretanha com o objetivo de reduzir a ocorrência de doenças gastrointestinais. No condado de Hampshire, foram registrados 10752 casos de infecção gastrointenstinal não-específica nos anos de 2001 a 2003. Um *caso* da doença é definido como qualquer chamada telefônica ao serviço de orientação médica *NHS Direct* relatando sintomas infecção gastrointenstinal. Cada caso é identificado pela localização residencial da pessoa (coordenadas geográficas) e pela data da chamada.

A Figura 6.1 mostra o mapa do condado com a localização espacial dos casos nos três anos de estudo. Os eventos estão concentrados na região sul, área de mais alta densidade populacional.

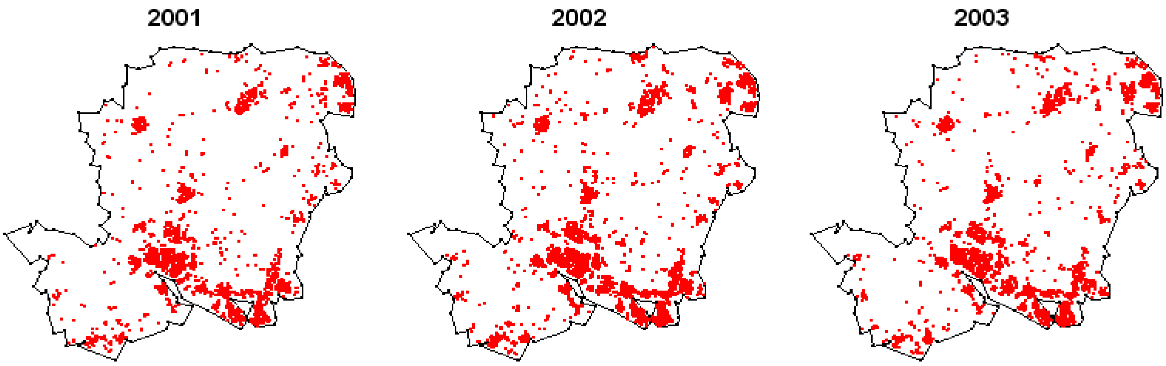


Figura 6.1: Mapa do contorno do condado de Hampshire e eventos observados em cada ano.

Diggle *et al.* (2005b) analisaram os casos diários dos dois primeiros anos, com foco na vigilância sanitária em tempo real para detecção precoce de variações localizadas não-explicadas na intensidade espaço-temporal $\lambda(s,t)$ na localização espacial s no tempo t. O modelo proposto por eles é um processo de Cox log-gaussiano não-estacionário com decomposição multiplicativa da log-intensidade em $log\,[\lambda(s,t)] = \mu(t)+\zeta(s)+\phi(s,t)$. Os componentes $\mu(t)$ e $\zeta(s)$ descrevem, respectivamente, as variações puramente temporal e puramente espacial na incidência normal da doença e são tratados como determinísticos. $S(s,t)=\exp\{\phi(s,t)\}$ é um componente estocástico não-observável que representa desvios espaço-temporalmente localizados, sendo modelado como um processo de Cox log-gaussiano estacionário, cujos parâmetros são estimados pelo método dos momentos propostos em Brix e Diggle (2001).

O padrão espacial de chamadas ao serviço de assistência não segue necessariamente aquele da população sob risco da doença. Portanto, o uso de contagens de população de censos para estimar a intensidade populacional $\lambda_0(s)=exp\{\zeta(s)\}$ não é adequado. Diggle *et al.* (2005b) usam a distribuição espacial de todos os casos dos dois anos de estudo para estimar o padrão de variação espacial normal da incidência da doença. Nesta tese, os casos do primeiro ano de observação foram usados para estimar a distribuição espacial da população sob risco e os modelos do Capítulo 4 foram aplicados aos casos dos dois anos seguintes.

Nesta tese, considera-se inicialmente a análise espaço-temporal dos totais de casos nos 24 meses dos anos de 2002 e 2003 (Figura 6.2). A discretização espacial é definida pela interseção da região de estudo com uma grade regular com 270 células sobreposta a ela (Figura 6.3), totalizando 168 células válidas. Como para as células sobrepostas à borda da região a interseção não é total, definimos a *área efetiva* $a_{[i]}$ como a proporção da célula i que se sobrepõe à região, para $i=1,...,168$.

Figura 6.2: Totais de casos mensais nos três anos do estudo.

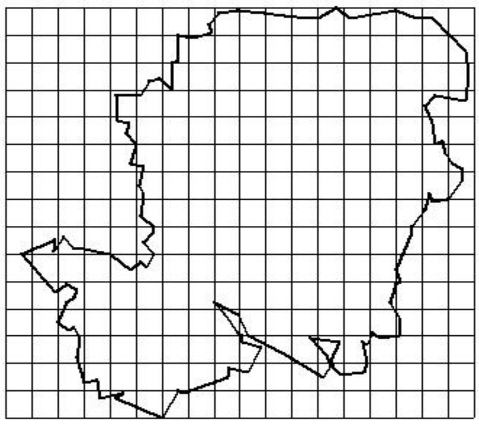


Figura 6.3: Grade regular com 270 células sobreposta à região de estudo.

A distribuição espacial de todos os casos de 2001 é usada na estimação da intensidade populacional $\lambda_{0[i]}$ em cada célula i, através de

$$\hat{\lambda}_{0[i]} = \frac{\sum_{t=1}^{12} y_{[i,t]}}{a_{[i]}} + \delta, \qquad i = 1, ..., 168, \quad t = 1, ..., 24,$$

onde $y_{[i,t]}$ é o número de casos na i-ésima célula da grade no t-ésimo mês e $\delta = 10^{-4}$ é uma correção necessária para que o modelo não atribua intensidade nula às células sem casos em 2001.

A intensidade do processo em cada célula i e mês t, $\lambda_{[i,t]}$, é modelada por

$$log\left[\lambda_{[i,t]}\right] = log\left[a_{[i]}\right] + log[\hat{\lambda}_{0[i]}] + \mu_{[t]} + \phi_{[i,t]}, \qquad i=1,...,168, \;\; t=1,...,24,$$

para a qual a tendência temporal $\mu_{[t]}$ é modelada por

$$\mu_{[t]} = \beta_0 + \beta_1 t, \qquad t=1,...,24.$$

Definindo $\phi_{[\cdot,t]} = (\phi_{[1,t]},...,\phi_{[N,t]})'$, a equação de evolução no tempo é dada por

$$\phi_{[\cdot,t]} = \eta\phi_{[\cdot,t-1]} + \omega_{[\cdot,t]}, \qquad \omega_{[\cdot,t]} \sim N\left[\underline{0}\,;(1-\eta^2)\sigma^2 R_\theta\right], \qquad t=2,...,24,$$

onde $0<\eta<1$, $\sigma^2>0$, $\theta>0$, $\underline{0}$ é um vetor de comprimento 168 com elementos iguais a zero e $R_\theta = [R_{i,j}]_{\{i,j=1,...,168\}}$, com $R_{i,j} = \exp\{\theta\,\|s_i - s_j\|\}$, é a matriz 168×168 de correlações espaciais entre as células, modeladas pela função de correlação exponencial. Assume-se a priori que $\phi_{[\cdot,1]} \sim N\,[\underline{0}\,;\sigma^2 R_\theta]$.

Foram escolhidas as prioris de referência do Capítulo 3 para $\boldsymbol{\beta} = (\beta_0,\beta_1)'$, σ^2 e θ, e $U[0;1]$ para η. No processo de amostragem via MCMC, foram geradas, para cada modelo, duas cadeias de tamanho 100 mil, definidas por diferentes valores iniciais das quantidades a serem estimadas. Como as duas cadeias convergiram para o mesmo ponto, a amostra final de cada parâmetro foi formada por 500 valores tomados a cada 100 no terceiro quarto das duas cadeias.

Os histogramas das amostras a posteriori dos hiperparâmetros são mostrados na Figura 6.4. O coeficiente linear do tempo, β_1, foi estimado pontualmente por -0,02, mostrando a tendência descrescente da intensidade dos casos. O parâmetro de correlação temporal entre os efeitos espaciais, η, mostrou-se de valor moderado, com média a posteriori igual a 0,55. O parâmetro θ, relacionado à correlação espacial na função exponencial, foi estimado pontualmente em 0,22, valor que significa uma correlação igual a 0,33 para os pares de áreas mais próximas entre si, ou seja, áreas adjacentes a norte, sul, leste ou oeste.

A Figura 6.5 mostra os mapas das médias a posteriori dos efeitos espaciais $\phi_{[i,t]}$ em cada mês. Estes efeitos parecem não ter uma estrutura espacial. De fato, o índice de autocorrelação espacial I de Moran (Bailey e Gatrell, 1995) foi significante (a 5%) em menos da metade dos meses. Isto sugere que a estimação da população sob risco usando os próprios casos incorporou toda a informação espacial da dispersão da doença, restando apenas um ruído branco no espaço. Enquanto não houver um modo mais eficaz de estimar a densidade espacial da população sob risco, sem usar os próprios casos da doença, não há como aplicar modelos mais elaborados para estes dados. Por exemplo, poderiam ser utilizados os casos de outra doença, não relacionada a infecções gastrointentinais, que sejam reportados pelo mesmo tipo de sistema telefônico, para estimar a distribuição espacial das chamadas telefônicas.

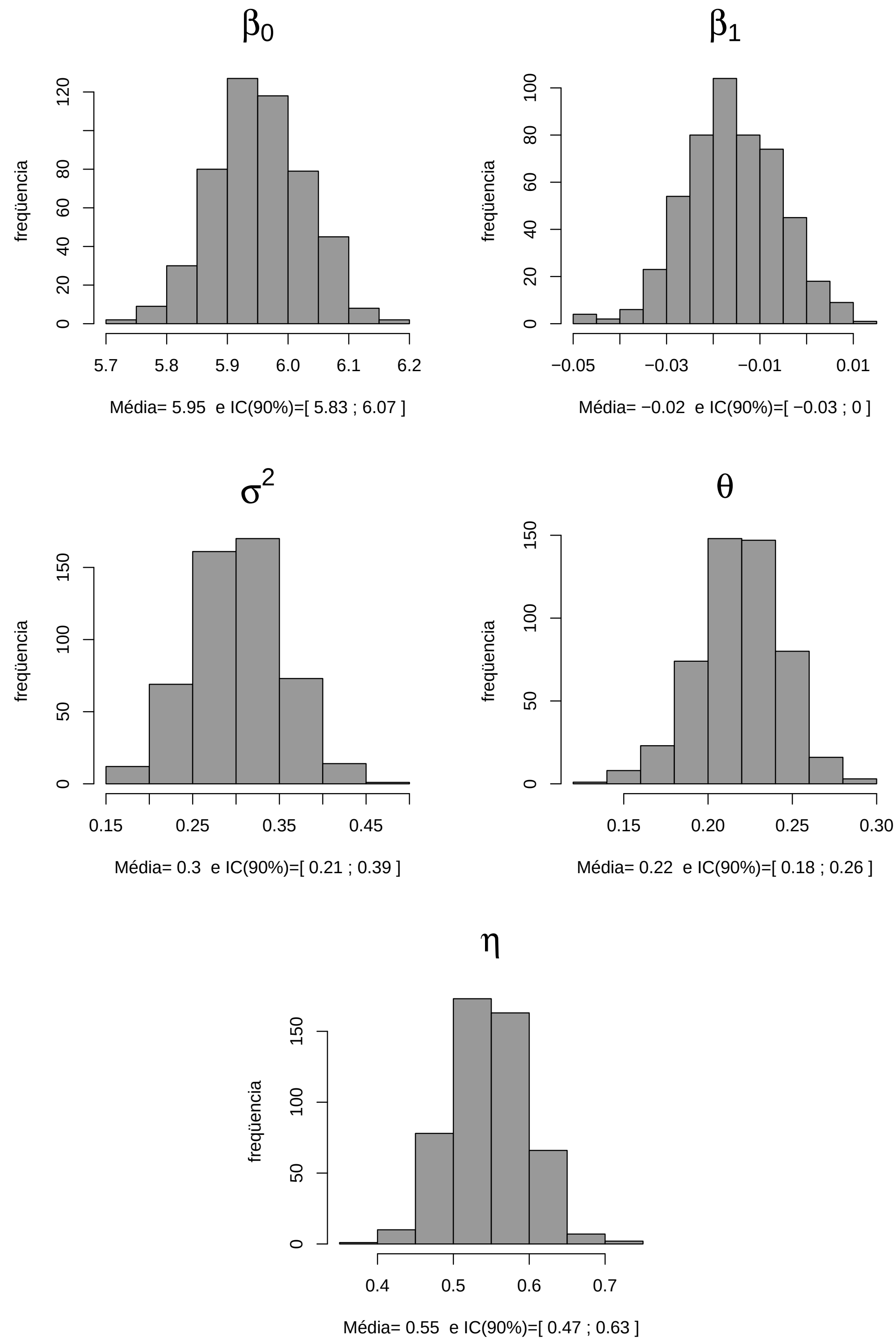


Figura 6.4: Histogramas das amostras a posteriori dos hiperparâmetros.

