



**Universidade Federal do Rio de Janeiro
Instituto de Matemática**

**Modelos para Processos Espaço-Temporais
Inflacionados de Zeros**

Marcus Vinicius Morais Fernandes

Orientadores: Alexandra M. Schmidt e Helio S. Migon

Rio de Janeiro
Fevereiro/2006

Livros Grátis

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

Milhares de livros grátis para download.

MODELOS PARA PROCESSOS ESPAÇO-TEMPORAIS INFLACIONADOS DE ZEROS

Marcus Vinicius Morais Fernandes

Dissertação de Mestrado submetida ao Programa
de Pós-Graduação em Estatística do Instituto de
Matemática da Universidade Federal do Rio de
Janeiro - UFRJ, como parte dos requisitos necessá-
rios à obtenção do título de Mestre em Estatística.

Rio de Janeiro
Fevereiro/2006

**MODELOS PARA PROCESSOS
ESPAÇO-TEMPORAIS INFLACIONADOS DE
ZEROS**

Marcus Vinicius Morais Fernandes

Dissertação de Mestrado submetida ao programa de Pós-Graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro-UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Aprovada por:

Orientadora, Professora Alexandra M. Schmidt - IM/UFRJ

Orientador, Professor Helio S. Migon - IM-COPPE/UFRJ

Professor Dani Gamerman - IM/UFRJ

Professor Josemar Rodrigues - DEs/UFSCar

Rio de Janeiro

Janeiro 2006

Abstract

MODELING ZERO INFLATED SPATIO-TEMPORAL PROCESSES

Marcus Vinicius Morais Fernandes

Orientadores: Alexandra M. Schmidt e Helio S. Migon

We consider models for spatio-temporal processes which assume either non-negative values and often are observed as zero, or assume discrete values but are also inflated by zeros. Typically, in the first case, the spatial observations are obtained at fixed locations (point-referenced data) over a region D ; whereas in the second, the region D is divided into a finite number of regular or irregular subregions (areal level) resulting on observations for each subregion. Our main idea is based on those of zero-inflated models, by assuming that the value observed at location or subregion \mathbf{s} and time t , $Y_t(\mathbf{s})$, is a realization of a mixture between a Bernoulli distribution with a probability of success $\theta_t(\mathbf{s})$ and a probability density function or probability function $p(Y_t(\mathbf{s}) | \cdot)$.

For both continuous and discrete cases, we include in the model, spatio-temporal latent processes to account for possible extra variation present in the mean structure of $\theta_t(\mathbf{s})$ and/or of $p(Y_t(\mathbf{s}) | \cdot)$. One of the main contributions lies in the fact that in the continuous case the observations are modelled in their original scale without the need of considering any transformation to attain normality of the data. Our proposed model is applied for two different examples. In the context of point-referenced data we model the amount of rainfall over the city of Rio de Janeiro during 75 weeks; whereas in the areal data level case, we consider weekly cases of dengue fever in the city of Rio de Janeiro during the summer of 2001-2002; we investigate further the effect of rainfall in the cases of dengue fever using a transfer function model.

Resumo

MODELOS PARA PROCESSOS ESPAÇO-TEMPORAIS INFLACIONADOS DE ZEROS

Marcus Vinicius Morais Fernandes

Orientadores: Alexandra M. Schmidt e Helio S. Migon

Nós consideramos modelos para processos espaço-temporais que tanto assumem valores não-negativos e frequentemente são observados como zero, quanto assumem valores discretos mas também são inflacionados de zeros. Tipicamente, no primeiro caso, as observações são obtidas em localizações fixas em uma região D ; enquanto no segundo, a região D é dividida em um número finito de sub-regiões regulares ou irregulares, resultando em observações para cada sub-região. Nossa idéia principal é baseada nos modelos inflacionados de zeros, assumindo que o valor observado na região ou localização \mathbf{s} e tempo t , $Y_t(\mathbf{s})$, é a realização de uma mistura entre uma distribuição Bernoulli com uma probabilidade de sucesso $\theta_t(\mathbf{s})$ e uma função densidade de probabilidade ou função de probabilidade $p(Y_t(\mathbf{s}) | \cdot)$.

Para ambos casos, contínuos e discretos, nós incluímos no modelo, processos espaço-temporais latentes para considerar possíveis variações presentes na estrutura da média de $\theta_t(\mathbf{s})$ e/ou de $p(Y_t(\mathbf{s}) | \cdot)$. Uma das principais contribuições consiste no fato de que em casos contínuos, as observações são modeladas na escala original, sem a necessidade de considerar qualquer transformação para obter a normalidade dos dados. O nosso modelo proposto é aplicado para três diferentes exemplos. No contexto de Geostatística nós modelamos a quantidade de chuva para a cidade do Rio de Janeiro durante 75 semanas; enquanto para o caso de dados de área, nós consideramos casos semanais de dengue na cidade do Rio de Janeiro durante os anos de 2001-2002; investigamos adiante o efeito da chuva em casos de dengue usando um modelo de função de transferência.

Agradecimentos

Agradeço em primeiro lugar a Deus, por me dar forças a continuar.

Àqueles que não estão mais aqui comigo e que tenho certeza, onde quer que estejam, estão muito contentes em me verem terminando esse Mestrado.

Às minhas três tias: Mariluce, Cândida e Ana, por terem feito e por fazerem, até hoje, de tudo para que minha vida seja um pouco menos complicada, nestes últimos 15 anos. Amo muito vocês.

Aos meus primos, Cláudia, Fernanda, Flávio, André e Carla por também terem me dado suporte em todos esses anos.

À Maria Inês e Dalvinha por toda ajuda que me prestam no dia a dia.

Aos meus orientadores, Alexandra M. Schmidt e Helio S. Migon, pela parceria deste último ano. Foi um prazer trabalhar com dois excelentes profissionais. Agradeço também aos demais professores do Departamento que tenham contribuído de alguma forma para a minha formação e à CAPES por ter financiado meus estudos em um período do Mestrado.

Aos meus amigos da turma de mestrado de 2004, por terem sido tão companheiros nestes últimos tempos, em especial a Adelmo, Carla, Carolina, Gilmar, Marcelo e Valmária.

À Romy e Mariane pelas frequentes colaborações e ao Gustavo por ter cedido uma parte de seu material para minha dissertação.

Enfim, à todos que, mesmo anonimamente, tenham torcido por mim.

Muito Obrigado!

Sumário

| | | |
|----------|-----------------------------------------------------------------------|-----------|
| 1 | Introdução | 1 |
| 2 | Revisão da Literatura | 4 |
| 2.1 | Inferência Bayesiana | 4 |
| 2.1.1 | Estimação de Parâmetros | 6 |
| 2.1.2 | Predição | 7 |
| 2.2 | Modelos Lineares Generalizados | 7 |
| 2.3 | Modelos Dinâmicos | 9 |
| 2.4 | Modelos Espaciais | 11 |
| 2.4.1 | Geoestatística | 12 |
| 2.4.2 | Dados de área | 13 |
| 2.5 | Modelos Espaço-Temporais | 14 |
| 2.6 | Comparação de Modelos | 15 |
| 2.6.1 | <i>Deviance Information Criterion</i> | 15 |
| 2.6.2 | Comparação de Modelos usando a Distribuição Preditiva | 16 |
| 3 | Modelos para Processos Espaço-Temporais Inflacionados de Zeros | 19 |

| | | |
|----------|------------------------------------------------------------------|-----------|
| 3.1 | Modelo Geral | 20 |
| 3.2 | Modelo para Variáveis Mistas | 23 |
| 3.3 | Modelo para Variáveis Discretas | 25 |
| 3.4 | Especificação de Prioris | 27 |
| 3.5 | Interpolação Espacial e Previsão Temporal | 29 |
| 4 | Aspectos Computacionais | 32 |
| 4.1 | Simulação Estocástica | 32 |
| 4.2 | Amostrador de Gibbs | 32 |
| 4.3 | Algoritmo de Metropolis-Hastings | 34 |
| 4.4 | Método da Rejeição Adaptativa com passos de Metropolis | 34 |
| 4.5 | Inferência em Modelos Lineares Dinâmicos Generalizados | 37 |
| 4.5.1 | Exemplos | 40 |
| 5 | Aplicações | 43 |
| 5.1 | Modelando a chuva na cidade do Rio de Janeiro | 43 |
| 5.1.1 | Conjunto de dados | 44 |
| 5.1.2 | Modelos Utilizados | 45 |
| 5.1.3 | Procedimento de Inferência | 47 |
| 5.1.4 | Resultados | 56 |
| 5.2 | Modelando a Dengue na Cidade do Rio de Janeiro | 71 |
| 5.2.1 | Conjunto de dados | 72 |
| 5.2.2 | Modelos Utilizados | 72 |

| | | |
|----------|------------------------------------------------------------------------------------------|------------|
| 5.2.3 | Procedimento de Inferência | 74 |
| 5.2.4 | Resultados | 78 |
| 5.3 | Investigando o Efeito da Chuva nos Casos de Dengue na Cidade do Rio de Janeiro | 92 |
| 5.3.1 | Conjunto de dados | 92 |
| 5.3.2 | Modelo utilizado | 93 |
| 5.3.3 | Procedimento de Inferência | 95 |
| 5.3.4 | Resultados | 97 |
| 5.4 | Considerações Finais | 100 |
| 6 | Conclusões | 102 |
| | Referências Bibliográficas | 105 |
| A | Localização Geográfica das Estações Monitoradoras da Geo-Rio | 110 |
| B | Cálculo das Distribuições Condicionais Completas a Posteriori | 112 |
| C | Cálculos para o Modelo de Função de Transferência | 118 |

Capítulo 1

Introdução

Recentemente o estudo de modelos espaço-temporais tem sido bastante difundido em razão da grande demanda de informações a nível espacial e temporal. Os avanços tecnológicos têm contribuído para isto, principalmente quando se trabalha sob o enfoque da estatística Bayesiana, além da facilidade no armazenamento de grandes bancos de dados.

Diversas áreas de pesquisa usufruem destes modelos a fim de obter respostas cada vez mais precisas sobre os fenômenos que estudam. Em estudos meteorológicos, pode-se tentar explicar o comportamento pluviométrico de um determinado mês em uma determinada região. Podemos ver como se distribuem os casos de homicídios nos bairros de uma cidade, durante um determinado ano. Também, a evolução de uma determinada doença, durante seu período de epidemia, em uma determinada área da cidade.

Em particular, casos como estes, podem ter muitas observações assumindo o valor 0 (zero). Isto faz com que haja uma grande heterogeneidade nos dados, e, portanto, uma grande sobredispersão.

Sob a presença de sobredispersão, a modelagem usual dos dados, não é certamente a melhor escolha. É necessário, portanto, a utilização de modelos que levem em conta a informação de inflação de zeros. Esta inflação de zeros pode ocorrer sob

dois enfoques distintos: quando a quantidade de zeros existente é superior àquela esperada pelo pesquisador, ou, simplesmente, quando observamos quantidades puramente não-negativas.

O presente trabalho tem como objetivo expor uma nova classe de modelos que agreguem especialmente estas duas características: dependência espaço-temporal e excesso de zeros. Estes serão chamados de modelos espaço-temporais inflacionados de zeros. São baseados na idéia originalmente proposta por Velarde, Migon, and Pereira (2004).

Velarde et al. (2004) utilizaram uma modelagem espaço-temporal para analisar o comportamento pluviométrico no estado de Goiás. Expandiram para o contexto espaço-temporal os modelos propostos por Lambert (1992), que foi a pioneira na utilização de modelos de regressão com covariáveis em mistura de distribuições, em especial colocando massa no ponto zero, usando as distribuições Bernoulli e Poisson. Estes modelos ficaram bastante conhecidos, como *Zero Inflated Poisson (ZIP) models*.

Em seu trabalho, Velarde et al. (2004) colocaram massa de probabilidade no ponto zero, afim de se modelar, também, períodos secos. Para os demais valores utilizaram a distribuição Exponencial. Ou seja, a variável de interesse (chuva) é modelada por uma mistura de distribuições: uma Bernoulli e uma Exponencial. Foram utilizados Efeitos Auto-Regressivos Condicionais (CAR) para a modelagem dos efeitos espaço-temporais.

Aqui, generalizamos estas idéias. Pois também trabalharemos com misturas de distribuições, mas utilizaremos, tanto distribuições contínuas (criando uma variável mista) quanto distribuições discretas. Na âmbito da modelagem espacial, trabalharemos tanto com dados de área quanto com observações distribuídas continuamente no espaço, permitindo no caso de variáveis contínuas, diferentemente de Velarde et al. (2004), previsões para quaisquer localizações no espaço (medidas ou não).

No Capítulo 2, faremos uma breve revisão da literatura, abordando conceitos de modelagem estatística, desde os modelos lineares até os modelos espaço-temporais.

No Capítulo 3, introduziremos toda a conceituação dos modelos espaço-temporais inflacionados de zeros, distinguindo-os em duas classes: Modelos Contínuos e Modelos Discretos.

Como o procedimento de inferência seguirá o Paradigma de Bayes, no Capítulo 4, trataremos de todos os aspectos computacionais ligados a esse procedimento, os métodos de simulação estocástica e algoritmos computacionais.

Já no Capítulo 5, apresentaremos algumas aplicações para estes modelos. Uma relacionada a dados contínuos usando observações de quantidade de chuva na cidade do Rio de Janeiro; outra para dados discretos, desta vez utilizando notificações de casos de dengue, também na cidade do Rio de Janeiro; e a última que seria outra modelagem da dengue no Rio de Janeiro, mas desta vez utilizando a chuva como variável explicativa, com o uso de funções de transferência.

O Capítulo 6 apresenta as conclusões e considerações finais desse estudo.

Capítulo 2

Revisão da Literatura

Neste Capítulo faremos, brevemente, uma revisão de todos os conceitos que serão abordados ao longo deste trabalho. Começaremos pela inferência Bayesiana e depois falaremos de várias classes de modelos existentes na literatura estatística.

2.1 Inferência Bayesiana

Nesta Seção serão descritos os principais conceitos necessários para se fazer inferência estatística sob o enfoque Bayesiano.

Assim como na inferência frequentista paramétrica, a inferência Bayesiana trabalha na presença de observações \mathbf{Y} , cujos valores são descritos através de uma distribuição de probabilidades com densidade ou função de probabilidade $p(\mathbf{Y} | \theta)$. A quantidade θ é desconhecida, podendo ser um escalar, um vetor ou uma matriz. Para uma leitura extensiva sobre o assunto veja, Migon and Gamerman (1999).

Podemos dizer que o principal objetivo de um estudo de inferência é a determinação do valor de θ . Entretanto, é bastante provável que o pesquisador saiba caracterizá-lo. Neste caso, é possível e recomendável, que esse conhecimento prévio a respeito da quantidade, seja incorporado à análise, e é nesse ponto que o método

Bayesiano se diferencia do frequentista. Enquanto o segundo não admite o uso dessa informação por não ser observável, o primeiro incorpora à análise através de uma densidade $p(\theta)$. Esta densidade, conhecida como densidade a priori, e a função de verossimilhança, são os ingredientes da inferência Bayesiana.

A função de verossimilhança de θ é a função que associa o valor $p(\mathbf{Y} | \theta)$ para cada θ . Essa função é geralmente denotada por $l(\theta; \mathbf{Y})$ e, considerando Θ como o espaço paramétrico, podemos defini-la da seguinte maneira:

$$\begin{aligned} l(\cdot; \mathbf{Y}) : \Theta &\rightarrow \mathfrak{R}^+ \\ \theta &\rightarrow l(\theta; \mathbf{Y}) = p(\mathbf{Y} | \theta). \end{aligned}$$

A densidade a priori, $p(\theta)$, contém a distribuição de probabilidade de θ antes de se observar o valor de \mathbf{Y} . Desta forma, é razoável que o processo de inferência se baseie na distribuição de probabilidade de θ , após observar o valor de \mathbf{Y} , que passa a fazer parte do conjunto de informação disponível. Esta distribuição é conhecida como distribuição a posteriori de θ .

Obtemos, portanto, uma regra de atualização de probabilidades, começando com a priori $p(\theta)$ e terminando com a posteriori $p(\theta | \mathbf{Y})$, que pode ser obtida via teorema de Bayes da seguinte forma:

$$p(\theta | \mathbf{Y}) = \frac{p(\mathbf{Y} | \theta)p(\theta)}{p(\mathbf{Y})},$$

onde:

$$p(\mathbf{Y}) = \int_{\Theta} p(\mathbf{Y}, \theta) d\theta = \int_{\Theta} p(\mathbf{Y} | \theta)p(\theta) d\theta.$$

Como podemos notar, $p(\mathbf{Y})$ não depende de θ e, portanto, pode ser considerada constante. Desta maneira pode-se reescrever o teorema de Bayes, em sua forma mais usual:

$$p(\theta | \mathbf{Y}) \propto p(\mathbf{Y} | \theta)p(\theta).$$

A fórmula acima é válida para quantidades discretas, contínuas, escalares, vetores e matrizes.

2.1.1 Estimação de Parâmetros

Depois de obtida a distribuição a posteriori, é necessária a estimação do parâmetro θ . Um estimador, δ , pode ser definido como uma regra de decisão ótima a respeito de uma dada função perda esperada, $L(\delta, \theta)$. A estimação pode ser pontual ou intervalar.

Estimação Pontual

Na metodologia Bayesiana a estimação pontual lida sempre com a minimização de uma função perda, $L(\delta, \theta) = h(\delta - \theta)$, para alguma função h e estimador $\delta = \hat{\theta}$. Três funções perda, são geralmente utilizadas, (i) a função perda quadrática, $L(\delta, \theta) = (\delta - \theta)^2$, (ii) a função perda absoluta, $L(\delta, \theta) = |\delta - \theta|$ e (iii) a função perda zero-um, $L(\delta, \theta) = \lim_{\epsilon \rightarrow 0} I_{|\theta - \delta|}(|\epsilon, \infty|)$. Os estimadores obtidos com a minimização destas funções são, respectivamente:

- (i) média a posteriori: $\hat{\theta} = E(\theta | \mathbf{Y})$;
- (ii) mediana a posteriori: $\hat{\theta} : \int_{-\infty}^{\hat{\theta}} p(\theta | \mathbf{Y}) d\theta = 0.5$;
- (iii) moda a posteriori: $\hat{\theta} : p(\hat{\theta} | \mathbf{y}) = \sup_{\theta} p(\theta | \mathbf{Y})$.

Estimação por Intervalo

Às vezes sintetizar toda a informação da distribuição posteriori em um único valor, pode não ser apropriado. É necessário se obter, também, informações sobre a precisão deste valor. Portanto, uma alternativa seria a utilização de intervalos de credibilidade, onde pretende-se concentrar maior massa de probabilidade em um pequeno intervalo para θ .

Desta forma seja θ , uma quantidade desconhecida definida em Θ . Uma região em $C \subset \Theta$ é uma região com $100(1 - \alpha)\%$ de credibilidade para θ , se $P(\theta \in C | \mathbf{Y}) \geq 1 - \alpha$. E neste caso, $1 - \alpha$, é chamado de nível de credibilidade. Desta forma

obtemos um intervalo de 95% de credibilidade, calculando diretamente os quantis 2,5% e 97,5% da posteriori de θ , isto é:

$$\int_{-\infty}^{\hat{\theta}} p(\theta | \mathbf{Y})d\theta = 0.025 \quad \text{e} \quad \int_{-\infty}^{\hat{\theta}} p(\theta | \mathbf{Y})d\theta = 0.975.$$

2.1.2 Predição

Outro aspecto importante que se segue, é o de se obter um meio de fazer previsões para observações futuras. Se nós queremos prever $\tilde{\mathbf{Y}}$, baseados no vetor \mathbf{Y} , cuja função de probabilidade é dada por $p(\tilde{\mathbf{Y}} | \theta)$, devemos obter a distribuição preditiva, de maneira que:

$$\begin{aligned} p(\tilde{\mathbf{Y}} | \mathbf{Y}) &= \int_{\Theta} p(\tilde{\mathbf{Y}}, \theta | \mathbf{Y})d\theta \\ &= \int_{\Theta} p(\tilde{\mathbf{Y}} | \theta, \mathbf{Y})p(\theta | \mathbf{Y})d\theta \\ &= \int_{\Theta} p(\tilde{\mathbf{Y}} | \theta)p(\theta | \mathbf{Y})d\theta, \end{aligned}$$

onde a última igualdade vem da independência entre $\tilde{\mathbf{Y}}$ e \mathbf{Y} , uma vez que θ é conhecido. Segue também da última equação que:

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

2.2 Modelos Lineares Generalizados

Os modelos lineares generalizados são uma extensão dos modelos lineares clássicos. Um modelo linear é baseado em um vetor de observações \mathbf{Y} com n componentes, que são uma realização de uma variável aleatória \mathbf{Y} cujo componentes são independentemente distribuídas com médias $\boldsymbol{\mu}$. Um modelo linear pode ser descrito como:

$$\mathbf{Y} = \boldsymbol{\mu} + \boldsymbol{\epsilon}. \tag{2.1}$$

A parte sistemática do modelo é uma especificação para $\boldsymbol{\mu}$ em função de um número pequeno de parâmetros, β_1, \dots, β_p . Tal especificação é feita da seguinte

forma:

$$\mu_i = \sum_{j=1}^p x_{ij}\beta_j \quad i = 1, \dots, n.$$

Ou na forma matricial,

$$E(\mathbf{Y}) = \boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta},$$

onde \mathbf{X} é uma matriz $n \times p$, com as covariáveis ou regressoras do modelo.

Para a parte aleatória, supõe-se independência e variância constante dos erros. Em um modelo linear clássico, tem-se que:

$$\boldsymbol{\epsilon} \sim N(0, \sigma^2 I).$$

Portanto o Modelo Linear Clássico pode ser resumido na forma:

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2 I) \tag{2.2}$$

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \tag{2.3}$$

$$Var(\mathbf{Y}) = \sigma^2 I. \tag{2.4}$$

A generalização dos modelos lineares inclui a especificação de três principais aspectos:

(i) As componentes de \mathbf{Y} tem distribuição normal com variância constante e são independentes.

(ii) Na parte sistemática, as covariáveis, x_1, x_2, \dots, x_p , produzem um preditor linear η , dado por:

$$\eta = \sum_{j=1}^p x_j \beta_j.$$

(iii) A ligação entre as componentes sistemática e aleatória é feita através de uma função, de tal maneira que:

$$\boldsymbol{\mu} = \boldsymbol{\eta}.$$

Os modelos lineares generalizados ou MLG's, permitem duas extensões.

A primeira extensão está na função de ligação, que é a parte do modelo que determina a relação entre a média da variável resposta e as covariáveis. A função de ligação, agora, poderá ser qualquer função monótona diferenciável e geralmente é denotada por $g(\boldsymbol{\mu})$.

A outra extensão reside na distribuição especificada para a componente aleatória. Nos MLG's esta pode ser da família exponencial, da qual a distribuição normal faz parte.

Assume-se que, se \mathbf{Y} tem uma distribuição na família exponencial, para $a(\cdot), b(\cdot)$ e $c(\cdot)$ específicos, ela assume a seguinte forma:

$$f_Y(Y; \eta, \phi) = \exp \left\{ \frac{Y\eta - b(\eta)}{a(\phi)} + c(Y, \phi) \right\} \quad (2.5)$$

O parâmetro ϕ é chamado de parâmetro de dispersão e se é conhecido, chamamos essa família de família exponencial linear de parâmetro canônico η . Utilizando (2.5) e algumas relações, pode-se obter expressões para a média e variância de \mathbf{Y} , de modo que:

$$E(\mathbf{Y}) = b'(\eta) \quad (2.6)$$

$$Var(\mathbf{Y}) = a(\phi)b''(\eta). \quad (2.7)$$

2.3 Modelos Dinâmicos

Frequentemente há o interesse de se modelar fenômenos que se caracterizam por uma evolução temporal. Ou seja, as observações não são mais estáticas, elas variam no tempo. É necessário, portanto, que se incorpore esta característica ao modelo, fazendo com que os parâmetros possuam uma estrutura de evolução temporal. Os MLG's não comportam este tipo de especificação e por isso introduziremos uma nova classe de modelos chamada de Modelos Lineares Dinâmicos (MLD). Para uma leitura mais extensiva sobre o assunto, veja West and Harrison (1997).

Um MLD, como é comumente conhecido, é caracterizado por um par de

equações, chamadas de equação das observações e equação de evolução, que são representadas, respectivamente por:

$$Y_t = F_t' \boldsymbol{\theta}_t + \epsilon_t, \quad \epsilon_t \sim N(0, V_t) \quad (2.8)$$

$$\boldsymbol{\theta}_t = G_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t \quad \mathbf{w}_t \sim N(0, W_t), \quad (2.9)$$

onde Y_t é uma sequência de observações no tempo, condicionalmente independentes dado a sequência de parâmetros $\boldsymbol{\theta}_t$, F_t é uma matriz de valores conhecidos das variáveis independentes, $\boldsymbol{\theta}_t$ é um vetor de dimensão $p \times 1$ e G_t é uma matriz de dimensão $p \times p$ que descreve a evolução temporal dos parâmetros. V_t e W_t são as variâncias dos erros associados às observações e ao vetor de parâmetros, respectivamente.

A inferência nesta classe de modelos pode ser feita de maneira sequencial, seguindo o paradigma bayesiano. Se $(F, G, V, W)_t$ são conhecidos utiliza-se o Filtro de Kalman, que fornece a distribuição condicional de $\boldsymbol{\theta}_t$, dada a informação disponível até o tempo t , D_t , de maneira eficiente.

Assim como nos Modelos Lineares Generalizados, podemos também estender os conceitos dos Modelos Lineares Dinâmicos para os Modelos Lineares Dinâmicos Generalizados (MLDG). Tal aplicação foi introduzida por West, Harrison, and Migon (1985). Novamente é permitido que as observações sigam qualquer distribuição da família exponencial. A equação das observações, descrita em (2.8) será agora substituída por:

$$p(Y_t | \eta_t) \propto \exp \left\{ \frac{Y_t \eta_t - b(\eta_t)}{a(\phi_t)} \right\} \quad (2.10)$$

$$g(\mu_t) = F_t' \boldsymbol{\theta}_t \quad (2.11)$$

onde, $\mu_t = E[Y_t | \eta_t] = b'(\eta_t)$. O modelo se completa com a inclusão da equação de evolução, já descrita em (2.9).

O procedimento de inferência é relativamente complexo, e precisa ser feito de forma aproximada. Estimadores usando *Linear Bayes* foram utilizados por West

et al. (1985). Porém há outros tipos de abordagens que serão descritos, posteriormente, no Capítulo 4.

2.4 Modelos Espaciais

Métodos envolvendo estatística espacial vêm sendo muito aplicados desde a última década, com a publicação de Cressie (1993). Um fator preponderante nesta procura é o grande desenvolvimento de Sistemas de Informações Geográficas (SIG), que têm facilitado a criação de bases de dados georeferenciados. Entretanto o avanço computacional e, por consequência, de técnicas envolvendo Cadeias de Markov, no contexto bayesiano, aumentaram ainda mais o interesse de pesquisadores na área. Boas referências para leitura são Schmidt, Nobre, and Ferreira (2002) e Banerjee, Carlin, and Gelfand (2004).

Os modelos espaciais, utilizam a estrutura espacial dos dados, com intuito de trazer maior precisão nas estimativas. A referência espacial deverá ser incorporada, ao modelo, de acordo com o tipo de observação associada ao espaço. Basicamente, nós temos três tipos de dados espacialmente referenciados:

(i) *dados de superfície aleatória*: onde $Y(\mathbf{s})$ é um vetor aleatório na localização $\mathbf{s} \in \mathbb{R}^p$, onde \mathbf{s} varia continuamente sobre D , um subconjunto fixo de \mathbb{R}^p que contém um retângulo p -dimensional de volume positivo. Compreende uma área da estatística espacial chamada geostatística;

(ii) *dados de área*: aonde D é novamente um subconjunto fixo de \mathbb{R}^p (com forma regular ou irregular), mas agora particionado em um número finito de áreas com vizinhanças bem definidas;

(iii) *dados de processos pontuais*: aonde o dado aleatório de interesse é a própria localização espacial do evento.

Descreveremos mais especificamente os dois primeiros casos.

2.4.1 Geoestatística

Em geostatística a localização \mathbf{s} varia continuamente sobre $D \subset \mathbb{R}^p$. Aqui nós assumimos que a covariância entre variáveis aleatórias em duas localizações, depende da distância entre as localizações. Para descrever essa dependência espacial, usualmente assume-se que uma variável aleatória Z segue um processo Gaussiano.

Um processo estocástico é dito gaussiano se essa variável aleatória segue uma distribuição normal multivariada, para quaisquer localizações $s, s' \in D$, com parâmetros dados por $E(Z(s)) = \mu(s)$ e $cov\{Z(s), Z(s')\} = c(s, s')$.

O processo é estacionário se $E(Z(s))$ é a mesma para todo s e $c(s, s')$ depende somente de $s - s'$. É também considerado isotrópico e, portanto, homogêneo, se essa correlação só depende de $\|s - s'\|$, a distância euclidiana entre os pontos.

Quando o processo é homogêneo, sua variância é constante ao longo de D , isto é, $V(Z(s)) = \sigma^2 \forall s \in D$. Dessa forma, a função de covariância pode ser reescrita como $c(s, s') = \sigma^2 \rho(\|s - s'\|; \phi)$, onde $\rho(\cdot)$ denota uma função de correlação válida, ou seja, uma função que faça da matriz de covariância uma matriz positiva definida. Existem algumas funções de correlação na literatura. Os dois principais tipos são:

(a) Família Exponencial Potência:

$$\rho(d; \phi) = \exp(-\phi d^\lambda), \quad (2.12)$$

onde ϕ é parâmetro de escala e d é a distância euclidiana entre dois pontos quaisquer em D .

(b) Família Matérn:

$$\rho(d; \phi, \lambda) = \frac{1}{2^{\lambda-1} \Gamma(\lambda)} (2\phi\sqrt{\lambda d}) k_\lambda (2\phi\sqrt{\lambda d}), \quad (2.13)$$

onde $\phi > 0$ é o parâmetro de escala e λ é o parâmetro de forma. A função $\Gamma(\cdot)$ é a função Gama usual e k_λ é a função modificada de Bessel do terceiro tipo de ordem λ . Em especial, utilizaremos a função de correlação exponencial, obtida assumindo, $\lambda = 1$, em (2.12) ou $\lambda = 0,5$ em (2.13).

2.4.2 Dados de área

Em modelos para dados de área, as regiões geográficas são subdivididas em um número finito de regiões, denotadas por D_i , $i = 1, \dots, n$ e os dados são geralmente somas ou médias de variáveis sobre esse local. A possibilidade de uma resposta ocorrer entre localizações é excluída. Os modelos devem especificar, de alguma maneira, que o valor obtido para área i é influenciado pelos valores das regiões vizinhas. Desta maneira, devemos introduzir um critério de vizinhança para as regiões.

Supondo que temos n subregiões, podemos usar a idéia de modelos autoregressivos temporais de primeira ordem. Dois modelos, bastante conhecidos, que possuem uma estrutura de vizinhança, são os Modelos Autoregressivos Simultâneos (SAR) e os Modelos Autoregressivos Condicionais (CAR). Neste trabalho, estaremos apenas interessados neste último.

O Modelo CAR, originalmente introduzido por Besag (1974), especifica a distribuição condicional do processo na área i dados os vizinhos, como:

$$(Z_i | Z_j, j \neq i) \sim N(\mu_i + \sum_{j=1}^n c_{ij}(Z_j - \mu_j), \sigma^2), \quad (2.14)$$

onde σ^2 é a variância condicional e c_{ij} são constantes conhecidas ou desconhecidas, tais que $c_{ii} = 0$ para $i = 1, \dots, n$.

Um jeito de se construir $C = (C_{ij})$ é fazer $C = \rho W$; ρ denota o parâmetro de correlação espacial e W uma matriz de vizinhança. Usaremos $\rho = 1$ e para W existem algumas especificações (Schmidt and Ferreira (2006) utilizam diferentes estruturas para essa matriz), mas trabalharemos com a seguinte estrutura:

$$W_{ij} = \begin{cases} 1 & \text{se } i \sim j \\ 0 & \text{caso contrário} \end{cases}$$

onde $i \sim j$ denota i vizinho de j .

Com a parametrização acima, obtemos, então, uma distribuição imprópria. Uma forma de se evitar tal situação, é criar uma restrição ao modelo, em que os

efeitos, devem somar zero. De um modo geral, temos que a priori para os efeitos espaciais, é tal que:

$$(Z_i | Z_j, j \neq i) \sim N(\mu_i, v_i), \quad (2.15)$$

onde:

$$\mu_i = \frac{\sum_{j \in \delta_i} Z_j}{n_i}, \quad v_i = \frac{\sigma^2}{n_i},$$

com n_i denotando o número e δ_i o conjunto, de vizinhos de i . Deste modo, a distribuição priori conjunta para \mathbf{Z} é dada por:

$$(\mathbf{Z} | \sigma^2) \propto \frac{1}{\sigma^{2n}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j \leq i} W_{ij} (Z_i - Z_j) \right\} \quad (2.16)$$

2.5 Modelos Espaço-Temporais

Nos últimos anos, a modelagem espaço-temporal vêm recebendo especial atenção. O motivo pelo interesse em tais tipos de modelos está na abundância de dados que são tanto indexados pelo tempo, quanto pelo espaço.

A dimensão tempo, poderá ser de natureza contínua, assumindo valores na reta real, e até mesmo intervalos dentro dela, ou poderá assumir valores discretizados. Estaremos interessados especificamente neste último caso, onde t pode ser, por exemplo, um dia, um mês ou uma semana ($t = 1, \dots, T$).

Como feito na Seção anterior, distinguimos dois tipos de efeitos espaço-temporais, baseados na sua estrutura espacial. Novamente para dados distribuídos continuamente numa região utiliza-se processos Gaussianos e para dados de área, usamos efeitos condicionais autoregressivos.

A partir de agora, consideraremos $Y_t(s)$ um campo aleatório espacial, no tempo t e na localização s . Se observamos dados com localizações s_i, i, \dots, n e tempos $t = 1, \dots, T$, podemos alocá-los em uma matriz, digamos Y , de dimensão $T \times n$, onde cada coluna da matriz representa uma série temporal para a localização s_i e

cada linha representa observações de uma determinada área ou ponto, no tempo t .

(Banerjee et al., 2004), descrevem um modelo geral, assumindo o caso Gaussiano, da seguinte maneira:

$$Y_t(s) = \mu_t(s) + \epsilon_t(s), \quad (2.17)$$

onde $\mu_t(s)$ denota a estrutura de média e $\epsilon_t(s)$, o erro espacial e temporalmente estruturado. Se associarmos um vetor de covariáveis $F_t(s)$ a $Y_t(s)$, podemos reescrever $\mu_t(s)$, como $\mu_t(s) = \theta_t(s)F_t'(s)$.

De maneira análoga à Seção 2.2, estes modelos podem ser generalizados para distribuições pertencentes à família exponencial. Os modelos descritos anteriormente servirão como base para os modelos propostos no Capítulo 3.

2.6 Comparação de Modelos

A evolução da estruturação dos modelos acarreta ao pesquisador um processo de tomada de decisão, para a escolha do modelo a ser usado, ou seja, aquele que se apresenta mais adequado ao estudo. A seguir descreveremos alguns critérios que serão adotados ao longo deste trabalho, para comparar os diversos modelos propostos ao ajuste das observações no Capítulo 5.

