

Universidade Federal do Rio de Janeiro

ANÁLISE ESPAÇO-TEMPORAL DA DISTRIBUIÇÃO  
DOS CASOS DE DENGUE NA CIDADE DO RIO DE  
JANEIRO NO PERÍODO DE 1986 A 2002

Gustavo da Silva Ferreira

2004

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



UFRJ

ANÁLISE ESPAÇO-TEMPORAL DA DISTRIBUIÇÃO DOS CASOS  
DE DENGUE NA CIDADE DO RIO DE JANEIRO  
NO PERÍODO DE 1986 A 2002

Gustavo da Silva Ferreira

Dissertação de Mestrado 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 Mestre em Ciências Estatísticas.

Orientadores:

Alexandra M. Schmidt e Dani Gamerman

Rio de Janeiro

Junho de 2004

# ANÁLISE ESPAÇO-TEMPORAL DA DISTRIBUIÇÃO DOS CASOS DE DENGUE NA CIDADE DO RIO DE JANEIRO NO PERÍODO DE 1986 A 2002

Gustavo da Silva Ferreira

Orientadores:

Alexandra M. Schmidt e Dani Gamerman

Dissertação de Mestrado 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 Mestre em Ciências Estatísticas.

Aprovada por :

---

Presidente, Prof<sup>a</sup>. Alexandra M. Schmidt

---

Prof. Paulo Justiniano Ribeiro Jr.

---

Prof. Marco A. R. Ferreira

---

Prof. Dani Gamerman

Rio de Janeiro

Junho de 2004

Ferreira, Gustavo da Silva

Análise espaço-temporal da distribuição dos casos de dengue na cidade do Rio de Janeiro no período de 1986 a 2002 / Gustavo da Silva Ferreira. - Rio de Janeiro: UFRJ/IM, 2004.

xi, 113f.: il.; 31cm.

Orientador: Alexandra M. Schmidt e Dani Gamerman

Dissertação (mestrado) - UFRJ/IM/ Programa de Pós-graduação em Estatística, 2004.

Bibliografia: p.105-108.

1. Dengue 2. Estatística Bayesiana. 3. Estatística Espacial 4. Modelos Lineares Generalizados Dinâmicos I. Schmidt, Alexandra M. e Gamerman, Dani II. Universidade Federal do Rio de Janeiro, Instituto de Matemática. III. Título.

## Resumo

# ANÁLISE ESPAÇO-TEMPORAL DA DISTRIBUIÇÃO DOS CASOS DE DENGUE NA CIDADE DO RIO DE JANEIRO NO PERÍODO DE 1986 A 2002

Gustavo da Silva Ferreira

Orientadores:

Alexandra M. Schmidt e Dani Gamerman

Séries de notificações de dengue na cidade do Rio de Janeiro foram analisadas neste estudo de 3 formas distintas: Análise Temporal, Análise Espacial e Análise Espaço-Temporal. A primeira utiliza as idéias de Modelos Lineares Generalizados Bayesianos e analisa a série de dados mensais de dengue referentes ao período de 1986 a 2002. Nesta análise busca-se estudar o efeito de fatores sazonais e anuais no aumento e diminuição dos casos de dengue. Na análise espacial, os casos de dengue referentes ao período que compreende a epidemia 2001-2002 são agregados no tempo e analisados. Os modelos empregados foram baseados nas idéias de utilização de Campos Aleatórios Markovianos no mapeamento de doenças (Mollié, 1996). A distribuição espacial dos riscos de dengue foi analisada segundo várias combinações de covariáveis e especificações das prioris para os efeitos aleatórios dos modelos. A análise espaço-temporal analisa a evolução dos riscos de dengue nos bairros ao longo de 70 semanas consecutivas no período 2001-2002. Os modelos empregados nesta análise combinam os modelos anteriores com as idéias de Modelos Lineares Generalizados Dinâmicos (West, Harrison e Migon, 1985).

*Palavras-chave:* dengue; risco relativo; efeitos aleatórios; estruturas de vizinhança; modelos lineares generalizados dinâmicos.

# Abstract

## SPATIO-TEMPORAL ANALYSIS OF THE DISTRIBUTION OF THE NUMBER OF CASES OF DENGUE FEVER IN THE CITY OF RIO DE JANEIRO DURING THE PERIOD FROM 1986 UNTIL 2002

Gustavo da Silva Ferreira

Supervisors:

Alexandra M. Schmidt and Dani Gamerman

Here, notifications of dengue fever cases are analyzed in the city of Rio de Janeiro under three different settings: temporal, spatial and spatio-temporal data. The purely temporal analysis is based on the ideas of Bayesian Generalized Linear Models and considers monthly observations of the number of cases of dengue during the period comprising the years from 1986 until 2002. The aim is to search for seasonal and annual factors which might be present in the increase and decrease of the number of cases of the disease. In the purely spatial analysis, the number of cases observed during the epidemic period, from December 2001 until May 2002, are aggregated in time and analyzed. The proposed models are based on the ideas of Markov Random Fields as described in (Mollié, 1996). The spatial distribution of dengue's relative risks are fitted according to the use and combination of different explanatory variables and different prior specifications for the random effects considered in the model. Finally, the spatio-temporal analysis studies the temporal evolution of the relative risks of dengue in the areas of Rio de Janeiro during 70 consecutive weeks starting from December 2001 until May 2002. In this case the fitted models combine the purely temporal and spatial models described above with those of Generalized Dynamic Linear Models (West, Harrison e Migon, 1985).

# Agradecimentos

À Deus, por ter me capacitado a enfrentar todas as dificuldades provenientes da realização deste curso.

À Carla, por ser minha razão maior de estar nesta cidade.

À meus pais, Glênio e Graziela, que só me permitiram vir ao Rio de Janeiro com a promessa de um regresso rápido, mas que sempre me ajudaram muito quando necessitei.

Aos meus irmãos e demais familiares pela torcida realizada em Porto Alegre e Canoas.

Aos meus orientadores, Alexandra Schmidt e Dani Gamerman, além dos demais professores do Departamento de Métodos Estatísticos da UFRJ, pela transmissão do conhecimento.

À minha orientadora de graduação, Jandyra Fachel, por ter me incentivado a seguir em frente.

Aos meus amigos e colegas Leonardo, Esther, Juan, Geraldo e Manoela, assim como aos veteranos da turma do mestrado de 2001.

Aos demais colegas que passaram pela Casa dos Mestrandos, em espe-



cial, Carlos Abanto e Fernando Pulgati.

À Márcia Fontes e Cecília Nicolai, da Secretaria Municipal de Saúde do Rio de Janeiro (SMS-RJ), Roberto Medronho e Ivan Ortiz Valencia, do Núcleo de Estudos em Saúde Coletiva (NESC-UFRJ) e Antonio Ponce de Leon e Washington Junger, do Departamento de Epidemiologia da UERJ, pelos dados fornecidos e pelas considerações que contribuíram neste estudo.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	O Dengue . . . . .	1
1.2	Infestação do <i>Aedes aegypti</i> . . . . .	3
1.3	A Distribuição do Dengue . . . . .	5
1.4	Objetivos do Estudo . . . . .	7
<b>2</b>	<b>Inferência Bayesiana</b>	<b>10</b>
2.1	Teorema de Bayes . . . . .	10
2.2	Estimação de Parâmetros . . . . .	12
2.3	Distribuição Preditiva . . . . .	13
2.4	Inferência via Simulação Estocástica . . . . .	14
2.4.1	Inferência via MCMC . . . . .	14
2.4.2	Software WinBugs . . . . .	19

<b>3</b>	<b>Modelos Lineares</b>	<b>20</b>
3.1	Modelos Lineares Generalizados . . . . .	21
3.2	Modelos Lineares Generalizados Dinâmicos . . . . .	23
3.2.1	Modelos Lineares Dinâmicos . . . . .	24
3.2.2	O Modelo Linear Generalizado Dinâmico . . . . .	26
3.3	Modelos Hierárquicos . . . . .	27
<b>4</b>	<b>Análise Temporal</b>	<b>28</b>
4.1	Dados Utilizados . . . . .	29
4.2	Modelo Proposto . . . . .	31
4.2.1	Comparação entre os Modelos . . . . .	34
4.3	Resultados . . . . .	35
4.3.1	Modelo sem Efeito de Sobredispersão . . . . .	36
4.3.2	Modelo com Efeito de Sobredispersão . . . . .	38
4.4	Conclusões . . . . .	46
<b>5</b>	<b>Análise Espacial</b>	<b>48</b>
5.1	Dados Utilizados . . . . .	49
5.1.1	Covariáveis . . . . .	50
5.2	Modelo Proposto . . . . .	54

5.2.1	Estruturas de vizinhanças . . . . .	59
5.2.2	Comparação entre os Modelos . . . . .	61
5.3	Resultados . . . . .	62
5.4	Conclusões . . . . .	80
<b>6</b>	<b>Análise Espaço-Temporal</b>	<b>83</b>
6.1	Dados Utilizados . . . . .	84
6.2	Modelo Proposto . . . . .	85
6.2.1	Comparação entre os Modelos . . . . .	89
6.3	Resultados . . . . .	90
6.4	Conclusões . . . . .	101
<b>7</b>	<b>Considerações Finais</b>	<b>103</b>
	<b>Referências Bibliográficas</b>	<b>105</b>
	<b>Apêndice</b>	<b>109</b>

# Capítulo 1

## Introdução

### 1.1 O Dengue

Nas décadas de 1950 e 1960, o Brasil e mais 17 países das Américas conseguiram eliminar o mosquito transmissor do dengue, o *Aedes aegypti*, de seus territórios. Na época, utilizou-se uma estratégia de campanha nacional, centralizada, verticalizada, com estruturação militar, onde a disciplina e a hierarquia eram características marcantes. Entretanto, alguns destes países não obtiveram o mesmo sucesso fazendo com que o Brasil enfrentasse centenas de re-infestações, as quais foram detectadas precocemente e eliminadas (Tauil, 2002).

Em 1976, foi detectada uma infestação que não pôde ser eliminada, disseminando-se para outros estados, como o Rio Grande do Norte e o Rio de Janeiro. Após isso, o *Aedes aegypti* re-infestou várias Unidades da Federação além de todos os 18 países das Américas que o eliminaram de seus territórios no passado (Tauil, 2002).

O número de municípios brasileiros nos quais tem sido assinalada a presença de *Aedes aegypti*, vetor da dengue e da febre amarela, tem aumentado rapidamente nos últimos anos. Atualmente, esse vetor pode ser encontrado em todos os estados do Brasil (Santos e Carvalho, 2000).

Entre as doenças reemergentes, o dengue é a que se constitui em problema mais grave de saúde pública (Tauil, 2002).

O dengue é transmitido ao homem por algumas espécies do gênero *Aedes*, sendo o *Aedes aegypti* a mais importante. Além do homem, não se conhece nenhum outro hospedeiro de importância significativa que atue como reservatório (Medronho, 1995).

O *Aedes aegypti* possui habitat domiciliar e peridomiciliar, com curto raio de distanciamento (100 a 200 m) dos locais de oviposição (Medronho, 1995). Somente as fêmeas se alimentam de sangue, as quais possuem hábitos alimentares diurnos e utilizam-se preferencialmente de depósitos artificiais de água limpa para colocar os seus ovos, que por sua vez, têm uma alta capacidade de resistir à dessecação, mantendo-se viáveis na ausência de água por até 450 dias. O mosquito tem mostrado uma grande capacidade de adaptação a diferentes situações ambientais consideradas desfavoráveis. Adultos já foram encontrados em altitudes elevadas e larvas em água poluída (Tauil, 2002).

Assim, o dengue apresenta um ciclo epidemiológico urbano e seus principais elementos são o homem - o hospedeiro do vírus - e o *Aedes aegypti* - o vetor de transmissão (Rebêlo *et al.*, 1999).

A possibilidade de transmissão do vírus do dengue através de gerações de mosquitos pode manter o vírus na natureza em condições adversas, como por exemplo, na ausência ou diminuição do número de suscetíveis (SESRJ,

1988 *apud* Medronho, 1995).

Finalmente, deve ser ressaltado que os programas de erradicação do *Aedes aegypti* na América Latina inicialmente visavam o controle da febre amarela. Posteriormente, verificou-se ser impossível controlar o vírus da febre amarela na natureza, e a erradicação mundial desta doença foi considerada uma meta inatingível. Em consequência desse fato, aliado à descoberta da vacina contra a febre amarela, o controle do *Aedes aegypti* tornou-se uma baixa prioridade em muitos países (Slosek, 1986 *apud* Medronho, 1995). Isto fez com que o mosquito retornasse para algumas áreas das quais fora totalmente erradicado (Medronho, 1995).

Além disso, a persistência da circulação do vírus é favorecida pelas elevadas densidades das populações humanas, migrações e nascimentos que repõem o estoque de indivíduos suscetíveis criando oportunidades para o ciclo de transmissão (Teixeira *et al.*, 2003).

## 1.2 Infestação do *Aedes aegypti*

Outro aspecto relevante em estudos de dengue diz respeito à infestação do mosquito em uma determinada cidade ou região. Existem dois índices amplamente utilizados em campanhas de combate ao mosquito. São eles:

- (i) *Índice de Infestação Predial* (IIP): é a relação, expressa em porcentagem, entre o número de imóveis positivos, isto é, onde foram encontradas larvas e/ou pupas da espécie em avaliação, e o número de imóveis pesquisados;

- (ii) *Índice de Breteau*: é a relação entre o número de recipientes positivos e o número de imóveis pesquisados.

Na literatura, há referência de que com um índice de infestação predial menor que 1% e um índice de Bretau abaixo de 5% não haveria transmissão de dengue. Porém, há relatos de transmissão de dengue em Cingapura com índice de Bretau abaixo de 5% (Kuno, 1995 *apud* Tauil, 2002). Portanto, baixos índices de infestação pelo *Aedes aegypti* reduzem o risco de transmissão de dengue, porém não o eliminam.

Além disso, nenhum dos índices é suficientemente capaz de medir a intensidade de infestação (Gomes, 1998 *apud* Tauil, 2002). No caso do IIP, um prédio será considerado como infestado independentemente de possuir 1 (um) ou vários recipientes positivos para larvas. No caso do índice de Bretau, não se diferencia o tipo de reservatório, contabilizando-se da mesma forma um tonel de água com larvas e um prato de xaxim com larvas, embora o número de larvas num tonel seja em geral, muitas vezes maior que no prato de xaxim (Tauil, 2002).

Cardoso *et al.* (1997), ao estudar o tempo necessário para que o *Aedes aegypti* e o *Aedes albopictus* fossem novamente detectados na área urbana do município de Catanduva - SP, mostraram que as ovitrampas <sup>1</sup> apresentaram positividade para *Aedes aegypti* dois meses após os trabalhos de controle, enquanto o Índice de Breteau veio a positivar-se somente no quarto mês após o término dos referidos trabalhos.

---

<sup>1</sup>As ovitrampas consistiam de vidros pretos nos quais foram instaladas paletas presas por prendedores de madeira à borda do recipiente e mergulhadas em água. As paletas das ovitrampas permaneceram expostas por uma semana e as larvas encontradas foram posteriormente identificadas.



Além disso, a utilização desses indicadores como medida da probabilidade da ocorrência de casos de dengue é inadequada uma vez que componentes ambientais e sociais, bem como fatores individuais associados ao hospedeiro e as populações dos agentes são fundamentais na análise de risco de transmissão de dengue <sup>2</sup>.

Estes fatos comprovam que as dificuldades de combater o *Aedes aegypti* são muito grandes nos dias de hoje. Outros exemplos recentes podem ser lembrados, entre eles o de Cuba que, após reduzir drasticamente os índices de infestação, enfrentou o reaparecimento do *Aedes aegypti* transmitindo dengue em fins de 2001 e início de 2002. Outro exemplo é Cingapura, modelo internacional de combate ao vetor do dengue, que ainda apresenta um resíduo de infestação, insuficiente para impedir transmissão esporádica da doença (Tauil, 2002).

### 1.3 A Distribuição do Dengue

Os fatores determinantes do ressurgimento do dengue são complexos e ainda não completamente conhecidos (Teixeira *et al.*, 2003).

O cenário propiciado pelo processo industrial moderno de embalagens descartáveis contribui para a multiplicação dos mosquitos quando estas embalagens de plástico, alumínio, vidro ou isopor não são adequadamente recolhidas após sua utilização. O fantástico aumento da produção de veículos automotores contribui igualmente para a multiplicação do vetor, na medida em que aumenta o número de pneus usados dispostos inadequadamente no

---

<sup>2</sup>SUCEN - SP, 2000 - 2001: Elaborado por Mariza Pereira e Silvana Gazola Santucci - [www.sucen.sp.gov.br](http://www.sucen.sp.gov.br)

meio ambiente, comportando-se como recipientes prioritários para a postura de ovos pelos mosquitos, e permitindo o transporte passivo de ovos, larvas e insetos adultos facilitando a sua disseminação (Tauil, 2002).

Além destes fatores, o fluxo rural-urbano intenso de pessoas nos últimos trinta anos resultou numa concentração populacional muito elevada nos grandes centros urbanos. Mais de 80% da população brasileira vive hoje em área urbana. As cidades, pressionadas por essa demanda, não conseguiram oferecer condições satisfatórias de habitação e de saneamento básico a uma fração importante dos seus habitantes, os quais vivem em favelas, invasões, mocambos ou cortiços, onde o abastecimento de água e a coleta de dejetos são irregulares (Tauil, 2002).

Por estes motivos, a distribuição do *Aedes aegypti* se torna cada vez mais abrangente. O rápido crescimento e urbanização das populações nas áreas tropicais ampliaram a faixa de ocorrência do dengue, em razão da difusão do mosquito em áreas antes livres da doença (Rebêlo *et al.*, 1999).

Apesar disto, a distribuição dos riscos de dengue em relação às distintas situações sociais e econômicas ainda é uma questão controversa.

Estas características juntamente com a constatação da exposição ao risco de dengue em diferentes espaços intra-urbanos o diferencia da grande maioria das doenças infecciosas e parasitárias (Teixeira *et al.*, 2003).

Vários estudos sobre a distribuição do mosquito *Aedes aegypti* no Brasil têm sido realizados. Entre eles, podemos citar um estudo realizado nos municípios do estado do Maranhão (Rebêlo *et al.*, 1999) e o trabalho realizado por Santos e Carvalho (2000) a respeito da distribuição do mosquito *Aedes aegypti* em um bairro do Rio de Janeiro, onde propuseram a utilização de

técnicas de análise de dados espaciais na vigilância e controle do mosquito.

Santos e Carvalho (2000) já haviam ressaltado a necessidade de estudos mais consistentes em áreas urbanas, que possibilitem uma melhor compreensão dos fatores que possam influenciar na ocorrência e distribuição de *Aedes aegypti* e que orientem as ações de controle. Os autores ressaltaram também a necessidade de identificar os locais com base em indicadores de risco e não apenas com base na densidade da população do mosquito.

A realização de investigações sobre o papel da população, considerando-se condições econômicas e ambientais, pode acrescentar elementos para o debate de estratégias de prevenção mais adequadas (Teixeira *et al.*, 2003).

## 1.4 Objetivos do Estudo

O presente estudo busca analisar a distribuição espaço-temporal dos casos de dengue na cidade do Rio de Janeiro - Brasil a partir de notificações semanais de casos efetuadas durante os anos de 2001 e 2002 e de notificações mensais referentes ao período de 1986 à 2002. Aspectos sócio-econômicos e ambientais foram utilizados em busca de possíveis relações com a distribuição dos casos de dengue. Os dados foram obtidos junto à Secretaria Municipal de Saúde do Rio de Janeiro (SMS-RJ) e junto ao IBGE.

A análise estatística empregada utiliza a abordagem bayesiana, onde combinamos a informação dada pela verossimilhança com informações a priori e obtemos a distribuição dos riscos de dengue a posteriori.

Os modelos utilizados estão baseados nas idéias de utilização de Cam-

pos Aleatórios Markovianos no mapeamento de doenças (Mollié, 1996), Modelos Lineares Generalizados Bayesianos e Modelos Lineares Generalizados Dinâmicos (West *et al.*, 1985). Os princípios dos Modelos Lineares Generalizados podem ser encontrados em McCullagh e Nelder (1989).

Nas três análises empregadas neste estudo, o número de casos de dengue é modelado segundo uma distribuição de Poisson, onde o fator de risco é incorporado na composição do valor esperado de casos para um determinado bairro ou mês (ou mês e bairro, no caso espaço-temporal).

Assim, os principais objetivos do estudo são apresentados abaixo:

- analisar a distribuição espacial do risco de dengue nos bairros da cidade;
- busca de fatores sociais, geográficos ou econômicos que estejam relacionados com o risco de dengue de uma determinada área;
- analisar possíveis tendências, ciclos e peculiaridades que possam existir na série histórica de notificações de dengue na cidade;
- analisar a distribuição espaço-temporal do risco de dengue nos bairros da cidade ao longo de uma epidemia;
- detecção de épocas em que possíveis fatores ou desvios ocorreram, fazendo com que houvessem discrepâncias acentuadas entre os casos de dengue destes períodos e os casos esperados segundo os modelos utilizados.

Este estudo está dividido de forma a contemplar questões pertinentes da distribuição dos casos de dengue no Rio de Janeiro no tempo, no espaço e no espaço-tempo.

No Capítulo 2 são apresentados os fundamentos básicos da Inferência Bayesiana e dos métodos computacionais de Monte Carlo Via Cadeias de Markov (MCMC).

No Capítulo 3 apresentam-se alguns conceitos básicos a respeito de Modelos Lineares, Modelos Lineares Generalizados, Modelos Lineares Dinâmicos, Modelos Lineares Generalizados Dinâmicos e Modelos Hierárquicos.

No Capítulo 4 são apresentados os dados utilizados na análise puramente temporal, o modelo utilizado e os resultados obtidos a respeito de possíveis efeitos sazonais ou padrões temporais dos casos de dengue.

O Capítulo 5 analisa a distribuição dos casos de dengue durante a grande epidemia do verão de 2001/2002 e analisa a possível influência de fatores sociais e geográficos da cidade do Rio de Janeiro.

O Capítulo 6 complementa o estudo ao analisar os casos de dengue a partir de modelos que utilizam efeitos espaço-temporais sob diferentes cenários e a evolução dos casos ao longo de setenta semanas nos anos de 2001 e 2002.

Por fim, o Capítulo 7 apresenta as considerações finais e comenta a importância dos resultados obtidos.

# Capítulo 2

## Inferência Bayesiana

As conclusões obtidas através da inferência bayesiana a respeito de um determinado parâmetro  $\theta$ , ou de uma variável não observada  $\mathbf{y}$ , são baseadas em especificações probabilísticas. Tais especificações são feitas condicionalmente a uma amostra de valores observados, relacionados de alguma forma com as quantidades de interesse.

### 2.1 Teorema de Bayes

Todo o processo de inferência é baseado na atualização da informação através do Teorema de Bayes.

A descrição apresentada a seguir define o Teorema de Bayes no contexto da inferência bayesiana (Migon e Gamerman, 1999):

Seja  $\mathbf{H}$  a informação disponível inicialmente para alguma quantidade de interesse. Seja  $\theta$  o vetor de parâmetros desconhecidos que assume valores

em  $\Theta$ . Supondo que a informação inicial pode ser expressa em termos probabilísticos  $P(\theta|H)$ , temos uma descrição completa da incerteza a respeito de  $\theta$ .

Para atualizar a informação a respeito de  $\theta$ , podemos observar uma amostra  $\mathbf{x}$  de um vetor aleatório  $\mathbf{X}$  relacionado com  $\theta$ . Assim, a informação disponível para a inferência passará a ser  $H^* = H \cap \{\mathbf{X} = \mathbf{x}\}$ .

Precisamos ainda conhecer a distribuição amostral de  $\mathbf{X}$  para realizar a atualização. Esta distribuição, conhecida como *função de verossimilhança*, associa para cada valor de  $\theta$ , a plausibilidade de  $\mathbf{x}$  ser observado.

A partir da especificação dos elementos da inferência, temos:

$$p(\theta|H^*) = p(\theta|\mathbf{x}, H) = \frac{p(\theta, \mathbf{x}|H)}{p(\mathbf{x}|H)} = \frac{p(\mathbf{x}|\theta, H)p(\theta|H)}{p(\mathbf{x}|H)} \quad (2.1)$$

onde

$$p(\mathbf{x}|H) = \int p(\theta, \mathbf{x}|H)d\theta.$$

Como  $p(\mathbf{x}|H)$  não depende de  $\theta$  e como  $\mathbf{H}$  é comum a todos os termos, podemos reescrever o teorema da seguinte forma:

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta)p(\theta). \quad (2.2)$$

O resultado apresentado em (2.2) é conhecido como Teorema de Bayes e se constitui como a base de todos os procedimentos da inferência bayesiana.

## 2.2 Estimação de Parâmetros

O processo de estimação de parâmetros pode ser realizado a partir de uma estimação pontual ou de uma estimação por intervalo.

### (i) Estimação Pontual

No contexto da inferência bayesiana, a estimação pontual de um determinado parâmetro  $\theta$  pode ser vista como um problema de decisão. Os elementos que compõem este problema de decisão são (Migon e Gamerman, 1999):

- espaço de parâmetros  $\Theta$ ;
- espaço de resultados do experimento  $\Omega$ ;
- espaço de ações possíveis  $A$ .