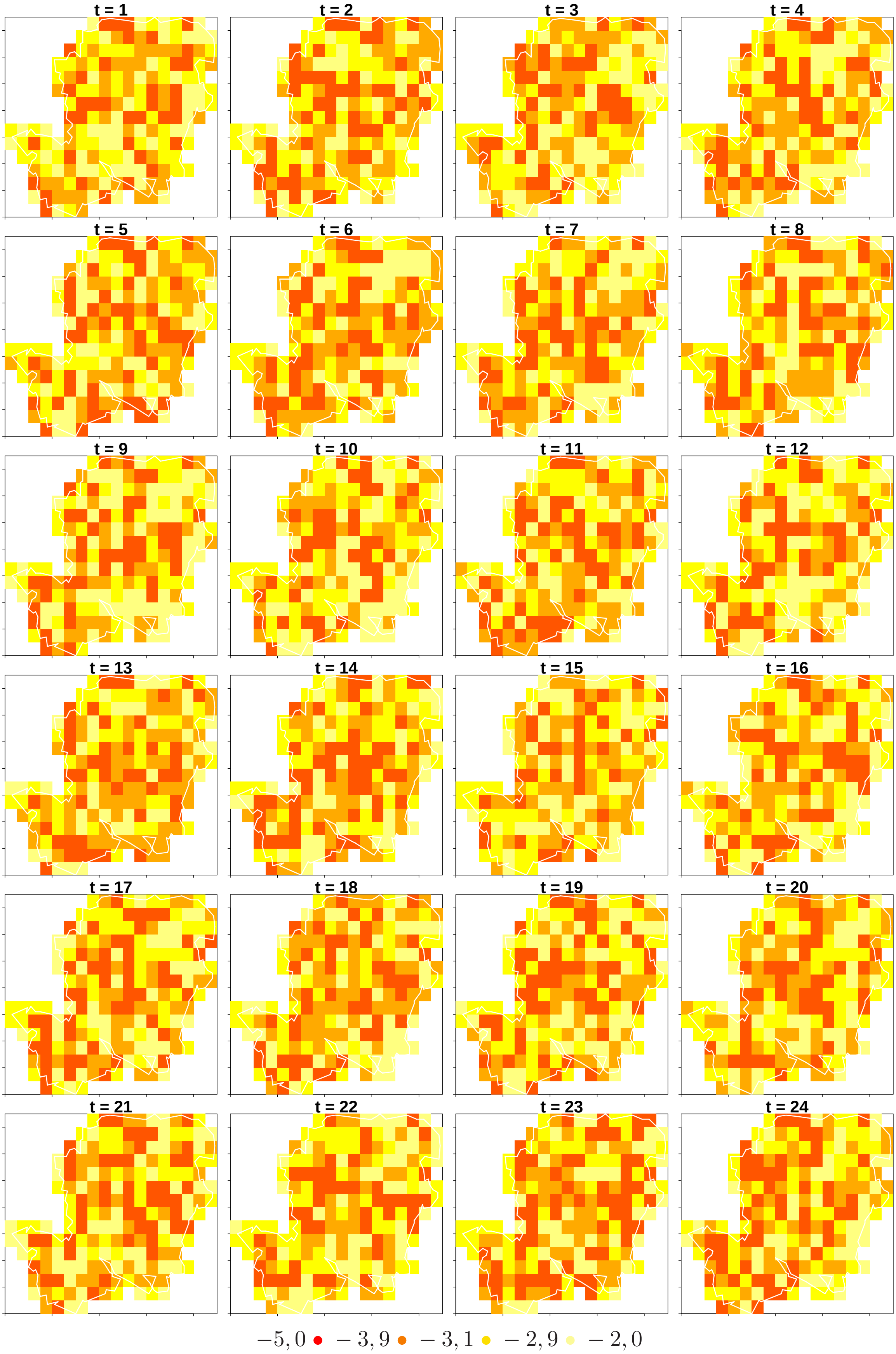


Figura 6.5: Mapas das médias a posteriori dos efeitos espaciais $\phi_{[i,t]}$ para os 24 meses de observação.

6.3 Evolução Espaço-Temporal de Impulsos Elétricos no Intestino Delgado

O intestino delgado finaliza o processo de digestão, absorve os nutrientes e conduz os resíduos para o intestino grosso. As células nervosas existentes na parede do intestino delgado emitem sinais que controlam os movimentos coordenados de contração de sua parede muscular, fazendo com que o conteúdo resultante da digestão seja empurrado ao longo do trato intestinal.

Dois padrões de atividade elétrica são importantes neste processo: as *ondas lentas* (*slow-waves*) e os *impulsos* (*spike potentials*). Uma onda lenta age como um sinal de marca-passo que induz o músculo à contração. Os impulsos superimpostos às ondas lentas determinam a força e duração da contração muscular.

Uma questão de interesse sobre este processo é saber se existem áreas com incidência mais alta de impulsos, comparadas com outras áreas. Outra questão é o entendimento das características temporais e espaciais da ocorrência de impulsos durante sucessivas ondas lentas. Especificamente, deseja-se saber se as áreas com atividades elétricas mais intensas são as mesmas ao longo das ondas lentas sucessivas. Desse modo, a modelagem da distribuição espaço-temporal dos impulsos pode ajudar no entendimento do mecanismo de geração e propagação dos movimentos intestinais.

No experimento descrito em Faes *et al.* (2006), um segmento do intestino delgado foi removido de sete gatos e suas atividades elétricas espontâneas foram observadas durante o período de um minuto, usando-se 240 eletrodos dispostos em uma grade regular $(10{\times}24)$ na superfície do tecido. A Figura 6.6 ilustra a atividade elétrica para um gato, medida pelo número de impulsos em cada célula da grade em 13 sucessivas ondas lentas. Este foi o único dos sete conjuntos de dados disponibilizado pelos autores do artigo original.

Este é um exemplo de um processo pontual original que foi observado, por razões de instrumento de medição, com os dados já discretizados na forma de contagens. Embora os impulsos possam ocorrer em qualquer ponto do tecido, sua medição através de um número limitado de eletrodos fez com que eles pudessem ser registrados apenas como contagens na área de percepção do eletrodo.

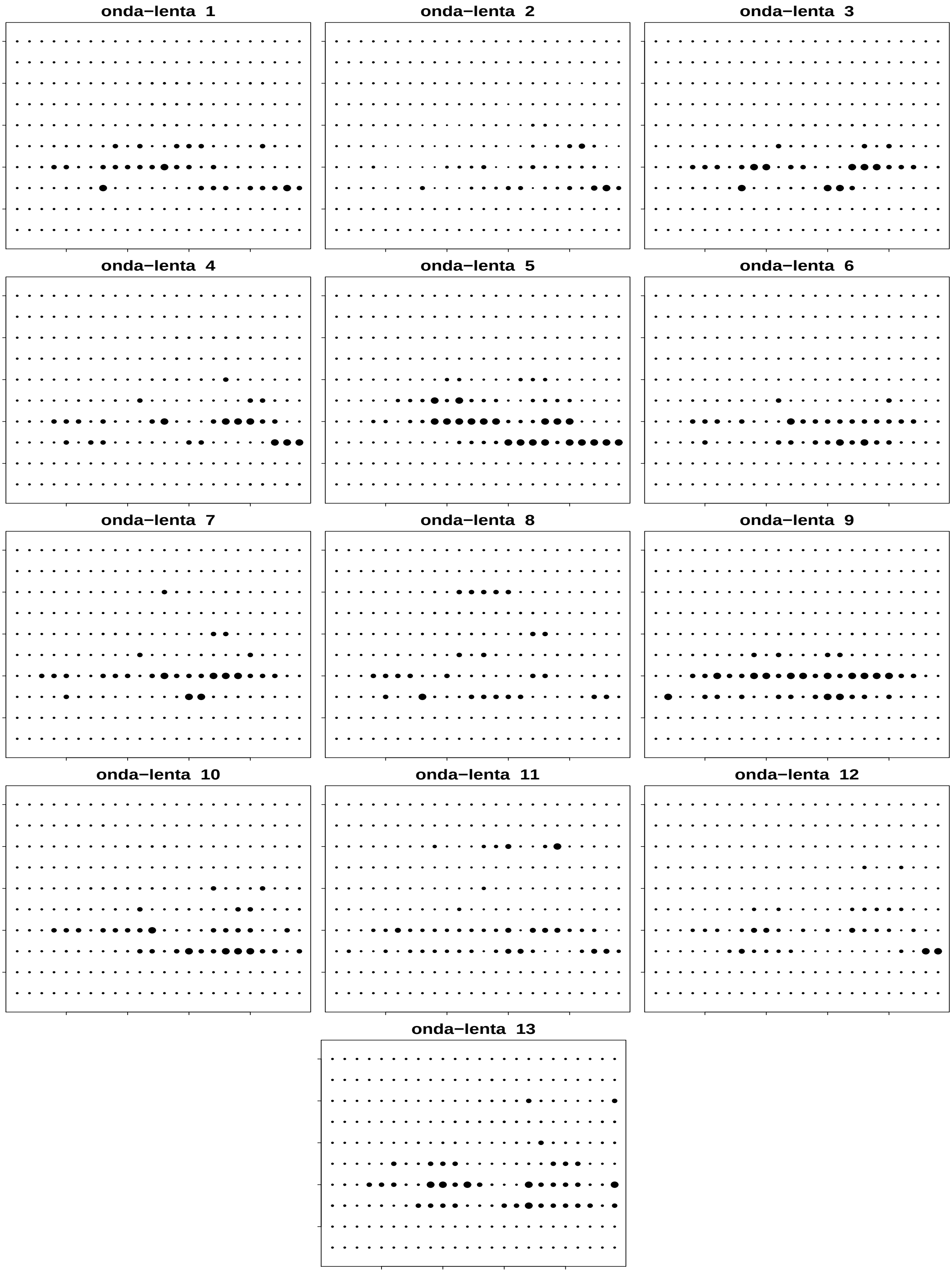


Figura 6.6: Número de impulsos na grade espacial com 10×24 células no intestino de um gato, durante 13 ondas-lentas sucessivas. A área do círculo é proporcional ao número de impulsos na célula, que varia de 0 a 5 impulsos.

Para tratar desses dados, foi adotado o procedimento descrito a seguir. Seja $y_{[i,t]}$ o número de impulsos ocorridos na *i*-ésima célula da grade espacial e na *t*-ésima onda lenta. Assume-se que o logaritmo da intensidade $\lambda_{[i,t]}$ do processo é modelada por

$$log[\,\lambda_{[i,t]}\,] \;=\; \mu_{[t]} \;+\; \phi_{[i,t]}, \qquad i=1,...,240, \;\; t=1,...,13.$$

Três modelos foram ajustados à tendência temporal μ_t:

Modelo 1 (Constante) : $\mu_{[t]} = \mu$;

Modelo 2 (Linear): $\mu_{[t]} = \beta_0 + \beta_1 \cdot t, \quad t=1,...,13$;

Modelo 3 (Dinâmica de 1^a Ordem) : $\mu_{[t]} = \mu_{[t-1]} + \upsilon_t, \quad \upsilon_t \sim N[0;\omega^2]$, t=1,...,13.

Para os efeitos espaço-temporais $\phi_{[\cdot,t]} = \left(\phi_{[1,t]},...,\phi_{[240,t]}\right)'$ foi escolhido o modelo com processos gaussianos autoregressivos estacionários no tempo:

$$\phi_{[\cdot,t]} \;=\; \eta\,\phi_{[\cdot,t-1]} \;+\; \omega_{[\cdot,t]}, \qquad \text{com} \qquad \omega_{[\cdot,t]} \sim N\left[\,\underline{0}\,;\,(1-\eta^2)\sigma^2 R_\theta\,\right], \quad t=2,...,13,$$

$$\phi_{[\cdot,1]} \;\sim\; N\left[\,\underline{0}\,;\,\sigma^2 R_\theta\,\right],$$

com $0<\eta<1$ e elementos da matriz de correlações espaciais R_θ definidos pela função de correlação exponencial.

