Douglas Roberto Mesquita Azevedo

Modelagem espaço-temporal para Campos Aleatórios Gaussianos Transformados.

Belo Horizonte

2016, Março

Douglas Roberto Mesquita Azevedo

Modelagem espaço-temporal para Campos Aleatórios Gaussianos Transformados.

Universidade Federal de Minas Gerais – UFMG

Instituto de Ciências Exatas

Programa de Pós-Graduação

Orientador: Marcos Oliveira Prates

Belo Horizonte

2016, Março

Este trabalho é dedicado à minha família pela confiança e pelo apoio incondicional.

Agradecimentos

Em primeiro lugar agradeço aos meus pais que mais uma vez suportaram a distância e não mediram esforços para que eu pudesse terminar mais uma etapa da minha vida. O apoio e confiança de vocês me motiva a seguir adiante.

Agradeço aos meus irmãos pela confiança e apoio incondicional em todos os momentos. Agradeço também pelas frases de incentivo e ao bom humor de sempre.

Agradeço ao Marcos por ter sido um excelente orientador, por dedicar seu tempo para me transmitir um pouco do seu conhecimento (as vezes rabiscado no rodapé de um artigo) e pela oportunidade de trabalhar no InfoSAS o qual contribuiu de forma grandiosa para o meu crescimento profissional.

Agradeço aos meus amigos de longa data: Douglas, Gulliti, Maicon e Mateus pelas conversas, risadas e amizade. Espero que dessa vez vocês paguem pelo menos o churrasco!

Agradeço aos colegas de Mestrado Estevão, Fernanda, Gabi e Juliana pelas manhãs, tardes, noites e as vezes finais de semana de estudo. Sem vocês o caminho seria muito mais difícil (e os teoremas também!).

Agradeço ao amigo e colega de trabalho Luís pelos ensinamentos e por nunca ter negado ajuda. Aprendi muito contigo neste período, muito obrigado!

Agradeço à Larissa pela paciência nos últimos dias e pelo carinho e companheirismo de sempre. Agradeço pelo apoio nas horas em que precisei e por sempre estar disposta a ajudar. E agradeço por ser uma cozinheira de mão cheia (mãe, não fique com inveja!).

Agradeço aos meus professores da graduação e pós-graduação pelo conhecimento compartilhado e à Maiza pelo papo matinal e pela água para o chimarrão.

Agradeço também a todos que não foram citados mas que torcem por mim e que de alguma maneira contribuem ou contribuíram para o meu crescimento.

"Todos estes que aí estão
Atravancando meu caminho,
Eles passarão.
Eu passarinho!"
(Mário Quintana)

Resumo

Modelos capazes de capturar características espaciais e temporais tem grande aplicabilidade em diversas áreas do conhecimento. Modelos espaço-temporais não separáveis são descritos na literatura e visam entender estes tipos de características, entretanto estes modelos comumente são complexos e com uma interpretação pouco intuitiva. No presente trabalho, a metodologia dos Campos Aleatórios Markovianos Gaussianos Transformados – *TGMRF* é generalizada para o contexto espaço-temporal através de uma estrutura não-separável. A estrutura de dependência utilizada é flexível e fornece uma interpretação simples para os parâmetros associados ao espaço, tempo e espaço-tempo. Outra vantagem dos *TGMRF's* é a flexibilidade na escolha direta da distribuição marginal do parâmetro de interesse. A construção dos modelos espaço-temporais utilizando os *TGMRF's* proposta neste trabalho fornece novas classes de modelos, por exemplo Campos Aleatórios Markovianos Gama em que é possível modelar diretamente as intensidades Poisson no espaço-tempo. O modelo proposto é aplicado a um conjunto de dados em que o interesse é entender as características do *habitat* da espécie de gastrópode *Nenia Tridens*. Para isso considerou-se importantes variáveis ambientais além das dimensões espaço, tempo e espaço-tempo.

Palavras-chaves: Estruturas de dependência não separáveis, Campos Aleatórios Gama, Modelos Lineares Generalizados Mistos, Confundimento espacial.

Abstract

Models that are capable of capturing the spatial and temporal characteristics of the data are applicable in many science fields. Non-separable spatio-temporal models were introduced in the literature to capture these features, however, these models are usually complicated in its interpretation and construction. In this work, we introduce a class of non-separable Transformed Gaussian Markov Random Fields (TGMRF) where the dependence structure is not only flexible but also provides simple interpretation to the spatial, temporal and spatio-temporal parameter in the random effects. Another advantage is that the TGMRF settings allow specialists to define any desired margins. Therefore, the construction of spatio-temporal models using the TGMRF framework leads to a new class of models such as spatio-temporal gamma random fields, that can be direclty used to model Poisson intensity for space-time data. The proposed models were applied to the abundance data of *Nenia Tridens* to pick out important environmental variables that affect their abundance and also study possible spatial and temporal trends.

Key-words: Non-separable dependency structures, Gamma Random Fields (TGRF), Generalized Linear Mixed Models (GLMM), Spatial confounding.

Lista de ilustrações

Figura 1 – Mapa de municípios brasileiros infestados pelo vírus da Dengue para os anos de 1995 e 2010. Fonte: R7 Notícias 10
Figura 2 – Confundimento espacial. (a): Covariável $\boldsymbol{X}$. (b): Efeitos espaciais $\boldsymbol{\epsilon}$. . . 14
Figura 3 – Grafo não direcionado e matriz de proximidade. 17
Figura 4 – Exemplo de regiões e matriz de precisão do modelo *CAR*. 18
Figura 5 – Exemplo de regiões e matriz de precisão do modelo *MCAR*. 20
Figura 6 – Tipos de vizinhança. (a): vizinhança dentro de variáveis. (b): vizinhança dentro de regiões. (c): vizinhança entre variáveis. 21
Figura 7 – Ilustração da quantidade $b_{ij,kl}$. 22
Figura 8 – Exemplo de regiões e matriz de precisão do modelo proposto por Rodrigues (2012). 25
Figura 9 – Caixa: Limites inferiores e superiores de $\boldsymbol{\rho}$. Sólido: Verdadeiro espaço de $\boldsymbol{\rho}$. 26
Figura 10 – Curvas dos modelos empregados para um exemplo simulado a partir do modelo LN. 45
Figura 11 – Cenários de simulação com coeficientes atemporais. 50
Figura 12 – Cenários de simulação com coeficientes temporais. 51
Figura 13 – Cenário 1: *Boxplots* das médias *a posteriori* dos coeficientes de regressão segundo o modelo gerador e modelo de ajuste. 54
Figura 14 – Cenário 5: *Boxplots* das médias *a posteriori* dos coeficientes de regressão segundo o modelo gerador e modelo de ajuste. 58
Figura 15 – Látice e contagens do número de caramujos da espécie *Nenia Tridens* para os anos de 2010 à 2014 e estrutura de vizinhança adotada. 61
Figura 16 – Cadeias do parâmetro ρ_t. (a): LN. (b): GI. (c): GSH. (d): GSC 63
Figura 17 – Densidades *a posteriori* para os coeficientes associados às covariáveis segundo os modelos ajustados. 64
Figura 18 – Intervalos *HPD* e mediana *a posteriori* dos coeficientes associados às covariaveis segundo os modelos ajustados. 66
Figura 19 – Contagem de gastrópodes e elevações no período em análise. 68

Lista de tabelas

Tabela 1 – Esperança e variância dos modelos utilizados. 45
Tabela 2 – Cenário 1: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas. 53
Tabela 3 – Cenário 1: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$. 55
Tabela 4 – Cenário 5: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas. 57
Tabela 5 – Cenário 5: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$. 59
Tabela 6 – Análise descritiva das covariáveis utilizadas na aplicação. 62
Tabela 7 – Estudo 1: Médias *a posteriori* para os coeficientes associados às covariáveis e intervalos *HPD*. 64
Tabela 8 – Estudo 2: Médias *a posteriori* para os coeficientes associados às covariáveis e intervalos *HPD*. 65
Tabela 9 – Cenário 2: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$. 71
Tabela 10 – Cenário 2: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas. 72
Tabela 11 – Cenário 3: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$. 73
Tabela 12 – Cenário 3: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas. 74
Tabela 13 – Cenário 4: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$. 75
Tabela 14 – Cenário 4: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas. 76
Tabela 15 – Cenário 6: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$. 77
Tabela 16 – Cenário 6: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas. 78

Sumário

1 INTRODUÇÃO 10

2 OBJETIVOS 12

3 MODELOS LINEARES GENERALIZADOS MISTOS 13

4 MODELOS ESPACIAIS E ESPAÇO-TEMPORAIS 16
4.1 *Conditional Autoregressive - CAR* 16
4.2 *Multivariate Conditional Autoregressive - MCAR* 18
4.3 Estruturas de precisão mais convenientes 21

5 CAMPOS ALEATÓRIOS MARKOVIANOS GAUSSIANOS TRANSFORMADOS 27
5.1 *TGMRF's* em modelos hierárquicos espaciais 28

6 *TGMRF'S* EM MODELOS HIERÁRQUICOS ESPAÇO-TEMPORAIS 31

7 REGRESSÃO POISSON 33
7.1 Regressão Poisson Espacial 33
7.2 Regressão Poisson Espaço-Temporal 35

8 INFERÊNCIA 38
8.1 Métodos de amostragem 38
8.2 Modelo de amostragem adotado 42
8.3 Modelos de ajuste 43

9 SELEÇÃO DE MODELOS 46
9.1 *Deviance Information Criterion - DIC:* 46
9.2 *Watanabe-Akaike Information Criterion - WAIC:* 47
9.3 *Logarithm of the Pseudo-Marginal Likelihood - LPML:* 48

10 ESTUDO DE SIMULAÇÃO 50

11 ABUNDÂNCIA DE *NENIA TRIDENS* 61

12 CONCLUSÕES 69

13 APÊNDICE - ESTUDO DE SIMULAÇÃO 71

REFERÊNCIAS 79

1 Introdução

Conhecer e entender as fontes de variação em conjuntos de dados é de grande importância nas mais diversas áreas do conhecimento, podendo ajudar na identificação de padrões e auxiliar o pesquisador na tomada de decisões. Em muitos casos, o conjunto de dados possui estrutura espacial ou espaço-temporal, tornando imprescindíveis modelos capazes de identificar a influência destas diferentes estruturas nos conjuntos de dados. Por exemplo, pode-se ter interesse em identificar algum padrão espacial relacionado a infestação do vírus da dengue. Além disso, é plausível e interessante investigar a evolução temporal da infestação deste vírus. Na Figura 1 observa-se o mapa dos municípios brasileiros infestados pela dengue nos anos de 1995 e 2010.

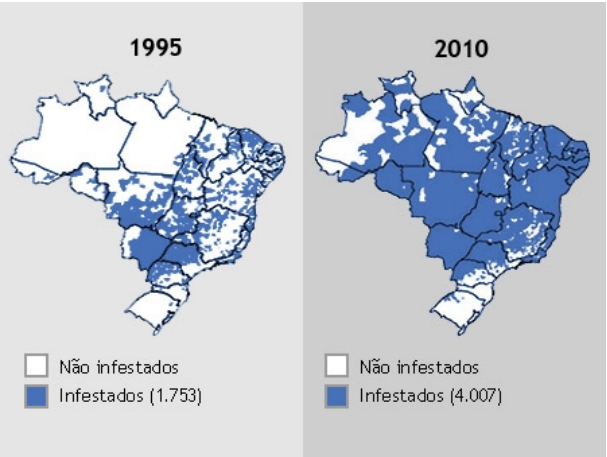


Figura 1 – Mapa de municípios brasileiros infestados pelo vírus da Dengue para os anos de 1995 e 2010. Fonte: R7 Notícias [1]

Claramente em 1995 os municípios infestados estavam concentrados na região Centro-Oeste, além de vários municípios nas regiões Sudeste e Nordeste. Já em 2010, a infestação se espalha pelas regiões Centro-Oeste, Nordeste, Sudeste e Norte. Muitas covariáveis podem influenciar na disseminação do vírus, como por exemplo variáveis climáticas ou sócio-econômicas. Além disso, o alastramento pode ter algum padrão espacial, temporal ou espaço-temporal. Entender e quantificar a influência de cada uma destas componentes é de grande valor para a saúde pública, pois assim, pode-se especficar os motivos do alastramento do vírus e criar políticas visando o seu controle.

Uma classe de modelos capaz de acomodar tais características é a dos Modelos Lineares Generalizados Mistos *GLMM* (BRESLOW; CLAYTON, 1993). A ideia do modelo é explicar linearmente uma função da média condicional através de possíveis efeitos fixos

[1] Disponível em: http://noticias.r7.com/saude/noticias/em-quatro-meses-2010-se-torna-o-segundo-ano-com-maior-numero-de-mortes-por-dengue-20100711.html. Acesso em março de 2016

e/ou aleatórios. A função da média condicional é chamada de função de ligação e seu papel é mapear o suporte da média condicional nos número reais. Diferentes escolhas para a função de ligação geram diferentes ajustes e interpretações, tornando a escolha dessa função um passo de suma importância na modelagem (LI; DUAN, 1989; CZADO; SANTNER, 1992).

Ao se trabalhar com um conjunto de dados com estrutura espacial, pode-se assumir que existe um efeito aleatório relacionado a cada região em estudo. Esse efeito pode apresentar algum tipo de estrutura, em que cada região terá um nível de dependência com suas regiões vizinhas. Uma forma de incorporar estruturas de dependência espacial para o efeito aleatório é utilizar o modelo *Conditional Autoregressive* (BESAG, 1974, CAR). Neste modelo a estrutura de vizinhança é dada por um grafo não direcionado representado por uma matriz composta de zeros e uns, em que o número um significa que as regiões são vizinhas e zero significa que as regiões não são vizinhas. A partir da matriz de vizinhança deriva-se a estrutura de dependência entre as regiões em estudo.

A ideia do modelo *CAR* pode ser adaptada para modelar conjuntos de dados com estrutura espacial em que diversas variáveis dependentes são observadas. Esses modelos são chamados de *Multivariate Conditional Autoregressive* (*MCAR*). Uma possível aplicação dessa abordagem é no contexto espaço-temporal, em que a variável dependente é observada em todas regiões do espaço em diferentes momentos do tempo. Alguns trabalhos foram desenvolvidos com o intuito de generalizar a estrutura da matriz de precisão do modelo *CAR* para o modelo *MCAR*, tentando levar em conta não só a dependência espacial mas também a dependência entre as variáveis (GELFAND; VOUNATSOU, 2003; CARLIN; BANERJEE, 2003; JIN; CARLIN; BANERJEE, 2005).

Sain et al. (2011) apresenta uma forma alternativa de definir a estrutura de precisão a qual leva em consideração não só a dependência espacial e entre as variáveis mas também a interação destes dois tipos de estrutura. Rodrigues (2012) propõe modificações nesta estrutura de dependência visando uma melhor interpretação para a esperança e variância dos efeitos aleatórios.

Os *Transformed Gaussian Markov Random Fields* (PRATES et al., 2015, TGMRF) aparecem como uma maneira alternativa para modelagem de dados espaciais. A abordagem consiste em escolher diretamente a distribuição marginal da média no modelo de interesse e, então, introduzir dependência espacial através do uso de uma cópula Gaussiana. Esse método apresenta duas vantagens: 1) é possível fazer escolhas de distribuições apropriadas para as marginais da média; 2) não possui confundimento espacial visto que a estrutura do modelo separa as marginais da estrutura de dependência espacial. Além disso, existe uma correspondência entre a estrutura tradicional de *GLMM* e a introduzida nos *TGMRF's*, sendo vantajosa a escolha dos *TGMRF's* para alguns casos de interesse.

2 Objetivos

O presente trabalho tem como objetivo principal generalizar os resultados de Prates et al. (2015) para o caso espaço-temporal. Para isso é necessário o uso de uma estrutura de dependência não separável adequada. Esta estrutura de dependência espaço-temporal deve ser simples e interpretável e atualmente não existe na literatura uma estrutura popularizada. Neste trabalho a estrutura de dependência adotada será dada por uma adaptação da estrutura proposta por Rodrigues (2012). Essa formulação permite uma fácil interpretação dos parâmetros de dependência espacial, temporal e espaço-temporal. Em contrapartida existe uma dificuldade na estimação dos parâmetros associados a esta estrutura.

A utilização dos *TGMRF's* traz uma maior flexibilidade na modelagem visto que é possível escolher distribuições adequadas para a marginal do parâmetro de média de interesse. Além disso, a estrutura imposta pelos *TGMRF's* evita o confundimento espacial e também o confundimento temporal.

Dada a flexibilidade dos *TGMRF's* uma grande variedade de modelos podem ser adotados. Desta forma, o presente trabalho irá comparar diferentes critérios de seleção de modelos a fim de identificar dentre uma variedade de modelos aquele cujo ajuste aos dados é o melhor.

Por fim a metodologia é aplicada à contagem de caramujos da espécie *Nenia Tridens*. Esta espécie é responsável pela recomposição da fertilidade do solo e é de interesse dos especialistas caracterizar o seu *habitat* ideal. Para isso o estudo conta com diversas covariáveis consideradas importantes para o desenvolvimento da espécie.

3 Modelos Lineares Generalizados Mistos

A classe dos Modelos Lineares Generalizados Mistos (BRESLOW; CLAYTON, 1993, GLMM) é uma classe bastante ampla capaz de acomodar diversos tipos de estruturas. A ideia básica desta abordagem é explicar uma função da média condicional através de possíveis efeitos fixos e aleatórios. De forma geral,

$$\boldsymbol{Y}|\boldsymbol{\mu},\boldsymbol{\gamma} \sim f(\boldsymbol{Y};\boldsymbol{\mu},\boldsymbol{\gamma})$$
$$g(\boldsymbol{\mu}) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\epsilon},$$

em que $\boldsymbol{Y}$ é o conjunto de dados observados, $f(\boldsymbol{Y};\boldsymbol{\mu},\boldsymbol{\gamma})$ é uma distribuição pertencente a família exponencial associada aos dados, $g(.)$ é chamada de função de ligação cujo papel é mapear o domínio da média condicional nos números reais, $\boldsymbol{\mu}$ é um parâmetro de média, $\boldsymbol{\gamma}$ são outros possíveis parâmetros associados à distribuição $f(.)$, $\boldsymbol{X}$ é uma matriz de covariáveis de dimensão $n \times q$, $\boldsymbol{\beta}$ é um vetor de coeficientes de dimensão q associado às covariáveis $\boldsymbol{X}$, $\boldsymbol{\epsilon}$ são possíveis efeitos aleatórios associados ao modelo e $\boldsymbol{Z}$ é a matriz de delineamento associada aos efeitos aleatórios (STROUP, 2012).

Na abordagem dos *GLMM's* a escolha da função de ligação é um passo muito importante na modelagem sendo que escolhas equivocadas podem levar à conclusões erradas a cerca do modelo (LI; DUAN, 1989; CZADO; SANTNER, 1992). O grande problema é que para alguns modelos poucas opções de funções de ligação estão descritas na literatura, um limitante para a análise.

Uma possível aplicação dos *GLMM's* são os modelos espaciais em que o objetivo é considerar a relação entre as regiões do espaço na análise. Estas estruturas podem ser inseridas através de efeitos aleatórios e para este tipo de análise dá-se o nome de *Spatial Generalized Linear Mixed Models* (*SGLMM*).

Uma limitação da abordagem dos *SGLMM's* é o chamado confundimento espacial que ocorre quando existe uma combinação linear de variáveis regressoras que estão fortemente correlacionadas com os efeitos aleatórios das regiões do espaço. Desta forma, o modelo não é capaz de identificar o efeito das variáveis regressoras ou dos efeitos espaciais (REICH; HODGES; ZADNIK, 2006).

Exemplo 1: Seja $\boldsymbol{Y} = Y_1, \ldots, Y_{900}$ variáveis aleatórias observadas em 900 regiões em que $Y_i|\mu_i \sim Poisson(\mu_i)$, $i = 1, \ldots, 900$. Sejam β_0 e β_1 o intercepto e um coeficiente de regressão associado a uma covariável $\boldsymbol{X}$, seja $\boldsymbol{\epsilon} = \epsilon_1, \ldots, \epsilon_{900}$ efeitos aleatórios relacionados às 900 regiões do espaço e considere a função de ligação logarítmica. Desta forma tem-se o seguinte Modelo Linear Generalizado Misto:

$$Y_i|\mu_i \sim Poisson(\mu_i)$$
$$\log(\mu_i) = \beta_0 + \beta_1 X_i + \epsilon_i.$$

Considere também que a covariável $\boldsymbol{X}$ esteja trazendo a mesma informação que o efeito espacial $\boldsymbol{\epsilon}$ tal como na Figura 2. Apesar de $\boldsymbol{X}$ e $\boldsymbol{\epsilon}$ estarem em escalas diferentes, o padrão de cores no mapa é muito semelhante.

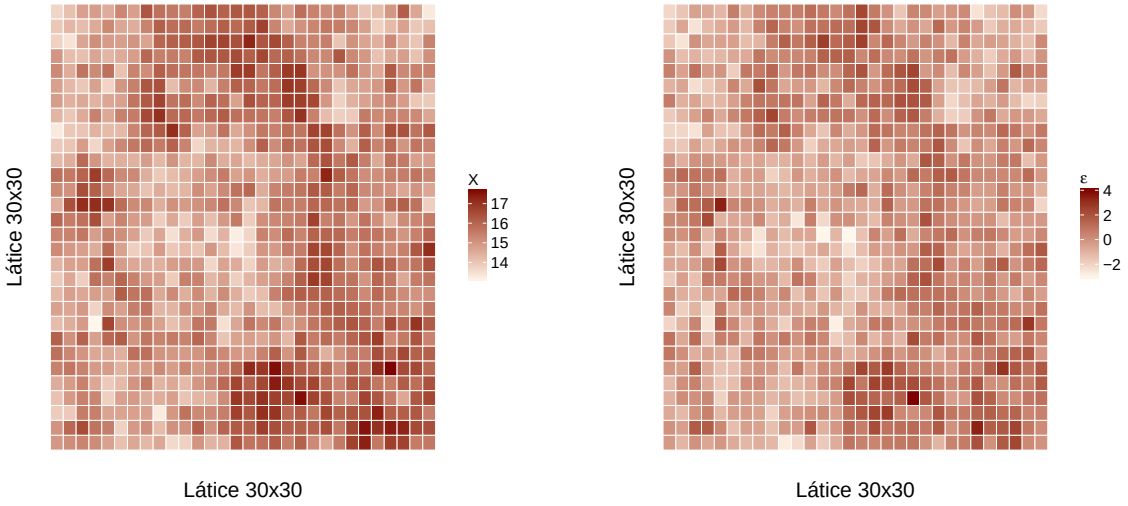


Figura 2 – Confundimento espacial. (a): Covariável $\boldsymbol{X}$. (b): Efeitos espaciais $\boldsymbol{\epsilon}$.

Suponha ainda que $\boldsymbol{X} = \boldsymbol{\epsilon} + b + \boldsymbol{e}$, em que $\boldsymbol{\epsilon} \sim N_{900}(\boldsymbol{0}, \sigma^2 \boldsymbol{I_{900}})$, b é uma constante aditiva e $\boldsymbol{e}$ um termo de erro. Pode-se reescrever cada $\epsilon_i = \sigma\varepsilon_i$ em que $\varepsilon_i \sim N(0,1)$, ou seja, $\boldsymbol{X}$ é uma combinação linear de $\boldsymbol{\varepsilon}$ mais um ruído. Se $\boldsymbol{e} = \boldsymbol{0}$ fica claro que os parâmetros não serão bem estimados dado que o modelo se torna não identificável,

$$\begin{aligned} log(\mu_i) &= \beta_0 + \beta_1 \boldsymbol{X_i} + \epsilon_i \\ &= \beta_0 + \beta_1(\epsilon_i + b) + \epsilon_i \\ &= \beta_0 + \beta_1(\sigma\varepsilon_i + b) + \sigma\varepsilon_i \\ &= \beta_0 + \beta_1 b + (\beta_1 + 1)\,\sigma\varepsilon_i. \end{aligned} \tag{3.1}$$

Neste sentido o confundimento espacial se assemelha ao problema da multicolinearidade em um modelo de regressão, porém ocorre entre as covariáveis e o efeito espacial. A relação entre as covariáveis e o efeito espacial não necessariamente é linear o que dificulta a evidenciação do problema. Caso $\boldsymbol{e} \approx \boldsymbol{0}$ então o confundimento espacial tem uma semelhança com a chamada multicolinearidade severa.

Uma outra aplicação dos modelos *GLMM's* se dá quando os dados são observados em diferentes regiões do espaço e em diferentes instantes de tempo, caracterizando a abordagem espaço-temporal. Neste caso pode-se pensar em duas formas de modelagem:

- Considerando efeitos aleatórios separáveis para os efeitos espacial, temporal e espaço-temporal com diferentes estruturas.
- Considerando um efeito aleatório não separável para os efeitos espacial, temporal e espaço-temporal com apenas um tipo de estrutura.

A abordagem separável é mais simples. Nela é possível utilizar uma estrutura puramente espacial para a componente espacial, uma estrutura puramente temporal para a componente temporal e uma estrutura para modelar o efeito espaço-temporal. Neste caso supõem-se e insere-se no modelo três tipos de estruturas, o que pode não ser adequado em alguns casos. Por exemplo, quando não existe efeito espacial e se insere esta componente há um aumento desnecessário da complexidade do modelo.

A abordagem não separável tem como maior desafio a obtenção de uma estrutura adequada para o efeito aleatório. Dada uma estrutura adequada, tem-se um modelo mais parcimonioso visto que o número de parâmetros é menor. Esta estrutura deve ser simples e facilmente interpretável o que torna a análise mais efetiva.

Nos modelos espaço-temporais além do confundimento espacial existe também o confudimento temporal que ocorre quando existe uma combinação linear de variáveis regressoras que estão fortemente correlacionadas com os efeitos temporais.

Alguns tipos de estruturas são comumente utilizadas para modelagem espacial e espaço-temporal em conjunto com os modelos *GLMM*, como por exemplo os modelos CAR e MCAR. Na Seção 4 alguns modelos espaciais e espaço-temporais são descritos.

4 Modelos Espaciais e Espaço-temporais

Problemas que envolvem a observação de uma variável aleatória em diferentes regiões do espaço são muito comuns nas mais diversas áreas do conhecimento. Para dados com estrutura espacial observados em regiões bem definidas (dados de área), um modelo comumente utilizado é o *CAR*. Entretanto, ele foi construído para o caso univariado e não contempla o caso espaço-temporal. A generalização do modelo *CAR* é chamada de *MCAR*. No modelo *MCAR* é possível trabalhar com o caso multivariado incluindo como caso particular o modelo espaço-temporal. O grande problema do *MCAR* está na definição da estrutura de covariâcias que não se dá de forma simples.

4.1 *Conditional Autoregressive - CAR*

O modelo *CAR* é utilizado para modelar dados com estrutura espacial, em que apenas uma variável dependente é observada em cada região. Suponha que exista interesse em descrever uma variável aleatória que foi observada em n regiões do espaço. Na abordagem Bayesiana tem-se um efeito aleatório relacionado à cada região em estudo para o qual assume-se uma estrutura *CAR*. O modelo *CAR* mais utilizado é o Autonormal, em que se atribui distribuições condicionais Gaussianas para os efeitos aleatórios nas n regiões.

Sejam $Y_1, \ldots, Y_n$, variáveis aleatórias observadas em n regiões e seja $\mathbf{\Theta} = (\theta_1, \ldots, \theta_n)^T$, efeitos aleatórios com média zero, relacionados às n regiões. O modelo *CAR* define as distribuições condicionais

$$(\theta_i|\theta_{-\mathbf{i}}) \sim N\left(\sum_{j\sim i} b_{ij}\theta_j, \tau_i^2\right), \tag{4.1}$$

em que a notação $\theta_{-\mathbf{i}}$ representa o vetor $\mathbf{\Theta}$ a menos do elemento θ_i, b_{ij} é um peso que relaciona os efeitos das regiões i e j, e $j \sim i$ indica vizinhança entre as regiões i e j. Utilizando o Lema de Brook (BROOK, 1964) é possível derivar a distribuição conjunta para $\mathbf{\Theta} = (\theta_1, \ldots, \theta_n)^T$:

$$\pi(\mathbf{\Theta}) \propto \exp\left\{-\frac{1}{2}\mathbf{\Theta}^T\mathbf{D}^{-1}(\mathbf{I_n} - \mathbf{B})\,\mathbf{\Theta}\right\}, \tag{4.2}$$

em que $\mathbf{D}$ é matriz diagonal com entradas τ_i^2 e $\mathbf{B}$ é uma matriz quadrada $n \times n$ com entradas b_{ij}.