Uma regra de decisão  $\delta$  é uma função definida em  $\Omega$  que assume valores em  $A$ , tal que  $\delta : \Omega \rightarrow A$ .

Associamos então uma função perda  $L$  a cada regra de decisão  $\delta(x)$ ,  $x \in \Omega$ , e a cada possível valor de  $\theta \in \Theta$ . Assim, temos uma medida de quanto perdemos quando tomamos a decisão  $\delta(x)$  e o valor verdadeiro do parâmetro é  $\theta$ .

O risco associado a uma determinada regra de decisão  $\delta(x)$  é obtido tomando-se a esperança da função perda  $L(\delta, \theta)$  com relação a distribuição a posteriori de  $\theta$ , ou seja,

$$R(\delta) = E_{\theta|x}(L(\delta, \theta)).$$

Um estimador pontual de  $\theta$  é dado pela regra de decisão que minimiza o risco esperado segundo uma função perda especificada.



Os estimadores pontuais associados às perdas quadráticas, perdas absolutas e perdas zero-um são a média, a mediana e a moda da distribuição a posteriori de  $\theta$ , respectivamente.

### (i) **Estimação por Intervalo**

Sob o ponto de vista bayesiano, a forma mais adequada de avaliar a informação disponível a respeito de uma quantidade desconhecida  $\theta$  é através da distribuição a posteriori (Migon e Gamerman, 1999). Assim, sumarizar a informação desta distribuição em um único valor não dá ao pesquisador uma medida de quão preciso ele é. Uma alternativa para este problema é calcular intervalos de credibilidade para estes valores.

Um intervalo de credibilidade bayesiano é definido da seguinte maneira:

- seja  $\theta$  uma quantidade desconhecida em  $\Theta$ . Uma região  $C \subset \Theta$  é um Intervalo de Credibilidade Bayesiano  $100(1 - \alpha)\%$  para  $\theta$  se  $p(\theta \in C|x) \geq 1 - \alpha$ .

Desta forma, o Intervalo de Credibilidade Bayesiano, com nível de credibilidade  $(1 - \alpha)$ , é denotado por  $IC(1 - \alpha)\%$ .

## **2.3 Distribuição Preditiva**

O processo de predição sob a perspectiva bayesiana é tratado naturalmente através da análise da distribuição preditiva.

A distribuição preditiva para um determinado valor  $y$ , baseada num conjunto de dados  $\mathbf{y}$ , é dada por:

$$p(y|\mathbf{y}) = \int p(y, \theta|\mathbf{y})d\theta = \int p(y|\theta, \mathbf{y})p(\theta|\mathbf{y})d\theta = \int p(y|\theta)p(\theta|\mathbf{y})d\theta \quad (2.3)$$

onde a última igualdade segue da independência entre  $y$  e  $\mathbf{y}$  dado  $\theta$ .

Analisando a forma de obtenção da distribuição preditiva, podemos ver que ela representa a verossimilhança esperada com relação a distribuição a posteriori de  $\theta$ , ou seja,

$$p(y|\mathbf{y}) = E_{\theta|\mathbf{y}}[p(y|\theta)].$$

## 2.4 Inferência via Simulação Estocástica

No contexto da inferência estatística, a simulação estocástica tem o objetivo de estimar características probabilísticas de modelos ou distribuições de interesse, as quais não poderiam ser obtidas analiticamente.

Na inferência bayesiana, os métodos de simulação estão relacionados ao processo de obtenção de amostras de distribuições a posteriori. Nos casos em que lidamos com problemas altamente complexos, em virtude do grande número de parâmetros envolvidos, geralmente temos de recorrer aos métodos de *Monte Carlo via Cadeias de Markov* para concluirmos o processo inferencial.

### 2.4.1 Inferência via MCMC

Os métodos computacionais baseados nos métodos de Monte Carlo via Cadeias de Markov (MCMC) têm sido amplamente utilizados na inferência

bayesiana quando o pesquisador está interessado em simular amostras de uma determinada distribuição a posteriori, cuja função densidade de probabilidade não possui forma conhecida ou é demasiadamente complicada.

Em linhas gerais, a idéia básica do método é construir uma cadeia de Markov cuja distribuição de equilíbrio é igual à distribuição de interesse. Após um número finito de simulações desta cadeia, esperamos atingir a distribuição de equilíbrio e, a partir deste ponto, dar origem a uma amostra da distribuição de interesse.

Considere que  $\theta_1, \dots, \theta_p$  possuem densidade conjunta  $p(\boldsymbol{\theta}) = p(\theta_1, \dots, \theta_p)$  e que  $q(\theta, \theta^*)$  define a distribuição condicional das transições do estado  $\theta$ .

Assim, é possível contruir uma cadeia com probabilidades de transição invariantes no tempo, onde cada estado pode ser obtido de um outro estado com um número finito de iterações. Assim, independentemente do estágio inicial, uma trajetória pode ser gerada e, conseqüentemente, pode-se alcançar a distribuição de equilíbrio  $p(\boldsymbol{\theta})$ .

Dentre os métodos mais utilizados para a construção da cadeia de Markov, temos o Amostrador de Gibbs, proposto por Geman e Geman (1984) e popularizado por Gelfand e Smith (1990), e o método de Metropolis-Hastings, proposto por Metropolis *et al.* (1953) e Hastings (1970). Gamerman (1997) apresenta uma descrição detalhada dos métodos de simulação baseados nos métodos de Monte Carlo via Cadeias de Markov.

## Amostrador de Gibbs

O amostrador de Gibbs é um esquema iterativo de amostragem de uma cadeia de Markov, cujas probabilidades de transição dos estados são realizadas a partir das distribuições condicionais completas.

Considere  $p_i(\theta_i|\boldsymbol{\theta}_{-i})$  como a função densidade condicional de  $\theta_i$ , dados os valores de todos os outros  $\theta_j$ 's ( $j \neq i$ ), e assumamos que é possível amostrar valores destas distribuições para cada  $i = 1, \dots, p$ .

A partir disto, especificamos valores iniciais  $\boldsymbol{\theta}^0 = (\theta_1^0, \dots, \theta_p^0)$  para todas as quantidades desconhecidas do modelo. Assim, na  $j$ -ésima iteração a cadeia se encontra no estado  $\boldsymbol{\theta}^{(j)}$  e a posição da cadeia na iteração seguinte ( $j + 1$ ) é obtida a partir dos seguintes passos:

- gerar  $\theta_1^{(j+1)}$  de  $p_1(\theta_1|\theta_2^{(j)}, \dots, \theta_p^{(j)})$ ;
- gerar  $\theta_2^{(j+1)}$  de  $p_2(\theta_2|\theta_1^{(j+1)}, \theta_3^{(j)}, \dots, \theta_p^{(j)})$ ;
- repetir sucessivamente os passos anteriores para  $i = 3, \dots, p$ , onde no último passo geramos  $\theta_p^{(j+1)}$  de

$$p_p(\theta_p|\theta_1^{(j+1)}, \theta_2^{(j+1)}, \dots, \theta_{p-1}^{(j+1)})$$

obtendo-se então o vetor  $\boldsymbol{\theta}^{(j+1)} = (\theta_1^{(j+1)}, \dots, \theta_p^{(j+1)})$ .

Sob certas condições de regularidade, a distribuição limite de  $\boldsymbol{\theta}^{(j)}$  tende a  $p(\boldsymbol{\theta})$ .

## Metropolis-Hastings

Outro método bastante utilizado para a construção da cadeia é o algoritmo de Metropolis-Hastings, que é baseado na utilização de distribuições auxiliares.

Considere  $q(\boldsymbol{\theta}, \cdot)$  como um núcleo arbitrário de transição e assumamos que na iteração  $(j)$  a cadeia se encontra no estado  $\boldsymbol{\theta}^{(j)}$ . Assim, a posição da cadeia na iteração  $(j + 1)$  será denotada por  $\boldsymbol{\theta}^{(j+1)}$  e é obtida a partir dos seguintes passos:

- proponha um movimento da cadeia para o estado  $\boldsymbol{\theta}^*$  a partir de  $q(\boldsymbol{\theta}^{(j)}, \cdot)$ ;
- aceitar o movimento proposto com probabilidade

$$\alpha(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^*) = \min \left\{ 1, \frac{p(\boldsymbol{\theta}^*)/q(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}^{(j)})/q(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(j)})} \right\}$$

e fazer  $\boldsymbol{\theta}^{(j+1)} = \boldsymbol{\theta}^*$  ou rejeitar o movimento com probabilidade  $1 - \alpha(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^*)$  e fazer  $\boldsymbol{\theta}^{(j+1)} = \boldsymbol{\theta}^{(j)}$  caso contrário.

Apesar do movimento da cadeia se fazer em blocos para todos os parâmetros do modelo, na prática é muito difícil encontrar funções  $q(\boldsymbol{\theta}, \cdot)$  apropriadas quando trabalha-se com modelos de alta dimensionalidade. Uma variação do algoritmo resolve este problema movendo uma componente do vetor de parâmetros  $\boldsymbol{\theta}$  por vez. Assim, forma-se uma transição completa após o ciclo de  $p$  transições a partir de funções  $q(\theta_i, \cdot)$ , para  $i = 1, \dots, p$ .

Após decidido o método a ser utilizado e obtida uma simulação da cadeia, devemos nos certificar se a convergência foi obtida. Somente após esta confirmação poderemos formar a amostra da distribuição a posteriori das quantidades desconhecidas do modelo.

Existem várias formas de se realizar uma análise a respeito da convergência da cadeia. Uma das formas é baseada em uma inspeção gráfica, onde analisa-se a trajetória de uma ou mais cadeias em períodos distintos de tempo. Neste estudo será utilizado um critério semelhante a este, onde o comportamento de duas cadeias com valores iniciais distintos será observado e concluir-se-á que a convergência foi alcançada quando ambas as cadeias permanecerem em torno de um mesmo ponto. Critérios mais formais também podem ser utilizados, como os métodos propostos por Gelman e Rubin (1992) e Geweke (1992).

Após a obtenção da amostra, deve-se analisar a autocorrelação existente entre  $\theta^{(j)}$  e  $\theta^{(j+1)}$ . Como estamos lidando com uma amostra de uma cadeia de Markov, temos uma amostra aleatória, mas não independente. Isto não afeta as estimativas dos parâmetros, mas tem influência sobre as variâncias das estimativas resultantes desse procedimento de amostragem (Gamerman, 1997). Portanto, nos casos em que, após verificada a convergência, for constatada uma forte correlação serial na cadeia, recomenda-se a retirada de uma amostra sistemática de seus valores para compor uma nova amostra. A forma como a amostragem sistemática será realizada pode ser baseada em um gráfico contendo a função de autocorrelação da cadeia.

Após a verificação destes itens, a inferência tem prosseguimento a partir do método de Monte Carlo. A idéia básica deste método consiste em estimar o valor de uma integral a partir da obtenção de seu valor esperado com relação a alguma distribuição de probabilidade. Para se obter, por exemplo, uma estimativa para o valor esperado a posteriori de um parâmetro  $\theta$  do modelo, basta tomar a média das  $i$ -ésimas componentes dos valores amostrados, ou

seja,

$$\hat{\theta} = \frac{\sum_j \theta_j}{n}.$$

## 2.4.2 Software WinBugs

O pacote estatístico **WinBugs**<sup>1</sup>, versão em ambiente **Windows** do pacote **BUGS** (Bayesian Using Gibbs Sampling), implementado por Thomas *et al.* (1992), consiste de um conjunto de procedimentos que permite a especificação de diversos tipos de modelos, estimando seus parâmetros via MCMC.

O amostrador de Gibbs implementado no **WinBugs** utiliza duas rotinas diferentes, a saber: **ARS** (Adaptive Rejection Sampling) e **ARMS** (Adaptive Rejection Metropolis Sampling). A primeira pode ser usada para amostrar de forma eficiente qualquer distribuição condicional cuja função densidade é log-côncava. A segunda generaliza a rotina **ARS** para o caso de funções que não são log-côncavas.

No caso do algoritmo Metropolis-Hastings, o software utiliza como densidade de transição  $q(\theta^{(j)}, \cdot)$  uma distribuição gaussiana centrada no valor atual do parâmetro  $\theta^{(j)}$ .

Todo o processo inferencial utilizado neste trabalho foi baseado nos métodos de Monte Carlo via Cadeias de Markov e foram implementados no software **WinBugs versão 1.4**.

---

<sup>1</sup>Disponível através do website <http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml>

# Capítulo 3

## Modelos Lineares

Existem várias formas de analisar-se a série de casos de uma certa doença. Se desejamos verificar o efeito de covariáveis na série, podemos utilizar as idéias dos modelos lineares para estimar os parâmetros relacionados com cada efeito e realizar previsões.

A classe dos modelos lineares tem por objetivo analisar a influência de covariáveis em uma determinada variável  $\mathbf{Y}$  através de uma relação linear. Em linhas gerais, estamos interessados em estudar  $\mathbf{Y}$  de forma condicional aos valores de  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_p)$ , ou seja,

$$E(\mathbf{Y}|\mathbf{X}) = \beta_1 \mathbf{X}_1 + \dots + \beta_p \mathbf{X}_p.$$

O modelo descrito acima é conhecido como um modelo linear ou modelo linear de regressão (Migon e Gamerman, 1999). Supondo uma amostra  $Y_1, \dots, Y_n$  independente e normalmente distribuída, temos que a função de verossimilhança é dada por:

$$p(Y_1, \dots, Y_n | \boldsymbol{\beta}, \sigma^2) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n e_i^2 \right\}$$



onde  $e_i = Y_i - (\beta_1 x_{1i} + \dots + \beta_p x_{pi})$ ,  $i = 1, \dots, n$ .

Se desejamos estimar os valores de  $\boldsymbol{\beta}$  por máxima verossimilhança, devemos minimizar o termo  $\sum_{i=1}^n e_i^2$ . Seguindo a notação matricial, temos que:

$$p(\mathbf{Y}|\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-n} \exp \left\{ -\frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2} \right\}$$

onde  $\mathbf{Y}' = (Y_1, \dots, Y_n)$ ,  $\mathbf{X}' = (\mathbf{x}'_1, \dots, \mathbf{x}'_n)$  e  $\boldsymbol{\beta}' = (\beta_1, \dots, \beta_p)$ . Assim, temos que o estimador de máxima verossimilhança de  $\boldsymbol{\beta}$  é dado por

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

Na abordagem bayesiana, a utilização dos modelos lineares deve ser completada com a especificação de distribuições a priori para os parâmetros do modelo.

Em geral, especificam-se distribuições a priori  $\text{Normal}(\mathbf{A}, \phi^{-1}\mathbf{B})$  para o vetor  $\boldsymbol{\beta}$  e distribuições  $\text{Gama}(c, d)$  para o parâmetro  $\phi = \sigma^{-2}$ . Desta forma, obtém-se distribuições a posteriori T-Student e Gama para  $\boldsymbol{\beta}|\mathbf{Y}$  e  $\phi|\mathbf{Y}$ , respectivamente.

### 3.1 Modelos Lineares Generalizados

A modelagem do número de casos de dengue em um determinado período através de um modelo linear nem sempre é recomendada, uma vez que estes modelos requerem a suposição de normalidade da série. Pelo fato da série de casos da doença ser discreta, deve-se buscar uma nova forma de relacionamento linear entre as covariáveis e a série de dados.

Uma extensão dos modelos lineares permite modelar observações descritas por membros da família exponencial. Esta classe de modelos é conhe-

cida como *Modelos Lineares Generalizados* (MLG). Uma descrição detalhada destes modelos pode ser encontrada em McCullag e Nelder (1989).

Por definição, uma família de distribuições com densidade  $p(x|\theta)$  pertence a família exponencial se sua densidade pode ser escrita como

$$p(x|\theta, V) = a(x, V) \exp \left\{ \frac{f(x)\theta - b(\theta)}{V} \right\}, x \in \chi \subset \Re \quad (3.1)$$

onde  $\theta$  é o parâmetro natural da distribuição,  $V$  é o parâmetro associado à precisão da distribuição e  $\chi$  não depende de  $\theta$ .

Outros aspectos importantes desta família são que a função  $b(\theta)$  é duas vezes diferenciável em  $\theta$  e que

$$\mu = E[f(x)|\theta] = b'(\theta).$$

Assim,  $\mu$  e o parâmetro  $\theta$  se relacionam através da transformação inversa

$$\theta = b'^{-1}(\mu).$$

No contexto epidemiológico, modelar o número de casos mensais  $y_t$  de uma doença conforme uma distribuição de Poisson seria uma escolha natural.

Neste caso teríamos

$$p(y_t|\lambda_t) = \frac{e^{-\lambda_t} \lambda_t^{y_t}}{y_t!} = \frac{1}{y_t!} \exp \{-\lambda_t + y_t \log(\lambda_t)\}, y_t = 0, 1, 2, \dots$$

onde  $\theta_t = \log(\lambda_t)$ ,  $V = 1$ ,  $b(\theta) = \lambda$ ,  $a(y_t, V) = \frac{1}{y_t!}$  e  $f(\cdot)$  é a função identidade.

Neste caso, a função de ligação entre a média da série  $\lambda$  e o parâmetro  $\theta$  seria a função  $\exp(\cdot)$ . Assim, o preditor linear seria da seguinte forma:

$$\theta_t = \log(\lambda_t) = \mathbf{X}'_t \boldsymbol{\beta}.$$

Se fosse utilizada uma abordagem clássica de inferência, os parâmetros  $\beta$ 's poderiam ser estimados por máxima verossimilhança e testes seriam aplicados para sua significância e seriam preditos os valores da série.

Em contrapartida, na abordagem bayesiana, são especificadas distribuições a priori para o vetor de parâmetros  $\beta$  e faz-se a inferência através da análise da distribuição a posteriori  $p(\beta|y_t)$ , obtida através da atualização da informação via Teorema de Bayes. O processo de predição decorre naturalmente através da análise da distribuição

$$p(y^*|y_t) = \int \dots \int p(y_t|\beta)p(\beta|y_t)d\beta_1\dots d\beta_p,$$

onde leva-se em conta na predição a incerteza a respeito dos parâmetros do modelo.

## 3.2 Modelos Lineares Generalizados Dinâmicos

A utilização de modelos pertencentes a classe dos modelos lineares generalizados para a modelagem de séries compostas por contagens pode não ser suficiente para captar certas estruturas presentes nos dados. Isto acontece porque, em geral, séries epidemiológicas são dependentes no tempo.

Uma das formas de contornar este problema é incluir no modelo covariáveis que representem características temporais, tais como sazonalidade e tendência.

Outra forma pode ser a utilização de Modelos Lineares Dinâmicos (MLD), que atualizam os parâmetros do modelo de forma sequencial com base nas observações passadas.

### 3.2.1 Modelos Lineares Dinâmicos

A classe dos Modelos Lineares Dinâmicos (MLD) é muito utilizada para a análise de séries temporais. Por ser uma classe ampla, ela permite a modelagem de dados com as mais variadas características e restrições. Uma análise detalhada desta classe de modelos é apresentada em West e Harrison (1997).

De uma forma geral, os modelos lineares dinâmicos são modelos onde a variação do parâmetro no tempo e os dados disponíveis são descritos probabilisticamente (Migon e Gamerman, 1999). Estes modelos são caracterizados pela *equação observacional* (3.2) e pela *equação de sistema* (3.3), que apresentam a seguinte forma:

$$Y_t = \mathbf{x}'_t \boldsymbol{\beta}_t + \epsilon_t, \epsilon_t \sim Normal(0, \sigma_t^2) \quad (3.2)$$

$$\boldsymbol{\beta}_t = \mathbf{G}_t \boldsymbol{\beta}_{t-1} + \mathbf{w}_t, \mathbf{w}_t \sim Normal(\mathbf{0}, \mathbf{W}_t) \quad (3.3)$$

onde  $Y_t$  representa a série,  $\mathbf{x}'_t$  um vetor de valores conhecidos,  $\boldsymbol{\beta}_t$  os parâmetros do modelo,  $\mathbf{G}_t$  uma matriz representando a evolução de cada parâmetro no tempo,  $\sigma_t^2$  o erro observacional e  $\mathbf{W}_t$  a variância do vetor de parâmetros.

Assim, esta classe de modelos pode ser especificada através da quádrupla

$$\{\mathbf{x}_t, \mathbf{G}_t, \sigma_t^2, \mathbf{W}_t\}.$$

Analisando as equações que caracterizam um modelo linear dinâmico, podemos notar que a distribuição de  $Y_t$  é independente das observações passadas dado o conhecimento de  $\boldsymbol{\beta}_t$ . Isto significa que a dinâmica temporal é sumarizada no processo de evolução dos parâmetros.

Utilizando-se uma abordagem bayesiana, podemos realizar previsões e inferências para os parâmetros através da utilização da informação atualizada

no tempo  $t - 1$ . Assim, a distribuição a posteriori em  $t - 1$  é utilizada pelo modelo como distribuição a priori em  $t$ .

A título de ilustração do processo de inferência, considere o MLD caracterizado pela quádrupla  $\{1, 1, \sigma_t^2, W_t\}$ . Suponha também que  $D_{t-1}$  represente toda a informação do passado disponível até o instante  $t - 1$ . Desta forma, o processo de inferência seguirá os passos:

(i) Posteriori em  $t - 1$ :

$$p(\beta_{t-1}|D_{t-1})$$

(ii) Priori em  $t$ :

$$p(\beta_t|D_{t-1}) = \int p(\beta_t|\beta_{t-1})p(\beta_{t-1}|D_{t-1})d\beta_{t-1}$$

(iii) Preditiva:

$$p(Y_t|D_{t-1}) = \int p(Y_t|\beta_t)p(\beta_t|D_{t-1})d\beta_t$$

(iv) Posteriori em  $t$ :

$$p(\beta_t|D_t) \propto p(Y_t|\beta_t)p(\beta_t|D_{t-1})$$

A utilização da classe dos Modelos Lineares Dinâmicos em séries temporais permite que mudanças bruscas sejam suavizadas além de permitir a acomodação natural de informações subjetivas (Ferreira e Gamerman, 1998). Porém, a utilização destes modelos no contexto de séries epidemiológicas nem sempre é possível, já que estes dados, em geral, são constituídos de valores discretos.

### 3.2.2 O Modelo Linear Generalizado Dinâmico

West *et al.* (1985) formalizaram uma extensão dos Modelos Lineares Dinâmicos para dados que pertençam à família exponencial, tornando possível a utilização destes modelos para os mais variados tipos de problemas, como a análise de séries epidemiológicas.

A classe dos Modelos Lineares Generalizados Dinâmicos é caracterizada pelas seguintes componentes:

- Modelo observacional: a equação observacional do MLD é substituída pelo modelo observacional conforme apresentado em (3.1);
- Função de Ligação: relaciona a média da série com os parâmetros do modelo

$$g(\boldsymbol{\mu}_t) = \boldsymbol{\theta}_t = \mathbf{x}'_t \boldsymbol{\beta}_t,$$

onde

$$\boldsymbol{\mu}_t = E[\mathbf{Y}_t | \boldsymbol{\theta}_t] = b'(\boldsymbol{\theta}_t).$$

- Equação de evolução:

$$\boldsymbol{\beta}_t = \mathbf{G}_t \boldsymbol{\beta}_{t-1} + \mathbf{w}_t, \mathbf{w}_t \sim Normal(\mathbf{0}, \mathbf{W}_t)$$

O processo de inferência é equivalente ao modo apresentado para o MLD. Entretanto, este processo é prejudicado pois o cálculo das integrais necessárias para a atualização dos parâmetros não pode ser obtido exatamente. Assim, vários métodos para lidar com este problema foram propostos na literatura, tais como os métodos de MCMC. Ferreira e Gamerman (2000) apresentam os princípios de alguns destes métodos.

### 3.3 Modelos Hierárquicos

Assim como podemos especificar distribuições a priori para os parâmetros de um modelo em estágios, podemos também utilizar esta estratégia no processo de modelagem de fenômenos. No caso dos modelos lineares, a utilização de estruturas hierárquicas pode ser muito bem-vinda.

Para exemplificar uma modelagem em dois estágios, suponha que  $\mathbf{Y}$  é uma variável de interesse e que desejamos analisá-la através de um modelo linear. Desta forma, temos que

$$\mathbf{Y}|\boldsymbol{\beta}, \phi \sim Normal(\mathbf{X}\boldsymbol{\beta}, \phi^{-1}I),$$

onde  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$  é o vetor de parâmetros e  $\mathbf{X}$  é uma matriz de covariáveis.