Foram escolhidas as prioris de referência do Capítulo 3 para μ, $\boldsymbol{\beta}=(\beta_0,\beta_1)'$, σ^2, θ e ω^2 e $U[0;1]$ para η. No processo de amostragem via MCMC, foram geradas, para cada modelo, duas cadeias de tamanho 100 mil, definidas por diferentes valores iniciais das quantidades a serem estimadas. Como as duas cadeias convergiram para o mesmo ponto, a amostra final de cada parâmetro foi formada de 1000 valores tomados a cada 100 na segunda metade das duas cadeias.

A Tabela 6.1 mostra a média e o intervalo de credibilidade de 90% das amostras a posteriori dos parâmetros β_0, β_1, μ, σ^2, θ, η e ω^2 relativos a cada um dos três modelos (e outros dois modelos definidos seguir). Os histogramas destas amostras a posteriori (figuras 6.7 a 6.9) mostram distribuições a posteriori unimodais bem comportadas em todos os modelos. A medida de correlação temporal dos efeitos espaciais entre duas ondas lentas sucessivas é dada pelo parâmetro η, estimado pontualmente por 0,8, 0,7 e 0,75, respectivamente para os três modelos. O parâmetro σ^2 foi estimado pontualmente em 0,6 para os três modelos. Estes resultados semelhantes eram esperados, pois σ^2 mede a variabilidade entre os efeitos espaciais em cada tempo t. Do mesmo modo, para θ, parâmetro relacionado à correlação (puramente) espacial na função exponencial, não se esperava resultados diferentes entre os modelos. De fato, ele foi estimado pontualmente em 0,15, valor que significa uma correlação igual a 0,86 entre os pares de áreas mais próximas entre si, ou seja, áreas adjacentes (distância entre centróides igual a 1 unidade) e uma correlação igual a 0,02 entre os pares de áreas mais distantes entre si, ou seja, áreas localizadas nos vértices opostos nas diagonais da região de estudo (distância entre centróides igual a 26 unidades).

A intensidade dos impulsos ao longo das sucessivas ondas lentas é caracterizada pelos efeitos temporais μ_t (Figura 6.12, para os modelos 2, 3 e outros dois modelos definidos seguir). No primeiro modelo, μ foi estimado por -0,3. No segundo modelo, a estimativa de μ_t é uma combinação das estimativas de β_0 e β_1, e gerou um tendência temporal linear decrescente, com nível médio um pouco menor que a estimativa do modelo 1. O histograma da amostra a posteriori do coeficiente linear β_1 mostra que o valor zero tem alta densidade, o que significa que este efeito do tempo em μ_t não parece ser significativo. Para o modelo 3, as estimativas de μ_t se mostraram aproximadamente constantes ao longo das ondas lentas.

A Figura 6.13 mostra os envelopes de estimação (média amostral e intervalos de credibilidade de 90%) dos efeitos ϕ_{it}, agrupados nas 13 ondas lentas (para os modelos 1, 2, 3 e outros dois modelos definidos seguir). A forma destes envelopes não se mostrou como esperado nos estudos simulados, nos quais a amplitude dos intervalos aumentou com a média a posteriori dos efeitos. Nesta aplicação, no entanto, há uma inexplicada inversão desta relação a partir de certo valor da média a posteriori.

A Tabela 6.2 mostra os resultados dos critérios de seleção de modelos *DIC* e *EPD* para estes três modelos (e outros dois modelos que serão apresentados a seguir). Embora os valores sejam muito parecidos, o modelo com tendência temporal dinâmica de primeira ordem obteve os menores valores, sendo, portanto, o escolhido dentre este três modelos segundo este critérios.

Entretanto, antes de escolher o modelo 3 como o mais adequado para este conjunto de dados (dentre as alternativas testadas), decidiu-se verificar a necessidade de se usar uma estrutura autorregresiva nos efeitos espaço-temporais $\phi_{[i,t]}$.

Desse modo, considerando o modelo dinâmico de primeira ordem para a tendência temporal $\mu_{[t]}$ (modelo 3), foram ajustados dois modelos mais simples para os efeitos espaço-temporais $\phi_{[i,t]}$. Um destes modelos assume que estes efeitos são puramente espaciais, ou seja, $\phi_{[i,t]} = \zeta_{[i]}, \forall t,$

Modelo 3b (Puramente espaciais): $\zeta_{[\cdot]} = \left(\zeta_{[1]}, ..., \zeta_{[240]}\right)' \sim N\left[\underline{0}\,; \sigma^2 R_\theta\right].$

O outro modelo ajustado é um caso particular do modelo 3, tomando-se o parâmetro de correlação temporal η igual a zero. Ou seja, este modelo assume que os efeitos $\phi_{[i,t]}$ são independentes no tempo:

Modelo 3c (Efeitos livres): $\phi_{[\cdot,t]} \sim N\left[\underline{0}\,; \sigma^2 R_\theta\right]$, independentes para $t = 1, ..., 13.$

Os resultados da estimação dos hiperparâmetros (figuras 6.10 e 6.11), da tendência temporal (Figura 6.12) e dos efeitos espaciais (Figura 6.13) são bastante semelhantes aos resultados dos demais modelos.

Os valores dos *DIC* e *EPD* destes dois modelos (Tabela 6.2), se comparados aos valores obtidos do modelo 3, levam à conclusão de que este modelo é mais adequado a este conjunto de dados. Além disso, deve-se notar que os resultados de estimação do parâmetro de correlação temporal η dos efeitos espaciais $\boldsymbol{\phi}$ no modelo 3 descartam o modelo 3c (que assume $\eta = 0$), pois η é estimado com valores distantes de zero. Desse modo, os modelos propostos nesta tese, com efeitos espaciais específicos em cada tempo e com estrutura autoregressiva, levam a um melhor ajuste.

Entretanto, o padrão espacial dos efeitos espaço-temporais não parece se modificar significativamente entre os modelos 1, 2, 3 e 3c, como pode ser visto nas figuras 6.14 e 6.15. Assim, pode-se concluir que os diferentes resultados de estimação da tendência temporal não afetaram as estimativas dos efeitos espaço-temporais. A Figura 6.16 mostra que, para o modelo escolhido, a variabilidade dos efeitos $\boldsymbol{\phi}$ também têm estrutura espacial.

Somando a estes efeitos espaço-temporais a tendência temporal, são obtidas as estimativas do logaritmo das intensidades $\lambda_{[i,t]}$, ou seja, do número esperado de impulsos nas células da grade em cada onda lenta.

Faes *et al.* (2006) não são conclusivos sobre a tendência temporal da intensidade dos impulsos, talvez por terem chegado à mesma conclusão que este estudo de que não há efeito aparente do tempo na seqüencia analisada de onda lentas. Assim como em Faes *et al.* (2006), a inspeção visual dos mapas dos efeitos espaço-temporais leva à conclusão de que os impulsos elétricos claramente tendem a ocorrer em algumas áreas e não em outras, o que significa que as contrações musculares não estão distribuídas de maneira homogênea na parede intestinal. Além disso, as áreas com maior ocorrência de impulsos são as mesmas ao longo das ondas-lentas, ou seja, guardam uma dependência temporal. Segundo os fisiologistas co-autores do artigo original, esta constatação tem importantes implicações no entendimento da motilidade intestinal nos mamíferos.

Tabela 6.1: Médias a posteriori e Intervalo de Credibilidade de 90% para os hiperparâmetros.

	Modelos				
	1	2	3	3b	3c
μ	-0,21 [-1,58;1,21]				
β_0		-0,46 [-0,77;-0,13]			
β_1		-0.01 [-0,03;0,01]			
σ^2	0,57 [0,45;0,70]	0,52 [0,40;0,64]	0,62 [0,50;0,75]	0,62 [0,48;0,76]	0,63 [0,47;0,79]
θ	0,14 [0,10;0,19]	0,14 [0,10;0,19]	0,14 [0,09;0,20]	0,14 [0,09;0,20	0,16 [0,10;0,23]
η	0,79 [0,71;0,87]	0,69 [0,62;0,76]	0,74 [0,67;0,82]		
ω^2			0,10 [0,07;0,14]	0,10 [0,07;0,13]	0,10 [0,06;0,13]

Tabela 6.2: Resultados dos critérios de seleção de modelos.

Modelo	1	2	3	3b	3c
DIC	7580	7542	7510	7691	7779
EPD	5015	5011	5008	5082	5091

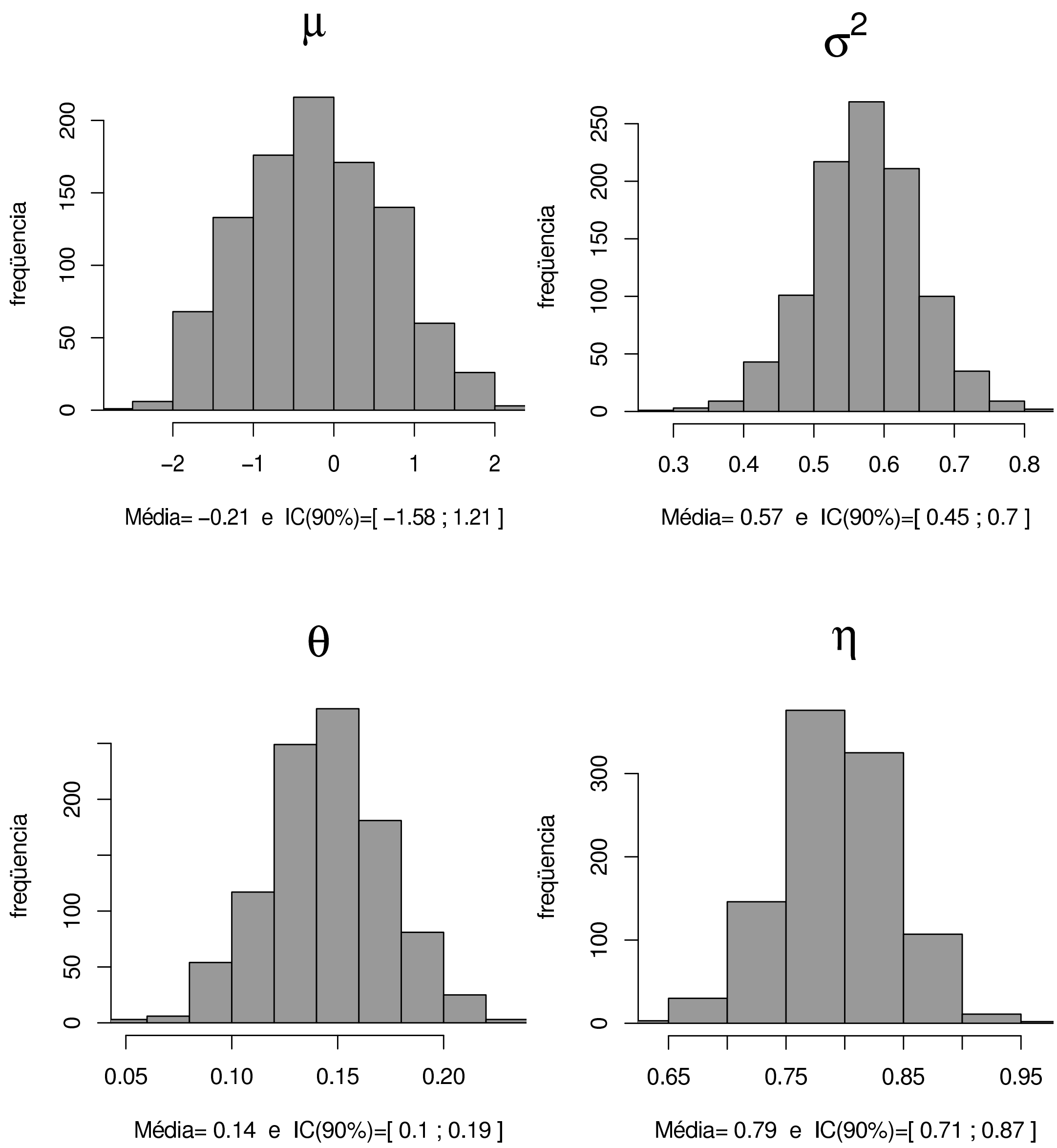


Figura 6.7: Histogramas das amostras a posteriori dos hiperparâmetros do modelo 1.