A Equação (4.2) se assemelha ao núcleo de uma distribuição normal multivariada com vetor de médias $\mathbf{0}$ e matriz de covariâncias $\mathbf{\Sigma} = (\mathbf{I_n} - \mathbf{B})^{-1}\mathbf{D}$. Porém, para isso a

matriz $\mathbf{\Sigma}^{-1}$ precisa ser simétrica e positiva definida.

Para garantir simetria, deve-se ter $\frac{b_{ij}}{\tau_i} = \frac{b_{ji}}{\tau_j}$ $\forall i, j$. Uma forma muito comum de garantir simetria é definir uma matriz de proximidade $\mathbf{W}$ vista como um grafo não direcionado, em que $w_{ij} \neq 0$ se e somente se i é conectado a j. Geralmente $\mathbf{W}$ é composta de zeros e uns, sendo que 1 indica que as regiões são vizinhas e 0 indica que as regiões não são vizinhas. Na Figura 3 pode-se ver um exemplo de um grafo não direcionado e sua respectiva matriz de proximidade.

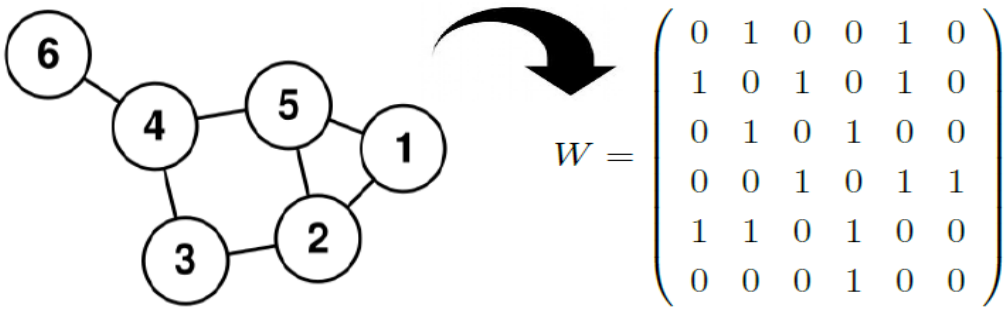


Figura 3 – Grafo não direcionado e matriz de proximidade.

Em seguida, define-se $b_{ij} = \frac{w_{ij}}{w_{i+}}$, onde w_{i+} é a soma dos elementos da linha i da matriz $\mathbf{W}$, ou seja, representa o número de vizinhos da região i. Além disso define-se que a variância em cada região depende apenas do número de vizinhos desta região, ou seja, $\tau_i^2 = \frac{\tau^2}{w_{i+}}$. Veja que com esta configuração a matriz $\mathbf{\Sigma}^{-1}$ é simétrica:

$$\frac{b_{ij}}{\tau_i} = \frac{\frac{w_{ij}}{w_{i+}}}{\frac{\tau^2}{w_{i+}}} = \frac{w_{ij}}{\tau^2} = \frac{w_{ji}}{\tau^2} = \frac{\frac{w_{ji}}{w_{j+}}}{\frac{\tau^2}{w_{j+}}} = \frac{b_{ji}}{\tau_j}.$$

Pode-se reescrever $\mathbf{\Sigma}^{-1} = \frac{1}{\tau^2}(\mathbf{D_w} - \mathbf{W})$, em que $\mathbf{D_w}$ é uma matriz diagonal com entradas w_{i+}. Nesta configuração a matriz $\mathbf{\Sigma}^{-1}$ é simétrica, porém, não é positiva definida e, portanto, não se tem uma distribuição conjunta própria.

Apesar disto, é possível utilizar esta configuração em uma abordagem Bayesiana como distribuição *a priori* para algum efeito aleatório contanto que a distribuição *a posteriori* seja própria. Esta configuração do modelo *CAR* é chamada de modelo Autoregressivo Intrínseco (ICAR), com a notação $CAR(1, \tau^2)$.

Desta forma, sejam $Y_1, \ldots, Y_n$, variáveis aleatórias observadas em n regiões, onde cada $Y_i|\mu_i$ tem distribuição $f(Y_i; \mu_i, \boldsymbol{\gamma})$. Considere $\theta_1, \ldots, \theta_n$, efeitos aleatórios relacionados às n regiões, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ um vetor de parâmetros de média e $\boldsymbol{\gamma}$ um vetor de outros possíveis parâmetros. O interesse está em estimar o vetor de parâmetros de média $\boldsymbol{\mu}$. Utilizando as ideias do modelo *GLMM* vistas na Seção 3, pode-se explicar uma função da

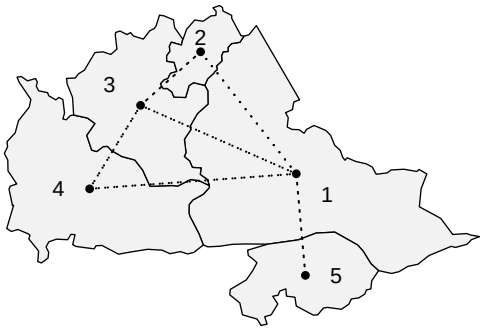
média condicional através de efeitos fixos e efeitos aleatórios. Abaixo o modelo hierárquico associado ao modelo *GLMM*:

$$\pi(Y_i|\mu_i) = f(Y_i; \mu_i, \boldsymbol{\gamma})$$
$$g(\mu_i|Y_i) = \mathbf{X_i}\boldsymbol{\beta} + \theta_i$$
$$\theta_i \sim N\left(\sum_{j\sim i} b_{ij}\theta_j, \tau_i^2\right), \beta_k \sim N\left(0, \sigma_\beta^2\right),$$

em que $\mathbf{X_i}\boldsymbol{\beta}$ são efeitos lineares de covariáveis e θ_i é um termo relacionado a região i.

Se $\boldsymbol{\Sigma}^{-1} = \frac{1}{\tau^2}(\mathbf{D_w} - \rho\mathbf{W})$, sendo $\rho \in (\lambda_{min}^{-1}, \lambda_{max}^{-1})$, onde λ_{min} e λ_{max} são respectivamente o menor e o maior autovalores de $\mathbf{D_w^{\frac{1}{2}}WD_w^{\frac{1}{2}}}$. Então, neste caso, $\boldsymbol{\Sigma}^{-1}$ é uma matriz positiva definida (BANERJEE; CARLIN; GELFAND, 2014). Esse modelo possui distribuição conjunta própria o que garante uma distribuição *a posteriori* própria quando utilizada como distribição *a priori*. Esta configuração é chamada de $CAR(\rho, \tau^2)$ e ρ é chamado de parâmetro de dependência espacial responsável por indicar a intensidade da relação entre as regiões em estudo, quanto maior o valor de ρ maior a dependência entre as regiões em estudo.

Exemplo 2: Na Figura 4 tem-se um exemplo de conjunto de regiões e também da matriz de precisão do modelo *CAR* associada a ele. Considerou-se regiões vizinhas aquelas que possuem fronteira comum. As retas pontilhadas representam a vizinhança espacial definida por tais fronteiras.



$$\boldsymbol{\Sigma^{-1}} = \frac{1}{\tau^2}\begin{pmatrix} 4 & -\rho & -\rho & -\rho & -\rho \\ -\rho & 2 & -\rho & 0 & 0 \\ -\rho & -\rho & 3 & -\rho & 0 \\ -\rho & 0 & -\rho & 2 & 0 \\ -\rho & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Figura 4 – Exemplo de regiões e matriz de precisão do modelo *CAR*.

Na Seção 4.2 alguns aspectos básicos do modelo *CAR* multivariado são apresentados.

4.2 *Multivariate Conditional Autoregressive - MCAR*

Quando mais de uma variável dependente é observada em diferentes regiões e o interesse é modelar estas variáveis conjuntamente, não é possível utilizar o modelo *CAR*.

A extensão do modelo *CAR* para quando mais de uma variável dependente é observada em cada região é chamada de *MCAR*.

Seja n o número de regiões em análise e p o número de variáveis observadas. Seja $\mathbf{Y}$ um vetor de observações ordenados por regiões, ou seja,

$$\mathbf{Y} = (Y_{11}, Y_{12}, \ldots, Y_{1p}, Y_{21}, \ldots, Y_{2p}, \ldots, Y_{n1} \ldots, Y_{np}).$$

Seja $\theta_{\mathbf{i}} = (\theta_{i1}, \ldots, \theta_{ip})$, vetor de p efeitos aleatórios relacionados à cada variável na região i. Definem-se distribuições condicionais para cada $\theta_{\mathbf{i}}$ as quais podem ser vistas na Expressão (4.3)

$$(\theta_{\mathbf{i}}|\theta_{-\mathbf{i}}) \sim N_p \left(\sum_{j \sim i} \mathbf{B_{ij}} \theta_{\mathbf{j}}, \boldsymbol{\Sigma}_{\boldsymbol{i}} \right), \tag{4.3}$$

onde $\mathbf{B_{ij}}$ é uma matriz $p \times p$ assim como $\boldsymbol{\Sigma}_{\boldsymbol{i}}$.

Novamente, através do Lema de Brook é possível encontrar a distribuição conjunta de $\boldsymbol{\Theta} = (\theta_{\mathbf{1}}, \theta_{\mathbf{2}}, \ldots, \theta_{\mathbf{n}})$:

$$\Pi(\boldsymbol{\Theta}) \propto \exp\left\{ -\frac{1}{2} \boldsymbol{\Theta}^T \boldsymbol{\Gamma}^{-1} \left(\mathbf{I_{np}} - \tilde{\mathbf{B}} \right) \boldsymbol{\Theta} \right\}, \tag{4.4}$$

em que, $\boldsymbol{\Gamma}$ é uma matriz $np \times np$ bloco diagonal com blocos $\boldsymbol{\Sigma}_{\boldsymbol{i}}$ e $\tilde{\mathbf{B}}$ é formada por blocos $\mathbf{B_{ij}}$, de forma análoga à matriz $\mathbf{B}$ do caso univariado.

Novamente é necessário garantir que $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}}^{-1} = \boldsymbol{\Gamma}^{-1} \left(\mathbf{I_{np}} - \tilde{\mathbf{B}} \right)$ seja simétrica e positiva definida. Um caso interessante é quando se tem $\mathbf{B_{ij}} = b_{ij} \mathbf{I_p}$, sendo $b_{ij} = \frac{w_{ij}}{w_{i+}}$ e $\boldsymbol{\Sigma}_{\boldsymbol{i}} = w_{i+}^{-1} \boldsymbol{\Sigma}$. Desta forma garante-se simetria de modo análogo a (4.1) e a matriz de covariâncias é dada por:

$$\boldsymbol{\Sigma}_{\boldsymbol{\Theta}}^{-1} = \boldsymbol{\Gamma}^{-1} \left(\mathbf{I_{np}} - \tilde{\mathbf{B}} \right) = (\mathbf{D_w} - \mathbf{W}) \otimes \boldsymbol{\Sigma}^{-1}. \tag{4.5}$$

Basicamente a matriz resultante é o produto de *Kronecker*, denotado por $\otimes$, entre a matriz de covariâncias do modelo *ICAR* do caso univariado e uma matriz $\boldsymbol{\Sigma}^{-1}$. Neste caso, $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}}^{-1}$ é singular e o modelo descrito é uma extensão multivariada do *ICAR*, chamado de *MCAR*(1, $\boldsymbol{\Sigma}$).

Para tornar a distribuição conjunta própria é necessário redefinir as distribuições condicionais como segue:

$$(\theta_{\mathbf{i}}|\theta_{-\mathbf{i}}) \sim N_p \left(R_i \sum_{j \sim i} \mathbf{B_{ij}} \theta_{\mathbf{j}}, \boldsymbol{\Sigma}_{\boldsymbol{i}} \right). \tag{4.6}$$

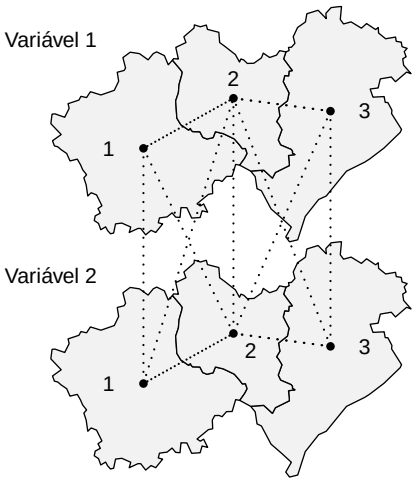
Então é possível reescrever $\mathbf{\Gamma}^{-1}\left(\mathbf{I_{np}} - \tilde{\mathbf{B}}\right)$ como $\mathbf{\Gamma}^{-1}\left(\mathbf{I_{np}} - \tilde{\mathbf{B}}_{\mathbf{R}}\right)$, onde agora $\tilde{\mathbf{B}}_{\mathbf{R}}$ terá blocos $\mathbf{R_i B_{ij}}$. Novamente se faz necessário assegurar que a matriz $\mathbf{\Sigma}_{\mathbf{\Theta}}^{-1}$ seja simétrica e positiva definida. Voltando ao caso especial onde $\mathbf{B_{ij}} = b_{ij}\mathbf{I_p}$, sendo $b_{ij} = \frac{w_{ij}}{wi+}$ e $\Sigma_i = w_{i+}^{-1}\mathbf{\Sigma}$, tem-se que uma solução é tomar $\mathbf{R_i} = \mathbf{R} = \rho\mathbf{I_p}$. As restrições para ρ são as mesmas do caso univariado. Reescrevendo a matriz de covariâncias chega-se em

$$\mathbf{\Sigma}_{\mathbf{\Theta}}^{-1} = \mathbf{\Gamma}^{-1}\left(\mathbf{I_{np}} - \tilde{\mathbf{B}}_{\mathbf{R}}\right) = (\mathbf{D_w} - \rho\mathbf{W}) \otimes \mathbf{\Sigma}^{-1}. \tag{4.7}$$

Neste caso, a matriz $\mathbf{\Sigma}_{\mathbf{\Theta}}^{-1}$ é não singular. Este modelo é chamado de $MCAR(\rho, \mathbf{\Sigma})$. É importante observar que, nesta formatação, o mesmo parâmetro de dependência espacial está sendo utilizado para todas as variáveis em estudo. Não existe nenhum parâmetro de dependência entre as variáveis e também não existe nenhum parâmetro de dependência para a interação regiões e variáveis.

Para acrescentar outros parâmetros de dependência deve-se alterar $\mathbf{R_i}$. Porém, considerar $\mathbf{R_i}$ mais geral torna a expressão intratável. Neste caso, é possível garantir simetria porém é muito complicado garantir não singularidade (GELFAND; VOUNATSOU, 2003). Para tornar $\mathbf{R_i}$ mais geral Gelfand e Vounatsou (2003) e Carlin e Banerjee (2003) propõem reordenar os dados por variável. Desta forma, é possível incorporar um parâmetro de dependência espacial para cada variável em análise, porém, ainda não é simples incorporar um parâmetro de dependência para a interação de regiões e variáveis.

Exemplo 3: Na Figura 5 tem-se um exemplo de conjunto de regiões e também da respectiva matriz de precisão do modelo *MCAR*. Novamente considerou-se regiões vizinhas aquelas que possuem fronteira comum. As retas pontilhadas representam três tipos de vizinhança que são apresentadas na Seção 4.3 ou mais especificamente na Figura 6. É importante observar que a modelagem *MCAR* abordada neste trabalho apenas quantifica a dependência espacial desconsiderando os demais tipos de vizinhança.



$$\mathbf{\Sigma}_{\mathbf{\Theta}}^{-1} = \begin{pmatrix} 1 & -\rho & -\rho \\ -\rho & 2 & -\rho \\ -\rho & -\rho & 1 \end{pmatrix} \otimes \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1}.$$

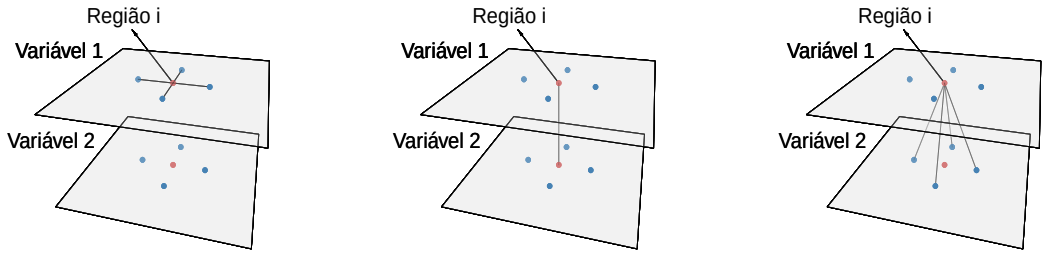
Figura 5 – Exemplo de regiões e matriz de precisão do modelo *MCAR*.

Na Seção 4.3 alguns aspectos básicos de outros dois tipos de estruturas de precisão são apresentados.

4.3 Estruturas de precisão mais convenientes

Sain et al. (2011) propõem uma forma alternativa de definir a estrutura de precisão. Esta estrutura leva em conta não só a dependência espacial mas também a dependência entre as variáveis e a dependência das regiões entre as variáveis. Para isso definem-se três tipos de vizinhança. A Figura 6 exibe os três tipos de vizinhança tomando como base uma região i qualquer e considerando o caso bivariado. Em cada um dos gráficos tem-se cinco regiões representadas pelos pontos e dois planos que representam duas diferentes variáveis.

Na Figura 6a tem-se a vizinhança dentro de variáveis (*within variable*), essa vizinhança fornece a estrutura espacial. Na Figura 6b tem-se a vizinhança dentro de regiões (*within location*), que fornece a estrutura de vizinhança entre a mesma região nas diferentes variáveis observadas. Na Figura 6c tem-se a vizinhança entre variáveis (*cross-variable*), que fornece a estrutura de vizinhança da interação entre região e variável.



(a) (b) (c)

Figura 6 – Tipos de vizinhança. (a): vizinhança dentro de variáveis. (b): vizinhança dentro de regiões. (c): vizinhança entre variáveis.

Sain et al. (2011) propõem expressar a média condicional dos efeitos aleatórios em função de três quantidades diferentes. Seja $\boldsymbol{\Theta} = (\theta_{11}, \ldots, \theta_{1p}, \theta_{21}, \ldots, \theta_{2p}, \ldots, \theta_{n1} \ldots, \theta_{np})$, onde θ_{ij} é o efeito aleatório relacionado à região i na variável j. A média condicional de θ_{ij} é dada por:

$$E\left(\theta_{ij}|\theta_{-\{ij\}}\right) = \mu_{ij} + \underbrace{\sum_{k\neq i} b_{ij,kj}\left(\theta_{kj} - \mu_{kj}\right)}_{A} + \underbrace{\sum_{l\neq j} b_{ij,il}\left(\theta_{il} - \mu_{il}\right)}_{B} + \underbrace{\sum_{k,l\neq i,j} b_{ij,kl}\left(\theta_{kl} - \mu_{kl}\right)}_{C}, \tag{4.8}$$

e a variância condicional é dada por:

$$Var\left(\theta_{ij}|\theta_{-\{ij\}}\right) = \tau_{ij}^2, \tag{4.9}$$

em que cada μ_{ij} representa a média da variável j na região i e $b_{ij,kl}$ é um peso que relaciona as regiões i e k de acordo com as variáveis j e l. Este peso define qual a importância da variável l observada na região k na média da variável j observada na região i tal como mostra a Figura 7.

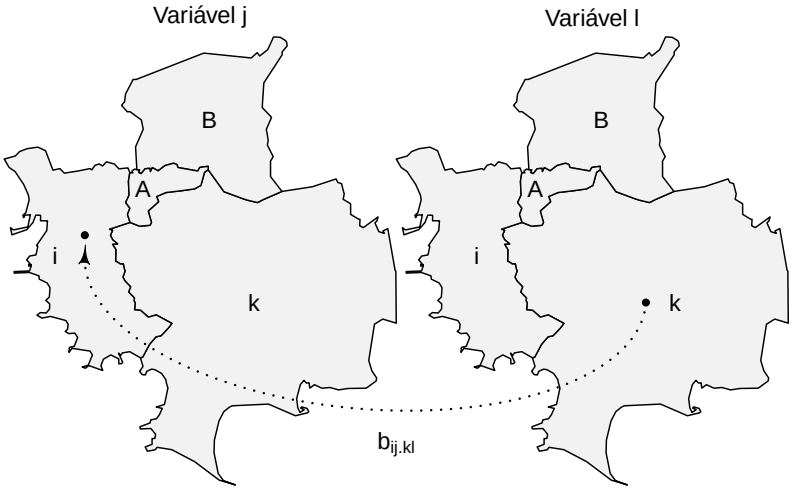


Figura 7 – Ilustração da quantidade $b_{ij,kl}$.

O somatório A em (4.8) leva em conta apenas a dependência entre os vizinhos espaciais de i dentro de uma mesma variável j, esse tipo de vizinhança está representada na Figura 6a. O somatório B leva em conta a vizinhança da região i com ela mesma nas diferentes variáveis e está representada na Figura 6b. Por último, o somatório C leva em conta a vizinhança da região i dentro de uma variável j com seus vizinhos espaciais nas demais variáveis e está representada na Figura 6c.

De maneira geral, cada região i tem $v_i = d_i p + p - 1$ vizinhos, onde d_i é o número de vizinhos espaciais da região i e p o número de variáveis/tempos. O número de vizinhos não está sendo levado em consideração em (4.8) e (4.9).

Rodrigues (2012) define a matriz de covariâncias de forma que a média e a variância condicional dependam do número de vizinhos da seguinte maneira:

$$E\left(\theta_{ij}|\theta_{-\{ij\}}\right) = \frac{1}{v_i}\left(\mu_{ij} + \sum_{k\neq i} b_{ij,kj}\left(\theta_{kj} - \mu_{kj}\right) + \sum_{l\neq j} b_{ij,il}\left(\theta_{il} - \mu_{il}\right) + \sum_{k,l\neq i,j} b_{ij,kl}\left(\theta_{kl} - \mu_{kl}\right)\right), \tag{4.10}$$

e

$$Var\left(\theta_{ij}|\theta_{-\{ij\}}\right) = \frac{\tau_j^2}{v_i}. \tag{4.11}$$

Esta configuração é mais interessante pois nesse caso a esperança condicional de um efeito aleatório relacionado a uma variável j em uma região i é uma média dos seus vizinhos e a variância também depende do número de vizinhos.

Pode-se pensar em definir pesos $b_{ij,kl}$ de acordo com o tipo de vizinhança. Por exemplo, se há indícios de que a vizinhança espacial é mais influente nesse conjunto de dados, basta tomar os pesos que entram no somatório A de (4.10) maiores que os demais.

Define-se que a distribuição conjunta dos efeitos aleatórios é Normal np-variada com vetor de médias $\mathbf{0}$ e matriz de precisão $\mathbf{Q}$ de dimensão $np \times np$ dada por:

$$\mathbf{T}^{-1}\begin{pmatrix} \mathbf{A_1} + \mathbf{I_p}(v_1 - 1) & \mathbf{B_{12}}\delta_{12} & \dots & \mathbf{B_{1(n-1)}}\delta_{1(n-1)} & \mathbf{B_{1n}}\delta_{1n} \\ \mathbf{B_{21}}\delta_{21} & \mathbf{A_2} + \mathbf{I_p}(v_2 - 1) & \dots & \mathbf{B_{2(n-1)}}\delta_{2(n-1)} & \mathbf{B_{2n}}\delta_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{B_{(n-1)1}}\delta_{(n-1)1} & \mathbf{B_{(n-1)2}}\delta_{(n-1)2} & \dots & \mathbf{A_{n-1}} + \mathbf{I_p}(v_{n-1} - 1) & \mathbf{B_{(n-1)n}}\delta_{(n-1)n} \\ \mathbf{B_{n1}}\delta_{n1} & \mathbf{B_{n2}}\delta_{n2} & \dots & \mathbf{B_{n(n-1)}}\delta_{n(n-1)} & \mathbf{A_n} + \mathbf{I_p}(v_n - 1) \end{pmatrix}, \tag{4.12}$$

em que $\boldsymbol{\delta}$ é uma matriz de proximidade composta de 0's e 1's, $\mathbf{T} = \text{diag}\left(\tau_{11}^2, \dots, \tau_{1p}^2, \dots, \tau_{n1}^2, \dots, \tau_{np}^2\right)$, $\mathbf{A_i}$ e $\mathbf{B_{ij}}$ são matrizes quadradas $p \times p$ dadas por

$$\mathbf{A_i} = \begin{pmatrix} 1 & \dots & -b_{i1,ip} \\ \vdots & \ddots & \vdots \\ -b_{ip,i1} & \dots & 1 \end{pmatrix} \qquad \mathbf{B_{ij}} = \begin{pmatrix} -b_{i1,j1} & \dots & -b_{i1,jp} \\ \vdots & \ddots & \vdots \\ -b_{ip,j1} & \dots & -b_{ip,jp} \end{pmatrix}.$$

Para garantir simetria de $\mathbf{Q}^{-1}$, assume-se variâncias iguais para cada variável/tempo. Ou seja, a variância não muda de acordo com a região. Desta forma, $\tau_{ij}^2 = \tau_j^2$ para todo j e então $\mathbf{T} = \mathbf{I_n} \bigotimes \text{diag}(\boldsymbol{\tau})$, em que $\boldsymbol{\tau} = \left[\tau_1^2, \dots, \tau_p^2\right]$.

$\mathbf{Q}^{-1}$ é simétrica se seus blocos são simétricos. Para os blocos das diagonais tem-se,

$$\text{diag}\left(\boldsymbol{\tau}\right)^{-1}\mathbf{A_i} = \begin{pmatrix} \frac{1}{\tau_1^2} & \dots & \frac{-b_{i1,ip}}{\tau_1^2} \\ \vdots & \ddots & \vdots \\ \frac{-b_{ip,i1}}{\tau_p^2} & \dots & \frac{1}{\tau_p^2} \end{pmatrix}.$$

Este bloco é simétrico se $\frac{-b_{ij,il}}{\tau_j^2} = \frac{-b_{il,ij}}{\tau_l^2}$. Definir $b_{ij,il} = \frac{\rho_{jl}\tau_j}{\tau_l}$, com $\rho_{lj} = \rho_{jl}$, é uma maneira de tornar o bloco simétrico. Desta forma, obtem-se $\text{diag}(\boldsymbol{\tau})^{-1}\mathbf{A_i} = \text{diag}(\boldsymbol{\tau})^{-\frac{1}{2}}\mathbf{A}\text{diag}(\boldsymbol{\tau})^{-\frac{1}{2}}$, em que

$$\mathbf{A} = \begin{pmatrix} 1 & \dots & -\rho_{1p} \\ \vdots & \ddots & \vdots \\ -\rho_{1p} & \dots & 1 \end{pmatrix}.$$

De forma análoga pode-se obter a matriz $\mathbf{B}$ dada por,

$$\mathbf{B} = \begin{pmatrix} \psi_{11} & \dots & \psi_{1p} \\ \vdots & \ddots & \vdots \\ \psi_{1p} & \dots & \psi_{pp} \end{pmatrix},$$

sendo, $b_{ij,kl} = \frac{\psi_{jl}\tau_j}{\tau_l}$. Desta forma a matriz de precisão em (4.12) pode ser reescrita como segue:

$$\mathbf{Q} = \left(\mathbf{I_n} \otimes \mathbf{T}^{-\frac{1}{2}}\right) \begin{pmatrix} \mathbf{A} + \mathbf{I_p}(v_1 - 1) & \dots & -\delta_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ -\delta_{1n}\mathbf{B} & \dots & \mathbf{A} + \mathbf{I_p}(v_n - 1) \end{pmatrix} \left(\mathbf{I_n} \otimes \mathbf{T}^{-\frac{1}{2}}\right). \tag{4.13}$$

Após algumas simplificações,

$$\mathbf{Q} = \left(\mathbf{I_n} \otimes \mathbf{T}^{-\frac{1}{2}}\right) \left[\mathbf{I_n} \otimes (\mathbf{A} - \mathbf{I_p}) + \mathbf{V} \otimes \mathbf{I_p} - \boldsymbol{\delta} \otimes \mathbf{B}\right] \left(\mathbf{I_n} \otimes \mathbf{T}^{-\frac{1}{2}}\right), \tag{4.14}$$

sendo $\mathbf{V} = \text{diag}(v_i)$.