2.6.1 *Deviance Information Criterion*

Spiegelhalter, Best, Carlin, and Linde (2002) propuseram um método de comparação de modelos com o objetivo de superar as dificuldades encontradas na utilização de certos critérios, como o AIC (Akaike Information Criterion), em modelos complexos e com estrutura hierárquica. Este critério, chamado de *Deviance Information Criterion* (DIC), baseia-se na distribuição a posteriori da estatística deviance, isto é:

$$D(\theta) = -2 \log\{p(Y | \theta)\} + 2 \log\{f(Y)\}, \quad (2.18)$$

onde $f(Y)$ denota a distribuição marginal da variável Y .

Este método combina uma parte que mede o ajuste do modelo em si, com outra que avalia a complexidade do modelo em questão. O DIC pode ser calculado como:

$$DIC = \bar{D} + p_D, \quad (2.19)$$

onde \bar{D} é a esperança a posteriori da estatística deviance, $E_{\theta|y}[D]$, e é justamente a parte que avalia o ajuste do modelo. A componente p_D , pode ser interpretada como sendo o número efetivo de parâmetros no modelo, que mede a complexidade do mesmo. Podemos calculá-lo da seguinte forma:

$$p_D = \bar{D} - \hat{D}, \quad (2.20)$$

onde $\hat{D} = -2 \log\{p(Y | \bar{\theta})\}$, ou seja, uma estimativa pontual da deviance, utilizando a média a posteriori de θ , $\bar{\theta}$. Portanto o DIC também pode ser reescrito, como:

$$DIC = 2\bar{D} - \hat{D} \quad (2.21)$$

Os melhores ajustes são obtidos com os modelos que possuem os menores valores de DIC.

2.6.2 Comparação de Modelos usando a Distribuição Preditiva

Uma alternativa para esta escolha de modelos é a utilização de métodos que fazem esta comparação via distribuições preditivas. Ou seja, considera-se o melhor modelo, aquele que faz as melhores previsões para os valores observados.

Gelfand and Ghosh (1998), propuseram um método, em que a posteriori de uma função perda é minimizada. Na prática, calcula-se uma medida que é baseada na replicação dos dados observados e considera-se duas quantidades, uma de bondade de ajuste, e uma de capacidade preditiva dos modelos. Este critério é conhecido

como EPD e é obtido através da estatística $D(M)$, cuja fórmula está descrita em (2.24).

$$P(M_j) = \sum_{t=1}^T \sum_{i=1}^n \text{Var}(Y_{novo,t}(s_i)), \quad (2.22)$$

$$G(M_j) = \sum_{t=1}^T \sum_{i=1}^n (E(Y_{novo,t}(s_i)) - Y_{obs,t}(s_i))^2, \quad (2.23)$$

$$D(M_j) = P(M_j) + \frac{1}{2}G(M_j). \quad (2.24)$$

Onde $Y_{obs,t}(s)$, denota o valor observado para a localização s e instante t , $Y_{novo,t}(s)$, denota o valor predito para a localização s e instante t e M_j é o modelo utilizado.

Waller, Carlin, Xia, and Gelfand (1997) utilizam também a distribuição preditiva para a comparação de modelos espaço-temporais. Portanto, trabalharemos com a distribuição preditiva descrita como:

$$p(Y_{novo} | Y_{obs}) = \int p(Y_{novo} | \Theta)p(\Theta | Y_{obs})d\Theta, \quad (2.25)$$

onde Θ é o conjunto de parâmetros do modelo estudado. Para um dado modelo M_j , teremos:

$$p(Y_{novo} | Y_{obs}, M_j) = \int p(Y_{novo} | \Theta^{(i)}, M_j)p(\Theta^{(i)} | Y_{obs}, M_j)d\Theta^{(j)}. \quad (2.26)$$

Waller et al. (1997) propõem comparar diversos modelos, computando-se uma função de discrepância $d(Y_{novo}, Y_{obs})$, tal que:

$$E[d(Y_{novo}, Y_{obs}) | Y_{obs}, M_j].$$

O modelo selecionado será aquele que minimiza a expressão acima. Uma proposta para esta função de discrepância, quando estamos trabalhando com Modelos Lineares Generalizados, seria a função deviance. Por exemplo, para um modelo-espaço temporal, utilizando a distribuição Gama, teríamos:

$$d(Y_{novo}, Y_{obs}) = 2 \sum_{t=1}^T \sum_{i=1}^n \left\{ -\log \left(\frac{Y_{obs,t}(s_i)}{Y_{novo,t}(s_i)} \right) + \frac{Y_{obs,t}(s_i) - Y_{novo,t}(s_i)}{Y_{novo,t}(s_i)} \right\}. \quad (2.27)$$

Baseando-se também na idéia de Gelfand and Ghosh (1998), podemos calcular o Erro Quadrático Médio, para cada modelo ajustado (M_i). Considerando o contexto de modelos espaço-temporais, teremos:

$$EQM(M_j) = \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n (E(Y_{novo,t}(s_i) | Y_{obs,t}(s_i)) - Y_{obs,t}(s_i))^2. \quad (2.28)$$

Capítulo 3

Modelos para Processos

Espaço-Temporais Inflacionados de Zeros

Como dito anteriormente, a prática de modelagem de dados que possuem uma evolução temporal e são espacialmente referenciados, está bastante disseminada hoje em dia. Este trabalho tem como objetivo lidar com dados que apresentam esta característica e que, além disso, possuem em excesso, valores medidos iguais a 0 (zero).

Tais exemplos podem ser observados, em estudos epidemiológicos, por exemplo, onde temos a contagem da ocorrência de uma doença em uma determinada região, durante um período específico. Este período pode agregar épocas onde há um maior número de casos e épocas onde o número de casos observados da doença é menor, como no caso de doenças epidêmicas.

Outros exemplos podem ser encontrados na observação de fenômenos meteorológicos, em estudos socio-econômicos, na agronomia e em outras diversas áreas.

O excesso de zeros nos dados pode ocasionar um problema de sobredispersão,

pois estes apresentam uma variabilidade muito maior do que a distribuição, que se pretende utilizar, pode prever.

Para tentar contornar este problema, a idéia, utilizada aqui, é a de incluir massa de probabilidade no ponto zero, inflacionando suas possibilidades de existir no modelo. Isso pode ser feito através de uma mistura de distribuições, onde, em uma primeira etapa, utiliza-se uma distribuição Bernoulli com probabilidade $1 - \theta$ de se obter um valor igual a zero. Depois, teremos com probabilidade θ , um valor que virá da distribuição $p(Y | \cdot)$.

Neste Capítulo, introduziremos um modelo geral para lidar com tal situação e depois dividiremos o problema em duas seções: a modelagem de variáveis mistas e a modelagem de variáveis discretas.

3.1 Modelo Geral

Como estamos trabalhando com modelos espaço-temporais, teremos $\{Y_t(s) : s \in D \subset R^2; t = 1, 2, \dots\}$ um campo aleatório espacial no tempo t e localização s . Considerando o excesso de zeros nos dados, modelamos a distribuição de probabilidades para $Y_t(s)$, da seguinte forma:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} (1 - \theta_t(s)) + \theta_t(s)p(Y_t(s) = 0 | \lambda_t(s)) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) \neq 0. \end{cases} \quad (3.1)$$

Trata-se da mistura de uma distribuição de Bernoulli, com uma função densidade de probabilidade (fdp) ou função de probabilidade (fp) $p(Y_t(s) | \lambda_t(s))$. Temos $1 - \theta_t(s)$ de probabilidade de se obter um valor 0 e $\theta_t(s)$ de se obter um valor de $p(Y_t(s) | \lambda_t(s))$. Note que para se obter a probabilidade total de um valor nulo, deve-se somar a $1 - \theta_t(s)$ a possibilidade de se obter um valor nulo em $p(Y_t(s) | \lambda_t(s))$.

Este tipo de conceituação pode se estender para qualquer outro valor que se deseje modelar em especial. Entretanto, em nosso caso e de maneira geral, o

interesse está no valor zero.

Nosso foco principal nesse estudo está em modelar variáveis não negativas, e portanto, podemos reescrever (3.1) da seguinte maneira:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} (1 - \theta_t(s)) + \theta_t(s)p(Y_t(s) = 0 | \lambda_t(s)) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) > 0. \end{cases} \quad (3.2)$$

No segundo nível de hierarquia do modelo, devemos introduzir covariáveis e aplicar estruturas aos parâmetros $\theta_t(s)$ e $\lambda_t(s)$, a fim de se buscar meios que representem a estrutura espaço-temporal inerente aos dados.

Dentro do contexto dos MLDG's, mencionados no Capítulo anterior, modelaremos o parâmetro da Bernoulli, através de uma regressão logística, tal que:

$$\text{logit}(\theta_t(s)) = \log \frac{\theta_t(s)}{1 - \theta_t(s)} = F'_{1t} \gamma_t + Z_t(s) \quad (3.3)$$

$$\gamma_t = G \gamma_{t-1} + w_{\gamma t}, \quad w_{\gamma t} \sim N(0, W_{\gamma}). \quad (3.4)$$

O termo F_{1t} contém as covariáveis que possam explicar $\theta_t(s)$, sendo γ_t seus respectivos efeitos, desconhecidos. É permitido estruturas variadas para essas covariáveis, tais como: tendência e sazonalidade, que podem ser facilmente incorporadas. Podemos ter inclusive γ variando no tempo, com matriz de evolução G .

A componente $Z_t(s)$, $t = 1, \dots, T$ e $s = 1, \dots, n$, representa o efeito espaço-temporal. Sua finalidade é de tentar incluir uma estrutura, tanto espacial quanto temporal no *logit*. Essa estrutura vai estar diretamente correlacionada com o tipo de referência espacial dos dados em estudo. Se forem dados de área, trabalharemos com uma priori CAR, caso estejamos trabalhando com pontos em uma região, usaremos processos gaussianos.

O parâmetro da distribuição $p(Y_t(s) | \lambda_t(s))$, pode ser modelado por:

$$g(E(Y_t(s) | \lambda_t(s))) = g(\lambda_t(s)) = F_{2t}' \boldsymbol{\alpha}_t + S_t(s) \quad (3.5)$$

$$\boldsymbol{\alpha}_t = H \boldsymbol{\alpha}_{t-1} + w_{\boldsymbol{\alpha}t}, \quad w_{\boldsymbol{\alpha}t} \sim N(0, W_{\boldsymbol{\alpha}}) \quad (3.6)$$

onde $g(\lambda_t(s))$ é a função de ligação da distribuição $p(Y_t(s) | \cdot)$.

Novamente, F_{2t} representa as covariáveis existentes, que podem, ou não, ser as mesmas de F_{1t} . H é a matriz de evolução de $\boldsymbol{\alpha}$, que também pode possuir estruturas diferentes da matriz G . Finalmente $S_t(s)$, $t = 1, \dots, T$ e $s = 1, \dots, n$, descrevem os efeitos espaços-temporais do modelo. Mais uma vez, o tipo de efeito a ser utilizado dependerá do tipo de dado observado.

Assumindo as observações $\mathbf{Y} = (\mathbf{Y}(s_1), \dots, \mathbf{Y}(s_n))'$, onde $\mathbf{Y}(s_i) = (Y_1(s_i), \dots, Y_T(s_i))'$, a verossimilhança será tal que:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}) = \prod_{i=1}^n \prod_{t=1}^T [(1 - \theta_t(s_i)) + \theta_t(s_i) p(Y_t(s_i) = 0 | \lambda_t(s_i))]^{1 - I(Y_t(s_i) > 0)} [\theta_t(s_i) p(Y_t(s_i) | \lambda_t(s_i))]^{I(Y_t(s_i) > 0)}, \quad (3.7)$$

onde $I(Y_t(s_i) > 0)$ é uma variável indicadora de ocorrência de um valor maior que zero no tempo t e localização s_i , $t = 1, \dots, T$ e $i = 1, \dots, n$.

Analisando a expressão (3.7), notamos que não há independência entre os parâmetros $\theta_t(s)$ e $\lambda_t(s)$. Desta forma é necessário a inclusão de uma variável latente para que possamos fatorar a verossimilhança e obtermos assim um facilitador do procedimento de inferência.

Para isso, consideraremos $X_t(s)$ uma variável indicadora do evento $[Y_t(s) = 0]$ ser proveniente da distribuição Bernoulli e não de $p(Y_t(s) | \lambda_t(s))$. Podemos obter uma verossimilhança aumentada, da seguinte forma:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) = \prod_{i=1}^n \prod_{t=1}^T \theta_t(s_i)^{1 - X_t(s_i)} (1 - \theta_t(s_i))^{X_t(s_i)} \prod_{s=1}^n \prod_{t=1}^T p(Y_t(s_i) | \lambda_t(s_i))^{1 - X_t(s_i)}.$$

Observamos que com este artifício conseguimos fatorar a verossimilhança e por consequência obtemos independência entre $\theta_t(s)$ e $\lambda_t(s)$. Desta maneira temos também que, marginalmente, $X_t(s) \sim \text{Bern}(1 - \theta_t(s))$.

3.2 Modelo para Variáveis Mistas

A modelagem de excesso de zeros para variáveis mistas ocorre quando a distribuição $p(Y_t(s) | \lambda_t(s))$ é contínua. Esta modelagem comporta-se de uma maneira especial, pois neste caso teremos, $Pr[Y_t(s) = 0] = 0$. Portanto, as únicas ocorrências do evento $[Y_t(s) = 0]$ estarão ligadas à distribuição de Bernoulli.

Este tipo de abordagem é bem interessante, podendo ser utilizada para dados contínuos provenientes de variáveis não-negativas, em que ocasionalmente são medidos valores nulos. Este é o caso, por exemplo, de estudos de precipitação pluviométrica. Feuerverger (1979) já havia utilizado uma mistura de distribuições, para modelar o comportamento da chuva na Austrália, em uma abordagem clássica.

Vellarde *et al* (2004) também seguiram essa estratégia e fizeram, agora dentro do contexto bayesiano, uma modelagem espaço-temporal, da chuva no estado de Goiás, utilizando também massa de probabilidade no ponto 0.

Farewell (1986) e Meeker (1987), misturaram zeros com distribuições contínuas censuradas para modelar dados de sobrevivência.

Seguindo a conceituação anterior, um modelo completo para variáveis mistas pode ser resumido a seguir:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} 1 - \theta_t(s) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) > 0. \end{cases} \quad (3.8)$$

$$\log \frac{\theta_t(s)}{1 - \theta_t(s)} = F'_{1t} \gamma_t + Z_t(s)$$

$$\gamma_t = G \gamma_{t-1} + w \gamma_t, \quad w \gamma_t \sim N(0, W_\gamma)$$

$$g(\lambda_t(s)) = F'_{2t} \alpha_t + S_t(s)$$

$$\alpha_t = H \alpha_{t-1} + w \alpha_t, \quad w \alpha_t \sim N(0, W_\alpha)$$

E os efeitos espaço-temporais são tais que:

$$Z_t(s), S_t(s) \sim \begin{cases} CAR(\sigma^2) & \text{se } s \text{ é uma subregião em } D \\ GP(0, \sigma^2 \exp(-\phi \|s - s'\|)) & \text{se } s \text{ é pontual.} \end{cases}$$

Notamos que há uma pequena diferença agora, pois a probabilidade do evento $[Y_t(s) = 0]$ é $1 - \theta_t(s)$ e, portanto, o evento complementar, $[Y_t(s) > 0]$, no caso de variáveis não negativas, possui probabilidade $\theta_t(s)$. Por isso Feuerverger (1979) e Vellarde *at al* (2004), conseguiram modelar estações secas e chuvosas em seus respectivos trabalhos sobre estudos pluviométricos.

Diversas distribuições podem ser usadas em $p(Y_t(s) \mid \lambda_t(s))$, entre elas, a Gama, a Exponencial e a Log-Normal.

Assumindo-se que foram observados os valores de $\mathbf{Y} = (\mathbf{Y}(s_1), \dots, \mathbf{Y}(s_n))'$, onde $\mathbf{Y}(s_i) = (Y_1(s_i), \dots, Y_T(s_i))'$, a verossimilhança desse modelo será dada por:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}) = \prod_{i=1}^n \prod_{t=1}^T [1 - \theta_t(s_i)]^{1 - I(Y_t(s_i) > 0)} [\theta_t(s_i) p(Y_t(s_i) \mid \lambda_t(s_i))]^{I(Y_t(s_i) > 0)}, \quad (3.9)$$

Como dito anteriormente os eventos $[Y_t(s) = 0]$ serão provenientes apenas da Bernoulli e neste caso a variável latente $X_t(s)$, definida na Seção anterior, terá uma interpretação bastante específica, que podemos verificar nas distribuições marginais descritas em (3.10).

$$\begin{aligned} Pr(X_t(s) = 0 \mid Y_t(s) = 0, \theta_t(s), \lambda_t(s)) &= 0 \quad \text{e} \\ Pr(X_t(s) = 1 \mid Y_t(s) = 1, \theta_t(s), \lambda_t(s)) &= 1. \end{aligned} \quad (3.10)$$

Os valores de $X_t(s)$ serão conhecidos e podem ser definidos como $X_t(s) = I(Y_t(s) = 0)$. Teremos novamente uma verossimilhança, tal que:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) = \prod_{i=1}^n \prod_{t=1}^T \theta_t(s_i)^{1 - X_t(s_i)} (1 - \theta_t(s_i))^{X_t(s_i)} \prod_{s=1}^n \prod_{t=1}^T p(Y_t(s_i) \mid \lambda_t(s_i))^{1 - X_t(s_i)}.$$

Desta forma obtemos uma verossimilhança aumentada, onde a primeira parcela do produtório acima está associada à regressão logística e a segunda parte, incide sobre a modelagem de valores maiores que zero.

3.3 Modelo para Variáveis Discretas

A modelagem de variáveis discretas inflacionadas de zeros é mais utilizada do que em variáveis mistas, especialmente quando trata-se de dados de contagem. Aqui está inserida uma classe de modelos, muito conhecida na literatura estatística, os chamados *Zero Inflated Poisson (ZIP) models*.

Os *ZIP models*, como são conhecidos, ou modelos inflacionados de zeros, já vêm sendo usados há algum tempo, mas com Heilbron (1989) e principalmente com Lambert (1992) é que ganharam mais destaque.

Estudos em modelos de contagem para dados inflacionados de zeros, começaram a ser feitos por Cohen (1963) e também por Johnson and Kotz (1969). Estes porém, ainda não incluíam covariáveis para os parâmetros do modelo.

Heilbron (1989) propôs um modelo Poisson alterado de zeros (ZAP), em estudos sobre o comportamento humano. Mas foi Lambert (1992), que inicialmente utilizou modelos de regressão com os modelos ZIP, aplicando covariáveis tanto no parâmetro da Bernoulli, quanto no parâmetro da Poisson. Foi utilizada uma aplicação em contagem de defeitos em um processo de manufatura.

Mais recentemente, Rideout, Demetrio, and Hinde (1998) e Rodrigues (2003), também analisaram distribuições com excesso de zeros. Este último com uma abordagem bayesiana.

Ainda dentro do contexto bayesiano, Agarwal, Gelfand, and Citron-Pousty (2002) utilizaram modelos inflacionados de zeros, com efeitos espaciais.

Em nosso trabalho, tal qual o modelo estabelecido na primeira Seção deste Capítulo, teremos a seguinte estrutura:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} (1 - \theta_t(s)) + \theta_t(s)p(Y_t(s) = 0 | \lambda_t(s)) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) > 0. \end{cases} \quad (3.11)$$

As estruturas para $\theta_t(s)$ e $\lambda_t(s)$ também são as mesmas apresentadas em (3.3) e (3.5), respectivamente.

Não teremos mais a interpretação específica para $\theta_t(s)$, como tínhamos no caso das variáveis mistas. Talvez a interpretação mais interessante esteja em $1 - \theta_t(s)$ que representa a chance de um valor medido zero, não ser proveniente da distribuição de probabilidades, $p(Y_t(s) | \lambda_t(s))$, que estivermos utilizando.

Novamente diversas distribuições poderão ser utilizadas para $Y_t(s)$, entre elas a Poisson, a Binomial Negativa (Yau, Wang, & Lee, 2003) e a Poisson Generalizada (J.F. & Biswas, 2003). Como já dissemos, os modelos que utilizam a distribuição Poisson são comumente chamados de *ZIP models*, logo também teremos os equivalentes ZINB e ZIGP, para as outras distribuições.

A verossimilhança para este caso será a mesma descrita em (3.7), assim como a verossimilhança aumentada:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) = \prod_{i=1}^n \prod_{t=1}^T \theta_t(s_i)^{1-X_t(s_i)} (1 - \theta_t(s_i))^{X_t(s_i)} \prod_{s=1}^n \prod_{t=1}^T p(Y_t(s_i) | \lambda_t(s_i))^{1-X_t(s_i)}.$$

Com a introdução desta variável latente, $X_t(s)$, teremos condicionalmente:

$$Pr(Y_t(s) = 0 | X_t(s) = 1, \theta_t(s), \lambda_t(s)) = 1,$$

$$Pr(Y_t(s) = y > 0 | X_t(s) = 1, \theta_t(s), \lambda_t(s)) = 1,$$

$$Pr(Y_t(s) = y | X_t(s) = 0, \theta_t(s), \lambda_t(s)) = p(Y_t(s) | \lambda_t(s)),$$

$$Pr(X_t(s) = 0 | Y_t(s) = y > 0, \theta_t(s), \lambda_t(s)) = 1,$$

$$Pr(X_t(s) = 1 | Y_t(s) = y > 0, \theta_t(s), \lambda_t(s)) = 0,$$

$$Pr(X_t(s) = 0 | Y_t(s) = 0, \theta_t(s), \lambda_t(s)) = \frac{\theta_t(s)p(Y_t(s) = 0 | \lambda_t(s))}{(1 - \theta_t(s)) + \theta_t(s)p(Y_t(s) = 0 | \lambda_t(s))},$$

$$Pr(X_t(s) = 1 | Y_t(s) = 0, \theta_t(s), \lambda_t(s)) = \frac{1 - \theta_t(s)}{(1 - \theta_t(s)) + \theta_t(s)p(Y_t(s) = 0 | \lambda_t(s))}.$$

Mais uma vez teremos, marginalmente, $X_t(s) \sim Bern(1 - \theta_t(s))$. É importante ressaltar, que neste caso, alguns valores de \mathbf{X} , não são observáveis. Mais especificamente, quando ocorrer o evento $[Y_t(s) = 0]$. Isto pode ser visto claramente nas duas

últimas probabilidades condicionais anteriores. Sendo assim, quando estivermos usando o MCMC, estes valores de \mathbf{X} , se tornarão parâmetros do modelo e, portanto, devem ser sorteados.

3.4 Especificação de Prioris

Nesta Seção discutiremos a especificação de prioris para os modelos descritos nas seções anteriores.

Portanto, assumindo o modelo descrito em (3.2) e estruturas para $\theta_t(s)$ e $\lambda_t(s)$, descritas em (3.3) e (3.5), respectivamente, teremos o seguinte vetor paramétrico, para dados de Geoestatística:

$$\Theta_1 = \{\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\gamma}, W_{\boldsymbol{\gamma}}, \boldsymbol{\alpha}, W_{\boldsymbol{\alpha}}, \mathbf{Z}, \sigma_1^2, \phi_1, \mathbf{S}, \sigma_2^2, \phi_2\}.$$

E para dados de área:

$$\Theta_2 = \{\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\gamma}, W_{\boldsymbol{\gamma}}, \boldsymbol{\alpha}, W_{\boldsymbol{\alpha}}, \mathbf{Z}, \sigma_Z^2, \mathbf{S}, \sigma_S^2\}.$$

Considerando como p , o número de possíveis covariáveis para o *logit* e q , o número de covariáveis para a modelagem de valores maiores que zero, os vetores de instante inicial, $\boldsymbol{\gamma}_1$ e $\boldsymbol{\alpha}_1$, assumem prioris normais p -variadas e q -variadas, respectivamente. Logo, $\boldsymbol{\gamma}_1 \sim N_p(\mu_{\boldsymbol{\gamma}}, V_{\boldsymbol{\gamma}})$ e $\boldsymbol{\alpha}_1 \sim N_q(\mu_{\boldsymbol{\alpha}}, V_{\boldsymbol{\alpha}})$. Para as variâncias dos erros de evolução, $W_{\boldsymbol{\gamma}}$ e $W_{\boldsymbol{\alpha}}$, adotamos prioris Gama Invertida, para os elementos da diagonal de $W_{\boldsymbol{\gamma}}$ e $W_{\boldsymbol{\alpha}}$.

Para um caso particular, onde não temos variação temporal nos efeitos das covariáveis, poderemos adotar prioris normais, independentes, para os parâmetros em questão. Portanto, $\gamma_i \sim N(0, \sigma_{\gamma}^2)$ para $i = 1, \dots, p$ e $\alpha_j \sim N(0, \sigma_{\alpha}^2)$ para $j = 1, \dots, q$. Usualmente, valores para σ_{γ}^2 e σ_{α}^2 , são fixados tais que $\sigma_{\gamma}^{-2} \rightarrow 0$ e $\sigma_{\alpha}^{-2} \rightarrow 0$. Em geral estes modelos não são sensíveis à escolha destes valores, excetuando-se os modelos para variáveis discretas com utilização da variável latente

X. Neste caso precisaremos de prioris informativas e, portanto, valores baixos para σ_γ^2 . Isto será discutido com mais detalhe no Capítulo 5, quando estivermos fazendo uma aplicação do modelo.

Para os efeitos espaço-temporais \mathbf{Z} e \mathbf{S} , as prioris dependerão do tipo de dado utilizado e, portanto, do tipo de estrutura espacial utilizada. Como dito anteriormente, podemos ter efeitos Condicionais Auto-Regressivos (CAR) ou Processos Gaussianos.

Tomemos como exemplo, S_{it} , o efeito espaço-temporal na área i , no tempo t , quando estamos trabalhando com dados de área, teremos $S_{it} \sim CAR(\sigma_S^2)$, isto é:

$$(S_{it} \mid S_{jt}, j \neq i) \sim N(\mu_{it}, v_i), \quad (3.12)$$

$$\mu_{it} = \frac{\sum_{j \in \delta_i} S_{jt}}{n_i} \quad \text{e} \quad v_i = \frac{\sigma_S^2}{n_i}. \quad (3.13)$$

onde n_i denota o número e δ_i o conjunto, de vizinhos de i . Lembrando que as expressões acima para μ_{it} e v_i são decorrentes do uso de uma estrutura de vizinhança 0-1. A especificação se completa ao atribuirmos uma distribuição a priori Gama Invertida para o parâmetro σ_S^2 ($\sigma_S^2 \sim GI(a_{\sigma_S^2}, b_{\sigma_S^2})$).

Para os efeitos espaço-temporais, $Z_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$, considerando que estamos trabalhando com dados de Geoestatística, teremos, $Z_t(s) \sim GP(0, \sigma_1^2 \exp(-\phi_1 \|s - s'\|))$, como priori para o tempo t e localização s . Portanto:

$$\mathbf{Z}_t \sim N_n(0, \Sigma) \quad t = 1, \dots, T \quad (3.14)$$

Onde, $\mathbf{Z}_t = (Z_t(s_1), \dots, Z_t(s_n))'$, $t = 1, \dots, T$ e $\Sigma_{ij} = \sigma_1^2 \exp\{-\phi_1 \|s_i - s_j\|\}$, $i, j = 1, \dots, n$.

Para o parâmetro σ_1^2 , aplicaremos uma priori Gama Invertida, tal que $\sigma_1^2 \sim GI(a_{\sigma_1}, b_{\sigma_1})$ e para ϕ_1 , o parâmetro que controla o decaimento da função de correlação espacial, utilizaremos uma distribuição Gama, tal que $\phi_1 \sim Ga(a_{\phi_1}, b_{\phi_1})$, onde os valores de b_{ϕ_1} e a_{ϕ_1} , são calculados de forma que tenhamos:

$$E(\phi_1) = \frac{a_{\phi_1}}{b_{\phi_1}} = \frac{3}{0,5 \text{ } dmax} \quad \text{e} \quad Var(\phi_1) = \frac{a_{\phi_1}}{b_{\phi_1}^2} = 1, \quad (3.15)$$

onde d_{\max} é a distância máxima entre as localizações utilizadas (Banerjee et al., 2004).

Inferência a Posteriori

De posse da verossimilhança e com as prioris especificadas, podemos aplicar o Teorema de Bayes para obtermos a distribuição a posteriori conjunta dos parâmetros. Dessa forma, teremos:

$$p(\Theta | \mathbf{Y}) \propto p(\mathbf{Y} | \Theta)p(\Theta). \quad (3.16)$$

Onde:

(a) Para dados de Geoestatística:

$$p(\Theta_1) = p(\theta | \gamma, W_\gamma, \mathbf{Z})p(\gamma | W_\gamma)p(\lambda | \alpha, W_\alpha, \mathbf{S})p(\alpha | W_\alpha)p(\mathbf{Z} | \sigma_1^2, \phi_1)p(\sigma_1^2)p(\phi_1) \\ p(\mathbf{S} | \sigma_2^2, \phi_2)p(\sigma_2^2)p(\phi_2);$$

(b) Para dados de área:

$$p(\Theta_2) = p(\theta | \gamma, W_\gamma, \mathbf{Z})p(\gamma | W_\gamma)p(\lambda | \alpha, W_\alpha, \mathbf{S})p(\alpha | W_\alpha)p(\mathbf{Z} | \sigma_Z^2)p(\sigma_Z^2)p(\mathbf{S} | \sigma_S^2)p(\sigma_S^2).$$

Quando estivermos trabalhando com a variável latente \mathbf{X} será necessário o cálculo da posteriori aumentada $p(\Theta | \mathbf{Y}, \mathbf{X})$ antes do cálculo da posteriori desejada $p(\Theta | \mathbf{Y})$. Isto pode ser feito de maneira sucessiva, no que se chama de substituições sucessivas, utilizando as distribuições $p(\mathbf{X} | \Theta, \mathbf{Y})$ e $p(\Theta | \mathbf{Y})$. Mais detalhes sobre a convergência desta posteriori podem ser obtidos em Tanner (1996).

3.5 Interpolação Espacial e Previsão Temporal

Quando estamos trabalhando com dados de Geoestatística, o maior interesse encontra-se na previsão do processo em localizações não medidas. Detalhes sobre este tópico, podemos encontrar de forma mais detalhada, por exemplo, em Diggle and Ribeiro Jr (2000).

Basicamente, quando falamos previsão em k pontos não-medidos ou em interpolação espacial para k pontos, estamos nos referindo a $\mathbf{Y}_u = (Y_t(s_j), \dots, Y_t(s_k))'$, $t = 1, \dots, T$, um vetor de localizações não medidas no tempo t . Podemos, então, obter a distribuição preditiva, da seguinte maneira:

$$p(\mathbf{Y}_u | \mathbf{Y}) = \int_{\Theta} p(\mathbf{Y}_u | \mathbf{Y}, \Theta) p(\Theta | \mathbf{Y}) d\Theta, \quad (3.17)$$

onde $p(\Theta | \mathbf{Y})$ é a posteriori dos parâmetros, obtida em (3.16) e $p(\mathbf{Y}_u | \mathbf{Y}, \Theta)$ segue o modelo adotado em (3.2). A integral, descrita em (3.17), é analiticamente intratável, de forma que a saída para este problema, está na utilização de métodos de Monte Carlo (Gamerman, 1997), obtendo-se amostras a posteriori de Θ e aproximando-se esta integral por:

$$p(\mathbf{Y}_u | \mathbf{Y}) \approx \frac{1}{G} \sum_{i=1}^G p(Y_u | \Theta^{(i)}, \mathbf{Y}), \quad (3.18)$$

onde G denota o número total de iterações da amostra obtida de $p(\Theta | \mathbf{Y})$. Para obtermos o valor de $p(Y_u | \Theta^i, \mathbf{Y})$, para a iteração i , é necessário o conhecimento do vetor de amostras a posteriori ($\Theta^{(i)}$). Neste momento deve ser feito o cálculo da posteriori dos efeitos espaço-temporais das localizações não medidas. Considere, portanto, $\mathbf{Z}'_t^{(i)}$, o vetor de efeitos espaço-temporais para as localizações não medidas s_1, \dots, s_k , no instante t e como anteriormente, $\mathbf{Z}_t^{(i)}$ o vetor de efeitos espaço-temporais para as localizações medidas. Novamente, pela teoria da normal multivariada, teremos:

$$\left(\begin{array}{c} \mathbf{Z}'_t^{(i)} \\ \mathbf{Z}_t^{(i)} \end{array} \mid \sigma_1^{2(i)}, \phi_1^{(i)} \right) \sim N_{n+k} \left(\begin{array}{c} \mu' \\ \mu \end{array}, \left(\begin{array}{cc} \Delta^* & \Omega \\ \Omega' & \Delta \end{array} \right) \right), \quad (3.19)$$

onde:

Δ^* é a matriz $k \times k$ de covariância entre os pontos não medidos;

Ω é a matriz $k \times n$ de covariância entre pontos medidos e não medidos;

Δ é matriz $n \times n$ de covariância entre os pontos medidos.

Portanto, obtemos marginalmente:

$$\mathbf{Z}'_t^{(i)} \mid \mathbf{Z}_t^{(i)}, \sigma_1^{2(i)}, \phi_1^{(i)} \sim N_k(\mu' + \Omega' \Delta^{-1}(\mathbf{Z}_t^{(i)} - \mu); \Delta^* - \Omega' \Delta^{-1} \Omega).$$

Pode ser de nosso interesse, também, fazer previsões temporais para k instantes a frente. Novamente, será utilizada a distribuição preditiva, tal que:

$$p(\mathbf{Y}_{T+k} \mid \mathbf{Y}_1, \dots, \mathbf{Y}_T) = \int_{\Theta} p(\mathbf{Y}_{T+k} \mid \mathbf{Y}_1, \dots, \mathbf{Y}_{T+k-1}, \Theta_{T+k}) p(\Theta_{T+k} \mid \mathbf{Y}_1, \dots, \mathbf{Y}_{T+k-1}) d\Theta,$$

que pode ser aproximada por:

$$p(\mathbf{Y}_{T+k} \mid \mathbf{Y}_1, \dots, \mathbf{Y}_T) \approx \frac{1}{G} \sum_{i=1}^G p(Y_{T+k} \mid \Theta_{T+k}^{(i)}, \mathbf{Y}). \quad (3.20)$$

A distribuição $p(\mathbf{Y}_{T+k} \mid \mathbf{Y}_1, \dots, \mathbf{Y}_{T+k-1}, \Theta_{T+k})$ é obtida de forma sequencial, pelo próprio modelo utilizado.

Capítulo 4

Aspectos Computacionais

4.1 Simulação Estocástica

Como visto anteriormente, o processo de inferência bayesiana fundamenta-se numa distribuição, conhecida como distribuição à posteriori. Quando esta distribuição não é conhecida, é necessário utilizar métodos de simulação para obter amostras da distribuição e nesse contexto, faz-se necessário o uso de técnicas de simulação estocástica, tais como, Métodos de Monte Carlo via Cadeias de Markov.

4.2 Amostrador de Gibbs

O amostrador de Gibbs foi originalmente proposto, por Geman and Geman (1984), dentro do contexto de reconstrução de imagens, explorando as condicionais completas através de um algoritmo iterativo que define uma cadeia de Markov. Gelfand and Smith (1990) compararam este amostrador com esquemas de simulação estocástica e uma revolução ocorreu na inferência bayesiana.

O amostrador de Gibbs é essencialmente um esquema iterativo de amostragem de uma cadeia de Markov cujo núcleo de transição é formado pelas distribuições

condicionais completas, que seriam as distribuições marginais condicionais das componentes θ_i a partir da densidade conjunta $\pi(\theta_1, \dots, \theta_p)$. O problema que se coloca é o da geração de uma amostra de π , a partir de $\pi_i(\theta_i) = \pi(\theta_i | \theta_{-i})$, $i = 1, \dots, p$ que é a condicional completa do parâmetro θ_i , e p denota o número de parâmetros. Este método é extremamente útil quando não se consegue amostrar diretamente da densidade conjunta dos parâmetros, e as condicionais completas destes são conhecidas. O algoritmo de Gibbs fornece, então, uma forma alternativa de geração baseada em sucessivas gerações das distribuições condicionais completas. Podemos descrevê-lo da seguinte forma:

(i) inicialize o contador de iterações da cadeia $j = 1$ e arbitre valores iniciais $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_p^{(0)})$.