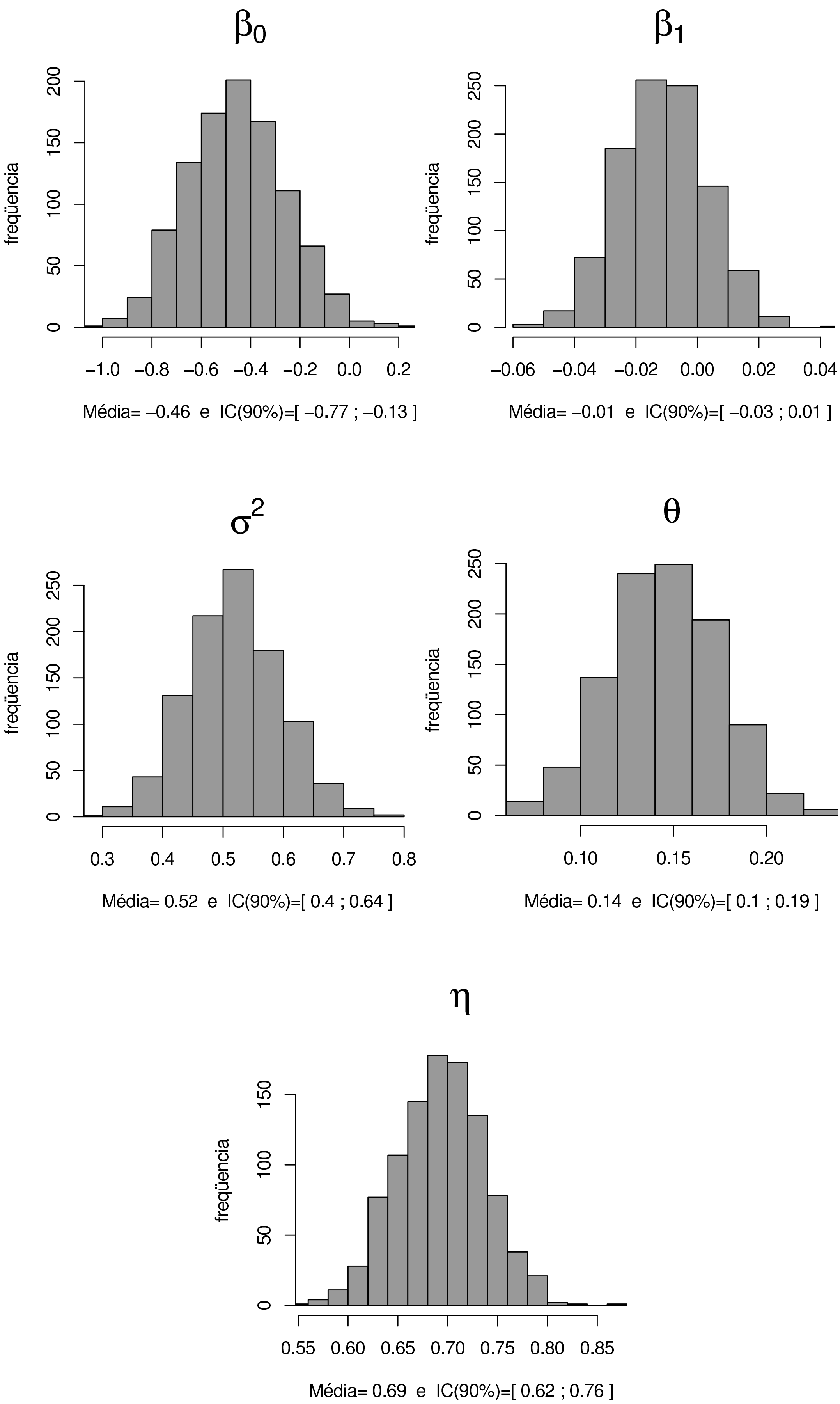


Figura 6.8: Histogramas das amostras a posteriori dos hiperparâmetros do modelo 2.

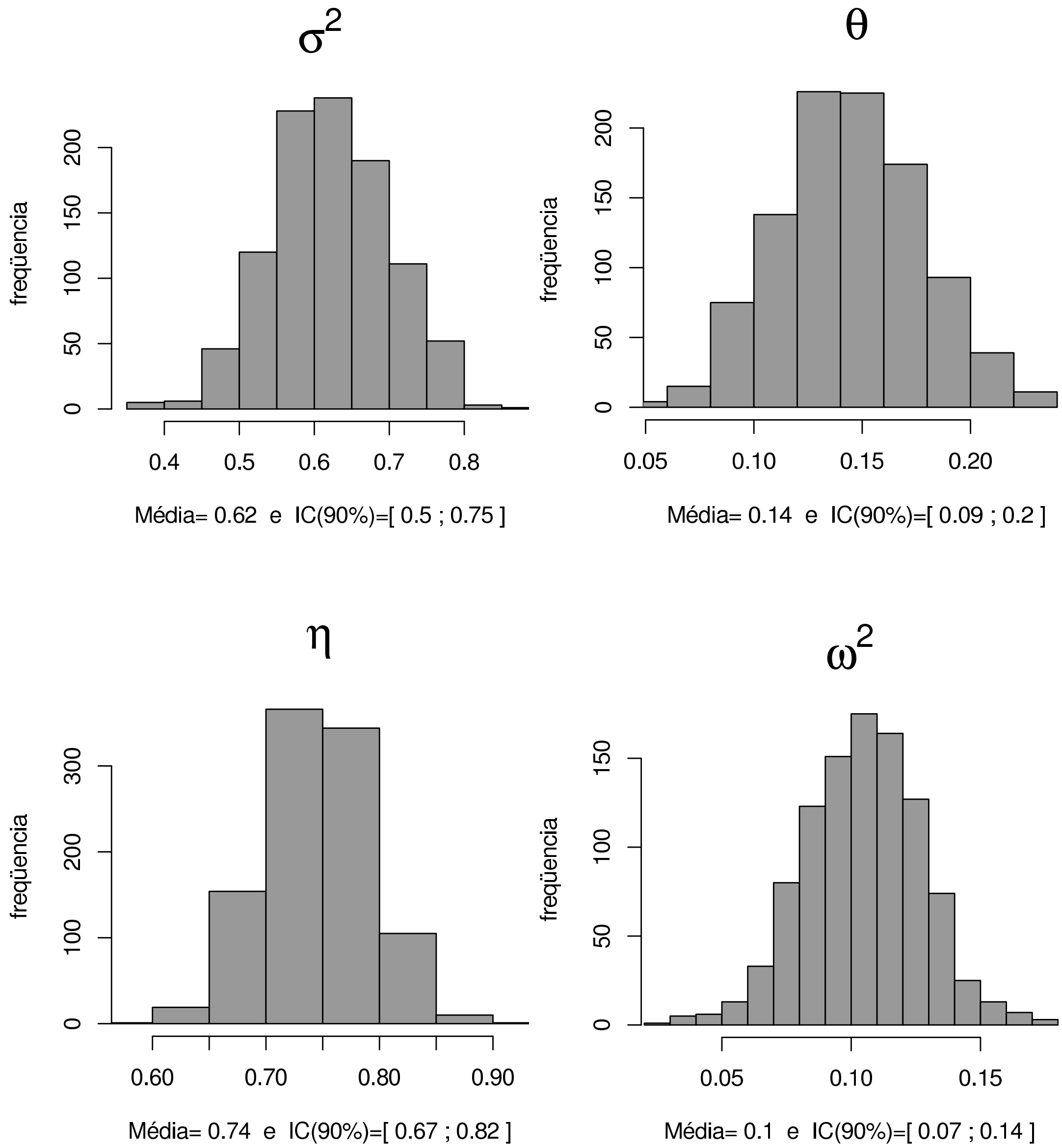


Figura 6.9: Histogramas das amostras a posteriori dos hiperparâmetros do modelo 3.

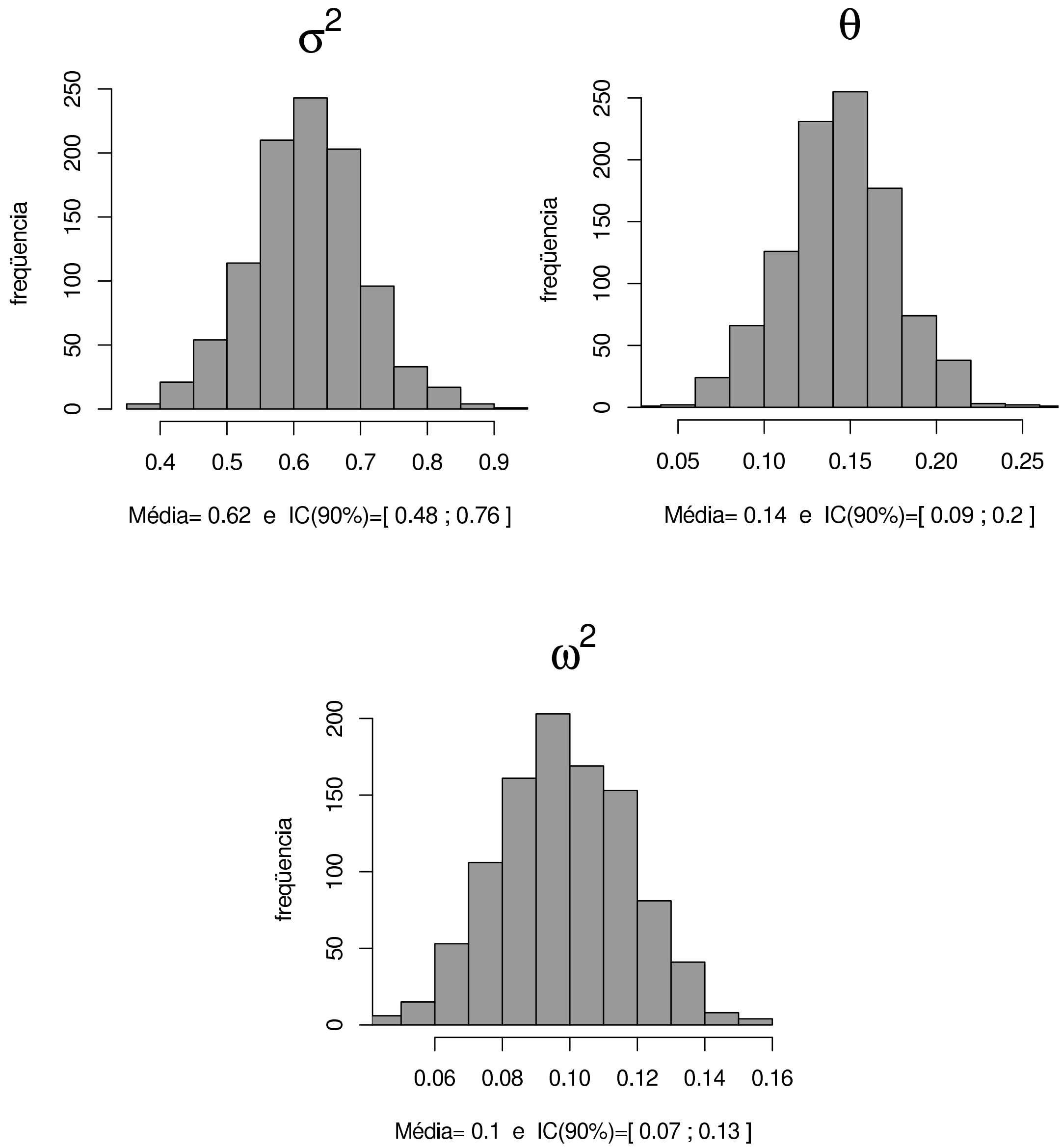


Figura 6.10: Histogramas das amostras a posteriori dos hiperparâmetros do modelo 3b.