Neste trabalho a estrutura de precisão utilizada é a dada na Equação (4.14). Além disso, o trabalho foca em análises espaço-temporais. Desta forma, fez-se as seguintes adaptações:

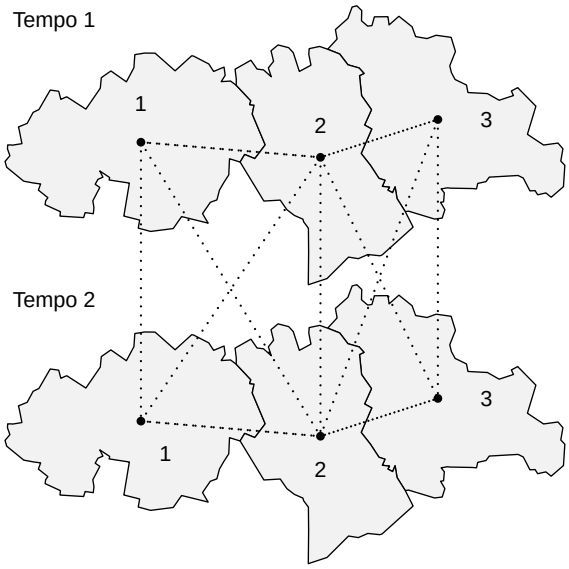
$$\rho_s = \psi_{jj}, \quad j = 1, \dots, t$$

$$\rho_t = \rho_{jl} = \rho_{lj}, \quad j = 2, \dots, t-1; \quad l = j-1, j+1$$

$$\rho_{st} = \psi_{jl} = \psi_{lj}, \quad j = 2, \dots, t-1; \quad l = j-1, j+1$$

em que ρ_t é responsável por captar a dependência entre as observações no tempo, ρ_s é responsável por captar a dependência das regiões com suas regiões vizinhas, ρ_{st} é responsável por captar a dependência espaço-temporal das observações e t é o número de tempos na análise. Note que da forma como os parâmetros foram definidos a dependência temporal ocorre apenas de um tempo para outro assim como num modelo autorregressivo de ordem 1.

Exemplo 4: Na Figura 8 tem-se um exemplo de conjunto de regiões e respectiva matriz de precisão do modelo proposto por Rodrigues (2012). Novamente considerou-se regiões vizinhas aquelas que possuem fronteira comum. As retas pontilhadas representam a vizinhança espacial, temporal e espaço-temporal. Considere $\boldsymbol{\Upsilon} = \left(\boldsymbol{I_n} \otimes \boldsymbol{T}^{-\frac{1}{2}}\right)$



$$\boldsymbol{Q} = \boldsymbol{\Upsilon} \begin{pmatrix} 3 & -\rho_t & -\rho_s & -\rho_{st} & 0 & 0 \\ -\rho_t & 3 & -\rho_{st} & -\rho_s & 0 & 0 \\ -\rho_s & -\rho_{st} & 5 & -\rho_t & -\rho_s & -\rho_{st} \\ -\rho_{st} & -\rho_s & -\rho_t & 5 & -\rho_{st} & -\rho_s \\ 0 & 0 & -\rho_s & -\rho_{st} & 3 & -\rho_t \\ 0 & 0 & -\rho_{st} & -\rho_s & -\rho_t & 3 \end{pmatrix} \boldsymbol{\Upsilon}.$$

Figura 8 – Exemplo de regiões e matriz de precisão do modelo proposto por Rodrigues (2012).

Nesta configuração tem-se que $\mathbf{Q}$ é simétrica porém não é simples garantir inversibilidade pois o espaço conjunto de $\boldsymbol{\rho} = (\rho_s, \rho_t, \rho_{st})$ é complexo além de variar de acordo com o número de regiões e tempos. Para garantir inversibilidade, utilizou-se um critério suficiente porém não necessário que se baseia na dominância diagonal (RUE; HELD, 2005). Se $Q_{ii} - \sum_{j:j\neq i} |Q_{ij}| > 0, \forall i$, então $\boldsymbol{Q}$ é positiva definida e portanto possui inversa. Dada a relação entre os parâmetros $\rho_s, \rho_t,$ e ρ_{st} não se sabe o domínio exato de $\boldsymbol{\rho}$, utilizando a restrição apresentada somos apenas capazes de determinar os limites do domínio deste vetor de parâmetros

A Figura 9 ilustra o problema apresentado. A caixa representa os limites inferiores e superiores de cada um dos parâmetros e o sólido representa o verdadeiro espaço dos parâmetros.

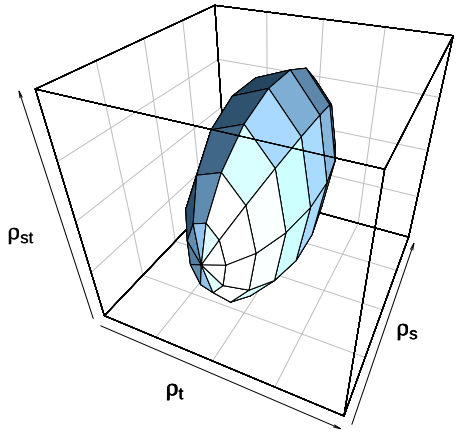


Figura 9 – Caixa: Limites inferiores e superiores de $\boldsymbol{\rho}$. Sólido: Verdadeiro espaço de $\boldsymbol{\rho}$.

No próximo capítulo, as ideias básicas de Campos Aleatórios (Markovianos) Gaussianos (Transformados) serão introduzidas.

5 Campos Aleatórios Markovianos Gaussianos Transformados

Um Campo Aleatório Gaussiano do inglês *Gaussian Random Field (GRF)* é um processo estocástico no qual um vetor aleatório $\mathbf{Y}$ de dimensão n possui distribuição marginal Gaussiana k-variada para $k \leq n$. De forma geral, $\mathbf{Y} \sim Normal(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e as variâncias condicionais entre os elementos de $\mathbf{Y}$ são dadas pela matriz de precisão $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$ (RUE; HELD, 2005).

O vetor aleatório $\mathbf{Y}$ pode ser representado por um grafo não direcionado $\mathbf{G}$ em que $\mathbf{Q_{ij}} \neq 0$ se e somente se Y_i é conectado a Y_j em $\mathbf{G}$. Nos casos em que $\mathbf{Y}$ é um processo Markoviano, $\mathbf{Y}$ é chamado de Campo Aleatório Markoviano Gaussiano do inglês *Gaussian Markov Random Field (GMRF)*. Em um *GMRF* a distribuição condicional de qualquer Y_i depende apenas dos elementos aos quais Y_i é conectado no grafo $\mathbf{G}$.

Os *GMRF's* são comumente utilizados nos casos onde é mais fácil definir a matriz de precisão $\mathbf{Q}$ do que a estrutura de covariâncias $\boldsymbol{\Sigma}$ e possuem muitas aplicações na literatura. Por exemplo, podem ser utilizados em conjunto com Modelos Lineares Generalizados Mistos para modelar efeitos cuja distribuição assumida é a distribuição Gaussiana.

Apesar das inúmeras aplicações e das boas características, os *GMRF's* possuem algumas limitações tais como a não alocação de assimetria ou caudas pesadas visto que as distribuições marginais são Gaussianas, o que torna o modelo pouco flexível (PRATES, 2011).

Campos Aleatórios Gaussianos Transformados do inglês *Transformed Gaussian Random Fields (TGRF)*, são uma possível solução para as limitações dos *GRF's*. Nesta classe, a distribuição marginal é qualquer distribuição que se queira, podendo dessa forma inserir caudas pesadas e assimetria ao modelo, características que são comumente observadas em dados reais. Além disso, os *TGRF* possuem algumas das propriedades dos *GRF*. Se o processo a ser modelado possui a propriedade de Markov, então tem-se os Campos Aleatórios Markovianos Gaussianos Transformados do inglês *Transformed Gaussian Markov Random Fields (TGMRF)*, que mantêm algumas das propriedades dos *GMRF*.

Como o nome sugere, os *TGRF's* são processos estocásticos obtidos através de transformações marginais aplicadas aos *GRF's*. Para definir um *TGRF* de dimensão n, primeiramente define-se um conjunto $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^T$ com distribuição normal multivariada com vetor de médias $\mathbf{0}$ e matriz de correlações dada por $\boldsymbol{\Psi}$, ou seja $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \boldsymbol{\Psi})$ e, consequentemente, $\boldsymbol{\epsilon}$ é um *GRF*. Seja $\mathbf{Z} = (Z_1, \ldots, Z_n)^T$ e $Z_i = F_i^{-1}\{\Phi(\epsilon_i)\}$, $i = 1, \ldots, n$,

em que $F_i(x)$ é uma função de distribuição acumulada absolutamente contínua em relação a x e Φ é a distribuição acumulada de uma $N(0,1)$. Desta forma, cada Z_i tem distribuição marginal f_i (função densidade associada a F_i) e então denota-se $\mathbf{Z} \sim TGRF_n(\mathbf{F}, \mathbf{\Psi})$ em que $\mathbf{F} = (F_1, \ldots, F_n)$. Nesse contexto, assim como no contexto dos *GRF's*, a matriz $\mathbf{Q} = \mathbf{\Psi}^{-1}$ traz uma interpretação mais intuitiva das distribuições condicionais e pode ser utilizada nos casos onde é mais intuitivo definir $\mathbf{Q}$ do que $\mathbf{\Psi}$. A propriedade de Markov é herdada pelo *TGRF*, ou seja, para $i \neq j$, $Z_i \perp Z_j | \mathbf{Z}_{-\{ij\}}$ se e somente se $Q_{ij} = 0$, em que $\mathbf{Z}_{-\{ij\}}$ é o vetor $\mathbf{Z}$ sem a i-ésima e a j-ésima observações, neste caso, denota-se $\mathbf{Z} \sim TGMRF_n(\mathbf{F}, \mathbf{Q})$.

Os *TGMRF's* podem ser utilizados para, por exemplo, modelar diretamente as intensidades Poisson ou então as taxas de uma Bernoulli de acordo com uma distribuição de interesse. Na Seção 5.1 tem-se a aplicação dos *TGMRF's* em modelos espaciais.

5.1 *TGMRF's* em modelos hierárquicos espaciais

Modelos Lineares Generalizados Mistos são comumente utilizados para modelagem de dados com estrutura espacial. *GLMM's* possuem níveis hierárquicos em que em um primeiro nível tem-se a verossimilhança, em um segundo nível uma função da média condicional é explicada por possíveis efeitos fixos e aleatórios e nos demais níveis os efeitos do modelo (fixos ou aleatórios) podem ser modelados através de estruturas convenientes. Como visto na Seção 4.2, uma forma de modelar o efeito espacial é através do modelo *CAR*.

Considere $\mathbf{Y} = (Y_1, \ldots, Y_n)$, um vetor aleatório observado em n regiões, $\mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_n})$ uma matriz de covariáveis com dimensão $q \times n$ e $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$ efeitos aleatórios relacionados às n regiões em estudo cuja distribuição conjunta será denotada por H.

Se a distribuição de cada Y_i pertence à família exponencial com média $\mu_i = E(Y_i | \mathbf{X}, \boldsymbol{\epsilon})$, então podemos modelar a esperança condicional segundo a abordagem dos *GLMM's*. Sendo $g(.)$ uma função de ligação adequada, o modelo para a média condicional pode ser visto como:

$$g(\mu_i) = \mathbf{X_i}^T \boldsymbol{\beta} + \epsilon_i, \tag{5.1}$$

em que $\boldsymbol{\beta}$ é o vetor de coeficientes de regressão associado às covariáveis do modelo. Utilizando a abordagem dos *GRF's*, H é uma distribuição $N_n(\mathbf{0}, \mathbf{\Sigma})$ e a dependência entre os componentes de $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$ fica determinada pela matriz $\mathbf{\Sigma^{-1}}$.

Em particular, se a distribuição de cada $Y_i | \mu_i$ é $Poisson(\mu_i)$ e a função de ligação

adotada é $g(.) = \log(.)$, então

$$\log(\mu_i) = \mathbf{X_i}^T\boldsymbol{\beta} + \epsilon_i, \tag{5.2}$$

e portanto $\mu_i = \exp\left(\mathbf{X_i}^T\boldsymbol{\beta} + \epsilon_i\right)$. Como $\boldsymbol{\epsilon}$ é um *GRF*, a distribuição conjunta de $\boldsymbol{\mu}$ é $Log - Normal(\mathbf{X}^T\boldsymbol{\beta}, \boldsymbol{\Sigma})$. Ou seja, a escolha da função de ligação logarítimica implica em modela $\boldsymbol{\mu}$ através de uma distribuição Log-Normal.

Uma alternativa a essa abordagem é utilizar os *TGRF's*, em que se escolhe a distribuição marginal de $\boldsymbol{\mu}$ e não a função de ligação $g(.)$. Por exemplo, no caso da regressão Poisson a modelagem da média condicional é realizada de forma direta para $\boldsymbol{\mu}$ e então pode-se transformar um Campo Aleatório Gaussiano em, por exemplo, um Campo Aleatório Gama.

A notação utilizada para essa abordagem é dada por:

$$\boldsymbol{\mu} \sim TGRF_n(\mathbf{F}, \boldsymbol{\Sigma}), \tag{5.3}$$

em que $\mathbf{F} = (F_1, \ldots, F_n)$, F_i é a distribuição acumulada relacionada à distribuição marginal (f_i) de μ_i e $\boldsymbol{\Sigma}$ é a matriz de covariâncias de $\boldsymbol{\mu}$. O modelo convencional dos *GLMM's* pode ser recuperado, basta tomar $\mathbf{F_i}$ como distribuição acumulada associada a $g^{-1}\left(\mathbf{X}^T\boldsymbol{\beta} + \boldsymbol{\epsilon}\right)$, o que nem sempre é simples. No caso da regressão Poisson basta tomar cada F_i como sendo a distribuição acumulada de uma $Log - N(\mathbf{X_i}^T\boldsymbol{\beta}, \Sigma_{ii})$.

Se o processo em análise possui a propriedade de Markov o modelo (5.3) pode ser reescrito como:

$$\boldsymbol{\mu} \sim TGMRF_n(\mathbf{F}, \mathbf{Q}). \tag{5.4}$$

Comumente, a estrutura de dependência espacial é dada pela matriz de precisão $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$, visto que neste caso é mais comum definir a matriz de precisão que rege a dependência entre as regiões em estudo. A matriz $\mathbf{Q}$ pode ser, por exemplo, a estrutura definida pelo modelo *CAR* visto na Seção 4.1.

Seja $\boldsymbol{\xi} = (\boldsymbol{\beta}, \nu, \mathbf{X}, \rho)$, um modelo hierárquico utilizando *TGMRF* pode ser definido como segue,

$$\begin{gathered} Y_i|\mu_i \sim \pi\left(y|\mu_i\right), i = 1, \ldots, n, \\ \boldsymbol{\mu} \sim TGMRF_n\left(\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_\rho\right), \\ \pi\left(\boldsymbol{\beta}\right), \pi\left(\nu\right), \pi\left(\rho\right), \end{gathered}$$

em que $\mathbf{F}_{\boldsymbol{\xi}}$ indica que $\mathbf{F}$ pode depender da matriz de covariáveis $\mathbf{X}$, do vetor de coeficientes de regressão $\boldsymbol{\beta}$, de um possível parâmetro de dispersão ν e do parâmetro de dependência espacial ρ. Já a matriz $\mathbf{Q}_{\rho}$ depende apenas do parâmetro de dependência espacial ρ e da estrutura de vizinhança. Neste modelo, os parâmetros de interesse são $\boldsymbol{\varphi} = (\boldsymbol{\beta}, \nu, \rho, \boldsymbol{\mu})$.

Dada a forma como os *TGMRF's* são obtidos, o modelo não sofre com o problema do cofundimento espacial já que em momento algum as estruturas latentes e fixas se misturam. A componente espacial está totalmente ligada ao *GRF* enquanto as covariáveis aparecem apenas no momento da transformação deste *GRF*.

No Capítulo 6 a teoria dos *TGMRF's* será generalizada para o caso em que as observações são coletadas em regiões do espaço em diferentes momentos do tempo, ou seja, para o caso espaço-temporal.

6 *TGMRF's* em modelos hierárquicos espaço-temporais

A análise de dados observados ao longo do tempo é importante na identificação de padrões temporais. Quando as observações são avaliadas em diferentes regiões, tem-se então além do padrão temporal os padrões espacial e espaço-temporal. Avaliar esses padrões conjuntamente não é um trabalho simples. A metodologia dos *TGMRF's* pode ser utilizada para modelar dados com estrutura espaço-temporal, sendo necessária a obtenção de uma estrutura de dependência adequada.

Como visto na Seção 4.3 a estrutura proposta por Rodrigues (2012) se dá de forma simples e consegue alocar efeitos temporais, espaciais e espaço-temporais em sua definição. Neste trabalho a estrutura de dependência espaço-temporal é dada pela matriz $\mathbf{Q} = \mathbf{\Sigma}^{-1}$ e optou-se pela estrutura dada na Equação (4.14) definida na Seção 4.3.

Os *TGMRF's* espaço-temporais são obtidos de forma similar ao que foi visto na Seção 5.1. Considere $\mathbf{Y} = (Y_{11}, \ldots, Y_{1t}, Y_{21}, \ldots, Y_{2t}, \ldots, Y_{n1} \ldots, Y_{nt})$, um vetor aleatório observado em n regiões e t tempos distintos, $\mathbf{X} = (\mathbf{X_{11}}, \ldots, \mathbf{X_{1t}}, \mathbf{X_{21}}, \ldots, \mathbf{X_{2t}}, \ldots, \mathbf{X_{n1}}, \ldots, \mathbf{X_{nt}})$ uma matriz de covariáveis com dimensão $q \times (nt)$ e $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_{1t}, \epsilon_{21}, \ldots, \epsilon_{2t}, \ldots, \epsilon_{n1}, \ldots, \epsilon_{nt})$ vetor de efeitos aleatórios relacionados às n regiões sob estudo em cada um dos t tempos cuja distribuição conjunta é $N_{nt}(\mathbf{0}, \mathbf{\Sigma})$.

Se a distribuição de cada Y_{ij} pertence à família exponencial com média $\mu_{ij} = E(Y_{ij}|\mathbf{X}, \boldsymbol{\epsilon})$, então novamente pode-se modelar a esperança condicional segundo a abordagem dos *TGMRF's*. Desta forma,

$$\boldsymbol{\mu} \sim TGRF_{nt}(\mathbf{F}, \boldsymbol{Q}), \tag{6.1}$$

em que $\mathbf{F} = (F_{11}, \ldots, F_{1t}, F_{21}, \ldots, F_{2t}, \ldots, F_{n1}, \ldots, F_{nt})$, F_{ij} é a distribuição acumulada relacionada à distribuição marginal (f_{ij}) de μ_{ij} e $\boldsymbol{Q}$ é a matriz de precisão de $\boldsymbol{\mu}$.

Se o processo em que se está trabalhando possui a propriedade de Markov o modelo (6.1) pode ser reescrito como:

$$\boldsymbol{\mu} \sim TGMRF_{nt}(\mathbf{F}, \mathbf{Q}). \tag{6.2}$$

Seja $\boldsymbol{\xi} = (\boldsymbol{\beta}, \nu, \mathbf{X}, \rho_s, \rho_t, \rho_{st})$, um modelo hierárquico espaço-temporal utilizando *TGMRF* pode ser definido como segue:

$$Y_{ij}|\mu_{ij} \sim \pi\left(y|\mu_{ij}\right), i = 1, \ldots, n; j = 1, \ldots t$$

$$\boldsymbol{\mu} \sim TGMRF_{nt}\left(\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_{\rho_s,\rho_t,\rho_{st}}\right)$$
$$\pi\left(\boldsymbol{\beta}\right),\ \pi\left(\nu\right),\ \pi\left(\rho_s, \rho_t, \rho_{st}\right),$$

em que $\mathbf{F}_{\boldsymbol{\xi}}$ indica que $\mathbf{F}$ pode depender da matriz de covariáveis $\mathbf{X}$, do vetor de coeficientes de regressão $\boldsymbol{\beta}$, de um possível parâmetro de dipersão ν e dos parâmetros de dependência espacial, temporal e espaço-temporal denotados por ρ_s, ρ_t e ρ_{st}. Já a matriz $\mathbf{Q}_{\rho_s,\rho_t,\rho_{st}}$ depende do parâmetro de dependência espacial (ρ_s), temporal (ρ_t), espaço-temporal (ρ_{st}) e da estrutura de vizinhança. Neste modelo, os parâmetros de interesse são $\boldsymbol{\varphi} = (\boldsymbol{\beta}, \nu, \rho_s, \rho_t, \rho_{st}, \boldsymbol{\mu})$.

Esta modelagem não sofre com os problemas de confundimento espacial e/ou temporal dado que a estrutura espaço-temporal não está diretamente ligada à distribuição marginal. Desta forma, a metodologia dos *TGMRF's* pode ser aplicada em modelos espaço-temporais e abre um grande leque de possibilidades de análises. Dado um conjunto de dados de interesse, é possível ajustar uma grande variedade de modelos não se limitando aos modelos fornecidos pela abordagem dos *GLMM's* nos casos em que não há variedades de funções de ligação descritas na literatura.

No Capítulo 7, o modelo de Regressão Poisson é apresentado. Este modelo foi utilizado para a aplicação da técnica dos *TGMRF's* deste trabalho.

7 Regressão Poisson

Quando se observa uma variável aleatória de contagem o modelo probabilístico comumente empregado é o Poisson cujo parâmetro de média está definido em $\mathbb{R}^+$. Desta forma, segundo a abordagem dos *TGMRF's*, pode-se pensar em modelar a média condicional através de um Campo Aleatório Gama ou então qualquer distribuição definida nos números reais positivos. Na Seção 7.1 apresenta-se a Regresão Poisson no contexto espacial e na Seção 7.2 apresenta-se uma versão espaço-temporal dessa abordagem

7.1 Regressão Poisson Espacial

Sejam $Y_1, \ldots, Y_n$, variáveis aleatórias observadas em n regiões de um determinado espaço em que $Y_i|\mu_i \sim Poisson(\mu_i)$, para $i = 1, \ldots, n$. Seja $\mathbf{X}$ uma matriz de covariáveis, com dimensão $q \times n$. Os *TGMRF's* introduzem efeitos espaciais através da distribuição conjunta de $\boldsymbol{\mu}$. Seja $\boldsymbol{\xi} = (\boldsymbol{\beta}, \nu, \mathbf{X}, \rho)$, então tem-se:

$$
\begin{gathered}
Y_i|\mu_i \sim Poisson(\mu_i), i = 1, \ldots, n \\
\boldsymbol{\mu} \sim TGMRF_n\left(\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_\rho\right) \\
\pi\left(\boldsymbol{\beta}\right), \pi\left(\nu\right), \pi\left(\rho\right),
\end{gathered}
$$

em que $\boldsymbol{\beta}$ é um vetor de coeficientes de dimensão $q \times 1$, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$ segue um modelo *CAR* com matriz de precisão $\mathbf{Q}_\rho^*$ e $\mathbf{F}_{\boldsymbol{\xi}} = (F_1, \ldots, F_n)$ sendo F_i a distribuição acumulada de qualquer distribuição de probabilidade f_i definida em $\mathbb{R}^+$.

A inversa da matriz $\mathbf{Q}_\rho^*$ não é uma matriz de correlações válida. Desta forma define-se:

$$
\mathbf{Q}_\rho = \boldsymbol{\Lambda}^{\frac{1}{2}} \mathbf{Q}_\rho^* \boldsymbol{\Lambda}^{\frac{1}{2}}, \tag{7.1}
$$

em que $\boldsymbol{\Lambda} = \text{ diag } (\lambda_1^2, \ldots, \lambda_n^2)$ e λ_i^2 é a i-ésima entrada da diagonal de $\mathbf{Q}_\rho^{*-1}$ tornando ${\mathbf{Q}_\rho}^{-1}$ uma matriz de correlações válida.

Dessa forma, flexibiliza-se a modelagem. Por exemplo, o modelo convencional de *GLMM* utiliza função de ligação logarítmica, o que implica em modelar μ_i através de uma distribuição marginal $Log - Normal$. Este modelo pode ser recuperado utilizando a abordagem dos *TGMRF's*, basta tomar F_i como sendo a distribuição $Log - N(\mathbf{X_i^T}\boldsymbol{\beta}, \nu\lambda_i^2)$. Além disso, pode-se pensar em modelar μ_i através de outras distribuições, como por exemplo a distribuição Gama ou a distribuição Weibull, entre outras. Sendo assim, pode-se

escolher o modelo que melhor se ajusta aos dados não se limitando aos ajustes concedidos pelos modelos convencionais.

Para estimação dos parâmetros do modelo, propõe-se o uso do algoritmo *Gibbs Sampling*, tornando necessário que se conheça a distribuição condicional completa dos parâmetros do modelo. Seja $\boldsymbol{\varphi} = (\boldsymbol{\mu}, \boldsymbol{\beta}, \rho, \nu)$, tem-se que a distribuição *a posteriori* conjunta para o vetor $\boldsymbol{\varphi}$ é:

$$\pi\left(\boldsymbol{\varphi}|\mathbf{Y}, \mathbf{X}\right) \propto \left[\prod_{i=1}^{n} \frac{\exp\left\{-\mu_i\right\} \mu_i^{y_i}}{y_i!}\right] h\left(\boldsymbol{\mu}|\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_{\rho}\right) \left[\prod_{j=1}^{q} \pi(\beta_j)\right] \pi(\nu)\pi(\rho), \tag{7.2}$$

em que,

$$h\left(\boldsymbol{\mu}|\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_{\rho}\right) = (2\pi)^{-\frac{n}{2}} |\mathbf{Q}_{\rho}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^T \mathbf{Q}_{\rho} \boldsymbol{\epsilon}\right\} \prod_{i=1}^{n} \frac{f_i(\mu_i|\boldsymbol{\xi})}{\phi(\epsilon_i)}. \tag{7.3}$$

A Equação (7.3) é a densidade de $\boldsymbol{\mu}$ obtida através do método do Jacobiano para a mudança de variável $\epsilon_i = \Phi^{-1}\left\{F_i(\mu_i)\right\}$, sendo $\boldsymbol{\epsilon}$ um *GMRF*. Temos que f_i é a densidade associada a F_i, ϕ é a densidade de uma distribuição Normal Padrão. Sejam as seguintes distribuições a *priori*,

$\beta_j \sim N(0, \sigma_j^2),\ j = 1, \ldots, q,$

$\nu \sim Gama(a, b)\ I(\nu > 0),$

$\rho \sim U\left(\lambda_{min}^{-1}, \lambda_{max}^{-1}\right).$

As condicionais completas são obtidas a partir de (7.2). Considere $\pi\left(\theta|.\right)$ a distribuição de θ dado os dados e todos os demais parâmetros, então:

$$\pi(\mu_i|.) \propto \exp\left\{y_i \log\left(\mu_i\right) - \mu_i - \frac{1}{2}\boldsymbol{\epsilon}^T \mathbf{Q}_{\rho} \boldsymbol{\epsilon}\right\} \frac{f_i(\mu_i|\boldsymbol{\xi})}{\phi(\epsilon_i)}, i = 1, \ldots, n,$$

$$\pi(\beta_j|.) \propto \exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^T \mathbf{Q}_{\rho} \boldsymbol{\epsilon} - \frac{1}{2\sigma_j^2}\beta_j^2\right\} \prod_{i=1}^{n} \frac{f_i(\mu_i|\boldsymbol{\xi})}{\phi(\epsilon_i)}, j = 1, \ldots, q,$$

$$\pi(\nu|.) \propto \exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^T \mathbf{Q}_{\rho} \boldsymbol{\epsilon} + (a-1)\log(\nu) - b\nu\right\} \prod_{i=1}^{n} \frac{f_i(\mu_i|\boldsymbol{\xi})}{\phi(\epsilon_i)} I(\nu > 0)$$

e

$$\pi(\rho|.) \propto |\mathbf{Q}_{\rho}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^T \mathbf{Q}_{\rho} \boldsymbol{\epsilon}\right\} \prod_{i=1}^{n} \frac{f_i(\mu_i|\boldsymbol{\xi})}{\phi(\epsilon_i)} I\left(\rho \in \left(\lambda_{min}^{-1}, \lambda_{max}^{-1}\right)\right).$$

As distribuições condicionais completas não são densidades conhecidas. Prates et

al. (2015) propõe o uso do algoritmo Amostrador de *Gibbs* em que a cada iteração as condicionais completas são amostradas através do algoritmo *Adaptive Rejection Metropolis Sampling* (GILKS; BEST; TAN, 1995; **??**).