(ii) obtenha um novo valor $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_p^{(j)})$ a partir de sucessivas gerações de valores

$$\begin{aligned}\theta_1^{(j)} &\sim \pi(\theta_1 | \theta_2^{(j-1)}, \dots, \theta_p^{(j-1)}) \\ \theta_2^{(j)} &\sim \pi(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_p^{(j-1)}) \\ &\vdots \\ \theta_p^{(j)} &\sim \pi(\theta_p | \theta_1^{(j)}, \dots, \theta_{p-1}^{(j)})\end{aligned}$$

(iii) mude o contador j para $j + 1$ e retorne até convergência.

Sob certas condições de regularidade pode-se mostrar que o vetor $(\theta_1^{(j)}, \dots, \theta_p^{(j)})$ converge em distribuição para uma amostra da distribuição a posteriori $\pi(\theta_1, \dots, \theta_p | \mathbf{Y})$, quando j tende a infinito.

Em geral, no procedimento de inferência, para se evitar valores da amostra, antes que a cadeia tenha atingido sua distribuição estacionária, descarta-se as primeiras iterações, fazendo uma espécie de aquecimento (*burn in*). A quantidade de iterações a se descartar é peculiar a cada estudo. Também para se evitar a autocorrelação destas amostras, guarda-se apenas os valores entre um intervalo específico de iterações.

4.3 Algoritmo de Metropolis-Hastings

O Algoritmo de Metropolis-Hastings foi inicialmente proposto por Metropolis, Rosenbulth, Rosenbulth, Teller, and Teller (1953), em um periódico de Físioquímica, e foi posteriormente extendido por Hastings (1970). Tal como o amostrador de Gibbs, ele também tem como finalidade gerar amostras de uma distribuição de probabilidades. Para isso utiliza a idéia de uma distribuição auxiliar, conhecida como densidade de transição. Desta densidade de transição, é gerado um valor proposto para o parâmetro e este é preferido em relação ao valor corrente da cadeia, de acordo com uma determinada probabilidade α .

Considerando $q(\theta, \cdot)$ a densidade proposta, $\pi(\cdot)$ uma distribuição de probabilidades (por exemplo, a condicional completa do parâmetro) e $\theta^{(j-1)}$ o valor corrente da cadeia, podemos resumir o algoritmo de Metropolis-Hastings, a seguir:

- (i) na iteração j , sorteie um valor θ^* para $\theta^{(j)}$, utilizando $q(\theta^{(j-1)}, \cdot)$;
- (ii) aceite mover a cadeia para o valor proposto com probabilidade:

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

e faça $\theta^{(j)} = \theta^*$ ou rejeite mover a cadeia com probabilidade $1 - \alpha(\theta^*, \theta^{(j-1)})$ e, portanto, faça $\theta^{(j)} = \theta^{(j-1)}$.

4.4 Método da Rejeição Adaptativa com passos de Metropolis

O Método da Rejeição Adaptativa com passos de Metropolis (ARMS - Adaptive Rejection Metropolis Sampling) é outro método para se amostrar eficientemente de distribuições de probabilidade complexas. Foi proposto por Gilks, Best, and Tan (1995). É basicamente uma generalização do Método da Rejeição Adap-

tativa (ARS) (Gilks, 1992), que por sua vez, foi desenvolvido do método original proposto por Gilks and Wild (1992).

O método da rejeição adaptativa funciona construindo-se uma função envelope sobre o log da densidade objetivo, aonde é aplicado o passo de rejeição. Entretanto, se um ponto é rejeitado pelo ARS, ele é incluído na construção da função envelope fazendo com que essa se aproxime cada vez mais da verdadeira log-densidade, como podemos observar na Figura 4.1 . Gilks e Wild (1992) propuseram a construção desta função através de tangentes e Gilks (1992) através de secantes. Os dois métodos assumem a log-concavidade da distribuição, o que é uma limitação principalmente quando se pretende amostrar condicionais completas provenientes de modelos que não estão na família exponencial, por exemplo. A generalização do ARS para o ARMS inclui um passo de Metropolis para acomodar os casos onde as distribuições não são log-côncavas.

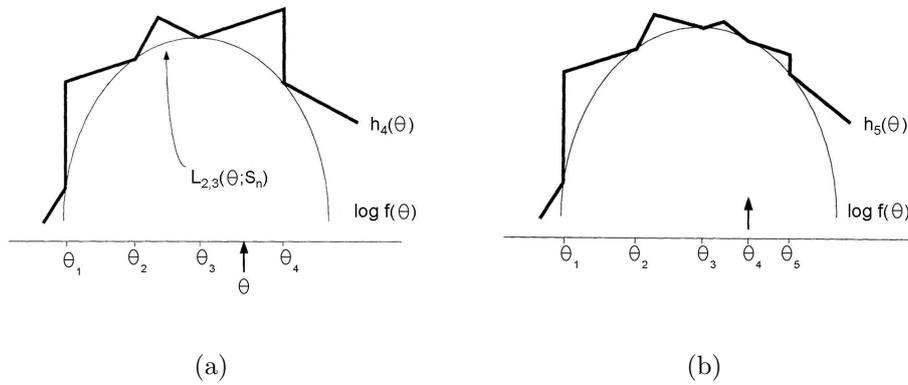


Figura 4.1: (a) Função de rejeição adaptativa $h_4(\theta)$, construída através da equação(4.1), para uma função log-côncava $f(\theta)$: θ é amostrado de $h_4(\theta)$;(b) Atualizando S_5 e a função de rejeição $h_5(\theta)$, após a incorporação de θ .

Assim como no ARS, o ARMS precisa de um conjunto inicial de pontos para a construção da função envelope (no método ARMS esta função pode não ser exatamente um envelope para a log-densidade (Figura 4.2)). Podemos denotar como $S_n = \{\theta_i; i = 0, \dots, n + 1\}$, o conjunto desses valores iniciais (em geral 4 ou 5). Para $1 \leq i \leq n$, seja $L_{ij}(\theta; S_n)$ uma linha reta entre os pontos $[\theta_i, \ln f(\theta_i)]$ e $[\theta_j, \ln f(\theta_j)]$,

onde $\ln f(\cdot)$ é o logarítmo da densidade objetivo. No ARS, definimos uma função linear, em partes, $h_n(\theta)$, como:

$$h_n(\theta) = \min[L_{i-1,i}(\theta; S_n), L_{i+1,i+2}(\theta; S_n)], \quad \theta_i \leq \theta \leq \theta_{i+1}. \quad (4.1)$$

E no ARMS, esta função é definida como:

$$h_n(\theta) = \max[L_{i,i+1}(\theta; S_n), \min\{L_{i-1,i}(\theta; S_n), L_{i+1,i+2}(\theta; S_n)\}], \quad \theta_i \leq \theta \leq \theta_{i+1}. \quad (4.2)$$

A função $h_n(\theta)$ é o que chamamos de função envelope.

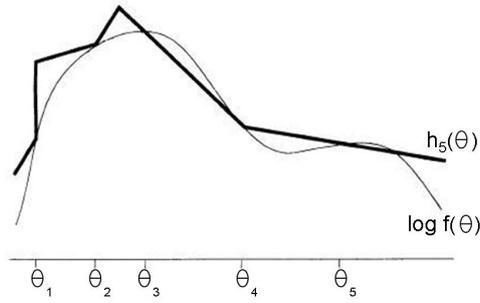


Figura 4.2: Função de rejeição adaptativa $h_5(\theta)$ para uma função não-log-côncava $f(\theta)$, construída a partir de (4.2).

De posse destes elementos podemos montar o algoritmo ARMS, da seguinte forma:

- (i) Inicialize n e S_n independente de $\theta^{(j-1)}$;
- (ii) Amostre θ^* de $g_n(\theta) \propto \exp[h_n(\theta)]$;
- (iii) Amostre U de uma uniforme $(0,1)$;
- (iv) Se

$$U > \frac{f(\theta^*)}{\exp[h_n(\theta^*)]},$$

então:

passo de rejeição ARS:

faça $S_{n+1} = S_n \cup \{\theta^*\}$;

rearrume os pontos em S_{n+1} em ordem ascendente e volte para (i);

caso contrário:

passo de aceitação ARS:

faça $\theta_A = \theta^*$;

(vi) Se

$$U > \min \left[1, \frac{f(\theta_A) \min\{f(\theta^{(j-1)}), \exp[h_n(\theta^{(j-1)})]\}}{f(\theta^{(j-1)}) \min\{f(\theta_A), \exp[h_n(\theta_A)]\}} \right],$$

então:

passo de rejeição do Metropolis-Hastings:

faça $\theta^j = \theta^{(j-1)}$;

caso contrário:

passo de aceitação do Metropolis-Hastings:

faça $\theta^j = \theta_A$;

(vii) Faça $j = j + 1$ e retorne a (i).

4.5 Inferência em Modelos Lineares Dinâmicos Generalizados

Para fazer inferência sobre modelos lineares dinâmicos generalizados (MLDG), Gamerman (1998) desenvolveu uma abordagem usando métodos de Monte Carlo via cadeias de Markov, utilizando uma densidade proposta para o Algoritmo de Metropolis-Hastings, baseada numa versão ajustada dos modelos lineares dinâmicos (MLD). O sorteio é feito através dos erros de evolução, que à priori, não são correlacionados. O conceito, aqui é estendido ao contexto dos modelos espaço-temporais.

Gamerman (1998) propõe utilizar observações ajustadas por mínimos quadrados ponderados, de forma a se obter um modelo linear dinâmico ajustado. Considere

remos, $\tilde{Y}_t(s)$ e $\tilde{V}_t(s)$, como as observações e variâncias ajustadas de um modelo linear dinâmico generalizado, tal como descrito em (2.10) e (2.11). Podemos obtê-las da seguinte maneira:

$$\tilde{Y}_t(s) = g(\mu_t(s)) + (Y_t(s) - \mu_t(s))g'(\mu_t(s)), \quad (4.3)$$

$$\tilde{V}_t(s) = b''(\eta_t(s))g'(\mu_t(s))^2. \quad (4.4)$$

Um modelo linear dinâmico ajustado pode ser criado com a equação das observações a seguir:

$$\tilde{Y}_t(s) = F'_t \theta_t + \tilde{v}_t(s), \quad \tilde{v}_t(s) \sim N(0, \tilde{V}_t(s)). \quad (4.5)$$

Completando-se com a equação de evolução:

$$\theta_t = G\theta_{t-1} + w_t. \quad (4.6)$$

Visando contornar o problema de convergência e autocorrelação existente no sorteio dos parâmetros de evolução, Gamerman (1998) utiliza a relação existente entre tais componentes com os erros de evolução, através da seguinte estrutura:

$$w_t = \theta_t - G\theta_{t-1} \quad (4.7)$$

Fazendo $w_1 = \theta_1$, podemos recompor os parâmetros de evolução, como:

$$\theta_1 = w_1, \quad \theta_t = \sum_{j=1}^t G^{t-j} w_j \quad (t = 2, \dots, n). \quad (4.8)$$

Reescrevendo a equação (4.5) como função de w_t , teremos:

$$\tilde{Y}_t(s) = F'_t \sum_{j=1}^t G^{t-j} w_j + \tilde{v}_t(s), \quad (4.9)$$

onde:

$$\tilde{v}_t(s) \sim N(0, \tilde{V}_t(s)), \quad w_t \sim N(0, W) \quad (t = 2, \dots, n), \quad w_1 \sim N(a, R), \quad (4.10)$$

onde a e R são os parâmetros da priori para w_1 .

Podemos obter a distribuição conjunta como:

$$p(\mathbf{Y}, \mathbf{w}, \mathbf{V}, W) = \left[\prod_{t=1}^T \prod_{i=1}^n p(Y_t(s_i) | w_t, \tilde{V}_t(s_i)) \right] \left[\prod_{t=2}^T p(w_t | W) \right] p(w_1) p(W). \quad (4.11)$$

Através de (4.9), notamos que dependem de w_t , os parâmetros de evolução do tempo t até o tempo T . Portanto pode-se obter a condicional completa para w_t da seguinte forma:

$$\pi(w_t) \propto \prod_{j=t}^T \prod_{i=1}^n \exp \left\{ -\frac{1}{2\tilde{V}_j(s_i)} (Y_j(s_i) - k_{tj} - H'_{tj} w_t)^2 - \frac{1}{2} w'_t W^{-1} w_t \right\}, \quad (4.12)$$

onde:

$$H'_{tj} = F'_t G^{j-t}, \quad k_{tj} = F'_t (G^{j-t} w_1 + \dots + G^{j-t+1} w_{t-1} + G^{j-t-1} w_{t+1} + \dots + w_t). \quad (4.13)$$

Após alguns cálculos, podemos mostrar que a condicional completa (ajustada) para w_t , com as observações ajustadas, é tal que:

$$w_t \sim N(\mu_w, \Sigma_w), \quad (4.14)$$

onde para $t = 2, \dots, T$:

$$\begin{aligned} \mu_w &= \Sigma_w \sum_{j=t}^T \sum_{i=1}^n \tilde{V}_j(s_i)^{-1} H_{tj} (Y_t(s_i) - k_{tj}), \\ \Sigma_w &= \left(W^{-1} + \sum_{j=t}^T \sum_{i=1}^n \tilde{V}_j(s_i)^{-1} H_{tj} H'_{tj} \right)^{-1}. \end{aligned}$$

Se $t = 1$, temos:

$$\begin{aligned} \mu_w &= \Sigma_w \left\{ R^{-1} a + \sum_{j=1}^T \sum_{i=1}^n \tilde{V}_j(s_i)^{-1} H_{1j} (Y_1(s_i) - k_{1j}) \right\}, \\ \Sigma_w &= \left(R^{-1} + \sum_{j=1}^T \sum_{i=1}^n \tilde{V}_j(s_i)^{-1} H_{1j} H'_{1j} \right)^{-1}. \end{aligned}$$

Estas condicionais completas servirão como propostas para o sorteio de w_t . A condicional completa para W , possui forma fechada, Gama Invertida, tal que:

$$\pi(W) \sim IG \left(a_w + \frac{(T-1)}{2}, \sum_{t=2}^T \frac{w_t^2}{2} + b_w \right), \quad (4.15)$$

onde a_w e b_w são os parâmetros da priori Gama Invertida para W .

4.5.1 Exemplos

Podemos calcular as expressões para as observações e variâncias ajustadas, para algumas distribuições, utilizando (4.3) e (4.4). Então, considerando $Y_t(s)$ com distribuição Bernoulli de parâmetro $\eta_t(s)$, teremos:

$$\begin{aligned}\tilde{Y}_t(s) &= \log\left(\frac{\eta_t(s)}{1-\eta_t(s)}\right) + \frac{(Y_t(s) - \eta_t(s))}{\eta_t(s)(1-\eta_t(s))}, \\ \tilde{V}_t(s) &= \frac{1}{\eta_t(s)(1-\eta_t(s))}.\end{aligned}$$

Para a distribuição Gama, isto é, $Y_t(s) \sim Ga(\delta_t(s), \eta_t(s))$, teremos :

$$\begin{aligned}\tilde{Y}_t(s) &= \log\left(\frac{\delta_t(s)}{\eta_t(s)}\right) + \left(Y_t(s) - \frac{\delta_t(s)}{\eta_t(s)}\right) \frac{\eta_t(s)}{\delta_t(s)}, \\ \tilde{V}_t(s) &= \frac{\eta_t(s)}{\delta_t(s)}.\end{aligned}$$

E para a distribuição Poisson, isto é, $Y_t(s) \sim Po(\eta_t(s))$, teremos:

$$\begin{aligned}\tilde{Y}_t(s) &= \log(\eta_t(s)) + \eta_t(s)^{-1}(Y_t(s) - \eta_t(s)), \\ \tilde{V}_t(s) &= \eta_t(s)^{-1}.\end{aligned}$$

Para todos estes casos, assume-se que o parâmetro $\eta_t(s)$ possua alguma estrutura, por exemplo:

$$\begin{aligned}g(\eta_t(s)) &= \theta_t + \gamma F'_t(s) + Z_t(s) \\ \theta_t &= \theta_{t-1} + w_t,\end{aligned}$$

onde θ_t é um nível que varia no tempo, $F_t(s)$ são covariáveis e $Z_t(s)$ são os efeitos espaço-temporais.

O sorteio de θ_t , $t = 1, \dots, T$ poderá ser feito como em Gamerman (1998). Para isso, obtemos o seguinte modelo linear ajustado:

$$\tilde{Y}_t(s) \sim N(\theta_t + \gamma F'_t(s) + Z_t(s), \tilde{V}_t(s)).$$

Podemos reconstruir o modelo da seguinte forma:

$$Y_t^*(s) = \tilde{Y}_t(s) - \gamma F_t'(s) - Z_t(s).$$

Portanto:

$$\begin{aligned} Y_t^*(s) &= \theta_t + v_t(s) & v_t &\sim N(0, \tilde{V}_t(s)), \\ \theta_t &= \theta_{t-1} + w_t & w_t &\sim N(0, W). \end{aligned}$$

A condicional completa para w_t , nas observações ajustadas seguindo (4.12) será:

$$\pi(w_t) \propto \prod_{j=t}^T \prod_{i=1}^n \exp \left\{ -\frac{(Y_j^*(s_i) - \theta_j)^2}{2\tilde{V}_j(s_i)} - \frac{w_t^2}{2} \right\}. \quad (4.16)$$

Se fizermos, $\hat{Y}_j(s) = Y_j^*(s) - \epsilon_j$, onde $\epsilon_j = \theta_j - w_t$, podemos reescrever (4.16), como:

$$\pi(w_t) \propto \exp \left\{ -\frac{w_t^2}{2} \left(\sum_{j=t}^T \sum_{i=1}^n \frac{1}{\tilde{V}_j(s_i)} + W^{-1} \right) - w_t \sum_{j=t}^T \sum_{i=1}^n \frac{\hat{Y}_j(s_i)}{\tilde{V}_j(s_i)} \right\}. \quad (4.17)$$

Logo a proposta $q(w_t^c, w_t^p)$ será normal com média μ_w e variância Σ_w , onde:

$$\Sigma_w = \left(\sum_{j=t}^T \sum_{i=1}^n \frac{1}{\tilde{V}_j(s_i)} + W^{-1} \right)^{-1} \quad \text{e} \quad \mu_w = \Sigma_w \sum_{j=t}^T \sum_{i=1}^n \frac{\hat{Y}_j(s_i)}{\tilde{V}_j(s_i)}. \quad (4.18)$$

Para $t=1$, os parâmetros ficarão:

$$\Sigma_w = \left(\sum_{j=1}^T \sum_{i=1}^n \frac{1}{\tilde{V}_j(s_i)} + R^{-1} \right)^{-1} \quad \text{e} \quad \mu_w = \Sigma_w \sum_{j=1}^T \sum_{i=1}^n \left(\frac{\hat{Y}_j(s_i)}{\tilde{V}_j(s_i)} + R^{-1}a \right), \quad (4.19)$$

onde a probabilidade α de aceitação no passo de Metropolis-Hastings, ficará:

$$\alpha = \min \left\{ 1, \frac{\pi(w_t^p)q(w_t^p, w_t^c)}{\pi(w_t^c)q(w_t^c, w_t^p)} \right\}.$$

A condicional completa para w_t , no caso do modelo de Bernoulli, será tal que;

para $t > 1$:

$$\pi(w_t) \propto \prod_{j=t}^T \prod_{i=1}^n \eta_j(s_i)^{1-Y_j(s_i)} [1 - \eta_j(s_i)]^{Y_j(s_i)} \exp \left\{ -\frac{w_t^2}{2W} \right\}, \quad (4.20)$$

Para $t = 1$:

$$\pi(w_1) \propto \prod_{j=1}^T \prod_{i=1}^n \eta_j(s_i)^{1-Y_j(s_i)} [1 - \eta_j(s_i)]^{Y_j(s_i)} \exp \left\{ -\frac{(w_1 - a)^2}{2R} \right\}. \quad (4.21)$$

Podemos reconstruir os valores de θ_t , através de $\theta_t = \sum_{j=1}^t w_j$.

Para o modelo Gama, teremos:

Para $t > 1$:

$$\pi(w_t) \propto \prod_{j=t}^T \prod_{i=1}^n \left(\frac{\eta_t(s_i)^{\delta_t(s_i)}}{\Gamma(\delta_t(s_i))} \right) \exp\{-\eta_j(s_i)Y_j(s_i)\} \exp \left\{ -\frac{w_t^2}{2W} \right\}. \quad (4.22)$$

Para $t = 1$:

$$\pi(w_1) \propto \prod_{j=1}^T \prod_{i=1}^n \left(\frac{\eta_t(s_i)^{\delta_t(s_i)}}{\Gamma(\delta_t(s_i))} \right) \exp\{-\eta_j(s_i)Y_j(s_i)\} \exp \left\{ -\frac{(w_1 - a)^2}{2R} \right\}. \quad (4.23)$$

Para o modelo Poisson, teremos:

Para $t > 1$:

$$\pi(w_t) \propto \prod_{j=t}^T \prod_{i=1}^n \frac{\exp\{-\eta_j(s_i)\} \eta_j(s_i)^{Y_j(s_i)}}{Y_j(s_i)!} \exp \left\{ -\frac{w_t^2}{2W} \right\}. \quad (4.24)$$

Para $t = 1$:

$$\pi(w_1) \propto \prod_{j=1}^T \prod_{i=1}^n \frac{\exp\{-\eta_j(s_i)\} \eta_j(s_i)^{Y_j(s_i)}}{Y_j(s_i)!} \exp \left\{ -\frac{(w_1 - a)^2}{2R} \right\}. \quad (4.25)$$

Capítulo 5

Aplicações

Neste Capítulo faremos algumas aplicações utilizando toda a teoria abordada nos Capítulos anteriores. Serão três aplicações ao todo: uma para dados contínuos e duas para dados discretos.

Todas as rotinas computacionais foram executadas em Ox versão 3.2 (Doornick, 2002). Para a confecção de gráficos e das superfícies utilizadas, foi usado o software R versão 2.2.0 (R Development Core Team, 2005).

Para lidar com as referências espaciais e execução de mapas e superfícies foi utilizado o ArcView versão 9.0.

5.1 Modelando a chuva na cidade do Rio de Janeiro

A cidade do Rio de Janeiro é a capital do Estado do Rio de Janeiro, que situa-se na Região Sudeste do Brasil. Esta região caracteriza-se pela presença de clima tropical.

É um clima quente e úmido, e pode ser dividido basicamente em duas principais estações: verão (chuvoso), com médias térmicas e índices pluviométricos elevados, e inverno (menos chuvoso).

De frente a este cenário, podemos querer modelar o fenômeno pluviométrico da cidade. É comum encontrarmos na literatura, este tipo de modelagem sendo feita com distribuições estritamente positivas, principalmente a Exponencial (Velarde et al., 2004) e a Gama (Feuerverger, 1979). A questão que fica é, aonde incluiríamos os períodos secos, ou seja, os períodos onde não há precipitação e, portanto, temos valor medido igual a zero?

Alternativas como, transformações nos dados e distribuições truncadas, podem ser usadas. Sansó and Guenni (1999) utilizaram uma normal multivariada, truncada e transformada, para modelar a chuva na Venezuela.

Mas como dito anteriormente, Feuerverger (1979) e Velarde et al. (2004) utilizaram massa de probabilidade no ponto zero, para casos semelhantes.

Nossa aplicação está baseada na teoria desenvolvida na segunda Seção do Capítulo 3. Utilizaremos a precipitação na cidade do Rio de Janeiro, como fonte de dados.

A abordagem será semelhante à utilizada por Velarde et al. (2004), entretanto, utilizaremos dados de Geoestatística, considerando a chuva distribuída continuamente sobre a superfície da cidade. Algumas distribuições serão testadas e previsões para quaisquer pontos da superfície poderão ser calculadas. Velarde et al. (2004) utilizam a distribuição Exponencial e modelam os efeitos espaço-temporais seguindo distribuições CAR, não podendo, portanto, obter previsões para os pontos não medidos.

5.1.1 Conjunto de dados

Os dados para esta aplicação, foram obtidos através do site da **Geo-Rio**¹. Correspondem a dados das 32 estações pluviométricas que monitoram os índices para a cidade do Rio de Janeiro. Estes são divulgados em espaços de 30 em 30

¹<http://www.rio.rj.gov.br/georio>

minutos, mas para esta aplicação foram agregados em semanas.

Foram coletados índices de 75 semanas entre os anos de 2001 e 2002. A série começa em 14/01/01 e termina em 20/06/2002. A Figura 5.1 ilustra a série de duas estações distintas.

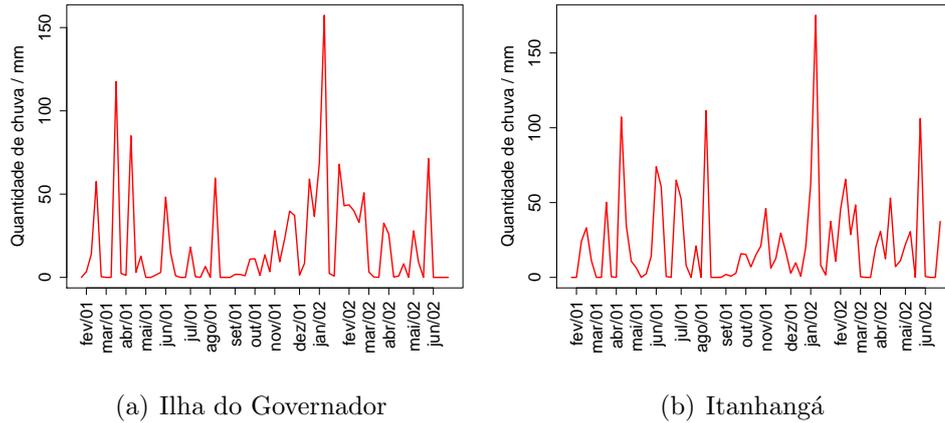


Figura 5.1: Quantidade de chuva medida em duas estações durante as 75 semanas consideradas.

A Tabela A.1 apresenta o nome das estações, com suas referências espaciais (latitude, longitude, UTM(Norte), UTM(Leste)) e a Figura 5.2 mostra o mapa da cidade onde encontram-se destacadas as localizações das 32 estações monitoradoras. Em verde, estão as 5 (cinco) estações que foram deixadas de fora da análise para validação dos resultados obtidos.

5.1.2 Modelos Utilizados

Seguindo toda a abordagem introduzida na Seção 3.2, um modelo para variável chuva, pode ser obtido como o apresentado em (3.8). Portanto:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} 1 - \theta_t(s) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) > 0. \end{cases}$$

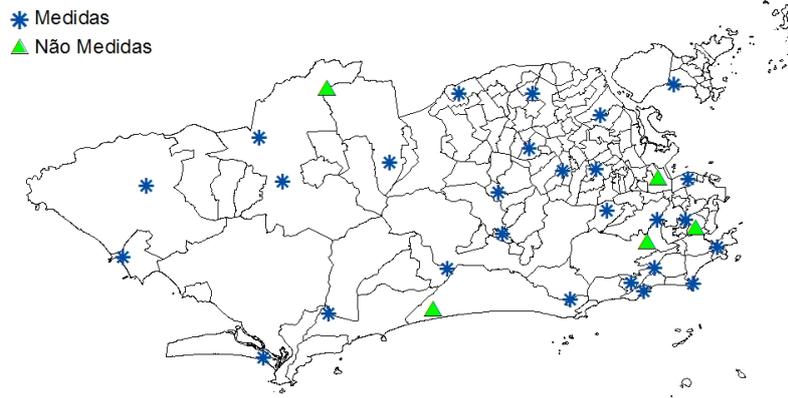


Figura 5.2: Localização das 32 estações monitoradoras da Geo-Rio.

A variável $Y_t(s)$ representa a quantidade de chuva medida pela estação s , na semana t . Aqui $\theta_t(s)$ possui uma interpretação bastante específica, pois como já visto anteriormente, $\theta_t(s)$ nos modelos para variáveis mistas, representa a chance de obtermos um valor não nulo. Logo, $\theta_t(s)$ é a probabilidade de ocorrência de chuva.

A densidade $p(Y_t(s) | \lambda_t(s))$ modelará os níveis de chuva, maiores que zero. Distribuições, como a Gama, Exponencial (caso particular da Gama) e a Lognormal, são prováveis candidatas, neste caso.

Seguindo a hierarquia do modelo, aplicaremos estruturas aos parâmetros. Como em (3.3), utilizaremos uma regressão logística para modelar a probabilidade de níveis não nulos de chuva, $\theta_t(s)$, ou seja:

$$\text{logit}(\theta_t(s)) = \log \frac{\theta_t(s)}{1 - \theta_t(s)} = \gamma_0 + \sum_{j=1}^p \gamma_j I(Y_{t-j}(s) > 0) + Z_t(s), \quad (5.1)$$

onde γ_0 representa uma média comum para todos os pontos, e γ_j $j = 1, \dots, p$ são fatores que visam incorporar a influência da existência de chuva, em p semanas anteriores, sobre a probabilidade de chuva na semana corrente. E, finalmente, $Z_t(s)$ representa o efeito espaço-temporal, da semana t , na localização s . Como estamos trabalhando com pontos no espaço (dados de Geostatística), este efeito terá como priori um processo Gaussiano, de modo que:

$$Z_t(\mathbf{s}) \sim GP(0, \sigma_1^2 \exp(-\phi_1 \|s - s'\|)),$$

onde σ_1^2 representa a variância do processo, ϕ_1 é o parâmetro que controla o quão rapidamente a correlação do processo decai para zero e $\| \mathbf{s} - \mathbf{s}' \|$ denota a distância euclidiana entre dois pontos medidos.

Para $\lambda_t(s)$ aplicaremos, inicialmente, a seguinte estrutura:

$$g(\lambda_t(s)) = \alpha_0 + S_t(s). \quad (5.2)$$

Aqui, $g(\lambda_t(s))$ é a função de ligação que dependerá da distribuição, $p(Y_t(s) | \lambda_t(s))$, adotada. Novamente α_0 representa um nível fixo no tempo e $S_t(s)$ o efeito espaço-temporal, que vai ter como priori, um processo Gaussiano, tal que:

$$S_t(\mathbf{s}) \sim GP(0, \sigma_2^2 \exp(-\phi_2 \| \mathbf{s} - \mathbf{s}' \|)).$$

Podemos notar que estamos assumindo que tanto, a probabilidade de chuva, quanto a quantidade de chuva, possuem variação espacial e temporal.

5.1.3 Procedimento de Inferência

Para $Y_t(s)$ testaremos três possibilidades, já citadas. Teremos portanto, (i) $Y_t(s) \sim Ga(\delta, \delta\lambda_t(s))$, (ii) $Y_t(s) \sim \text{Exp}(\lambda_t(s))$ e (iii) $Y_t(s) \sim \text{LN}(\lambda_t(s), \sigma^2)$, com suas respectivas densidades:

$$(i) f_Y(Y_t(s)) = \frac{(\delta\lambda_t(s))^\delta}{\Gamma(\delta)} Y_t(s)^{\delta-1} \exp\{-\delta\lambda_t(s)Y_t(s)\},$$

$$(ii) f_Y(Y_t(s)) = \lambda_t(s) \exp\{-\lambda_t(s)Y_t(s)\},$$

$$(iii) f_Y(Y_t(s)) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(\log Y_t(s) - \lambda_t(s))^2\right\} \frac{1}{Y_t(s)}.$$

Vale lembrar que a distribuição Exponencial é um caso particular da distribuição Gama. Em nossa parametrização, fazendo $\delta = 1$, na distribuição Gama, obteremos uma distribuição Exponencial.

Podemos obter a verossimilhança para o modelo de chuva, da seguinte maneira:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}, \delta; \mathbf{Y}) = \prod_{i=1}^{27} \prod_{t=1}^{75} [1 - \theta_t(s_i)]^{X_t(s_i)} [\theta_t(s_i) p(Y_t(s_i) | \lambda_t(s_i), \delta)]^{1-X_t(s_i)},$$

onde $X_t(s_i) = I(Y_t(s_i) = 0)$, a indicadora de chuva, na semana t e localização s_i , $i = 1, \dots, 27$. Por exemplo, para o modelo Gama, teríamos a seguinte log-verossimilhança:

$$\begin{aligned} l(\boldsymbol{\theta}, \boldsymbol{\lambda}, \delta; \mathbf{Y}) &= \sum_{i=1}^{27} \sum_{t=1}^{75} [X_t(s_i) \log(1 - \theta_t(s_i)) + X_t(s_i) \log \theta_t(s_i)] \\ &+ (1 - X_t(s_i)) \{ \delta \log(\delta \lambda_t(s_i)) - \log \Gamma(\delta) + (\delta - 1) \log Y_t(s_i) - \delta \lambda_t(s_i) Y_t(s_i) \}. \end{aligned}$$

Considerando $g(\lambda_t(s)) = \log\left(\frac{1}{\lambda_t(s)}\right)$, como a função de ligação do modelo Gama e utilizando (5.1) e (5.2), teremos:

$$\begin{aligned} l(\cdot; \mathbf{Y}) &= \sum_{i=1}^{27} \sum_{t=1}^{75} [X_t(s_i) (-\log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) + Z_t(s_i)\})) \\ &+ (1 - X_t(s_i)) (\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) + Z_t(s_i)) \\ &+ (1 - X_t(s_i)) (-\log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) + Z_t(s_i)\})) \\ &+ (1 - X_t(s_i)) [\delta \log(\delta) - \log(\Gamma(\delta)) - \delta(\alpha_0 + S_t(s_i)) + (\delta - 1) \log Y_t(s_i)] \\ &- (\delta Y_t(s_i) \exp\{-(\alpha_0 + S_t(s_i))\})]. \end{aligned}$$

De posse da verossimilhança, precisamos associar prioris aos parâmetros do modelo, afim de obtermos a distribuição posteriori.

A especificação das prioris seguirão a conceituação abordada no Capítulo 3. Portanto, consideraremos, como Θ , nosso vetor com todas as quantidades desconhecidas (parâmetros) do modelo. Logo,

$$\Theta = \{\boldsymbol{\theta}, \boldsymbol{\lambda}, \gamma_0, \gamma_1, \alpha_0, \mathbf{Z}, \sigma_1^2, \phi_1, \mathbf{S}, \sigma_2^2, \phi_2, \delta\}.$$

A priori para o parâmetro δ será uma Gama, tal que $\delta \sim G(0.1, 0.1)$. Para $\sigma_{\gamma_0}^2$, $\sigma_{\gamma_1}^2$ e $\sigma_{\alpha_0}^2$, que são as variâncias das prioris para γ_0 , γ_1 e α_0 , respectivamente,

assumiremos o valor 100. Como já especificado anteriormente essas prioris terão médias nulas. Para σ_1^2 e σ_2^2 , teremos $a_{\sigma_1} = 2, b_{\sigma_1} = 2, a_{\sigma_2} = 2$ e $b_{\sigma_2} = 2$. E para ϕ_1 e ϕ_2 , teremos $a_{\phi_1} = a_{\phi_2} = 0, 1178$ e $b_{\phi_1} = b_{\phi_2} = 0, 1085$, pois a distância euclidiana máxima entre as estações monitoradoras é de 55,28 (vide 3.15). Para os efeitos espaço-temporais \mathbf{Z} e \mathbf{S} , as distribuições a priori são normais.