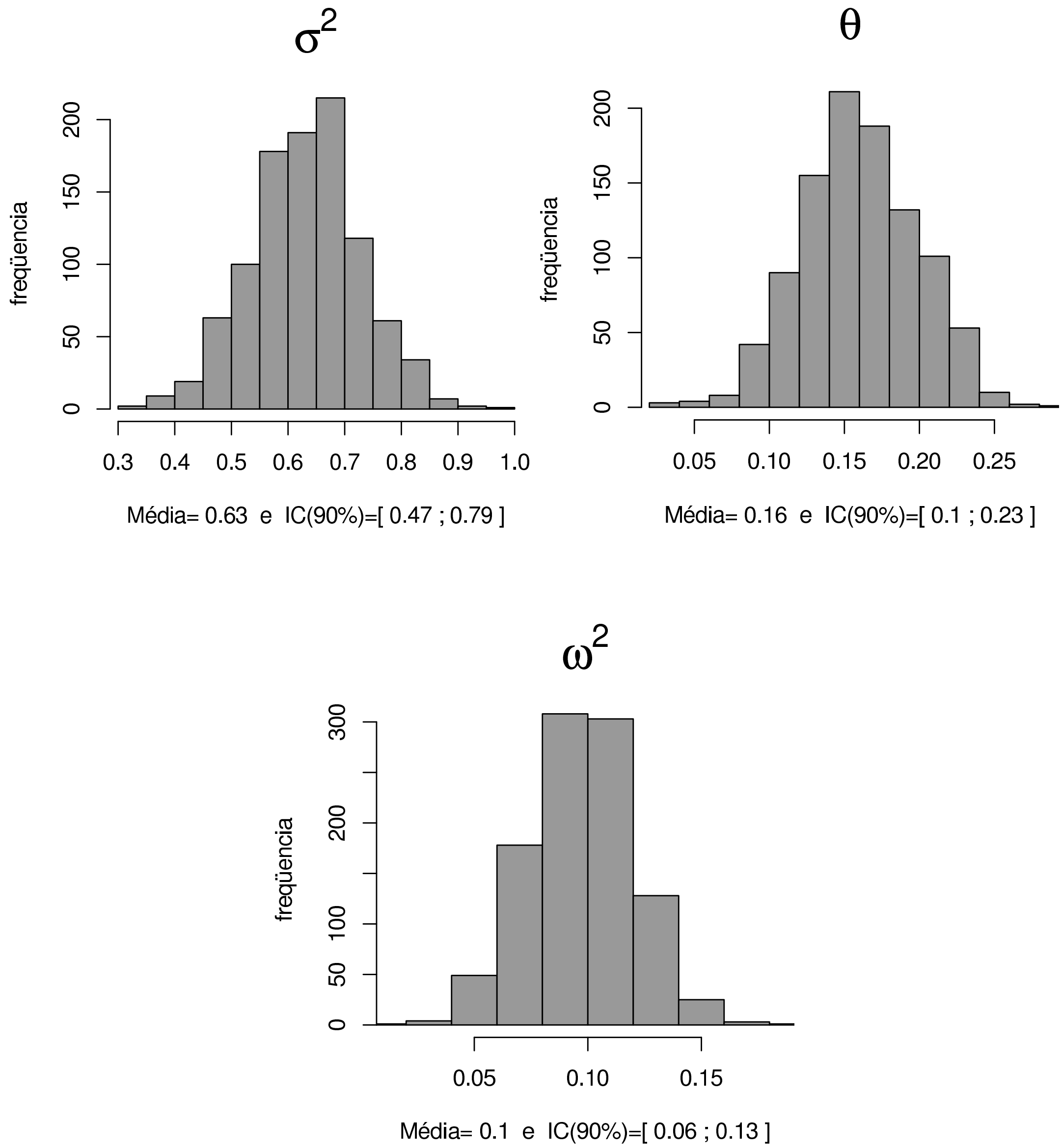


Figura 6.11: Histogramas das amostras a posteriori dos hiperparâmetros do modelo 3c.

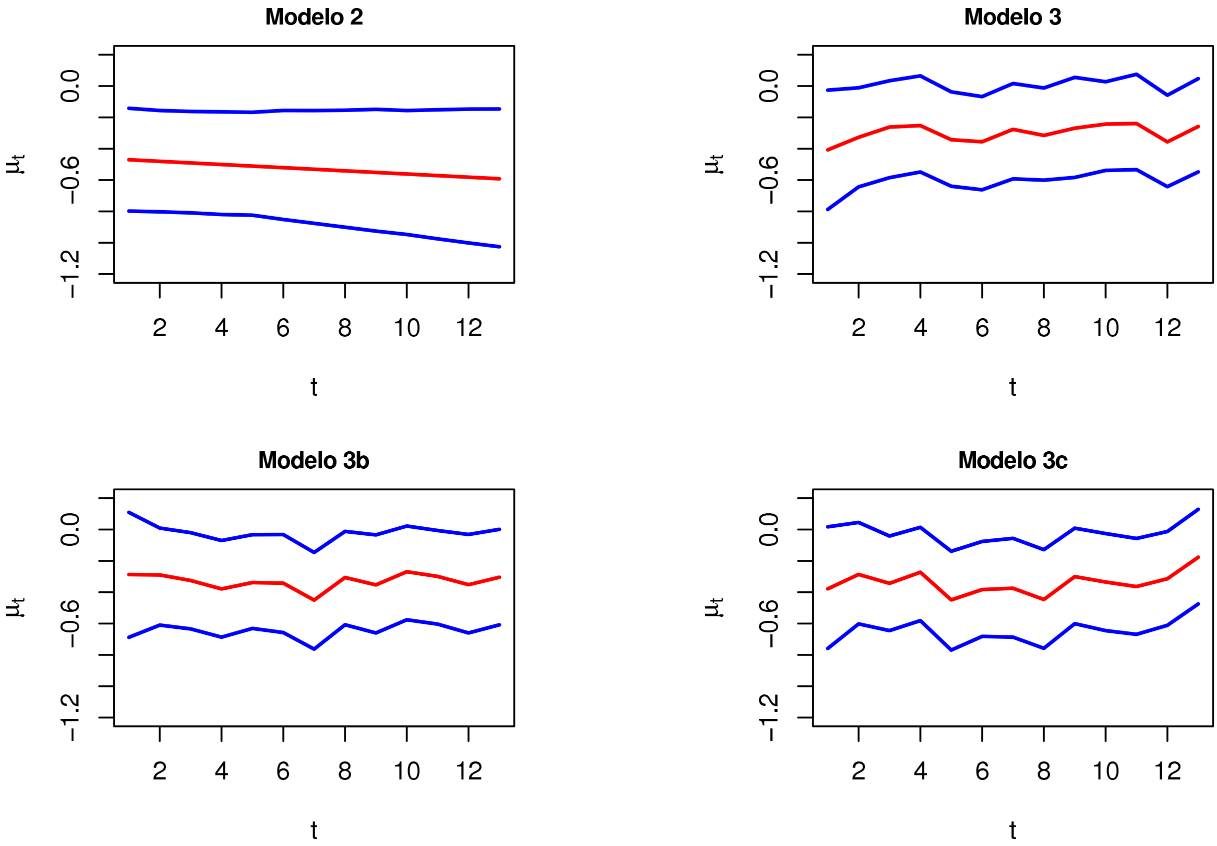


Figura 6.12: Médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul) dos efeitos temporais $\mu_{[t]}$.

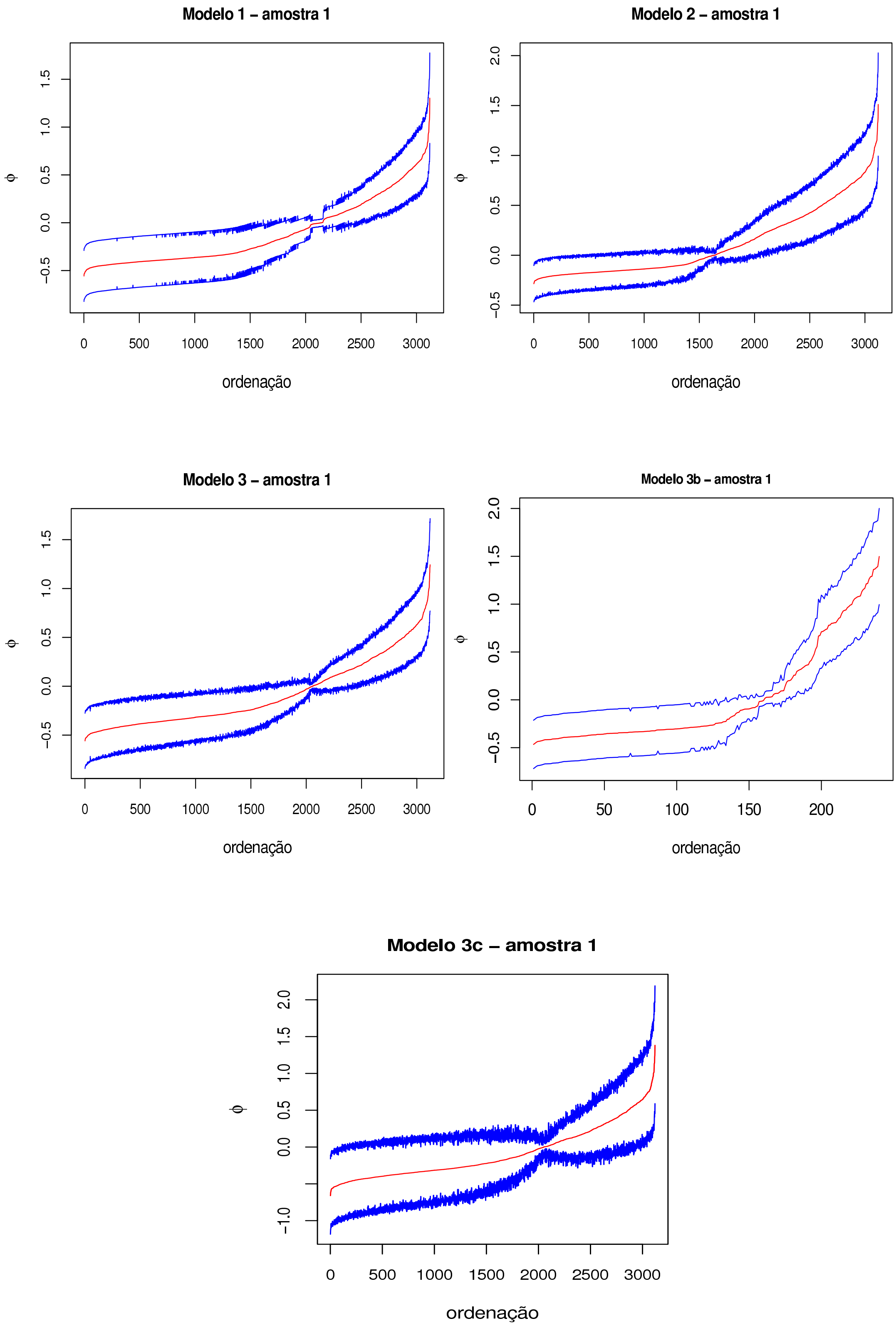


Figura 6.13: Médias a posteriori (em vermelho) e intervalos de 90% de credibilidade (em azul) dos efeitos espaço-temporais $\phi_{[i,t]}$. Os valores no eixo horizontal estão ordenados pela magnitude da média a posteriori.

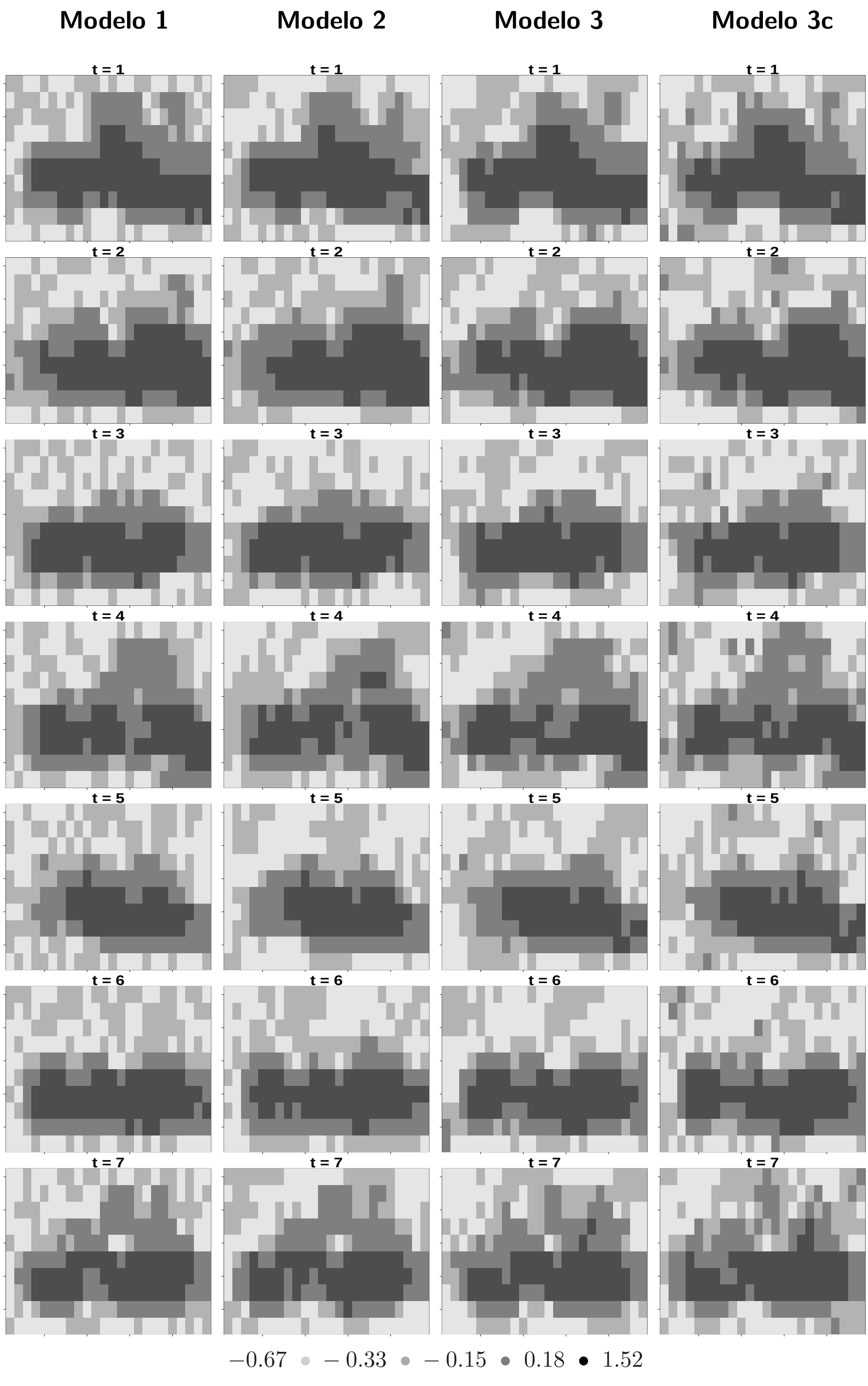


Figura 6.14: Mapas das médias a posteriori dos efeitos espaço-temporais $\phi_{[i,t]}$ dos modelos 1, 2, 3 e 3c, para $t = 1, ..., 7$.