7.2 Regressão Poisson Espaço-Temporal

O aumento da capacidade de armazenamento de dados viabiliza, a cada dia mais, a avaliação de variáveis ao longo do tempo. Quando os dados possuem uma natureza de contagem e ainda assim são observados em diferentes regiões do espaço, pode-se analisá-los segundo uma abordagem espaço-temporal.

Sejam $(Y_{11}, \ldots, Y_{1t}, Y_{21}, \ldots, Y_{2t}, \ldots, Y_{nt})$, variáveis aleatórias observadas em n regiões de um determinado espaço em t tempos distintos, em que $Y_{ij}|\mu_{ij} \sim Poisson(\mu_{ij})$, $i = 1, \ldots, n$ e $j = 1, \ldots, t$. Seja $\mathbf{X}$ uma matriz de covariáveis, com dimensão $q \times (nt)$. Seja $\boldsymbol{\xi} = (\boldsymbol{\beta}, \nu, \mathbf{X}, \boldsymbol{\rho})$ e seja $\boldsymbol{\rho} = (\rho_s, \rho_t, \rho_{st})$. Os *TGMRF's* introduzem efeitos espaciais através da distribuição conjunta de μ, então tem-se:

$$
\begin{gathered}
Y_{ij}|\mu_{ij} \sim Poisson(\mu_{ij}), i = 1, \ldots, n; j = 1, \ldots, t \\
\boldsymbol{\mu} \sim TGMRF_{nt}\left(\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_{\boldsymbol{\rho}}\right) \\
\pi\left(\boldsymbol{\beta}\right); \pi\left(\nu\right), \pi\left(\boldsymbol{\rho}\right),
\end{gathered}
$$

em que $\boldsymbol{\beta}$ é um vetor de coeficientes de dimensão $q \times 1$, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_{1t}, \epsilon_{21}, \ldots, \epsilon_{2t}, \ldots, \epsilon_{nt})$ segue um modelo *MCAR* com matriz de precisão $\mathbf{Q}^*_{\boldsymbol{\rho}}$ e $\mathbf{F}_{\boldsymbol{\xi}} = (F_{11}, \ldots, F_{1t}, F_{21}, \ldots, F_{2t}, \ldots, F_{nt})$ sendo F_{ij} qualquer distribuição definida em $\mathbb{R}^+$.

Novamente, a inversa da matriz $\mathbf{Q}^*_{\boldsymbol{\rho}}$ não é uma matriz de correlações válida. Desta forma define-se,

$$
\mathbf{Q}_{\boldsymbol{\rho}} = \boldsymbol{\Lambda}^{\frac{1}{2}} \mathbf{Q}^*_{\boldsymbol{\rho}} \boldsymbol{\Lambda}^{\frac{1}{2}}, \tag{7.4}
$$

em que $\boldsymbol{\Lambda} = \text{ diag } (\lambda_1^2, \ldots, \lambda_{nt}^2)$ e λ_k^2 é a k-ésima entrada da diagonal de $\mathbf{Q}^{*^{-1}}_{\boldsymbol{\rho}}$. Assim $\mathbf{Q}^{-1}_{\boldsymbol{\rho}}$ é uma matriz de correlações válida.

A estimação dos parâmetros desconhecidos do modelo é realizada através do algoritmo Amostrador de *Gibbs*, exigindo-se a definição das distribuições condicionais completas dos parâmetros. Seja $\boldsymbol{\varphi} = (\boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\rho}, \nu)$, tem-se que a distribuição *a posteriori* conjunta para o vetor $\boldsymbol{\varphi}$ é dada por:

$$
\pi\left(\boldsymbol{\varphi}|\mathbf{Y}, \mathbf{X}\right) \propto \left[\prod_{i=1}^{n}\prod_{j=1}^{t} \frac{\exp\left\{-\mu_{ij}\right\} \mu_{ij}^{y_{ij}}}{y_{ij}!}\right] h\left(\boldsymbol{\mu}|\mathbf{F}_{\boldsymbol{\xi}}, \mathbf{Q}_{\boldsymbol{\rho}}\right) \left[\prod_{k=1}^{q} \pi(\beta_k)\right] \pi(\nu)\pi(\boldsymbol{\rho}), \tag{7.5}
$$

em que,

$$h\left(\boldsymbol{\mu}|\mathbf{F}_{\boldsymbol{\xi}},\mathbf{Q}_{\boldsymbol{\rho}}\right)=(2\pi)^{-\frac{n\times t}{2}}|\mathbf{Q}_{\boldsymbol{\rho}}|^{\frac{1}{2}}\exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^{T}\mathbf{Q}_{\boldsymbol{\rho}}\boldsymbol{\epsilon}\right\}\prod_{i=1}^{n}\prod_{j=1}^{t}\frac{f_{ij}(\mu_{ij}|\boldsymbol{\xi})}{\phi(\epsilon_{ij})}. \tag{7.6}$$

A Equação (7.6) é a densidade de $\boldsymbol{\mu}$ obtida através do método do Jacobiano para a mudança de variável $\epsilon_{ij}=\Phi^{-1}\left\{F_{ij}(\mu_{ij})\right\}$, sendo $\boldsymbol{\epsilon}$ um *GMRF*. Na Equação (7.6), f_{ij} é a densidade associada a F_{ij}, ϕ é a densidade de uma variável aleatória com distribuição Normal Padrão. Sejam as seguintes distribuições a *priori*,

$\beta_k \sim N(0,\sigma_k^2)$, para $k=1,\ldots,q$,

$\nu \sim Gama(a,b)\ I(\nu>0)$,

e

$\boldsymbol{\rho} \sim U_3\left(\Omega_{\boldsymbol{\rho}}\right)$,

em que $\Omega_{\boldsymbol{\rho}}=I_{\rho_s}\left(\min\rho_s,\max\rho_s\right)\times I_{\rho_t}\left(\min\rho_t,\max\rho_t\right)\times I_{\rho_{st}}\left(\min\rho_{st},\max\rho_{st}\right)$ denota os limites do domínio de $\boldsymbol{\rho}$ calculados através do método da dominância diagonal descrito na Seção 4.3. U_3 denota a distribuição Uniforme definida no espaço $\Omega_{\boldsymbol{\rho}}$. As condicionais completas, obtidas a partir da Equação (7.5), são

$$\pi(\mu_{ij}|.)\propto\exp\left\{y_{ij}\log\left(\mu_{ij}\right)-\mu_{ij}-\frac{1}{2}\boldsymbol{\epsilon}^{T}\mathbf{Q}_{\boldsymbol{\rho}}\boldsymbol{\epsilon}\right\}\frac{f_{ij}(\mu_{ij}|\boldsymbol{\xi})}{\phi(\epsilon_{ij})}I(\mu_{ij}>0),$$

$i=1,\ldots,n$ e $j=1,\ldots,t$,

$$\pi(\beta_k|.)\propto\exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^{T}\mathbf{Q}_{\boldsymbol{\rho}}\boldsymbol{\epsilon}-\frac{1}{2\sigma_k^2}\beta_k^2\right\}\prod_{i=1}^{n}\prod_{j=1}^{t}\frac{f_{ij}(\mu_{ij}|\boldsymbol{\xi})}{\phi(\epsilon_{ij})},k=1,\ldots,q,$$

$$\pi(\nu|.)\propto\exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^{T}\mathbf{Q}_{\boldsymbol{\rho}}\boldsymbol{\epsilon}+(a-1)\log(\nu)-b\nu\right\}\prod_{i=1}^{n}\prod_{j=1}^{t}\frac{f_{ij}(\mu_{ij}|\boldsymbol{\xi})}{\phi(\epsilon_{ij})},$$

e

$$\pi(\boldsymbol{\rho}|.)\propto|\mathbf{Q}_{\boldsymbol{\rho}}|^{\frac{1}{2}}\exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}^{T}\mathbf{Q}_{\boldsymbol{\rho}}\boldsymbol{\epsilon}\right\}\prod_{i=1}^{n}\prod_{j=1}^{t}\frac{f_{ij}(\mu_{ij}|\boldsymbol{\xi})}{\phi(\epsilon_{ij})}\Omega_{\boldsymbol{\rho}}.$$

Estas distribuições condicionais completas não são densidades conhecidas. Uma forma de amostrar a partir delas é através do algoritmo *ARMS*, assim como no caso da Regressão Espacial Poisson. Para o vetor $\boldsymbol{\rho}$ uma maneira mais interessante de amostrar das condicionais completas é através do algoritmo *Metropolis-Hastings* visto que a utilização do *ARMS* necessita da inversão da matriz $\mathbf{Q}$ muitas vezes e a amostragem em bloco se torna complexa. Além disso, como não se sabe o verdadeiro domínio destes parâmetros, Sain et al. (2011) propõem o uso de um passo de amostragem por rejeição. Nele, primeiramente

gera-se uma proposta para $\boldsymbol{\rho}$ e em seguida testa-se a validade das mesmas, ou seja, se a matriz $\boldsymbol{Q}$ gerada por essa proposta é inversível e portanto válida. Computacionalmente não é difícil gerar configurações válidas para o vetor $\boldsymbol{\rho}$. Em geral, necessita-se de menos de sete iterações até uma configuração adequada ser aceita sendo que a verificação da adequabilidade das propostas é feita de maneira rápida sem a necessidade da inversão da matriz $\boldsymbol{Q}$.

No Capítulo 8, uma breve descrição dos métodos de amostragem e dos modelos de ajuste utilizados neste trabalho.

8 Inferência

Na abordagem Bayesiana utiliza-se a informação proveniente dos dados em conjunto com informações *a priori*. O objetivo é combinar tais quantidades para obter informações a respeito dos parâmetros de interesse. A distribuição originada a partir da combinação da verossimilhança e da distribuição *a priori*, chamada distribuição *a posteriori*, na maioria das vezes não é analiticamente tratável. Sendo assim, muitos métodos numéricos e de amostragem foram desenvolvidos com o intuito de aproximar tais distribuições (GILKS; RICHARDSON; SPIEGELHALTER, 1996).

Uma grande quantidade de métodos de amostragem com o objetivo de obter uma amostra representativa da distribuição de interesse está disponível na literatura. As técnicas de amostragem utilizadas neste trabalho são descritas na Seção 8.1. Nas Seções 8.2 e 8.3 apresenta-se a estrutura de amostragem utilizada no trabalho e as distribuições utilizadas para modelagem das intensidades Poisson, respectivamente.

8.1 Métodos de amostragem

Cada método de amostragem tem suas particularidades, vantagens e desvantagens motivando seu uso em casos de interesse. Quatro métodos de amostragem são descritos nesta Seção.

Amostragem por rejeição:

A Amostragem por Rejeição tem como ideia básica obter uma amostra de uma variável aleatória θ cuja distribuição é denotada por $f(\theta)$ a qual conhecemos pelo menos o núcleo. Para isso, utiliza-se uma distribuição auxiliar denotada por $g(\theta)$ da qual é simples se obter uma amostra e cujo espaço paramétrico seja o mesmo de $f(\theta)$. Escolhida a distribuição auxiliar, encontra-se a constante M que multiplicada pela distribuição $g(\theta)$ encapsula a distribuição $f(\theta)$. Para mais detalhes veja por exemplo (ALBERT, 2009).

Seja M tal que: $\dfrac{f(\theta)}{g(\theta)} \leq M$, $\forall \theta \in \Theta$, seja $\theta^{(i)}$ a observação de $f(\theta)$ gerada na i-ésima iteração e seja n o tamanho da amostra que se deseja, o algoritmo da Amostragem por Rejeição é dado por:

Algoritmo 1 Amostragem por rejeição

1: **função** AMOSTRAGEM POR REJEIÇÃO($f(\theta)$, $g(\theta)$, M, n)
2: **enquanto** $i \leq n$ **faça**
3: Gere θ^* de $g(\theta)$
4: Gere $u \sim U(0,1)$
5: **se** $u \leq \frac{f(\theta^*)}{M*g(\theta^*)}$ **então**
6: Faça $\theta^{(i)} = \theta^*$
7: $i \leftarrow i + 1$
8: **c.c.**
9: Retorne para a linha 3.
10: **fim se**
11: **fim enquanto**
12: **retorna** $\boldsymbol{\theta}$
13: **fim função**

Com isso, uma amostra de tamanho n da distribuição $f(\theta)$ foi obtida de forma indireta através da distribuição auxiliar $g(\theta)$.

Amostrador de Gibbs:

O Amostrador de Gibbs (GELFAND; SMITH, 1990) faz parte de um grupo de métodos chamados de Métodos *MCMC* do inglês *Markov Chain Monte Carlo*. O objetivo do método é obter através de um processo iterativo uma amostra da distribuição *a posteriori* de um ou de um grupo de parâmetros de interesse $\boldsymbol{\theta}$. Para isso deve-se conhecer as distribuições condicionais completas de tais parâmetros e saber gerar valores aleatórios destas distribuições.

De acordo com a abordagem Bayesiana, a distribuição condicional completa é dada pela distribuição do parâmetro de interesse condicionado aos dados e aos demais parâmetros do modelo. Dentro do processo iterativo, o Amostrador de Gibbs trabalha com os valores correntes dos demais parâmetros visto que os verdadeiros valores dos parâmetros são desconhecidos. O método constrói uma cadeia irredutível e aperiódica em que todos seus estados são recorrentes positivos e então, essa cadeia tem como distribuição estacionária a distribuição *a posteriori*.

O processo iterativo inicia com a escolha dos valores iniciais das cadeias dos parâmetros e então, iterativamente, caminha em direção à distribuição limite através da geração de valores aleatórios das condicionais completas. Seja $\pi_i(\theta_i|\boldsymbol{\theta_{-i}})$ a distribuição condicional completa de θ_i, seja k a dimensão de $\boldsymbol{\theta}$, seja $\theta^{(i)}$ a observação de θ gerada na i-ésima iteração e seja n o tamanho de amostra desejado, o algoritmo do amostrador de Gibbs é dado por:

Algoritmo 2 Amostrador de Gibbs

função AMOSTRADOR DE GIBBS($\pi_i(\theta_i|\boldsymbol{\theta}_{-i})$, k, n)
  Defina valores iniciais para a cadeia: $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots \theta_k^{(0)})$
  para i $\leftarrow$ 1 **até** n **faça**
    Gere $\theta_1^{(i)}$ de $\pi_1(\theta_1|\theta_2^{(i-1)}, \theta_3^{(i-1)}, \dots, \theta_k^{(i-1)})$
    Gere $\theta_2^{(i)}$ de $\pi_2(\theta_2|\theta_1^{(i)}, \theta_3^{(i-1)} \dots, \theta_k^{(i-1)})$
    $\vdots$
    Gere $\theta_{k-1}^{(i)}$ de $\pi_{k-1}(\theta_{k-1}|\theta_1^{(i)}, \dots, \theta_{k-2}^{(i)}, \theta_k^{(i-1)})$
    Gere $\theta_k^{(i)}$ de $\pi_p(\theta_p|\theta^{(i)})_1, \theta_2^{(i)}, \dots, \theta_{k-1}^{(i)})$
  fim para
  retorna $\{\boldsymbol{\theta_1}, \dots, \boldsymbol{\theta_k}\}$
fim função

Na maioria dos casos as distribuições condicionais completas não são distribuições conhecidas e então não é possível utilizar apenas o Amostrador de Gibbs. Porém, é possível utilizar o Amostrador de Gibbs em conjunto com outras técnicas. Desta forma, as observações das distribuições condicionais completas desconhecidas são geradas a partir de algum outro método de amostragem como por exemplo o algoritmo *Metropolis-Hastings* ou o algoritmo *ARMS*.

Metropolis-Hastings:

O Algoritmo *Metropolis-Hastings* (HASTINGS, 1970) assim como o Amostrador de Gibbs é um método *MCMC*. Este método utiliza uma distribuição de referência aqui denotada por $H(\theta^{(i)}|\theta^{(i-1)})$ para gerar candidatos da distribuição de interesse denotada por $f(\theta)$.

É importante que a distribuição proposta, também chamada de densidade de transição seja conhecida e que se possa gerar facilmente dela. É interessante também que $H(\theta^{(i)}|\theta^{(i-1)})$ possua caudas pesadas para que todos os valores do domínio da distribuição de interesse possam ser gerados durante o processo iterativo. Desta forma melhores propostas serão geradas durante as iterações do método.

O algoritmo *Metropolis-Hastings* funciona de forma similar ao método da Amostragem por Rejeição. Em cada iteração gera-se uma proposta que pode ser aceita ou não. Porém, caso a proposta não seja aceita o algoritmo utiliza a última observação gerada para dar sequência ao processo. Neste ponto é necessário avaliar a taxa de aceitação do método, o desejado é uma taxa de aceitação nem muito alta nem muito baixa. Diversos trabalhos tentam encontrar uma taxa de aceitação ótima. Pórem, essa taxa varia de acordo com a complexidade do modelo (ROBERTS; ROSENTHAL et al., 2001).

Seja $\theta^{(i)}$ a observação de θ gerada na i-ésima iteração e seja n o tamanho de amostra desejado. O algoritmo *Metropolis-Hastings* é dado por:

Algoritmo 3 Metropolis-Hastings

função METROPOLIS-HASTINGS($f(\theta)$,$H(\theta^{(i)}|\theta^{(i-1)})$, n)
  Defina o valor inicial para a cadeia: $\theta^{(0)}$
  para i $\leftarrow$ 1 **até** n **faça**
    Gere θ^* de $H(\theta^{(i)}|\theta^{(i-1)})$
    Calcule o salto: $R = \dfrac{f(\theta^{(i-1)})H(\theta^*|\theta^{(i-1)})}{f(\theta^*)H(\theta^{(i-1)}|\theta^*)}$
    Faça $K = \min\{1, R\}$
    Gere $u \sim U(0,1)$
    se $u \leq K$ **então**
      $\theta^{(i)} = \theta^*$
    c.c.
      $\theta^{(i)} = \theta^{(i-1)}$.
    fim se
  fim para
  retorna $\boldsymbol{\theta}$
fim função

O algoritmo *Metropolis-Hastings* pode ser utilizado em conjunto com o amostrador de Gibbs. Ao invés da necessidade de amostrar da distribuição condicional completa exata é possível amostrar indiretamente da distribuição de interesse utilizando em cada passo uma iteração do *Metropolis-Hastings*.

Amostragem por rejeição adaptativa com passo *Metropolis-Hastings*:

O Método da Amostragem por rejeição adaptativa com passo *Metropolis-Hastings* ou então *ARMS* do inglês *Adaptive Rejection Metropolis Sampling* é uma generalização do método da Amostragem por Rejeição Adaptativa ou *ARS* do inglês *Adaptive Rejection Sampling* (GILKS; WILD, 1992).

A metodologia do *ARMS* necessita da compreensão do método da Amostragem por Rejeição, do algoritmo *Metropolis Hastings* e também do método *ARS*, que é um caso particular do método *ARMS* quando a função objetivo $f(\theta)$ é log-côncava.

A ideia do método é criar uma função linear por partes $h_m(\theta)$ que se adapte à função de interesse. Essa função é utilizada para gerar propostas para o parâmetro de interesse. Em um primeiro momento compara-se a função $h_m(\theta)$ com a função $f(\theta)$ e caso a distância entre as curvas seja grande então o valor gerado possivelmente é rejeitado nesta etapa, caso contrário algumas quantidades são atualizadas e o valor gerado é novamente testado, agora em um passo *Metropolis-Hastings*.

Seja $g_m(\theta) = \dfrac{h_m(\theta)}{\int h_m(\theta)d\theta}$, m número de partes da função h_m, $S_m = \{\theta_i, i = 0, \ldots, m+1\}$ ordenado, θ_{atual} o último valor gerado para θ, seja $\theta^{(i)}$ a observação de θ gerada na i-ésima iteração e n o tamanho de amostra desejado, o algoritmo do *ARMS* é dado por:

Algoritmo 4 Adaptive Rejection Metropolis Sampling

função ARMS($f(\theta)$, n)
  Defina m, S_m, $h_m(\theta)$ e $g_m(\theta)$
  para i $\leftarrow$ 1 **até** n **faça**
    Gere θ^* de $g_m(\theta)$
    Gere $u \sim U(0,1)$
    se u $> \frac{f(\theta^*)}{\exp h_m(\theta^*)}$ **então**
      $S_{m+1} = S_m \bigcup \theta^*$
      Reordene S_{m+1}
      Faça $m = m + 1$
      Retorne para 4
    c.c.
      Faça $\theta^A = \theta^*$
    fim se
    Gere $u \sim U(0,1)$
    se $u > \min\left(1, \frac{f(\theta^A)\min\{f(\theta_{atual}), \exp h_m(\theta_{atual})\}}{f(\theta_{atual})\min\{f(\theta^A), \exp h_m(\theta^A)\}}\right)$ **então**
      Faça $\theta^{(i)} = \theta_{atual}$
    c.c.
      Faça $\theta^{(i)} = \theta^A$
    fim se
  fim para
  retorna $\boldsymbol{\theta}$
fim função

Assim como o algoritmo *Metropolis-Hastings*, o *ARMS* pode ser utilizado em conjunto com o Amostrador de *Gibbs* nos casos em que não se conhece a distribuição condicional completa dos parâmetros. Neste caso, o grupo de abscissas S_n pode ser escolhido de forma otimizada com base no conjunto de abscissas geradas na iteração anterior, diminuindo o número de avaliações da função $f(\theta)$.

8.2 Modelo de amostragem adotado

No Capítulo 6 a metodologia dos *TGMRF's* foi generalizada para o contexto espaço-temporal. O presente trabalho irá focar apenas na aplicação da teoria para dados *Poisson*.

Seja $\mathbf{Y} = (Y_{11}, \ldots, Y_{1t}, Y_{21}, \ldots, Y_{2t}, \ldots, Y_{nt})$, variáveis aleatórias observadas em n regiões de um determinado espaço em t tempos distintos, em que $Y_{ij}|\mu_{ij} \sim Poisson(\mu_{ij})$, $\forall i = 1, \ldots, n$ e $j = 1, \ldots, t$. Seja $\mathbf{Q}_{\boldsymbol{\rho}}$ a matriz definida em (7.4). Seja $\boldsymbol{\beta}$ um vetor de coeficientes de dimensão q, associados às covariáveis do modelo. Seja $\theta^{(i)}$ a observação de θ gerada na i-ésima iteração, seja $\boldsymbol{H}_{\boldsymbol{\eta}}(\Omega_{\boldsymbol{\eta}})$ uma distribuição que gere propostas para um vetor de parâmetros $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)$ em um espaço $\Omega_{\boldsymbol{\eta}} = I_{\eta_1}(\min \eta_1, \max \eta_1) \times I_{\eta_2}(\min \eta_2, \max \eta_2) \times I_{\eta_3}(\min \eta_3, \max \eta_3)$ e seja N o número de observações *a posteriori* desejado. O modelo de amostragem adotado é dado por:

Algoritmo 5 *TGMRF's* na Regressão Espaço-Temporal Poisson

1: Faça $\Omega_{\boldsymbol{\rho}} = I_{\rho_s}(\min \rho_s, \max \rho_s) \times I_{\rho_t}(\min \rho_t, \max \rho_t) \times I_{\rho_{st}}(\min \rho_{st}, \max \rho_{st})$
2: Defina valores iniciais para as cadeias:
$\boldsymbol{\beta^{(0)}} = \left\{\beta_0^{(0)}, \ldots, \beta_{q-1}^{(0)}\right\}$
$\boldsymbol{\mu^{(0)}} = \left\{\mu_{11}^{(0)}, \ldots, \mu_{1t}^{(0)}, \mu_{21}^{(0)}, \mu_{2t}^{(0)}, \ldots, \mu_{nt}^{(0)}\right\}$
$\boldsymbol{\rho^{(0)}} = \left\{\rho_s^{(0)}, \rho_t^{(0)}, \rho_{st}^{(0)}\right\}$
$\nu^{(0)}$
3: **para** i ← 1 **até** N **faça**
4: Defina Info = Falso
5: **enquanto** Info ≠ Verdadeiro **faça**
6: Gere $\boldsymbol{\rho}^*$ de $\boldsymbol{H_\rho}(\Omega_{\boldsymbol{\rho}})$
7: **se** $\mathbf{Q}_{\boldsymbol{\rho}^*}$ é positiva definida **então**
8: Info = Verdadeiro
9: **c.c.**
10: Info = Falso
11: **fim se**
12: **fim enquanto**
13: Faça $\boldsymbol{\rho^{(i)}}$ = METROPOLIS-HASTINGS($\pi(\boldsymbol{\rho}|.), \boldsymbol{H_\rho}(\Omega_{\boldsymbol{\rho}}), 1$)
14: **para** j ← 0 **até** q-1 **faça**
15: Faça $\beta_j^{(i)}$ = ARMS($\pi(\beta_j|.), 1$)
16: **fim para**
17: **para** k ← 1 **até** n **faça**
18: **para** l ← 1 **até** t **faça**
19: Faça $\mu_{kl}^{(i)}$ = ARMS($\pi(\mu_{kl}|.), 1$)
20: **fim para**
21: **fim para**
22: Toma $\nu^{(i)}$ = ARMS($\pi(\nu|.), 1$)
23: **fim para**
24: **retorna** $(\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\rho}, \nu)$

Utilizou-se como distribuição de proposta $\boldsymbol{H_\rho}\left(\Omega_{\rho}\right)$ a Normal Truncada no espaço Ω_{ρ} considerando independência entre os parâmetros.

Apesar do método *ARMS* ser robsuto para amostrar no intervalo $[-\infty, \infty]$ uma transformação foi aplicada para amostrar no domínio $[0, 1]$. Essa transformação impõe ao método *ARMS* um limite superior e um inferior bem definidos para amostragem o que beneficia o algoritmo.

8.3 Modelos de ajuste

Dada a flexibilidade dos *TGMRF's* é possível modelar o parâmetro de interesse com distribuições que respeitem seu domínio. No presente trabalho, quatro distribuições diferentes foram utilizadas para modelar as intensidades Poisson. A notação $\boldsymbol{X_{ij}}$ se refere ao conjunto de covariáveis relacionadas à região i no tempo j.

***Gamma Independent* - GI:**

Neste modelo a variância de μ_{ij} é dada por ν e a média é dada por $\exp(\boldsymbol{X_{ij}\beta})$, ou seja, os parâmetros de média e variância são independentes.

$$\mu_{ij} \sim Gama\left(\frac{\exp(\boldsymbol{X_{ij}\beta})^2}{\nu}, \frac{\exp(\boldsymbol{X_{ij}\beta})}{\nu}\right)$$

***Gamma Scale* - GSC:**

O nome deste modelo se dá pelo fato da média estar no parâmetro de escala da distribuição Gama. Além disso, ν funciona como um parâmetro de dispersão sendo que a variância é uma função da média e de ν.

$$\mu_{ij} \sim Gama\left(\frac{1}{\nu}, \frac{1}{\nu \exp(\boldsymbol{X_{ij}\beta})}\right)$$

***Gamma Shape* - GSH:**

Neste modelo tem-se a média no parâmetro de forma da distribuição Gama. Além disso, ν funciona como um parâmetro de dispersão. Novamente, a variância é uma função da média e de ν.

$$\mu_{ij} \sim Gama\left(\frac{\exp(\boldsymbol{X_{ij}\beta})}{\nu}, \frac{1}{\nu}\right)$$

***Log-Normal* - LN:**

Caso a modelagem fosse feita a partir de um *GLMM* com função de ligação log, então esta seria a distribuição modelando μ_{ij}. Neste caso, o modelo pode sofrer com o problema do confundimento espacial.

$$\mu_{ij} \sim LN\left(\boldsymbol{X_{ij}\beta}, \sqrt{\nu \boldsymbol{Q^{-1}_{ij,ij}}}\right)$$

Uma alternativa a esta modelagem seria utilizar a distribuição Log-normal em que os parâmetros não dependessem de $\boldsymbol{Q^{-1}}$

A parametrização da distribuição Gama utilizada é tal que:

$$f(y;\alpha,\beta) = \frac{\beta^\alpha y^{\alpha-1}\exp\{\beta y\}}{\Gamma(\alpha)},\ \alpha > 0 \text{ e } \beta > 0$$
$$E(y) = \tfrac{\alpha}{\beta}; V(y) = \tfrac{\alpha}{\beta^2}$$

A parametrização da distribuição Log-Normal utilizada é tal que:

$$f(y;\mu,\sigma) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{\frac{-\left(\log\left(y\right)-\mu\right)^2}{2\sigma^2}\right\}$$
$$E(y) = \exp\left\{\mu + 0.5\sigma^2\right\};\ \mathrm{V(y)} = \exp\left\{2\mu+\sigma^2\right\}\left(\exp\left\{\sigma^2\right\}-1\right)$$

A fim de comparação na Tabela 1 tem-se as médias e variâncias dos modelos empregados. Pode-se perceber que as médias adotadas para os modelos que utilizam a distribuição Gama são iguais diferindo apenas na variância.