Supondo que exista uma matriz  $\mathbf{X}^*$  que contenha covariáveis que explicam o vetor de parâmetros  $\boldsymbol{\beta}$ , podemos incluir o seguinte estágio no modelo:

$$\boldsymbol{\beta}|\boldsymbol{\beta}^*, \phi \sim Normal(\mathbf{X}^*\boldsymbol{\beta}^*, \phi^{-1}C_1),$$

onde  $\boldsymbol{\beta}^*$  é um vetor de parâmetros associado às covariáveis da matriz  $\mathbf{X}^*$ . Completando o modelo, temos

$$\boldsymbol{\beta}^*|\phi \sim Normal(\boldsymbol{\mu}, \phi^{-1}C_2)$$

$$\phi \sim Gama(a, b).$$

Neste caso, preservamos a linearidade do modelo e a inferência é realizada de forma simples, uma vez que esta modelagem produz distribuições a posteriori conhecidas. Dependendo do problema a ser analisado, mais estágios podem ser adicionados no modelo. Entretanto, quando utilizados muitos estágios, torna-se mais difícil interpretar os parâmetros envolvidos. Nestes casos, é comum especificar distribuições a priori não-informativas para os parâmetros presentes nos níveis mais elevados do modelo (Migon e Gamerman, 1999).

# Capítulo 4

## Análise Temporal

Análises da série de dengue podem ser úteis para a identificação de padrões e, conseqüentemente, contribuir com as autoridades sanitárias no que diz respeito à identificação de épocas mais propícias para a realização de campanhas junto à população e nas medidas de combate ao mosquito transmissor da doença.

Dentre os possíveis benefícios destes tipos de estudos, podemos destacar a questão da sazonalidade de notificações do dengue. Apesar de não ser difícil notar que o número de casos de dengue é mais acentuado em algumas épocas do ano, uma identificação da magnitude do impacto de cada período no aumento ou diminuição da ocorrência da doença pode ser de grande ajuda para as autoridades sanitárias.

Além disso, a influência de tendências e de possíveis fatores aleatórios, ou de origem desconhecida, podem possuir papéis fundamentais para o entendimento do comportamento da série.



O princípio básico empregado nesta análise consiste em utilizar um modelo Poisson para o número de casos mensais de dengue, onde a média da distribuição é relacionada com efeitos sazonais e outros efeitos aleatórios.

O capítulo está organizado da seguinte forma: Na seção 4.1 apresentamos a série de notificações mensais de dengue que será analisada. Na seção 4.2 apresentamos os dois modelos propostos para analisar a série. A seção 4.3 apresenta e compara os resultados obtidos com a utilização de ambos os modelos. Por fim, a seção 4.4 apresenta as conclusões da análise.

## 4.1 Dados Utilizados

Serão considerados nesta análise as notificações mensais de dengue ocorridas na cidade do Rio de Janeiro entre os anos de 1986 e 2002. Os dados foram obtidos junto à Secretaria Municipal de Saúde do Rio de Janeiro (SMS-RJ).

A série de dados possui grandes picos nos períodos em que ocorreram as epidemias de 1986-1987, 1990-1991 e 2001-2002 (ver figura 4.1). Apesar de não compreenderem períodos epidêmicos, os anos de 1995, 1996, 1998 e 1999 também apresentaram elevados índices de notificações de dengue. A figura 4.2 apresenta o logaritmo do número de casos, onde podemos observar o comportamento da série nos períodos entre as epidemias.

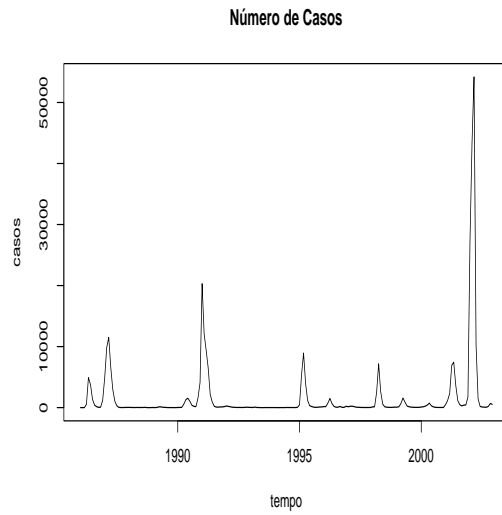


Figura 4.1: Notificações mensais de dengue no Rio de Janeiro (1986-2002).

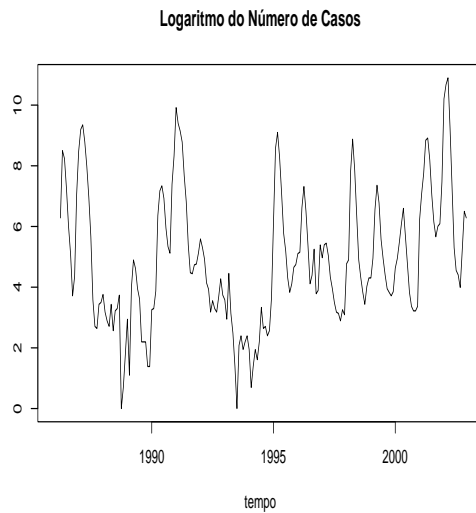


Figura 4.2: Logaritmo da série de casos de dengue (1986-2002).

Analisando-se a distribuição dos casos por mês (ver figura 4.3), a fim de identificar qual o padrão sazonal da série, podemos visualizar uma maior ocorrência de casos nos primeiros meses do ano. Destes, os meses de Março, Abril e Maio foram os que apresentaram os índices mais elevados. Entretanto, deve-se salientar a grande variabilidade existente entre os casos dentro de um mesmo mês.

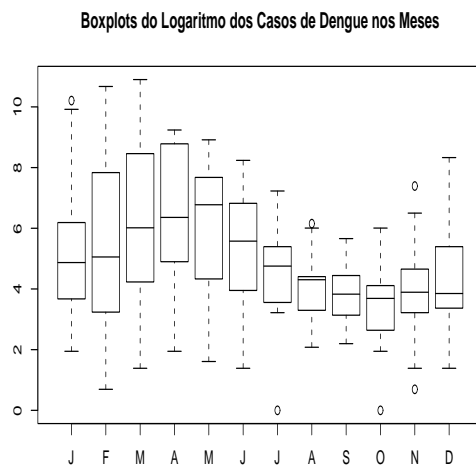


Figura 4.3: Boxplot's com as notificações de dengue ocorridas em cada mês durante o período 1986-2002.

## 4.2 Modelo Proposto

Para a modelagem da série  $y_t = \text{Número de casos de dengue no mês } t$ ,  $t = 1, \dots, 204$ , vamos utilizar um modelo Poisson, cuja média será relacionada com fatores que caracterizem efeitos temporais através da função de ligação  $\log(\cdot)$  de um Modelo Linear Generalizado *log-linear* (McCullagh e Nelder, 1989).

Desta forma, o modelo básico é dado por:

$$y_t | \lambda_t \sim \text{Poisson}(\lambda_t) \quad t = 1, \dots, 204.$$

O parâmetro  $\lambda_t$ , por sua vez, é relacionado com efeitos temporais conforme descrito a seguir:

$$\log \lambda_t = \mu + A_{[t/12]} + S_t, \quad (4.1)$$

onde  $\mu$  representa o nível médio,  $A_{[t/12]}$  representa o efeito anual e  $S_t$  representa o efeito sazonal. Estes efeitos, por sua vez, são modelados da seguinte forma:

- Efeitos anuais:

$$\sum_j^{17} A_j = 0$$

$$A_j \sim \text{Normal}(0, 10^3), \quad j = 1, \dots, 16$$

onde  $j$  representa o ano relacionado à cada observação. Os efeitos dos 17 anos da série foram incorporados no modelo a fim de capturar-se efeitos de tendência anual e para uma possível identificação de efeitos de epidemias.

- Efeitos Sazonais:

$$S_t = - \sum_{i=1}^{11} S_{t-i} + w_t, \quad t = 12, \dots, 204$$

$$S_t \sim \text{Normal}(0, \tau^{-2}), \quad t = 1, \dots, 11$$

$$w_t \sim \text{Normal}(0, \tau_w^{-2}), \quad t = 12, \dots, 204$$

$$\tau^2 \sim \tau_w^2 \sim \text{Gamma}(0,05; 0,005)$$

Os efeitos sazonais representam a influência de cada mês no número de casos notificados de dengue, onde o efeito  $w_t$  é responsável por capturar a variabilidade entre os efeitos sazonais a cada ciclo de 12 meses.

As distribuições a priori para os parâmetros  $\tau^2$  e  $\tau_w^2$  possuem grande variância, sendo consideradas como prioris não-informativas.

- Efeito Médio:

$$\mu \sim Normal(0, 10^3)$$

O efeito médio indica o nível médio geral da série, considerando todos os 204 meses. A grande variabilidade da distribuição de  $\mu$  indica que não estamos agregando informações a respeito do nível da série a priori.

Sabemos que possíveis variações não esperadas pelo modelo Poisson podem estar presentes nos dados. Estes efeitos podem ser causados por observações discrepantes ou simplesmente pela não incorporação de alguma covariável importante no modelo. Este tipo de efeito é conhecido como efeito de *sobredispersão*.

Se esta variação se faz presente, todo o processo de inferência pode ficar prejudicado. Para contornar isto, adicionamos ao modelo um efeito aleatório

$$\phi_t \sim Normal(0, \tau_\phi^{-2}), \quad t = 1, \dots, 204$$

$$\tau_\phi^2 \sim Gama(0.05, 0.005)$$

a fim de captar toda a variação não explicada pelas covariáveis do modelo.

Assim, o modelo descrito em (4.1) é alterado para

$$\log \lambda_t = \mu + A_{[t/12]} + S_t + \phi_t. \quad (4.2)$$

Desta forma, a função de verossimilhança da série fica da seguinte forma:

$$p(\mathbf{y}|\mu, \mathbf{A}, \mathbf{S}, \phi) = \prod_{t=1}^{204} \frac{\exp\{\mu + A_{[t/12]} + S_t + \phi_t\}^{y_t} \exp\{-\exp(\mu + A_{[t/12]} + S_t + \phi_t)\}}{y_t!}.$$

Para ilustrar o procedimento de inferência relacionado à utilização deste modelo, considere o vetor  $\boldsymbol{\theta} = (\mu, \mathbf{A}, \mathbf{S}, \boldsymbol{\phi}, \tau^2, \tau_w^2, \tau_\phi^2)$  como aquele que contém todas as quantidades desconhecidas do modelo. Assim, temos que a distribuição a posteriori de  $\boldsymbol{\theta}$  é proporcional a:

$$\begin{aligned} p(\boldsymbol{\theta}|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \\ &\propto p(\mathbf{y}|\mu, \mathbf{A}, \mathbf{S}, \boldsymbol{\phi}, \tau^2, \tau_w^2, \tau_\phi^2)p(\mu, \mathbf{A}, \mathbf{S}, \boldsymbol{\phi}, \tau^2, \tau_w^2, \tau_\phi^2) \\ &\propto p(\mu)p(\mathbf{A})p(\mathbf{S}|\tau^2, \tau_w^2)p(\tau^2)p(\tau_w^2)p(\boldsymbol{\phi}|\tau_\phi^2)p(\tau_\phi^2). \end{aligned}$$

Como esta distribuição não possui forma conhecida, propusemos a utilização de algoritmos baseados no método de Monte Carlo via Cadeias de Markov (MCMC). Entretanto, em virtude das distribuições condicionais completas para a maior parte dos parâmetros também não possuem formas conhecidas, é necessário a utilização de outros métodos de amostragem. Neste caso, foi utilizado o software `WinBugs versão 1.4` para a obtenção das amostras das distribuições a posteriori dos parâmetros.

### 4.2.1 Comparação entre os Modelos

Os modelos com e sem sobredispesão serão comparados a fim de que seja possível verificar a relevância da incorporação de um efeito aleatório no modelo.

O critério de seleção adotado destacará o modelo que apresentar o menor valor do DIC - *Deviance Information Criterion* (Spiegelhalter *et al.*, 2002). O Apêndice apresenta uma breve descrição deste critério.

### 4.3 Resultados

Para a obtenção das amostras das distribuições a posteriori dos parâmetros dos modelos por meio do algoritmo MCMC, monitorou-se a trajetória de duas cadeias partindo de valores iniciais distintos. Entretanto, em virtude da grande autocorrelação existente entre os valores amostrados, foram incluídas na amostra da posteriori apenas 1 (um) valor a cada 1000 valores gerados. A partir disto, concluiu-se que a convergência havia sido obtida após amostrarmos 400 valores para ambos os modelos considerados. A figura 4.4 apresenta a evolução das cadeias evidenciando a convergência dos parâmetros  $\tau^2$ ,  $\tau_\phi^2$  e  $\tau_w^2$  para o modelo contendo efeito de sobredispersão.

Para compor a amostra, observou-se as cadeias por mais 1000 iterações, totalizando duas amostras de tamanho 1000 para cada parâmetro de ambos os modelos.

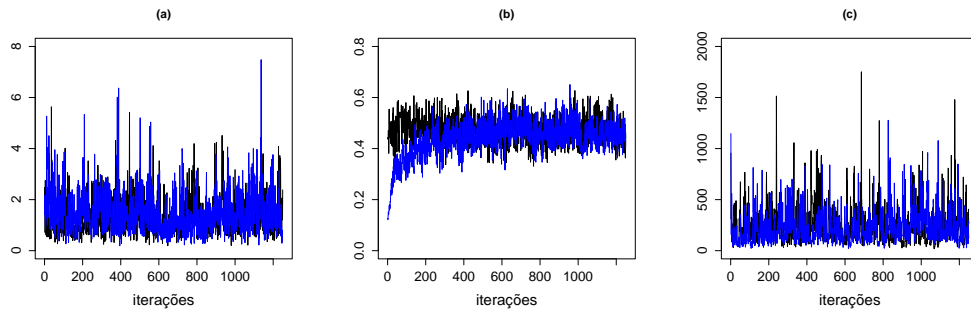


Figura 4.4: Evolução das cadeias referentes aos parâmetros  $\tau^2$  (a),  $\tau_\phi^2$  (b) e  $\tau_w^2$  (c) no modelo com sobredispersão.

### 4.3.1 Modelo sem Efeito de Sobredispersão

O ajuste do modelo sem sobredispersão é apresentado na figura 4.5 em escala logarítmica.

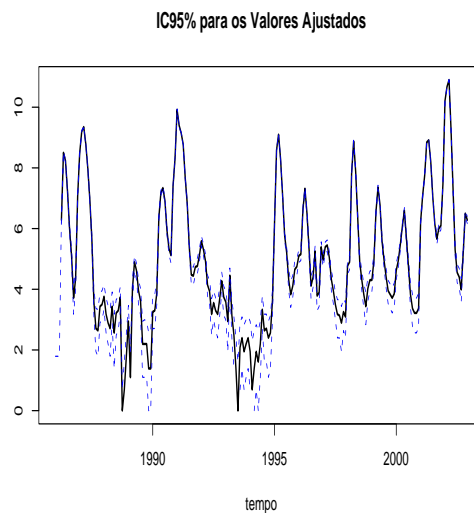


Figura 4.5: IC95% para os valores ajustados pelo modelo na escala logarítmica (linha tracejada) em comparação com a série original.

A tabela 4.1 apresenta um sumário das distribuições a posteriori dos parâmetros do modelo.

Os efeitos anuais estimados pelo modelo sem sobredispersão apresentaram valores elevados para os anos epidêmicos e valores negativos para os demais anos (ver figura 4.6). Para todos os anos, os intervalos de 95% de credibilidade a posteriori evidenciaram a significância dos efeitos.

Os efeitos sazonais estimados pelo modelo não apresentaram um padrão claro. Apesar dos primeiros meses do ano apresentarem efeitos maiores do que os efeitos dos demais meses, encontramos dificuldade em fazer uma com-



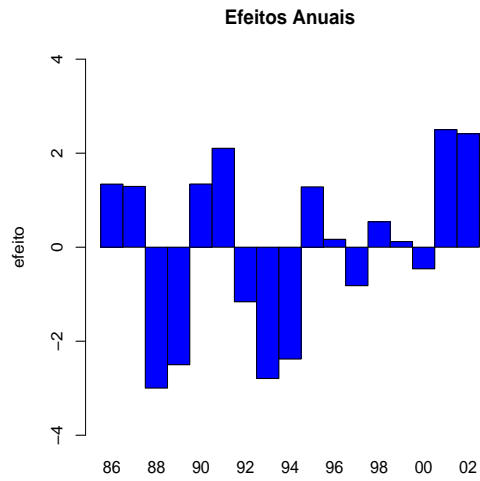


Figura 4.6: Média a posteriori dos efeitos anuais.

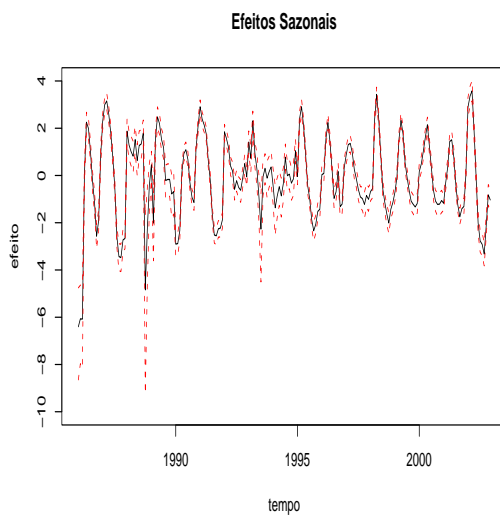


Figura 4.7: Média a posteriori e respectivos IC95% dos efeitos sazonais (linha tracejada).

Tabela 4.1: Sumário da distribuição a posteriori dos parâmetros do modelo sem sobredispersão.

Parâmetro	Média	Desvio Padrão	IC95%
$\mu$	4,905	0,036	(4,832 ; 4,970)
$\tau^2$	0,08192	0,03852	(0,02694 ; 0,1784)
$\tau_w^2$	0,0402	0,0046	(0,0317 ; 0,0496)

paração mês a mês.

Contudo, apesar deste padrão não ser muito claro, pode-se notar que ele foi ficando mais evidente ao longo da série. No início da série, onde havia sido detectada apenas a entrada do sorotipo I do vírus na cidade (em 1986), o padrão sazonal se mostrava bastante confuso. À medida que os anos passaram e novos sorotipos do vírus foram detectados na cidade ao longo das grandes epidemias de 1990/1991 e 2001/2002, os indícios da formação de um padrão sazonal pareceram tornar-se mais claros (ver figura 4.7).

Isto indicou que poderiam existir outros efeitos, não incorporados ao modelo Poisson considerado, que poderiam ser responsáveis por mascarar um possível padrão sazonal ou alguma observação discrepante.

### 4.3.2 Modelo com Efeito de Sobredispersão

O ajuste do modelo com sobredispersão é apresentado na figura 4.8 em escala logarítmica. Assim como no modelo sem efeitos de sobredispersão, podemos ver que a série praticamente não foi suavizada.

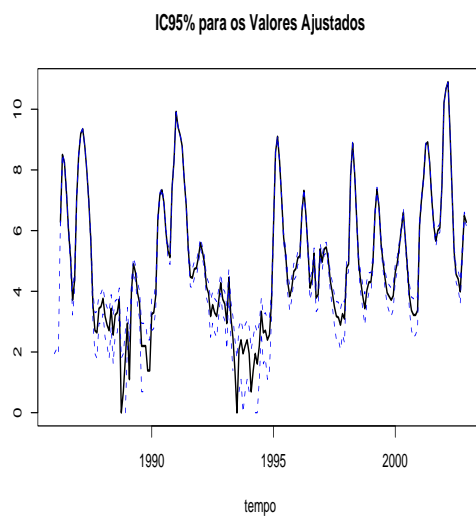


Figura 4.8: IC95% para os valores ajustados pelo modelo com sobredispersão na escala logarítmica (linha tracejada) em comparação com a série original.

Comparando-se o ajuste dos modelos com e sem sobredispersão, vemos que o modelo com sobredispersão apresentou valores ajustados mais próximos da série original do que os ajustados pelo modelo sem sobredispersão (figura 4.9).

Os efeitos anuais estimados pelo modelo com sobredispersão também apresentaram valores elevados para os anos epidêmicos e valores negativos para os demais anos (ver figura 4.10). Exceto para o ano de 1986, os intervalos de 95% de credibilidade evidenciaram a significância de todos os efeitos anuais.

Por sua vez, os efeitos sazonais estimados pelo modelo apresentaram um padrão bastante claro. A figura 4.11 apresenta as estimativas dos efeitos sazonais e seus respectivos intervalos de credibilidade, onde pode-se verificar a significância dos mesmos.

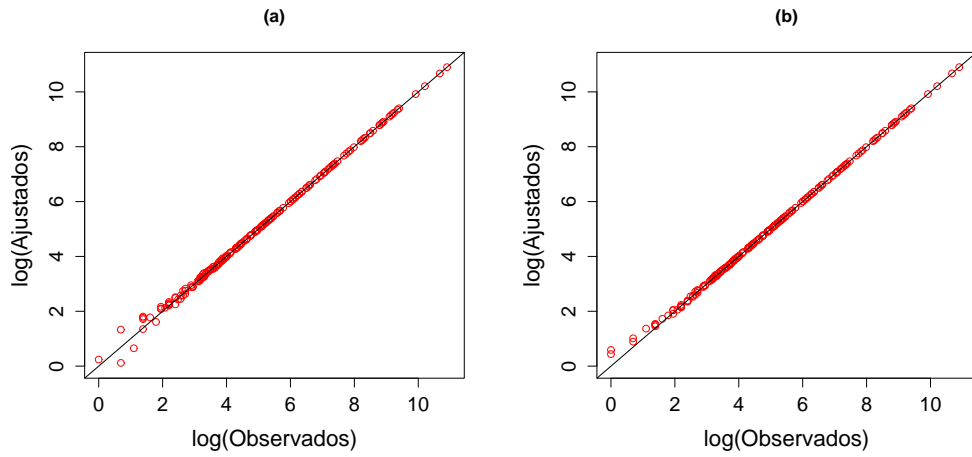


Figura 4.9: Valores observados versus valores ajustados do modelo sem efeitos de sobredispersão (a) e do modelo contendo efeito de sobredispersão (b) na escala logarítmica.

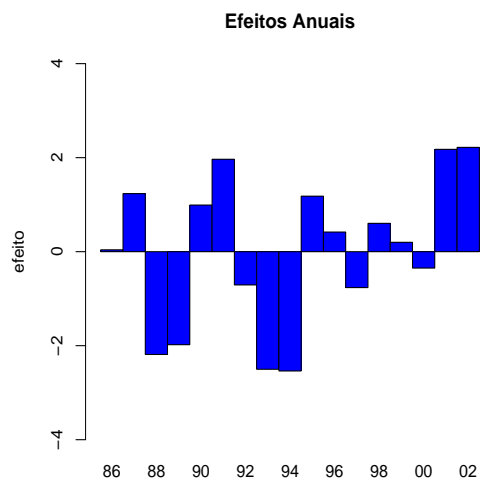


Figura 4.10: Média a posteriori dos efeitos anuais do modelo com sobredispersão.

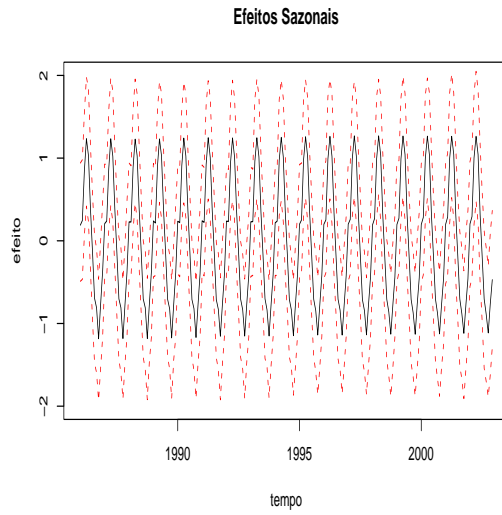


Figura 4.11: Média a posteriori e respectivos IC95% dos efeitos sazonais do modelo com sobredispersão (linha tracejada).

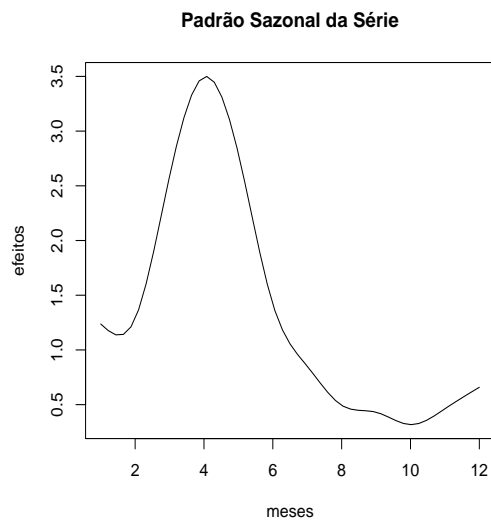


Figura 4.12: Padrão sazonal estimado a partir do modelo com sobredispersão na escala original.

Na figura 4.12 é apresentado um gráfico contendo uma ilustração do padrão sazonal da série, o qual foi obtido a partir de uma interpolação linear entre as médias mensais dos efeitos estimados para cada um dos 17 anos. Analisando o padrão destes efeitos, pode-se verificar que o mês de Abril representa o período do ano em que há um maior número de notificações de dengue. Juntamente com o mês de Abril, destacam-se os meses de Março e Maio como aqueles em que há um elevado efeito sazonal. Por outro lado, os meses de Setembro, Outubro e Novembro foram os que apresentaram os menores efeitos sazonais estimados.

Comparando os resultados obtidos pelos dois modelos, vemos que existem diferenças significativas entre as conclusões advindas de cada uma das análises.