Portanto, via teorema de Bayes, podemos calcular a posteriori para todos os parâmetros de interesse, sendo $\Theta_{-\theta}$ o vetor paramétrico Θ exceto o parâmetro θ :

$$p(\Theta | \mathbf{Y}) \propto p(\mathbf{Y} | \boldsymbol{\theta}, \boldsymbol{\lambda}, \delta) p(\boldsymbol{\theta} | \gamma_0, \gamma_1, \mathbf{Z}) p(\boldsymbol{\lambda} | \alpha_0, \mathbf{S}) p(\mathbf{Z} | \phi_1, \sigma_1^2) \\ p(\mathbf{S} | \phi_2, \sigma_2^2) p(\gamma_0) p(\gamma_1) p(\phi_1) p(\sigma_1^2) p(\alpha_0) p(\phi_2) p(\sigma_2^2) p(\delta).$$

A seguir estão apresentadas as condicionais completas para cada parâmetro separadamente:

(i) Condicional completa para γ_0 :

$$p(\gamma_0 | \mathbf{Y}, \mathbf{X}, \Theta_{-\gamma_0}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_0}^2} (\gamma_0 - \mu_{\gamma_0})^2 \right\} \exp \left\{ \gamma_0 \sum_{i=1}^{27} \sum_{t=1}^{75} (1 - X_t(s_i)) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) + Z_t(s_i)\}) \right\};$$

(ii) Condicional completa para γ_1 :

$$p(\gamma_1 | \mathbf{Y}, \mathbf{X}, \Theta_{-\gamma_1}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_1}^2} (\gamma_1 - \mu_{\gamma_1})^2 \right\} \exp \left\{ \gamma_1 \sum_{i=1}^{27} \sum_{t=1}^{75} (1 - X_t(s_i)) I(Y_{t-1}(s_i) > 0) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) + Z_t(s_i)\}) \right\};$$

(iii) Condicional completa para α_0 :

$$p(\alpha_0 | \mathbf{Y}, \mathbf{X}, \Theta_{-\alpha_0}) \propto \exp \left\{ -\frac{1}{2\sigma_{\alpha_0}^2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} \exp \left\{ -\delta \alpha_0 \sum_{i=1}^{27} \sum_{t=1}^{75} (1 - X_t(s_i)) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} (\delta Y_t(s_i) \exp\{-\alpha_0 + S_t(s_i)\}) \right\};$$

(iv) Condicional completa para $Z_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$:

$$p(Z_t(s_i) | \mathbf{Y}, \mathbf{X}, \Theta_{-Z_t(s_i)}) \propto \exp \left\{ -\frac{1}{2VZ} (Z_t(s_i) - mZ)^2 \right\} \exp \{ Z_t(s_i) (1 - X_t(s_i)) \} \\ \exp \{ -\log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) + Z_t(s_i)\}) \};$$

(v) Condicional completa para $S_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$:

$$p(S_t(s_i) | \mathbf{Y}, \mathbf{X}, \Theta_{-S_t(s_i)}) \propto \exp \left\{ -\frac{1}{2VS} (S_t(s_i) - mS)^2 \right\} \exp \{ -\delta S_t(s_i)(1 - X_t(s_i)) \} \\ \exp \{ (-\delta Y_t(s_i) \exp \{ -(\alpha_0 + S_t(s_i)) \}) \};$$

(vi) Condicional completa para δ :

$$p(\delta | \mathbf{Y}, \mathbf{X}, \Theta_{-\delta}) \propto \delta^{(a_\delta-1)} \exp \{ -b_\delta \delta \} \exp \left\{ \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) [\delta \log(\delta) - \log(\Gamma(\delta)) \right. \\ \left. - \delta(\alpha_0 + S_t(s_i)) + (\delta - 1) \log Y_t(s_i)] \right. \\ \left. - \sum_{s=1}^{27} \sum_{t=1}^{75} (\delta Y_t(s_i) \exp \{ -(\alpha_0 + S_t(s_i)) \}) \right\};$$

(vii) Condicional completa para σ_1^2 :

$$p(\sigma_1^2 | \mathbf{Y}, \mathbf{X}, \Theta_{-\sigma_1^2}) \propto \sigma_1^{2-(a_1+1)} \exp \left\{ -\frac{b_1}{\sigma_1^2} \right\} \sigma_1^{2(-\frac{n}{2})} \left\{ -\frac{1}{2\sigma_1^2} \sum_{t=1}^T \mathbf{Z}_t' R^{-1} \mathbf{Z}_t \right\},$$

onde:

$$R_{ij} = \exp \{ -\phi_1 || s_i - s_j || \}; i, j = 1, \dots, n;$$

(viii) Condicional completa para σ_2^2 :

$$p(\sigma_2^2 | \mathbf{Y}, \mathbf{X}, \Theta_{-\sigma_2^2}) \propto \sigma_2^{2-(a_2+1)} \exp \left\{ -\frac{b_2}{\sigma_2^2} \right\} \sigma_2^{2(-\frac{n}{2})} \exp \left\{ -\frac{1}{2\sigma_2^2} \sum_{t=1}^T \mathbf{S}_t' W^{-1} \mathbf{S}_t \right\},$$

onde

$$W_{ij} = \exp \{ -\phi_2 || s_i - s_j || \}; i, j = 1, \dots, n;$$

(ix) Condicional completa para ϕ_1 :

$$p(\phi_1 | \mathbf{Y}, \mathbf{X}, \Theta_{-\phi_1}) \propto \phi_1^{(\alpha_1-1)} \exp \{ -\beta_1 \phi_1 \} | \Sigma |^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \mathbf{Z}_t' \Sigma^{-1} \mathbf{Z}_t \right\},$$

onde

$$\Sigma_{ij} = \sigma_1^2 \exp \{ -\phi_1 || s_i - s_j || \}; i, j = 1, \dots, n;$$

(x) Condicional completa para ϕ_2 :

$$p(\phi_2 | \mathbf{Y}, \mathbf{X}, \Theta_{-\phi_2}) \propto \phi_2^{(\alpha_2-1)} \exp\{-\beta_2\phi_2\} | \Psi |^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^T \mathbf{S}_t' \Psi^{-1} \mathbf{S}_t\right\},$$

onde

$$\Psi_{ij} = \sigma_2^2 \exp\{-\phi_2 || s_i - s_j ||\}; i, j = 1, \dots, n.$$

Podemos notar que apenas as condicionais completas para σ_1^2 e σ_2^2 possuem forma conhecida. Desta maneira, temos que, $\sigma_1^2 \sim IG\left(a_1 + \frac{n}{2}, b_1 + \sum_{t=1}^T \mathbf{Z}_t' R^{-1} \mathbf{Z}_t\right)$ e $\sigma_2^2 \sim IG\left(a_2 + \frac{n}{2}, b_2 + \sum_{t=1}^T \mathbf{S}_t' W^{-1} \mathbf{S}_t\right)$. Para os demais parâmetros, precisaremos de métodos de simulação estocástica, descritos no Capítulo 4.

Como mencionado no início desta Seção, utilizou-se três distribuições diferentes para $Y_t(s)$, a Gamma, a Exponencial e a Log-normal. Também foram experimentadas três possibilidades na modelagem da probabilidade de chuva ($\theta_t(s)$). Em (5.1) fizemos, $p = 1, 2$ ou 3 , testando a significância da existência de chuva em uma, duas ou três semanas anteriores, respectivamente. Para a modelagem de níveis de chuva maiores que zero, utilizou-se a mesma estrutura. Totaliza-se, portanto, 9 modelos, que podem ser visualizados na Tabela 5.1.

Tabela 5.1: Modelos utilizados

| Modelo | $Y_t(s)$ | <i>logit</i> |
|--------|-----------------------------------|------------------------------------------------------------------|
| 1 | | $\gamma_0 + \gamma_1 I_1 + Z_t(s)$ |
| 2 | $\exp(\lambda_t(s))$ | $\gamma_0 + \gamma_1 I_1 + \gamma_2 I_2 + Z_t(s)$ |
| 3 | | $\gamma_0 + \gamma_1 I_1 + \gamma_2 I_2 + \gamma_3 I_3 + Z_t(s)$ |
| 4 | | $\gamma_0 + \gamma_1 I_1 + Z_t(s)$ |
| 5 | $Ga(\delta, \delta \lambda_t(s))$ | $\gamma_0 + \gamma_1 I_1 + \gamma_2 I_2 + Z_t(s)$ |
| 6 | | $\gamma_0 + \gamma_1 I_1 + \gamma_2 I_2 + \gamma_3 I_3 + Z_t(s)$ |
| 7 | | $\gamma_0 + \gamma_1 I_1 + Z_t(s)$ |
| 8 | $LN(\lambda_t(s), \sigma^2)$ | $\gamma_0 + \gamma_1 I_1 + \gamma_2 I_2 + Z_t(s)$ |
| 9 | | $\gamma_0 + \gamma_1 I_1 + \gamma_2 I_2 + \gamma_3 I_3 + Z_t(s)$ |

Onde $I_j = I(Y_{t-j}(s) > 0)$ para $j = 1, 2, 3$.

As condicionais completas para os modelos 2 e 9 encontram-se detalhadamente descritas no apêndice B.

A definição do sistema de amostragem empregado para cada uma destas condicionais completas, foi personalizado, de acordo com a característica específica de cada uma.

Portanto para γ_i , $i = 1, \dots, p$ e α_0 , que são parâmetros que possuem condicionais completas log-côncavas, utilizaremos o algoritmo ARMS. Para ϕ_1 , ϕ_2 e δ , utilizaremos o algoritmo de Metropolis-Hastings com proposta log-normal, sendo sua média o valor corrente da cadeia. Para os efeitos espaço-temporais $Z_t(s_i)$ e $S_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$, também empregaremos o algoritmo de Metropolis-Hastings, utilizando uma proposta baseada na distribuição condicional. Isto nos permite ter uma condicional completa proporcional a verossimilhança, acarretando uma melhor convergência da cadeia, e que utiliza a estrutura de correlação entre os efeitos.

Comparação dos Modelos

Utilizando os critérios definidos na Seção 2.6, podemos fazer a comparação dos nove modelos descritos anteriormente.

Então, para os modelos Gama estabelecidos, e isto inclui também os modelos Exponencial, ou seja, modelos 1, 2, 3, 4, 5 e 6, teremos a expressão $d(Y_{novo}, Y_{obs})$ (2.27), tal que:

$$\begin{aligned}
d(\mathbf{Y}_{novo}, \mathbf{X}_{novo}; \mathbf{Y}_{obs}, \mathbf{X}_{obs}) &= 2 \sum_{t=1}^T \sum_{i=1}^n \left\{ (1 - X_{obs,t}(s_i)) \log \left(\frac{(1 - X_{obs,t}(s_i))}{(1 - X_{novo,t}(s_i))} \right) \right. \\
&+ X_{obs,t}(s_i) \log \left(\frac{X_{obs,t}(s_i)}{X_{novo,t}(s_i)} \right) \left. \right\} \\
&+ 2 \sum_{t=1}^T \sum_{i=1}^n (1 - X_{obs,t}(s_i)) \left\{ -\log \left(\frac{Y_{obs,t}(s_i)}{Y_{novo,t}(s_i)} \right) \right. \\
&+ \left. \left\{ \frac{Y_{obs,t}(s_i) - Y_{novo,t}(s_i)}{Y_{novo,t}(s_i)} \right\} \right\}.
\end{aligned}$$

Onde $X_{obs,t}(s_i) = I(Y_{obs,t}(s_i) > 0)$ e $X_{novo,t}(s_i) = I(Y_{novo,t}(s_i) > 0)$.

A Tabela 5.2 contém os valores de DIC, EQM e de $E[d(\mathbf{Y}_{novo}, \mathbf{Y}_{obs}) | \mathbf{Y}_{obs}, M_i]$ $i = 1, \dots, 9$, para os nove modelos utilizados. Para os modelos lognormal não apresentaremos os valores de $E[d(\mathbf{Y}_{novo}, \mathbf{Y}_{obs}) | \mathbf{Y}_{obs}, M_i]$, por esta não pertencer à família exponencial.

Tabela 5.2: Comparação entre os modelos utilizados

| Modelo | DIC | \bar{D} | pD | $E[d(\cdot) \cdot]$ | EQM |
|--------|----------|-----------|--------|-----------------------|---------|
| 1 | 11797,43 | 11516,62 | 280,81 | 7847,23 | 118,57 |
| 2 | 11778,02 | 11497,80 | 280,22 | 7830,06 | 118,027 |
| 3 | 11670,92 | 11545,46 | 125,46 | 7838,10 | 117,21 |
| 4 | 10660,96 | 9796,04 | 864,91 | 3125,58 | 34,58 |
| 5 | 10656,47 | 9792,96 | 863,51 | 3163,04 | 35,18 |
| 6 | 10550,87 | 9688,32 | 862,55 | 2738,55 | 34,97 |
| 7 | 11076,27 | 10340,22 | 736,05 | – | 227,76 |
| 8 | 11012,46 | 10308,78 | 703,67 | – | 225,60 |
| 9 | 11013,66 | 10302,44 | 711,21 | – | 211,44 |

Podemos notar, que de maneira geral, os modelos Gama (4, 5 e 6) saíram-se melhor que os demais. Dentre estes, optou-se por escolher o modelo 4, com apenas uma semana de defasagem no *logit*, pois, embora os outros modelos possuam

menores valores para o DIC, não apresentaram estimativas significativas para os parâmetros γ_2 e γ_3 .

A Tabela 5.3 apresenta os valores das estimativas a posteriori dos parâmetros do modelo escolhido (média e intervalo de 95 % credibilidade). Foram rodadas 10.000 iterações descartando-se as primeiras 5.000 e salvando-se apenas as amostras de 5 em 5. A cadeia foi iniciada em dois pontos distintos a fim de monitorarmos sua convergência.

Tabela 5.3: Estimativas a posteriori do Modelo 4

| Parâmetro | Média | Quantil 2,5% | Quantil 97,5% |
|--------------|--------|--------------|---------------|
| α_0 | 2,0820 | 1,6998 | 2,4601 |
| γ_0 | 6,9341 | 3,5641 | 10,849 |
| γ_1 | 2,2424 | 0,3452 | 4,8013 |
| σ_1^2 | 168,01 | 57,658 | 450,64 |
| σ_2^2 | 2,8716 | 2,1829 | 3,8809 |
| ϕ_1 | 0,0063 | 0,0031 | 0,0107 |
| ϕ_2 | 0,0091 | 0,0067 | 0,0139 |
| δ | 4,3685 | 3,7566 | 5,0511 |

Todos os parâmetros apresentaram-se significativos. Destaca-se o valor de δ e seu intervalo de credibilidade, pois não há indícios de que pudéssemos utilizar a distribuição Exponencial, porque o valor 1 (um) não está incluso no intervalo. O valor de σ_1^2 também apresentou-se bastante elevado, bem como um largo intervalo de credibilidade.

Modelos com nível variando no tempo

Podemos também imaginar que os níveis γ_0 e α_0 , podem não ser necessariamente fixos no tempo. A Figura 5.1 evidencia que tanto a probabilidade, quanto o volume de chuva podem ser distintos para cada tempo (semana). Então, em (5.1)

e (5.2), podemos ter agora, γ_{0t} e α_{0t} , $t = 1, \dots, 75$, respectivamente. A evolução temporal pode ser feita através de um passeio aleatório. Portanto teríamos para a probabilidade de chuva:

$$\begin{aligned} \log \frac{\theta_t(s)}{1 - \theta_t(s)} &= \gamma_{0t} + \sum_{j=1}^p \gamma_j I(Y_{t-j}(s) > 0) + Z_t(s) \\ \gamma_{0t} &= \gamma_{0t-1} + w_{\gamma t} \quad w_{\gamma t} \sim N(0, W_\gamma). \end{aligned}$$

E para volumes de chuva maiores que zero, teríamos:

$$\begin{aligned} g(\lambda_t(s)) &= \alpha_{0t} + S_t(s) \\ \alpha_{0t} &= \alpha_{0t-1} + w_{\alpha t} \quad w_{\alpha t} \sim N(0, W_\alpha). \end{aligned}$$

Como o modelo Gama foi o que se adaptou melhor aos dados, quando estávamos trabalhando com os níveis fixos no tempo, optou-se por testar mais dois modelos, agora com níveis variando no tempo. Estes modelos estão descritos na Tabela 5.4.

Tabela 5.4: Modelos utilizados com níveis variando no tempo

| Modelo | $Y_t(s)$ | <i>logit</i> | $\log(\lambda_t(s)^{-1})$ |
|--------|-----------------------------------|---------------------------------------|---------------------------|
| 10 | $Ga(\delta, \delta \lambda_t(s))$ | $\gamma_0 + \gamma_1 I_1 + Z_t(s)$ | $\alpha_{0t} + S_t(s)$ |
| 11 | | $\gamma_{0t} + \gamma_1 I_1 + Z_t(s)$ | $\alpha_0 + S_t(s)$ |

Os modelos com essas especificações, inserem-se na classe dos MLDG's, definidos no Capítulo 2. Portanto, a estimação de γ_{0t} e de α_{0t} , $t = 1, \dots, 75$ e de W_γ e W_α , é bastante complexa e será feita como descrito na Seção 4.5. A Figura 5.3 ilustra a média a posteriori e o intervalo de credibilidade destes parâmetros.

Tabela 5.5: Comparação entre os modelos utilizados

| Modelo | DIC | \bar{D} | pD | $E[d(\cdot)]$ | EQM |
|--------|----------|-----------|--------|---------------|-------|
| 4 | 10660,96 | 9796,04 | 864,91 | 3125,58 | 34,58 |
| 10 | 10628,53 | 9745,76 | 882,77 | 3113,61 | 34,31 |
| 11 | 10659,19 | 9812,94 | 846,25 | 3305,50 | 35,61 |

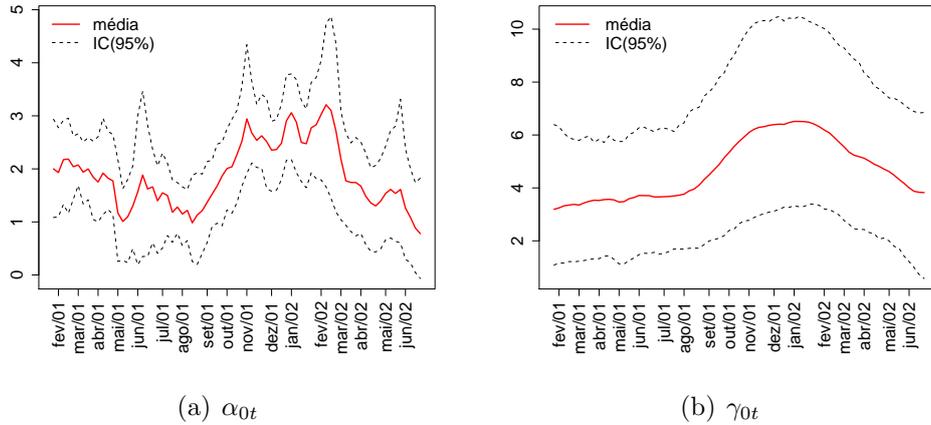


Figura 5.3: Média e Intervalo de 95% de Credibilidade a posteriori para os parâmetros variando no tempo.

Com a inclusão desses dois modelos, precisamos compará-los com o modelo já escolhido anteriormente. Para isto, novamente, foram utilizados os critérios da Seção 2.6. Foram rodadas 10.000 iterações, descartando-se as 5.000 primeiras e salvando-se apenas de 5 em 5 iterações. A Tabela 5.5 apresenta os valores de comparação para estes modelos.

Podemos notar que todos os valores se assemelham bastante, com o modelo 10 obtendo valores um pouco menores que os demais. Optou-se por escolher, então, este modelo, que inclui um nível variando no tempo, na parte que modela níveis de chuva maiores que zero.

5.1.4 Resultados

Definido o modelo final a ser utilizado, podemos agora, observar algumas características a partir dele. A Tabela 5.6 apresenta as estatísticas a posteriori dos parâmetros fixos do modelo; e a Figura 5.4, os respectivos histogramas.

Novamente, apuramos valores significativos para todos os parâmetros. O valor de δ continua significativo diferente de 1(un), supondo o modelo Gama. O parâmetro γ_1 ficou positivo, evidenciando que a existência de chuva na semana an-

Tabela 5.6: Estimativas a posteriori do Modelo 10

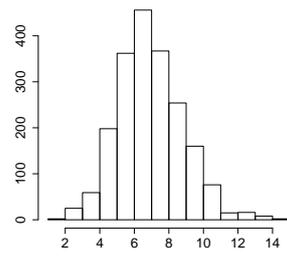
| Parâmetro | Média | Quantil 2, 5% | Quantil 97, 5% |
|--------------|--------|---------------|----------------|
| γ_0 | 5,6039 | 2,5382 | 8,5676 |
| γ_1 | 1,9180 | 0,5105 | 4,1342 |
| σ_1^2 | 100,69 | 33,643 | 205,49 |
| σ_2^2 | 2,5252 | 1,8810 | 3,3439 |
| ϕ_1 | 0,0066 | 0,0037 | 0,0108 |
| ϕ_2 | 0,0011 | 0,0080 | 0,0166 |
| δ | 4,5263 | 3,8796 | 5,2268 |
| W_α | 0,1730 | 0,0877 | 0,3182 |

terior aumenta a probabilidade de chuva na semana corrente. E σ_1^2 continua apresentando um valor alto, o que reflete uma variabilidade muito grande dos efeitos espaço-temporais do *logit*. É o parâmetro mais sensível deste modelo. Uma alternativa para solucionar este problema, seria a inclusão, quando possível, de mais covariáveis para $\theta_t(s)$.

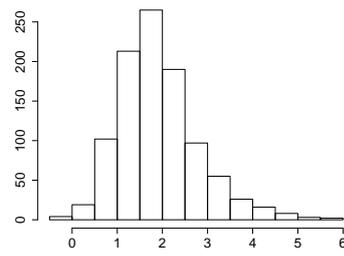
A Figura 5.3(a) ilustra a evolução no tempo da média a posteriori de α_{0t} . Nota-se um ligeiro aumento no nível perto do mês de dezembro de 2001 e um decaimento após março de 2002. Esta época, é justamente o período do verão no hemisfério sul, onde espera-se haver um aumento nos níveis de precipitação.

As Figuras 5.5 e 5.6 apresentam as estatísticas a posteriori dos efeitos espaço-temporais para $\theta_t(s)$ e $\lambda_t(s)$, respectivamente, para algumas estações monitoradoras.

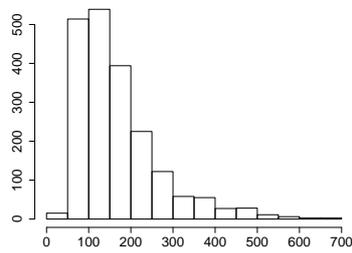
Embora as estimativas não sejam todas significativas, nota-se, pelas médias a posteriori, a grande variabilidade existente nos efeitos espaço temporais do *logit* (efeitos \mathbf{Z}). Como estes efeitos incidem sobre a probabilidade de chuva, esta grande oscilação acompanha a oscilação da existência de chuva. Nota-se em quase todas as estações valores mais altos desses efeitos entre os meses de outubro de 2001 e fevereiro de 2002.



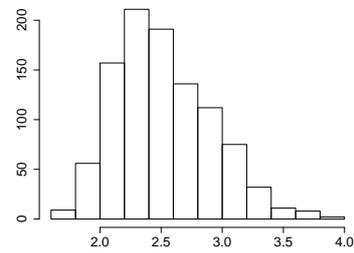
(a) γ_0



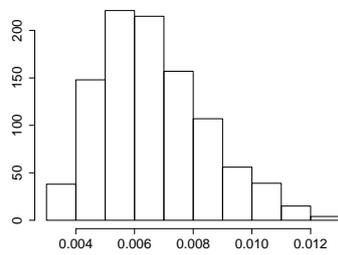
(b) γ_1



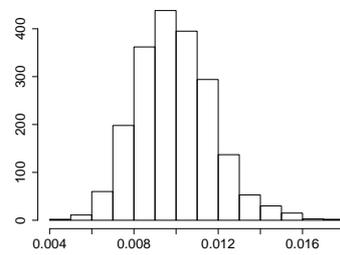
(c) σ_1^2



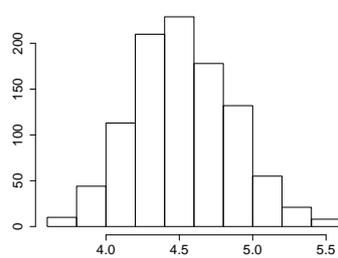
(d) σ_2^2



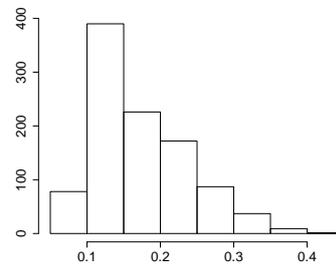
(e) ϕ_1



(f) ϕ_2

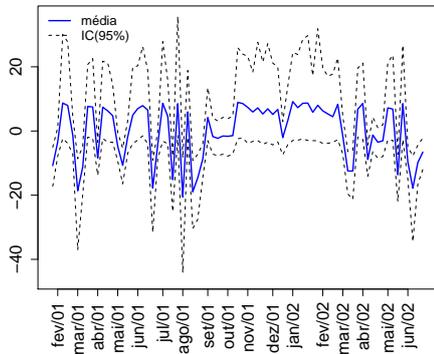


(g) δ

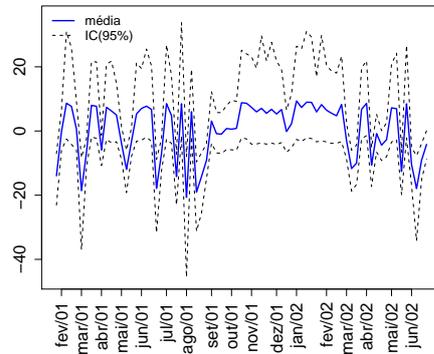


(h) W_α

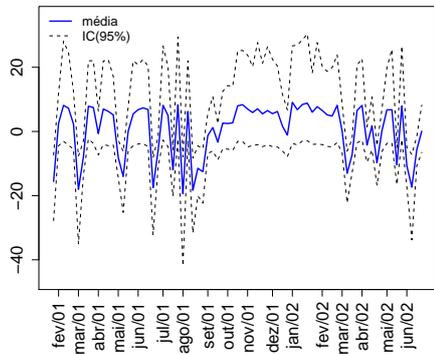
Figura 5.4: Histogramas das amostras das distribuições a posteriori dos parâmetros fixos.



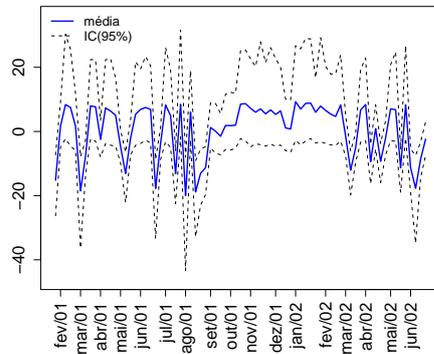
(a) Grajaú



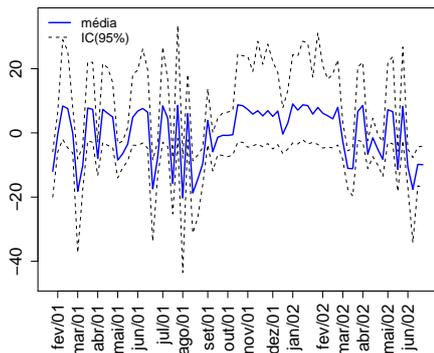
(b) Tanque



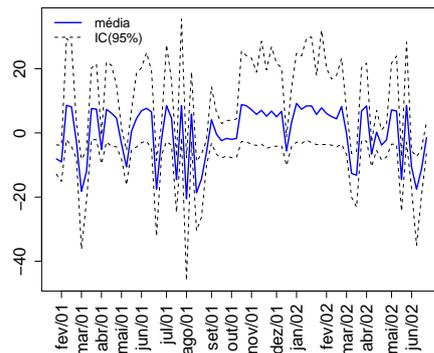
(c) Campo Grande



(d) Bangu



(e) Madureira



(f) Rocinha

Figura 5.5: Média a posteriori dos efeitos espaço-temporais para a probabilidade de chuva em algumas estações monitoradoras, juntamente com os intervalos de 95% de credibilidade.

Para os efeitos espaço-temporais nos níveis de chuva maiores que zero (efeitos **S**), notamos uma menor variabilidade. A Figura 5.6, apresenta a série, no tempo, para algumas estações. Aqui notamos que o comportamento destes efeitos é mais aleatório do que os efeitos **Z**.

A Figura 5.7, ilustra a comparação de grupos de estações, através de suas médias a posteriori, onde claramente notamos diferenças entre elas. No primeiro grupo, a estação de Santa Cruz possui valores menores que as demais, em algumas semanas dos meses de junho de 2001 e abril de 2002, e maiores nas semanas do mês de abril de 2001. No grupo 2 (Figura 5.7(b)) a estação de Guaratiba, que como a estação de Santa Cruz, está mais afastada das demais, na região oeste da cidade, apresenta valores de efeitos bem diferentes das demais, durante alguns períodos.

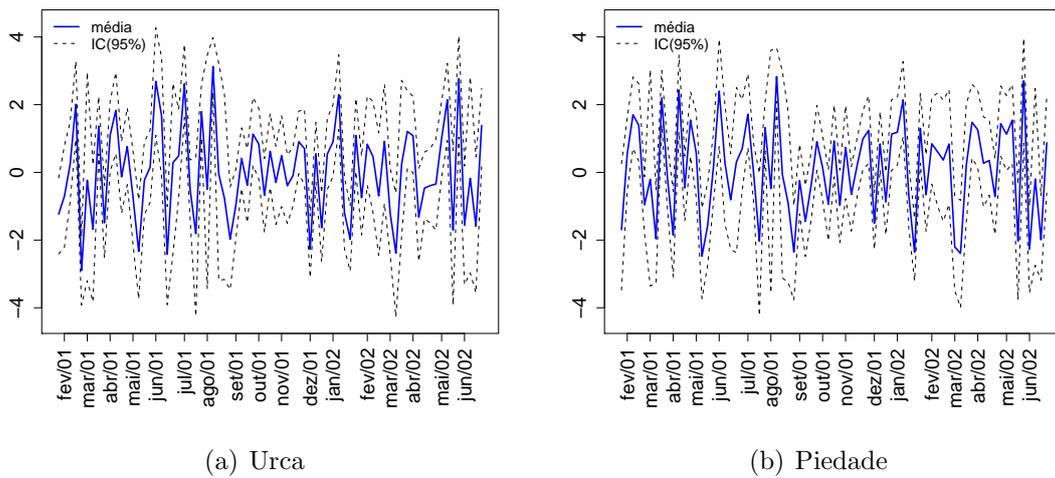
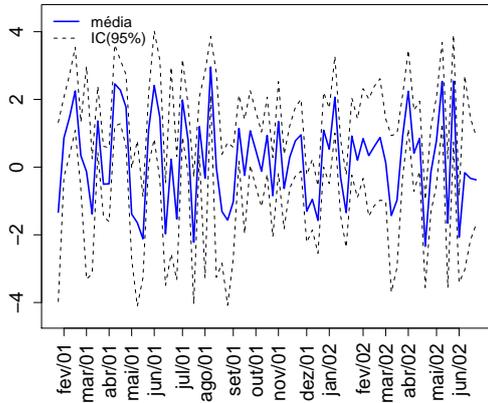
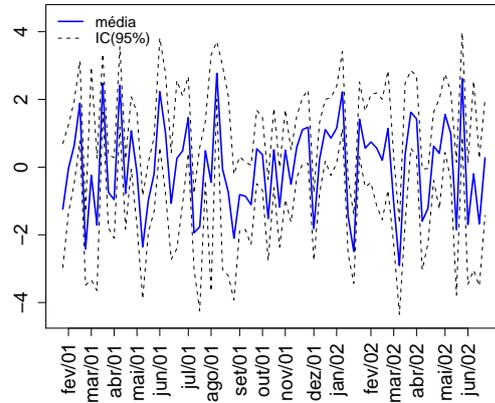


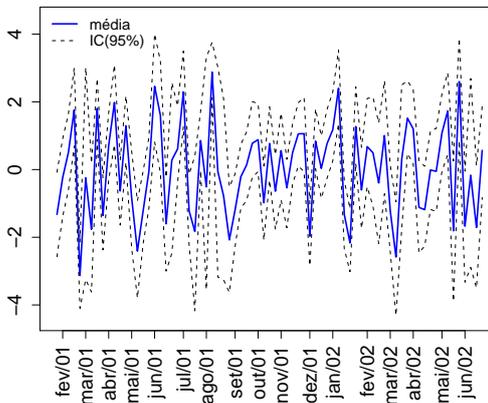
Figura 5.6: Média a posteriori dos efeitos espaço-temporais para os níveis de chuva maiores que 0, em algumas estações monitoradoras, juntamente com os intervalos de 95% de credibilidade.



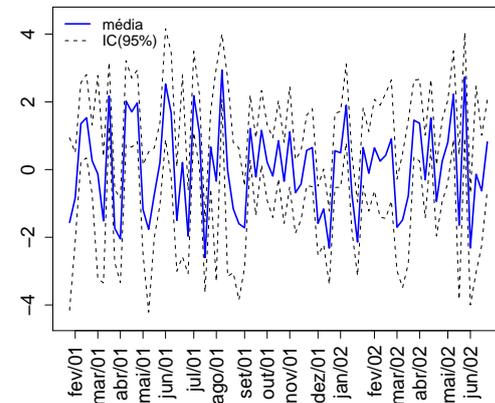
(c) Santa Cruz



(d) Ilha do Governador



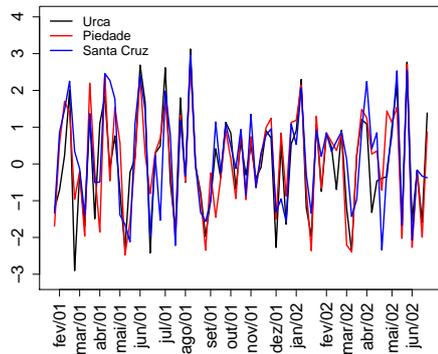
(e) Saúde



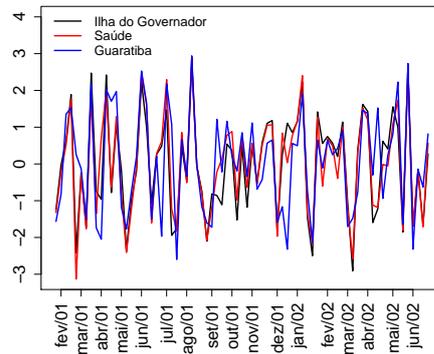
(f) Guaratiba

Figura 5.6: Média a posteriori dos efeitos espaço-temporais para os níveis de chuva maiores que 0, em algumas estações monitoradoras, juntamente com os intervalos de 95% de credibilidade - Continuação.

Obtidas todas as estimativas dos parâmetros, podemos agora obter a distribuição preditiva para algumas estações que foram deixadas de fora da análise, afim de compararmos com o verdadeiro valor observado. Estas estimativas foram feitas com base na equação (3.18). A Figura 5.8 apresenta gráficos com a média a posteriori e com os intervalos de credibilidade, bem como os valores reais de precipitação das estações de Sumaré, Mendanha, Itaúna, Laranjeiras e São Cristóvão.