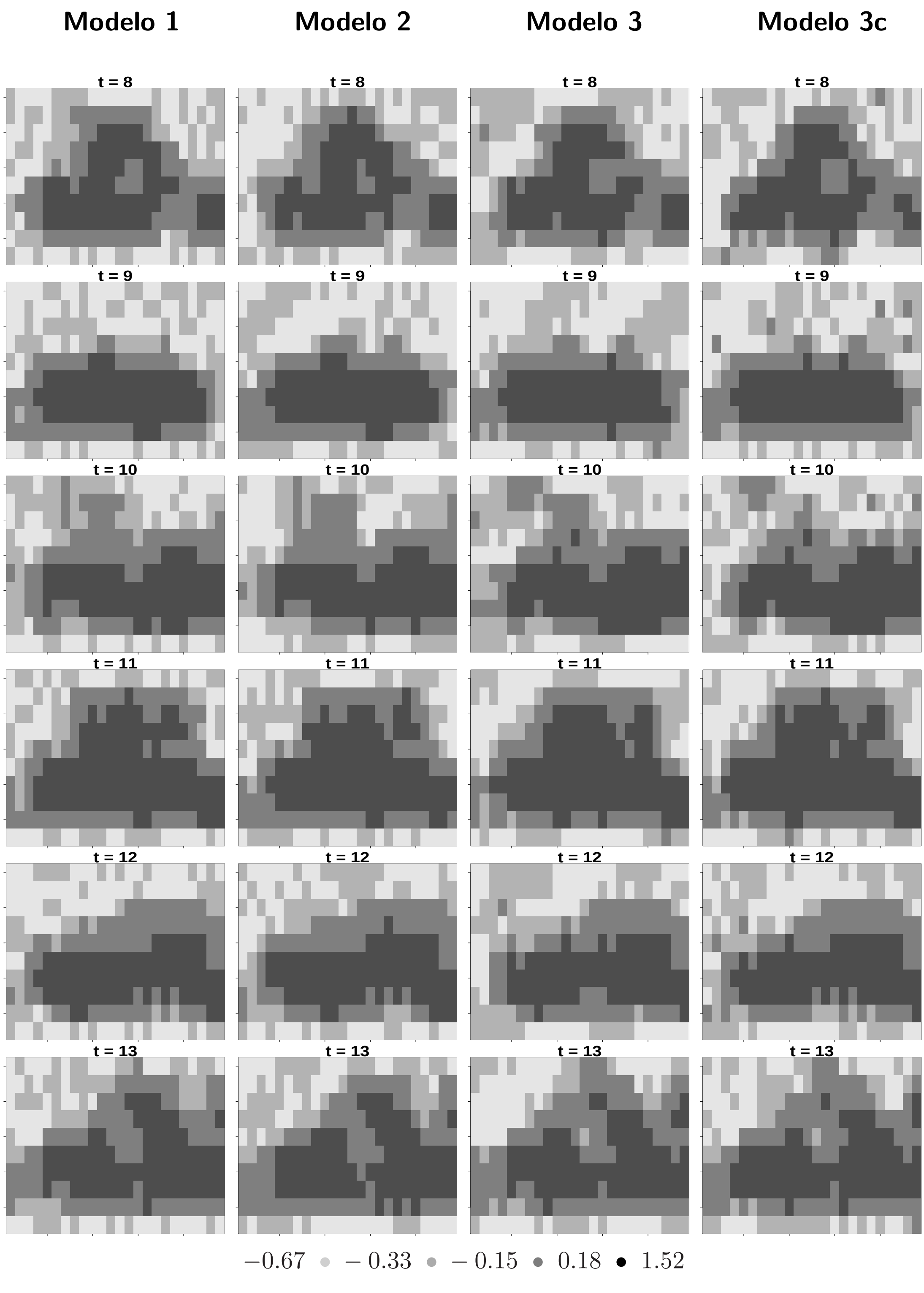


Figura 6.15: Mapas das médias a posteriori dos efeitos espaço-temporais $\phi_{[i,t]}$ dos modelos 1, 2, 3 e 3c, para $t = 8, ..., 13$.

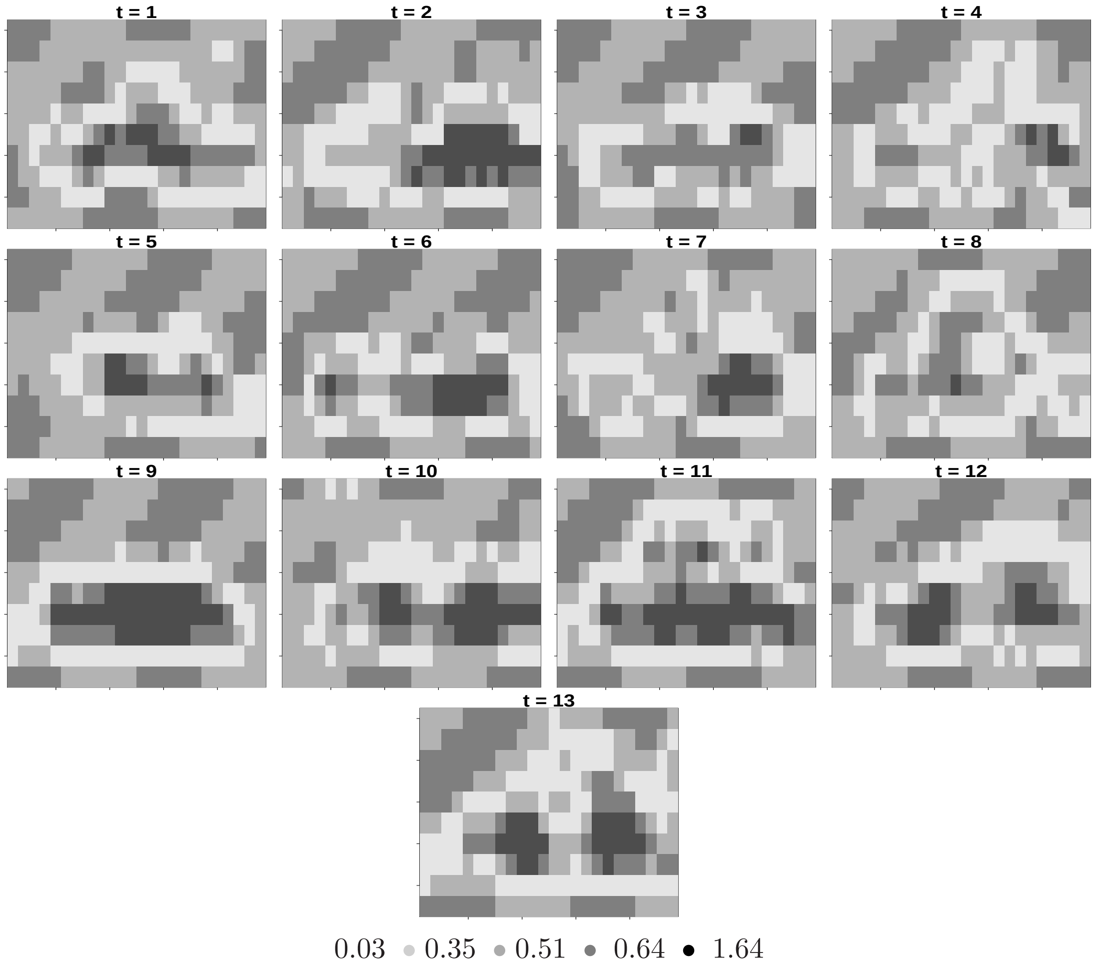


Figura 6.16: Mapas de variabilidade dos efeitos espaço-temporais $\phi_{[i,t]}$, $t=1,...,13$, do modelo 3.

Capítulo 7

Considerações Finais e Trabalhos Futuros

7.1 Considerações Finais

Nesta tese, foram propostos modelos espaço-temporais para a intensidade de processos pontuais especificados por uma seqüencia de superfícies de intensidades espaciais ligadas no tempo através de uma evolução dinâmica. A tendência temporal pode ser modelada livremente; por exemplo, pode ser assumida constante, ou ser escrita como uma função determinística de covariáveis medidas em cada intervalo de tempo, ou ainda, pode ser descrita por um modelo dinâmico, dentre outras possibilidades. A inferência é feita sob a abordagem bayesiana completa, através de métodos de simulação MCMC, como os amostradores de Gibbs e Metropolis-Hastings.

Os resultados de estudos simulados mostram que os modelos e métodos de estimação propostos são adequados para modelar conjuntos de dados gerados por processos pontuais espaço-temporais, pois as estimativas tomadas das amostras a posteriori tiveram uma boa concordância com os valores reais. Aplicamos os modelos aos dados de infecção gastrointenstinal em Hampshire e aos dados do estudo da evolução espacial e temporal de impulsos elétricos no intestino delgado de gatos.

Considerou-se a situação na qual toda estrutura de dependência espacial é devida apenas à heterogeneidade espacial na intensidade do processo, e não à interação direta entre os eventos. Entretanto, os modelos propostos também podem ser úteis na análise descritiva da distribuição espacial de eventos gerados de processos com interação espacial direta entre os eventos. De fato, Schoenberg (2005) conclui que, mesmo que o verdadeiro processo pontual espaço-temporal sendo estimado não seja Poisson, um estimador baseado na maximização da função de verossimilhança do processo de Poisson é consistente sob certas condições simples.

7.2 Trabalhos Futuros

Os resultados satisfatórios obtidos até o momento estimulam a extensão deste trabalho em diversas direções. Uma destas extensões é a utilização da discretização espacial através de Tesselagem de Voronoi, descrita na Seção 2.6, para aplicações nas quais a agregação dos eventos é bastante acentuada.

Pretende-se ampliar o conjunto de estudos simulados do Capítulo 5 com o acréscimo de mais casos particulares do modelo geral. Para analisar a capacidade de estimação dos modelos, seria interessante se fazer estudos nos quais o conjunto de dados é gerado de um modelo mais complexo e se ajusta um modelo mais simples, e vice-versa.

Outros trabalhos futuros de interesse são descritos a seguir.

7.2.1 Eficiência Computacional do Processo de Inferência

O esquema de amostragem individual dos efeitos espaço-temporais $\phi_{[i,t]}$ adotado neste trabalho demanda grande tempo computacional. Isto porque na geração da cadeia de cada $\phi_{[i,t]}$, os parâmetros de sua distribuição da proposta mudam sempre que as matrizes H_i e B_i são recalculadas, ou seja, sempre que θ (e/ou γ) são atualizados. O problema é que a inversão das N matrizes H_i de dimensão $N-1$, é muito demorada.

Uma solução é aproximar a matriz de correlações do processo gaussiano por uma matriz banda diagonal, ou seja, uma matriz com valores nulos exceto por aqueles localizados na diagonal principal e suas adjacências (Rue e Tjelmeland, 2002). A característica altamente esparsa da matriz banda diagonal permite o uso de algoritmos de inversão mais rápidos. Esta construção envolve dois passos: a permutação da matriz de correlações de modo que os elementos de maior valor fiquem na diagonal principal e em suas adjacências (Rue, 2001); a aproximação desta matriz permutada por uma matriz banda diagonal.

Knorr-Held e Rue (2002) propõem algoritmos para amostragem em blocos em modelos hieráquicos com *campos aleatórios markovianos gaussianos* (CAMG´s) com o objetivo de aumentar a eficiência do MCMC. Os resultados do artigo indicam que os maiores benefícios são obtidos quando cada parâmetro e seus respectivos hiperparâmetros são atualizados conjuntamente em um único bloco, juntamente com o uso dos métodos de amostragem rápida de CAMG´s de Rue (2001).