O efeito anual não significativo para o ano de 1986, estimado pelo modelo com sobredispersão, contrasta com aquele estimado pelo modelo sem sobredispersão, cujo valor indicava 1986 como um ano de epidemia. Uma possível explicação para isto está no fato de que no início de 1986, nos meses de Janeiro, Fevereiro e Março, não havia notificações de dengue no Rio de Janeiro. Porém, com a entrada do vírus na cidade, muitos casos foram notificados e uma epidemia começou a surgir, cujo ápice ocorreu no início do ano seguinte.

Acreditamos que esta característica peculiar fez com que 1986 fosse considerado como um ano não epidêmico pelo modelo com sobredispersão. Neste caso, o modelo considerou o surgimento da epidemia como resultado de efeitos aleatórios, ao contrário do modelo sem sobredispersão, que o considerou como um efeito anual. A figura 4.13 apresenta as estimativas dos efeitos aleatórios juntamente com seus respectivos intervalos de credibilidade. Analisando esta

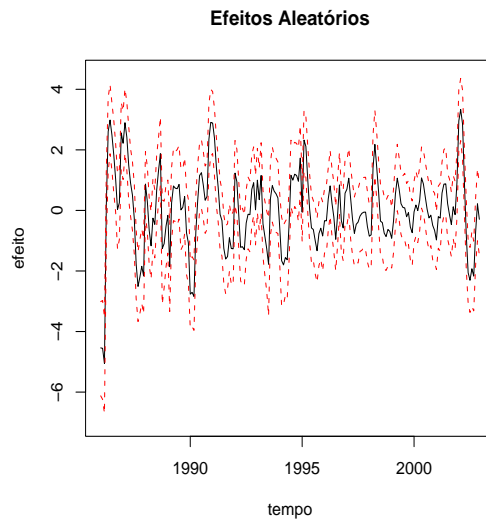


Figura 4.13: Média a posteriori e respectivos IC95% dos efeitos de sobredispersão (linha tracejada).

figura, podemos ver que grande parte dos efeitos aleatórios são significativos. Além disso, vemos que os efeitos aleatórios apresentam valores estimados muito baixos nos primeiros meses de 1986 e valores bastante elevados nos demais meses do ano.

O modelo sem sobredispersão não levava em consideração efeitos aleatórios e, portanto, teve suas estimativas dos efeitos anuais e sazonais influenciadas por eles. Por outro lado, o padrão sazonal tornou-se bastante claro após a inclusão destes efeitos no modelo.

Comparando os modelos através do DIC (tabela 4.2), vemos que o modelo com efeito de sobredispersão apresentou melhores resultados. Este fato apenas corrobora com as conclusões obtidas através da análise dos resultados apresentados pelos dois modelos.

Tabela 4.2: Comparação dos modelos através do critério DIC.

MODELO	$\bar{D}$	$\hat{D}$	$p_D$	DIC
Sem sobredispersão	1575,520	1375,300	200,221	1775,740
Com sobredispersão	1568,570	1367,860	200,713	1769,280

Tabela 4.3: Sumário da distribuição a posteriori dos parâmetros do modelo com sobredispersão.

Parâmetro	Média	Desvio Padrão	IC95%
$\mu$	4,834	0,108	(4,617 ; 5,049)
$\tau^2$	1,432	0,837	(0,401 ; 3,420)
$\tau_\phi^2$	0,466	0,053	(0,368 ; 0,580)
$\tau_w^2$	243,4	175,7	(44,01 ; 714,5)

A tabela 4.3 apresenta um sumário das distribuições a posteriori dos parâmetros do modelo. Comparando os valores apresentados nesta tabela com aqueles da tabela 4.1, pode-se notar o ganho significativo de precisão nas estimativas dos parâmetros. O valor elevado de  $\tau_w^2$ , estimado pelo modelo com sobredispersão, indicou a pouca variabilidade existente no padrão sazonal ao longo dos anos.

Os histogramas das amostras das distribuições a posteriori dos parâmetros do modelo com sobredispersão são mostrados na figura 4.14.



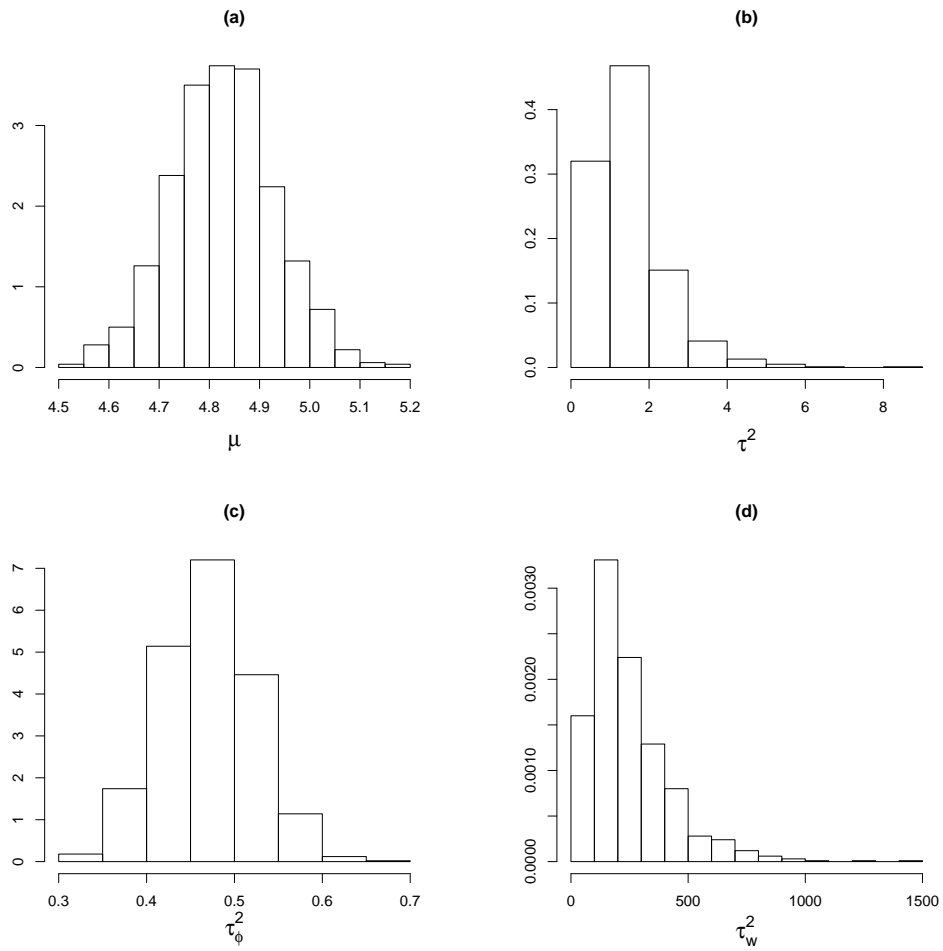


Figura 4.14: Histogramas das amostras das distribuições a posteriori dos parâmetros  $\mu$  (a),  $\tau^2$  (b),  $\tau_\phi^2$  (c) e  $\tau_w^2$  (d).

## 4.4 Conclusões

A utilização de Modelos Generalizados Bayesianos, aliada às técnicas de MCMC, permitiu a modelagem da série de notificações de dengue no Rio de Janeiro.

Através de uma abordagem bayesiana, incorporou-se no modelo fatores anuais e sazonais a fim de detectar-se efeitos cíclicos na série. Combinando a informação dos dados com as informações a priori, analisou-se a distribuição a posteriori das quantidades desconhecidas do modelo.

O modelo contendo efeitos de sobredispersão apresentou os melhores resultados segundo o DIC, além de evidenciar o padrão sazonal da série.

A identificação do padrão sazonal na série de casos de dengue no Rio de Janeiro tem uma grande importância para o planejamento de medidas de combate ao mosquito e de campanhas de prevenção entre a população.

O efeito de sobredispersão indicou a existência de efeitos significativos que causaram uma maior variabilidade do que a esperada pelo modelo Poisson. Assim, os efeitos aleatórios podem ser interpretados como provenientes do aumento de casos de dengue no início de epidemias e das quedas bruscas ocorridas ao seu final. Grande parte dos efeitos aleatórios mais elevados ocorreu em épocas de elevados índices de notificações de casos de dengue.

De uma forma geral, estes resultados ressaltam a importância da monitoração das infecções por dengue na comunidade durante os períodos inter-epidêmicos ou de transmissão esporádica ou silenciosa, conforme descrito em Medronho (1995). Quando se introduz um novo sorotipo do vírus, existe em geral um baixo nível de transmissão que pode durar de poucas sema-

nas a vários meses, antes que se inicie a transmissão epidêmica. O objetivo da vigilância ativa é detectar o vírus durante este período, muito antes do aumento da transmissão.

Isto mostra mais uma vantagem da incorporação de efeitos de sobre-dispersão em séries contendo grande variabilidade causada por fatores não explicados por efeitos cíclicos em modelos Poisson.

# Capítulo 5

## Análise Espacial

Neste capítulo, será analisada a distribuição espacial dos casos de dengue na cidade do Rio de Janeiro, a partir de notificações semanais de casos efetuadas durante os anos de 2001 e 2002. Aspectos sócio-econômicos e ambientais foram utilizados em busca de possíveis relações com a distribuição dos casos de dengue.

A análise estatística empregada utiliza a abordagem bayesiana, onde podemos combinar a informação dada pela verossimilhança com informações a priori e obter a distribuição dos riscos de dengue a posteriori.

Os modelos utilizados estão baseados nas idéias de utilização de Campos Aleatórios Markovianos no mapeamento de doenças (Mollié, 1996) e Modelos Lineares Generalizados Bayesianos.

Na análise empregada neste estudo, o número de casos de dengue é modelado segundo uma distribuição de Poisson, onde um fator de risco é incorporado na composição do valor esperado de casos para um determinado

bairro.

Através desta análise, buscamos identificar fatores sociais, geográficos ou econômicos que estejam relacionados com o risco de dengue de uma determinada área, além de avaliar a influência de diferentes especificações da distribuição a priori para efeitos espaciais embutidos nos modelos de risco da doença.

O capítulo está organizado da seguinte forma: Na seção 5.1 são apresentados os dados de dengue bem como as covariáveis utilizadas. Na seção 5.2 é apresentado o modelo e a forma como a inferência será realizada. Na seção 5.3 são apresentados os resultados obtidos. Finalmente, na seção 5.4 é discutida a relevância dos resultados obtidos e aponta-se para novos estudos sobre a distribuição dos riscos de dengue.

## 5.1 Dados Utilizados

Os dados que serão analisados aqui correspondem às 125.368 notificações de casos de dengue na cidade do Rio de Janeiro ocorridas entre a quadragésima oitava semana epidemiológica <sup>1</sup> de 2001 (início de dezembro) e a vigésima semana epidemiológica de 2002 (maio). Este período, que correspondeu à época de verão na cidade, ficou marcado pela ocorrência da maior epidemia de dengue da história do Rio de Janeiro. Por acreditarmos que no período epidêmico a influência dos riscos de dengue de cada bairro são determinantes

---

<sup>1</sup>Por convenção internacional, as semanas epidemiológicas são contadas de domingo à sábado. A primeira semana do ano é aquela que contém o maior número de dias de janeiro e a última a que contém o maior número de dias de dezembro. Fonte: Secretaria de Vigilância em Saúde - Ministério da Saúde ([www.saude.com.br](http://www.saude.com.br)).

para a disseminação da doença.

Os dados destas 25 semanas foram agregados no tempo e classificados segundo o bairro de moradia das pessoas infectadas. Desta forma, temos a informação de quantos casos ocorreram durante a epidemia em cada um dos 159 bairros da cidade.

O motivo de utilizar-se somente dados referentes ao período epidêmico foi o fato de acreditarmos outros períodos poderiam mascarar as influências dos riscos dos bairros. Caso fossem utilizados dados anteriores à epidemia, poderíamos concluir que um bairro não possui risco alto somente pelo fato dele ainda não ter sido exposto à doença. Por outro lado, os dados posteriores à epidemia podem mascarar os riscos devido ao efeito da diminuição da população de suscetíveis (imunização).

### 5.1.1 Covariáveis

As covariáveis utilizadas neste estudo dizem respeito às questões sócio-econômicas, de infra-estrutura e topografia da cidade do Rio de Janeiro. São elas:

- *Pontos Estratégicos (ESTRAT)*: As atividades de vigilância sanitária em nível municipal carecem de legislação de apoio e/ou de práticas de fiscalização para eliminarem os criadouros do mosquito em pontos considerados estratégicos. Estes são as borracharias, com pneus expostos; os cemitérios, com seus múltiplos vasos acumulando água; os depósitos de ferro velho a céu aberto retendo água de chuva; os terrenos baldios não cuidados, com múltiplos recipientes retendo água e as caixas d'água

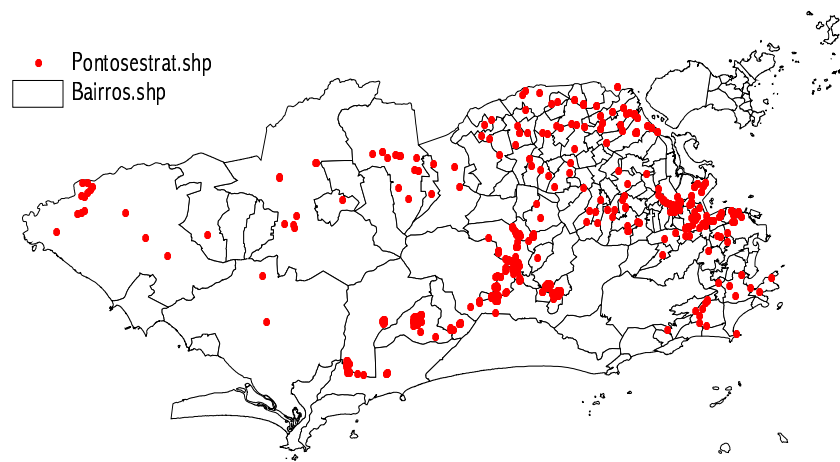


Figura 5.1: Pontos estratégicos apontados por agentes de saúde no Rio de Janeiro.

domiciliares descobertas (Tauil, 2002). Assim, utilizou-se o número de pontos estratégicos por bairro, dividido pela população do mesmo, como covariável. A figura 5.1 mostra os locais onde se encontram os pontos estratégicos do Rio de Janeiro;

- *Cota 100 m (C100)*: É o percentual da área de cada bairro coberta por elevações acima de 100 metros de altitude. Esta informação é relevante pois é esperado que o mosquito *Aedes aegypti* não ultrapasse barreiras com altura elevada;
- *Renda média e Salário médio (REND e SAL)*: A renda média ou o número de salários mínimos médio por chefe de família nos bairros pode servir como parâmetro para medir a influência de fatores econômicos no risco de dengue. Aqui, leva-se em consideração a renda em Reais de todos os indivíduos (inclusive aqueles com renda nula) referente ao ano de 2000;

- *Esperança de Vida ao Nascer (EVN)*: Número médio de anos que as pessoas viveriam a partir do nascimento (variável padronizada);
- *Taxa de Alfabetização de Adultos (ALFAB)*: Percentual de pessoas acima de 15 anos de idade que sabem ler e escrever (variável padronizada);
- *Índice de Desenvolvimento Humano Municipal (IDH)*: É obtido pela média aritmética simples de três índices, referentes às dimensões Longevidade (IDHM - Longevidade), Educação (IDHM - Educação) e Renda (IDHM - Renda). Os índices que compõem o IDH são relativos ao ano 2000 e levam em consideração variáveis relacionadas à esperança de vida ao nascer, taxa de alfabetização de adultos, taxa bruta de frequência escolar e renda per capita (variável padronizada);
- *Lixo não-coletado (LIXO)*: Mede o percentual do lixo do bairro que não é coletado. As condições gerais de infra-estrutura de um bairro podem ser medidas levando-se em conta indicadores relacionados à saneamento básico e coleta de lixo;
- *Água não-proveniente da rede geral (AGUA)*: A utilização de água proveniente de poços ou caixas-d'água a céu aberto propiciam ambientes favoráveis para a procriação dos mosquitos. Assim, o percentual de residências nos bairros que não possuem rede geral de água pode evidenciar a influência destes fatores no aumento do risco de dengue;
- *Favelas (FAV)*: A região da favela não somente constitui exemplo típico de zona de difícil acesso aos agentes de saúde, como também caracteriza-se pela irregularidade de suprimento de água potável (Santos & Carvalho, 2000). Entretanto, não existem dados disponíveis a respeito do total de população favelada por bairro na cidade do Rio de Janeiro.



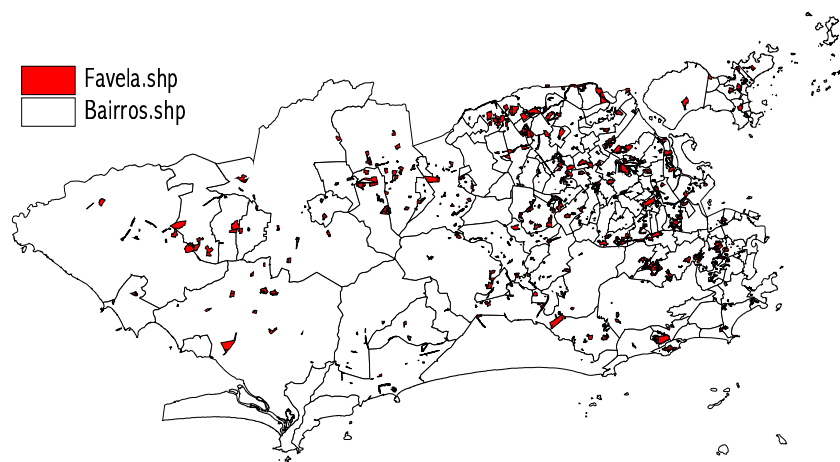


Figura 5.2: Localização das favelas no Rio de Janeiro - 2002.

Por este motivo, utilizamos uma informação alternativa que indica a proporção da área do bairro que é coberta por favelas (figura 5.2).

A tabela 5.1 apresenta a matriz de correlações entre as covariáveis.

A diversidade de covariáveis é necessária em virtude de ainda não existir um consenso entre quais fatores influem no aumento ou decréscimo, ainda que pequeno, do risco de dengue em uma determinada área ou região. Assim, o papel das covariáveis é fundamental nesta análise e a significância de algumas delas pode ajudar na identificação de fatores associados à doença e, conseqüentemente, fornecer o embasamento para o debate de novas estratégias de combate.

Tabela 5.1: Matriz de correlações entre as covariáveis.

$\rho$	ESTRAT	C100	REND	SAL	LIXO	AGUA	FAV	IDH	EVN	ALFAB
ESTRAT	1									
C100	0,20	1								
REND	-0,08	0,14	1							
SAL	-0,10	0,08	0,88	1						
LIXO	0,53	0,18	-0,14	-0,13	1					
AGUA	0,44	0,42	0,01	-0,13	0,38	1				
FAV	-0,09	0,05	-0,29	-0,35	0,02	-0,12	1			
IDH	-0,16	0,08	0,82	0,74	-0,4	-0,11	-0,42	1		
EVN	-0,15	0,03	0,70	0,64	-0,39	-0,17	-0,39	0,94	1	
ALFAB	-0,25	-0,15	0,57	0,56	-0,42	-0,28	-0,44	0,84	0,78	1

## 5.2 Modelo Proposto

Estamos interessados em analisar a variável  $Y_i = \text{Número de casos de dengue no bairro } i, i = 1, \dots, 159$ . Para isto, utilizaremos um modelo Poisson onde a média pode ser relacionada com covariáveis e/ou outros efeitos que ajudem a diminuir a incerteza a respeito da dinâmica da doença.

Desta forma, o modelo básico será dado por:

$$Y_i | r_i, e_i \sim \text{Poisson}(e_i r_i) \quad i = 1, \dots, 159,$$

onde  $r_i$  representa o risco de dengue e  $e_i$  representa o número esperado de casos de dengue para o bairro  $i$ .

O cálculo do número esperado de casos de dengue em cada bairro pode

ser realizado de várias formas. A forma utilizada aqui obtém os valores esperados para um bairro  $i$  a partir da divisão entre o número total de casos e a população total da cidade, multiplicada pela população de cada bairro. Isto é:

$$e_i = \frac{\sum y_i}{\sum pop_i} \times pop_i.$$

No segundo nível da hierarquia modelamos o logaritmo natural dos riscos  $r_i$  da seguinte forma:

$$\log(r_i) = \boldsymbol{\beta}'\mathbf{X}_i + b_i,$$

ou seja, os riscos serão relacionados com covariáveis através da função de ligação  $\log(\cdot)$  mais um efeito aleatório  $b_i$ , com o intuito de capturar outros efeitos não explicados pelas covariáveis.

Dando prosseguimento ao processo de inferência sob o enfoque bayesiano, devemos assumir distribuições a priori para os parâmetros do modelo a fim de obtermos as distribuições dos riscos de dengue a posteriori.

Para o vetor  $\boldsymbol{\beta}$  será especificada uma distribuição a priori Normal com vetor de médias  $\mathbf{0}$  e matriz de covariâncias  $\Sigma = \mathbf{I}\sigma_\beta^2$ , onde  $\sigma_\beta$  é igual a 1.000.

Ao vetor de efeitos aleatórios  $\mathbf{b}$  serão especificadas distribuições a priori que lhe permitam diferentes interpretações no modelo.

A primeira distribuição a priori que será considerada para estes efeitos não leva em consideração os efeitos aleatórios dos outros bairros da cidade. Em outras palavras, cada efeito é considerado como independente dos demais (Priori I). Assim,

$$b_i \sim Normal(0, \sigma^2),$$

onde  $\sigma^2$  possui distribuição a priori IG(0,05 ; 0,005).

A segunda maneira de especificar a distribuição a priori do efeito aleatório permite que o mesmo tenha uma interpretação de *efeito espacial*. Utilizando um modelo Condicional Autoregressivo Gaussiano Intrínseco (CAR) (Besag *et al.*, 1991), denotado por  $CAR(\sigma_b^2)$ , temos que, condicional aos demais efeitos, a distribuição de um efeito  $b_i$  é dada por:

$$[b_i | b_j = s_j, j \neq i] \sim Normal(m_i, v_i),$$

onde  $m_i = \frac{\sum_{j \in \delta_i} W_{ij} s_j}{\sum_{j \in \delta_i} W_{ij}}$ ,  $v_i = \frac{\sigma_b^2}{\sum_{j \in \delta_i} W_{ij}}$  e  $\delta_i$  representa o conjunto de áreas subjacentes à área  $i$ .

Assim, os efeitos aleatórios podem ser vistos como médias ponderadas dos efeitos dos bairros vizinhos. A especificação da matriz de pesos  $\mathbf{W}$ , que dará origem às Prioris II, III e IV, será discutida na seção 5.2.1.

Essa especificação condicional resulta na seguinte distribuição a priori conjunta para os efeitos aleatórios,

$$p(\mathbf{b} | \sigma_b^2) \propto \frac{1}{\sigma_b^n} \exp \left\{ -\frac{1}{2\sigma_b^2} \sum_{i=1}^n \sum_{j < i} w_{ij} (b_i - b_j)^2 \right\}$$

que possui vetor de médias  $\mathbf{0}$  e matriz de covariâncias contendo a diagonal igual a  $\frac{\sigma_b^2}{w_{i+}}$  e demais elementos iguais a  $-\frac{\sigma_b^2}{w_{ij}}$ .

Outra forma de especificar a distribuição dos efeitos aleatórios é conhecida como priori de *convolução* (Mollié, 1996). Neste caso, adicionamos outro efeito aleatório ao modelo, alterando-o para:

$$\log(r_i) = \boldsymbol{\beta}' \mathbf{X}_i + b_i + u_i,$$

onde  $\mathbf{b}$  terá uma distribuição a priori CAR, enquanto que os  $u_i$ 's terão distribuições a priori normais independentes com média zero.

A partir da especificação do modelo e das distribuições a priori dos parâmetros envolvidos, podemos seguir com a abordagem bayesiana e obter a distribuição a posteriori dos mesmos através do Teorema de Bayes.

Considere o vetor  $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{b}, \sigma_b^2)$  como aquele que contém todas as quantidades desconhecidas do modelo de riscos de dengue,  $\mathbf{y}$  o vetor contendo os casos de dengue nos bairros e  $\mathbf{x}$  a covariável associada. Assim, temos que a distribuição a posteriori de  $\boldsymbol{\theta}$  é dada por:

$$\begin{aligned} p(\boldsymbol{\theta}|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \\ &\propto p(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b}, \sigma_b^2)p(\boldsymbol{\beta}, \mathbf{b}, \sigma_b^2) \\ &\propto p(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b})p(\boldsymbol{\beta})p(\mathbf{b}|\sigma_b^2)p(\sigma_b^2) \end{aligned}$$

onde

$$p(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b}) = \left\{ \prod_{i=1}^{156} \frac{(e_i \exp \{\boldsymbol{\beta}'\mathbf{X}_i + b_i\})^{y_i} \exp \{-e_i \exp \{\boldsymbol{\beta}'\mathbf{X}_i + b_i\}\}}{y_i!} \right\}.$$

Como esta distribuição a posteriori não possui forma conhecida, propõe-se a utilização de algoritmos baseados nos métodos de Monte Carlo via Cadeias de Markov (MCMC). Mais especificamente, a partir das distribuições condicionais completas, utiliza-se o Amostrador de Gibbs (Gelfand & Smith, 1990) para obter-se uma amostra das distribuições a posteriori referentes à cada parâmetro. Entretanto, em virtude das distribuições condicionais completas dos parâmetros também não possuírem formas conhecidas (exceto no caso do parâmetro da variância), é necessária a utilização de outros métodos para a obtenção da amostra, como por exemplo, o método da rejeição adaptativa (Gilks & Wild, 1992). Para a obtenção das amostras das distribuições a posteriori destes modelos, utilizou-se o WinBugs versão 1.4.