Modelo	$E\left(\mu_{ij}\right)$	$\mathrm{V}(\mu_{ij})$
GI	$\exp\left\{\boldsymbol{X_{ij}\beta}\right\}$	ν
GSC	$\exp\left\{\boldsymbol{X_{ij}\beta}\right\}$	$\nu\exp\left\{\boldsymbol{X_{ij}\beta}\right\}^2$
GSH	$\exp\left\{\boldsymbol{X_{ij}\beta}\right\}$	$\nu\exp\left\{\boldsymbol{X_{ij}\beta}\right\}$
LN	$\exp\left\{\boldsymbol{X_{ij}\beta} + 0.5\nu\boldsymbol{Q^{-1}_{ij,ij}}\right\}$	$\exp\left\{2\boldsymbol{X_{ij}\beta} + \boldsymbol{\nu Q^{-1}_{ij,ij}}\right\}\left(\exp\left\{\nu\boldsymbol{Q^{-1}_{ij,ij}}\right\}-1\right)$

Tabela 1 – Esperança e variância dos modelos utilizados.

Para exemplificar a diferença entre estas distribuições, gerou-se um conjunto de dados a partir do modelo LN. Em seguida, ajustou-se os quatro modelos (GSC, GSH, GI, LN) a este conjunto de dados. A Figura 10 traz o ajuste dos quatro modelos para uma média de interesse (μ_{22}).

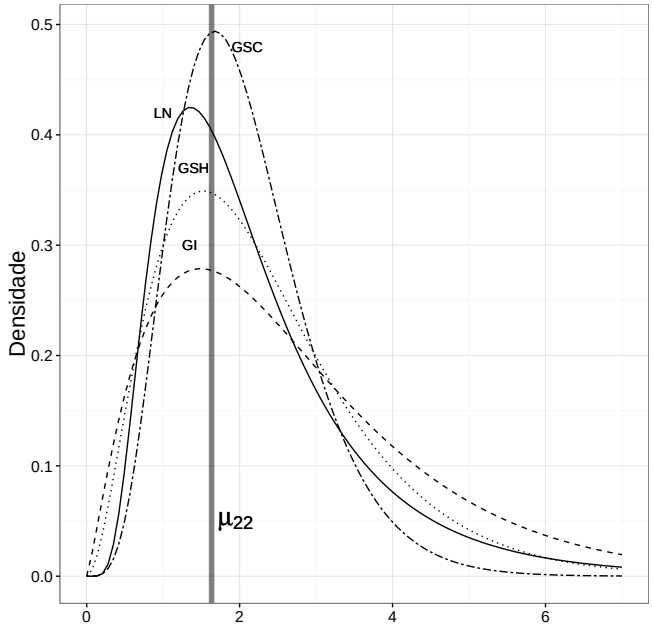


Figura 10 – Curvas dos modelos empregados para um exemplo simulado a partir do modelo LN.

Pode-se ver na Figura 10 que as curvas dos modelos adotados diferem uma das outras abrindo várias possibilidades de modelagens. Dado a disponibilidade de uma grande quantidade de modelos, medidas de ajuste se fazem necessárias para que se possa decidir pelo melhor modelo. Na Seção 9 os critérios de ajuste adotados são apresentados.

9 Seleção de modelos

Critérios de ajuste são muito utilizados em estatística. O objetivo destes critérios é, dado a disponibilidade de uma variedade de modelos, dar indícios ao pesquisador de qual dos modelos é o que melhor se ajusta aos dados. Por exemplo, em modelos de regressão quando existem muitas covariáveis, uma grande variedade de modelos pode ser ajustada. Para decidir sobre qual conjunto de covariáveis provê o melhor ajuste é possível observar qual gera o maior R^2, ou então penalizar o modelo de acordo com o número de covariáveis e usar o R^2 ajustado ou o AIC. As medidas R^2, R^2 ajustado e AIC são apenas algumas medidas de ajuste disponíveis na literatura.

Com o passar dos anos muitos critérios de ajuste foram criados mas ainda hoje ressalva-se que não existe um critério ótimo, ou seja, aquele que escolhe o modelo correto em todos os casos (GELMAN; HWANG; VEHTARI, 2014). No presente trabalho, alguns critérios serão testados em um estudo de simulação. O objetivo é identificar dentre quatro métodos qual é o mais eficaz no que diz respeito a escolher o modelo correto mais vezes. Para mais detalhes sobre os critérios de ajuste adotados neste trabalho veja por exemplo (GELMAN; HWANG; VEHTARI, 2014) e (CHRISTENSEN et al., 2011).

9.1 *Deviance Information Criterion - DIC:*

O método DIC (SPIEGELHALTER et al., 2002) é um dos critérios de ajuste mais utilizados dentro da estatística Bayesiana. Este critério considera em seu cálculo duas quantidades: Uma penaliza o modelo de acordo com a sua complexidade e a outra é responsável por evidenciar um bom ajuste do modelo aos dados.

Seja $\hat{\theta}_{Bayes}$ o estimador de Bayes para θ e seja $\log\left(y|\hat{\theta}_{Bayes}\right)$ o logaritmo da verossimilhança aplicada no estimador de Bayes, a medida DIC é dada por:

$$DIC = -2\log\left(y|\hat{\theta}_{Bayes}\right) + 2p_{DIC}, \tag{9.1}$$

em que a quantidade p_{DIC} é chamada de número efetivo de parâmetros e pode ser calculada de duas formas,

$$p_{DIC_1} = 2\left[\log\left(y|\hat{\theta}_{Bayes}\right) - E_{post}\left(\log\left(y|\theta\right)\right)\right], \tag{9.2}$$

e

$$p_{DIC_2} = 2\left[Var_{post}\left(\log\left(y|\theta\right)\right)\right]. \tag{9.3}$$

Prefere-se o modelo com menor *DIC* visto que quanto maior o logaritmo da verossimilhança em (9.1) menor o *DIC*. Caso o número efetivo de parâmetros seja muito grande, então, o valor da estatística cresce penalizando o ajuste. Desta forma o modelo com menor *DIC* tende a ser um modelo mais parcimonioso.

Assintoticamente as quantidades p_{DIC_1} e p_{DIC_2} retornam o número de parâmetros efetivos correto (GELMAN; HWANG; VEHTARI, 2014). A versão do número efetivo de parâmetros utilizada neste trabalho é a dada em (9.2) que não é analiticamente calculável para a maioria dos casos. Na equação (9.4) tem-se uma maneira de calcular o número de parâmetros efetivos com base em uma amostra da distribuição *a posteriori* de θ.

$$p_{DIC} = 2\left[\log\left(y|\hat{\theta}_{Bayes}\right) - \frac{1}{N}\sum_{j=1}^{N}\log\left(p\left(y|\theta^{(j)}\right)\right)\right], \tag{9.4}$$

em que N é o tamanho da amostra *a posteriori* e $\theta^{(j)}$ é o j-ésimo valor obtido da distribuição *a posteriori* de θ.

9.2 *Watanabe-Akaike Information Criterion - WAIC:*

O critério *WAIC* (WATANABE, 2010) utiliza as mesmas ideias do critério *DIC* em que duas quantidades mensuram o ajuste e a complexidade do modelo respectivamente. Seja $p_{post}(\theta) = p(\theta|y)$, a medida *WAIC* é dada por:

$$WAIC = -2\sum_{i=1}^{n}\log\left(\int p(y_i|\theta)\, p_{post}(\theta)\, d\theta\right) + 2p_{WAIC}. \tag{9.5}$$

Em que p_{WAIC} é o número efetivo de parâmetros e pode ser calculado de duas formas:

$$p_{WAIC_1} = 2\sum_{i=1}^{n}\left[\log\left(E_{post}p(y_i|\theta)\right) - E_{post}\left(\log\left(p(y_i|\theta)\right)\right)\right], \tag{9.6}$$

e

$$p_{WAIC_2} = \sum_{i=1}^{n} Var_{post}\left[\log\left(p(y_i|\theta)\right)\right]. \tag{9.7}$$

Valores pequenos indicam modelos mais parcimoniosos. Neste trabalho, as duas versões da medida *WAIC* foram utilizadas. Quando a forma de cálculo do número de parâmetros efetivos utilizados for a da Equação (9.6) a medida de ajuste será chamada de $WAIC_1$, quando utilizada a Equação (9.7) a medida será chamada de $WAIC_2$.

Na prática algumas das quantidades utilizadas para calcular $WAIC_1$ e $WAIC_2$ não são obtidas de forma trivial visto que na maioria dos casos não se conhece analiticamente

a distribuição *a posteriori* dos parâmetros de interesse. Porém, dada uma amostra de tamanho N da distribuição *a posteriori* de θ é possível calcular aproximações das medidas de ajuste computacionalmente (GELMAN; HWANG; VEHTARI, 2014).

$$WAIC = -2\sum_{i=1}^{n} \log\left(\frac{1}{N}\sum_{j=1}^{N} p\left(y_i|\theta^{(j)}\right)\right) + 2p_{WAIC}, \tag{9.8}$$

$$p_{WAIC_1} = 2\sum_{i=1}^{n}\left[\log\left(\frac{1}{N}\sum_{j=1}^{N} p\left(y_i|\theta^{(j)}\right)\right) - \frac{1}{N}\sum_{j=1}^{N}\log\left(p\left(y_i|\theta^{(j)}\right)\right)\right] \tag{9.9}$$

ou

$$p_{WAIC_2} = \sum_{i=1}^{n}\frac{1}{N-1}\sum_{j=1}^{N}\left[\log\left(p\left(y_i|\theta^{(j)}\right)\right) - \frac{1}{N}\sum_{j=1}^{N}\log\left(p\left(y_i|\theta^{(j)}\right)\right)\right]^2, \tag{9.10}$$

em que $\theta^{(j)}$ é o j-ésimo valor obtido da distribuição *a posteriori* de θ.

9.3 *Logarithm of the Pseudo-Marginal Likelihood - LPML:*

Outro possível critério de seleção de modelos é o *LPML* (GEISSER; EDDY, 1979), o qual se diferencia dos antes descritos. Seja $\boldsymbol{Y} = Y_1, \ldots, Y_n$ um vetor aleatório condicionalmente independente dado θ, seja M o modelo ajustado aos dados e seja $f(\boldsymbol{Y}|\theta, M)$ a verossimilhança associada ao modelo. A ideia do método é aproximar a distribuição $f(\boldsymbol{Y}|M)$, o que pode ser analiticamente difícil/impossível.

A distribuição $f(\boldsymbol{Y}|M) = \int f(\boldsymbol{Y}|\theta, M)\,\pi(\theta|M)\,d\theta$ mede o ajuste do modelo adotado aos dados e é utilizada para calcular o Fator de Bayes, por exemplo. Para aproximar tal distribuição faz-se:

$$\hat{f}(\boldsymbol{Y}|M) = \prod_{i=1}^{n} f_i(Y_i|Y_{-i}, M) = \prod_{i=1}^{n}\int_{\theta\in\Theta} f_i(Y_i|\theta, M)\,\pi(\theta|Y_{-i}, M)d\theta. \tag{9.11}$$

A distribuição $f_i(Y_i|Y_{-i}, M)$ vista na Equação (9.11) é chamada de Distribuição Condicional Preditiva Ordenada, do inglês *Conditional Predictive Ordinate Distribution (CPO)*. Então, a medida *LPML* é dada por:

$$LPML = \sum_{i=1}^{n}\log(CPO_i). \tag{9.12}$$

Ao contrário das demais medidas, opta-se pelo modelo com maior *LPML* por serem consideradas de melhor ajuste aos dados. Gelfand e Dey (1994) mostram como aproximar

a quantidade CPO_i que é utilizada para calcular a medida $LPML$ através de uma amostra *a posteriori* de θ:

$$LPML = -\sum_{i=1}^{n} \log\left(\frac{1}{N}\sum_{j=1}^{N}\frac{1}{f_i\left(y_i|\theta^{(j)}, M\right)}\right), \tag{9.13}$$

em que N é o tamanho da amostra e $\theta^{(j)}$ é o j-ésimo valor da obtido da distribuição *a posteriori* de θ.

10 Estudo de simulação

Este Capítulo tem como objetivos apresentar alguns resultados dos estudos de simulação realizados além de comparar a eficiência dos critérios de seleção de modelos.

Além do algoritmo descrito na Seção 8.2, implementou-se a opção de observar o comportamento das covariáveis em cada tempo em análise. Seja $\boldsymbol{X}$ uma covariável observada em todos os t tempos, é razoável investigar se o efeito das covariáveis é alterado ao longo do tempo. Desta forma, divide-se a covariável $\boldsymbol{X}$ em t novas covariáveis, ou seja, $\boldsymbol{X} = \boldsymbol{X_1}, \ldots, \boldsymbol{X_t}$ em que $\boldsymbol{X_i}$ é composta de zeros nas entradas correspondentes aos tempos diferentes do tempo i. A ideia é observar se uma dada variável mantém a sua influência na média ao longo do tempo ou se essa possui algum padrão de crescimento/decrescimento. Caso a influência não se mantenha, então realizar a análise sem esse passo introduz viés nas estimativas dos coeficientes. Durante a análise esta metodologia será chamada de Análise com Coeficientes Temporais, visto que para cada covariável em cada tempo existirá um coeficiente associado. A análise típica será chamada de Análise com Coeficientes Atemporais.

Na Figura 11 tem-se o diagrama explicitando cenários para a análise com coeficientes atemporais. Os resultados apresentados nesta seção encontram-se destacados em cinza.

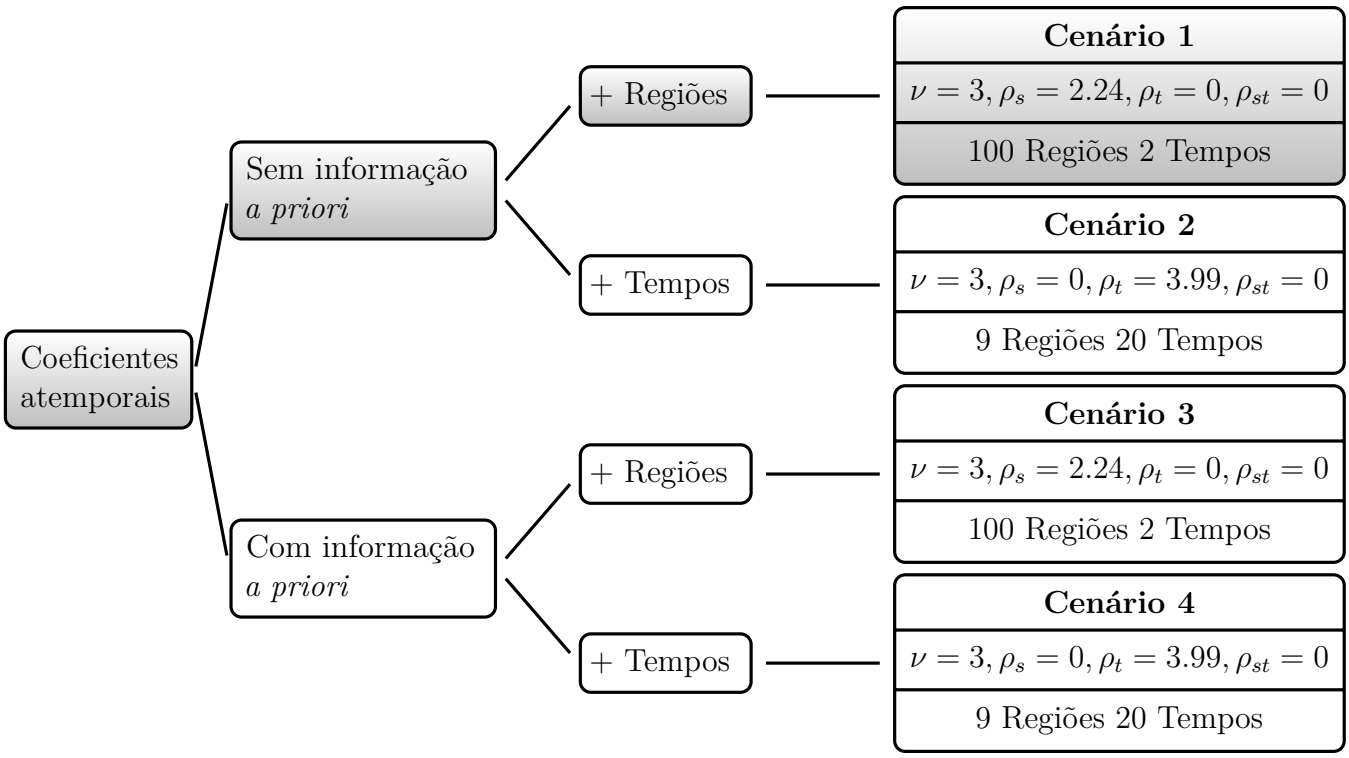


Figura 11 – Cenários de simulação com coeficientes atemporais.

O segundo nível da árvore diz respeito a informação *a priori* inserida para estimação do parâmetro de dependência em estudo. Quando o estudo é realizado em um número grande de regiões (+ Regiões) o parâmetro de interesse é o de dependência espacial ρ_s. Quando o estudo é realizado em um número grande de tempos (+ Tempos) o parâmetro de interesse é o de dependência temporal ρ_t. O objetivo principal é identificar se o método é capaz de captar quando um determinado conjunto de dados possui alta dependência espacial/temporal. Além disso, deseja-se verificar se os demais parâmetros são bem estimados dada a configuração adotada.

A Figura 12 apresenta o diagrama com os estudos realizados para a análise com coeficientes temporais. Para a análise com coeficientes temporais menos cenários foram gerados pois acredita-se que os resultados são similares aos encontrados. Novamente os resultados apresentados nesta seção encontram-se destacados em cinza.

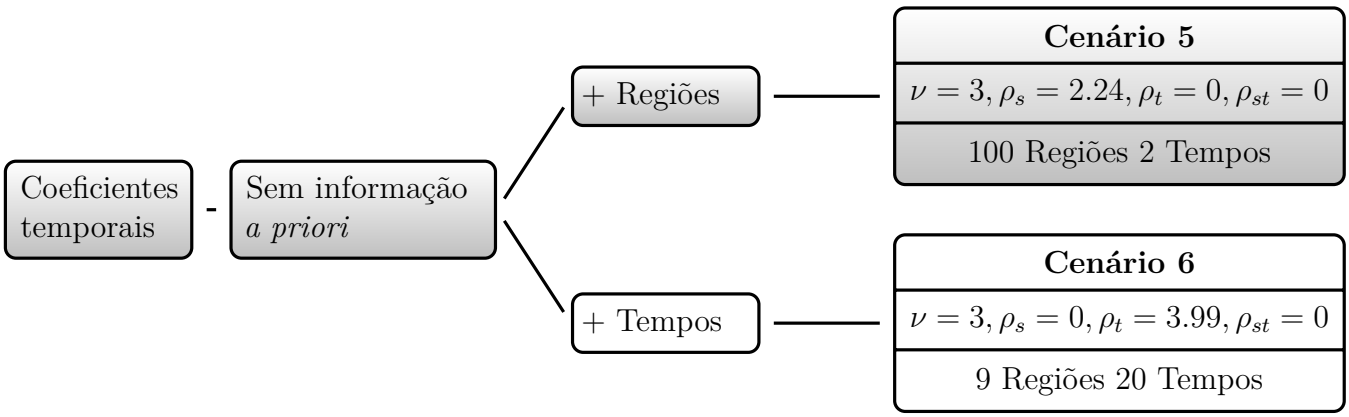


Figura 12 – Cenários de simulação com coeficientes temporais.

Para cada um dos cenários gerou-se 100 bases de dados utilizando como estruturas geradoras os quatro modelos descritos na Seção 8.3, totalizando 400 bases de dados por cenário. Para cada uma das bases utilizou-se os quatro modelos para ajuste.

Desta forma, gerou-se um total de 2400 bases de dados e ajustou-se um total de 9600 modelos. Para cada um destes modelos uma cadeia *MCMC* de tamanho 30000 foi utilizada seguindo o algoritmo descrito na Seção 8.2. Tomou-se como período de aquecimento da cadeia as primeiras 20000 observações e com as 10000 restantes selecionou-se amostras com $lag = 10$ visando diminuir a autocorrelação das cadeias. Desta forma, a amostra *a posteriori* tem tamanho 1000.

Para análise de convergência, utilizou-se o método de diagnóstico de *Geweke* (GEWEKE et al., 1991). A ideia do método é comparar duas partes da cadeia considerando independência entre elas e testar a hipótese de igualdade entre as médias. Os modelos em que pelo menos uma das cadeias não convergiu de acordo com o critério de diagnóstico de *Geweke* foram descartadas das análises.

Cenário 1:

Para este cenário fixou-se o parâmetro $\nu = 3$ para todos os modelos (exceto o modelo GSC) e por hora não será alvo de estudo. Para o modelo GSC reparametrizou-se o parâmetro $\nu = 1/3$. Perceba que este parâmetro tem diferentes interpretações em cada um dos modelos descritos na Seção 8.3, desta forma, os modelos devem ter estimativas diferentes para este parâmetro. As regiões de estudo foram dispostas tal como um látice 10x10 totalizando 100 regiões. Além disso, gerou-se observações para as 100 regiões em dois tempos diferentes. Dado o número de regiões, foram geradas as observações com um grande padrão espacial para observar se o método é capaz de recuperar bem este parâmetro. Para esta análise os limites calculados para os parâmetros de dependência foram, $\rho_s \in [-2.25, 2.25]$; $\rho_t \in [-5, 5]$; $\rho_{st} \in [-2.25, 2.25]$, e então definiu-se $\rho_s = 2.24$, $\rho_t = 0$ e $\rho_{st} = 0$ garantindo inversibilidade da matriz de precisão adotada. Além disso utilizou-se duas covariáveis e o intercepto.

Na Tabela 2 pode-se observar os resultados obtidos para este experimento. A coluna "Cob" se refere a porcentagem das vezes em que o intervalo *HPD* de 0.95 de probabilidade conteve o verdadeiro valor do parâmetro. A coluna "NB (%)" refere-se ao percentual de bases de dados dentre as 100 geradas em que todos os parâmetros convergiram segundo o critério de diagnóstico de *Geweke*.

Percebe-se na Tabela 2 que quando o modelo verdadeiro é o LN as estimativas pontuais do vetor $\boldsymbol{\beta}$ dos demais modelos difere bastante do verdadeiro parâmetro. Isso se deve ao fato de que a média do modelo LN é uma função da média dos demais modelos como visto na Tabela 1. Desta forma não é esperado que os outros modelos estimem corretamente o verdadeiro valor de $\boldsymbol{\beta}$. Pode-se observar também que em geral as estimativas pontuais de $\boldsymbol{\beta}$, tendem a ser próximas do verdadeiro valor dos parâmetros quando o modelo de ajuste é o modelo verdadeiro. Apesar disso, a taxa de cobertura dos intervalos *HPD* está abaixo da taxa nominal (0.95) o que pode ser indício de cadeias muito concentradas. Este resultado não é esperado, futuramente testar outro método de amostragem pode ser interessante para verificar se há melhorias na taxa de cobertura.

Percebe-se uma dificuldade de estimação do parâmetro ν por parte do modelo LN. As taxas de cobertura dos intervalos *HPD* para ν também encontram-se abaixo da taxa nominal de 0.95 para todos os modelos.

O parâmetro de dependência espacial não foi bem estimado. Aparentemente o modelo subestima o verdadeiro valor do parâmetro. Um complicador na análise é a relação entre os parâmetros de dependência. Conjuntamente os parâmetros de dependência devem garantir inversibilidade da matriz de precisão. Desta forma, valores altos (próximos ao limite superior definido pelo critério da dominância diagonal) podem ser alcançados somente se os demais parâmetros forem fixados em 0 o que não é interessante em uma análise espaço-temporal.

Modelo Verdadeiro	Parâmetro	Valor verdadeiro	Modelo de ajuste															
			LN				GI				GSH				GSC			
			Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %
LN	β_0	1.00	1.06	0.17	0.58	83	1.25	0.13	0.49	80	1.58	0.15	0.25	83	1.37	0.18	0.53	76
LN	β_1	0.70	0.69	0.07	0.99		0.40	0.07	0.12		0.53	0.06	0.24		0.67	0.08	0.95	
LN	β_2	0.30	0.31	0.18	0.96		0.20	0.13	0.70		0.24	0.15	0.92		0.34	0.18	0.93	
LN	ν	3.00	5.77	0.86	0.08		9.73	0.22	0.00		6.36	0.79	0.22		0.89	0.13	0.00	
LN	ρ_s	2.24	0.89	0.51	0.00		0.77	0.49	0.00		0.84	0.48	0.00		0.85	0.48	0.00	
LN	ρ_{st}	0.00	-0.13	0.59	1.00		-0.03	0.54	1.00		-0.13	0.56	1.00		-0.10	0.56	1.00	
LN	ρ_t	0.00	-0.09	1.43	1.00		0.01	1.27	1.00		-0.11	1.29	1.00		-0.04	1.30	1.00	
GI	β_0	1.00	0.86	0.12	0.59	75	0.95	0.11	0.68	78	0.91	0.12	0.64	75	0.93	0.12	0.66	79
GI	β_1	0.70	0.74	0.05	0.60		0.71	0.04	0.60		0.74	0.05	0.56		0.74	0.05	0.52	
GI	β_2	0.30	0.34	0.13	0.89		0.33	0.11	0.91		0.34	0.13	0.95		0.32	0.13	0.95	
GI	ν	3.00	1.18	0.38	0.08		2.90	0.87	0.77		0.83	0.22	0.00		0.19	0.05	0.00	
GI	ρ_s	2.24	0.35	0.58	0.00		0.42	0.56	0.00		0.42	0.56	0.00		0.39	0.55	0.00	
GI	ρ_{st}	0.00	-0.11	0.58	1.00		-0.11	0.57	1.00		-0.13	0.57	1.00		-0.10	0.57	1.00	
GI	ρ_t	0.00	-0.04	1.41	1.00		-0.03	1.36	1.00		-0.04	1.35	1.00		-0.06	1.38	1.00	
GSH	β_0	1.00	0.56	0.18	0.38	78	0.99	0.13	0.60	70	0.88	0.16	0.54	81	0.87	0.17	0.65	84
GSH	β_1	0.70	0.83	0.08	0.54		0.64	0.05	0.64		0.74	0.06	0.77		0.76	0.08	0.74	
GSH	β_2	0.30	0.35	0.19	0.95		0.25	0.12	0.93		0.32	0.16	0.94		0.31	0.18	0.92	
GSH	ν	3.00	4.46	0.87	0.62		8.13	0.99	0.10		2.61	0.50	0.73		0.68	0.12	0.00	
GSH	ρ_s	2.24	0.82	0.52	0.00		0.78	0.49	0.00		0.81	0.49	0.00		0.79	0.50	0.00	
GSH	ρ_{st}	0.00	-0.12	0.57	1.00		-0.09	0.55	1.00		-0.09	0.56	1.00		-0.10	0.55	1.00	
GSH	ρ_t	0.00	-0.02	1.35	1.00		-0.01	1.30	1.00		-0.02	1.31	1.00		-0.01	1.30	1.00	
GSC	β_0	1.00	0.82	0.13	0.63	92	1.08	0.12	0.60	82	1.00	0.13	0.78	77	0.94	0.13	0.74	86
GSC	β_1	0.70	0.70	0.06	0.95		0.60	0.05	0.43		0.66	0.05	0.75		0.70	0.06	0.95	
GSC	β_2	0.30	0.31	0.14	0.92		0.26	0.11	0.85		0.29	0.13	0.91		0.30	0.14	0.92	
GSC	ν	0.33	1.99	0.46	0.00		4.86	1.26	0.13		1.34	0.30	0.09		0.30	0.06	0.73	
GSC	ρ_s	2.24	0.56	0.56	0.00		0.55	0.54	0.00		0.62	0.53	0.00		0.56	0.54	0.00	
GSC	ρ_{st}	0.00	-0.10	0.58	1.00		-0.08	0.56	1.00		-0.11	0.56	1.00		-0.13	0.56	1.00	
GSC	ρ_t	0.00	-0.01	1.36	1.00		-0.03	1.32	1.00		-0.01	1.31	1.00		-0.04	1.32	1.00	

Tabela 2 – Cenário 1: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas.