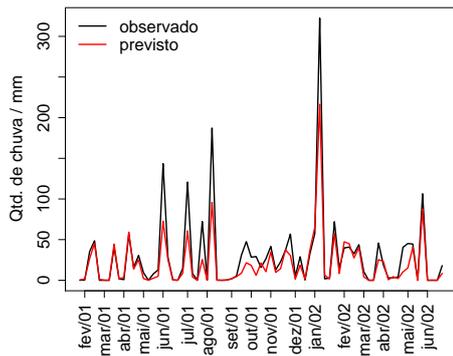


(a) Grupo 1

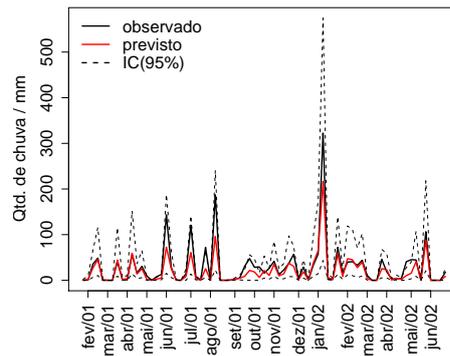


(b) Grupo 2

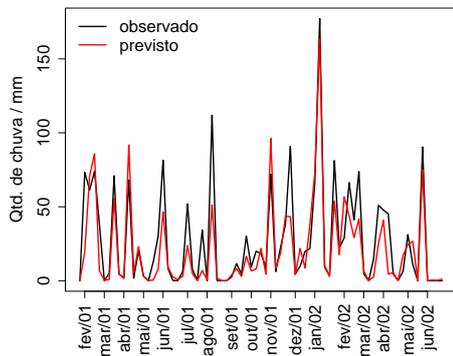
Figura 5.7: Comparação entre os efeitos espaço-temporais dos níveis de chuva maiores que zero.



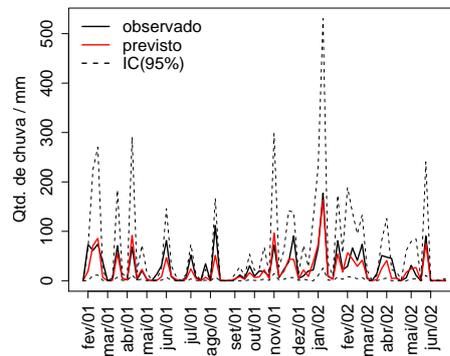
(a) Sumaré - Média



(b) Sumaré - IC(95%)

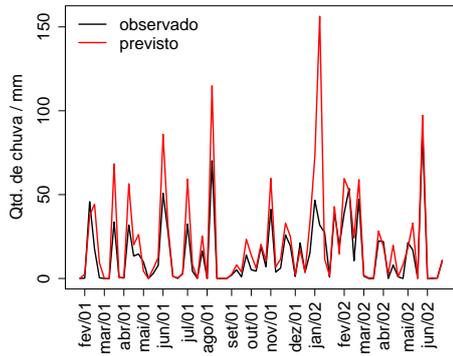


(c) Mendanha - Média

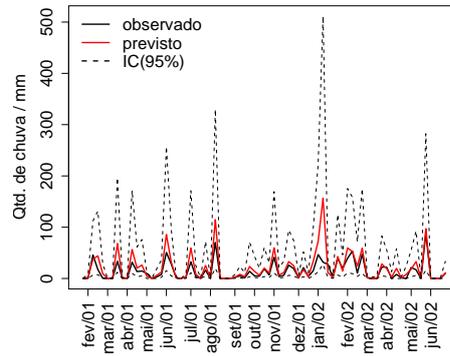


(d) Mendanha - IC(95%)

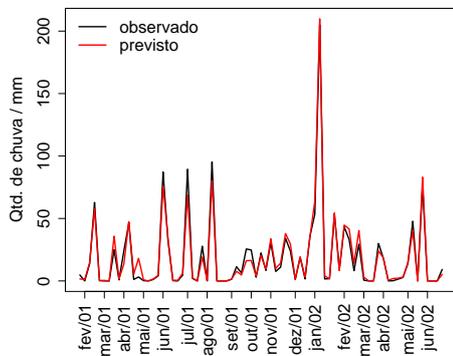
Figura 5.8: Valores previstos para as estações que não foram consideradas no procedimento de inferência.



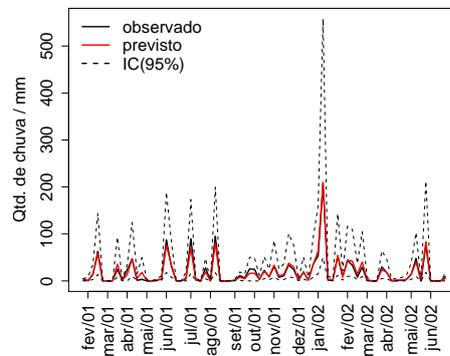
(e) Itaúna - Média



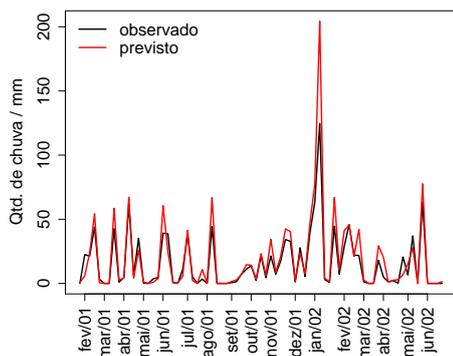
(f) Itaúna - IC(95%)



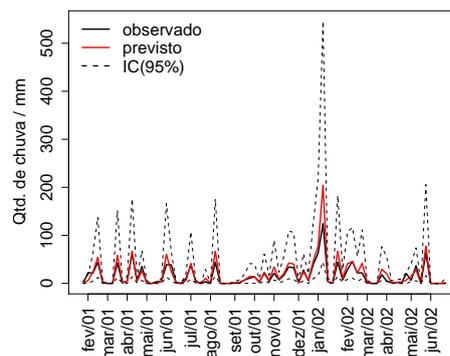
(g) Laranjeiras - Média



(h) Laranjeiras - IC(95%)



(i) São Cristóvão - Média



(j) São Cristóvão - IC(95%)

Figura 5.8: Valores previstos para as estações que não foram consideradas no procedimento de inferência - Continuação.

Podemos notar que as estimativas aproximam-se bastante dos valores reais, conseguindo captar os períodos onde não há chuva e os períodos de maior precipitação. A estação de Laranjeiras (Figuras 5.8(g) e 5.8(h)) destaca-se dentre as demais, obtendo valores bem precisos de suas estimativas. Para as demais também observou-se que todos os intervalos de credibilidade contiam os verdadeiros valores.

Uma das principais vantagens desses modelos é a possibilidade de se descrever a probabilidade a posteriori de $\theta_t(s)$, que em nosso caso, representa a probabilidade de ocorrência de chuva. A Figura 5.9 exibe os gráficos com a estimativa da posteriori de probabilidade de chuva para algumas estações monitoradoras.

Nota-se que há grande variabilidade temporal, por causa dos períodos secos e chuvosos, entretanto há também uma variabilidade entre as estações. Por exemplo, para as estações de Santa Tereza, Irajá, Grota Funda e Jardim Botânico, há uma chance maior que 80% de chuva, nas últimas semanas de dezembro de 2001. Já para a estação de Sepetiba, este valor não chega a 50%. Santa Tereza, nas primeiras semanas de 2001, apresenta uma probabilidade de chuva de quase 80%. Sepetiba, Irajá e Grota Funda, apresentam 0% de probabilidade.

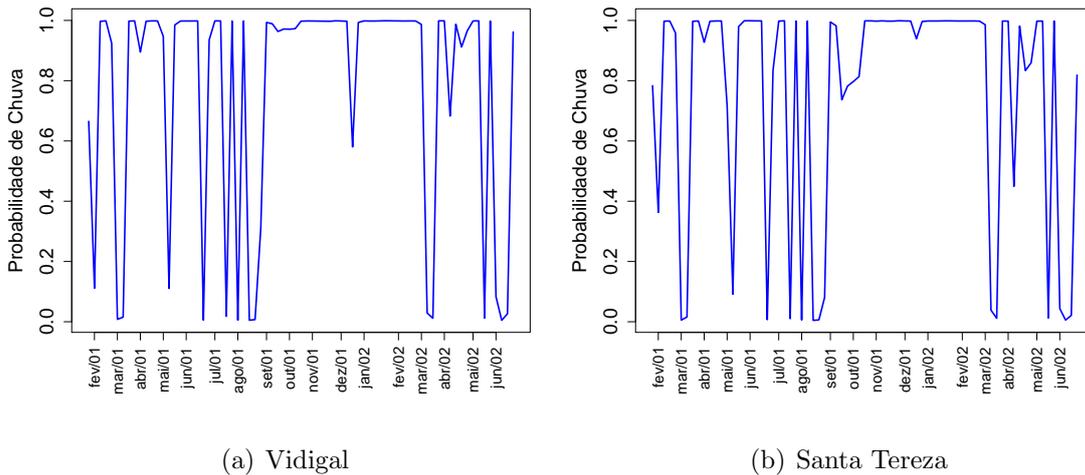
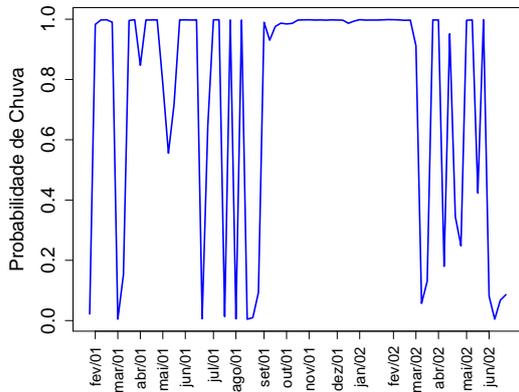
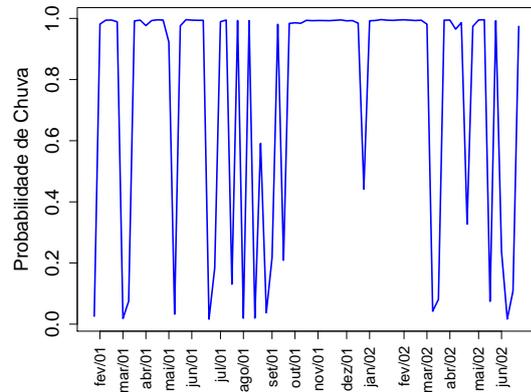


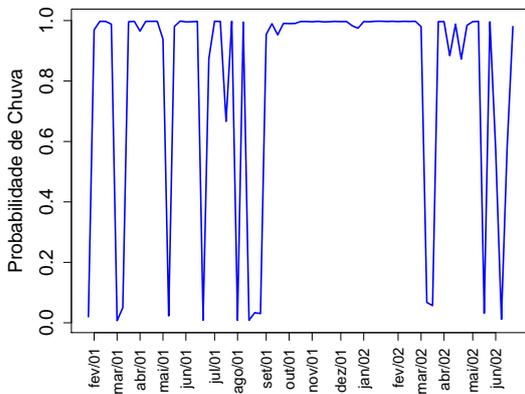
Figura 5.9: Estimativas da média a posteriori das probabilidades de Chuva para algumas estações monitoradoras.



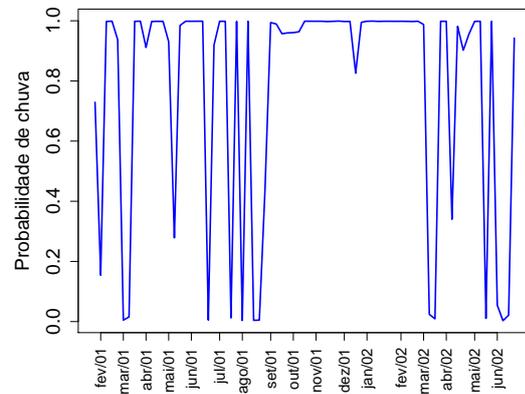
(c) Irajá



(d) Setpetiba



(e) Grota Funda



(f) Jardim Botânico

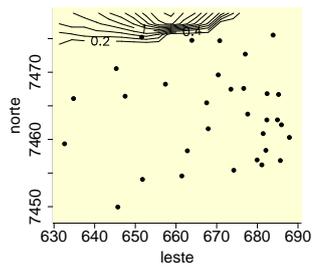
Figura 5.9: Estimativas da média a posteriori das probabilidades de Chuva para algumas estações monitoradoras - Continuação.

Além das estimativas para as estações não medidas, este modelo nos permite fazer previsões para quaisquer outros pontos da cidade. Por isso, uma grade de 400 pontos foi utilizada para a construção de superfícies de previsão, construídas no software R. A Figura 5.10 mostra as superfícies estimadas pelo modelo usando a mediana a posteriori e a Figura 5.11 o intervalo interquartilico, para algumas semanas.

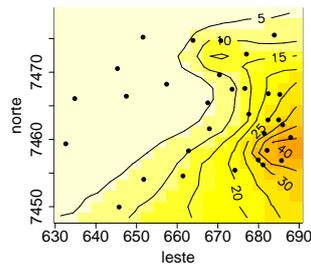
As superfícies construídas no R, possuem uma escala que vai do vermelho (valores maiores), passando pelo amarelo, chegando ao branco (valores menores). As Figu-

ras 5.10 e 5.11, mostram dois grupos de semanas (1^a semana a 12^a semana e 42^a semana a 53^a semana) que são comparáveis entre si.

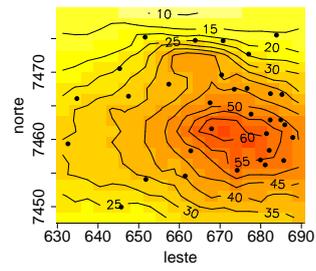
Notamos que as duas primeiras semanas são bem escassas de chuvas, onde visualizamos uma grande região clara. Destacamos a semana 6 onde o modelo estimou o valor zero para toda a superfície. Nas demais, nota-se regiões isoladas, onde há uma maior precipitação (pontos mais escuros) e na semana 51 (verão 2001/2001), uma grande região avermelhada, onde tivemos uma grande precipitação em toda a cidade.



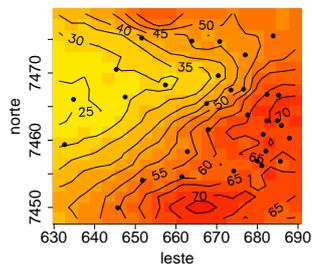
(a) Semana 1



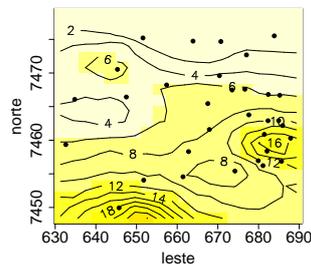
(b) Semana 2



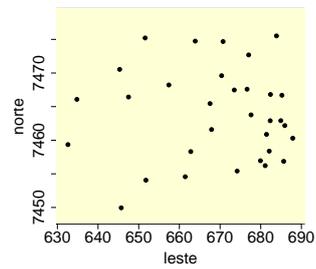
(c) Semana 3



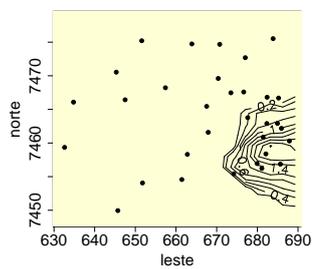
(d) Semana 4



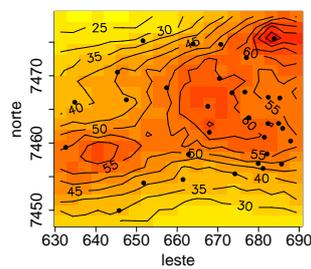
(e) Semana 5



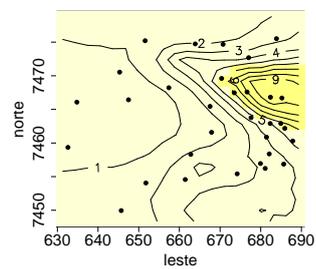
(f) Semana 6



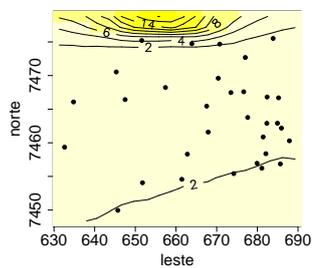
(g) Semana 7



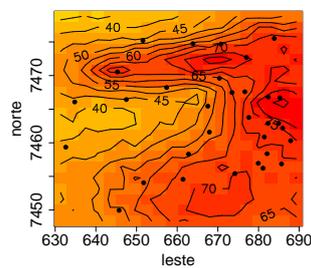
(h) Semana 8



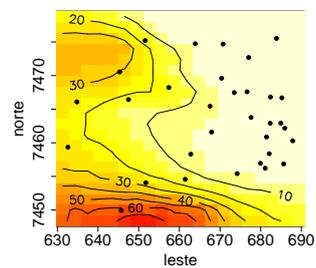
(i) Semana 9



(j) Semana 10

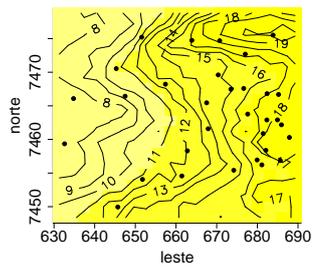


(k) Semana 11

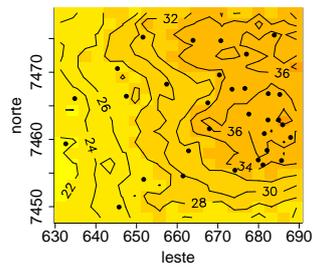


(l) Semana 12

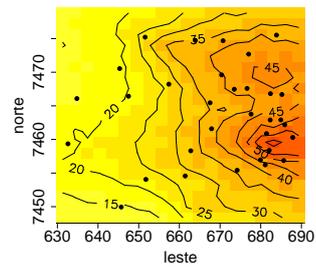
Figura 5.10: Superfícies com Mediana a posteriori da distribuição preditiva para a respectiva semana.



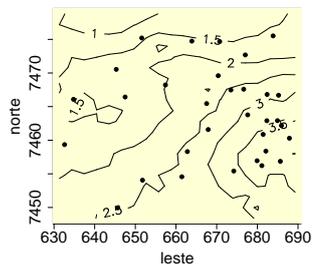
(m) Semana 42



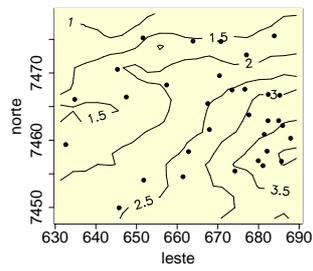
(n) Semana 43



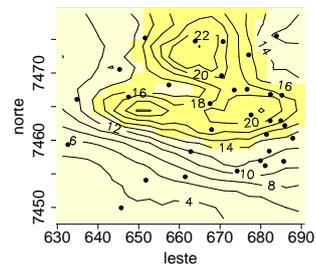
(o) Semana 44



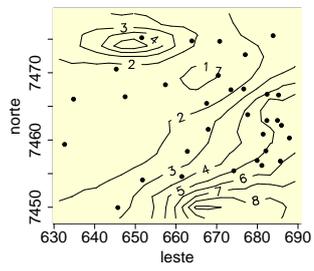
(p) Semana 45



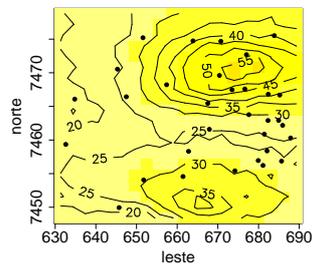
(q) Semana 46



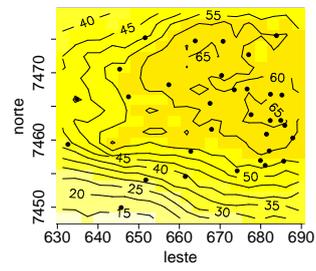
(r) Semana 47



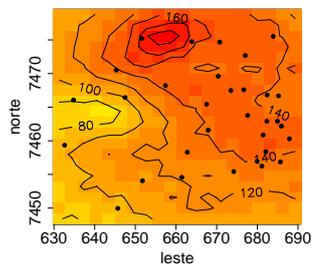
(s) Semana 48



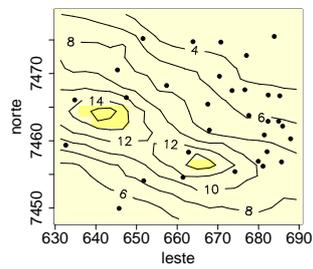
(t) Semana 49



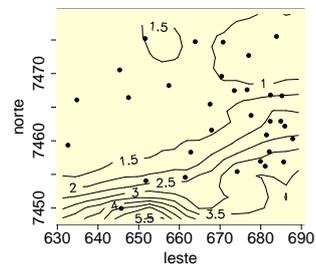
(u) Semana 50



(v) Semana 51

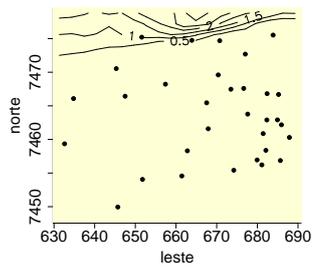


(w) Semana 52

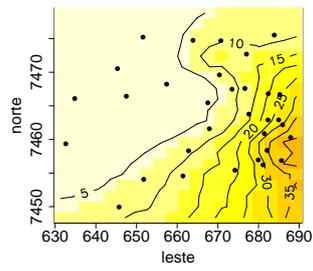


(x) Semana 53

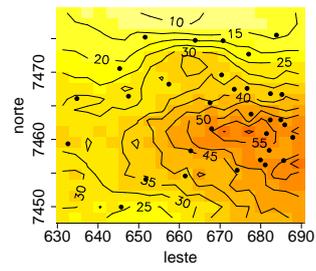
Figura 5.10: Superfícies com Mediana a posteriori da distribuição preditiva para a respectiva semana - Continuação.



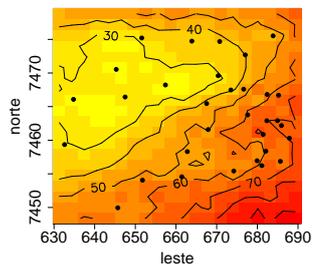
(a) Semana 1



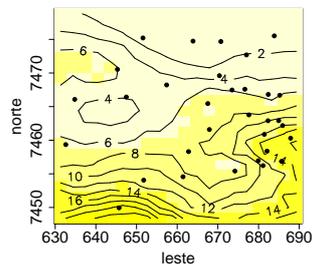
(b) Semana 2



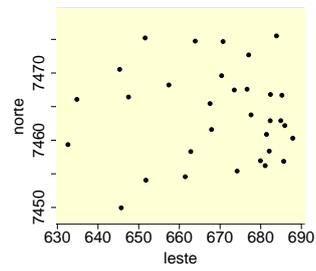
(c) Semana 3



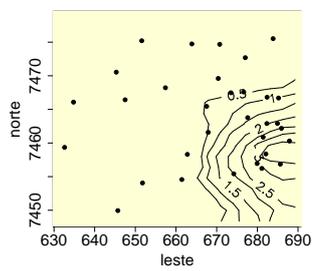
(d) Semana 4



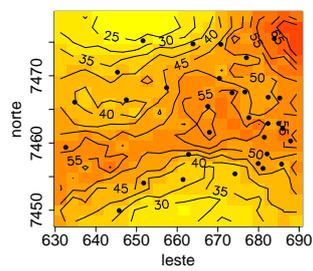
(e) Semana 5



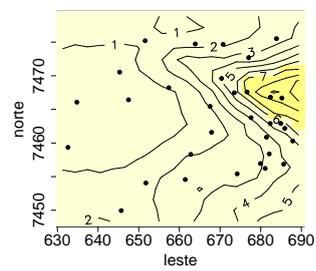
(f) Semana 6



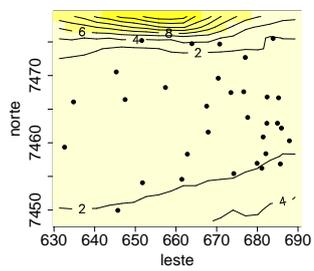
(g) Semana 7



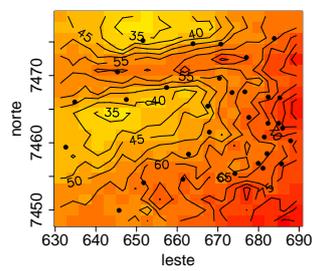
(h) Semana 8



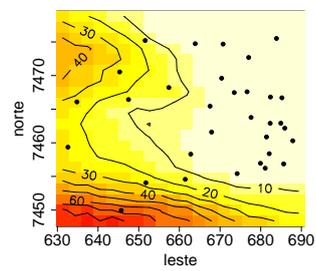
(i) Semana 9



(j) Semana 10

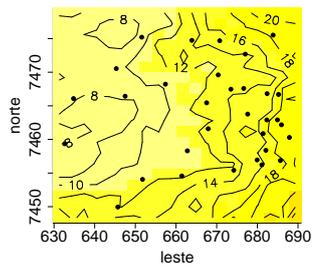


(k) Semana 11

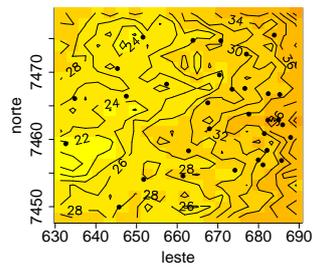


(l) Semana 12

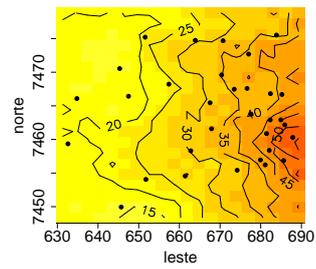
Figura 5.11: Superfícies com Intervalo Interquartilico a posteriori da distribuição preditiva para a respectiva semana.



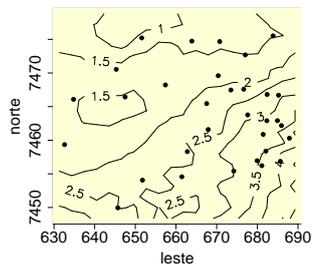
(m) Semana 42



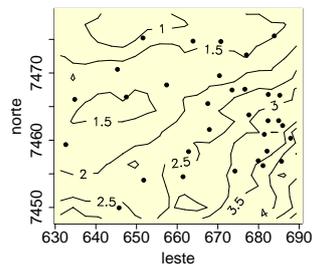
(n) Semana 43



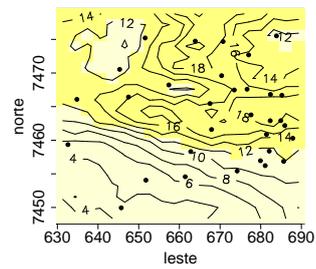
(o) Semana 44



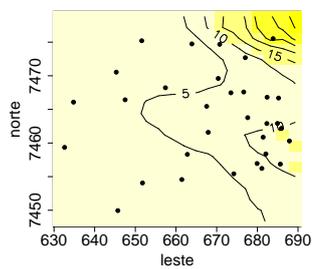
(p) Semana 45



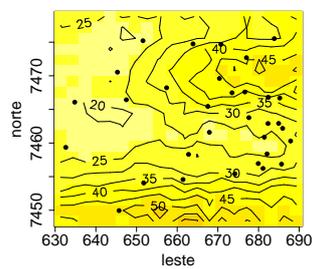
(q) Semana 46



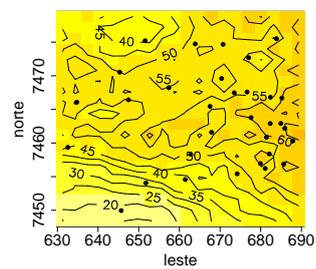
(r) Semana 47



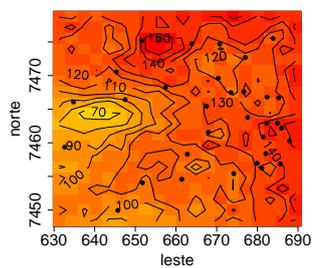
(s) Semana 48



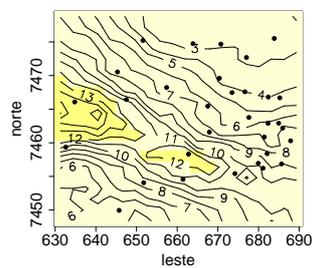
(t) Semana 49



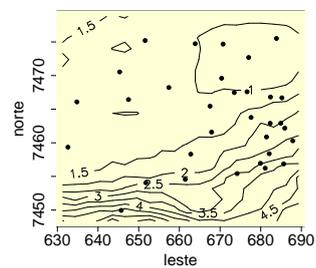
(u) Semana 50



(v) Semana 51



(w) Semana 52



(x) Semana 53

Figura 5.11: Superfícies com Intervalo Interquartilicoa posteriori da distribuição preditiva para a respectiva semana - Continuação.

5.2 Modelando a Dengue na Cidade do Rio de Janeiro

A dengue é uma doença febril aguda, causada por vírus, de evolução benigna, na forma clássica, e, grave, quando se apresenta na forma hemorrágica. A dengue é hoje a mais importante arbovirose que afeta o homem e constitui um sério problema de saúde pública no mundo, especialmente nos países de clima tropical, onde as condições do meio ambiente favorecem o desenvolvimento e a proliferação do mosquito transmissor da doença. O *Aedes aegypti* e o *Aedes albopictus*, são os mosquitos responsáveis pela transmissão da doença.

No Brasil, epidemias de dengue vêm se instalando desde a década de 80. Nos recentes anos de 2001 e 2002, aconteceu a maior delas, na cidade do Rio de Janeiro, quando foram registrados, cerca de 150.000 casos da doença. Ainda hoje o índice de infestação do mosquito é superior ao aconselhável pelo Ministério da Saúde, que todo ano, faz campanhas para evitar um cenário idêntico ao de 2002.

Por isso tudo, a dengue é um tema bastante recorrente em estudos, hoje em dia. Em especial, no estudo de modelos espaço-temporais, onde busca-se entender melhor a dinâmica da doença, tanto no espaço territorial quanto no tempo. Ferreira (2003) faz uma análise espaço-temporal dos casos de dengue na cidade do Rio de Janeiro, focando o estudo nos anos de 2001 e 2002.

Neste estudo, Ferreira (2003), cita a existência de semanas onde há pouca notificação de casos de dengue (períodos pré e pós-epidemicos) e que poderiam prejudicar a modelagem, utilizando a distribuição Poisson. Neste caso, a quantidade de zeros observada é maior do que a predita pela distribuição, apresentando uma heterogeneidade muito grande.

A aplicação que faremos agora, utiliza os mesmos dados utilizados por Ferreira (2003), entretanto adicionando algumas semanas não aproveitadas por haver excesso de zeros. Usaremos os modelos propostos na Seção 3.3, afim de considerar

toda essa quantidade de zeros.

5.2.1 Conjunto de dados

O conjunto de dados utilizado, corresponde às notificações de casos de dengue na cidade do Rio de Janeiro, entre os anos de 2001 e 2002. São, ao todo, 77 semanas, iniciando-se na primeira semana epidemiológica de 2001, em 1° de janeiro e terminando na 25ª semana epidemiológica de 2002, em 22 de junho.

Foram considerados 156 bairros da cidade, excluindo-se os bairros de Paquetá, Vasco da Gama e Parque Columbia. Para os bairros Joá e Grumari, onde temos observações em apenas determinadas semanas, optou-se por fazer imputação de valores, considerando o próprio modelo utilizado.

A Figura 5.12 expõe a série de notificações para dois bairros distintos nos anos de 2001 e 2002.

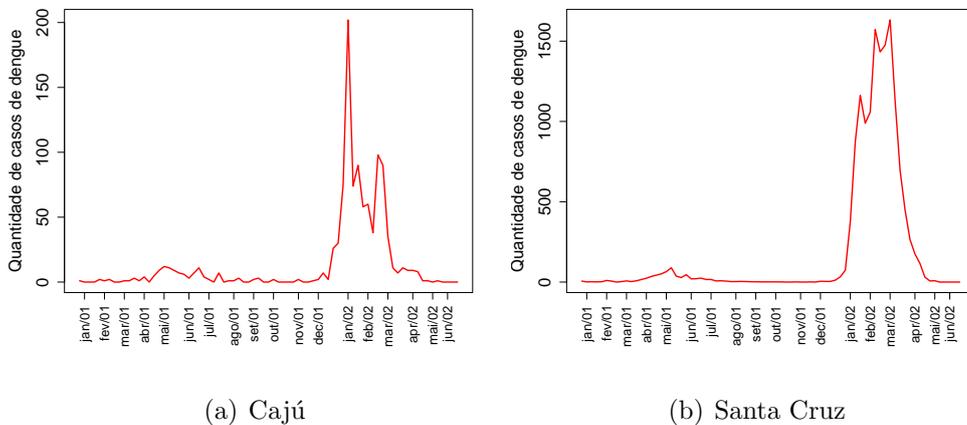


Figura 5.12: Quantidade de notificações de dengue durante as 77 semanas.

5.2.2 Modelos Utilizados

Como dito anteriormente, utilizaremos a idéia introduzida na Seção 3.3, para a modelagem de casos de dengue. Então consideremos Y_{it} , $i = 1, \dots, 156$ e $t =$

$1, \dots, 77$, como a notificação do número de casos de dengue no bairro i , e semana t . Podemos construir um modelo da seguinte forma:

$$p(Y_{it} | \theta_{it}, \lambda_{it}) = \begin{cases} (1 - \theta_{it}) + \theta_{it}p(Y_{it} = 0 | \lambda_{it}) & \text{se } Y_{it} = 0, \\ \theta_{it}p(Y_{it} | \lambda_{it}) & \text{se } Y_{it} > 0. \end{cases} \quad (5.3)$$

Vale lembrar que nesse caso $1 - \theta_{it}$, representa a probabilidade de obtermos valores iguais a zero, não provenientes da distribuição $p(Y_{it} | \lambda_{it})$, que neste caso, será uma Poisson. Como mencionado na Seção 3.3, outras distribuições também poderiam ser utilizadas, como a Binomial Negativa, mas nos concentraremos apenas nesta.

Devemos, agora atribuir estruturas para os parâmetros. Para a modelagem de θ_{it} , assumiremos:

$$\log \frac{\theta_{it}}{1 - \theta_{it}} = \gamma_0 + \gamma_1 I(Y_{it-1} > 0). \quad (5.4)$$

Teremos γ_0 como efeito fixo e γ_1 , como efeito da existência de casos de dengue na semana anterior. Note, que neste estágio, não estaremos utilizando efeitos espaço-temporais. Em modelos para variáveis discretas, os componentes do *logit* são muito influenciados pela quantidade de zeros observada, fato também observado por Agarwal et al. (2002). Neste caso, o modelo foi testado com a existência dos efeitos espaço-temporais no *logit*. Como não houve um bom comportamento dos parâmetros, em geral, optou-se por retirar estas componentes do modelo, deixando a incorporação destes efeitos apenas para a modelagem da média da Poisson.

Trabalharemos com a seguinte distribuição:

$$Y_{it} | \lambda_{it}, e_{it} \sim Po(\lambda_{it}e_{it}), \quad i = 1, \dots, 156 \quad t = 1, \dots, 77, \quad (5.5)$$

onde λ_{it} representa o risco relativo de dengue e e_{it} , o número esperado de casos de dengue, no bairro i , e semana t . O número de casos esperados de dengue

pode ser obtido da seguinte maneira:

$$e_{it} = \frac{\sum_{i=1}^n Y_{it}}{\sum_{i=1}^n pop_i} \times pop_i, \quad (5.6)$$

onde pop_i representa o contingente populacional do bairro i , sendo este considerado o mesmo, para todo o período em estudo.