Conforme foi apresentado, em virtude dos efeitos espaciais possuem distribuições a priori que dependem da informação de bairros vizinhos, bairros que são constituídos por ilhas distantes dos demais bairros não podem ser considerados na análise. Por este motivo, o bairro de *Paquetá* será desconsiderado da análise. Outros dois bairros, a saber *Vasco da Gama* e *Parque Colúmbia*, foram também excluídos por não haver informação disponível a respeito de covariáveis que pudessem ser utilizadas no modelo. Já os bairros de *Joá* e *Grumari*, apesar de não possuem informação a respeito do número de casos de dengue no período considerado, foram incluídos na análise e terão os riscos estimados a partir das distribuições a priori, das informações das covariáveis e dos efeitos espaciais dos bairros vizinhos. Assim, o número de bairros que serão analisados neste estudo será reduzido de 159 para 156.

Estes bairros são mostrados na figura 5.3, juntamente com os riscos de dengue estimados por máxima verossimilhança (MV). Estes por sua vez, são obtidos a partir do estimador de MV da média da distribuição Poisson

$$\hat{\mu} = (r_i \hat{e}_i) = y_i \implies \hat{r}_i = \frac{y_i}{e_i},$$

ou seja, são dados pela razão entre os casos observados e os casos esperados de dengue dos bairros. Mollié (1996) comenta algumas desvantagens relacionadas a utilização destes estimadores.

A figura 5.3 mostra os riscos estimados por máxima verossimilhança e destaca os bairros que, apesar de não possuem informações de casos de dengue, foram incluídos na análise.

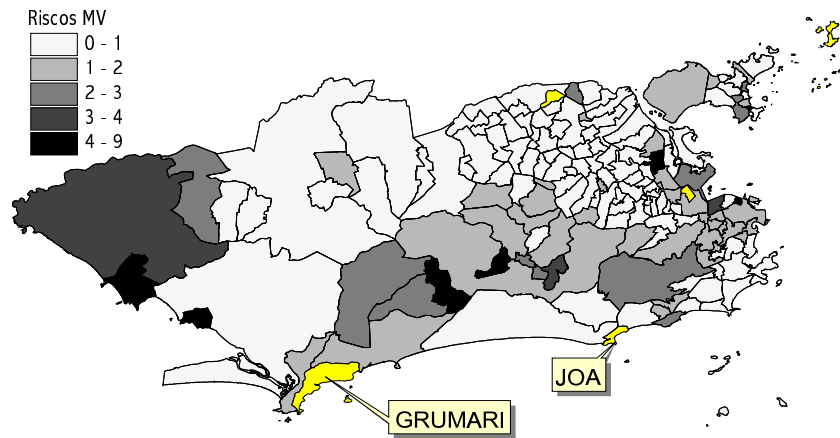


Figura 5.3: Riscos de dengue estimados por máxima verossimilhança e indicação dos bairros em que não há informação sobre casos de dengue (em amarelo).

### 5.2.1 Estruturas de vizinhanças

Para a especificação da matriz de pesos da estrutura de vizinhança  $\mathbf{W}$ , foram utilizados modelos contendo três diferentes estruturas para que fosse possível identificar qual delas se ajustaria melhor aos dados. Todas estas estruturas de vizinhança consideradas obedecem a restrição de simetria da matriz  $\mathbf{W}$  e são apresentadas a seguir.

- *Vizinhança Binária (Priori II)*

Esta é a estrutura de vizinhança mais simples e é muito utilizada em modelos espaciais onde leva-se em consideração apenas o fato de um bairro fazer fronteira com outro. Desta forma,  $W_{ij}$  é igual a 1, se o bairro  $i$  faz fronteira com o bairro  $j$ , e é igual a zero, caso contrário. Nesse caso, a média e a variância dos efeitos espaciais tornam-se:

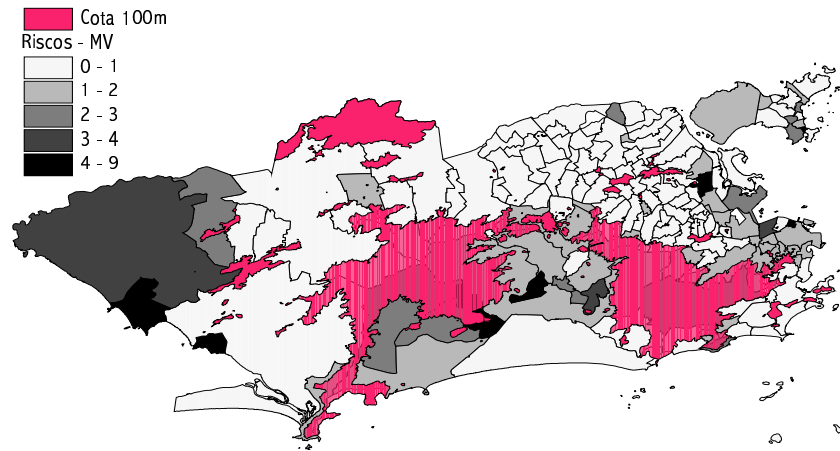


Figura 5.4: Cota 100 metros e sua relação com os riscos de dengue na cidade do Rio de Janeiro.

$$m_i = \frac{\sum_{j \in \delta_i} s_j}{n_i} \quad \text{e} \quad v_i = \frac{\sigma_b^2}{n_i},$$

onde  $n_i$  é o número de vizinhos do bairro  $i$ .

- *Vizinhança ponderada pelo tamanho da fronteira* (Priori III)

Em virtude dos diferentes tamanhos e das diferentes formas dos bairros do Rio de Janeiro, os quais são bastante irregulares, considerou-se bastante apropriada uma estrutura de vizinhança que levasse em consideração o tamanho da fronteira entre dois bairros. Assim,  $W_{ij}$  é igual ao comprimento da fronteira entre o bairro  $i$  e o bairro  $j$ , em km. Ao contrário da vizinhança binária, esta estrutura de vizinhança faz com que a influência do efeito espacial de um bairro em seu vizinho aumente ou diminua de acordo com a extensão de fronteira existente entre eles.

- *Vizinhança ponderada pelo tamanho da fronteira e pelas barreiras naturais* (Priori IV)



Esta estrutura de vizinhança é uma extensão da anterior e está baseada na discussão de Mollié (1996), onde sugere-se que fronteiras naturais, tais como montanhas, possam influenciar na análise espacial dos riscos relativos. Assim,  $W_{ij}$  é igual ao comprimento da fronteira entre o bairro  $i$  e o bairro  $j$ , em km, menos o comprimento da fronteira coberta por montanhas. Foram consideradas como montanhas as regiões da cidade que encontram-se acima de 100 metros de altitude. Este tipo de estrutura de vizinhança pode se mostrar relevante, uma vez que as montanhas representam uma barreira para o mosquito *Aedes aegypti* e podem fazer com que o efeito espacial do risco de dengue não se propague de um bairro a outro. Esta característica parece estar presente ao analisar-se os riscos de dengue estimados por máxima verossimilhança e as áreas da cidade cobertas pela Cota 100 metros (ver figura 5.4).

Para o cálculo dos comprimentos de fronteiras entre os bairros foi utilizado o software ArcView GIS 3.2a (Breslin *et al.*, 1999).

### 5.2.2 Comparação entre os Modelos

Através de uma análise preliminar, onde utilizou-se modelos lineares generalizados, foram selecionadas as combinações de covariáveis que mostravam-se significativas para explicar o padrão dos riscos observados de dengue nos bairros.

Nesta análise, observou-se que apenas as covariáveis ESTRAT, FAV, AGUA e ALFAB eram significativas quando analisadas individualmente. Além disso, os modelos que continham efeitos de três ou mais covariáveis não forneceram coeficientes significativos. O mesmo aconteceu quando incorporou-

se efeitos de interação nos modelos.

Sendo assim, na análise espacial desenvolvida neste capítulo, serão utilizados modelos aditivos contendo efeitos de apenas uma ou duas covariáveis. Ao todo, serão analisados 20 modelos segundo as 5 prioris para os efeitos aleatórios.

Em virtude de sua importância na literatura de dengue e por terem sido consideradas como significativas na aplicação do modelo preliminar, priorizar-se-á a avaliação dos efeitos das covariáveis FAV e ESTRAT nos modelos contendo duas covariáveis.

Os modelos, classificados segundo as covariáveis envolvidas e segundo a priori utilizada para os efeitos aleatórios, serão comparados a fim de que seja possível identificar quais os fatores que são mais importantes na análise dos riscos de dengue.

O critério de seleção adotado destacará os modelos que, segundo cada estrutura de dados, apresentarem os menores valores do DIC.

### **5.3 Resultados**

Para a obtenção das amostras das distribuições a posteriori dos parâmetros dos modelos por meio do algoritmo MCMC, monitorou-se a trajetória de duas cadeias partindo de valores iniciais distintos. Entretanto, em virtude da grande autocorrelação existente entre os valores amostrados, foram incluídas na amostra da posteriori apenas 1 (um) valor a cada 400 valores gerados. Adotando-se este procedimento, a convergência foi obtida após amostrar-se

pouco mais de 10 observações para praticamente todos os modelos considerados.

Para compor a amostra, observou-se as cadeias por mais 1000 iterações, totalizando duas amostras de tamanho 1000 para cada parâmetro de cada modelo.

Na figura 5.5 são mostradas as cadeias com os valores incluídos na amostra da distribuição a posteriori dos parâmetros  $\beta_0$ ,  $\beta_1$  e  $\sigma^2$  para o modelo FAV contendo a Priori I para os efeitos aleatórios.

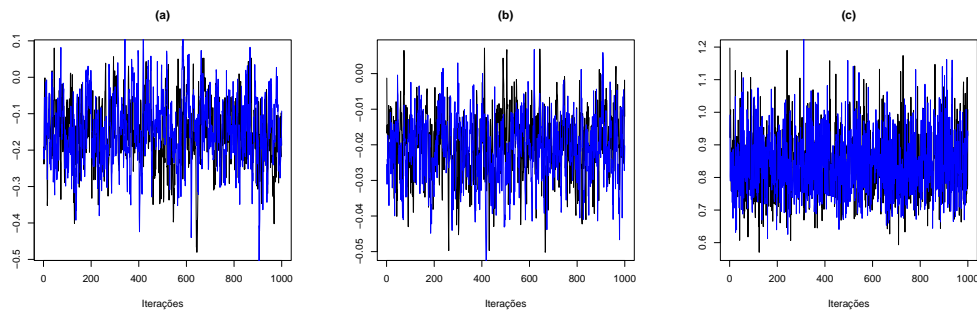


Figura 5.5: Evolução das cadeias após obtenção de convergência para os parâmetros  $\beta_0$  (a),  $\beta_1$  (b),  $\sigma^2$  (c).

Analisando o efeito das covariáveis, pode-se verificar o papel dos fatores sócio-econômicos e de infra-estrutura nos riscos de dengue de cada um dos bairros da cidade. Apesar de não se mostrarem como os fatores determinantes para um aumento ou decréscimo dos riscos de dengue, suas influências foram significativas (ver tabelas 5.2, 5.3, 5.4, 5.5 e 5.6).

As variáveis que apresentaram relação positiva com os riscos de dengue foram ESTRAT e AGUA, o que reforça a importância da eliminação de criadouros naturais além de destacar a necessidade de uma ampliação da

cobertura de rede de água na cidade.

As variáveis LIXO e EDUC apresentaram resultados contraditórios, evidenciando relações ora negativas, ora positivas, com os riscos de dengue. Já as variáveis IDH e EVN não se mostraram significantes em nenhum dos modelos testados.

As variáveis que apresentaram relação negativa com os riscos de dengue foram SAL, REND e FAV. Sobre as duas primeiras variáveis, vemos uma evidência de que bairros com menores rendimentos médios apresentam riscos de dengue mais elevados. Apesar disto, salientamos que os rendimentos médios dos chefes de família variam bastante dentro de um mesmo bairro e este fator não é captado por uma análise baseada somente em médias dos mesmos.

Tabela 5.2: Relação das covariáveis com os riscos de dengue para os modelos com a Priori I para os efeitos aleatórios. Utilizou-se como critério para significância o IC95% para os parâmetros relativos às covariáveis, exceto nos casos indicados por (\*) onde utilizou-se o IC90%.

COVARIÁVEL 1	RELAÇÃO C/RISCO	COVARIÁVEL 2	RELAÇÃO C/RISCO
ESTRAT	positiva	-	-
FAV	negativa	-	-
C100	positiva*	-	-
REND	ñ significativa	-	-
SAL	negativa*	-	-
LIXO	positiva*	-	-
AGUA	positiva	-	-
ALFAB	negativa	-	-
IDH	ñ significativa	-	-
EVN	ñ significativa	-	-
ESTRAT	positiva	FAV	negativa
ESTRAT	positiva	C100	ñ significativa
ESTRAT	positiva	REND	ñ significativa
ESTRAT	positiva	LIXO	ñ significativa
ESTRAT	positiva*	AGUA	positiva
FAV	negativa	REND	ñ significativa
FAV	negativa	LIXO	positiva*
FAV	negativa	AGUA	positiva
FAV	negativa	C100	positiva*

Tabela 5.3: Relação das covariáveis com os riscos de dengue para os modelos com a Priori II para os efeitos aleatórios. Utilizou-se como critério para significância o IC95% para os parâmetros relativos às covariáveis, exceto nos casos indicados por (\*) onde utilizou-se o IC90%.

COVARIÁVEL 1	RELAÇÃO C/RISCO	COVARIÁVEL 2	RELAÇÃO C/RISCO
ESTRAT	positiva	-	-
FAV	negativa	-	-
C100	positiva	-	-
REND	negativa	-	-
SAL	negativa	-	-
LIXO	positiva	-	-
AGUA	positiva	-	-
ALFAB	positiva*	-	-
IDH	ñ significativa	-	-
EVN	ñ significativa	-	-
ESTRAT	positiva	FAV	negativa
ESTRAT	positiva	C100	positiva
ESTRAT	positiva	REND	ñ significativa
ESTRAT	positiva	LIXO	negativa*
ESTRAT	positiva	AGUA	positiva
FAV	negativa	REND	negativa
FAV	negativa	LIXO	positiva*
FAV	negativa	AGUA	positiva
FAV	negativa	C100	positiva

Tabela 5.4: Relação das covariáveis com os riscos de dengue para os modelos com a Priori III para os efeitos aleatórios. Utilizou-se como critério para significância o IC95% para os parâmetros relativos às covariáveis, exceto nos casos indicados por (\*) onde utilizou-se o IC90%.

COVARIÁVEL 1	RELAÇÃO C/RISCO	COVARIÁVEL 2	RELAÇÃO C/RISCO
ESTRAT	positiva	-	-
FAV	negativa	-	-
C100	positiva	-	-
REND	negativa	-	-
SAL	negativa	-	-
LIXO	positiva	-	-
AGUA	positiva	-	-
ALFAB	positiva*	-	-
IDH	ñ significativa	-	-
EVN	ñ significativa	-	-
ESTRAT	positiva	FAV	negativa
ESTRAT	positiva	C100	positiva*
ESTRAT	positiva	REND	negativa
ESTRAT	positiva	LIXO	negativa
ESTRAT	positiva	AGUA	positiva
FAV	negativa	REND	negativa
FAV	negativa	LIXO	positiva*
FAV	negativa	AGUA	positiva
FAV	negativa	C100	positiva

Tabela 5.5: Relação das covariáveis com os riscos de dengue para os modelos com Priori IV para os efeitos aleatórios. Utilizou-se como critério para significância o IC95% para os parâmetros relativos às covariáveis.

COVARIÁVEL 1	RELAÇÃO C/RISCO	COVARIÁVEL 2	RELAÇÃO C/RISCO
ESTRAT	positiva	-	-
FAV	negativa	-	-
C100	ñ significativa	-	-
REND	negativa	-	-
SAL	negativa	-	-
LIXO	ñ significativa	-	-
AGUA	positiva	-	-
ALFAB	positiva	-	-
IDH	ñ significativa	-	-
EVN	ñ significativa	-	-
ESTRAT	positiva	FAV	negativa
ESTRAT	positiva	C100	ñ significativa
ESTRAT	positiva	REND	negativa
ESTRAT	positiva	LIXO	negativa
ESTRAT	positiva	AGUA	ñ significativa
FAV	negativa	REND	negativa
FAV	negativa	LIXO	ñ significativa
FAV	negativa	AGUA	positiva
FAV	negativa	C100	ñ significativa



Tabela 5.6: Relação das covariáveis com os riscos de dengue para os modelos com Priori V para os efeitos aleatórios. Utilizou-se como critério para significância o IC95% para os parâmetros relativos às covariáveis, exceto nos casos indicados por (\*) onde utilizou-se o IC90%.

COVARIÁVEL 1	RELAÇÃO C/RISCO	COVARIÁVEL 2	RELAÇÃO C/RISCO
ESTRAT	positiva	-	-
FAV	negativa	-	-
C100	ñ significativa	-	-
REND	negativa*	-	-
SAL	negativa	-	-
LIXO	positiva*	-	-
AGUA	positiva	-	-
ALFAB	ñ significativa	-	-
IDH	ñ significativa	-	-
EVN	ñ significativa	-	-
ESTRAT	positiva	FAV	negativa*
ESTRAT	positiva	C100	ñ significativa
ESTRAT	positiva	REND	negativa*
ESTRAT	positiva	LIXO	ñ significativa
ESTRAT	positiva*	AGUA	positiva*
FAV	negativa	REND	negativa
FAV	negativa*	LIXO	positiva*
FAV	negativa*	AGUA	positiva
FAV	negativa*	C100	ñ significativa

Quanto aos resultados obtidos a respeito dos dados de favelas, podemos dizer que os mesmos são de difícil interpretação. Uma relação negativa deste fator com os riscos de dengue contraria outros resultados obtidos relacionados com fatores presentes em regiões favelizadas. Regiões de favelas não somente constituem exemplo típico de zona de difícil acesso aos agentes de saúde, como também caracterizam-se pela irregularidade de suprimento de água potável (Santos & Carvalho, 2000). Além disso, sabemos que os moradores destas áreas dispõem de menos infra-estrutura urbana e geralmente possuem baixa renda. Com base na análise destes fatores esperaríamos riscos mais elevados de dengue nestas regiões.

Teixeria *et al.* (2003) analisaram a dinâmica do vírus do dengue em Salvador-BA e não encontraram relacionamentos positivos entre regiões com baixa infra-estrutura e riscos de dengue. Kypraios (2003), analisando os riscos de leptospirose no Rio de Janeiro, encontrou relação negativa com dados a respeito de favelas. Isto mostra que este tipo de resultado não está limitado somente ao estudo do dengue.

Por outro lado, Medronho (1995), ao estudar a incidência de dengue na cidade do Rio de Janeiro com dados agregados em nível de R.A.<sup>2</sup> (*Região Administrativa*), evidenciou a importância das condições de moradia, adensamento populacional, infraestrutura urbana e elevação da temperatura do ar na determinação do dengue. De acordo com simulações efetuadas, o autor concluiu que o grande impacto na diminuição do risco de ocorrência de dengue associa-se à melhoria das condições de vida da população. Neste estudo o autor utilizou dados das epidemias ocorridas no Rio de Janeiro em

---

<sup>2</sup>As R.A.'s correspondem a uma divisão administrativa da cidade e são compostas pela união de alguns bairros vizinhos. No Rio de Janeiro existem ao todo 32 R.A.'s.

1986/1987 e 1990/1991 agregados no tempo.

Apesar disto, acreditamos que o fato de estarmos estudando a área coberta por favelas ao invés da proporção de população favelada não invalida nossas conclusões. Um dos indícios disto está nos resultados encontrados analisando-se os riscos de dengue agregados em R.A.'s e sua relação com os dados existentes de população favelada por R.A.'s. Através de um modelo semelhante àquele proposto na seção 5.2, também chegamos às mesmas conclusões a respeito da relação inversa existente entre os riscos da doença e as favelas.

Comparando os diferentes modelos segundo o DIC (ver tabela 5.7), podemos perceber as vantagens e desvantagens de cada uma das 5 maneiras de modelar os efeitos aleatórios do modelo de riscos.

Baseando-se neste critério, o modelo que possui a covariável FAV e Priori I para os efeitos aleatórios se mostrou como o mais indicado para modelar os riscos de dengue. As distribuições a posteriori dos parâmetros  $\beta_0$ ,  $\beta_1$  e  $\sigma^2$  são mostradas na figura 5.6 e as médias à posteriori dos riscos estimados são apresentados na figura 5.7.

Os modelos com Prioris II, III e IV apresentaram riscos estimados bastante semelhantes aos obtidos pelos modelos com Priori I. A diferença mais perceptível encontra-se nos efeitos aleatórios estimados, onde os modelos com prioris CAR apresentaram um mapa mais suavizado (ver figuras 5.8 e 5.9). Este resultado é bastante interessante pois nos mostra que a componente dos riscos que não é explicada pelas covariáveis pode ser bem representada por efeitos espaciais. Além disso, nota-se que estes efeitos tem um papel importante nas estimativas dos riscos, dada a similaridade de ambos. Pode-se ver

Tabela 5.7: Tabela com DIC's para todos os modelos considerados segundo a priori utilizada para os efeitos aleatórios.

MODELO	PRIORI I	PRIORI II	PRIORI III	PRIORI IV	PRIORI V
SEM COV.	1513,03	1596,53	1565,28	1691,11	1513,21
ESTRAT	1513,13	1595,14	1565,29	1672,06	1513,65
FAV	<b>1510,96</b>	1579,53	1569,42	1661,51	1512,73
C100	1514,33	1591,94	1570,71	1679,26	1512,58
REND	1512,23	1583,68	1568,88	1671,25	1511,16
SAL	1512,96	1589,91	1568,06	1678,21	1511,54
LIXO	1513,57	1594,29	1573,96	1676,55	1514,04
AGUA	1513,23	1598,16	1569,13	1668,35	1511,77
ALFAB	1512,11	1588,40	1571,04	1677,53	1512,89
IDH	1512,35	1590,12	1569,98	1676,43	1512,68
EVN	1511,92	1589,70	1572,92	1677,76	1512,61
ESTRAT*FAV	1512,21	1584,18	1563,45	1662,91	1513,01
ESTRAT*C100	1512,49	1591,76	1563,38	1661,58	1513,14
ESTRAT*REND	1511,33	1585,77	1566,34	1665,32	1513,21
ESTRAT*LIXO	1512,33	1590,82	1565,63	1662,95	1512,47
ESTRAT*AGUA	1512,24	1583,57	1563,57	1662,51	1511,56
FAV*REND	1512,99	1582,68	1567,33	1662,54	1514,41
FAV*LIXO	1512,78	1578,61	1573,92	1671,35	1514,52
FAV*AGUA	1511,61	1588,72	1563,97	1667,56	1513,08
FAV*C100	1512,63	1584,26	1567,88	1671,36	1512,17

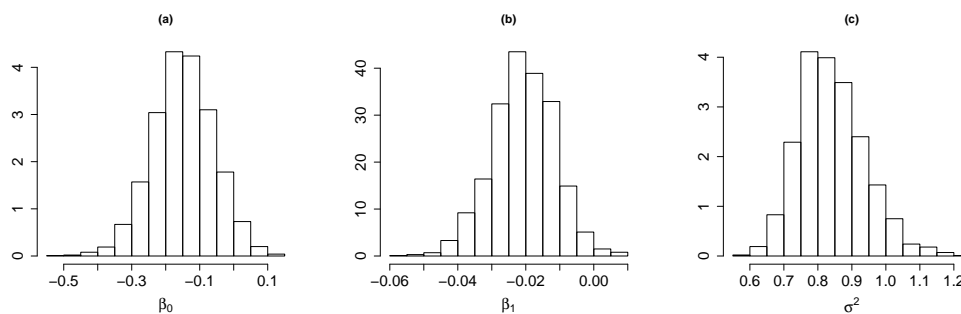


Figura 5.6: Histogramas das amostras das distribuições a posteriori dos parâmetros  $\beta_0$  (a),  $\beta_1$  (b),  $\sigma^2$  (c) para o modelo FAV contendo a Priori I para os efeitos aleatórios.