Outra solução para redução do tempo de processamento computacional é dispensar a demorada geração de longas cadeias no MCMC e buscar uma *aproximação analítica* dos momentos das densidades marginais a posteriori dos hiperparâmetros e dos componentes do processo latente, como proposto em Rue e Martino (2006) para modelos hierárquicos de CAMG´s. Atráves de exemplos, estes autores mostram que o custo computacional destes esquemas determinísticos são muito baixos se comparados à alternativa via MCMC, especialmente se há um número pequeno de hiperparâmetros. Os autores argumentam ainda que estes resultados podem ser aplicados para modelos hierárquicos de CAG´s, como o processo de Cox log-gaussiano, através da aproximação da matriz de covariâncias por uma matriz banda diagonal, como descrito anteriormente.

7.2.2 Análise de Resíduos

Assim como nos modelos de regressão usuais, nos modelos para processos pontuais o uso eficaz da análise de resíduos torna possível encontrar características dos dados que não foram capturadas pelo modelo.

Baddeley *et al.* (2005) apresentam uma análise dos resíduos para processos pontuais espaços-temporais em modelos ETAS (*Epidemic Type Aftershock-Sequences*), comumente utilizados para descrever ocorrências de terremotos. Nestes modelos, assume-se uma intensidade não homogênea para o processo e uma dependência direta entre os eventos. No caso de terremotos, esta dependência ocorre porque um abalo inicial provoca outros abalos em sua volta em um curto intervalo de tempo, gerando um arranjo pontual do tipo agregado. Os autores definem resíduos chamados de primeira e segunda ordens, com a função específica de auxiliar a investigação da direção apropriada na qual o modelo pode ser melhorado.

Nos modelos propostos nesta tese, a análise de resíduos poderia ser desenvolvida para verificar a adequação dos modelos para descrever as propriedades de primeira ordem, dado que os modelos assumem que não há efeito de segunda ordem.

Referências

[1] Abramowitz, M. and Stegun, I.A. (1972) *Handbook of Mathematical Functions*. Dover, New York.

[2] Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In *Proceedings of the Second International Symposium on Information Theory*. B. N. Petrov and F. Csáki (eds). Budapest: Akadémiai Kiadó, pp. 267–281.

[3] Anderson, B.O.O. and Moore, V.B. (1979) *Optimal Filtering*. Prentice-Hall, Englewood Cliffs.

[4] Assunção, J.J., Gamerman, D. and Assunção, R. (1999) Regional differences in factor productivities of Brazilian agriculture: a Bayesian spatial varying parameter approach. In *Proceedings of the XVII Latin American Meeting of the Econometric Society*, Cancun.

[5] Baddeley A., Gregori, P., Mateu, J., Stoica and R. Stoyan, D. (eds.) (2006) *Case Studies in Spatial Point Process Modelling*. New York: Springer.

[6] Baddeley, A.J. and Turner (2005) R. Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, **12**, 1–42.

[7] Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005) Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society Series B*, **67**, 617–666.

[8] Bailey, T.C. and Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Essex: Longman Scientific & Technical.

[9] Beněs, V., Bodlak, K., Møller, J. and Waagepetersen, R. (2002). *Bayesian analysis of log Gaussian processes for disease mapping*. Research Report **3**, Centre for Mathematical Physics and Statistics, University of Aarhus.

[10] Berger, J.O. (1985) *Statistical Decision Theory and Bayesian Analysis*, 2.ed. New York: Springer-Verlag.

[11] Berger, De Oliveira and Sansó (2001) Objective Bayesian Analysis of Spatially Correlated Data. *Journal of the American Statistical Association*, **96**, 1361–1374.

[12] Bernardo, J.M. and Smith, A.F.M. (1994) *Bayesian Theory*. New York: John Wiley.

[13] Box, G. E. P. (1976) Science and statistics. *Journal of the American Statistical Association*, **71**, 791–799.

[14] Brix, A. and Diggle, P. J. (2001). Spatiotemporal prediction for log-Gaussian Cox processes. *Journal of the Royal Statistical Society Series B*, **63**, 823–841.

[15] Brix, A. and Møller, J.(2001). Space-Time Multi Type Log Gaussian Cox Processes with a View to Modelling Weeds. *Scandinavian Journal of Statistics*, **28**, 471–488.

[16] Brooks, S. P. and Gelman, A. (1998). Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, **7**, 434–455.

[17] Carter, C. K. and Kohn, R. (1994) On Gibbs sampling for state space models. *Biometrika*, **81**, 541–553.

[18] Cox, D.R. (1955). Some statiscal models related with series of events. *Journal of Royal Statistical Society Series B* **17**, 129-164.

[19] Cox, D.R. and Isham, V. (1980) *Point Processes*. New York: Chapman and Hall.

[20] Cressie, N.A.C. (1993) *Statistics for Spatial Data* (rev. ed.). New York: John Wiley & Sons.

[21] Daley, D.J. and Vere-Jones, D. (2003) *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. 2.ed. New York: Springer-Verlag.

[22] Diggle, P.J. (2000) Overview of statistical methods for disease mapping and its relationship to cluster detection. In *Spatial Epidemiology: Methods and Applications*. P. Elliott, J.C. Wakefield, N.G. Best and D.G. Briggs (eds). Oxford: Oxford University Press, pp. 87–103.

[23] Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns*. 2.ed. London: Arnold.

[24] Diggle, P., Zheng, P. and Durr, P. (2005a) Nonparametric estimation of spatial segregation in a multivariate point process: bovine tuberculosis in Cornwall, UK. Applied Statistics, **54** (3), 645–658.

[25] Diggle, P., Rowlingson, B. and Su, T. (2005b) Point Process Methodology for On-line Spatio-temporal Disease Surveillance. *Environmetrics*, **16**, 423–434.

[26] Doornik, J.A. (2002). *Object-Oriented Matrix Programming Using Ox*. 3.ed. London: Timberlake Consultants and www.nuff.ox.ac.uk/Users/Doornik.Ox programming.

[27] Dorai-Raj, S.S. (2001) *First- and Second-Order Properties of Spatiotemporal Point Processes in the Space-Time and Frequency Domains*. Unpublished Ph.D. Thesis, Faculty of the Virginia Polytechnic Institute and State University.

[28] Faes, C., Aerts, M., Geys, H., Bijnens, L. Donck, L.V. e Lammers, W. J. (2006) GLMM Approach to Study the Spatial and Temporal Evolution of Spikes in Small Intestine. *Statistical Modelling*, **6**, 300–320.

[29] Fishman e Snyder (1976) The Statistical Analysis of Space-Time Point Processes. *IEEE Transactions on Information Theory*, **22**, 257–274.

[30] Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, **15**, 183–202.

[31] Gamerman, D. (1997). Sampling from the posterior distribution in generalized linear mixed models. *Statistics and Computing*, **7**, 57–68.

[32] Gamerman, D. (1992) A dynamic approach to the statistical analysis of point processes. *Biometrika*, **79**, 39–50.

[34] Gamerman, D. and Lopes. H.F. (2006) Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. 2.ed. New York: Chapman and Hall/CRC.

[34] Gamerman, D., Moreira, A.R.B. and Rue, H. (2003) Space-varying regression models: specifications and simulation. *Computational Statistics and Data Analysis*, **42**, 513–533.

[35] Gamerman, D., Salazar, E. and Reis, E.A. (2007) Dynamic Gaussian process priors, with applications to the analysis of space-time data (with discussion). In *Bayesian Statistics 8*. J.M. Bernardo, M.J. Bayarri, J.O. Berger, A.P. Dawid, D. Heckerman, A.F.M. Smith and M. West (eds). Oxford: Oxford University Press, pp. 1–25.

[36] Gelfand, A.E., Banerjee, S. and Gamerman, D. (2005) Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics*, **16**, 465–479.

[37] Gelfand, A.E. and Ghosh, S. (1998) Model choice: a minimum posterior predictive loss approach. *Biometrika*, **85**, 1–11.

[38] Gelfand, A.E., Kim, H., Sirmans, C.F. and Banerjee, S. (2003) Spatial modelling with spatial varying coefficient processes. *Journal of the American Statistical Association*, **98**, 387–396.

[39] Gelfand, A.E. and Smith, A.F.M. (1990) Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398–409.

[40] Gelman, A. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–511.

[41] Geweke, J. (1992) Evaluating the accuracy of sampling-based approaches to the calculation of the posterior moments (with discussion). In *Bayesian Statistics 4*. J. M. Bernardo *et al.* (eds). Oxford: Oxford University Press, pp. 169–193.

[42] Hastings, W.K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.

[44] Heikkinen, J. and Arjas, E. (1998). Non-parametric Bayesian estimation of a spatial Poisson intensity. *Scandinavian Journal of Statistics*, **25**, 435–450.

[44] Heikkinen, J. and Arjas, E. (1999). Modelling a Poisson forest in variable elevations: a nonparametric Bayesian approach. *Biometrics*, **55**, 738–745.

[45] Knorr-Held, L. and Rue, H. (2002). On block updating in Markov random field models for disease mapping. *Scandinavian Journal of Statistics*, **29** (4), 597–614.

[46] Lantuéjoul, C. (1994). Nonconditional simulation of stationary isotropic multigaussian random functions. In *Geostatistical Simulations*. M. Armstrong and P. Dowd (eds). Dordrecht: Kluwer Academic Publishers.

[47] Lewis, P.A.W. and Shedler, G.S. (1979) Simulation of non-homogenous Poisson processes by thinning. *Naval Research Logistics Quartely*, **26**, 403–413.

[48] Liu, H. e Brown, D. E. (2003). Criminal Incident Prediciton Using a Point-Pattern-Based Density Model. *International Journal of Forecasting*, **19**, 603–622.

[49] Matérn, B. (1986) Spatial Variation. 2.ed. Berlin: Springer-Verlag.

[50] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equation of state calculations by fast computing machine. *Journal of Chemical Physics*, **21**, 1087–1091.

[51] Migon, H.S. and Gamerman, D. (1999) *Statistical Inference: an Integrated Approach*. London: Arnorld.

[52] Migon, H.S., Gamerman, D., Lopes, H.F. and Ferreira, M.A.R. (2005) Dynamic Models. In *Handbook of Statistics*. D. Dey and C. R. Rao (eds), **25**, 553–588.

[53] Møller, J., Syversveen, A. and Waagepetersen, R. (1998). Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, **25**, 451–482.

[54] Møller, J. and Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes.* Chapman & Hall.

[55] Møller, J., and Waagepetersen, R.P. (2007). Modern Statistics for Spatial Point Processes. *Scandinavian Journal of Statistics*, **34**, 643–684.

[56] Ogata, Y. (1998). Space-time point process models for earthquake occurrences. *The Annals of the Institute of Statistical Mathematics*, **50**, 379–402.

[57] Paez, M.S. (2004) *Análise de modelos para a estimação e previsão de processos espaço-temporais*. Tese de Doutorado, Programa de Pós-Graduação em Estatística, Universidade Federal do Rio de Janeiro.

[58] Paez, M.S. and Diggle, P. (2006). *Cox processes in time for point patterns and their aggregations*. Relatório Técnico, Instituto de Matemática, Universidade Federal do Rio de Janeiro.

[59] R Development Core Team (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

[60] Ravines, R.E.R. (2006) *Um Esquema Eficiente de Amostragem em Modelos Dinâmicos Generalizados com Aplicações em Funções de Transferência*. Tese de Doutorado, Programa de Pós-Graduação em Estatística, Universidade Federal do Rio de Janeiro.

[61] Ribeiro Jr., P.J. and Diggle, P.J. (2001). geoR: A package for geostatistical analysis. *R-News*, **1**.

[62] Richardson, S. (2003). Spatial models in epidemiological applications. In *Highly Structured Stochastic Systems*. P.J. Green, N.L.Hjort and S. Richardson (eds).Oxford: Oxford University Press, 237–259.

[63] Rowlingson, B.S. and Diggle, P.J. (1993) Splancs: Spatial point pattern analysis code in S-plus. *Computers in Geosciences*, **19**, 627–655

[64] Rue, H. (2001). Fast Sampling of Gaussian Random Fields. *Journal of de Royal Statistical Society Series B*, **63**, 325–338.

[65] Rue, H. and Martino, S. (2006). Approximate Bayesian inference for hierarchical Gaussian Markov random fields models. Preprint Statistics 7, Norwegian University of Science and Technology.

[66] Rue, H. and Tjelmeland, H. (2002). Fitting Gaussian Markov Random Fields to Gaussian Fields. *Scandinavian Journal of Statistics*, **29**, 31–49.

[67] Salazar, E. (2006) *Análise fatorial espacial dinâmica*. Exame de Qualificação de Doutorado não publicado. Universidade Federal do Rio de Janeiro.

[68] Schlather, M. (2001). Simulation of stationary and isotropic random fields. *R-News*, **1**, 18–20.