Para uma compreensão visual do resultado na Figura 13 tem-se os *boxplots* com as estimativas pontuais dos coeficientes de regressão dado o modelo gerador e os modelos de ajuste nas 100 replicações resumidas na Tabela 2.

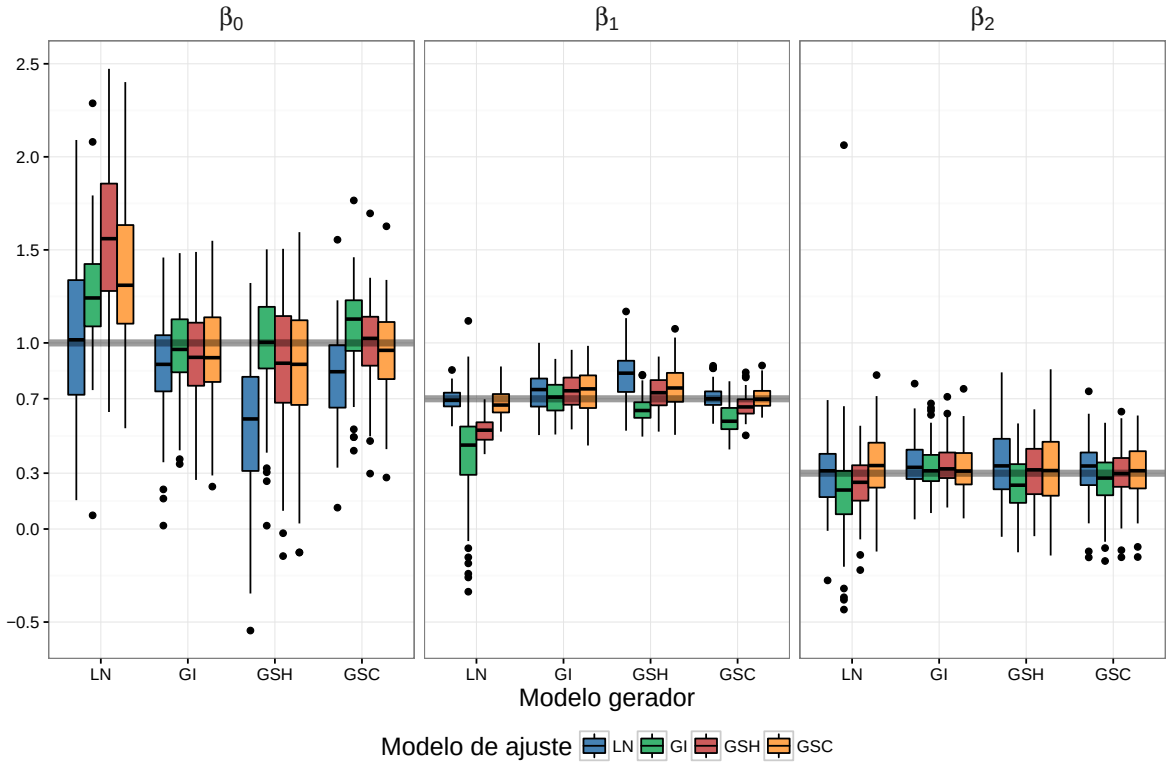


Figura 13 – Cenário 1: *Boxplots* das médias *a posteriori* dos coeficientes de regressão segundo o modelo gerador e modelo de ajuste.

É possível identificar na Figura 13 que as estimativas para o parâmetro β_0 são as que mais diferem entre os modelos Gama e o modelo LN. Isso é explicado pela expressão da média do modelo LN. Como a média do modelo LN é uma função da média dos demais pode-se identificar a relação entre os interceptos dos modelos Gama com o intercepto do modelo LN. Na Equação (10.1) tem-se esta relação dado ν e $\boldsymbol{Q}_{\boldsymbol{ij,ij}}^{\boldsymbol{-1}}$ fixos.

$$\begin{aligned} E(\mu_{ij}) &= \exp\left\{\boldsymbol{X_{ij}\beta} + 0.5\nu\boldsymbol{Q}_{\boldsymbol{ij,ij}}^{\boldsymbol{-1}}\right\} \\ &= \exp\left\{\beta_0 + \beta_1 X_{1ij} + ... + \beta_p X_{pij} + 0.5\nu\boldsymbol{Q}_{\boldsymbol{ij,ij}}^{\boldsymbol{-1}}\right\} \\ &= \exp\left\{\left(\beta_0 + 0.5\nu\boldsymbol{Q}_{\boldsymbol{ij,ij}}^{\boldsymbol{-1}}\right) + \beta_1 X_{1ij} + ... + \beta_p X_{pij}\right\} \\ &= \exp\left\{\beta_0^{Gama} + \beta_1 X_{1ij} + ... + \beta_p X_{pij}\right\} \end{aligned} \tag{10.1}$$

Medida	Modelo verdadeiro	Modelo ajustado				Número de bases	Taxa de acerto
		LN	GI	GSH	GSC		
DIC	LN	38	0	3	1	42	90.48
DIC	GI	0	33	0	0	33	100.00
DIC	GSH	1	7	27	0	35	77.14
DIC	GSC	27	0	6	18	51	35.29
DIC	-	-	-	-	-	-	72.05
LPML	LN	42	0	0	0	42	100.00
LPML	GI	21	8	4	0	33	24.24
LPML	GSH	35	0	0	0	35	0.00
LPML	GSC	51	0	0	0	51	0.00
LPML	-	-	-	-	-	-	31.06
$WAIC_1$	LN	33	0	9	0	42	78.57
$WAIC_1$	GI	0	28	5	0	33	84.85
$WAIC_1$	GSH	0	1	34	0	35	97.14
$WAIC_1$	GSC	12	0	17	22	51	43.14
$WAIC_1$	-	-	-	-	-	-	72.67
$WAIC_2$	LN	41	0	1	0	42	97.62
$WAIC_2$	GI	0	28	5	0	33	84.85
$WAIC_2$	GSH	11	4	20	0	35	57.14
$WAIC_2$	GSC	45	0	1	5	51	9.80
$WAIC_2$	-	-	-	-	-	-	58.39

Tabela 3 – Cenário 1: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$.

Desta forma, espera-se que a estimativa do intercepto dos modelos Gama seja algo próximo a $\beta_0 + 0.5\nu\boldsymbol{Q}_{ij,ij}^{-1}$ do modelo LN. Este fato justifica a diferença das estimativas dos coeficientes entre os modelos.

Na Tabela 3 tem-se o número de vezes em que o modelo gerador dos dados foi o que obteve o melhor ajuste segundo os critérios de ajuste *DIC, LPML,* $WAIC_1$ e $WAIC_2$. Foram computados apenas os casos em que todos os modelos competiram, isto é, todas as cadeias dos modelos convergiram concomitantemente.

Fica claro na Tabela 3 que o modelo LN é o modelo mais vezes selecionados pelos critérios de ajuste. Porém, a taxa de acerto dos critérios ao selecionar o modelo LN é baixa, aproximadamente 43%. O modelo GSC é o menos vezes selecionado inclusive nos casos em que ele é o modelo gerador dos dados. Esse resultado coincide com o encontrado por Prates et al. (2015) em que o modelo GSC é aparentemente bem estimado pelo modelo LN.

Neste cenário o critério de ajuste com o melhor resultado foi o critério $WAIC_1$ que selecionou o modelo correto em aproximadamente 72.5% das vezes e o pior critério foi o

LPML cuja taxa de acerto foi próxima de 31%. Esse cálculo foi feito somando o número de acertos da cada método dividido pelo número total de bases analisadas.

Os resultados dos Cenários 2, 3 e 4 encontram-se no Apêndice deste trabalho nas Tabelas 10, 12 e 14. Em aspectos gerais, os resultados se mantiveram parecidos. A taxa de cobertura teve uma ligeira melhora para os modelos em que se inseriu informação *a priori* para o parâmetro de dependência em estudo. As tabelas com as contagens do número de vezes em que os critérios de ajuste selecionaram o modelo correto encontram-se no Apêndice deste trabalho nas Tabelas 9, 11 e 13.

Cenário 5:

Para este cenário fixou-se o parâmetro $\nu = 3$ para todos os modelos assim como no Cenário 1. As regiões do estudo foram dispostas tal como um látice 10x10 totalizando 100 regiões. Além disso, gerou-se observações para as 100 regiões em dois tempos diferentes. Para esta análise os limites calculados para os parâmetros de dependência foram iguais aos limites do Cenário 1. Definiu-se $\rho_s = 2.24$, $\rho_t = 0$ e $\rho_{st} = 0$ garantindo inversibilidade da matriz de precisão adotada. Além disso utilizou-se duas covariáveis variando o seu efeito ao longo do tempo e o intercepto, ou seja, no total estimou-se cinco coeficientes. Denotou-se β_i^j o coeficiente associado a covariável i no tempo j.

Inseriu-se propositalmente no modelo uma covariável que perde importância do instante de tempo 1 para o instante de tempo 2 e uma que ganha importância com o passar do tempo e estão respectivamente associadas aos coeficientes β_1 e β_2. Desta forma, não se considera fixo o efeito das covariáveis. Este tipo estudo é importante para quando o pesquisador tem a intenção de observar a importância de uma dada covariável ao longo do tempo e identificar padrões específicos tais como crescimento, decrescimento ou até mesmo, acompanhar ao longo do tempo o desaparecimento da importância de uma determinada covariável.

Na Tabela 4 pode-se perceber semelhanças com os resultados do Cenário 1 no que diz respeito às estimativas pontuais do vetor $\boldsymbol{\beta}$, à taxa de cobertura dos intervalos *HPD*, à dificuldade de estimação do parâmetro ν e aos parâmetros de dependência espacial, temporal e espaço-temporal.

O mesmo raciocínio pode ser utilizado para criar coeficientes relacionados as regiões em estudo para quando se tem o interesse de testar a influência das covariáveis em diferentes regiões. Ou seja, permitir que o efeito das covariáveis varie nas diferentes regiões o que muitas vezes pode fazer sentido para o pesquisador. Para tanto, ambas as abordagens necessitam de um número grande de regiões/tempos para que os coeficientes sejam bem estimados.

Modelo Verdadeiro	Parâmetro	Valor verdadeiro	Modelo de ajuste															
			LN				GI				GSH				GSC			
			Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %
LN	β_0	1.00	0.92	0.09	0.48	84	1.52	0.09	0.07	81	1.71	0.11	0.09	81	1.27	0.13	0.45	66
LN	β_1^1	0.70	0.71	0.08	0.94		0.49	0.06	0.16		0.54	0.06	0.30		0.68	0.09	0.95	
LN	β_1^2	0.30	0.30	0.09	0.93		0.22	0.05	0.43		0.23	0.07	0.73		0.28	0.08	0.85	
LN	β_2^1	-0.35	-0.36	0.11	0.96		-0.20	0.07	0.35		-0.27	0.09	0.84		-0.35	0.11	0.97	
LN	β_2^2	1.60	1.59	0.13	0.86		1.25	0.09	0.20		1.24	0.08	0.10		1.49	0.13	0.82	
LN	ν	3.00	5.86	0.98	0.17		28.82	1.02	0.00		10.33	1.65	0.09		0.93	0.31	0.03	
LN	ρ_s	2.24	0.82	0.52	0.00		0.69	0.48	0.00		0.78	0.47	0.00		0.79	0.47	0.00	
LN	ρ_{st}	0.00	-0.05	0.57	1.00		0.01	0.51	1.00		-0.01	0.54	1.00		-0.07	0.54	1.00	
LN	ρ_t	0.00	-0.07	1.36	1.00		-0.05	1.19	1.00		-0.12	1.23	1.00		-0.00	1.24	1.00	
GI	β_0	1.00	0.84	0.07	0.34	80	0.99	0.07	0.59	73	0.96	0.07	0.50	84	0.96	0.06	0.29	79
GI	β_1^1	0.70	0.74	0.05	0.62		0.70	0.03	0.86		0.71	0.04	0.77		0.72	0.05	0.58	
GI	β_1^2	0.30	0.34	0.06	0.80		0.31	0.04	0.93		0.32	0.05	0.93		0.32	0.06	0.76	
GI	β_2^1	-0.35	-0.37	0.07	0.85		-0.35	0.06	0.95		-0.36	0.07	0.90		-0.36	0.07	0.82	
GI	β_2^2	1.60	1.67	0.07	0.65		1.60	0.04	0.68		1.61	0.04	0.65		1.62	0.07	0.63	
GI	ν	3.00	1.16	0.39	0.19		3.23	1.03	0.85		1.09	0.27	0.04		0.17	0.05	0.00	
GI	ρ_s	2.24	0.25	0.57	0.00		0.42	0.55	0.00		0.44	0.55	0.00		0.30	0.56	0.00	
GI	ρ_{st}	0.00	-0.05	0.59	1.00		-0.07	0.57	1.00		-0.07	0.56	1.00		-0.08	0.56	1.00	
GI	ρ_t	0.00	-0.07	1.45	1.00		-0.05	1.36	1.00		-0.10	1.35	1.00		-0.06	1.40	1.00	
GSH	β_0	1.00	0.62	0.10	0.22	76	1.07	0.10	0.51	76	0.92	0.09	0.39	72	0.89	0.09	0.25	77
GSH	β_1^1	0.70	0.80	0.08	0.66		0.66	0.04	0.63		0.72	0.05	0.74		0.76	0.08	0.75	
GSH	β_1^2	0.30	0.34	0.08	0.89		0.28	0.05	0.76		0.33	0.07	0.92		0.32	0.08	0.94	
GSH	β_2^1	-0.35	-0.39	0.10	0.91		-0.29	0.07	0.75		-0.34	0.09	0.93		-0.36	0.10	0.91	
GSH	β_2^2	1.60	1.79	0.11	0.49		1.56	0.05	0.63		1.63	0.06	0.68		1.67	0.11	0.68	
GSH	ν	3.00	3.85	0.87	0.75		11.90	3.37	0.21		2.64	0.53	0.71		0.57	0.12	0.00	
GSH	ρ_s	2.24	0.66	0.54	0.00		0.69	0.50	0.00		0.75	0.50	0.00		0.66	0.52	0.00	
GSH	ρ_{st}	0.00	-0.04	0.57	1.00		-0.02	0.54	1.00		-0.01	0.56	1.00		-0.03	0.55	1.00	
GSH	ρ_t	0.00	-0.11	1.36	1.00		-0.04	1.26	1.00		-0.07	1.31	1.00		-0.10	1.30	1.00	
GSC	β_0	1.00	0.86	0.07	0.38	84	1.17	0.07	0.36	80	1.14	0.08	0.39	69	0.96	0.07	0.52	75
GSC	β_1^1	0.70	0.70	0.06	0.95		0.60	0.04	0.32		0.64	0.05	0.55		0.70	0.06	0.96	
GSC	β_1^2	0.30	0.31	0.07	0.89		0.27	0.04	0.55		0.28	0.06	0.86		0.32	0.07	0.91	
GSC	β_2^1	-0.35	-0.35	0.08	0.98		-0.29	0.06	0.64		-0.31	0.08	0.93		-0.35	0.08	0.97	
GSC	β_2^2	1.60	1.61	0.09	0.71		1.49	0.04	0.22		1.52	0.05	0.42		1.61	0.09	0.81	
GSC	ν	0.33	2.02	0.46	0.00		8.60	2.10	0.34		2.54	0.49	0.01		0.29	0.07	0.72	
GSC	ρ_s	2.24	0.50	0.56	0.00		0.42	0.53	0.00		0.58	0.52	0.00		0.48	0.53	0.00	
GSC	ρ_{st}	0.00	-0.06	0.58	1.00		-0.01	0.55	1.00		-0.10	0.55	1.00		-0.07	0.56	1.00	
GSC	ρ_t	0.00	-0.09	1.37	1.00		-0.07	1.32	1.00		-0.04	1.27	1.00		-0.08	1.32	1.00	

Tabela 4 – Cenário 5: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas.

Na Figura 14 selecionou-se os coeficientes de regressão das covariáveis cuja influência varia ao longo do tempo e fez-se os *boxplots* com as estimativas pontuais de cada um dos coeficientes de regressão dado o modelo gerador e os modelos de ajuste nas 100 replicações resumidas na Tabela 4.

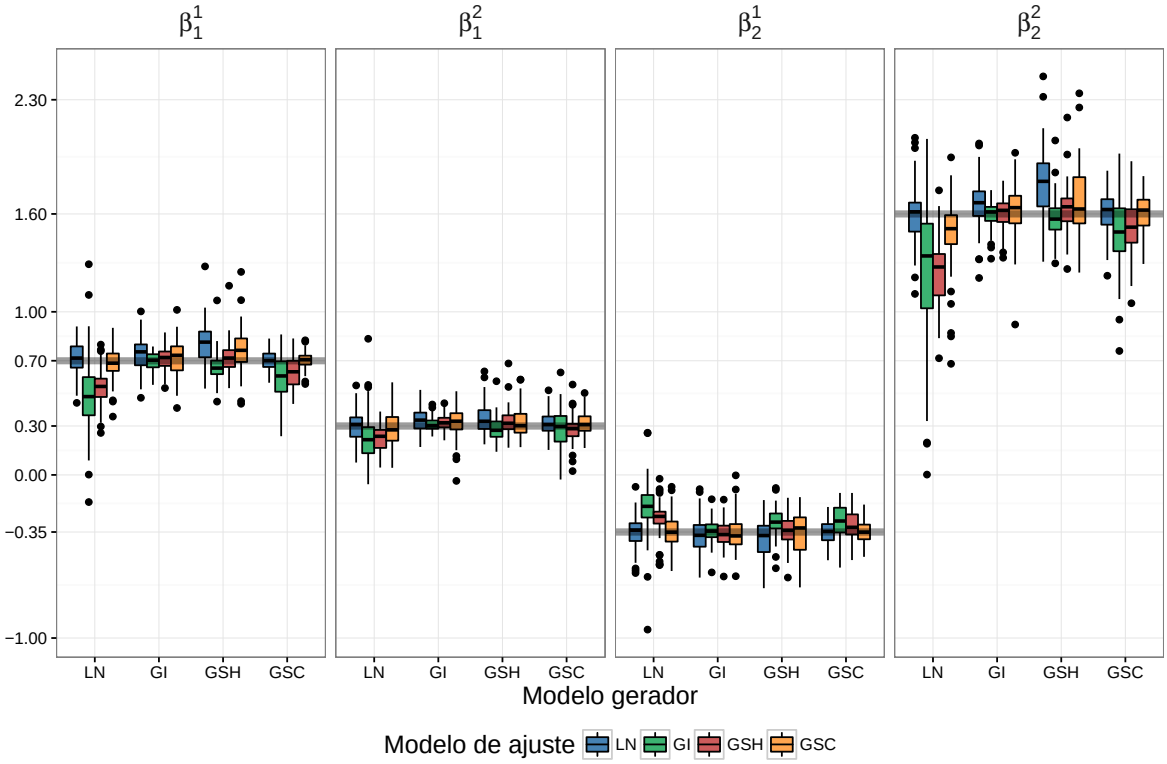


Figura 14 – Cenário 5: *Boxplots* das médias *a posteriori* dos coeficientes de regressão segundo o modelo gerador e modelo de ajuste.

Percebe-se na Figura 14 que em geral os coeficientes são bem estimados pelos modelos. É interessante observar que a influência das covariáveis varia de um instante de tempo para o outro e que a análise convencional com coeficientes atemporais iria estimar apenas um coeficiente. Desta forma, a má especificação do modelo introduziria viés na análise.

Ressalva-se que é interessante testar as duas configurações de modelagem. Em um primeiro momento testa-se a configuração habitual de um coeficiente por covariável para não inserir parâmetros desnecessários no modelo e em seguida testa-se os coeficientes para cada instante de tempo para evitar a má especificação do modelo.

Na Tabela 5 tem-se o número de vezes em que o modelo gerador dos dados foi o que obteve o melhor ajuste segundo os critérios de ajuste *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Medida	Modelo verdadeiro	Modelo ajustado				Número de bases	Taxa de acerto
		LN	GI	GSH	GSC		
DIC	LN	31	0	3	0	34	91.18
DIC	GI	0	35	1	0	36	97.22
DIC	GSH	0	3	32	0	35	91.43
DIC	GSC	24	0	6	1	31	3.23
DIC	-	-	-	-	-	-	72.79
LPML	LN	34	0	0	0	34	100.00
LPML	GI	20	5	11	0	36	13.89
LPML	GSH	34	0	1	0	35	2.86
LPML	GSC	31	0	0	0	31	0.00
LPML	-	-	-	-	-	-	29.41
$WAIC_1$	LN	15	0	19	0	34	44.12
$WAIC_1$	GI	0	30	6	0	36	83.33
$WAIC_1$	GSH	0	0	35	0	35	100.00
$WAIC_1$	GSC	11	0	16	4	31	12.90
$WAIC_1$	-	-	-	-	-	-	61.76
$WAIC_2$	LN	33	0	1	0	34	97.06
$WAIC_2$	GI	0	24	12	0	36	66.67
$WAIC_2$	GSH	0	0	35	0	35	100.00
$WAIC_2$	GSC	30	0	1	0	31	0.00
$WAIC_2$	-	-	-	-	-	-	67.65

Tabela 5 – Cenário 5: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Vê-se na Tabela 5 que para o Cenário 5 o método com a maior taxa de acerto é o *DIC*. Novamente o modelo em que os critérios menos acertam é o modelo GSC. É possível perceber que o método *LPML* tende a sempre selecionar o modelo LN se mostrando um critério não muito interessante a ser adotado. Este resultado diverge do que foi obtido por Prates et al. (2015) em que o método *LPML* mostrou-se um bom critério para seleção de modelos.

Os resultados do Cenário 6 encontram-se no Apêndice deste trabalho na Tabela 16 e são similares aos resultados encontrados no Cenário 5. O grande diferencial desta análise se dá no grande número de coeficientes estimados (23 ao todo). Para simplificar a tabela escolheu-se apenas 4 coeficientes para apresentação dos resultados. Na Tabela 15 tem-se o número de vezes em que o modelo gerador dos dados foi o que obteve o melhor ajuste segundo os critérios de ajuste *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Como resultado geral das simulações tem-se que os intervalos *HPD* em geral são pouco amplos visto que as estimativas pontuais são boas (*boxplot*) e a taxa de cobertura dos intervalos em geral está abaixo da taxa nominal. Também foi possível verificar que os

parâmetros de dependência tanto temporal quanto espacial em geral são subestimados pelo método. Uma possível razão para isto é a relação entre as quantidades de dependência visto que em conjunto elas devem garantir inversibilidade da matriz de precisão.

Quanto aos critérios de seleção o método com a maior taxa de acerto geral foi o método $WAIC_1$ com uma taxa de acerto sendo próxima a 72.50%. O pior dos critérios empregados foi o *LPML* com uma taxa de acerto em torno de 30.50%. Desta forma, definiu-se como método de seleção de modelos o critério $WAIC_1$ à ser aplicado ao conjunto de dados de contagem de gastrópodes da espécie *Nenia Tridens*. No Capítulo 11 a metodologia dos *TGMRF's* é aplicada a este conjunto de dados.

11 Abundância de *Nenia Tridens*

Gastrópodes constituem uma grande classe de moluscos e são muito importantes ecologicamente, tendo um papel fundamental na recomposição da fertilidade do solo. É de interesse dos pesquisadores caracterizar o *habitat* mais atrativo para essas espécies.

Os dados utilizados para a aplicação da metodologia foram obtidos em uma floresta localizada na região nordeste de Porto Rico (Montanhas Luquillo). Realizou-se a contagem de gastrópodes da espécie *Nenia tridens* entre os anos de 2010 e 2014, em quarenta pontos de observação sendo que as contagens foram realizadas quatro vezes nos períodos chuvosos. Ou seja, para cada ano em cada região foram realizadas quatro contagens totalizando oitocentas contagens no período. Os quarenta pontos de observação são dispostos tal como um látice de dimensões 8×5.

A variável resposta deste estudo é o menor número de gastrópodes vivos da espécie *Nenia tridens* encontrados nas quatro contagens realizadas por região/ano. Essa variável indica o número mínimo de gastrópodes da espécie *Nenia tridens* que vive próximo de cada um dos pontos de observação. Desta forma, tem-se duzentas observações, uma para cada região no período de 5 anos.

Na Figura 15 pode-se ver o látice que descreve os pontos de coleta e também as contagens de caramujos nos cinco anos em estudo. Além disso, pode-se ver a definição adotada para vizinhança dos pontos de coleta (Figura 15 ano de 2012).

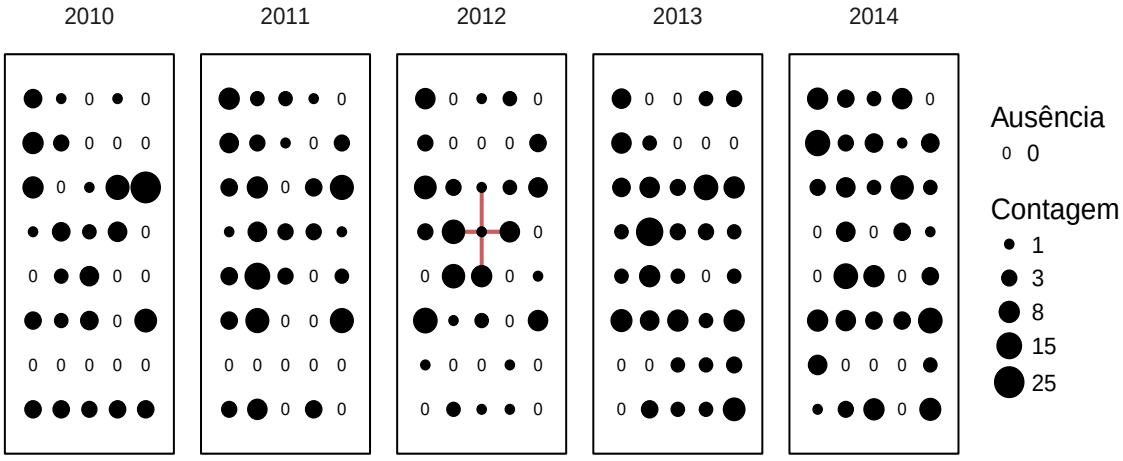


Figura 15 – Látice e contagens do número de caramujos da espécie *Nenia Tridens* para os anos de 2010 à 2014 e estrutura de vizinhança adotada.

Os pesquisadores observam algumas covariáveis durante a contagem dos gastrópodes

as quais se quer testar a importância no modelo. As estatísticas descritivas das covariáveis e da variável resposta podem ser vistas na Tabela 6.

Variável	Mínimo	25%	Média	Mediana	75%	Máximo	Desvio padrão
Nenia Tridens	0.00	0.00	2.80	1.00	4.00	25.00	4.27
Elevação	3.46	3.67	3.78	3.75	3.90	4.18	0.17
Inclinação	0.08	0.16	0.23	0.22	0.27	0.55	0.11
Vegetação total	0.00	8.00	15.84	14.00	21.00	74.00	11.02
Sierra Palm	0.00	1.00	7.14	5.00	10.25	49.00	7.31
Detritos	1.00	2.00	2.80	3.00	3.00	5.00	0.79
Abertura da Copa	0.00	4.16	7.82	7.28	10.40	24.96	4.67

Tabela 6 – Análise descritiva das covariáveis utilizadas na aplicação.

Estas covariáveis descrevem o *habitat* encontrado pelos moluscos na região em estudo. Para a análise, as covariáveis foram padronizadas e escaladas. Abaixo tem-se uma breve descrição das covariáveis utilizadas no estudo:

- **Inclinação:** Inclinação do terreno medido em graus.
- **Elevação:** Altitude em relação ao nível do mar medida em metros.
- **Vegetação total:** Volume ocupado por toda vegetação rasteira.
- ***Sierra Palm*:** Volume ocupado pela vegetação rasteira da espécie *Sierra Palm*.
- **Detritos:** Grau em que o solo é coberto por detritos no intervalo [0, 5].
- **Abertura da copa:** Abertura da copa das árvores no ponto de observação.