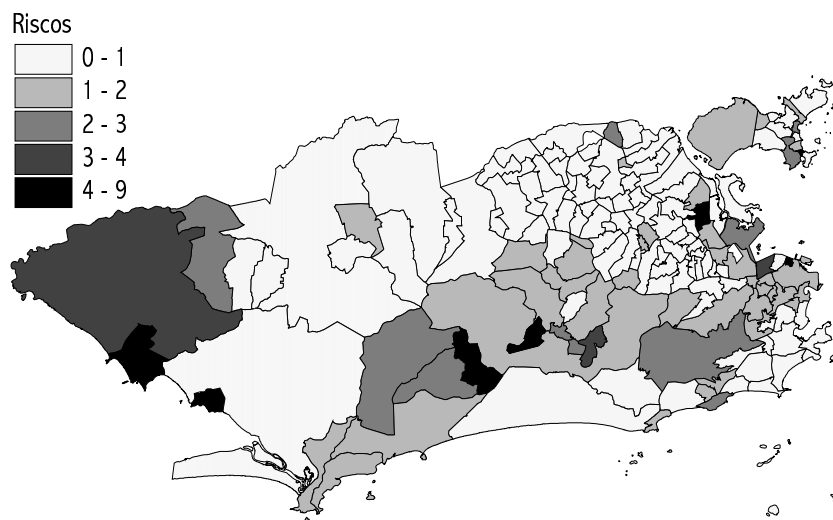


Figura 5.7: Riscos de dengue estimados pelo modelo com a covariável FAV e com a Priori I para os efeitos aleatórios.

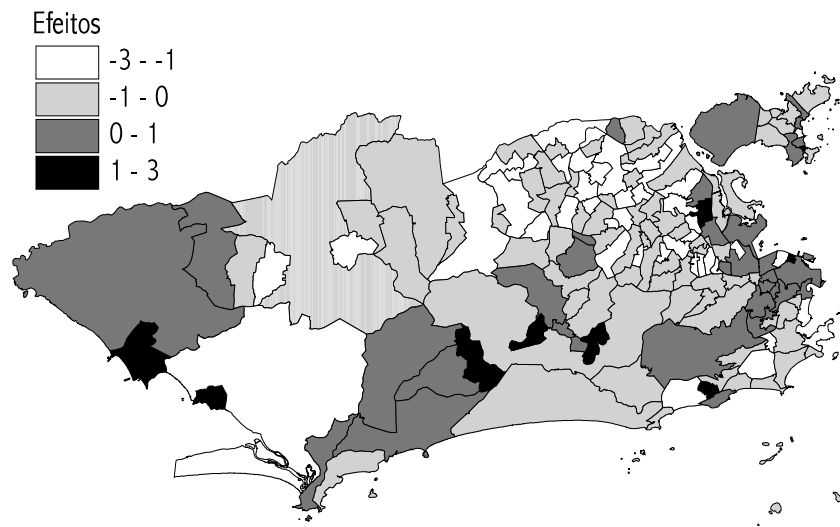


Figura 5.8: Médias a posteriori dos efeitos aleatórios estimados pelo modelo com a covariável FAV utilizando-se a Priori I.

que a zona oeste, a zona central e o eixo coberto por montanhas apresentam tanto efeitos espaciais altos, como também riscos de dengue altos. No que diz respeito à significância, os efeitos aleatórios para todos os modelos contendo as Prioris I, II, III e IV se mostraram significativos. Um mapa com as variâncias dos efeitos para um dos modelos é apresentada na figura 5.10.

Dentre os modelos com efeitos espaciais, aquele cuja estrutura de vizinhança da priori levava em consideração a extensão da fronteira entre os bairros apresentou os melhores resultados segundo o critério do DIC. Por outro lado, a utilização da informação referente à Cota 100 metros na composição da estrutura de vizinhança não respondeu de maneira satisfatória, conforme esperávamos. Além de produzirem os modelos com DIC mais elevados, algumas covariáveis perderam a significância, como a própria variável C100. A hipótese mais provável é que tenha havido um confundimento entre a informação da estrutura de vizinhança e a informação da variável C100.



Figura 5.9: Médias a posteriori dos efeitos espaciais estimados pelo modelo com covariáveis ESTRAT e FAV utilizando-se a Priori III.

Entretanto, não descartamos a importância de ampliar-se estudos relacionados à influência das montanhas ou estudos que busquem novas maneiras de incorporar a informação relacionada às barreiras naturais do deslocamento do vírus do dengue em uma área urbana.

A partir destes resultados, analisou-se os modelos que utilizam uma priori de convolução (Priori V) para os efeitos espaciais  $b_i$ 's, onde a estrutura de vizinhança escolhida foi a que leva em consideração os comprimentos de fronteira entre os bairros.

Os resultados obtidos através destes modelos apresentaram DIC's em patamares semelhantes àqueles obtidos pelos modelos que utilizavam a Priori I para os efeitos aleatórios. Para os modelos que utilizavam as covariáveis C100, REND, SAL e AGUA os valores do DIC foram os mais baixos. Entretanto, os intervalos de 95% de credibilidade indicaram que a maior parte dos efeitos espaciais  $b_i$ 's e dos efeitos  $u_i$ 's não eram significativos. Enquanto o

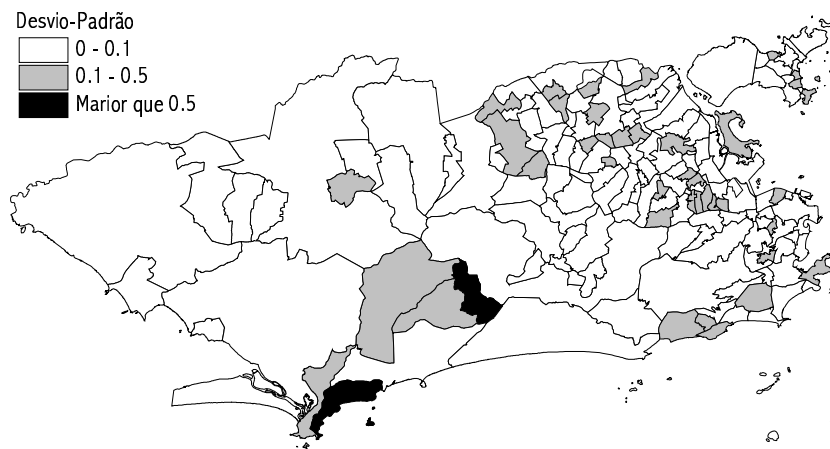


Figura 5.10: Desvios-padrões dos efeitos espaciais estimados pelo modelo com covariáveis ESTRAT e FAV utilizando-se a Priori III.

modelo com covariáveis ESTRAT e FAV com Priori III teve 82,7% de bairros com efeitos espaciais significativos (129 bairros entre 156), o modelo com a covariável SAL e Priori V apresentou 21,15% de bairros com efeitos espaciais significativos (32 de 156) e 7,05% de bairros com efeitos  $u_i$ 's significativos (11 de 156).

As figuras 5.11, 5.12 e 5.13 apresentam os histogramas das amostras das distribuições a posteriori de alguns efeitos aleatórios dos modelos FAV com Priori I e do modelo SAL com Priori V. A figura 5.14, por sua vez, apresenta os intervalos de credibilidade para todos os efeitos aleatórios do modelo SAL com Priori V.



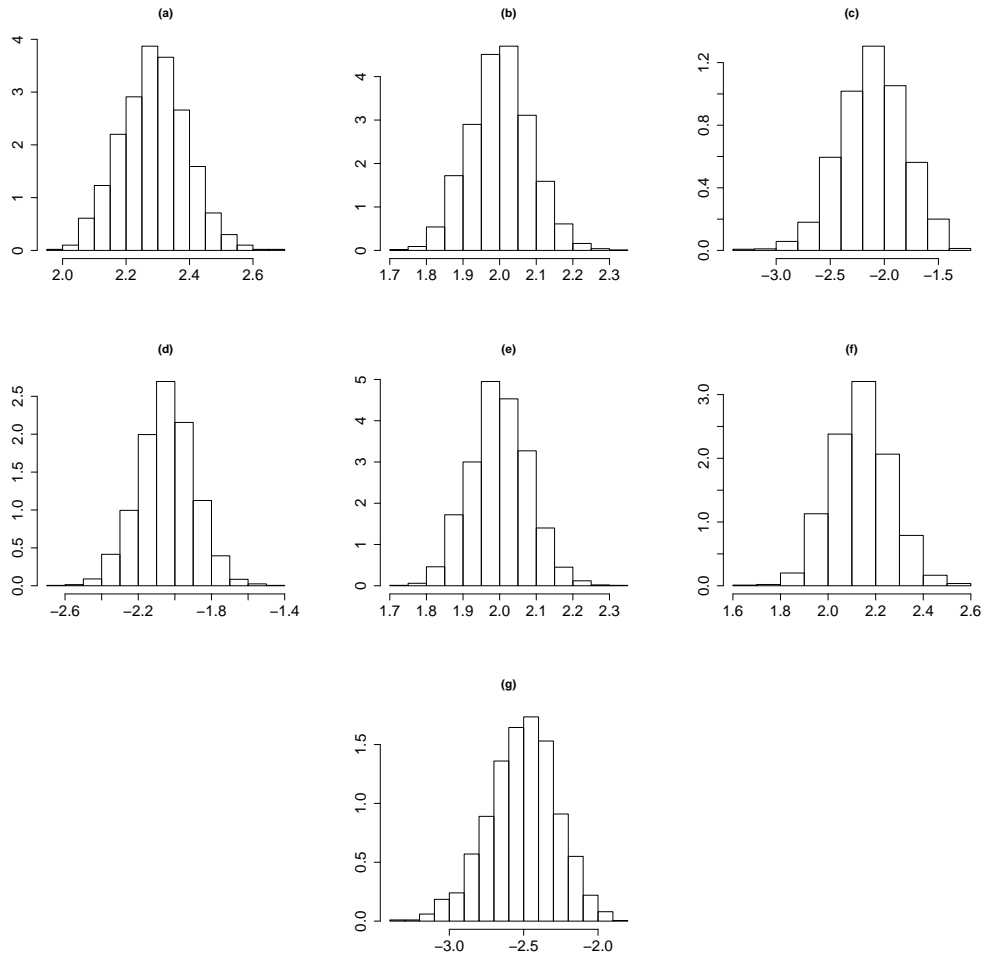


Figura 5.11: Histogramas das amostras das distribuições a posteriori dos efeitos aleatórios dos bairros Saúde (a), Grajaú (b), Quintino Bocaiúva (c), Guadalupe (d), Gardênia Azul (e), Itanhangá (f) e Grumari (g) para o modelo FAV com Priori I para os efeitos aleatórios.

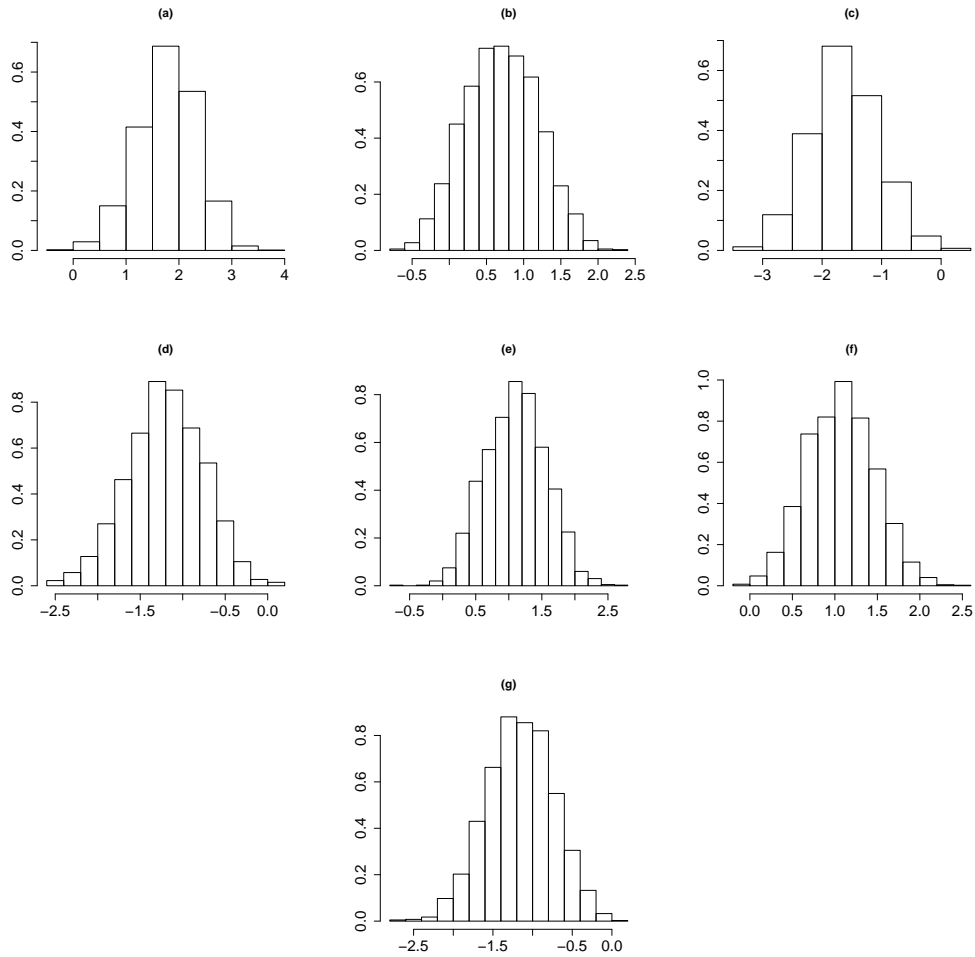


Figura 5.12: Histogramas das amostras das distribuições a posteriori dos efeitos espaciais dos bairros Saúde (a), Grajaú (b), Quintino Bocaiúva (c), Guadalupe (d), Gardênia Azul (e), Itanhangá (f) e Grumari (g) para o modelo SAL com Priori V para os efeitos aleatórios.

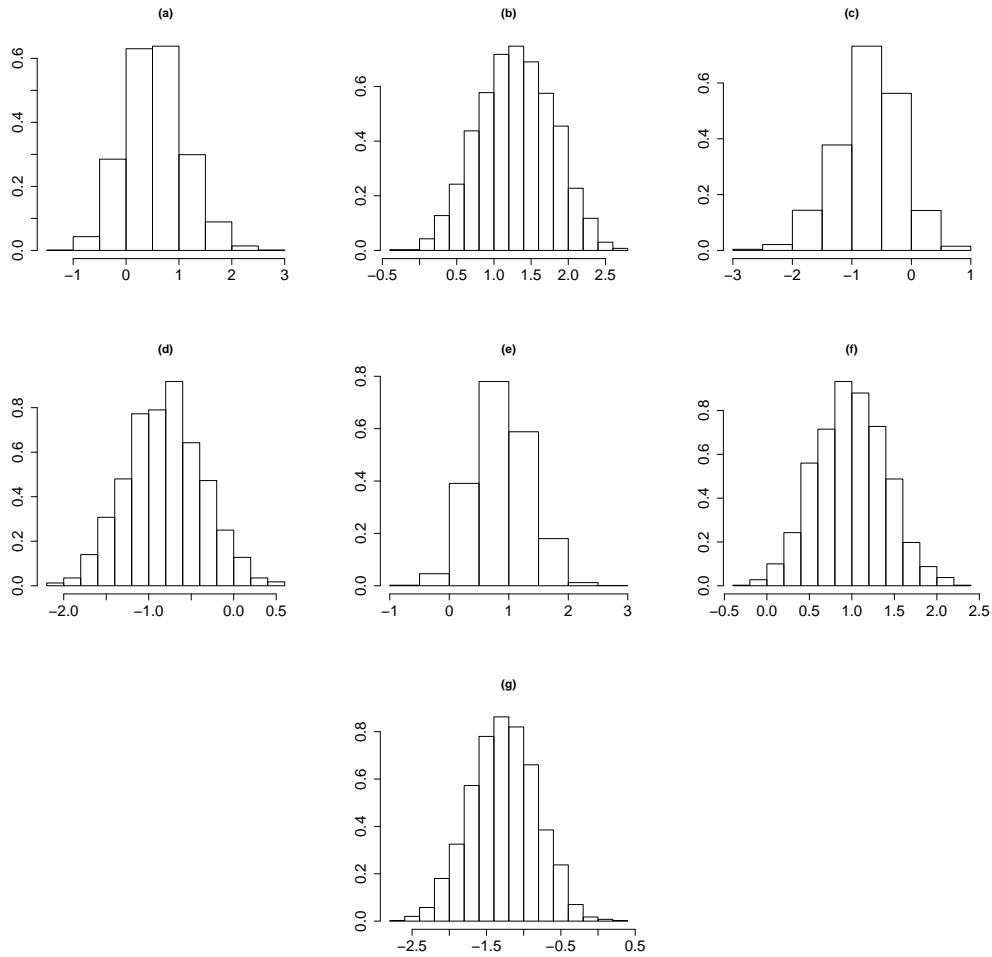


Figura 5.13: Histogramas das amostras das distribuições a posteriori dos efeitos aleatórios dos bairros Saúde (a), Grajaú (b), Quintino Bocaiúva (c), Guadalupe (d), Gardênia Azul (e), Itanhangá (f) e Grumari (g) para o modelo SAL com Priori V para os efeitos aleatórios.

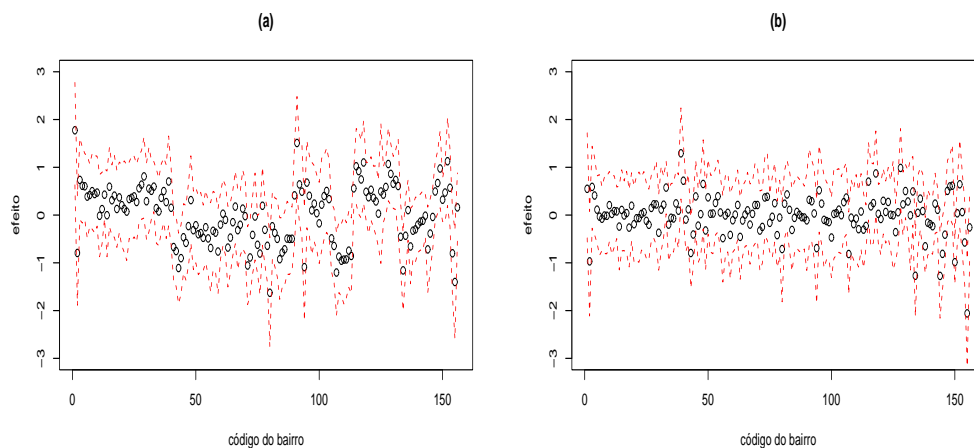


Figura 5.14: Estimativas a posteriori e respectivos IC95% dos efeitos espaciais (a) e efeitos independentes (b) utilizando o modelo SAL com priori de convolução (Priori V).

## 5.4 Conclusões

Em virtude da dificuldade de identificar-se fatores que influenciam na distribuição espacial dos riscos de dengue, conforme discutido anteriormente, consideramos satisfatórios os resultados obtidos neste capítulo. A identificação de uma pequena, porém significativa, influência de fatores sócio-econômicos e de infra-estrutura nos riscos da doença são um passo importante para a diminuição dos seus efeitos.

Dentre os resultados de difícil explicação, destacamos a influência negativa do fator associado às favelas nos riscos de dengue. Apesar desta influência já ter sido encontrada na literatura, sua correspondência com os fatores associados às características de uma área coberta por favelas não se confirmaram, já que estes fatores indicaram uma relação positiva com os riscos. Uma possível causa desta discrepância pode estar relacionada ao

problema da sub-notificação de casos de dengue. Estudos relacionados ao conhecimento do grau de informação da população e à forma como lidam com a doença, como o estudo realizado por Lenzi *et al.* (2000) em uma área favelizada do Rio de Janeiro, podem contribuir para a resolução desta questão.

Quanto aos resultados obtidos com os diferentes modelos, vimos que a utilização da Priori I para os efeitos aleatórios forneceu os DIC's em níveis mais baixos para a maioria dos casos. Apesar de se tratarem de modelos menos complexos, lembramos que estes modelos têm a desvantagem de não fornecerem ao pesquisador uma possível explicação do que representam os efeitos aleatórios. A utilização de prioris CAR para os efeitos aleatórios forneceu bons resultados, conseguindo traduzir estes efeitos como efeitos espaciais devidos à influência dos bairros vizinhos na estimação dos riscos de dengue. A combinação da utilização de um efeito espacial com um efeito aleatório não estruturado, apesar de gerar bons resultados quando comparada com o DIC dos demais modelos, indicou que sua utilização deve ser questionada, em virtude da não significância da maior parte dos efeitos.

A grande variedade de resultados produzidos através da utilização das prioris CAR, tanto no que diz respeito à significância de covariáveis, quanto no que diz respeito à influência dos efeitos espaciais, evidenciam a importância da escolha da estrutura de vizinhança na especificação das distribuições a priori para os efeitos aleatórios.

Outro ponto importante a ser destacado é a agregação dos dados no tempo e no espaço. Diferentes formas de agregação podem levar o pesquisador a diferentes conclusões. No caso da epidemia 2001/2002, os casos de dengue surgiram a partir de poucos pontos da cidade, alastrando-se para todos os

demais ao longo do período de verão. Acreditamos que a utilização de dados referentes ao começo da epidemia poderiam mascarar, tanto um possível efeito espacial, quanto os efeitos das covariáveis.

A agregação espacial pode influenciar nos resultados relacionados aos efeitos das covariáveis. A informação das covariáveis, quando agregadas para descrever grandes regiões, produzem grande variabilidade dentro das mesmas. No caso do Rio de Janeiro este problema é agravado ainda mais devido às desigualdades sócio-econômicas e de infra-estrutura existentes na cidade. Assim, à medida que a agregação aumenta, a informação das covariáveis discriminará menos as regiões e, conseqüentemente, os resultados poderão tornar-se enganosos.

Assim, acreditamos que estudar o efeito de diferentes formas de agregação dos dados no tempo e no espaço podem ser relevantes para os estudos da distribuição dos riscos da doença.

# Capítulo 6

## Análise Espaço-Temporal

Este capítulo visa generalizar a análise realizada no capítulo anterior através de uma análise dos riscos de dengue ao longo de setenta semanas consecutivas.

Através de uma análise dos riscos no tempo, podemos destacar os períodos em que a influência espacial foi determinante e quais os períodos em que esta distribuição foi aparentemente aleatória.

O período em que deflagrou-se a epidemia, com a explosão do número de casos de dengue notificados, possui grande importância no estudo da doença, pois uma análise realizada neste período poderia indicar a forma como o vírus se propagou ao longo dos bairros.

O capítulo está organizado da seguinte forma: Na seção 6.1 são apresentados os dados utilizados na análise. Na seção 6.2 apresentamos os diferentes modelos para analisar a influência de efeitos espaço-temporais nos riscos de dengue. A seção 6.3 apresenta os principais resultados obtidos sob cada modelo. Por fim, a seção 6.4 apresenta as conclusões a respeito da

utilização destes modelos e analisa os resultados mais importantes.

## 6.1 Dados Utilizados

Os dados analisados aqui representam as 148.160 notificações semanais de casos de dengue na cidade do Rio de Janeiro registradas durante 70 semanas a partir da primeira semana epidemiológica do ano de 2001.

Apesar de estarmos considerando a semana epidemiológica como o indicador de tempo, as semanas referentes ao ano de 2002 serão adicionadas à contagem das semanas referentes ao ano de 2001. Assim, como o ano de 2001 possui 52 semanas epidemiológicas, as 18 semanas epidemiológicas de 2002 serão tratadas como as semanas 53, 54, ... e 70, respectivamente.

Assim como na análise espacial, os dados estão classificados segundo o bairro de moradia do paciente infectado. A figura 6.1 apresenta os totais de casos ao longo das semanas consideradas na análise, onde é possível notar o aumento elevado de casos ocorrido entre as semanas 48 e 56.

As semanas 71 e 72, que foram utilizadas na análise espacial quando agregou-se os dados semanais, não serão consideradas neste capítulo. Como estas duas semanas são referentes ao período pós-epidemia, onde o número de casos de dengue na cidade foi baixo, uma caracterização da distribuição espacial seria prejudicada em virtude da pouca informação.



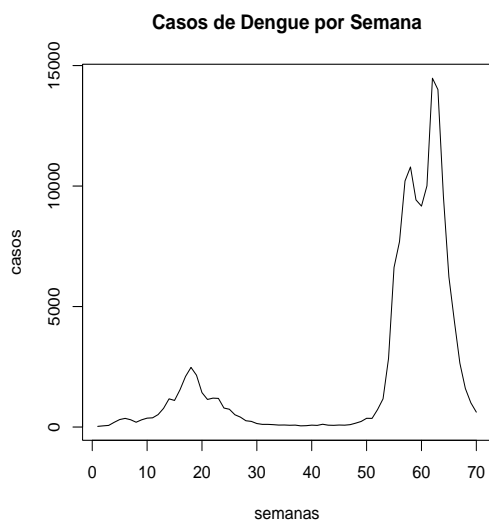


Figura 6.1: Número de casos de dengue ao longo das 70 semanas consideradas na análise.

## 6.2 Modelo Proposto

Nesta análise espaço-temporal, estamos interessados em analisar a variável  $Y_{it}$  = número de casos de dengue no bairro  $i$  na semana  $t$ ,  $i = 1, \dots, 156$  e  $t = 1, \dots, 70$ . Assim como nas análises anteriores, utilizaremos um modelo Poisson onde a média é relacionada à efeitos de ordem espacial e/ou temporal.