Para a modelagem do risco, teremos a seguinte estrutura:

$$\log(\lambda_{it}) = \alpha_0 + S_{it}, \quad (5.7)$$

onde α_0 é um nível fixo e S_{it} , $i = 1, \dots, 156$, $t = 1, \dots, 77$, são os efeitos espaço-temporais. Para estes efeitos, como estamos trabalhando com dados de área, teremos uma distribuição CAR, associada, portanto:

$$S_{it} \sim CAR(\sigma_S^2). \quad (5.8)$$

5.2.3 Procedimento de Inferência

De posse dos valores observados $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_{156})'$, onde $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{i77})'$, $i = 1, \dots, 156$, a verossimilhança aumentada desse modelo, de acordo com (??), será dada por:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) = \prod_{i=1}^{156} \prod_{t=1}^{77} \theta_{it}^{1-X_{it}} (1 - \theta_{it})^{X_{it}} \prod_{i=1}^{156} \prod_{t=1}^{77} p(Y_{it} | \lambda_{it})^{1-X_{it}}. \quad (5.9)$$

Podemos, então, obter a seguinte log-verossimilhança:

$$\begin{aligned} l(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) &= \sum_{i=1}^{156} \sum_{t=1}^{77} [(1 - X_{it}) \log \theta_{it} + X_{it} \log(1 - \theta_{it}) \\ &+ (1 - X_{it}) \log \left(\frac{(\lambda_{it} e_{it})^{Y_{it}} \exp\{-\lambda_{it} e_{it}\}}{Y_{it}!} \right)]. \end{aligned}$$

Depois de obtida a verossimilhança, precisamos atribuir prioris aos parâmetros, afim de aplicarmos o Teorema de Bayes. Novamente adotaremos a prática estabelecida no Capítulo 3. Portanto, consideraremos, como $\boldsymbol{\Theta}$, nosso vetor paramétrico,

$$\boldsymbol{\Theta} = \{\boldsymbol{\theta}, \boldsymbol{\lambda}, \gamma_0, \gamma_1, \alpha_0, \mathbf{S}, \sigma_S^2, \mathbf{X}\}.$$

Agarwal et al. (2002) discutem prioris bem informativas para os efeitos que estão no *logit*, ou seja, para γ_0 e γ_1 , quando estamos trabalhando na presença da variável latente. O procedimento adotado foi estimar valores para as médias das prioris desses dois parâmetros e atribuímos variâncias pequenas. Essas estimativas foram obtidas através da média a posteriori desses parâmetros, resultante do mesmo modelo, sem a utilização da variável latente. Portanto, teremos como prioris $\gamma_0 \sim N(1, 1; 0, 1)$ e $\gamma_1 \sim N(4, 9; 0, 1)$.

Para α_0 , teremos uma priori normal vaga, fazendo $\sigma_\alpha^2 = 100$. Para os efeitos temporais S_{it} , $i = 1, \dots, 156$, $t = 1, \dots, 77$, teremos uma priori CAR, como descrito em (3.12) e (3.13). Para σ_S^2 teremos uma Gama Invertida com $a_{\sigma_S^2}$ e $b_{\sigma_S^2} = 2$.

Como dito anteriormente, as variáveis latentes X_{it} , não são observadas e precisam ser estimadas a cada iteração do MCMC. Como descrito na Seção 3.3, é fácil ver que, $X_{it} \sim \text{Bern}(1 - \theta_{it})$.

Obtemos a posteriori para este modelo, de maneira que:

$$p(\Theta | \mathbf{Y}) \propto p(\mathbf{Y} | \boldsymbol{\theta}, \boldsymbol{\lambda})p(\boldsymbol{\theta} | \gamma_0, \gamma_1)p(\mathbf{X} | \boldsymbol{\theta})p(\boldsymbol{\lambda} | \alpha_0, \mathbf{S})p(\mathbf{S} | \sigma_S^2)p(\gamma_0)p(\gamma_1)p(\alpha_0)p(\sigma_S^2). \quad (5.10)$$

Desta forma, podemos obter as condicionais completas para cada parâmetro:

(i) Condicional completa para γ_0 :

$$p(\gamma_0 | \mathbf{Y}, \Theta_{-\gamma_0}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_0}^2}(\gamma_0 - \mu_{\gamma_0})^2 \right\} \exp \left\{ nT\gamma_0 - \gamma_0 \sum_{i=1}^{156} \sum_{t=1}^{77} X_{it} - \sum_{i=1}^{156} \sum_{t=1}^{77} \log(1 + \exp(\gamma_0 + \gamma_1 I(Y_{it-1} > 0))) \right\};$$

(ii) Condicional completa para γ_1 :

$$p(\gamma_1 | \mathbf{Y}, \Theta_{-\gamma_1}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_1}^2}(\gamma_1 - \mu_{\gamma_1})^2 \right\} \exp \left\{ \gamma_1 \sum_{i=1}^{156} \sum_{t=1}^{77} I(Y_{it-1} > 0) - \gamma_1 \sum_{i=1}^{156} \sum_{t=1}^{77} X_{it} I(Y_{it-1} > 0) - \sum_{i=1}^{156} \sum_{t=1}^{77} \log(1 + \exp(\gamma_0 + \gamma_1 I(Y_{it-1} > 0))) \right\};$$

(iii) Condicional completa para X_{it} :

$$p(X_{it} | \mathbf{Y}, \Theta_{-X_{it}}) \propto \exp \left\{ \sum_{i=1}^{156} \sum_{t=1}^{77} X_{it} [-\log(1 + \exp(\gamma_0 + \gamma_1 I(Y_{it-1} > 0)))] \right. \\ \left. + (1 - X_{it}) [(\gamma_0 + \gamma_1 I(Y_{it-1} > 0)) - \log(1 + \exp(\gamma_0 + \gamma_1 I(Y_{it-1} > 0)))] \right\};$$

(iv) Condicional completa para α_0 :

$$p(\alpha_0 | \mathbf{Y}, \Theta_{-\alpha_0}) \propto \exp \left\{ -\frac{1}{2\sigma_{\alpha_0}^2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} \exp \left\{ \alpha_0 \sum_{i=1}^{156} \sum_{t=1}^{77} (1 - X_{it}) Y_{it} \right. \\ \left. - \sum_{i=1}^{156} \sum_{t=1}^{77} (1 - X_{it}) \exp \{ \alpha_0 + S_{it} \} e_{it} \right\};$$

(v) Condicional completa para S_{it} , $i = 1, \dots, 156$ e $t = 1, \dots, 77$:

$$p(S_{it} | \mathbf{Y}, \mathbf{X}, \Theta_{-S_{it}}) \propto \exp \left\{ -\frac{n_i}{2\sigma_S^2} (S_{it} - \mu_{it})^2 \right\} \exp \left\{ (1 - X_{it}) Y_{it} S_{it} \right. \\ \left. - (1 - X_{it}) \exp \{ \alpha_0 + S_{it} \} e_{it} \right\};$$

(vi) Condicional completa para σ_S^2

$$p(\sigma_S^2 | \mathbf{Y}, \Theta_{-\sigma_S^2}) \propto \sigma_S^{2-(a_2+1)} \exp \left\{ -\frac{b_2}{\sigma_S^2} \right\} \sigma_S^{2\left(-\frac{156 \times 77}{2}\right)} \exp \left\{ \sum_{i=1}^{156} \sum_{t=1}^{77} -\frac{n_i}{2\sigma_S^2} (S_{it} - \mu_{it})^2 \right\}.$$

A condicional completa para σ_S^2 possui forma fechada, conhecida, isto é, $\sigma_S^2 \sim IG \left(a_2 + \frac{nT}{2}, b_2 + \sum_{i=1}^{156} \sum_{t=1}^{77} \frac{n_i (S_{it} - \mu_{it})^2}{2} \right)$.

Para γ_0 , γ_1 e α_0 aplicamos o algoritmo ARMS, pois esses parâmetros possuem condicionais completas log-côncavas. E para os efeitos espaço-temporais, utilizamos Metroropolis-Hastings, empregando uma normal como densidade proposta. Esta normal segue a mesma especificação da priori CAR. Com isso teremos novamente uma condicional completa proporcional à verossimilhança, agilizando a convergência destes parâmetros.

A Tabela 5.7 apresenta as estimativas a posteriori dos parâmetros fixos do modelo

Tabela 5.7: Estimativas a posteriori do Modelo Poisson com nível (α_0) fixo

| Parâmetro | Média | Quantil 2,5% | Quantil 97,5% |
|--------------|---------|--------------|---------------|
| γ_0 | 1,1668 | 1,0081 | 1,3495 |
| γ_1 | 4,9002 | 4,2482 | 5,8501 |
| α_0 | -0,3153 | -0,3303 | -0,3006 |
| σ_S^2 | 2,6007 | 2,4166 | 2,8049 |

Modelo com nível variando no tempo

Novamente, se analisarmos as Figuras 5.12(a) e 5.12(b), que ilustram a quantidade de dengue em dois bairros da cidade, notamos que há um acréscimo grande de notificações perto do verão de 2001/2002, que foi o auge da epidemia. Esta tendência se replica para quase todos os bairros. Portanto, é de se esperar que o nível (para quantidade de casos maiores que zero) desta série, não seja exatamente o mesmo durante todo o tempo. Por isso, substituiremos a equação (5.7) por uma, com nível variando no tempo. Conseqüentemente:

$$\log(\lambda_{it}) = \alpha_{0t} + S_{it} \quad (5.11)$$

$$\alpha_{0t} = \alpha_{0t-1} + w_t \quad w_t \sim N(0, W_\alpha) \quad (5.12)$$

Novamente, a estimação dos parâmetros α_{0t} e W_α seguirá a conceituação vista na Seção 4.5. Para compararmos este modelo com o outro, onde temos um nível fixo, utilizaremos o DIC, o EQM e um EPD aproximado, baseado na distribuição normal. A Tabela 5.8 apresenta os valores para a comparação entre os modelos. O Modelo 1 é modelo que utiliza o nível fixo no tempo (α_0) e o Modelo 2 utiliza o nível variando (α_{0t}).

Apesar do DIC ser menor para o Modelo 1, o valor de \bar{D} é menor para o Modelo 2. Os valores de EPD e EQM também saíram-se melhores para o Modelo 2. Por esta razão e também pelo próprio comportamento da série optou-se pela escolha do Modelo 2.

Tabela 5.8: Comparação entre os modelos Poisson

| Modelo | DIC | \bar{D} | pD | G | P | EPD | EQM |
|--------|----------|-----------|---------|--------|----------|-----------|-------|
| 1 | 52471,59 | 45065,51 | 7406,08 | 402932 | 238683,3 | 641615,30 | 33,54 |
| 2 | 51369,73 | 43552,53 | 7817,20 | 262168 | 274177,1 | 536345,50 | 21,82 |

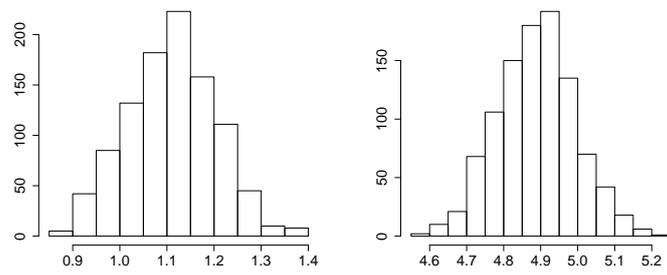
5.2.4 Resultados

Após escolhido o modelo com nível variando no tempo, podemos analisar alguns resultados provenientes do mesmo. Primeiramente, a Tabela 5.9 apresenta as estimativas a posteriori dos parâmetros fixos do modelo e a Figura 5.13 seus respectivos histogramas.

Tabela 5.9: Estimativas a posteriori do Modelo Poisson com nível (α_{0t}) variando no tempo

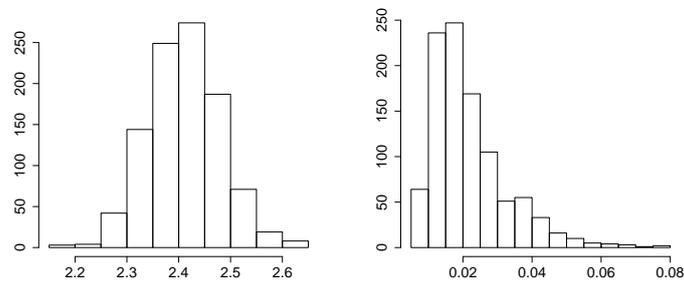
| Parâmetro | Média | Quantil 2,5% | Quantil 97,5% |
|--------------|--------|--------------|---------------|
| γ_0 | 1,1105 | 0,9319 | 1,2893 |
| γ_1 | 4,8845 | 4,6973 | 5,0560 |
| σ_S^2 | 2,4228 | 2,2894 | 2,5542 |
| W_α | 0,0280 | 0,0083 | 0,0488 |

Podemos notar que todos se apresentaram significativos. O valor positivo de γ_1 , indica que a existência de casos de dengue na semana anterior aumenta as chances de não termos excesso de zeros na semana corrente. A Figura 5.14 apresenta a média e intervalo de credibilidade a posteriori de α_{0t} . Notamos que há pequenas oscilações, com uma leve tendência a um decaimento com o decorrer do tempo.



(a) γ_0

(b) γ_1



(c) σ_S^2

(d) W

Figura 5.13: Histogramas das amostras das distribuições a posteriori dos parâmetros fixos.

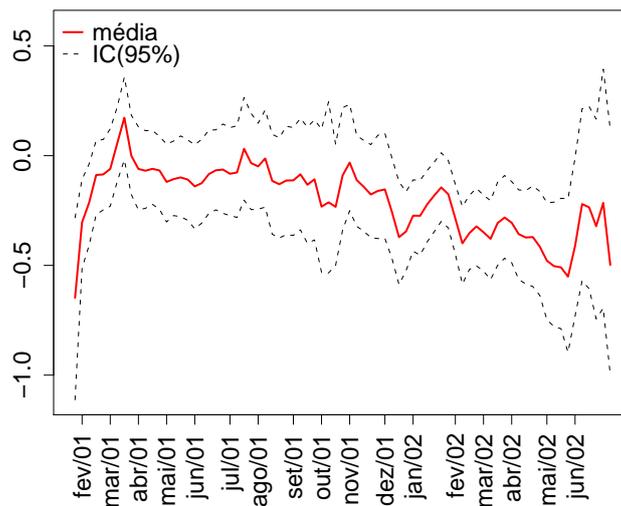
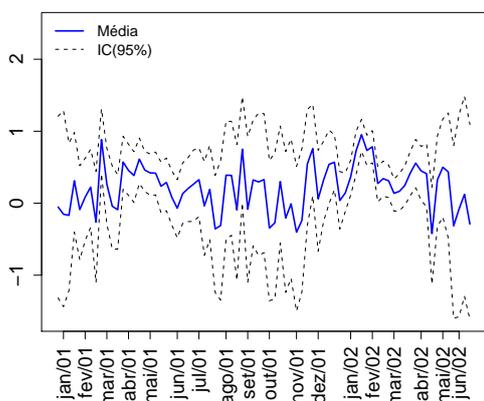
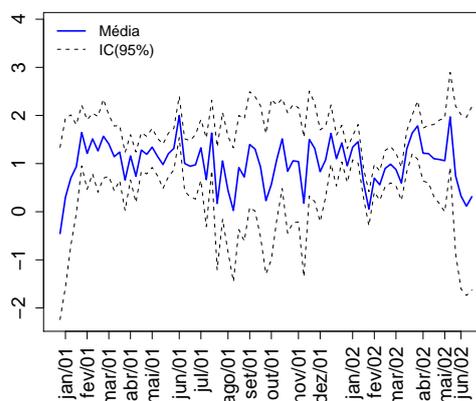


Figura 5.14: Média a posteriori e Intervalos de Credibilidade para o nível da série (α_{0t}).

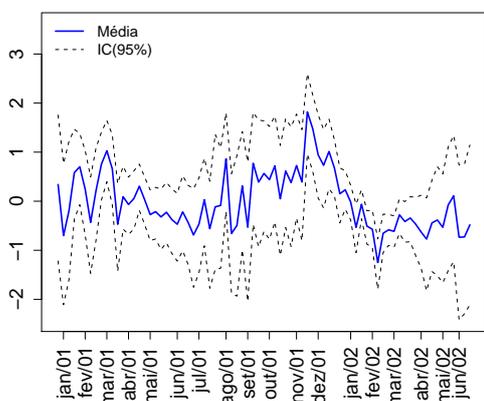
A Figura 5.15 apresenta a série de efeitos espaço-temporais para alguns bairros. Observamos que, principalmente nos bairros situados na zona oeste da cidade, como Santa Cruz (5.15(f)) e Pedra de Guaratiba (5.15(g)), os efeitos se acentuam no período do verão de 2001/2002, atingindo valores bem altos. Já os bairros da Tijuca e de Ramos, não apresentam essa elevação. O bairro de Marechal Hermes (5.15(c)), já apresenta valores negativos para estes efeitos no mesmo período considerado.



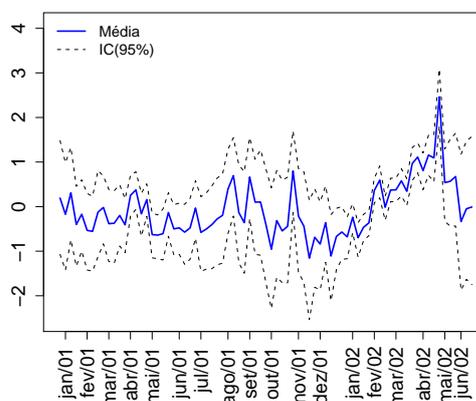
(a) Tijuca



(b) Ramos

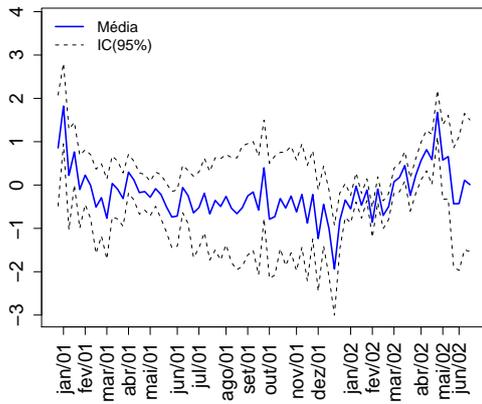


(c) Marechal Hermes

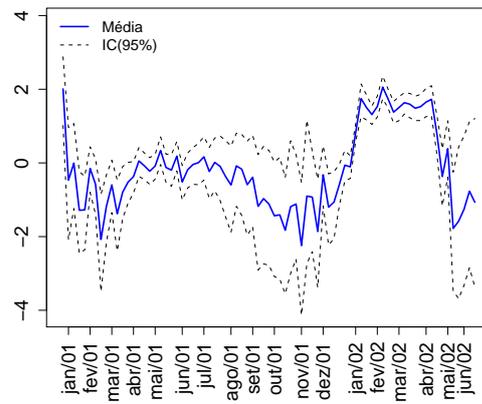


(d) Botafogo

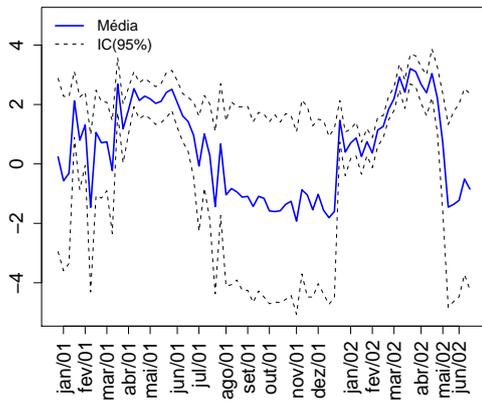
Figura 5.15: Média a posteriori dos efeitos espaço-temporais para o risco de dengue, em alguns bairros, juntamente com os intervalos de 95% de credibilidade.



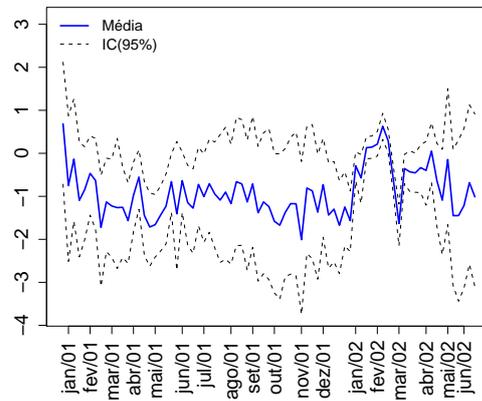
(e) Barra da Tijuca



(f) Santa Cruz



(g) Pedra de Guaratiba

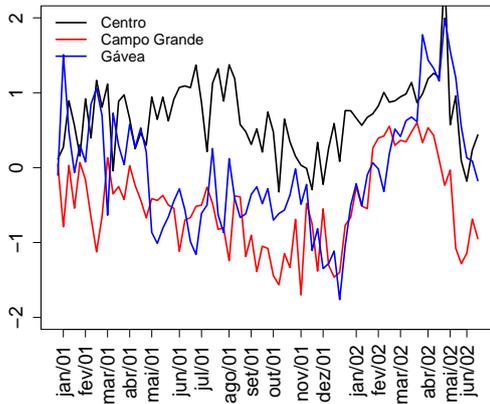


(h) Cosmos

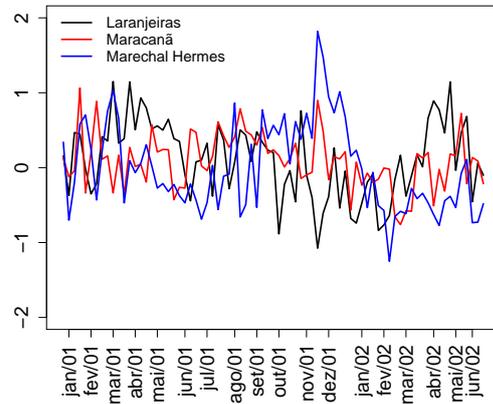
Figura 5.15: Média a posteriori dos efeitos espaço-temporais para o risco de dengue, em alguns bairros, juntamente com os intervalos de 95% de credibilidade - Continuação.

Na Figura 5.16 podemos notar as diferenças entre os bairros, através da média a posteriori dos efeitos (S) espaço-temporais para os bairros. Notamos no grupo 1 (5.16(a)), que o Centro possui efeitos bem mais elevados que os demais, durante quase todo o período e o bairro de Campo Grande e a Gávea, com um grande aumento no período da epidemia. O mesmo movimento é observado em Santa Cruz, que no grupo 3 (5.16(c)) apresentou uma variabilidade maior que os demais bairros.

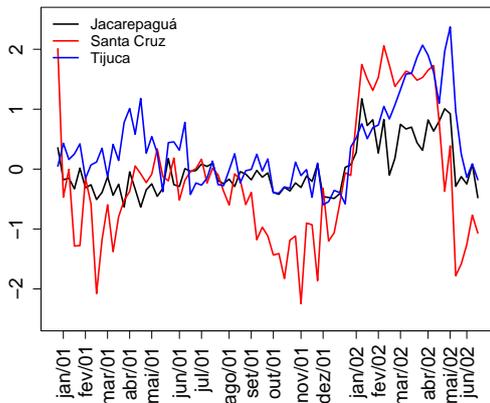
Ainda na Figura 5.16, observamos que no grupo 4 (5.16(d)), o bairro de Parada de Lucas apresenta uma evolução bem diferenciada dos demais, com valores negativos de efeitos no final da série.



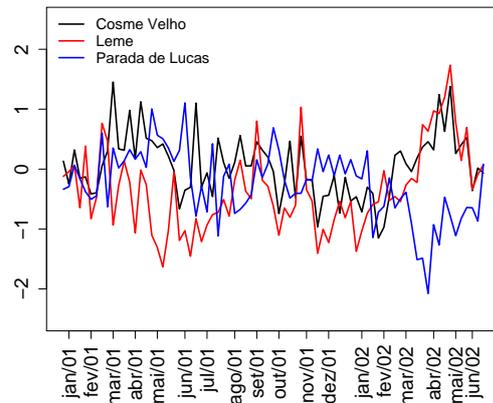
(a) Grupo 1



(b) Grupo 2



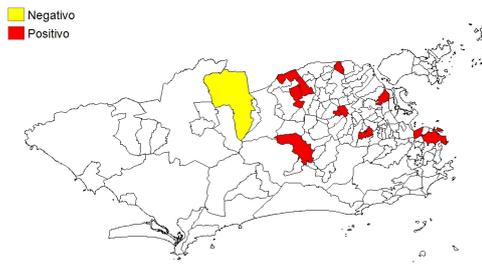
(c) Grupo 3



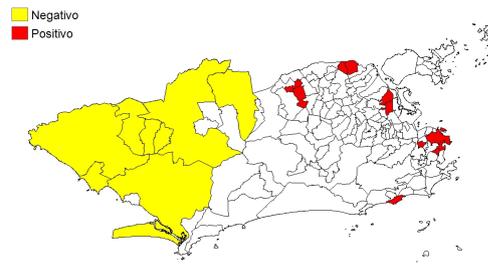
(d) Grupo 4

Figura 5.16: Comparação entre efeitos espaço-temporais para alguns bairros.

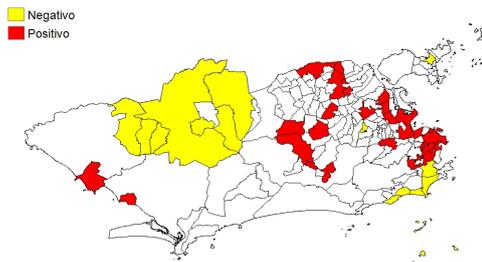
Visto isso, na Figura 5.17, podemos visualizar nos mapas da cidade, os bairros que obtiveram efeitos positivos (em vermelho) e negativos (em amarelo), durante o período estudado. Em branco aparecem os que não tiveram efeitos significativos. Há uma tendência de que os bairros da zona norte e sul da cidade, tenham efeitos positivos no início do período e que esses efeitos positivos vão se propagando-se para zona oeste da cidade no final do período.



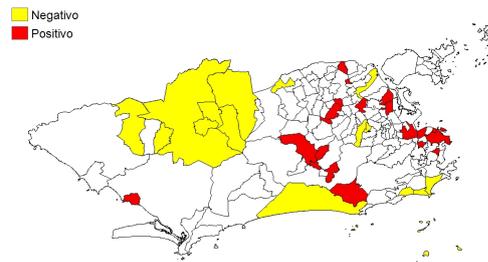
(a) Semana 6



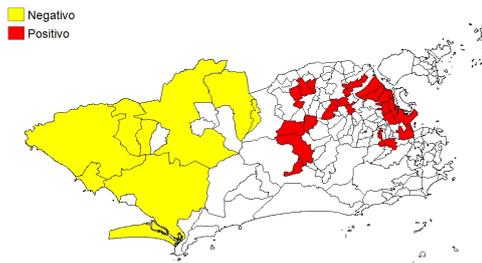
(b) Semana 8



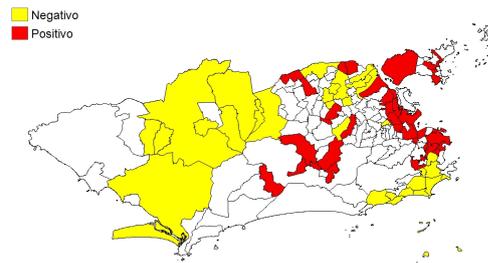
(c) Semana 20



(d) Semana 22



(e) Semana 50



(f) Semana 54

Figura 5.17: Mapas com a influência dos efeitos espaço-temporais.

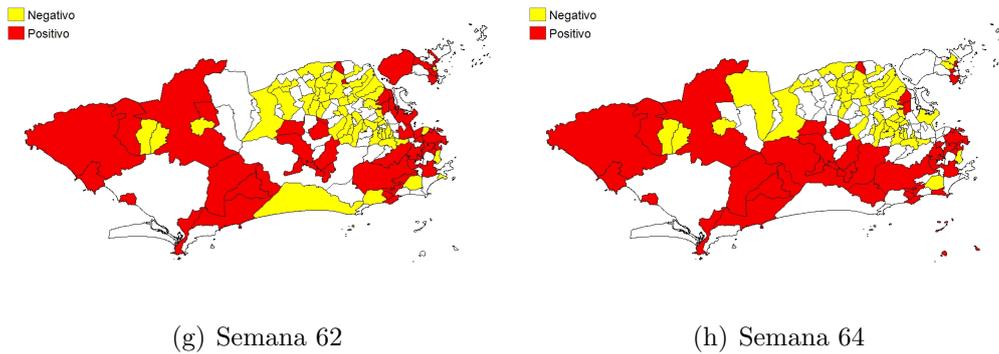
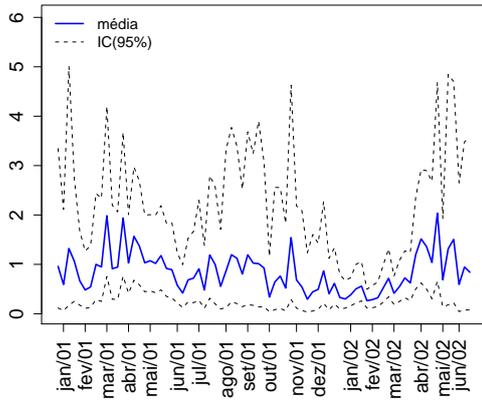


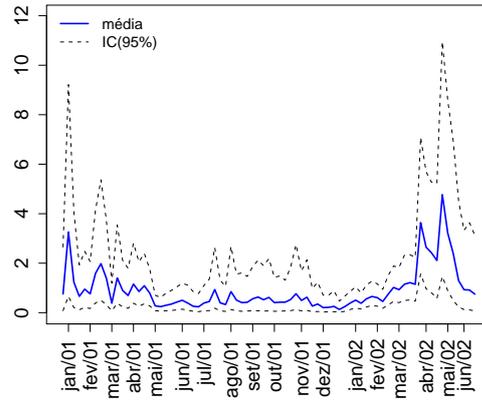
Figura 5.17: Mapas com a influência dos efeitos espaço-temporais - Continuação.

Após obtermos os valores dos efeitos espaço-temporais (S_{it}), podemos também calcular os valores de risco de dengue (λ_{it}). A Figura 5.18 ilustra estes valores para alguns bairros da cidade. Podemos observar que o bairro de Laranjeiras (5.18(a)), possuía risco de dengue maior, no início da série e no final da série. Já o bairro do Maracanã (5.18(d)) teve um pico nas semanas do mês de novembro de 2001.

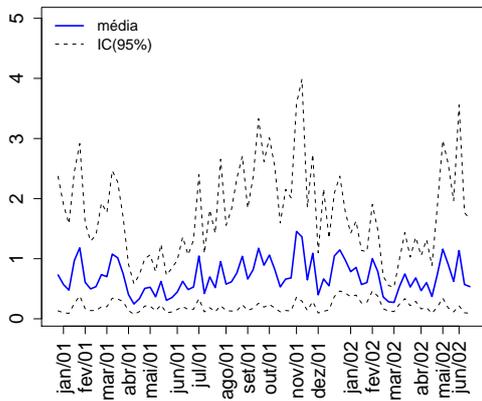
Novamente, os bairros da zona oeste, como Jacarepaguá (5.18(f)), Campo Grande (5.18(g)) e Grumari (5.18(h)) apresentaram riscos de dengue bem mais elevados nas semanas referentes a epidemia nos meses de janeiro, fevereiro e março de 2002. O bairro da Gávea apresentou altos valores de risco nas últimas semanas do estudo.



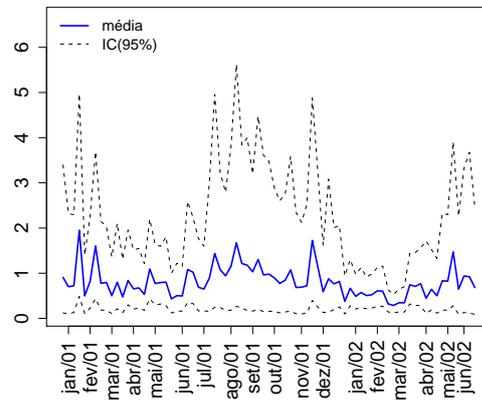
(a) Laranjeiras



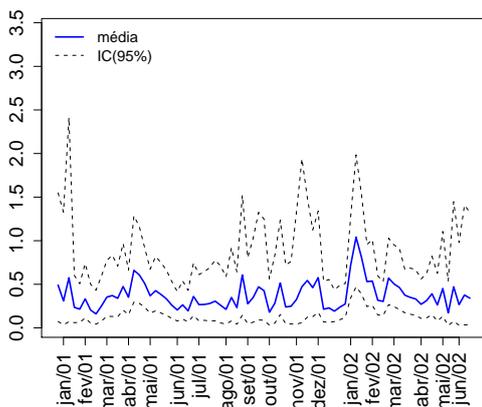
(b) Gávea



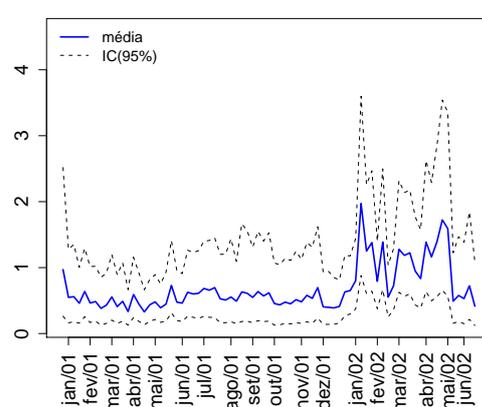
(c) Piedade



(d) Maracanã

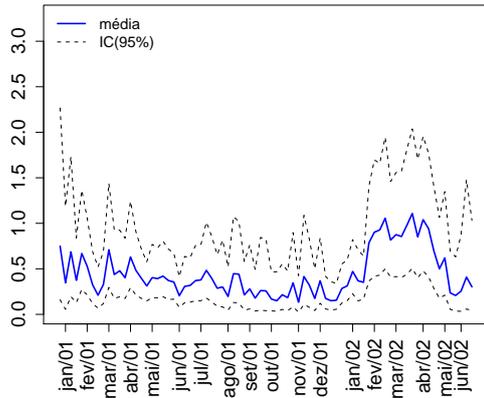


(e) Bangu

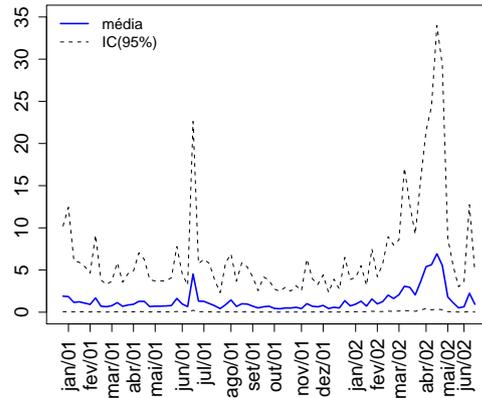


(f) Jacarepaguá

Figura 5.18: Valores dos riscos relativos de dengue estimados para alguns bairros.



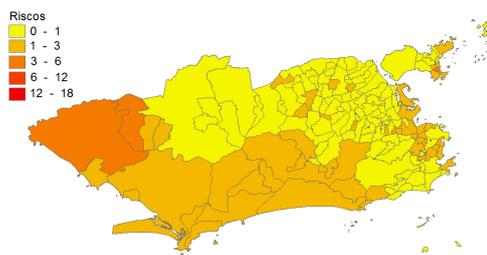
(g) Campo Grande



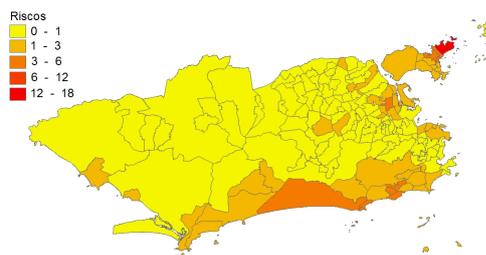
(h) Grumari

Figura 5.18: Valores dos riscos relativos de dengue estimados para alguns bairros - Continuação.

Por isso, também ilustramos em mapas, na Figura 5.19 a evolução dos riscos de dengue na cidade do Rio de Janeiro. Notamos que nas primeiras semanas os riscos são maiores nas regiões norte e sul e no centro da cidade. Com o passar do tempo, notamos que com o provável movimento de pessoas da região oeste para estas regiões, a epidemia se alastra para a região oeste, chegando a riscos bastante elevados em torno da semana 64.