Analisou-se a base de dados de duas formas. Num primeiro estudo, todas as covariáveis foram consideradas fixas no tempo. Em um segundo estudo, duas das covariáveis foram consideradas fixas no tempo (Inclinação e Elevação) e as demais foram consideradas covariáveis que variam ao longo do tempo. Desta forma, é possível identificar se as covariáveis ganham/perdem importância na modelagem. Sendo assim, o segundo modelo efetivamente conta com $2 + 4 * 5 = 22$ covariáveis.

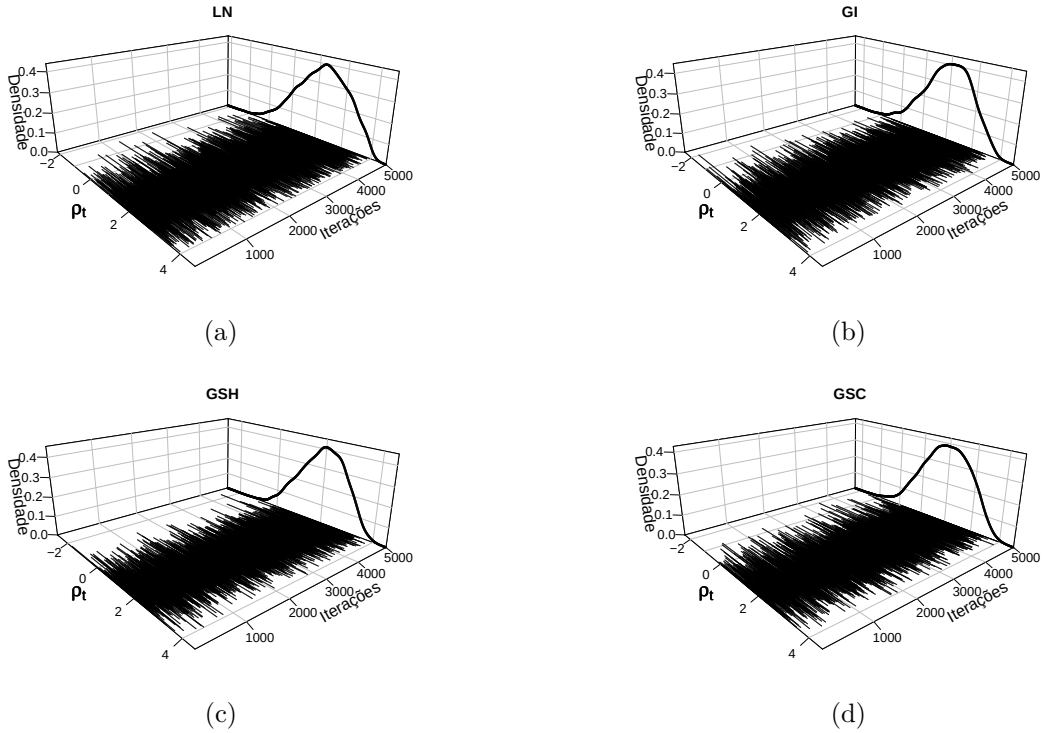
Adotando-se a modelagem dos *TGMRF's* tem-se $Y_{1,1}, \ldots, Y_{1,5}, Y_{2,1}, \ldots, Y_{2,5}, \ldots, Y_{40,5}$ variáveis aleatórias independentes dado μ_{ij} em que $Y_{ij}|\mu_{ij} \sim Poisson(\mu_{ij})$. Como explicado anteriormente, pode-se modelar μ_{ij} através de qualquer distribuição contínua cujo domínio seja os números reais positivos. Nesta aplicação, as distribuições escolhidas foram as mesmas do estudo de simulação da Seção 10 e que foram explicadas na Seção 8.3.

Para análise do banco de dados, foram computadas 300000 iterações do algoritmo descrito na Seção 8.2. Após análise das cadeias, tomou-se um período de aquecimento

de 200000 iterações. Com as 100000 observações restantes, tomou-se um *lag* de tamanho 20 para diminuir a autocorrelação das cadeias, totalizando uma amostra *a posteriori* de tamanho 5000. A análise de convergência foi verificada pelo critério de *Geweke*. Como critério de seleção de modelos optou-se pelo Método $WAIC_1$ visto que produziu os melhores resultados durante o estudo de simulação.

Estudo 1:

Ajustou-se o modelo utilizando as 6 covariáveis descritas na Tabela 6. O objetivo da análise é entender quais covariáveis são importantes para explicar a abundância de gastrópodes da espécie *Nenia Tridens* e também quais são os efeitos de dependência encontrados na base de dados. Primeiramente avaliou-se os parâmetros de dependência estimados pelo modelo e percebeu-se que o parâmetro de dependência temporal obteve estimativas positivas. Esse resultado pode ser observado na Tabela 7. A Figura 16 mostra as cadeias do parâmetro ρ_t e a sua densidade empírica segundo os quatro modelos ajustados.



(a) (b)

(c) (d)

Figura 16 – Cadeias do parâmetro ρ_t. (a): LN. (b): GI. (c): GSH. (d): GSC

Na Figura 16 pode-se perceber que o parâmetro ρ_t tende a ser positivo. Como visto no Capítulo 10 os parâmetros de dependência são normalmente subestimados, sugerindo que a real dependênica temporal é maior que a estimada. Os demais parâmetros de dependência são estimados próximos a zero como pode ser visto Na Tabela 7.

Pode-se perceber na Tabela 7 que a estimativa para o intercepto do modelo LN difere bastante da estimativa dos demais modelos assim como foi visto no estudo de simulação. Para as demais covariáveis as estimativas pontuais foram semelhantes. Uma

Parâmetro	Modelo ajustado							
	LN		GI		GSH		GSC	
	Média	*HPD*	Média	*HPD*	Média	*HPD*	Média	*HPD*
Intercepto	0.66	(0.43,0.87)	1.22	(1.03,1.41)	1.20	(1.01,1.4)	1.23	(1,1.43)
Elevação	-0.25	(-0.48,-0.04)	-0.18	(-0.3,-0.06)	-0.20	(-0.37,-0.04)	-0.12	(-0.31,0.06)
Inclinação	0.16	(-0.04,0.37)	0.10	(-0.01,0.21)	0.13	(-0.02,0.29)	0.12	(-0.09,0.32)
Vegetação total	0.20	(0,0.42)	0.10	(0,0.21)	0.13	(-0.02,0.26)	0.16	(-0.07,0.38)
Sierra Palm	-0.09	(-0.29,0.12)	-0.06	(-0.17,0.06)	-0.07	(-0.23,0.09)	-0.07	(-0.29,0.16)
Detritos	0.25	(0.05,0.45)	0.12	(0,0.23)	0.17	(0.03,0.34)	0.20	(0.01,0.4)
Abertura da copa	-0.00	(-0.18,0.18)	-0.02	(-0.13,0.09)	-0.01	(-0.16,0.14)	0.03	(-0.15,0.2)
ρ_s	0.07	(-1.41,1.42)	0.07	(-1.2,1.4)	0.08	(-1.18,1.36)	0.11	(-1.16,1.47)
ρ_t	2.07	(0.32,3.94)	1.87	(0.04,3.59)	1.84	(0.05,3.52)	1.83	(0.06,3.54)
ρ_{st}	0.10	(-0.94,1.11)	0.08	(-0.89,1.13)	0.05	(-0.93,1.15)	0.07	(-0.92,1.1)
ν	11.44	(7.16,15.85)	15.19	(8.73,22.58)	4.42	(2.97,5.94)	0.77	(0.55,1)
$WAIC_1$	2590.96	-	3064.46	-	3033.69	-	2978.01	-

Tabela 7 – Estudo 1: Médias *a posteriori* para os coeficientes associados às covariáveis e intervalos *HPD*.

maneira de decidir se determinada covariável é importante para o modelo é observando se o *HPD* não contém o valor 0. Neste estudo, percebe-se que dependendo do modelo escolhido, diferentes variáveis são consideradas importantes. Na Figura 17 tem-se as densidades *a posteriori* dos coeficientes de regressão segundo os modelos de ajuste. Em cinza tem-se o valor zero como ponto de referência.

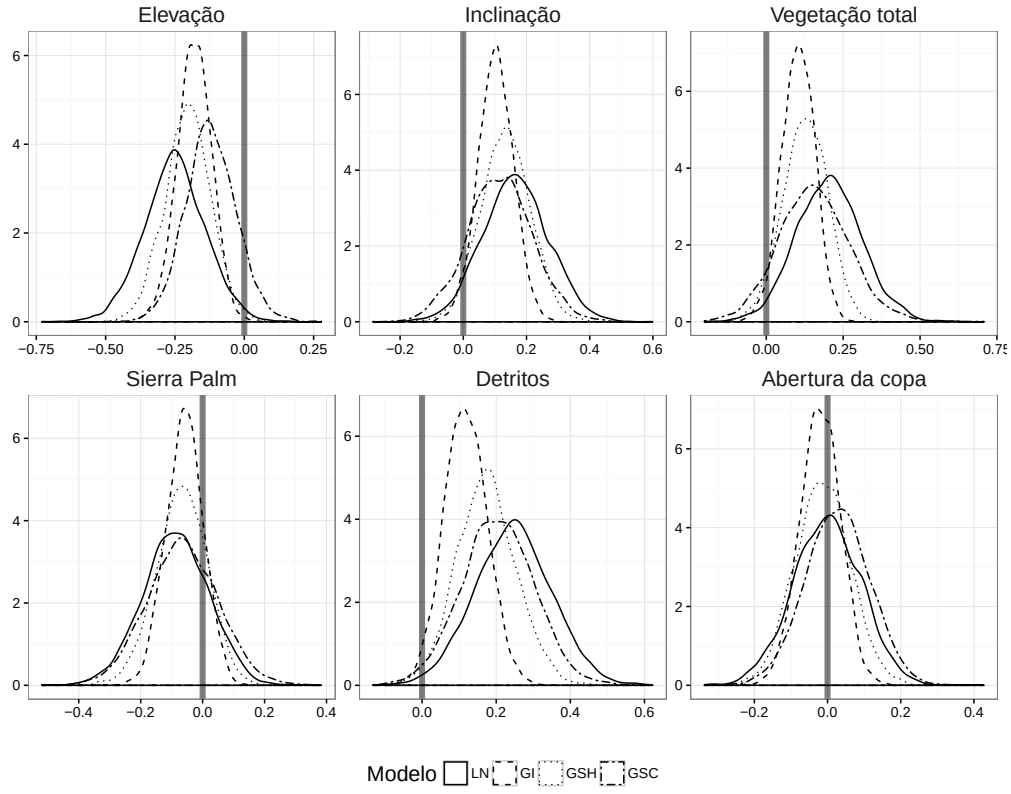


Figura 17 – Densidades *a posteriori* para os coeficientes associados às covariáveis segundo os modelos ajustados.

Percebe-se na Figura 17 que em geral as densidades *a posteriori* do modelo GI são mais concentradas em torno do seu valor central. Já as densidades *a posteriori* do modelo LN aparentam ter cauda mais pesada do que os demais modelos. Apesar da cauda mais pesada, o modelo LN é o modelo que mais considera covariáveis importantes segundo o intervalo *HPD*. Como pode-se ver na Tabela 7, o modelo com o menor $WAIC_1$ foi o modelo LN que considera importante as covariáveis *Elevação*, *Vegetação total* e *Detritos*.

Estudo 2:

Neste estudo, ajustou-se o modelo utilizando as 6 covariáveis descritas na Tabela 6. As covariáveis *Elevação* e *Inclinação* se mantém constante ao longo do tempo. Desta forma, achou-se interessante utilizar apenas um coeficiente associado a estas covariáveis independente do tempo. As demais covariáveis foram consideradas com efeitos distintos em cada um dos tempos. Efetivamente utilizou-se 22 covariáveis neste estudo. Na Tabela 8 pode-se ver as médias *a posteriori* e intervalos *HPD* para os parâmetros do modelo.

Parâmetro	Modelo ajustado							
	LN		GI		GSH		GSC	
	Média	*HPD*	Média	*HPD*	Média	*HPD*	Média	*HPD*
Intercepto	0.64	(0.35,0.92)	1.11	(0.88,1.33)	1.13	(0.89,1.37)	1.20	(0.94,1.48)
Elevação	-0.26	(-0.49,-0.03)	-0.21	(-0.35,-0.08)	-0.21	(-0.37,-0.03)	-0.16	(-0.36,0.05)
Inclinação	0.16	(-0.07,0.37)	0.12	(0,0.23)	0.14	(-0.02,0.3)	0.09	(-0.12,0.29)
Vegetação total - 2010	-0.02	(-0.76,0.73)	0.01	(-0.31,0.31)	-0.08	(-0.52,0.38)	-0.08	(-0.83,0.73)
Vegetação total - 2011	0.20	(-0.14,0.54)	0.08	(-0.09,0.24)	0.08	(-0.15,0.3)	0.32	(-0.08,0.78)
Vegetação total - 2012	0.38	(-0.29,1)	0.14	(-0.2,0.48)	0.17	(-0.31,0.64)	0.37	(-0.37,1.1)
Vegetação total - 2013	0.09	(-0.34,0.48)	-0.01	(-0.25,0.2)	0.01	(-0.28,0.29)	0.16	(-0.33,0.66)
Vegetação total - 2014	0.41	(-0.2,1.04)	0.33	(0,0.62)	0.33	(-0.09,0.71)	0.23	(-0.29,0.76)
Sierra Palm - 2010	0.12	(-0.45,0.74)	0.12	(-0.17,0.38)	0.13	(-0.27,0.52)	0.11	(-0.5,0.71)
Sierra Palm - 2011	-0.06	(-0.38,0.27)	-0.02	(-0.22,0.16)	-0.03	(-0.32,0.22)	-0.08	(-0.47,0.32)
Sierra Palm - 2012	-0.18	(-0.89,0.52)	0.04	(-0.33,0.41)	-0.01	(-0.55,0.53)	-0.24	(-0.99,0.56)
Sierra Palm - 2013	-0.16	(-0.65,0.27)	-0.19	(-0.46,0.08)	-0.17	(-0.52,0.17)	-0.09	(-0.65,0.44)
Sierra Palm - 2014	-0.35	(-1.11,0.42)	-0.64	(-1.12,-0.16)	-0.56	(-1.19,0.05)	-0.13	(-0.69,0.48)
Detritos - 2010	0.57	(0.04,1.07)	0.23	(-0.02,0.45)	0.37	(0.05,0.7)	0.58	(0.1,1.1)
Detritos - 2011	0.68	(0.15,1.18)	0.38	(0.04,0.7)	0.51	(0.07,0.95)	0.64	(0.08,1.26)
Detritos - 2012	-0.16	(-0.71,0.37)	-0.09	(-0.35,0.16)	-0.12	(-0.48,0.26)	-0.15	(-0.78,0.48)
Detritos - 2013	0.39	(-0.01,0.77)	0.27	(0.02,0.5)	0.33	(0.03,0.61)	0.39	(-0.03,0.81)
Detritos - 2014	-0.07	(-0.55,0.42)	-0.05	(-0.28,0.19)	-0.07	(-0.37,0.27)	-0.03	(-0.43,0.35)
Abertura da copa - 2010	-0.14	(-0.92,0.63)	-0.03	(-0.43,0.36)	-0.08	(-0.58,0.43)	-0.38	(-1.09,0.29)
Abertura da copa - 2011	-0.15	(-0.58,0.26)	-0.12	(-0.4,0.15)	-0.18	(-0.58,0.2)	-0.12	(-0.51,0.23)
Abertura da copa - 2012	0.12	(-0.39,0.59)	0.01	(-0.26,0.31)	0.08	(-0.3,0.49)	0.12	(-0.33,0.6)
Abertura da copa - 2013	-0.20	(-0.63,0.23)	-0.19	(-0.46,0.07)	-0.21	(-0.57,0.14)	-0.12	(-0.58,0.29)
Abertura da copa - 2014	0.09	(-0.42,0.63)	-0.06	(-0.38,0.26)	-0.03	(-0.46,0.37)	0.19	(-0.24,0.61)
ρ_s	0.17	(-1.18,1.5)	0.06	(-1.14,1.31)	0.15	(-1.07,1.39)	0.15	(-1.08,1.41)
ρ_t	2.03	(0.26,3.79)	1.81	(-0.07,3.49)	1.80	(0.03,3.62)	1.81	(0.02,3.62)
ρ_{st}	-0.04	(-1.05,1.04)	0.00	(-1,1.01)	-0.07	(-1.04,0.95)	0.02	(-0.99,1.02)
ν	12.72	(8.12,17.73)	11.71	(6.54,17.44)	4.14	(2.76,5.51)	0.74	(0.54,0.97)
$WAIC_1$	2710.79	-	3318.63	-	3217.31	-	3060.53	-

Tabela 8 – Estudo 2: Médias *a posteriori* para os coeficientes associados às covariáveis e intervalos *HPD*.

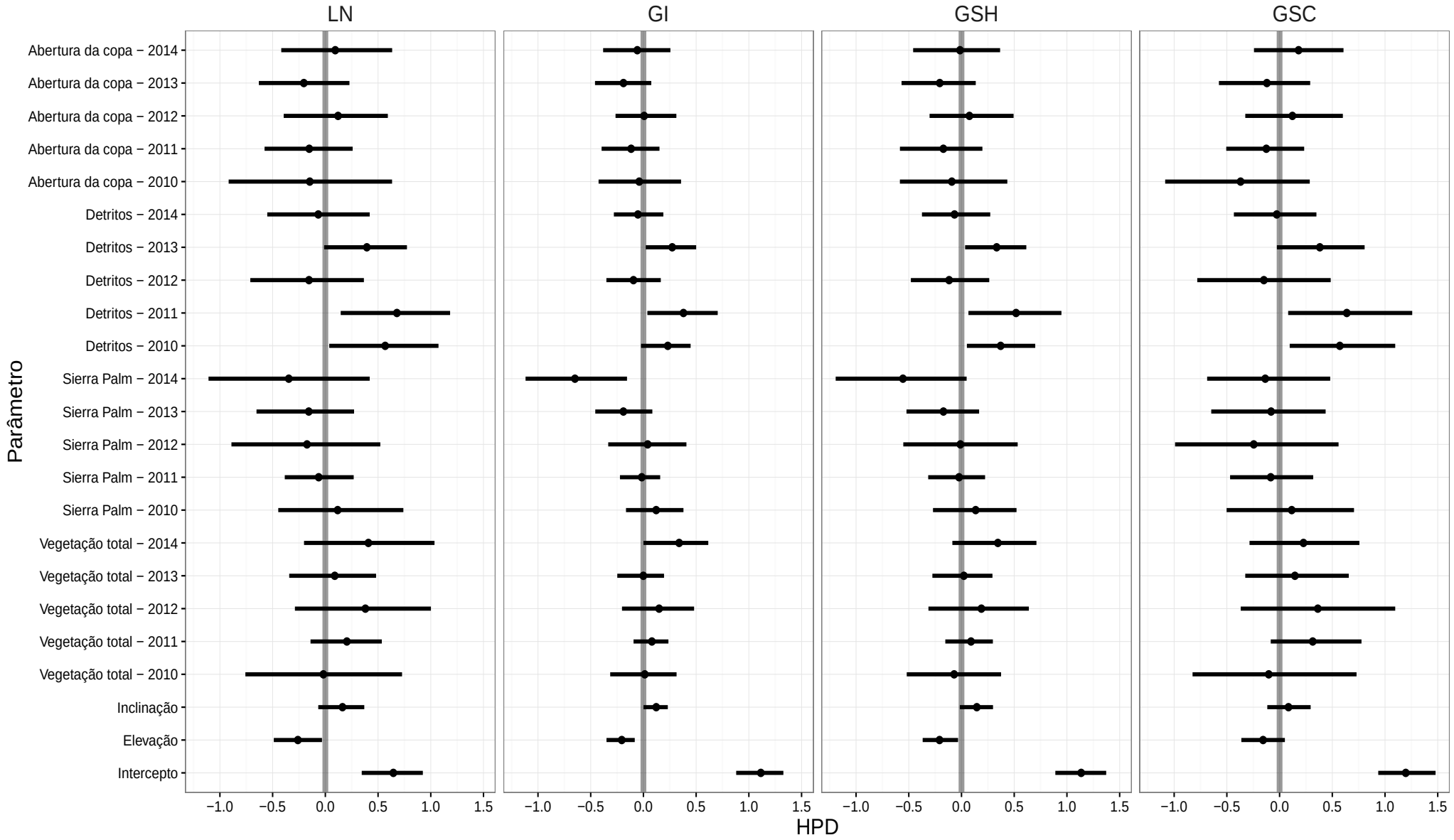


Figura 18 – Intervalos *HPD* e mediana *a posteriori* dos coeficientes associados às covariaveis segundo os modelos ajustados.

Pode-se perceber na Tabela 8 que o parâmetro de dependência temporal novamente obteve estimativas positivas sugerindo esta característica. Para os parâmetros de dependência espacial e espaço-temporal novamente não verificou-se uma forte dependência.

Para uma visualização mais adequada dos resultados da Tabela 8, na Figura 18 tem-se um gráfico que apresenta os intervalos *HPD* dos coeficientes associados às covariáveis segundo os quatro modelos ajustados. Nesta figura tem-se o intervalo *HPD* dos coeficientes realçados com a linha horizontal. Os pontos centrais representam a média *a posteriori* dos coeficientes. A linha central em cinza representa o valor 0 ao qual se pode utilizar como base para decisão da importância da covariável no modelo.

Percebe-se na Figura 18 que a maioria dos intervalos *HPD* contém o valor zero, indicando que tais covariáveis não são importantes para o modelo. Nota-se também que as covariáveis não tem nenhuma tendência de crescimento/decrescimento clara, indicando que possivelmente a modelagem considerando a dinâmica temporal não é necessária (veja por exemplo a covariável *sierra palm* em todos os modelos). Neste caso, a inserção de um maior número de coeficientes não melhora em grande escala o ajuste do modelo. A inserção destas covariáveis no modelo aumenta a sua complexidade e por não trazer muita melhoria ao ajuste acaba tornando o modelo menos parcimonioso.

Pode-se observar que os modelos tendem a concordar quanto a direção da importância das covariáveis variando apenas a intensidade. Por exemplo, a covariável *Sierra Palm - 2014* possui um coeficiente negativo segundo todos os modelos. Porém, a intensidade varia de acordo com o modelo. O modelo LN tem uma estimativa pontual próxima a -0.35 e um amplo intervalo *HPD* já para o modelo GI esse coeficiente é estimado próximo a -0.64 e é considerado importante para o modelo visto que o *HPD* não contém o valor zero.

Novamente o modelo com o menor $WAIC_1$ foi o LN. Neste estudo o modelo LN considera importante as covariáveis *Elevação*, *Detritos-2010* e *Detritos-2011*.

Para ambos os estudos o modelo LN foi o que obteve menor $WAIC_1$. Quando comparados os dois modelos vencedores chega-se a conclusão que o modelo LN abordado no Estudo 1 é o que tem melhor ajuste visto que os valores do critério $WAIC_1$ observados foram 2590.96 para o Estudo 1 e 2710.79 para o Estudo 2. Este resultado vai ao encontro do que se é esperado visto que as covariáveis utilizadas não possuem nenhum tipo de tendência como citado anteriormente.

O modelo LN abordado no Estudo 1 aponta *Elevação*, *Vegetação total* e *Detritos* como covariáveis importantes. Do ponto de vista ecológico este resultado é bastante interessante. A abundância da espécie *Nenia Tridens* varia ao longo do tempo de forma suave. Ou seja, fixado um determinado ponto de observação, se a contagem de caramujos tende a aumentar ao longo do tempo, esse crescimento se dá de forma suave e o mesmo acontece quando a tendência é de decrescimento.

Além disso, as contagens de gastrópodes decrescem com relação à elevação em relação ao nível do mar. Este resultado é bastante conhecido em ecologia e em geral é observado para elevadas altitudes. Na Figura 19 tem-se as contagens de gastrópodes em conjunto com as respectivas elevações dos locais de coleta.

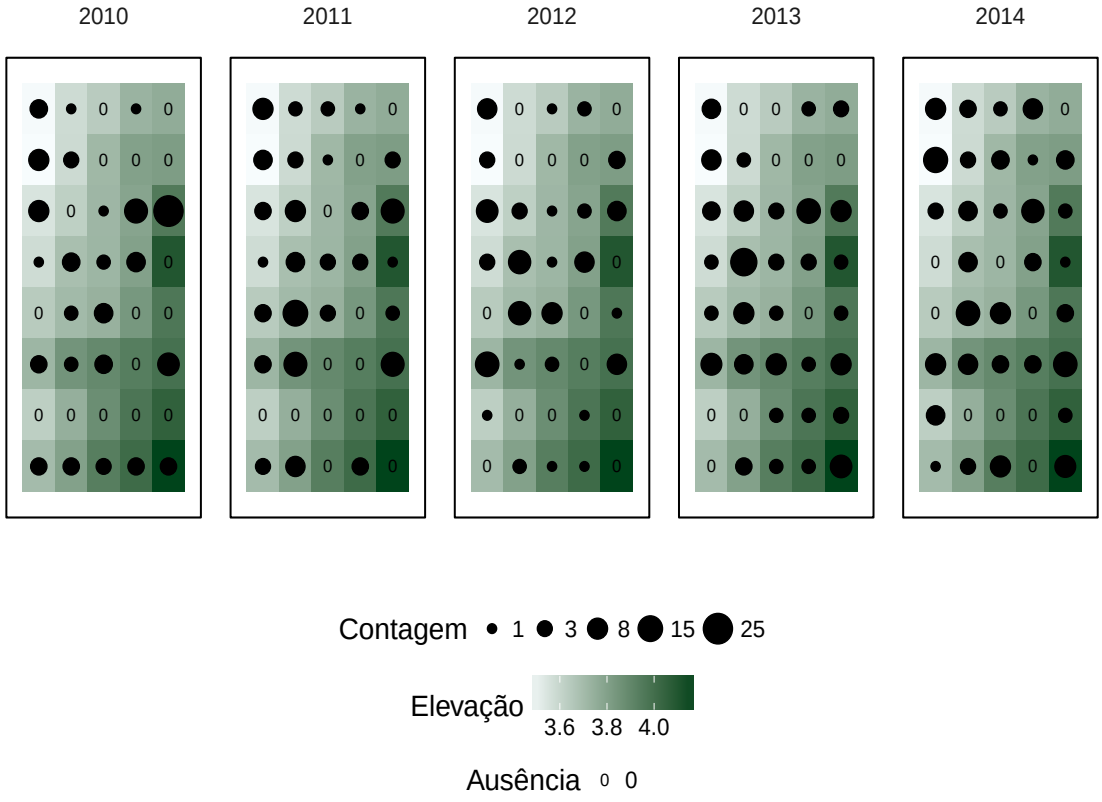


Figura 19 – Contagem de gastrópodes e elevações no período em análise.

Percebe-se na Figura 19 que em geral as contagens dos gastrópodes em regiões mais altas efetivamente são menores do que em regiões mais baixas como esperado. A relação com as covariáveis *Vegetação total* e *Detritos* se justifica já que gastrópodes preferem matas fechadas e ambientes com mais detritos. Esta combinação fornece um ambiente mais úmido e com uma temperatura mais adequada além de conter grande quantidade de alimentos favorecendo ainda mais o desenvolvimento da espécie.

12 Conclusões

A análise espaço-temporal para dados de área é um assunto atual e de grande aplicabilidade. Tal análise pode ser incorporada em modelos *GLMM* através de efeitos aleatórios, com uma estrutura introduzida de forma separável ou não separável. Essa última é mais realista porém mais complexa visto que requer uma estrutura de dependência conveniente e com interpretações simples. Pode-se citar duas grandes limitações dos modelos *GLMM*: 1) Podem sofrer com os problemas do confundimento espacial e temporal. 2) Para alguns modelos poucas opções de função de ligação estão disponíveis na literatura tornando menos flexível a modelagem em alguns casos.