Desta forma, o modelo básico será dado por:

$$Y_{it} | r_{it}, e_{it} \sim \text{Poisson}(e_{it} r_{it}) \quad i = 1, \dots, 156 \quad e \quad t = 1, \dots, 70,$$

onde  $r_{it}$  representa o risco de dengue e  $e_{it}$  representa o número esperado de casos de dengue para o bairro  $i$  na semana  $t$ .

O número esperado de casos de dengue é calculado de forma semelhante à apresentada na seção 5.2. O único diferencial é que o número de casos esperado por bairro é calculado para cada uma das 70 semanas consideradas.

A seguir, modelamos o logaritmo natural dos riscos  $r_{it}$  da seguinte forma:

$$\log(r_{it}) = \beta_t + b_{it} \quad (6.1)$$

onde  $t = 1, \dots, 70$  e  $i = 1, \dots, 156$ . Ou seja, os riscos serão relacionados a um efeito de nível (temporal) mais um efeito aleatório (espacial ou espaço-temporal).

A relação expressa em (6.1) será realizada através de três modelos distintos, os quais são apresentados a seguir:

- **Modelo 1**

A primeira forma de relacionar os riscos de dengue com efeitos de ordem espaço-temporal é considerar um efeito médio fixo mais um efeito aleatório que varia dependendo do bairro e da semana de referência. Assim, os riscos de dengue são relacionados com os efeitos da seguinte forma:

$$\log(r_{it}) = \beta + b_{it}. \quad (6.2)$$

Para dar prosseguimento ao processo de inferência, assumimos uma distribuição a priori Normal com média zero e desvio padrão igual a 1.000 para o parâmetro  $\beta$ . Assim como na análise espacial, a distribuição a priori para o efeito aleatório  $\mathbf{b}$  pode ser especificada de várias formas.

A primeira distribuição que será considerada não leva em consideração os efeitos aleatórios dos outros bairros da cidade em qualquer semana considerada (Priori I). Assim,

$$b_{it} \sim Normal(0, \sigma^2),$$

onde  $\sigma^2$  possui distribuição a priori  $IG(0,05 ; 0,005)$ .

A segunda forma de especificar a distribuição a priori do efeito aleatório, que permite que o caracterizemos como um efeito espacial, é utilizar a priori  $CAR(\sigma^2)$ . Assim, condicional aos demais efeitos em uma dada semana  $t$ , a distribuição de um efeito  $b_{it}$  será dada por:

$$[b_{it} | b_{jt} = s_{jt}, j \neq i] \sim Normal(m_{it}, v_i)$$

onde  $m_{it} = \frac{\sum_{j \in \delta i} W_{ij} s_{jt}}{\sum_{j \in \delta i} W_{ij}}$ ,  $v_i = \frac{\sigma^2}{\sum_{j \in \delta i} W_{ij}}$  e  $\delta i$  representa o conjunto de áreas subjacentes à área  $i$ .

Assim, para cada semana, os efeitos aleatórios podem ser vistos como efeitos espaciais dos bairros vizinhos. Assim como na seção 5.2.1, a especificação da matriz de pesos  $\mathbf{W}$  dará origem às Prioris II, III, IV e V.

A Priori V caracteriza a utilização de uma priori de convolução para os efeitos aleatórios, adicionando na equação (6.2) um efeito aleatório com distribuição

$$u_{it} \sim Normal(0, \sigma_u^2).$$

- **Modelo 2**

A segunda forma de relacionar os riscos de dengue com efeitos de ordem espaço-temporal é analisar os efeitos espaciais e temporais de forma independente. Para isto, consideramos um efeito de nível para captar os efeitos temporais e um efeito aleatório para captar a estrutura espacial ao longo dos bairros.

Assim, os riscos de dengue são relacionados com os efeitos da seguinte forma:

$$\log(r_{it}) = \beta_t + b_i. \tag{6.3}$$

A forma de especificar a distribuição a priori para o parâmetro  $\beta_t$  pode ser feita de duas formas. A primeira delas considera que os  $\beta$ 's são independentes e possuem distribuição Normal com média zero e desvio-padrão igual a 1.000. A segunda forma de especificar a distribuição a priori leva em consideração os conceitos de Modelos Lineares Generalizados Dinâmicos, modificando a relação (6.3) para :

$$\log(r_{it}) = \beta_t + b_i$$

$$\beta_t = \beta_{t-1} + e_t, \quad e_t \sim N(0, \sigma_\beta^2).$$

Em outras palavras, o nível da série é modelado segundo um passeio aleatório, onde

$$\beta_0 \sim N(0, \sigma_\beta^2).$$

Com relação a distribuição a priori para o efeito aleatório  $\mathbf{b}$ , serão utilizadas as mesmas Prioris I, II, III, IV e V descritas na análise espacial.

- **Modelo 3**

A terceira forma de relacionar os riscos de dengue com efeitos de ordem espaço-temporal combina o efeito de nível do Modelo 2 com os efeitos aleatórios do Modelo 1.

Assim, os riscos de dengue são relacionados com os efeitos da seguinte forma:

$$\log(r_{it}) = \beta_t + b_{it}.$$

Neste caso, as distribuições a priori especificadas para  $\beta_t$  serão as mesmas utilizadas no Modelo 2 (aleatória e dinâmica) e as distribuições a priori para o efeito aleatório  $\mathbf{b}$  serão as mesmas utilizadas no Modelo 1 (Priori I, II, III, IV e V).

A partir da especificação do modelo e das distribuições a priori dos parâmetros envolvidos, mais uma vez podemos seguir com a abordagem bayesiana e obter a distribuição a posteriori dos mesmos através do Teorema de Bayes.

Para ilustrar o procedimento de inferência, considere o vetor  $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{b}, \sigma^2, \sigma_\beta^2)$  como aquele que contém todas as quantidades desconhecidas do Modelo 3. Assim, temos que a distribuição a posteriori de  $\boldsymbol{\theta}$  é dada por:

$$\begin{aligned} p(\boldsymbol{\theta}|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \propto p(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b}, \sigma^2, \sigma_\beta^2)p(\boldsymbol{\beta}, \mathbf{b}, \sigma^2, \sigma_\beta^2) \\ &\propto p(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b}) \prod_{t=1}^{70} \left\{ p(\beta_t|\beta_{t-1}, \sigma_\beta^2)p(\mathbf{b}_t|\sigma_b^2) \right\} p(\sigma^2)p(\sigma_\beta^2) \end{aligned}$$

onde

$$p(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b}) = \left\{ \prod_{t=1}^{70} \prod_{i=1}^{156} \frac{(e_{it} \exp \{\beta_t + b_{it}\})^{y_{it}} \exp \{-e_{it} \exp \{\beta_t + b_{it}\}\}}{y_{it}!} \right\}.$$

Assim como as distribuições a posteriori dos modelos da análise espacial, esta distribuição também não possui forma conhecida. Assim, para implementar o algoritmo MCMC e realizar a inferência através de amostras da distribuição a posteriori dos parâmetros, utilizou-se o software WinBugs versão 1.4.

### 6.2.1 Comparação entre os Modelos

Os modelos, classificados segundo o tipo de priori para os efeitos aleatórios e segundo o tipo de estrutura espaço-temporal, serão comparados a fim de que seja possível destacar o modelo mais apropriado para analisar os dados.

Assim como nas análises anteriores, o critério de seleção de modelos utilizado foi o DIC.

## 6.3 Resultados

Para a obtenção das amostras das distribuições a posteriori dos parâmetros dos modelos por meio do algoritmo MCMC, monitorou-se a trajetória de duas cadeias partindo de valores iniciais distintos. Para minimizar problemas devidos à autocorrelação existente entre os valores amostrados, foram incluídas na amostra da posteriori 1 (um) valor a cada 20 valores gerados. A partir disto, a convergência foi obtida rapidamente (aproximadamente após amostrarmos 20 valores) para todos os modelos considerados.

Comparando-se o desempenho dos diferentes modelos através do critério DIC, vimos que o Modelo 3 foi o que apresentou os melhores resultados. O Modelo 2, por outro lado, apresentou valores do DIC demasiadamente elevados. A tabela 6.1 apresenta o valor do DIC para os modelos considerados segundo cada uma das prioris para os efeitos aleatórios.

Analisando-se o desempenho das diferentes estruturas de vizinhança nas distribuições a priori para os efeitos aleatórios, vimos que o critério do DIC destacou os modelos com Priori II. A partir destes resultados, analisou-se os modelos que utilizam uma priori de convolução (Priori V) para os efeitos aleatórios, onde a estrutura de vizinhança binária foi a escolhida.

A utilização da Priori V para os efeitos aleatórios dos modelos resultou nos valores de DIC mais baixos dentre todos os modelos considerados. Entretanto, a utilização desta priori forneceu estimativas dos efeitos aleatórios não significativas. Desta forma, apesar de ter fornecido os valores de DIC mais baixos, acreditamos que a utilização de uma priori de convolução não se faz necessária, uma vez que apenas 1 (um) efeito aleatório foi suficiente para captar a estrutura espacial dos dados. A figura 6.2 apresenta as estimativas

Tabela 6.1: Tabela com DIC's para todos os modelos considerados segundo a priori utilizada para os efeitos aleatórios.

MODELO	PRIORI I	PRIORI II	PRIORI III	PRIORI IV	PRIORI V
$\beta + b_{it}$	43942,7	45132,7	45540,1	45656,9	41878,2
$\beta_t + b_i$	75152,8	75233,2	75214,6	75344,2	72711,4
$\beta_t + b_i$ (din)	75152,1	72624,2	72620,3	72635,4	72596,5
$\beta_t + b_{it}$	44631,1	44091,6	44305,4	44446,5	41960,1
$\beta_t + b_{it}$ (din)	44649,2	42050,1	42184,8	42291,7	41914,5

a posteriori dos efeitos aleatórios do Modelo 1 com Priori V referente aos bairros Saúde, Jacaré e Tauá.

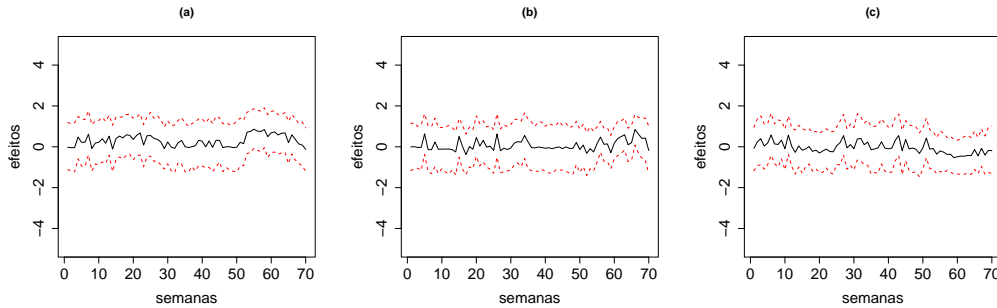


Figura 6.2: Médias a posteriori e respectivos IC95% dos efeitos aleatórios dos bairros Saúde (a), Jacaré (b) e Tauá (c) para o Modelo 1 e Priori V para os efeitos aleatórios.

Analisando o resultado das demais prioris para os efeitos aleatórios, vemos que a Priori I forneceu resultados melhores para os Modelos 1 e 2, enquanto que a Priori II forneceu os melhores resultados para o Modelo 3. A Priori III forneceu os melhores resultados para o Modelo 2 com estrutura dinâmica, enquanto que a Priori IV não foi superior em nenhum dos modelos

considerados.

A combinação de modelo e priori que forneceu o melhor resultado foi a combinação entre o Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

O histograma das amostras das distribuições a posteriori dos parâmetros  $\sigma^2$  e  $\sigma_\beta^2$  referentes a este modelo são apresentados na figura 6.3.

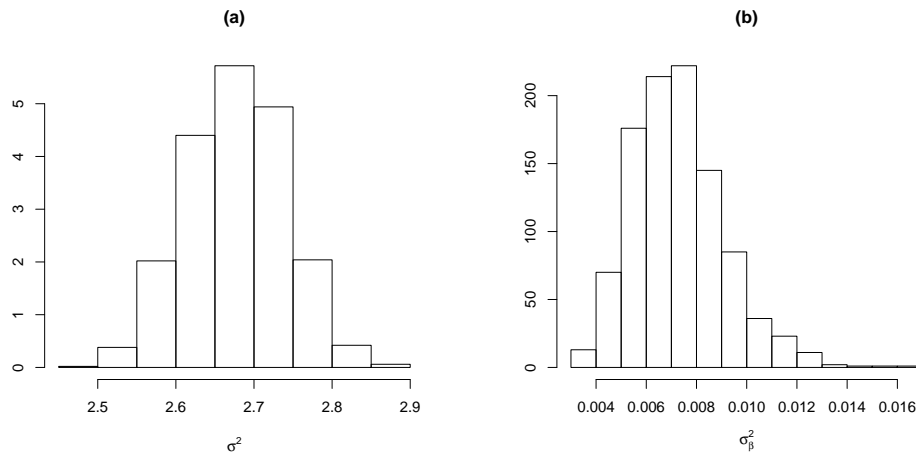


Figura 6.3: Histogramas das amostras das distribuições a posteriori dos parâmetros  $\sigma^2$  e  $\sigma_\beta^2$  referentes ao Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

Analisando-se as estimativas da média a posteriori obtidas a partir do modelo (figura 6.4), podemos ver a mudança de comportamento ocorrida com o início da epidemia de dengue. Pode-se notar que a partir da semana 48, há um aumento e depois uma queda brusca do nível da série. Após isso, o nível se mantém em patamares mais baixos do que aqueles apresentados nas primeiras semanas.

O efeito da epidemia também pode ser visualizado através dos mapas



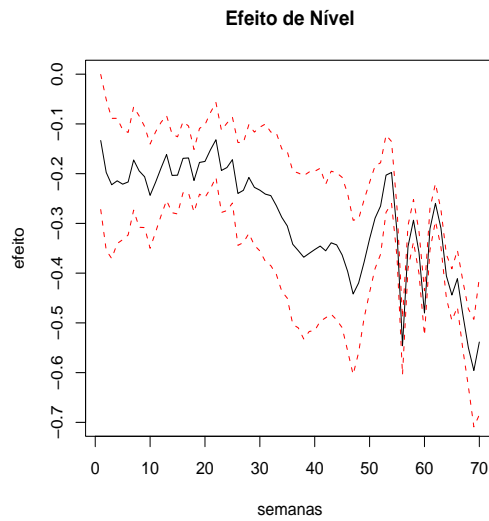


Figura 6.4: Média a posteriori do nível da série e respectivos IC95% referentes ao Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

com as estimativas dos riscos de dengue a posteriori para as semanas 48 a 59 (figuras 6.5 e 6.6), onde pode-se notar a expansão da epidemia ao longo da cidade. Podemos ver também que os maiores riscos se concentravam na zona norte e leste da cidade, evoluindo para as demais regiões ao longo da epidemia. Na figura 6.7, onde é apresentada a evolução dos riscos de dengue em alguns bairros da cidade, pode-se ver o aumento significativo do risco nos bairros Gardênia Azul e Itanhangá a partir da semana 48.

A análise do efeito espacial na distribuição espaço-temporal dos riscos de dengue pode ser feita através de mapas com as estimativas a posteriori dos efeitos espaciais para algumas semanas. As figuras 6.8 e 6.9 mostram as estimativas a posteriori dos efeitos espaciais das semanas 48 a 59.

Analisando o mapa com os efeitos espaciais referentes à semana 48, vemos claramente que a zona leste da cidade apresentava os efeitos espaciais

mais elevados, enquanto que a zona oeste apresentava os efeitos mais baixos. À medida que a epidemia foi surgindo, os efeitos espaciais distribuíram-se ao longo dos bairros que compõem a cadeia de montanhas do município no sentido leste-oeste. Além disso, destacou-se a elevação dos efeitos espaciais na região oeste, em torno do bairro de Santa Cruz.

A figura 6.10 apresenta a evolução dos efeitos espaciais para alguns bairros da cidade. Pode-se notar que os efeitos espaciais dos bairros Gardênia Azul e Itanhangá aumentaram significativamente durante o período da epidemia. Por outro lado, os bairros de Quintino Bocaiúva e Guadalupe tiveram uma queda nas estimativas dos efeitos. Isto pode ser explicado devido a posição geográfica destes bairros. Enquanto que os bairros de Quintino Bocaiúva e Guadalupe encontram-se em regiões onde os riscos de dengue eram mais elevados no início da epidemia, os bairros de Gardênia Azul e Itanhangá se encontram em regiões que possuíam riscos mais baixos (figura 6.11). Comparando-se a evolução dos efeitos espaciais com a evolução dos riscos de dengue, vemos que a componente espacial parece ser o fator determinante para a elevação e queda dos riscos de dengue.

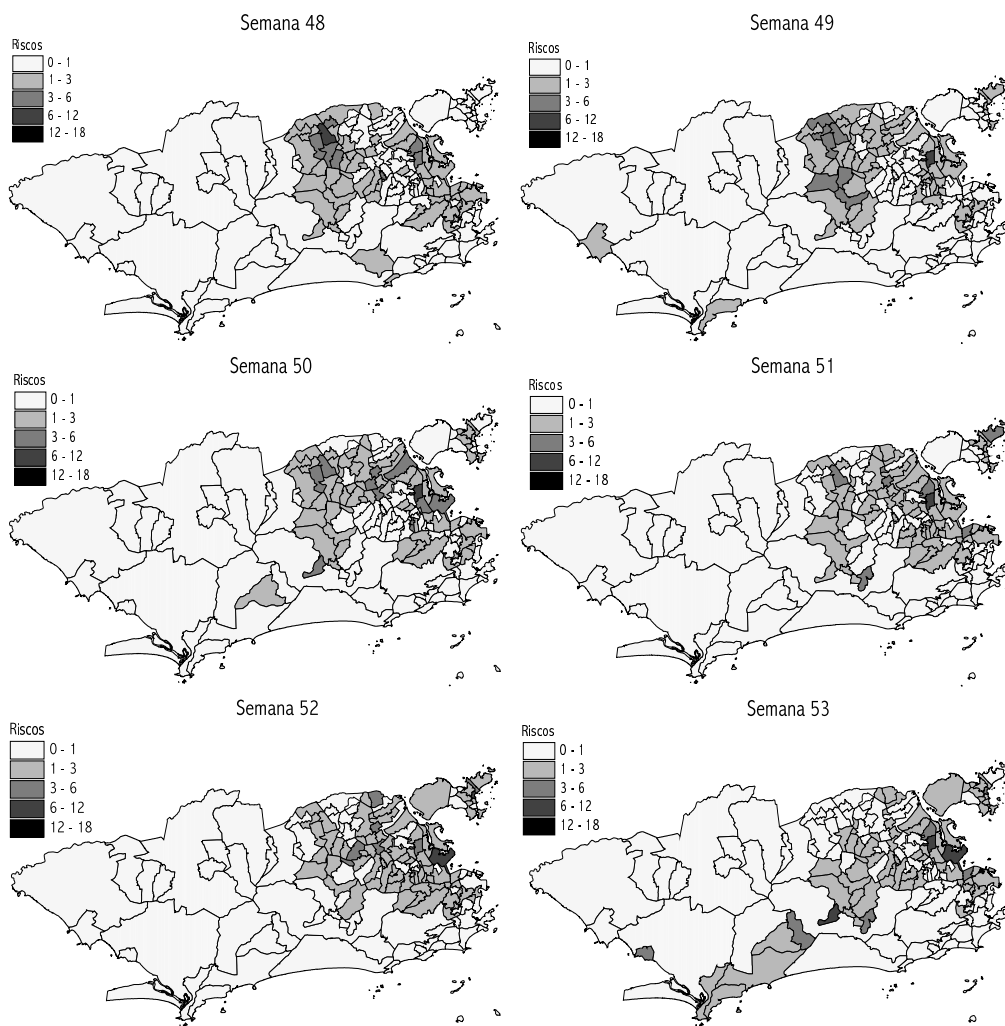


Figura 6.5: Mapas contendo as médias a posteriori dos riscos de dengue para os bairros do Rio de Janeiro entre as semanas 48 e 53 estimadas pelo Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

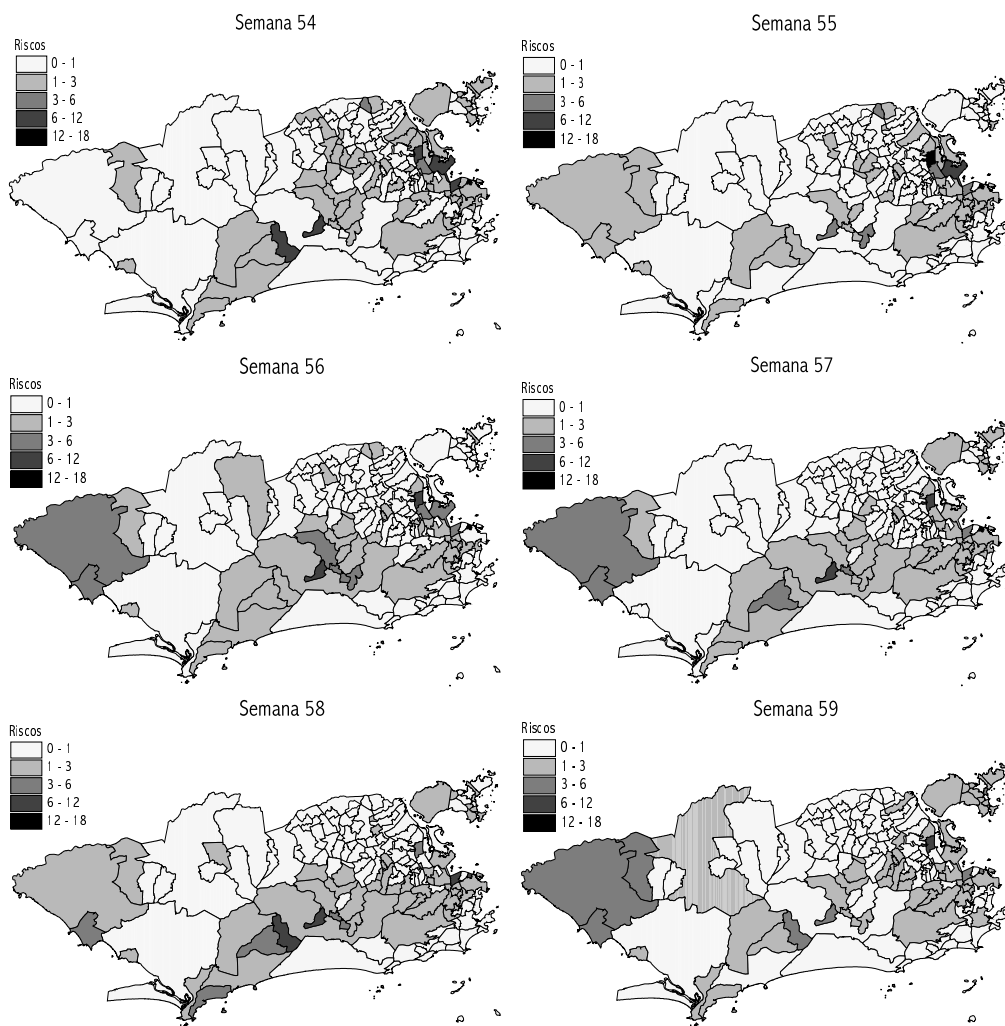


Figura 6.6: Mapas contendo as médias a posteriori dos riscos de dengue para os bairros do Rio de Janeiro entre as semanas 54 e 59 estimadas pelo Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

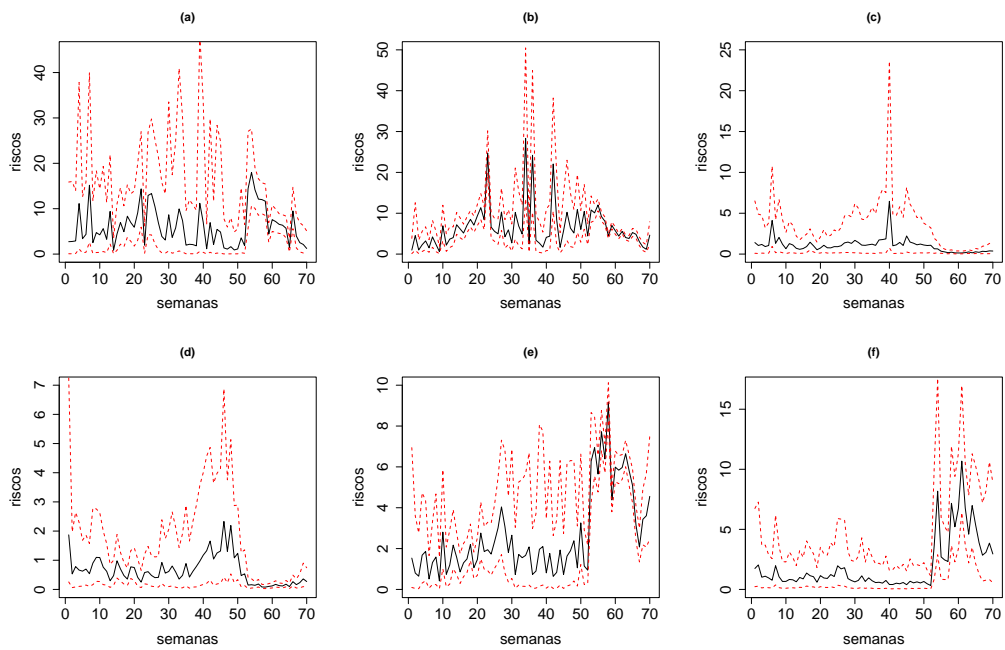


Figura 6.7: Médias a posteriori e respectivos IC95% dos riscos de dengue dos bairros Saúde (a), Grajaú (b), Quintino Bocaiúva (c), Guadalupe (d), Gardênia Azul (e) e Itanhangá (f) para o Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