[69] Schoenberg, F.P. (2005) Consistent parametric estimation of the intensity of a spatial-temporal point process. *Journal of Statistical Planning and Inference* **128**, 79–93.

[70] Schoenberg, F.P., Brillinger, D.R. and Guttorp, P. (2002) Point Processes, Spatial-Temporal. In *Encyclopedia of Environmetrics*, **3**, 1573-1577.

[71] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Linde, A. van der (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B*, **64**, 583–639.

[72] Waagepetersen, R. (2004). Convergence of posteriors for discretized LGCPs. *Statistics and Probability Letters*, **66**, 229–235.

[73] West, M. and Harrison, P.J. (1997) *Bayesian Forecasting and Dynamic Models*. 2.ed. New York: Springer-Verlag.

[74] West, M., Harrison, P.J. and Migon, H.S. (1985) Dynamic generalised linear models and Bayesian forecasting (with discussion). *Journal of the American Statistical Association*, **80**, 73–97.

Apêndice A

A.1 O Filtro de Kalman

Considere o modelo dinâmico linear normal definido por

$$\begin{aligned} y_t &= F_t^{'}\theta_t &+ \epsilon_t, \qquad & \epsilon_t \sim N[0;V_t], \\ \theta_t &= G_t\theta_{t-1} &+ \omega_t, \qquad & \omega_t \sim N[0;W_t], \end{aligned}$$

com $\{F_t, G_t, V_t, W_t\}$ conhecido $\forall t{=}1,...,T$. Defina $D_t = D_{t-1} \cap \{y_t\}$, que denota a informação até o tempo t; D_0 representa a informação a priori de que $\theta_1 \sim N[a_1; R_1]$.

Os três passos do Filtro de Kalman (Anderson e Moore, 1979) são descritos a seguir. Sucessivamente para $t = 1, 2, ..., T$, faça:

1. **Evolução:** Passagem de $p(\theta_{t-1} \,|\, D_{t-1})$ para $p(\theta_t \,|\, D_{t-1})$.

   Denotando $\theta_{t-1} \,|\, D_{t-1} \sim N[m_{t-1}; C_{t-1}]$, sabendo que $\theta_t \,|\, \theta_{t-1} \sim N[G_t\theta_{t-1}; W_t]$ e que

   $$p(\theta_t \,|\, D_{t-1}) = \int p(\theta_t \,|\, \theta_{t-1})\, p(\theta_{t-1} \,|\, D_{t-1})\, d\theta_{t-1},$$

   tem-se que esta distribuição a priori $\theta_t \,|\, D_{t-1} \sim N[a_t; R_t]$,
   com $a_t = G_t m_{t-1}$ e $R_t = G_t C_{t-1} G_t' + W_t$.

2. **Previsão:** Obtenção de $p(y_t \,|\, D_{t-1})$.

   Como $y_t \,|\, \theta_t \sim N[F_t^{'}\theta_t; V_t]$ e $\theta_t \,|\, D_{t-1} \sim N[a_t; R_t]$, de

   $$p(y_t \,|\, D_{t-1}) = \int p(y_t \,|\, \theta_t)\, p(\theta_t \,|\, D_{t-1})\, d\theta_t$$

   tem-se que esta distribuição preditiva é $y_t \,|\, D_{t-1} \sim N[f_t; Q_t]$, com $f_t = F_t{}' a_t$ e $Q_t = F_t{}' R_t F_t + V_t$.

3. **Atualização:** Passagem de $p(\theta_t \,|\, D_{t-1})$ para $p(\theta_t \,|\, D_t)$.

   Como $y_t \,|\, D_{t-1} \sim N[f_t; Q_t]$ e $\theta_t \,|\, D_{t-1} \sim N[a_t; R_t]$, de

$$p(\theta_t \,|\, D_t) \propto p(\theta_t \,|\, D_{t-1})\, p(y_t \,|\, D_{t-1})$$

   tem-se que esta distribuição a posteriori é $\theta_t \,|\, D_t \sim N[m_t; C_t]$,

   com $m_t = a_t + R_t F_t Q_t^{-1}(y_t - f_t)$, $C_t = R_t - R_t F_t Q_t^{-1} F_t' R_t$.

A.2 O Algoritmo FFBS

O algoritmo *FFBS-Forward Filtering Backward Smoothing* (Carter e Kohn, 1994; Frühwirth-Schnatter, 1994), é um esquema MCMC de amostragem da distribuição a posteriori em modelos dinâmicos.

Considere o modelo dinâmico linear normal definido por

$$\begin{aligned} y_t &= F_t{}'\theta_t \quad + \epsilon_t, & \epsilon_t &\sim N[0,;V], \\ \theta_t &= G_t\theta_{t-1} + \omega_t, & \omega_t &\sim N[0;W] \quad \text{e} \quad \theta_1 \sim N[a_1;R_1]. \end{aligned}$$

Baseado no fato de que a densidade a posteriori pode ser escrita como

$$\pi(\theta_1, ..., \theta_T, V, W \,|\, D_T) \;\propto\; \pi(\theta_1, ..., \theta_T \,|\, V, W, D_T) \cdot \pi(V, W \,|\, D_T) \tag{A.1}$$

com $D_T = \{y_1, ..., y_T\}$, o FFBS atualiza os parâmetros em dois blocos, um deles formado por $(\theta_1, ..., \theta_T)$e o outro formado por (V, W).

Além disso, pode ser mostrado que

$$\pi(\theta_1, ..., \theta_T \,|\, V, W, D_T) \;=\; \pi(\theta_T \,|\, V, W, D_T) \prod_{t=1}^{T-1} \pi(\theta_t \,|\, \theta_{t+1}, V, W, D_t),$$

onde

$$\theta_T \mid V, W, D_T \sim N[m_T; C_T] \tag{A.2}$$

$$\text{e, para } t = 1, ..., T-1, \quad \theta_t \mid \theta_{t+1}, V, W, D_t \sim N[m_t^*; C_t^*],$$

$$m_t^* = C_t^*(G_t' W^{-1} \theta_{t+1} + C_t^{-1} m_t) \quad \text{e} \quad C_t^* = (G_t' W^{-1} G_t + C_t^{-1})^{-1},$$

onde m_t e C_t são, respectivamente, a média e a variância da distribuição atualizada de θ_t no Filtro de Kalman. Desse modo, a atualização do bloco $(\theta_1, ..., \theta_T)$ é feita em duas etapas: a primeira etapa é a análise seqüencial "para frente" do Filtro de Kalman e a segunda etapa é a suavização "para trás" da equação (A.2).

Resumindo, os passos do esquema FFBS de amostragem da posteriori (A.1) são:

1. **Inicialização:** inicialize o contador de iterações da cadeia em $j=1$ e atribua valores iniciais $(\boldsymbol{\theta}^{(0)}, V^{(0)}, W^{(0)})$;

2. **Amostragem de** $(\theta_1, ..., \theta_T)$**:**
   (a) Amostre θ_T da distribuição $\theta_T \mid V^{(j-1)}, W^{(j-1)}, D_T$ (do Filtro de Kalman) e faça $t=T-1$;
   (b) Amostre θ_t da distribuição $\theta_t \mid \theta_{t+1}^{(j)}, V^{(j-1)}, W^{(j-1)}, D_t$ dada em (A.2);
   (c) Decresça t para $t-1$ e retorne ao passo (b) até $t=1$;

3. **Amostragem de** V **e** W**:** $V^{(j)}$ e $W^{(j)}$ são amostrados sucessivamente de suas respectivas condicionais completas $p(V \mid \boldsymbol{\theta}^{(j)}, W^{(j-1)}, D_T)$ e $p(W \mid \boldsymbol{\theta}^{(j)}, V^{(j)}, D_T)$;

4. **Atualização:** Mude o contador de j para $j+1$ e retorne ao passo 2 até que a convergência da cadeia seja atingida.

A.3 Algoritmo de Gamerman (1997)

Sejam Y_i, $i = 1, 2, ..., n$, variáveis aleatórias com função de densidade de probabilidade pertencente à Família Exponencial da forma $p(y_i \mid \eta_i) \propto \exp\{y_i \eta_i - b(\eta_i)\}$, onde η_i é um parâmetro desconhecido e $b(\eta_i)$ uma função conhecida. Defina $\mu_i \doteq E(Y_i \mid \eta_i) = b'(\eta_i)$ (a primeira derivada) e assuma que as Y_i´s são condicionalmente independentes dadas as médias μ_i's.

Considere o modelo de regressão $g(\mu_i) = \boldsymbol{x}_i' \boldsymbol{\beta}$, no qual g é uma função que projeta μ_i na reta real, $\boldsymbol{x}_i$ são os valores das variáveis regressoras e $\boldsymbol{\beta}$ são os coeficientes de regressão desconhecidos. Desse modo, o objetivo da inferência é estimar $\boldsymbol{\beta}$.

Com escolha de uma distribuição a priori normal para β, com vetor de médias μ_0 e matriz de variâncias Σ_0, a distribuição a posteriori é dada por

$$\pi(\boldsymbol{\beta}|\boldsymbol{y}) \;\propto\; \exp\left\{-\frac{1}{2}(\boldsymbol{\beta}-\mu_0)'\Sigma_0^{-1}(\boldsymbol{\beta}-\mu_0) + \sum_{i=1}^{n}[y_i\eta_i - b(\eta_i)]\right\},$$

na qual o termo da verossimilhança depende dede $\boldsymbol{\beta}$ através de $\eta_i,\ i=1,...,n$.

A idéia da versão bayesiana do algoritmo IRLS, adaptado por West *et al.* (1985), é modelar $g(y_i) \sim N(M_i, V_i)$ através da expansão de Taylor de $g(y_i)$ em torno de M_i, aproximada por $g(y_i) \cong g(M_i) + (y_i - M_i)\, g'(M_i)$. Definem-se as variáveis de trabalho

$$\tilde{y}_i = g(M_i) + (y_i - M_i)\, g'(M_i),$$

que são tais que

$$\tilde{M}_i = E(\tilde{y}_i) = g(M_i) = \boldsymbol{x}_i'\boldsymbol{\beta} \qquad \text{e} \qquad \tilde{V}_i = Var(\tilde{y}_i) = [g'(M_i)]^2\, Var(\tilde{y}_i).$$

Fazendo $\tilde{\boldsymbol{y}} = (\tilde{y}_1, ..., \tilde{y}_n)'$ e $\tilde{V} = diag(\tilde{V}_1, ..., \tilde{V}_n)$, a combinação do modelo de regressão $\tilde{\boldsymbol{y}}\,|\,\boldsymbol{\beta} \sim N[X\boldsymbol{\beta}; \tilde{V}]$ com a distribuição a priori resulta na seguinte aproximação para distribuição a posteriori

$$\boldsymbol{\beta} \mid \tilde{\boldsymbol{y}} \sim N[\mu_\beta; \Sigma_\beta], \;\; \text{com} \;\; \mu_\beta = \Sigma_\beta(\Sigma_0^{-1}\mu_0 + X'\tilde{V}^{-1}\tilde{\boldsymbol{y}}) \;\; \text{e} \;\; \Sigma_\beta = (\Sigma_0^{-1} + X'\tilde{V}^{-1}X)^{-1}. \tag{A.3}$$

Gamerman (1997) propõe usar esta densidade a posteriori como a densidade da proposta no algoritmo de Metropolis-Hastings.

Os passos deste esquema de amostragem são:

1. Inicialize o contador de iterações da cadeia em $j=1$ e atribua valor inicial $\boldsymbol{\beta}^{(0)}$;

2. Obtenha um novo valor $\boldsymbol{\beta}^*$ para $\boldsymbol{\beta}$ gerado da distribuição em (A.3), cuja função de densidade de probabilidade é denotada por $q(\boldsymbol{\beta}^*\,|\,\boldsymbol{\beta}^{(j-1)})$;

3. Avalie a probabilidade de aceitação do novo valor, dada por

$$\alpha(\boldsymbol{\beta}^{(j-1)}, \boldsymbol{\beta}^*) = min\left\{ 1,\; \frac{\pi(\boldsymbol{\beta}^* \mid \boldsymbol{y}) \;\cdot q(\boldsymbol{\beta}^{(j-1)} \mid \boldsymbol{\beta}^*)}{\pi(\boldsymbol{\beta}^{j-1} \mid \boldsymbol{y}) \cdot q(\boldsymbol{\beta}^* \mid \boldsymbol{\beta}^{(j-1)})} \right\}.$$

   Se o novo valor é aceito, $\boldsymbol{\beta}^{(j)} = \boldsymbol{\beta}^*$; caso contrário, $\boldsymbol{\beta}^{(j)} = \boldsymbol{\beta}^{(j-1)}$;

4. Mude o contador de j para $j+1$ e retorne ao passo 2 até que a convergência da cadeia seja atingida.