(a) Semana 1



(b) Semana 2

Figura 5.19: Mapas com os riscos relativos de dengue estimados para alguns bairros da cidade do Rio de Janeiro.

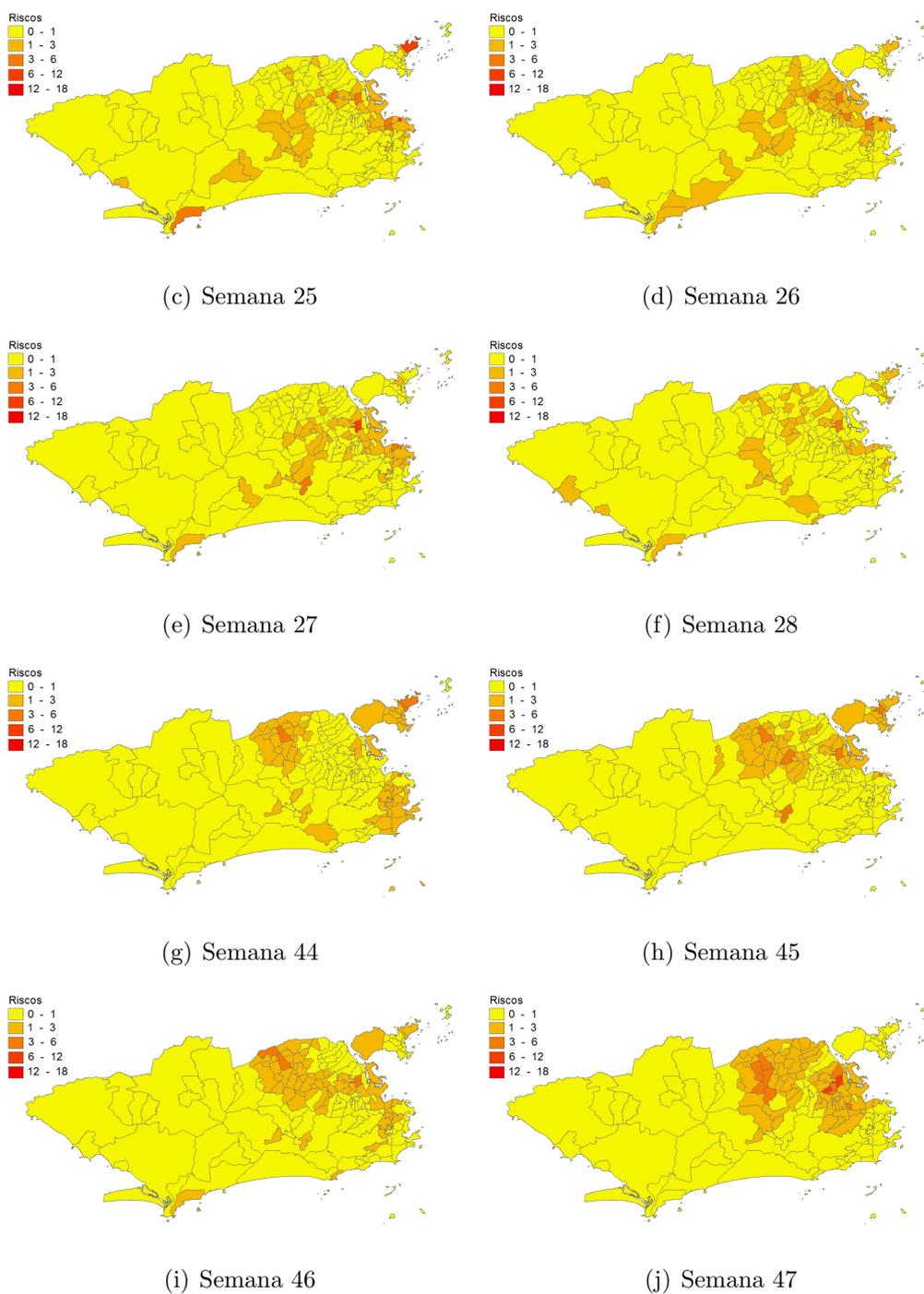
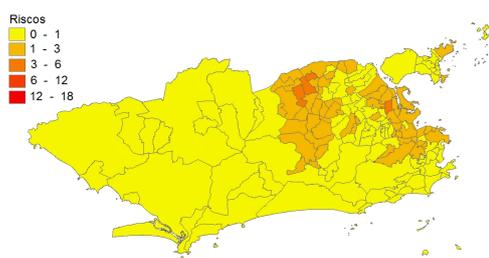
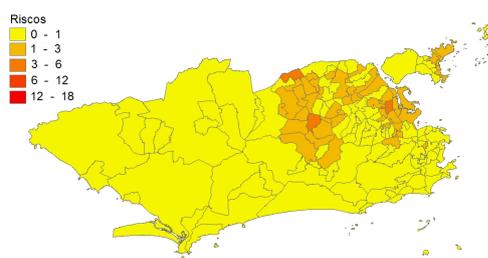


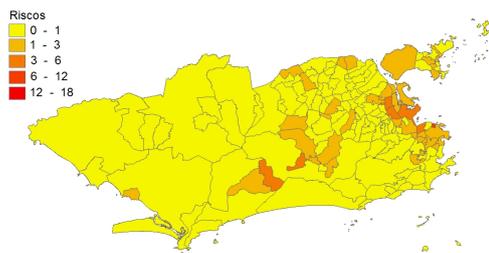
Figura 5.19: Mapas com os riscos relativos de dengue estimados para alguns bairros da cidade do Rio de Janeiro - Continuação.



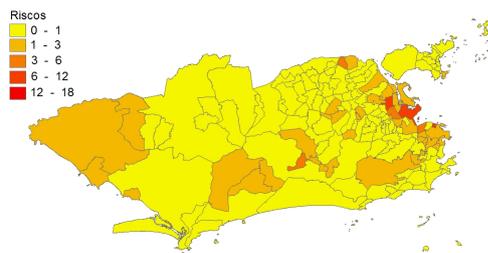
(k) Semana 48



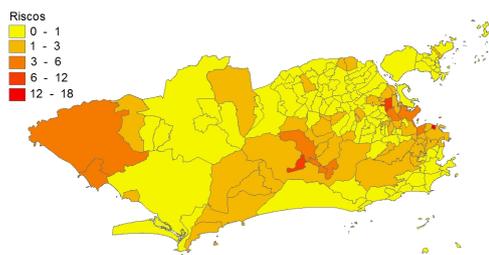
(l) Semana 49



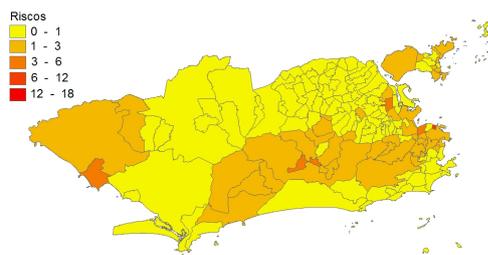
(m) Semana 54



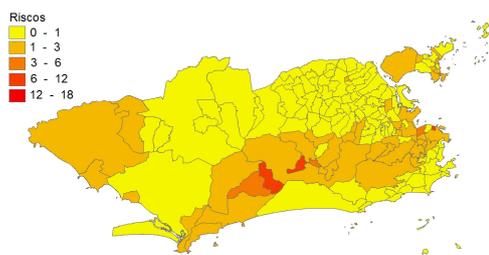
(n) Semana 55



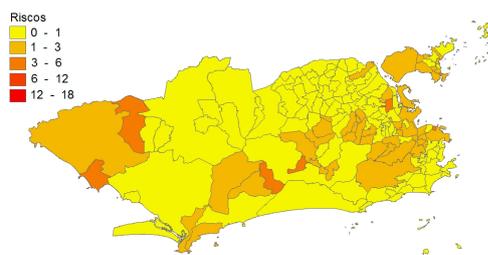
(o) Semana 56



(p) Semana 57

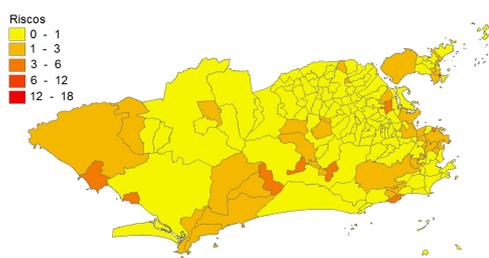


(q) Semana 58

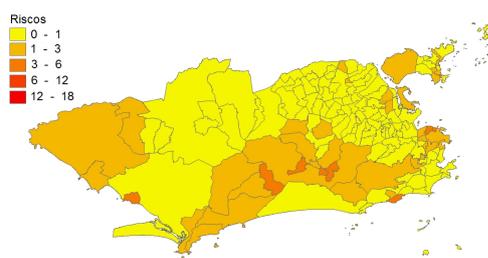


(r) Semana 59

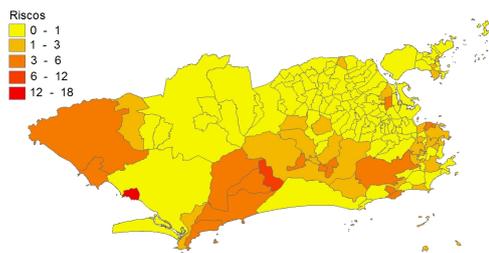
Figura 5.19: Mapas com os riscos relativos de dengue estimados para alguns bairros da cidade do Rio de Janeiro - Continuação.



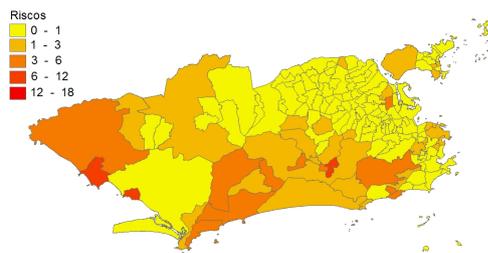
(s) Semana 62



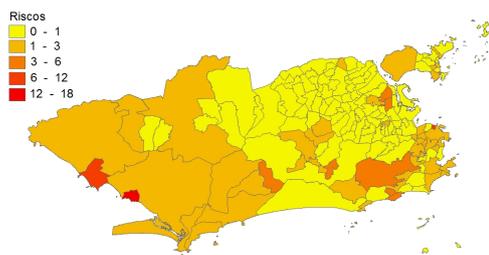
(t) Semana 63



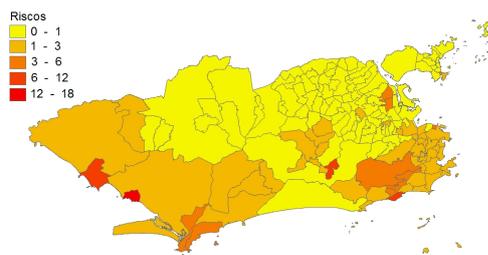
(u) Semana 64



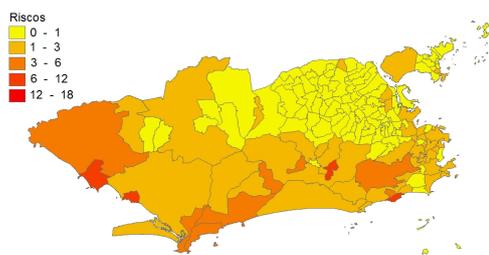
(v) Semana 65



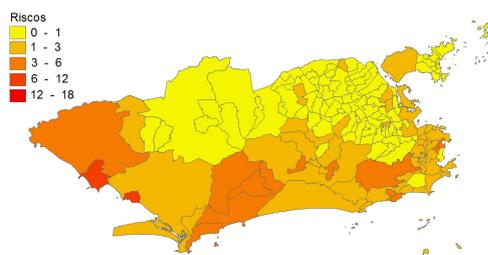
(w) Semana 66



(x) Semana 67



(y) Semana 68



(z) Semana 69

Figura 5.19: Mapas com os riscos relativos de dengue estimados para alguns bairros da cidade do Rio de Janeiro - Continuação.

Uma forma de analisar o excesso de zeros é utilizando a distribuição condicional da variável latente, como visto na Seção 3.3. Portanto a quantidade

$$\frac{1 - \theta_t(s)}{(1 - \theta_t(s)) + \theta_t(s)p(Y_t(s) = 0 | \lambda_t(s))}$$

indica a probabilidade de que um zero ocorrido seja da distribuição degenerada e não da Poisson. Ou seja, que este zero, esteja em excesso no modelo. A Figura 5.20, ilustra estas probabilidades para algumas semanas do estudo.

Notamos que as primeiras semanas possuem altas probabilidades, pois também são semanas com poucas notificações. Já no período de epidemia, notamos que apenas alguns bairros possuem altos índices de excesso de zeros, o que pode ser indícios de sub-notificações.

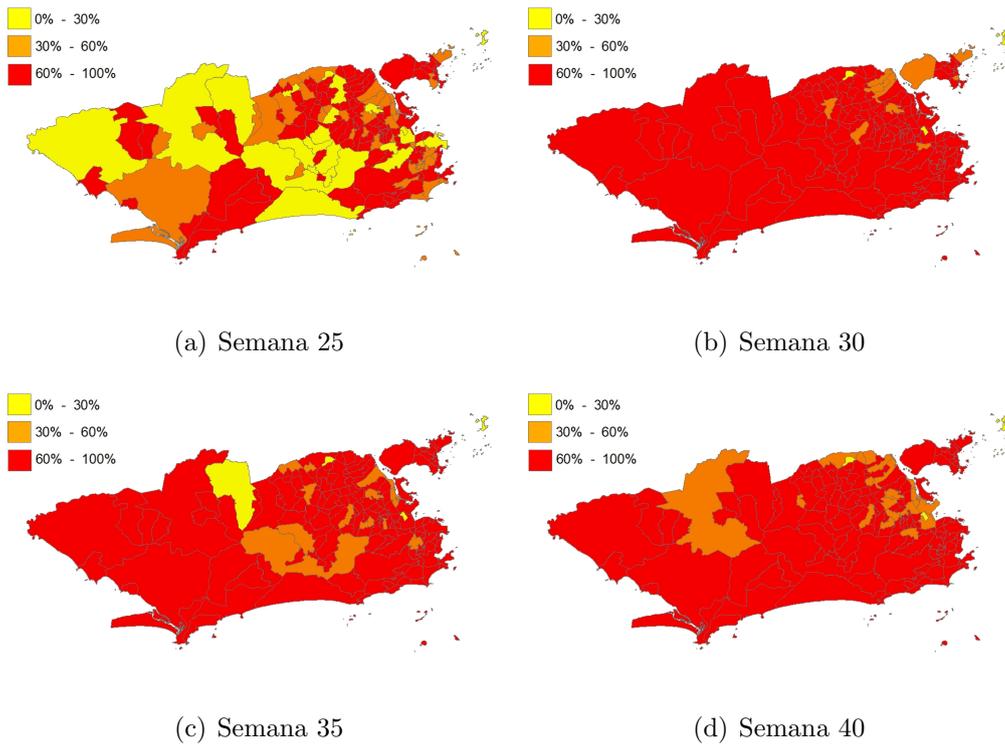
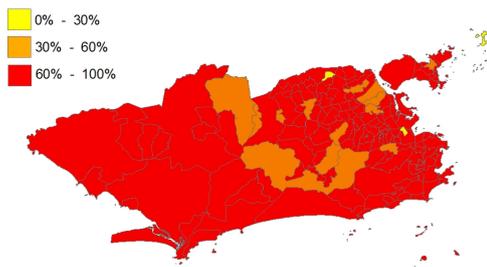
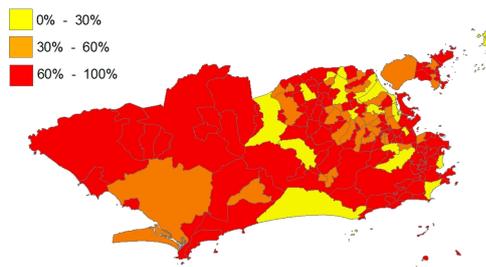


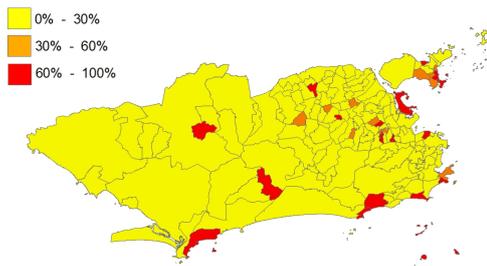
Figura 5.20: Mapas com as probabilidades de excesso de zeros para os bairros da cidade do Rio de Janeiro.



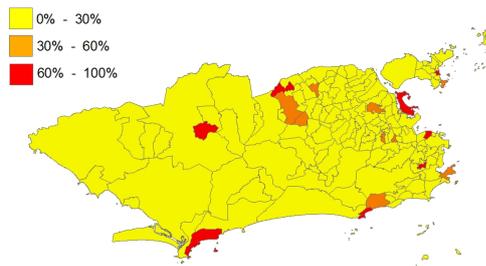
(e) Semana 45



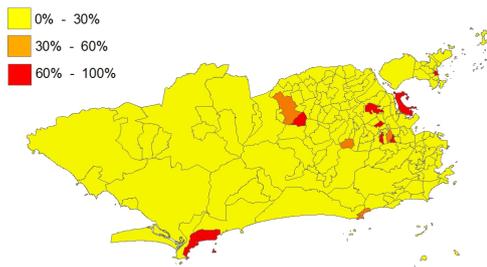
(f) Semana 50



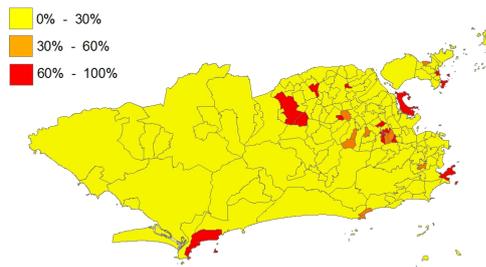
(g) Semana 55



(h) Semana 58



(i) Semana 62



(j) Semana 64

Figura 5.20: Mapas com as probabilidades de excesso de zeros para os bairros da cidade do Rio de Janeiro - Continuação.

5.3 Investigando o Efeito da Chuva nos Casos de Dengue na Cidade do Rio de Janeiro

Como visto anteriormente, a dinâmica da transmissão da dengue é um tema recorrente em estudos. Donalisio and Glasser (2002) citam que "A disseminação do dengue acompanha o homem e seus empreendimentos, migrações, deslocamentos e aglomerações no decorrer da história". Uma questão que logo aparece nesses estudos, é a da grande ocorrência de casos de dengue nos períodos mais quentes, em geral no verão, onde ocorrem grandes índices pluviométricos.

A transmissão da dengue é feita principalmente pelo mosquito *Aedes aegypti*, cuja larva se reproduz em ambientes encharcados e com temperaturas elevadas. Donalisio and Glasser (2002) também citam que questões climáticas como a temperatura e a quantidade de chuva influenciam na proliferação da doença. Desta forma a precipitação, seguida do aumento da temperatura, constitui um estado ambiental-climático ótimo para a ocorrência de focos da dengue.

Nesta aplicação, modelaremos as notificações de casos de dengue em cada bairro, com uma informação a mais, a quantidade de chuva observada naquele bairro. Como existe um período, para a reprodução do mosquito, entre a ocorrência da chuva e sua completa formação, espera-se que quantidades de chuva passadas influenciem na quantidade de casos de dengue no presente período. Entretanto esse tempo de defasagem, em geral é desconhecido. Uma alternativa estatística para este problema seria o uso de funções de transferência que utilizam valores passados de covariáveis de maneira iterativa (West et al., 1985).

5.3.1 Conjunto de dados

O Conjunto de dados para esta aplicação é uma junção dos dados utilizados nas duas aplicações anteriores. Portanto, serão usados dados semanais de notificações de casos de dengue e de precipitação pluviométrica, iniciando-se em janeiro

de 2001 e terminando em junho de 2002, em um total de 75 semanas.

A Figura 5.21 ilustra a quantidade de casos de dengue e os valores de precipitação para o bairro de Copacabana no período avaliado.

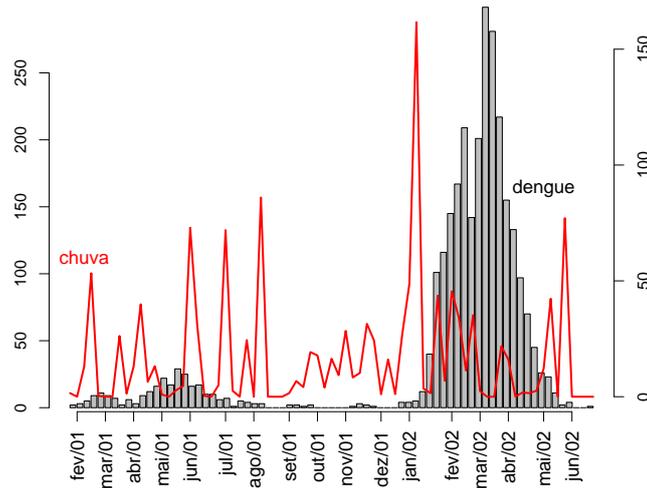


Figura 5.21: Série de casos de dengue e precipitação pluviométrica para o bairro de Copacabana entre janeiro de 2001 e junho de 2002.

Podemos notar que em janeiro de 2002, tivemos uma semana com um alto índice pluviométrico e algumas semanas depois, ocorreram as semanas da epidemia de dengue.

5.3.2 Modelo utilizado

Seguindo a notação utilizada na Seção anterior, consideremos Y_{it} o número de casos de dengue no bairro i , na semana t . A modelagem de Y_{it} , novamente levará em consideração o excesso de zeros, nos dados, e teremos um modelo tal e qual o utilizado em (5.3). Portanto:

$$p(Y_{it} | \theta_{it}, \lambda_{it}) = \begin{cases} 1 - \theta_{it} + \theta_{it}p(Y_{it} = 0 | \lambda_{it}) & \text{se } Y_{it} = 0, \\ \theta_{it}p(Y_{it} | \lambda_{it}) & \text{se } Y_{it} > 0. \end{cases} \quad (5.13)$$

Para a modelagem de θ_{it} , também adotaremos o mesmo procedimento e, consequentemente, teremos:

$$\log \frac{\theta_{it}}{1 - \theta_{it}} = \gamma_0 + \gamma_1 I(Y_{it-1} > 0), \quad (5.14)$$

onde $I(Y_{it-1} > 0)$ é uma variável indicadora de existência de casos de dengue na semana anterior.

Entretanto, para λ_{it} , que seria o risco de dengue no bairro i , na semana t , levaremos em consideração a quantidade de chuva observada. Como dito anteriormente, utilizaremos funções de transferência (veja, Box, Jenkins, and Reinsel (1994)), que considera o efeito de níveis passados de covariáveis. A estrutura usada é de uma função de transferência de primeira ordem. A modelagem de λ_{it} será da seguinte maneira:

$$Y_{it} \mid \lambda_{it}, e_{it} \sim Po(e_{it} \lambda_{it}) \quad i = 1, \dots, n \text{ e } t = 1, \dots, T \quad (5.15)$$

$$\log \lambda_{it} = \alpha_0 + \epsilon_{it} \quad (5.16)$$

$$\epsilon_{it} = \rho \epsilon_{it-1} + \beta \text{chuva}_{it}. \quad (5.17)$$

Note que toda a variação espaço-temporal está inserida na função de transferência. Neste momento não estaremos utilizando efeitos espaço-temporais. A variável chuva_{it} indica a quantidade de chuva no bairro i , na semana t e β é seu efeito associado.

Esta quantidade de chuva, em cada bairro, será obtida com a utilização da distribuição preditiva de um dos modelos empregados na primeira aplicação deste trabalho, onde modelamos a quantidade de chuva na cidade do Rio de Janeiro. Desta forma, teremos:

$$p(\text{chuva}_t(s) \mid \theta_t^c(s), \lambda_t^c(s)) = \begin{cases} 1 - \theta_t^c(s) & \text{se } \text{chuva}_t(s) = 0, \\ \theta_t^c(s) p(\text{chuva}_t(s) \mid \lambda_t^c(s)) & \text{se } \text{chuva}_t(s) > 0, \end{cases}$$

onde $\theta_t^c(s)$ e $\lambda_t^c(s)$, são os parâmetros da modelagem, de probabilidade de chuva e de níveis de chuva maiores que zero, respectivamente. Estes parâmetros também seguem as mesmas estruturas apresentadas na primeira aplicação.

Para a obtenção dos valores de chuva para cada bairro, é necessário fazer a troca de suporte (Gelfand, Zhu, & Carlin, 2001), ou seja, devemos obter estimativas para áreas, por exemplo, através de estimativas em pontos de uma grade previsão. Tal procedimento foi utilizado por Nobre, Schmidt, and Lopes (2005) na modelagem de incidência de malária no Pará. A troca de suporte pode ser feita calculando-se a seguinte integral:

$$\text{chuva}_{it} = |B_i|^{-1} \int_{B_i} \text{chuva}_t(s) d(s) \quad \forall s \in i, \quad (5.18)$$

onde B_i representa a área de cada bairro. A integral descrita em (5.18), pode ser aproximada pela expressão a seguir:

$$\text{chuva}_{it} = \frac{1}{m_i} \sum_{j=1}^{m_i} \hat{Y}_t(s_j), \quad (5.19)$$

onde m_i denota o número de pontos da grade de previsão que estão dentro dos limites territoriais do bairro i e $\hat{Y}_t(s)$ são os valores de quantidade de chuva preditos, pela grade, para o instante t .

5.3.3 Procedimento de Inferência

Com os valores observados $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_{156})'$, onde $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{i75})'$, a verossimilhança desse modelo, será dada por:

$$L(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) = \prod_{i=1}^{156} \prod_{t=1}^{75} \theta_{it}^{1-X_{it}} (1 - \theta_{it})^{X_{it}} \prod_{i=1}^{156} \prod_{t=1}^{75} p(Y_{it} | \lambda_{it})^{1-X_{it}}. \quad (5.20)$$

E por conseguinte, teremos a seguinte log-verossimilhança:

$$\begin{aligned} l(\boldsymbol{\theta}, \boldsymbol{\lambda}; \mathbf{Y}, \mathbf{X}) &= \sum_{i=1}^{156} \sum_{t=1}^{75} [(1 - X_{it}) \log \theta_{it} + X_{it} \log(1 - \theta_{it}) \\ &+ (1 - X_{it}) \log \left(\frac{(\lambda_{it} e_{it})^{Y_{it}} \exp\{-\lambda_{it} e_{it}\}}{Y_{it}!} \right)]. \end{aligned}$$

Depois de seguidos processos de substituições, podemos reescrever a função de transferência, vista em (5.17) da seguinte maneira:

$$\epsilon_{it} = \rho^t \epsilon_{i0} + \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij}. \quad (5.21)$$

Desto forma reduzimos a quantidade de parâmetros a serem estimados. A reconstrução dos valores de ϵ_{it} , será feita utilizando a equação de recorrência (5.21). Considerando Θ como o vetor paramétrico, teremos:

$$\Theta = \{\boldsymbol{\theta}, \boldsymbol{\lambda}, \gamma_0, \gamma_1, \alpha_0, \boldsymbol{\epsilon}_0, \rho, \beta, \mathbf{X}\}.$$

Devemos atribuir prioris a estes parâmetros. Para γ_0, γ_1 e α_0 , teremos novamente prioris normais, como as definidas na Seção 5.2.3. Para $\boldsymbol{\epsilon}_0$ e β também teremos prioris normais, com média nula e variância 100 e para ρ teremos uma Uniforme, tal que, $\rho \sim U(0, 1)$.

Desto forma a posteriori para este modelo será dada por:

$$\begin{aligned} p(\Theta | \mathbf{Y}) &\propto p(\mathbf{Y} | \boldsymbol{\theta}, \boldsymbol{\lambda})p(\boldsymbol{\theta} | \gamma_0, \gamma_1)p(\mathbf{X} | \boldsymbol{\theta})p(\boldsymbol{\lambda} | \alpha_0, \boldsymbol{\epsilon}_0, \beta, \rho) \\ &\quad p(\gamma_0)p(\gamma_1)p(\alpha_0)p(\boldsymbol{\epsilon}_0)p(\beta)p(\rho). \end{aligned}$$

A seguir encontram-se as expressões para as distribuições condicionais completas dos parâmetros α_0 , ρ , β e ϵ_{i0} , $i = 1, \dots, 156$. Para os demais parâmetros as condicionais são as mesmas obtidas na Seção 5.2.3. Teremos, portanto:

(i) Para α_0 :

$$\begin{aligned} p(\alpha_0 | \mathbf{Y}, \Theta_{-\alpha_0}) &\propto \exp\left\{-\frac{1}{2\sigma_{\alpha_0}^2}(\alpha_0 - \mu_{\alpha_0})^2\right\} \exp\left\{\sum_{t=1}^{75} \sum_{i=1}^{156} (1 - X_{it}) [Y_{it}\alpha_0 \right. \\ &\quad \left. - \exp\{\alpha_0 + \rho^t \epsilon_{i0} + \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij}\} e_{it}]\right\}; \end{aligned}$$

(ii) Para β :

$$\begin{aligned} p(\beta | \mathbf{Y}, \Theta_{-\beta}) &\propto \exp\left\{-\frac{1}{2\sigma_{\beta}^2}(\beta - \mu_{\beta})^2\right\} \exp\left\{\sum_{t=1}^{75} \sum_{i=1}^{156} (1 - X_{it}) \left[Y_{it}\beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij} \right. \right. \\ &\quad \left. \left. - \exp\{\alpha_0 + \rho^t \epsilon_{i0} + \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij}\} e_{it} \right] \right\}; \end{aligned}$$

(iii) Para ϵ_{i0} $i = 1, \dots, n$:

$$p(\epsilon_{i0} | \mathbf{Y}, \Theta_{-\epsilon_{i0}}) \propto \exp \left\{ -\frac{1}{2\sigma_{\epsilon_{i0}}^2} (\epsilon_{i0} - \mu_{\epsilon_{i0}})^2 \right\} \exp \left\{ \sum_{t=1}^{75} (1 - X_{it}) \left[Y_{it} \rho^t \epsilon_{i0} - \exp\{\alpha_0 + \rho^t \epsilon_{i0} + \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij}\} e_{it} \right] \right\};$$

(iv) Para ρ :

$$p(\rho | \mathbf{Y}, \Theta_{-\rho}) \propto \exp \left\{ \sum_{t=1}^{75} \sum_{i=1}^{156} (1 - X_{it}) \left[Y_{it} (\rho^t \epsilon_{i0} + \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij}) - \exp\{\alpha_0 + \rho^t \epsilon_{i0} + \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{ij}\} e_{it} \right] \right\}.$$

Novamente usaremos simulação estocástica para obter amostras da distribuição a posteriori dos parâmetros. Para γ_0 e γ_1 , que possuem condicionais completas log-côncavas, utilizaremos o ARMS. Para \mathbf{X} , teremos, novamente, $X_{it} \sim \text{Bern}(1 - \theta_{it})$ $i = 1, \dots, 156$ e $t = 1, \dots, 75$.

Para α_0 , ϵ_0 e β , utilizaremos o Algoritmo de Metropolis-Hastings, com propostas baseadas nas condicionais completas ajustadas, seguindo Gamerman (1998). Tal procedimento também pode ser encontrado em Alves (2006), também em aplicações de funções de transferência. As expressões para estas propostas estão detalhadas no apêndice C. Para ρ usaremos *slice sampling* (Neal, 1997).

5.3.4 Resultados

Foram rodadas 40.000 iterações descartando-se as 20.000 primeiras e salvando apenas as amostras de 10 em 10. A Tabela 5.10 apresenta os valores das estatísticas a posteriori dos parâmetros fixos do modelo; a Figura 5.22 ilustra os respectivos histogramas.

Podemos notar que todos os parâmetros ficaram significativos dentro de um intervalo de credibilidade de 95%. E em especial, embora muito próximo do valor

Tabela 5.10: Estimativas a posteriori do Modelo Poisson com Função de Transferência

| Parâmetro | Média | Quantil 2,5% | Quantil 97,5% |
|------------|---------|--------------|---------------|
| γ_0 | 1,1090 | 0,3697 | 1,8432 |
| γ_1 | 4,9000 | 4,3291 | 5,5241 |
| α_0 | -0,3351 | -0,3797 | -0,2792 |
| ρ | 0,999 | 0,995 | 0,999 |
| β | 0,00031 | 0,000048 | 0,00058 |

zero, o parâmetro β mostrou-se positivo, o que confirma nossa hipótese de influência da chuva sobre o risco de dengue.

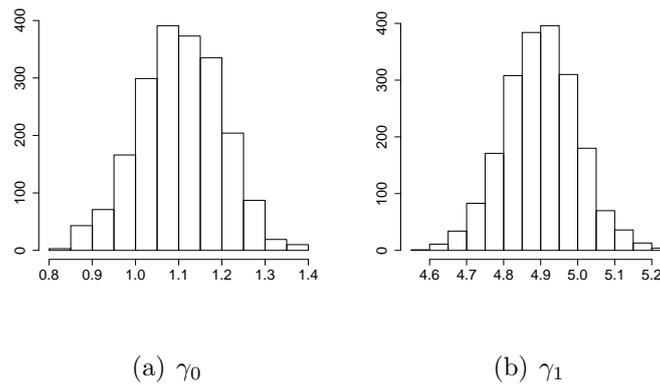
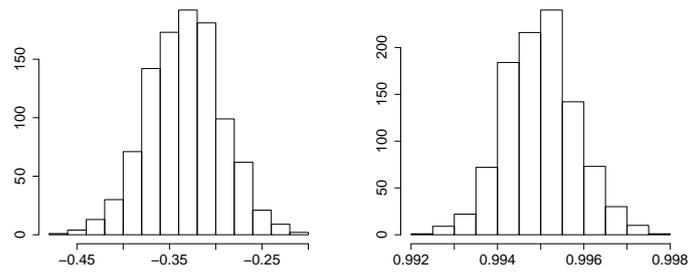
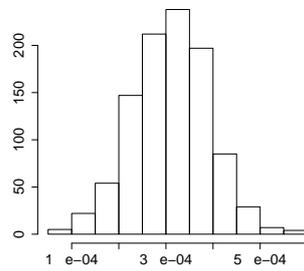


Figura 5.22: Histogramas das distribuições a posteriori dos parâmetros fixos do Modelo Poisson com Função de Transferência.



(c) α_0

(d) ρ



(e) ρ

Figura 5.22: Histogramas das distribuições a posteriori dos parâmetros fixos do Modelo Poisson com Função de Transferência - Continuação.

Comparação de Modelos

Visto que estamos modelando o risco relativo de dengue nas duas últimas aplicações, é natural pensarmos em fazer uma comparação entre os dois modelos. A Tabela 5.11 mostra os valores das estatísticas de comparação para o Modelo 2, da aplicação anterior, onde o risco de dengue era modelado por um nível variando no tempo, mais um efeito espaço-temporal, e para o Modelo 3, que é o modelo utilizado nesta aplicação, com a utilização dos valores observados de chuva.

Podemos notar claramente que, o Modelo 2 se ajusta melhor aos dados do que o Modelo 3, pois possui valores menores tanto de DIC, quanto de EPD e EQM. Esta diferença se torna mais evidente se levarmos em conta que o Modelo 2 possui duas semanas a mais de observações que o Modelo 3. Este fato destaca a necessidade de

Tabela 5.11: Comparação entre os modelos Poisson

| Modelo | DIC | \bar{D} | pD | G | P | EPD | EQM |
|--------|----------|-----------|---------|---------|----------|-----------|--------|
| 2 | 51369,73 | 43552,53 | 7817,20 | 262168 | 274177,1 | 536345,50 | 21,82 |
| 3 | 73667,55 | 73740,63 | 73,08 | 3384299 | 203589,9 | 3587889 | 289,25 |

explorar melhor a inclusão da chuva observada na modelagem do risco de dengue.