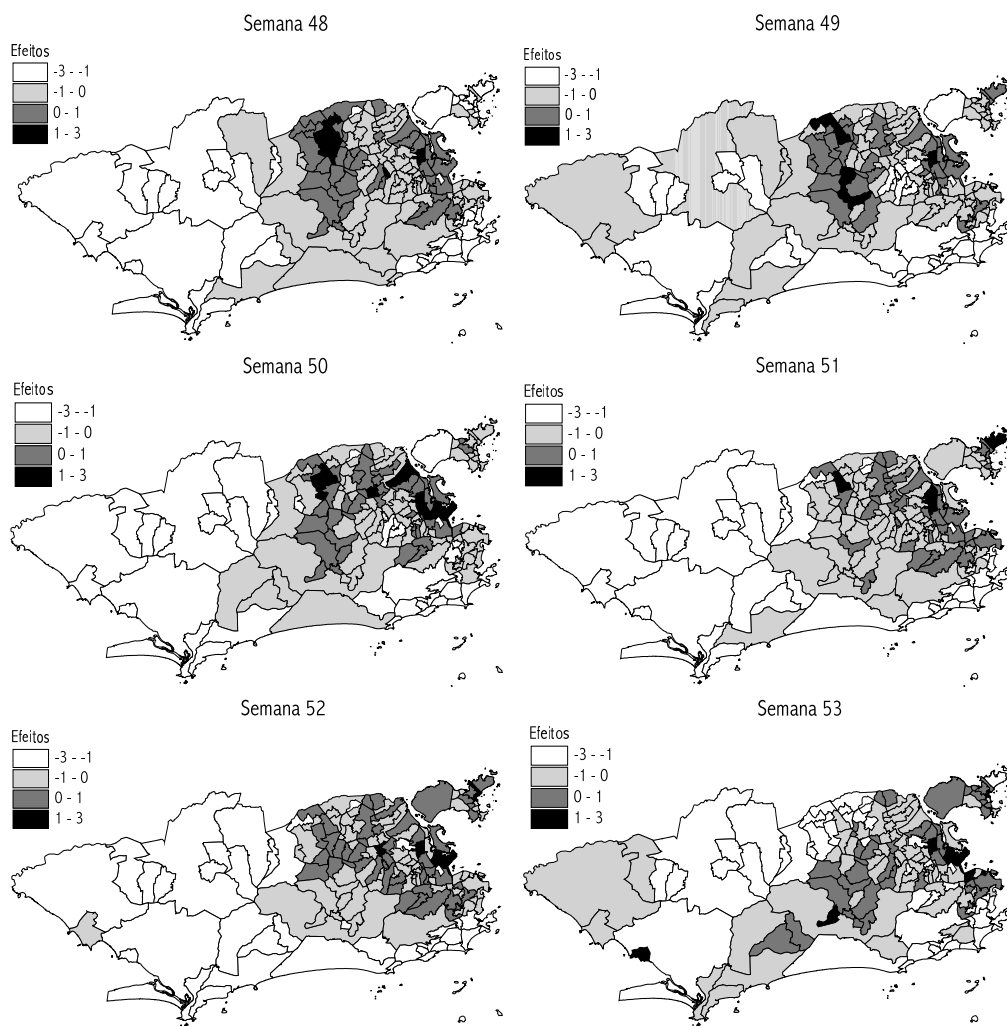


Figura 6.8: Mapas contendo as médias a posteriori dos efeitos espaciais dos bairros do Rio de Janeiro entre as semanas 48 e 53 estimados pelo Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

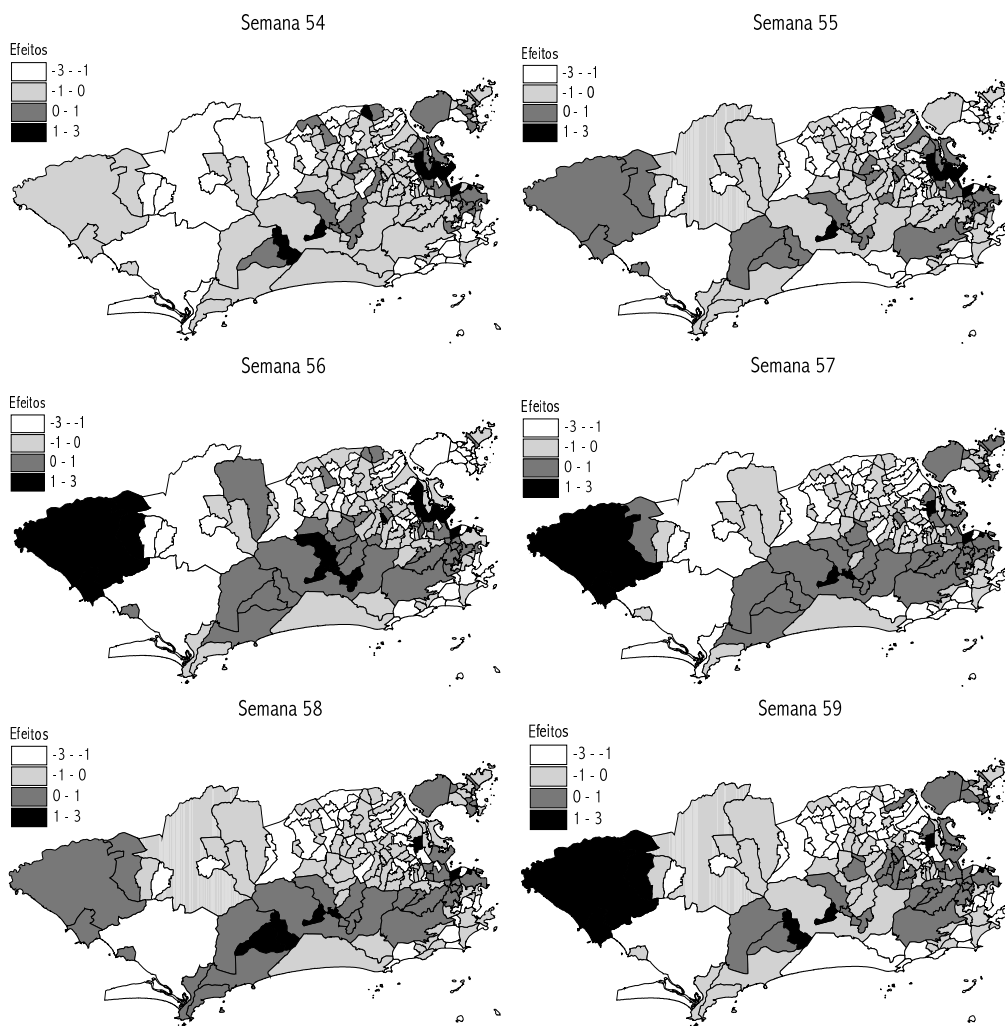


Figura 6.9: Mapas contendo as médias a posteriori dos efeitos espaciais dos bairros do Rio de Janeiro entre as semanas 54 e 59 estimados pelo Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.

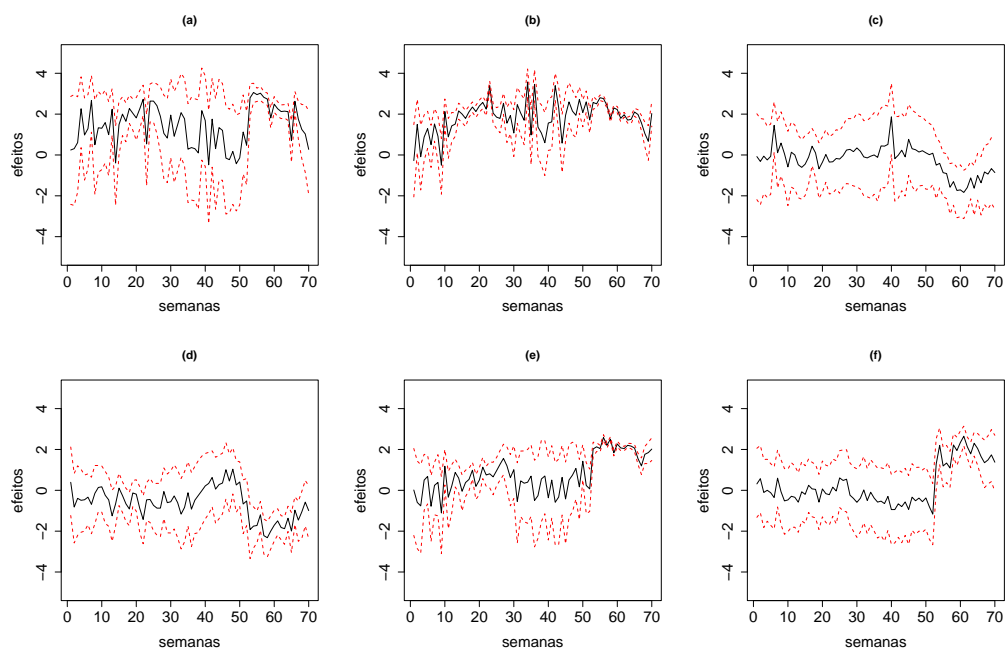


Figura 6.10: Médias a posteriori e respectivos IC95% dos efeitos espaciais dos bairros Saúde (a), Grajaú (b), Quintino Bocaiúva (c), Guadalupe (d), Gardênia Azul (e) e Itanhangá (f) para o Modelo 3 com estrutura dinâmica e Priori II para os efeitos aleatórios.



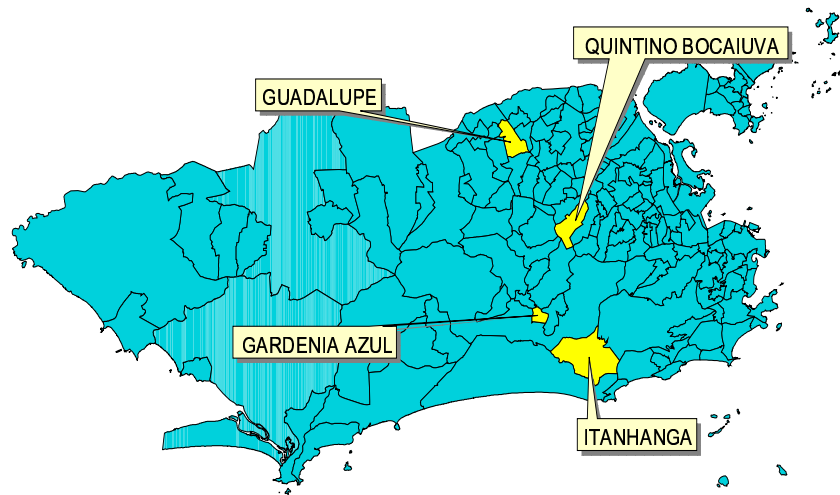


Figura 6.11: Mapa contendo a localização dos bairros Gardênia Azul, Guadalupe, Itanhanga e Quintino Bocaiuva.

## 6.4 Conclusões

Através de uma abordagem bayesiana, estudou-se a distribuição espaço-temporal dos riscos de dengue na cidade do Rio de Janeiro. Os modelos utilizados no capítulo anterior foram generalizados para que os riscos pudessem evoluir no tempo ao longo de 70 semanas, a partir do início do ano de 2001. Desta forma, incorporou-se ao modelo efeitos aleatórios variando no espaço-tempo com o objetivo de captar influências de bairros vizinhos nos riscos de dengue ao longo das semanas. A partir da combinação dos dados com as informações a priori, analisou-se a distribuição a posteriori das quantidades desconhecidas do modelo.

A utilização de uma estrutura dinâmica no efeito do nível da série foi satisfatória, resultando no melhor modelo segundo o critério do DIC.

Diferentemente do resultado obtido na análise espacial, o desempenho

da estrutura de vizinhança binária foi superior às demais. Este resultado pode estar associado à influência das covariáveis, as quais foram utilizadas somente na análise espacial. Apesar de produzir modelos com DIC em patamares mais baixos que aqueles produzidos pelas demais priors, as priors de convolução nos conduziram à estimativas não significativas dos efeitos aleatórios.

Em virtude da variedade de resultados, mais uma vez tornou-se claro a influência das diferentes formas de compor a matriz de pesos  $\mathbf{W}$  ao especificar-se a distribuição a priori CAR para os efeitos aleatórios.

No que diz respeito à evolução dos riscos de dengue, vimos que durante a epidemia, iniciada no final do ano de 2001, os riscos se alastraram pelos demais bairros partindo das zonas norte e leste e, paralelamente, elevando-se no extremo oeste da cidade. Neste período, evidenciou-se o fator espacial como determinante na proliferação da epidemia na cidade.

# Capítulo 7

## Considerações Finais

Neste estudo, analisou-se a distribuição espaço-temporal dos casos de dengue na cidade do Rio de Janeiro a partir de uma abordagem bayesiana.

Utilizando as idéias de Modelos Lineares Generalizados Bayesianos, analisou-se a série de casos mensais de dengue em busca de efeitos cíclicos e de outros possíveis fatores relacionados com os períodos epidêmicos.

Dentre os principais resultados obtidos, podemos destacar a identificação de um padrão sazonal na série de notificações mensais de dengue. Esta identificação foi possível a partir da utilização do modelo contendo efeitos de sobredispersão.

A distribuição espacial do dengue foi investigada a partir da análise dos riscos da doença agregados no tempo. À estes dados foram ajustados modelos utilizando efeitos aleatórios e covariáveis relacionadas à fatores sócio-econômicos e de infra-estrutura dos bairros.

Diferentes formas de especificação das distribuições a priori para os

efeitos aleatórios foram utilizadas. As priors condicionais auto-regressivas (CAR) foram utilizadas a partir de diferentes estruturas de vizinhança propostas neste estudo. Posteriormente, estes modelos foram generalizados para serem utilizados na análise espaço temporal.

Dentre os resultados obtidos nesta análise, destacam-se a significância da influência de diversos fatores sócio-econômicos e de infra-estrutura, além de evidenciarmos a importância de ações de combate ao vírus junto aos pontos estratégicos da cidade. Em virtude da grande variedade de resultados obtidos, evidenciamos também a importância da escolha da estrutura de vizinhança na especificação das distribuições a priori para os efeitos aleatórios dos modelos.

Na análise espaço-temporal, a utilização de uma estrutura dinâmica no efeito do nível da série foi satisfatória, resultando no melhor modelo segundo o critério do DIC.

Por fim, verificou-se uma associação entre a distribuição dos riscos de dengue na cidade e o padrão dos efeitos espaciais estimados pelos modelos. Na análise espaço-temporal, esta característica foi mais evidente durante o período epidêmico.

# Referências Bibliográficas

- [1] Besag, J., York, J. e Mollié A. (1991). Bayesian Image Restoration, with Two Applications on Spatial Statistics. *Annals of the Institute of Statistical Mathematics*, 43, 1-59.
- [2] Breslim, P., Frunzi, N., Napoleon, E. e Ormsby, T. (1999). *Getting to know ArcView GIS*. ESRI Press.
- [3] Cardoso, Jr. R.P., Scandar, S.A.S., Mello, N.V., Ernandes, S., Botti, M.V. e Nascimento, E.M.M. (1997). Detecção de *Aedes aegypti* e *Aedes albopictus* na zona urbana do município de Catanduva-SP, após controle de epidemia de dengue. *Revista da Sociedade Brasileira de Medicina Tropical*, 30(1), 37-40, jan-fev.
- [4] Ferreira, M.A.R. e Gamerman, D. (1998). Análise Bayesiana de Séries Epidemiológicas de Contagem via Modelos Dinâmicos Bayesianos Generalizados. *Cadernos de Saúde Coletiva*, 6(2).
- [5] Ferreira, M.A.R. e Gamerman, D. (2000). Dynamic Generalized Linear Models. Em *Generalized Linear Models: A Bayesian Perspective* (Editors Dey, Ghosh e Mallick), Marcel Dekker, New York.

- [6] Gamerman, D. (1997). *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference*. Chapman and Hall.
- [7] Gelfand, A.E. e Smith, A.F.M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398-409.
- [8] Gelman, A. e Rubin, D.B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 457-472.
- [9] Geman, S. e Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.
- [10] Geweke, J. (1992). Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. Em *Bayesian Statistics 4*, 169-193 (Editors Bernardo, J.M., Berger, J.O., Dawid, A.P. e Smith, A.F.M.).
- [11] Gilks, W.R. e Wild, P. (1992). Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics*, 41(2), 337-348.
- [12] Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57, 97-109.
- [13] Kypraios, T. (2003). Space-time Analysis of Leptospirosis Incidence in the City of Rio de Janeiro. Dissertação de Mestrado. Lancaster University, Inglaterra.
- [14] Lenzi, M.F., Coura, L.C., Gault, C.E. e Val, M.B. (2000). Estudo do dengue em área urbana favelizada do Rio de Janeiro: considerações iniciais. *Cadernos de Saúde Pública*, 16(3), 851-856, jul-set.

- [15] McCullagh, P. e Nelder, J.A. (1989). *Generalized Linear Models* 2nd ed. Chapman and Hall, Londres.
- [16] Medronho, R.A. (1995). *Geoprocessamento e Saúde: uma nova abordagem do espaço no processo saúde-doença*. FIOCRUZ/CICT/NECT.
- [17] Metropolis, N., Rosenbulth, A.W., Rosenbulth, M.N., Teller, A.H. e Teller, E. (1953). Equation of State Calculations by Fast Computing Machine. *Journal of Chemical Physics*, 21, 1089-1091.
- [18] Migon, H.S. e Gamerman, D. (1999). *Statistical Inference - an integrated approach*. Arnold.
- [19] Mollié, A. (1996). Bayesian Mapping of Disease. Em *Monte Carlo Markov Chain in Practice*. W.R. Gilks, S. Richardson and D.J. Spiegelhalter (editors).
- [20] Rebêlo, J.M.M., Costa, J.M.L., Silva, F.S., Pereira, Y.N.O., e Silva, J.M.S. (1999). Distribuição de *Aedes aegypti* e do dengue no Estado do Maranhão, Brasil. *Cadernos de Saúde Pública*, 15(3), 477-486, jul-set.
- [21] Santos, R.S. e Carvalho, M.S. (2000). Análise da distribuição espacial de larvas de *Aedes aegypti* na Ilha do Governador, Rio de Janeiro, Brasil. *Cadernos de Saúde Pública*, 16(1), 31-42, jan-mar.
- [22] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. e Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, B, 64, 583-639.
- [23] Tauil, P.L. (2002). Aspectos críticos do controle do dengue no Brasil. *Cadernos de Saúde Pública*, 18(3), 867-871, mai-jun.

- [24] Teixeira, M.G., Barreto, M.L., Costa, M.C.N., Ferreira, L.D.A. e Vasconcelos, P. (2003). Dinâmica de circulação do vírus do dengue em uma área metropolitana do Brasil. *Epidemiologia e Serviços de Saúde. Brasília, DF.*, 12(2), 87-97.
- [25] Thomas, A., Spiegelhalter, D.J. e Gilks, W.R. (1992). BUGS: A program to perform Bayesian inference using Gibbs Sampling. . Em *Bayesian Statistics 4*, 837-842 (Editors Bernardo, J.M., Berger, J.O., Dawid, A.P. e Smith, A.F.M.).
- [26] West, M. e Harrison, P.J. (1997). *Bayesian Forecasting and Dynamic Models* 2nd ed. Springer Verlag, New York.
- [27] West, M. e Harrison, P.J. and Migon, H.S. (1985). Dynamic generalized linear models and Bayesian forecasting (com discussão). *Journal of the American Statistical Association*, 80, 73-96.



# Apêndice

## Deviance Information Criterion (DIC)

O critério de comparação de modelos DIC (Spiegelhalter *et al.*, 2002) vem sendo amplamente utilizado para comparar modelos com diferentes níveis de complexidade.

Como este critério assume que a média a posteriori é uma boa estimativa dos parâmetros do modelo, nos casos em que estas distribuições são multimodais ou nos casos em que há uma acentuada assimetria, o DIC não é recomendado.

Ao utilizar-se o DIC, estamos computando as seguintes estatísticas:

- $\bar{D}$ : Representa a média a posteriori do *desvio*. O *desvio*, por sua vez, é definido como

$$D = -2 \ln(p(\mathbf{y}|\boldsymbol{\theta})),$$

onde  $p(\mathbf{y}|\boldsymbol{\theta})$  representa a função de verossimilhança.

- $\hat{D}$ : É uma estimativa pontual do *desvio* obtida a partir da utilização da média a posteriori de  $\boldsymbol{\theta}$ , denotada por  $\bar{\boldsymbol{\theta}}$ , na expressão do *desvio*. Ou seja,

$$\hat{D} = -2 \ln(p(\mathbf{y}|\bar{\boldsymbol{\theta}})).$$

O valor de  $\hat{D}$  pode mudar dependendo da parametrização utilizada na especificação das priors do modelo.

- $p_D$ : Representa o número efetivo de parâmetros e é dada por  $p_D = \bar{D} - \hat{D}$ . Para modelos não-hierárquicos, o valor de  $p_D$  pode ser aproximadamente igual ao verdadeiro número de parâmetros.

A partir destas estatísticas, o DIC é definido como:

$$DIC = \bar{D} + p_D = \hat{D} + 2p_D$$

O modelo que apresenta o menor DIC é considerado como aquele que melhor pode prever um novo conjunto de dados com a mesma estrutura dos dados observados.

O DIC é semelhante ao critério clássico de comparação de modelos AIC (Akaike's Information Criterion), mas difere substancialmente do critério BIC e do fator de Bayes.

## Código do WinBugs para Análise Temporal

```
model {
for (t in 1:12) { deng[t] ~ dpois(lambda[t]) log(lambda[t]) < - mu + ano[1]
+ s[t]+phi[t] }
for (t in 13:24) { deng[t] ~ dpois(lambda[t]) log(lambda[t]) < - mu + ano[2]
+ s[t]+phi[t] }
...
for (t in 193:204) { deng[t] ~ dpois(lambda[t]) log(lambda[t]) < - mu +
ano[17] + s[t]+phi[t] }

for (i in 1:16) { a[i] ~ dnorm(0,0.001)
ano[i] < - a[i] }
ano[17] < - -sum(a[1:16])
for (t in 1:11) { s[t] ~ dnorm(0,tau2) }
for (t in 12:204) { k[t-11] < - -sum(s[(t-11):(t-1)])
s[t] ~ dnorm(k[t-11],tau2.w) }
for (t in 1:204) { phi[t] ~ dnorm(0,tau2.phi) }
tau2 ~ dgamma(0.05,0.005)
tau2.phi ~ dgamma(0.05,0.005)
tau2.w ~ dgamma(0.05,0.005)
mu ~ dnorm(0,0.001)
}
```

## Código do WinBugs para Análise Espacial

```
model {  
  b[1:N] ~ car.normal(adj[], weights[], num[], tau)  
  for (i in 1 : N)  
  {  
    deng[i] ~ dpois(mu[i])  
    log(mu[i]) < - log(e[i]) + alfa0 + alfa1*favelas[i]+b[i]+u[i]  
    RR[i] < - exp(alfa0 + alfa1*favelas[i]+b[i]+u[i])  
    u[i] ~ dnorm(0,tau.u)  
  }  
  
  alfa0 ~ dnorm(0,0.001)  
  alfa1 ~ dnorm(0,0.001)  
  tau ~ dgamma(0.5, 0.0005)  
  tau.u ~ dgamma(8,1)  
  sigma2 < - 1 / tau  
  sigma2.u < - 1 / tau.u  
}
```

## Código do WinBugs para Análise Espaço-Temporal

```
model {
  for (i in 1 : N) {
    for (t in 1 : K) {
      deng[i,t] ~ dpois(mu[i,t])
    }
  }
  for (i in 1:N) {
    log(mu[i,1]) <- - log(e[i,1])+ alfa0[1] + b1[i]
    RR[i,1] <- - exp(alfa0[1] + b1[i])
    log(mu[i,2]) <- - log(e[i,2])+ alfa0[2] + b2[i]
    RR[i,2] <- - exp(alfa0[2] + b2[i])
    ...
    log(mu[i,70]) <- - log(e[i,70])+ alfa0[70] + b70[i]
    RR[i,70] <- - exp(alfa0[70] + b70[i])
  }
  b1[1:156] ~ car.normal(adj[], weights[], num[], tau)
  b2[1:156] ~ car.normal(adj[], weights[], num[], tau)
  b70[1:156] ~ car.normal(adj[], weights[], num[], tau)
  alfa0[1] ~ dnorm(0,tau.a)
  for (t in 2 : K) {
    alfa0[t] ~ dnorm(alfa0[t-1],tau.a)
  }
  tau ~ dgamma (0.5, 0.0005)
  tau.a ~ dgamma (0.5, 0.0005)
  sigma2 <- - 1/tau
  sigma2.a <- - 1/tau.a
}
```

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)