5.4 Considerações Finais

Neste capítulo foram feitas três aplicações em que utilizamos a teoria desenvolvida durante o trabalho.

Para a primeira aplicação, onde empregamos diversos modelos para a chuva na cidade do Rio de Janeiro, notamos que os dados ajustaram-se um pouco melhor ao modelo com o nível, de quantidade de chuva não-nula, variando no tempo, somando-se a ele os efeitos espaço-temporais. Identificamos, também, que para esta modelagem (quantidade de chuva não nula) a distribuição Gama, foi a que se ajustou melhor.

Para a modelagem da probabilidade de chuva, verificou-se que a melhor estrutura foi a de um nível fixo mais um fator, que mostrou-se positivo, especificando a existência de chuva na semana anterior, além, é claro, dos efeitos espaço-temporais. Observamos, ainda, que, com este modelo, de maneira geral, conseguimos fazer previsões para quaisquer pontos da cidade.

Na modelagem de notificações de casos de dengue, novamente o modelo que ajustou-se mais adequadamente aos dados, foi o com nível variando no tempo, na modelagem do risco relativo. Destacamos a evolução deste risco de uma região da cidade (centro-sul) para outra (oeste), no decorrer das semanas do estudo. Para a modelagem do excesso de zeros, neste caso, foi utilizado somente um nível fixo e

um fator de existência de casos de dengue na semana anterior.

Quando investigamos a influência da chuva neste mesmo risco relativo, em nossa terceira aplicação, percebemos que esta afeta significativamente o aumento do risco. Entretanto este modelo não se comportou de maneira mais adequada do que o anterior, sem a inclusão da chuva. Este fato pode ser explicado pela ausência da componente aleatória, presente no modelo anterior na forma dos efeitos espaço-temporais. Um modelo com a componente espaço-temporal na modelagem do risco, foi ajustado, mas de um modo geral, alguns parâmetros do modelo não convergiram.

Capítulo 6

Conclusões

O presente trabalho teve como objetivo explorar uma nova classe de modelos inflacionados de zeros com dependência espaço-temporal. Basicamente, trabalhou-se com mistura de distribuições, sempre uma Bernoulli, modelando a existência ou excesso do valor zero, com uma outra distribuição de interesse, que foi considerada discreta ou contínua.

A dependência espaço-temporal aparece no modelo sob a forma de efeitos aleatórios espacial e temporalmente estruturados, que a priori podem ser tanto modelados por Processos Gaussianos, quanto por efeitos Condicionais Autoregressivos (CAR), dependendo do tipo de observação que temos.

No segundo nível de hierarquia do modelo, atribuímos estruturas para $\theta_t(s)$ e para $\lambda_t(s)$, que são os parâmetros da Bernoulli e da distribuição associada.

Vimos que para a modelagem de $\theta_t(s)$ podemos atribuir covariáveis e um efeito espaço-temporal. A quantidade de covariáveis pode ser um fator preponderante na estimação destes efeitos. Para o caso de variáveis discretas, se observado um número grande de zeros, estes parâmetros podem não ser identificáveis pelo modelo e, portanto, não deverão ser utilizados. Resultado este, também obtido, numa versão puramente espacial por Agarwal et al. (2002).

Para $\lambda_t(s)$, também podemos incorporar covariáveis e efeitos espaço-temporais. Na modelagem de variáveis discretas é necessário a inclusão de uma variável latente, afim de obtermos independência entre $\theta_t(s)$ e $\lambda_t(s)$ e por conseguinte facilidades no processo de inferência, especialmente quando trabalhamos com Modelos Dinâmicos.

Observamos que para estas estruturas dinâmicas utilizou-se uma adaptação do método proposto por Gamerman (1998). Este método, porém, é de implementação bastante trabalhosa e consome um grande tempo computacional. Outras alternativas podem ser utilizadas para o sorteio destes parâmetros, utilizando a idéia do *Linear Bayes* (West et al., 1985), que vem sendo desenvolvida no projeto de doutorado de Romy Ravines (Ravines, 2005).

Na aplicação para a chuva no Rio de Janeiro, a inclusão de mais covariáveis na modelagem do *logit*, poderia diminuir a grande variabilidade dos efeitos espaço-temporais, já que apenas a inclusão do nível variando no tempo (γ_{0t}), não solucionou este problema. A disponibilidade e a definição de quais covariáveis seriam pertinentes, neste caso, é uma questão a ser estudada.

Considerando a aplicação para os casos de dengue, outros modelos podem ser avaliados. Pela Figura 5.12 notamos um aumento acentuado, no número de casos, em um determinado período, seguido de uma queda igualmente acentuada. Uma alternativa para este caso seria a utilização de um modelo de crescimento, com intervenção no período epidêmico, na modelagem $\lambda_t(s)$. Ou seja:

$$\log(\lambda_{it}) = \alpha_{0t} + S_t(s) \quad (6.1)$$

$$\alpha_{0t} = \alpha_{0t-1} + \beta_t + w_{1t} \quad (6.2)$$

$$\beta_t = \beta_{t-1} + w_{2t} \quad (6.3)$$

Outra alternativa seria a inclusão de fatores de desconto distintos para o período de epidemia e os demais.

Quanto à modelagem dos casos de dengue, utilizando a chuva, uma alternativa seria a inserção de estruturas espacialmente referenciadas a priori tanto para β

quanto para ρ na função de transferência, ou seja:

$$\log \lambda_{it} = \alpha_0 + \epsilon_{it} \quad (6.4)$$

$$\epsilon_{it} = \rho_i \epsilon_{it-1} + \beta_i \text{chuva}_{it}, \quad (6.5)$$

onde $\beta_i \sim CAR(\sigma^2)$ e $\delta_i = \log \frac{\rho_i}{1-\rho_i} \sim CAR(\sigma^2)$ seriam as prioris para β_i e δ_i , respectivamente. Outra opção, também, seria a inclusão de efeitos espaço-temporais na função de transferência. Desta maneira, poderíamos modelar o risco de dengue, da seguinte forma:

$$\log \lambda_{it} = \alpha_0 + \epsilon_{it} \quad (6.6)$$

$$\epsilon_{it} = \rho \epsilon_{it-1} + \beta \text{chuva}_{it} + S_{it}, \quad (6.7)$$

onde S_{it} representam os efeitos espaço-temporais no bairro i e semana t .

Todas alternativas apresentadas serão exploradas em projetos futuros.

Referências Bibliográficas

- Agarwal, D., Gelfand, A., & Citron-Pousty, S. (2002). Zero-inflated models with application to spatial count data. *Environmental and Ecological Statistics*, 9, 341-355.
- Alves, M. B. (2006). *Funções de transferência em modelos dinâmicos lineares generalizados* (Tese de Doutorado). Instituto de Matemática - UFRJ, RJ Brasil.
- Banerjee, S., Carlin, B., & Gelfand, A. (2004). *Hierarchical modelling and analysis for spatial data*. Chapman & Hall.
- Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussion). *Journal of the Royal Statistical Society Series B*, 36, 192-236.
- Box, G., Jenkins, G., & Reinsel, G. (1994). *Time series analysis - forecasting and control*. Prentice-Hall, Inc., 3a. ed.
- Cohen, A. (1963). Estimation in mixtures of discrete distributions. In *Proceedings of the international symposium on discrete distributions* (p. 373-378). New York: Pergamon Press.
- Cressie, N. (1993). *Statistics for Spatial Data. Revised Edition*. John Wiley & Sons, Inc.
- Diggle, P., & Ribeiro Jr, P. (2000). *Model based geostatistics*. Associação Brasileira de Estatística - ABE - 14º SINAPE.

- Donalisio, M., & Glasser, C. (2002). Vigilância Entomológica e Controle de Vetores do Dengue. *Revista Brasileira de Epidemiologia*, 5(3), 259-272.
- Doornick, J. (2002). *Object-oriented matrix programming using ox (3rd edition)*. London: Timberlake Consultants Press and Oxford. (www.nuff.ox.ac.uk/Users/Doornick.).
- Farewell, V. (1986). Mixture Models in Survival Analysis: Are They Worth the Risk? *Canadian Journal of Statistics*, 14, 257-262.
- Ferreira, G. (2003). *Análise Espaço-Temporal da Distribuição dos Casos de Dengue na Cidade do Rio de Janeiro no Período de 1986 à 2002* (Master's Dissertation). Depto de Métodos Estatísticos, UFRJ, Brazil (In Portuguese).
- Feuerverger, A. (1979). On Some Methods of Analysis of Weather Experiments. *Biometrika*, 66(3), 655-658.
- Gamerman, D. (1997). *Markov chain monte carlo - stochastic simulation for bayesian inference*. Chapman & Hall.
- Gamerman, D. (1998). Markov chain Monte Carlo for Dynamic Generalized Linear Models. *Biometrika*, 85, 215-227.
- Gelfand, A., & Ghosh, S. (1998). Model choice: A minimum posterior predictive loss approach. *Biometrika*, 85, 1-11.
- Gelfand, A., & Smith, A. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398-409.
- Gelfand, A., Zhu, L., & Carlin, B. (2001). On the change of the support problem for spatio-temporal data. *Biostatistics*, 2, 31-45.
- Geman, S., & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-41.

- Gilks, W. (1992). Derivative-free Adaptive Rejection Sampling for Gibbs Sampling. In *Bayesian statistics* (p. 641-649). eds. J. M. Bernardo, J.O. Berger, A.P. David and A.F.M. Smith, Oxford: Clarendon.
- Gilks, W., Best, N., & Tan, K. (1995). Adaptive Rejection Metropolis Sampling with Gibbs Sampling. *Applied Statistics*, *44*(2), 455-472.
- Gilks, W., & Wild, P. (1992). Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics*, *41*(2), 337-48.
- Hastings, W. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, *57*, 97-109.
- Heilbron, D. (1989). *Generalized linear models for altered zero probabilities and over dispersion in count data* (Tech. Rep.). Department of Epidemiology and Biostatistics, University of California, San Francisco.
- J.F., A., & Biswas, A. (2003). A bayesian analysis of zero-inflated generalized poisson model. *Computational Statistics & Data Analysis*, *42*, 37-46.
- Johnson, N., & Kotz, S. (1969). *Distributions in statistics: Discrete distributions*. Boston: Houghton Mifflin.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1-14.
- Meeker, W. (1987). Limited Failure Population Life Tests: Application to Integrated Circuit Reliability. *Technometrics*, *29*, 51-65.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. (1953). Equation of State Calculations by Fast Computing Machine. *Journal of Chemical Physics*, *21*, 1087-91.
- Migon, H., & Gamerman, D. (1999). *Statistical inference - an integrated approach*. Arnold.

- Neal, R. (1997). *Markov chain monte carlo methods based on 'slice' the density function* (Tech. Rep.). Department of Statistics, University of Toronto.
- Nobre, A., Schmidt, A., & Lopes, H. (2005). Spatio-Temporal Models for Mapping the Incidence of Malaria in Pará. *Environmetrics*, 16, 291-304.
- R Development Core Team. (2005). *R: A language and environment for statistical computing*. Vienna, Austria. (ISBN 3-900051-07-0)
- Ravines, R. (2005). *Um esquema eficiente de amostragem em modelos dinâmicos generalizados com aplicações em modelos de função de transferência* (Tech. Rep.). Departamento de Métodos Estatísticos, Universidade Federal do Rio de Janeiro, RJ.
- Rideout, M., Demetrio, C., & Hinde, J. (1998). *Models for count data with many zeros* (Tech. Rep.). International Biometric Conference, Cape Town.
- Rodrigues, J. (2003). Bayesian analysis of zero-inflated distributions. *Communications in Statistics*, 32(2), 281-289.
- Sansó, B., & Guenni, L. (1999). Venezuela rainfall data analysed by using a bayesian space-time model. *Applied Statistics*, 48, 345-362.
- Schmidt, A., Nobre, A. A., & Ferreira, G. (2002). Alguns Aspectos da Modelagem de Dados Espacialmente Referenciados. *Revista Brasileira de Estatística*, 63(220), 59-88.
- Schmidt, A. M., & Ferreira, G. S. (2006). *Spatial modelling of the relative risk of dengue fever in rio de janeiro for the epidemic period between 2001 and 2002* (Tech. Rep.). Depto de Métodos Estatísticos, UFRJ, Brasil.
- Spiegelhalter, D., Best, N., Carlin, B., & Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. B*, 64, 583-639.
- Tanner, M. (1996). *Tools for statistical inference*. Springer.

- Velarde, L. G. C., Migon, H. S., & Pereira, B. B. (2004). Space-time Modeling of Rainfall Data. *Environmetrics*, 15, 561-576.
- Waller, L., Carlin, B., Xia, H., & Gelfand, E. (1997). Hierarchical Spatio-temporal Mapping of Disease Rates. *Journal of the American Statistical Association*, 92(438), 607-617.
- West, M., & Harrison, P. (1997). *Bayesian forecasting and dynamic models*. Springer-Verlag New York, Second Edition.
- West, M., Harrison, P., & Migon, H. (1985). Dynamic generalized linear models and Bayesian forecasting (with discussion). *Journal of the American Statistical Association* 80, 73-96(80), 73-96.
- Yau, K., Wang, K., & Lee, A. (2003). Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. *Biometrical Journal*, 43(4), 437-452.

Apêndice A

Localização Geográfica das Estações Monitoradoras da Geo-Rio

Tabela A.1: Estações monitoradoras de chuva no Rio de Janeiro

| Nome | Long.(W) | Lat. (S) | UTM(E) | UTM(N) | Cota (m) |
|--------------------|----------|----------|--------|---------|----------|
| Vidigal | 43 13 58 | 22 59 32 | 681,1 | 7456,22 | 85 |
| Urca | 43 10 00 | 22 57 20 | 687,88 | 7460,3 | 90 |
| Rocinha | 43 14 44 | 22 59 08 | 679,93 | 7456,95 | 160 |
| Tijuca | 43 13 18 | 22 55 54 | 682,33 | 7462,9 | 340 |
| Santa Teresa | 43 11 47 | 22 55 52 | 684,94 | 7462,91 | 170 |
| Copacabana | 43 11 19 | 22 59 9 | 685,65 | 7456,86 | 90 |
| Grajaú | 43 16 04 | 22 55 25 | 677,62 | 7463,77 | 80 |
| Ilha do Governador | 43 12 36 | 22 49 03 | 683,87 | 7475,52 | 0 |
| Penha | 43 16 28 | 22 50 37 | 677,03 | 7472,69 | 111 |
| Madureira | 43 20 20 | 22 52 20 | 670,38 | 7469,6 | 45 |
| Irajá | 43 20 12 | 22 49 17 | 670,74 | 7474,68 | 20 |
| Bangu | 43 27 55 | 22 53 10 | 657,4 | 7468,22 | 15 |

Tabela A.1: Estações monitoradoras de chuva no Rio de Janeiro - Continuação

| Nome | Long.(W) | Lat. (S) | UTM(E) | UTM(N) | Cota (m) |
|-----------------|-----------------|-----------------|---------------|---------------|-----------------|
| Piedade | 43 18 31 | 22 53 28 | 673,49 | 7467,47 | 72 |
| Tanque | 43 21 59 | 22 54 37 | 667,53 | 7565,45 | 73 |
| Saúde | 43 11 39 | 22 53 50 | 685,21 | 7466,69 | 35 |
| Jardim Botânico | 43 13 24 | 22 58 20 | 682,09 | 7458,38 | 0 |
| Itanhangá | 43 18 08 | 22 59 54 | 674,2 | 7455,42 | 80 |
| Cidade de Deus | 43 21 44 | 22 56 42 | 667,91 | 7461,59 | 15 |
| Riocentro | 43 24 42 | 22 58 30 | 662,79 | 7458,31 | 0 |
| Guaratiba | 43 34 40 | 23 03 08 | 645,7 | 7449,95 | 0 |
| Gericinó | 43 35 00 | 22 51 58 | 645,31 | 7470,54 | 0 |
| Santa Cruz | 43 41 04 | 22 54 44 | 634,8 | 7466,08 | 15 |
| Cachambi | 43 16 41 | 22 53 23 | 676,64 | 7467,59 | 25 |
| Anchieta | 43 24 11 | 22 49 36 | 663,89 | 7474,74 | 50 |
| Grota Funda | 43 31 08 | 23 00 52 | 651,76 | 7454,06 | 10 |
| Campo Grande | 43 33 42 | 22 54 12 | 647,49 | 7466,43 | 30 |
| Sepetiba | 43 42 42 | 22 58 06 | 632,6 | 7459,35 | 62 |
| Sumaré | 43 14 09 | 22 57 03 | 681,42 | 7460,87 | 768 |
| Mendanha | 43 31 24 | 22 49 25 | 651,58 | 7475,21 | 736 |
| Itaúna | 43 25 29 | 23 00 33 | 661,43 | 7454,57 | 65 |
| Laranjeiras | 43 11 14 | 22 56 15 | 685,92 | 7462,18 | 60 |
| São Cristóvão | 43 13 18 | 22 53 48 | 682,39 | 7466,81 | 25 |
| São Conrado | 43 15 29 | 22 59 52 | 678,54 | 7455,61 | 75 |

Apêndice B

Cálculo das Distribuições Condicionais Completas a Posteriori

Neste apêndice descreveremos detalhadamente as distribuições condicionais completas de mais dois modelos utilizados na aplicação para dados de chuva na cidade do Rio de Janeiro.

Cálculo das Condicionais Completas para o Modelo 2

Como definido na Seção 5.1, o Modelo 2 é descrito da seguinte forma:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} 1 - \theta_t(s) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) > 0. \end{cases}$$

$$\text{logit}\theta_t(s) = \log \frac{\theta_t(s)}{1 - \theta_t(s)} = \gamma_0 + \gamma_1 I(Y_{t-1} > 0) + \gamma_2 I(Y_{t-2} > 0) + Z_t(s)$$

$$Y_t(s) | \lambda_t(s) \sim \text{Exp}(\lambda_t(s))$$
$$\frac{1}{\lambda_t(s)} = \alpha_0 + S_t(s).$$

(i) Condicional completa para γ_0 :

$$\begin{aligned}
p(\gamma_0 \mid \mathbf{Y}, \Theta_{-\gamma_0}) &\propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_0}^2} (\gamma_0 - \mu_{\gamma_0})^2 \right\} \exp \left\{ \gamma_0 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) \right\} \\
&\quad \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\
&\quad \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + Z_t(s_i)\}) \right\}.
\end{aligned}$$

(ii) Condicional completa para γ_1 :

$$\begin{aligned}
p(\gamma_1 \mid \mathbf{Y}, \Theta_{-\gamma_1}) &\propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_1}^2} (\gamma_1 - \mu_{\gamma_1})^2 \right\} \exp \left\{ \gamma_1 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) I(Y_{t-1}(s_i) > 0) \right\} \\
&\quad \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\
&\quad \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + Z_t(s_i)\}) \right\}.
\end{aligned}$$

(iii) Condicional completa para γ_2 :

$$\begin{aligned}
p(\gamma_2 \mid \mathbf{Y}, \Theta_{-\gamma_2}) &\propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_2}^2} (\gamma_2 - \mu_{\gamma_2})^2 \right\} \exp \left\{ \gamma_2 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) I(Y_{t-2}(s_i) > 0) \right\} \\
&\quad \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\
&\quad \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + Z_t(s_i)\}) \right\}.
\end{aligned}$$

(iv) Condicional completa para α_0 :

$$\begin{aligned}
p(\alpha_0 \mid \mathbf{Y}, \Theta_{-\alpha_0}) &\propto \exp \left\{ -\frac{1}{2\sigma_{\alpha_0}^2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} \exp \left\{ -\alpha_0 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) \right\} \\
&\quad \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} (Y_t(s_i) \exp\{-(\alpha_0 + S_t(s_i))\}) \right\}.
\end{aligned}$$

(v) Condicional completa para $Z_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$:

$$\begin{aligned}
p(Z_t(s_i) \mid \mathbf{Y}, \Theta_{-Z_t(s_i)}) &\propto \exp \left\{ -\frac{1}{2VZ} (Z_t(s_i) - mZ)^2 \right\} \exp \{ Z_t(s_i) X_t(s_i) \} \\
&\quad \exp \{ -\log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\
&\quad \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + Z_t(s_i)\}) \right\}.
\end{aligned}$$

(vi) Condicional completa para $S_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$:

$$p(S_t(s_i) \mid \mathbf{Y}, \Theta_{-S_t(s_i)}) \propto \exp \left\{ -\frac{1}{2VS} (S_t(s_i) - mS)^2 \right\} \exp \{ -S_t(s_i) X_t(s_i) \} \\ \exp \{ (Y_t(s_i) \exp \{ -(\alpha_0 + S_t(s_i)) \}) \}.$$

(vii) Condicional completa para σ_1^2 :

$$p(\sigma_1^2 \mid \mathbf{Y}, \Theta_{-\sigma_1^2}) \propto \sigma_1^{2-(a_1+1)} \exp \left\{ -\frac{b_1}{\sigma_1^2} \right\} \sigma_1^{2(-\frac{n}{2})} \left\{ -\frac{1}{2\sigma_1^2} \sum_{t=1}^T \mathbf{Z}_t' R^{-1} \mathbf{Z}_t \right\}.$$

Onde:

$$R_{ij} = \exp \{ -\phi_1 \mid \mid s_i - s_j \mid \mid \} \quad i, j = 1, s, n.$$

(viii) Condicional completa para σ_2^2 :

$$p(\sigma_2^2 \mid \mathbf{Y}, \Theta_{-\sigma_2^2}) \propto \sigma_2^{2-(a_2+1)} \exp \left\{ -\frac{b_2}{\sigma_2^2} \right\} \sigma_2^{2(-\frac{n}{2})} \exp \left\{ -\frac{1}{2\sigma_2^2} \sum_{t=1}^T \mathbf{S}_t' W^{-1} \mathbf{S}_t \right\}.$$

Onde

$$W_{ij} = \exp \{ -\phi_2 \mid \mid s_i - s_j \mid \mid \} \quad i, j = 1, s, n.$$

(ix) Condicional completa para ϕ_1 :

$$p(\phi_1 \mid \mathbf{Y}, \Theta_{-\phi_1}) \propto \phi_1^{(\alpha_1-1)} \exp \{ -\beta_1 \phi_1 \} \mid \Sigma \mid^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \mathbf{Z}_t' \Sigma^{-1} \mathbf{Z}_t \right\}.$$

Onde

$$\Sigma_{ij} = \sigma_1^2 \exp \{ -\phi_1 \mid \mid s_i - s_j \mid \mid \} \quad i, j = 1, s, n.$$

(x) Condicional completa para ϕ_2 :

$$p(\phi_2 \mid \mathbf{Y}, \Theta_{-\phi_2}) \propto \phi_2^{(\alpha_2-1)} \exp \{ -\beta_2 \phi_2 \} \mid \Psi \mid^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \mathbf{S}_t' \Psi^{-1} \mathbf{S}_t \right\}.$$

Onde

$$\Psi_{ij} = \sigma_2^2 \exp \{ -\phi_2 \mid \mid s_i - s_j \mid \mid \} \quad i, j = 1, s, n.$$

Condicionais Completas para o Modelo 9

Como definido na Seção 5.1, o Modelo 9 é descrito da seguinte forma:

$$p(Y_t(s) | \theta_t(s), \lambda_t(s)) = \begin{cases} 1 - \theta_t(s) & \text{se } Y_t(s) = 0, \\ \theta_t(s)p(Y_t(s) | \lambda_t(s)) & \text{se } Y_t(s) > 0. \end{cases}$$

$$\text{logit}\theta_t(s) = \log \frac{\theta_t(s)}{1 - \theta_t(s)} = \gamma_0 + \gamma_1 I(Y_{t-1} > 0) + \gamma_2 I(Y_{t-2} > 0) \\ + \gamma_3 I(Y_{t-3} > 0) + Z_t(s)$$

$$Y_t(s) | \lambda_t(s) \sim LN(\lambda_t(s), \sigma^2)$$

$$\lambda_t(s) = \alpha_0 + S_t(s).$$

(i) Condicional completa para γ_0 :

$$p(\gamma_0 | \mathbf{Y}, \Theta_{-\gamma_0}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_0}^2} (\gamma_0 - \mu_{\gamma_0})^2 \right\} \exp \left\{ \gamma_0 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\ \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + \gamma_3 I(Y_{t-3}(s_i) > 0) + Z_t(s_i)\}) \right\}.$$

(ii) Condicional completa para γ_1 :

$$p(\gamma_1 | \mathbf{Y}, \Theta_{-\gamma_1}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_1}^2} (\gamma_1 - \mu_{\gamma_1})^2 \right\} \exp \left\{ \gamma_1 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) I(Y_{t-1}(s_i) > 0) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\ \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + \gamma_3 I(Y_{t-3}(s_i) > 0) + Z_t(s_i)\}) \right\}.$$

(iii) Condicional completa para γ_2 :

$$p(\gamma_2 | \mathbf{Y}, \Theta_{-\gamma_2}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_2}^2} (\gamma_2 - \mu_{\gamma_2})^2 \right\} \exp \left\{ \gamma_2 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) I(Y_{t-2}(s_i) > 0) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\ \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + \gamma_3 I(Y_{t-3}(s_i) > 0) + Z_t(s_i)\}) \right\}.$$

(iv) Condicional completa para γ_3 :

$$p(\gamma_3 | \mathbf{Y}, \Theta_{-\gamma_3}) \propto \exp \left\{ -\frac{1}{2\sigma_{\gamma_3}^2} (\gamma_3 - \mu_{\gamma_3})^2 \right\} \exp \left\{ \gamma_3 \sum_{i=1}^{27} \sum_{t=1}^{75} X_t(s_i) I(Y_{t-3}(s_i) > 0) \right\} \\ \exp \left\{ -\sum_{i=1}^{27} \sum_{t=1}^{75} \log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \right. \\ \left. + \gamma_2 I(Y_{t-2}(s_i) > 0) + \gamma_3 I(Y_{t-3}(s_i) > 0) + Z_t(s_i)\}) \right\}.$$

(v) Condicional completa para α_0 :

$$p(\alpha_0 | \mathbf{Y}, \Theta_{-\alpha_0}) \propto \exp \left\{ -\frac{1}{2\sigma_{\alpha_0}^2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} \\ \exp \left\{ \sum_{i=1}^{27} \sum_{t=1}^{75} -X_t(s_i) \frac{1}{2\sigma^2} (\log Y_t(s_i) - (\alpha_0 + S_t(s_i)))^2 \right\}.$$

(vi) Condicional completa para σ^2 :

$$p(\alpha_0 | \mathbf{Y}, \Theta_{-\alpha_0}) \propto \sigma^{2-(a+1)} \exp \left\{ -\frac{b}{\sigma^2} \right\} \exp \left\{ \sum_{i=1}^{27} \sum_{t=1}^{75} -\frac{X_t(s_i)}{2} \log \sigma^2 \right\} \\ \exp \left\{ \sum_{i=1}^{27} \sum_{t=1}^{75} -X_t(s_i) \frac{1}{2\sigma^2} (\log Y_t(s_i) - (\alpha_0 + S_t(s_i)))^2 \right\}.$$

(vii) Condicional completa para $Z_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$:

$$p(Z_t(s_i) | \mathbf{Y}, \Theta_{-Z_t(s_i)}) \propto \exp \left\{ -\frac{1}{2VZ} (Z_t(s_i) - mZ)^2 \right\} \exp \{Z_t(s_i) X_t(s_i)\} \\ \exp \{-\log(1 + \exp\{\gamma_0 + \gamma_1 I(Y_{t-1}(s_i) > 0) \\ + \gamma_2 I(Y_{t-2}(s_i) > 0) + \gamma_3 I(Y_{t-3}(s_i) > 0) + Z_t(s_i)\})\}.$$

(viii) Condicional completa para $S_t(s_i)$, $t = 1, \dots, T$ e $i = 1, \dots, n$:

$$p(S_t(s_i) | \mathbf{Y}, \Theta_{-S_t(s_i)}) \propto \exp \left\{ -\frac{1}{2VS} (S_t(s_i) - mS)^2 \right\} \\ \exp \left\{ -X_t(s_i) \frac{1}{2\sigma^2} (\log Y_t(s_i) - (\alpha_0 + S_t(s_i)))^2 \right\}.$$

(ix) Condicional completa para σ_1^2 :

$$p(\sigma_1^2 | \mathbf{Y}, \Theta_{-\sigma_1^2}) \propto \sigma_1^{2-(a_1+1)} \exp \left\{ -\frac{b_1}{\sigma_1^2} \right\} \sigma_1^{2\left(-\frac{n}{2}\right)} \left\{ -\frac{1}{2\sigma_1^2} \sum_{t=1}^T \mathbf{Z}_t' R^{-1} \mathbf{Z}_t \right\}.$$

Onde:

$$R_{ij} = \exp\{-\phi_1 \parallel s_i - s_j \parallel\} \quad i, j = 1, s, n.$$

(x) Condicional completa para σ_2^2 :

$$p(\sigma_2^2 \mid \mathbf{Y}, \Theta_{-\sigma_2^2}) \propto \sigma_2^{2-(a_2+1)} \exp\left\{-\frac{b_2}{\sigma_2^2}\right\} \sigma_2^{2\left(-\frac{n}{2}\right)} \exp\left\{-\frac{1}{2\sigma_2^2} \sum_{t=1}^T \mathbf{S}_t' W^{-1} \mathbf{S}_t\right\}.$$

Onde

$$W_{ij} = \exp\{-\phi_2 \parallel s_i - s_j \parallel\} \quad i, j = 1, s, n.$$

(xi) Condicional completa para ϕ_1 :

$$p(\phi_1 \mid \mathbf{Y}, \Theta_{-\phi_1}) \propto \phi_1^{(\alpha_1-1)} \exp\{-\beta_1 \phi_1\} \mid \Sigma \mid^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \sum_{t=1}^T \mathbf{Z}_t' \Sigma^{-1} \mathbf{Z}_t\right\}.$$

Onde

$$\Sigma_{ij} = \sigma_1^2 \exp\{-\phi_1 \parallel s_i - s_j \parallel\} \quad i, j = 1, s, n.$$

(xii) Condicional completa para ϕ_2 :

$$p(\phi_2 \mid \mathbf{Y}, \Theta_{-\phi_2}) \propto \phi_2^{(\alpha_2-1)} \exp\{-\beta_2 \phi_2\} \mid \Psi \mid^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \sum_{t=1}^T \mathbf{S}_t' \Psi^{-1} \mathbf{S}_t\right\}.$$

Onde

$$\Psi_{ij} = \sigma_2^2 \exp\{-\phi_2 \parallel s_i - s_j \parallel\} \quad i, j = 1, s, n.$$

Apêndice C

Cálculos para o Modelo de Função de Transferência

Neste apêndice serão descritos todos os cálculos para as densidades propostas dos parâmetros ϵ_{i0} , $i = 1, \dots, 156$, β e α_0 , do modelo de função de transferência, abordado na Seção 5.3.

Para o modelo descrito em (5.13) e (5.17) consideremos, as seguintes observações e variâncias ajustadas:

$$\tilde{Y}_{it} = \log(\lambda_{it}e_{it}) + (\lambda_{it}e_{it})^{-1}(Y_{it} - \lambda_{it}e_{it}), \quad (\text{C.1})$$

$$\tilde{V}_{it} = (\lambda_{it}e_{it})^{-1}. \quad (\text{C.2})$$

(i) Proposta para ϵ_{i0} :

$$q(\epsilon_{i0}, \epsilon_{i0}^p) = p(\epsilon_{i0} | \hat{\mathbf{Y}}, \Theta_{-\epsilon_{i0}}) \propto \exp \left\{ - \sum_{t=1}^T \frac{(\hat{Y}_{it} - \rho^t \epsilon_{i0}^p)^2}{2\hat{V}_{it}} \right\} \exp \left\{ - \frac{(\epsilon_{i0}^p - m_{\epsilon_{i0}})^2}{2\sigma_{\epsilon_{i0}}^2} \right\},$$

onde:

$$\begin{aligned} \hat{Y}_{it} &= \tilde{Y}_{it} - \alpha_0 - \beta \sum_{j=1}^t \rho^{t-j} \text{chuva}_{it}, \\ \epsilon_{i0}^p | \epsilon_{i0}^c &\sim N(\mu_{\epsilon_{i0}}, V_{\epsilon_{i0}}). \end{aligned} \quad (\text{C.3})$$

Onde:

$$V_{\epsilon_{i0}} = \left(\sigma_{\epsilon_{i0}}^{-2} + \sum_{t=1}^T \frac{\rho^{2t}}{\hat{V}_{it}} \right)^{-1} \quad \mu_{\epsilon_{i0}} = V_{\epsilon_{i0}} \left(\sigma_{\epsilon_{i0}}^{-2} m_{\epsilon_{i0}} + \sum_{t=1}^T \frac{\hat{Y}_{it} \rho^t}{\hat{V}_{it}} \right). \quad (\text{C.4})$$

(ii) Proposta para β :

$$q(\beta, \beta^p) = p(\beta \mid \hat{\mathbf{Y}}, \Theta_{-\beta}) \propto \exp \left\{ - \sum_{t=1}^T \sum_{i=1}^n \frac{(\hat{Y}_{it} - \beta^p \sum_{j=1}^t \rho^{t-j} \text{chuva}_{it})^2}{2\hat{V}_{it}} \right\} \\ \exp \left\{ - \frac{(\beta^p - m_{\beta})^2}{2\sigma_{\beta}^2} \right\},$$

onde:

$$\hat{Y}_{it} = \tilde{Y}_{it} - \alpha_0 - S_{it} - \rho^t \epsilon_{i0}, \\ \beta^p \mid \beta^c \sim N(\mu_{\beta}, V_{\beta}). \quad (\text{C.5})$$

Onde:

$$V_{\beta} = \left(\sigma_{\beta}^{-2} + \sum_{t=1}^T \sum_{i=1}^n \frac{F^2(\text{chuva})}{\hat{V}_{it}} \right)^{-1} \quad \mu_{\beta} = V_{\beta} \left(\sigma_{\beta}^{-2} m_{\beta} + \sum_{t=1}^T \sum_{i=1}^n \frac{\hat{Y}_{it} F(\text{chuva})}{\hat{V}_{it}} \right). \quad (\text{C.6})$$

$$F(\text{chuva}) = \sum_{j=1}^t \rho^{t-j} \text{chuva}_{it}. \quad (\text{C.7})$$

(iii) Proposta para α_0 :

$$q(\alpha_0, \alpha_0) = p(\alpha_0 \mid \hat{\mathbf{Y}}, \Theta_{-\alpha_0}) \propto \exp \left\{ - \sum_{t=1}^T \sum_{i=1}^n \frac{(\hat{Y}_{it} - \alpha_0)^2}{2\hat{V}_{it}} \right\} \exp \left\{ - \frac{(\alpha_0^p - m_{\alpha_0})^2}{2\sigma_{\alpha_0}^2} \right\},$$

onde:

$$\hat{Y}_{it} = \tilde{Y}_{it} - \epsilon_{it}, \\ \alpha_0^p \mid \alpha_0^c \sim N(\mu_{\alpha_0}, V_{\alpha_0}). \quad (\text{C.8})$$

Onde:

$$V_{\alpha_0} = \left(\sigma_{\alpha_0}^{-2} + \sum_{t=1}^T \sum_{i=1}^n \frac{1}{\hat{V}_{it}} \right)^{-1} \quad \mu_{\alpha_0} = V_{\alpha_0} \left(\sigma_{\alpha_0}^{-2} m_{\alpha_0} + \sum_{t=1}^T \sum_{i=1}^n \frac{\hat{Y}_{it}}{\hat{V}_{it}} \right). \quad (\text{C.9})$$

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](#)