Para contornar tais problemas realizou-se a generalização da metodologia dos *TGMRF's* para o caso espaço-temporal. Da forma como esses modelos são obtidos, os efeitos latentes e fixos não se misturam na modelagem evitando assim os confundimentos espacial e temporal. Além disso, pode-se escolher distribuições adequadas para a marginal da média não se limitando às distribuições fornecidas através da escolha das funções de ligação. De forma similar aos *GLMM's* os *TGMRF's* necessitam de uma estrutura adequada para governar os efeitos espacial, temporal e espaço-temporal. Essa estrutura deve ser não separável e precisa ser um Campo Aleatório Gaussiano.

A estrutura de dependência adotada neste trabalho atende tais requisitos e ainda fornece uma interpretação simples dos parâmetros relacionados ao espaço, tempo e espaço-tempo. Em contrapartida, não é trivial definir o domínio de tais parâmetros para assegurar inversibilidade da matriz de precisão. Desta forma, criou-se limites para o domínio de cada parâmetro, com base no critério da dominância diagonal.

No estudo de simulação optou-se por explorar o modelo Poisson, apesar da metodologia valer para qualquer distribuição pertencente à família exponencial. Em cada cenário proposto verificou-se através de quatro métodos de seleção de modelos qual o critério mais eficaz dado que se conhece o verdadeiro modelo. Esse estudo serviu para identificar que o critério $WAIC_1$ é o mais eficaz dentre os testados.

Pôde-se perceber no estudo de simulação que os parâmetros de dependência são subestimados pelo método. É necessário um estudo para entender a relação entre estes parâmetros visto que um valor muito alto de um deles implica em valores próximos a zero dos demais. Obteve-se estimativas pontuais adequadas para os coeficientes de regressão. Porém, os intervalos *HPD* apresentam baixa amplitude o que reflete uma taxa de cobertura abaixo da taxa nominal. Pensando nisso é razoável estudar outros métodos de amostragem para observar se a taxa de cobertura dos intervalos *HPD* se torna mais adequada.

Como aplicação da metodologia em estudo, investigou-se as características do

habitat ideal da espécie de gastrópode *Nenia Tridens* através de dois modelos que foram comparados segundo o critério de ajuste $WAIC_1$. Dado o modelo vencedor encontrou-se uma dependência temporal alta no conjunto de dados. Além disso, três covariáveis foram consideradas importantes para o modelo as quais caracterizam um *habitat* úmido e fechado, ideal para o desenvolvimento da espécie.

Além das melhorias antes citadas, pretende-se testar outras estruturas de dependência em que seja mais simples identificar o domínio dos parâmetros sem perder a interpretabilidade dos mesmos. Além disso, dada a relação entre os parâmetros de dependência pode-se pensar em outras formas de blocagem para o *MCMC* o que pode melhorar o *mixing* das cadeias e possivelmente a amostragem destes parâmetros. Neste trabalho utilizou-se como critério de seleção de covariáveis a inclusão ou não do valor zero no intervalo *HPD*. Uma possível melhoria na abordagem seria a possibilidade do pesquisador utilizar seleção de variáveis via mistura, em que se observa a probabilidade de inclusão *a posteriori* das covariáveis. É razoável se pensar na implementação de um pacote em linguagem R para que a metodologia possa ser explorada por outros pesquisadores.

13 Apêndice - Estudo de simulação

Medida	Modelo verdadeiro	Modelo ajustado				Número de bases	Taxa de acerto
		LN	GI	GSH	GSC		
DIC	LN	27	0	3	0	30	90.00
DIC	GI	0	45	1	0	46	97.83
DIC	GSH	0	8	32	0	40	80.00
DIC	GSC	7	0	14	24	45	53.33
DIC	-	-	-	-	-	-	79.50
LPML	LN	30	0	0	0	30	100.00
LPML	GI	15	23	8	0	46	50.00
LPML	GSH	38	0	2	0	40	5.00
LPML	GSC	45	0	0	0	45	0.00
LPML	-	-	-	-	-	-	34.16
$WAIC_1$	LN	23	0	6	1	30	76.67
$WAIC_1$	GI	0	41	4	1	46	89.13
$WAIC_1$	GSH	0	1	39	0	40	97.50
$WAIC_1$	GSC	4	0	21	20	45	44.44
$WAIC_1$	-	-	-	-	-	-	76.40
$WAIC_2$	LN	29	0	1	0	30	96.67
$WAIC_2$	GI	0	43	2	1	46	93.48
$WAIC_2$	GSH	7	2	31	0	40	77.50
$WAIC_2$	GSC	28	0	2	15	45	33.33
$WAIC_2$	-	-	-	-	-	-	73.29

Tabela 9 – Cenário 2: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Modelo Verdadeiro	Parâmetro	Valor verdadeiro	Modelo de ajuste															
			LN				GI				GSH				GSC			
			Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %
LN	β_0	1.00	0.99	0.19	0.83	71	0.95	0.22	0.74	74	1.70	0.17	0.15	80	1.47	0.27	0.53	77
LN	β_1	0.70	0.71	0.09	0.92		0.10	0.12	0.07		0.53	0.07	0.34		0.65	0.10	0.81	
LN	β_2	0.30	0.26	0.21	0.92		0.06	0.20	0.59		0.23	0.17	0.88		0.37	0.23	0.90	
LN	ν	3.00	7.83	0.84	0.00		9.85	0.17	0.00		8.37	0.52	0.09		1.56	0.48	0.12	
LN	ρ_s	0.00	-0.04	0.84	1.00		0.10	0.77	1.00		-0.06	0.78	1.00		-0.03	0.77	1.00	
LN	ρ_{st}	0.00	-0.03	0.55	1.00		0.11	0.52	1.00		-0.04	0.53	1.00		-0.02	0.53	1.00	
LN	ρ_t	3.99	1.97	0.90	0.00		1.41	0.81	0.00		1.47	0.77	0.00		1.46	0.74	0.00	
GI	β_0	1.00	0.89	0.13	0.82	83	0.98	0.12	0.86	84	0.96	0.13	0.86	80	0.98	0.13	0.82	80
GI	β_1	0.70	0.72	0.05	0.92		0.70	0.05	0.86		0.72	0.05	0.89		0.70	0.05	0.89	
GI	β_2	0.30	0.33	0.13	0.96		0.32	0.12	0.99		0.34	0.13	0.99		0.33	0.13	0.96	
GI	ν	3.00	1.54	0.51	0.29		3.00	0.93	0.83		0.92	0.24	0.00		0.20	0.05	0.00	
GI	ρ_s	0.00	0.00	0.73	1.00		0.01	0.71	1.00		-0.02	0.72	1.00		-0.01	0.71	1.00	
GI	ρ_{st}	0.00	0.00	0.51	1.00		-0.01	0.50	1.00		-0.01	0.50	1.00		-0.01	0.50	1.00	
GI	ρ_t	3.99	0.63	1.04	0.00		0.71	0.99	0.00		0.75	1.00	0.00		0.64	1.01	0.00	
GSH	β_0	1.00	0.64	0.19	0.49	80	1.05	0.13	0.89	82	0.97	0.17	0.92	79	0.98	0.18	0.89	80
GSH	β_1	0.70	0.80	0.08	0.76		0.63	0.05	0.62		0.70	0.06	0.94		0.72	0.08	0.92	
GSH	β_2	0.30	0.34	0.20	0.95		0.26	0.13	0.89		0.33	0.17	0.96		0.30	0.19	0.95	
GSH	ν	3.00	5.75	1.10	0.24		8.79	0.87	0.01		3.01	0.58	0.94		0.73	0.13	0.00	
GSH	ρ_s	0.00	-0.03	0.76	1.00		-0.05	0.73	1.00		-0.04	0.75	1.00		-0.04	0.74	1.00	
GSH	ρ_{st}	0.00	-0.01	0.51	1.00		-0.01	0.50	1.00		-0.01	0.51	1.00		-0.02	0.50	1.00	
GSH	ρ_t	3.99	1.59	0.90	0.00		1.43	0.86	0.00		1.55	0.84	0.00		1.47	0.85	0.00	
GSC	β_0	1.00	0.85	0.14	0.83	86	1.16	0.12	0.61	82	1.08	0.14	0.89	80	1.00	0.14	0.90	83
GSC	β_1	0.70	0.71	0.06	0.97		0.56	0.05	0.32		0.65	0.05	0.76		0.70	0.06	0.99	
GSC	β_2	0.30	0.31	0.15	0.97		0.23	0.12	0.87		0.26	0.14	0.95		0.30	0.15	0.96	
GSC	ν	0.33	2.61	0.61	0.01		5.99	1.33	0.04		1.57	0.34	0.02		0.33	0.07	0.88	
GSC	ρ_s	0.00	-0.00	0.73	1.00		0.01	0.71	1.00		0.02	0.71	1.00		0.04	0.71	1.00	
GSC	ρ_{st}	0.00	-0.01	0.51	1.00		-0.01	0.49	1.00		-0.01	0.49	1.00		-0.00	0.50	1.00	
GSC	ρ_t	3.99	0.97	0.98	0.00		1.02	0.93	0.00		1.08	0.92	0.00		1.00	0.95	0.00	

Tabela 10 – Cenário 2: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas.

Medida	Modelo verdadeiro	Modelo ajustado				Número de bases	Taxa de acerto
		LN	GI	GSH	GSC		
DIC	LN	44	0	0	0	44	100.00
DIC	GI	0	51	0	0	51	100.00
DIC	GSH	16	13	13	0	42	30.95
DIC	GSC	46	0	1	0	47	0.00
DIC	-	-	-	-	-	-	58.70
LPML	LN	44	0	0	0	44	100.00
LPML	GI	47	3	1	0	51	5.88
LPML	GSH	42	0	0	0	42	0.00
LPML	GSC	47	0	0	0	47	0.00
LPML	-	-	-	-	-	-	25.54
$WAIC_1$	LN	44	0	0	0	44	100.00
$WAIC_1$	GI	0	48	3	0	51	94.12
$WAIC_1$	GSH	9	2	31	0	42	73.81
$WAIC_1$	GSC	42	0	3	2	47	4.26
$WAIC_1$	-	-	-	-	-	-	67.93
$WAIC_2$	LN	44	0	0	0	44	100.00
$WAIC_2$	GI	2	47	2	0	51	92.16
$WAIC_2$	GSH	37	2	3	0	42	7.14
$WAIC_2$	GSC	47	0	0	0	47	0.00
$WAIC_2$	-	-	-	-	-	-	51.09

Tabela 11 – Cenário 3: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Modelo Verdadeiro	Parâmetro	Valor verdadeiro	Modelo de ajuste															
			LN				GI				GSH				GSC			
			Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %
LN	β_0	1.00	0.94	0.21	0.71	83	1.11	0.17	0.72	81	1.65	0.19	0.23	81	1.11	0.26	0.91	74
LN	β_1	0.70	0.71	0.07	1.00		0.45	0.08	0.16		0.54	0.06	0.21		0.69	0.07	0.97	
LN	β_2	0.30	0.32	0.17	0.95		0.16	0.13	0.63		0.25	0.13	0.90		0.33	0.16	0.93	
LN	ν	3.00	3.54	0.59	0.89		9.75	0.25	0.00		6.68	0.96	0.20		1.00	0.25	0.00	
LN	ρ_s	2.24	2.07	0.06	0.00		2.04	0.05	0.00		2.05	0.06	0.00		2.06	0.06	0.00	
LN	ρ_{st}	0.00	-0.03	0.28	1.00		0.01	0.30	1.00		-0.00	0.26	1.00		0.00	0.26	1.00	
LN	ρ_t	0.00	0.04	0.95	1.00		-0.01	1.01	1.00		-0.04	0.84	1.00		-0.02	0.87	1.00	
GI	β_0	1.00	0.88	0.13	0.74	87	1.02	0.14	0.79	80	0.97	0.15	0.78	85	1.00	0.13	0.65	84
GI	β_1	0.70	0.75	0.05	0.55		0.71	0.05	0.82		0.73	0.05	0.71		0.72	0.05	0.56	
GI	β_2	0.30	0.30	0.12	0.90		0.28	0.11	0.92		0.30	0.12	0.94		0.28	0.12	0.90	
GI	ν	3.00	0.82	0.25	0.01		3.05	1.05	0.86		0.95	0.28	0.01		0.21	0.06	0.00	
GI	ρ_s	2.24	2.05	0.06	0.00		2.06	0.06	0.00		2.05	0.06	0.00		2.05	0.06	0.00	
GI	ρ_{st}	0.00	-0.02	0.30	1.00		-0.02	0.29	1.00		-0.02	0.29	1.00		-0.02	0.29	1.00	
GI	ρ_t	0.00	0.03	1.02	1.00		0.03	1.01	1.00		0.04	0.99	1.00		0.04	1.01	1.00	
GSH	β_0	1.00	0.63	0.20	0.53	78	1.09	0.14	0.73	82	0.97	0.19	0.75	77	0.88	0.20	0.72	87
GSH	β_1	0.70	0.80	0.07	0.69		0.62	0.05	0.59		0.71	0.06	0.84		0.73	0.06	0.78	
GSH	β_2	0.30	0.35	0.16	0.95		0.28	0.11	0.91		0.33	0.14	0.99		0.33	0.16	0.97	
GSH	ν	3.00	2.67	0.55	0.71		8.35	1.09	0.07		2.56	0.59	0.82		0.65	0.15	0.00	
GSH	ρ_s	2.24	2.06	0.06	0.00		2.06	0.06	0.00		2.06	0.06	0.00		2.06	0.06	0.00	
GSH	ρ_{st}	0.00	-0.02	0.28	1.00		-0.02	0.28	1.00		-0.02	0.27	1.00		-0.01	0.27	1.00	
GSH	ρ_t	0.00	0.01	0.96	1.00		-0.01	0.96	1.00		0.01	0.92	1.00		-0.01	0.92	1.00	
GSC	β_0	1.00	0.85	0.15	0.78	81	1.23	0.13	0.51	84	1.07	0.16	0.79	76	0.98	0.15	0.84	82
GSC	β_1	0.70	0.71	0.05	0.98		0.54	0.05	0.25		0.65	0.05	0.83		0.70	0.05	0.98	
GSC	β_2	0.30	0.30	0.13	0.95		0.24	0.10	0.83		0.29	0.12	0.89		0.29	0.13	0.93	
GSC	ν	0.33	1.31	0.31	0.02		6.24	1.44	0.02		1.45	0.38	0.04		0.33	0.09	0.91	
GSC	ρ_s	2.24	2.06	0.06	0.00		2.05	0.06	0.00		2.05	0.06	0.00		2.06	0.06	0.00	
GSC	ρ_{st}	0.00	-0.02	0.29	1.00		-0.02	0.28	1.00		-0.02	0.28	1.00		-0.02	0.28	1.00	
GSC	ρ_t	0.00	-0.00	0.99	1.00		0.00	0.98	1.00		0.00	0.98	1.00		0.01	0.97	1.00	

Tabela 12 – Cenário 3: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas.

Medida	Modelo verdadeiro	Modelo ajustado				Número de bases	Taxa de acerto
		LN	GI	GSH	GSC		
DIC	LN	40	0	0	0	40	100.00
DIC	GI	0	49	3	0	52	94.23
DIC	GSH	5	17	27	0	49	55.10
DIC	GSC	29	0	9	5	43	11.63
DIC	-	-	-	-	-	-	65.76
LPML	LN	40	0	0	0	40	100.00
LPML	GI	20	20	12	0	52	38.46
LPML	GSH	49	0	0	0	49	0.00
LPML	GSC	42	0	1	0	43	0.00
LPML	-	-	-	-	-	-	32.61
$WAIC_1$	LN	40	0	0	0	40	100.00
$WAIC_1$	GI	0	40	12	0	52	76.92
$WAIC_1$	GSH	3	2	44	0	49	89.80
$WAIC_1$	GSC	9	0	13	21	43	48.84
$WAIC_1$	-	-	-	-	-	-	78.80
$WAIC_2$	LN	40	0	0	0	40	100.00
$WAIC_2$	GI	1	43	8	0	52	82.69
$WAIC_2$	GSH	27	3	19	0	49	38.78
$WAIC_2$	GSC	35	0	2	6	43	13.95
$WAIC_2$	-	-	-	-	-	-	58.70

Tabela 13 – Cenário 4: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Modelo Verdadeiro	Parâmetro	Valor verdadeiro	Modelo de ajuste															
			LN				GI				GSH				GSC			
			Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %
LN	β_0	1.00	0.96	0.18	0.88	76	1.12	0.21	0.79	82	1.75	0.18	0.10	77	0.51	0.52	0.86	70
LN	β_1	0.70	0.69	0.07	0.91		0.19	0.10	0.09		0.54	0.06	0.19		0.62	0.08	0.73	
LN	β_2	0.30	0.31	0.17	0.92		0.06	0.15	0.48		0.23	0.14	0.81		0.40	0.20	0.79	
LN	ν	3.00	3.52	0.61	0.80		9.78	0.26	0.00		8.30	0.84	0.06		10.77	7.07	0.21	
LN	ρ_s	0.00	0.02	0.63	1.00		0.03	0.69	1.00		0.04	0.65	1.00		0.01	0.63	1.00	
LN	ρ_{st}	0.00	-0.01	0.34	1.00		0.02	0.37	1.00		-0.03	0.35	1.00		0.04	0.34	1.00	
LN	ρ_t	3.99	3.86	0.07	0.55		3.82	0.06	0.00		3.82	0.07	0.00		3.81	0.06	0.00	
GI	β_0	1.00	0.92	0.12	0.84	90	1.03	0.13	0.95	88	0.98	0.14	0.96	81	0.99	0.13	0.88	86
GI	β_1	0.70	0.72	0.05	0.73		0.68	0.05	0.90		0.72	0.05	0.85		0.71	0.05	0.72	
GI	β_2	0.30	0.29	0.12	0.97		0.29	0.11	0.99		0.31	0.12	0.99		0.30	0.13	0.95	
GI	ν	3.00	0.72	0.27	0.00		3.39	1.14	0.94		1.00	0.30	0.00		0.22	0.06	0.00	
GI	ρ_s	0.00	0.01	0.66	1.00		-0.00	0.65	1.00		0.01	0.64	1.00		-0.00	0.64	1.00	
GI	ρ_{st}	0.00	-0.01	0.35	1.00		-0.01	0.35	1.00		-0.01	0.34	1.00		-0.00	0.35	1.00	
GI	ρ_t	3.99	3.81	0.07	0.00		3.83	0.07	0.00		3.82	0.07	0.00		3.82	0.07	0.00	
GSH	β_0	1.00	0.66	0.19	0.65	89	1.05	0.13	0.82	79	0.98	0.18	0.95	81	0.89	0.19	0.88	83
GSH	β_1	0.70	0.79	0.07	0.75		0.62	0.05	0.62		0.70	0.06	0.96		0.72	0.07	0.87	
GSH	β_2	0.30	0.34	0.18	0.90		0.26	0.12	0.81		0.30	0.15	0.93		0.33	0.17	0.89	
GSH	ν	3.00	3.32	0.72	0.84		8.67	0.97	0.01		3.15	0.70	0.94		0.78	0.16	0.00	
GSH	ρ_s	0.00	0.04	0.63	1.00		0.03	0.61	1.00		0.05	0.59	1.00		0.06	0.60	1.00	
GSH	ρ_{st}	0.00	-0.03	0.34	1.00		-0.03	0.33	1.00		-0.03	0.32	1.00		-0.04	0.32	1.00	
GSH	ρ_t	3.99	3.81	0.07	0.01		3.83	0.07	0.00		3.83	0.07	0.00		3.83	0.07	0.00	
GSC	β_0	1.00	0.87	0.15	0.85	79	1.21	0.12	0.56	84	1.07	0.15	0.86	81	1.01	0.15	0.95	82
GSC	β_1	0.70	0.69	0.06	0.96		0.54	0.05	0.20		0.64	0.05	0.75		0.69	0.06	0.95	
GSC	β_2	0.30	0.30	0.14	0.95		0.22	0.11	0.81		0.27	0.13	0.94		0.28	0.14	0.95	
GSC	ν	0.33	1.54	0.39	0.04		6.81	1.40	0.00		1.77	0.44	0.00		0.38	0.09	0.89	
GSC	ρ_s	0.00	0.03	0.65	1.00		0.02	0.63	1.00		-0.00	0.63	1.00		0.05	0.62	1.00	
GSC	ρ_{st}	0.00	-0.01	0.35	1.00		0.00	0.34	1.00		0.01	0.34	1.00		-0.02	0.33	1.00	
GSC	ρ_t	3.99	3.81	0.07	0.00		3.82	0.07	0.00		3.82	0.07	0.00		3.83	0.07	0.00	

Tabela 14 – Cenário 4: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas.

Medida	Modelo verdadeiro	Modelo ajustado				Número de bases	Taxa de acerto
		LN	GI	GSH	GSC		
DIC	LN	2	0	0	0	2	100.00
DIC	GI	0	4	0	0	4	100.00
DIC	GSH	0	4	5	0	9	55.56
DIC	GSC	0	0	0	2	2	100.00
DIC	-	-	-	-	-	-	76.47
LPML	LN	2	0	0	0	2	100.00
LPML	GI	2	2	0	0	4	50.00
LPML	GSH	8	0	1	0	9	11.11
LPML	GSC	2	0	0	0	2	0.00
LPML	-	-	-	-	-	-	29.41
$WAIC_1$	LN	2	0	0	0	2	100.00
$WAIC_1$	GI	0	4	0	0	4	100.00
$WAIC_1$	GSH	0	0	9	0	9	100.00
$WAIC_1$	GSC	0	0	0	2	2	100.00
$WAIC_1$	-	-	-	-	-	-	100.00
$WAIC_2$	LN	2	0	0	0	2	100.00
$WAIC_2$	GI	0	4	0	0	4	100.00
$WAIC_2$	GSH	0	1	8	0	9	88.89
$WAIC_2$	GSC	2	0	0	0	2	0.00
$WAIC_2$	-	-	-	-	-	-	82.35

Tabela 15 – Cenário 6: Número de vezes em que o modelo foi selecionado corretamente segundo os critérios *DIC, LPML,* $WAIC_1$ e $WAIC_2$.

Modelo Verdadeiro	Parâmetro	Valor verdadeiro	Modelo de ajuste															
			LN				GI				GSH				GSC			
			Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %	Média	Desvio padrão	Cob	NB %
LN	β_0	1.00	0.85	0.12	0.54	39	1.43	0.10	0.12	42	1.77	0.12	0.09	35	1.25	0.28	0.29	38
LN	β_1	0.70	0.71	0.12	0.90		0.39	0.08	0.10		0.52	0.09	0.43		0.62	0.16	0.95	
LN	β_3^1	-1.21	-1.26	0.60	0.95		-0.74	0.30	0.31		-0.98	0.32	0.57		-1.37	0.59	0.92	
LN	β_3^{20}	0.63	0.65	0.46	0.97		0.53	0.26	0.64		0.44	0.32	0.83		0.84	0.57	0.84	
LN	ν	3.00	9.98	1.74	0.00		28.24	1.28	0.00		12.78	1.84	0.06		4.57	0.47	0.13	
LN	ρ_s	0.00	-0.01	0.81	1.00		-0.04	0.73	1.00		-0.01	0.76	1.00		0.06	0.75	1.00	
LN	ρ_{st}	0.00	-0.04	0.54	1.00		0.02	0.50	1.00		-0.04	0.52	1.00		0.05	0.52	1.00	
LN	ρ_t	3.99	1.80	0.91	0.00		1.25	0.76	0.00		1.38	0.74	0.00		1.28	0.75	0.00	
GI	β_0	1.00	0.79	0.08	0.37	49	0.93	0.08	0.62	45	0.88	0.08	0.48	42	0.93	0.07	0.64	50
GI	β_1	0.70	0.78	0.08	0.86		0.72	0.06	0.84		0.75	0.07	0.81		0.76	0.08	0.86	
GI	β_3^1	-1.21	-1.39	0.40	0.94		-1.25	0.23	0.87		-1.26	0.24	0.88		-1.26	0.33	0.92	
GI	$\beta_3^2 20$	0.63	0.69	0.29	0.96		0.65	0.22	0.96		0.65	0.25	0.98		0.63	0.30	0.98	
GI	ν	3.00	2.58	0.82	0.55		2.51	0.89	0.58		1.11	0.29	0.05		0.26	0.08	0.00	
GI	ρ_s	0.00	0.03	0.73	1.00		-0.01	0.71	1.00		0.03	0.71	1.00		0.05	0.72	1.00	
GI	ρ_{st}	0.00	-0.02	0.51	1.00		0.01	0.50	1.00		-0.03	0.50	1.00		0.01	0.50	1.00	
GI	ρ_t	3.99	0.52	1.02	0.00		0.54	1.01	0.00		0.71	0.97	0.00		0.43	1.02	0.00	
GSH	β_0	1.00	0.47	0.12	0.00	41	1.09	0.11	0.64	53	0.88	0.11	0.67	54	0.88	0.11	0.67	46
GSH	β_1	0.70	0.86	0.12	0.76		0.57	0.07	0.53		0.72	0.09	0.96		0.76	0.11	0.89	
GSH	β_3^1	-1.21	-1.57	0.59	0.98		-1.03	0.24	0.68		-1.23	0.29	0.94		-1.35	0.47	0.98	
GSH	β_3^{20}	0.63	0.89	0.47	0.93		0.55	0.27	0.91		0.67	0.33	0.94		0.85	0.47	0.96	
GSH	ν	3.00	7.65	1.78	0.17		11.41	3.31	0.21		2.84	0.59	0.81		0.87	0.18	0.00	
GSH	ρ_s	0.00	-0.03	0.74	1.00		-0.06	0.71	1.00		-0.05	0.71	1.00		-0.02	0.71	1.00	
GSH	ρ_{st}	0.00	-0.05	0.51	1.00		-0.00	0.50	1.00		-0.01	0.50	1.00		-0.02	0.50	1.00	
GSH	ρ_t	3.99	1.10	0.94	0.00		1.11	0.87	0.00		1.20	0.88	0.00		1.03	0.90	0.00	
GSC	β_0	1.00	0.80	0.08	0.37	41	1.06	0.08	0.67	43	0.96	0.08	0.89	55	0.96	0.08	0.87	38
GSC	β_1	0.70	0.72	0.08	0.93		0.62	0.07	0.65		0.68	0.07	0.89		0.72	0.08	1.00	
GSC	β_3^1	-1.21	-1.20	0.41	0.90		-1.10	0.22	0.53		-1.20	0.27	0.75		-1.30	0.37	0.95	
GSC	β_3^{20}	0.63	0.70	0.34	0.90		0.57	0.22	0.86		0.65	0.28	0.89		0.58	0.32	0.97	
GSC	ν	0.33	2.93	0.77	0.05		4.46	1.71	0.30		1.37	0.34	0.05		0.35	0.08	0.89	
GSC	ρ_s	0.00	0.00	0.73	1.00		-0.04	0.69	1.00		-0.03	0.70	1.00		-0.06	0.72	1.00	
GSC	ρ_{st}	0.00	-0.03	0.51	1.00		-0.01	0.49	1.00		-0.00	0.49	1.00		-0.01	0.49	1.00	
GSC	ρ_t	3.99	0.70	1.02	0.00		0.56	0.99	0.00		0.73	0.96	0.00		0.65	0.97	0.00	

Tabela 16 – Cenário 6: Média e desvio padrão *a posteriori*, taxa de cobertura e número de bases utilizadas.

Referências

ALBERT, J. *Bayesian Computation with R*. [S.l.]: Springer Science & Business Media, 2009. Citado na página 38.

BANERJEE, S.; CARLIN, B. P.; GELFAND, A. E. *Hierarchical modeling and analysis for spatial data*. [S.l.]: Crc Press, 2014. Citado na página 18.

BESAG, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, JSTOR, p. 192–236, 1974. Citado na página 11.

BRESLOW, N. E.; CLAYTON, D. G. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 88, n. 421, p. 9–25, 1993. Citado 2 vezes nas páginas 10 e 13.

BROOK, D. On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. *Biometrika*, JSTOR, p. 481–483, 1964. Citado na página 16.

CARLIN, B. P.; BANERJEE, S. Hierarchical multivariate car models for spatio-temporally correlated survival data. *Bayesian Statistics*, Citeseer, v. 7, p. 45–63, 2003. Citado 2 vezes nas páginas 11 e 20.

CHRISTENSEN, R. et al. *Bayesian ideas and data analysis: an introduction for scientists and statisticians*. [S.l.]: CRC Press, 2011. Citado na página 46.

CZADO, C.; SANTNER, T. J. The effect of link misspecification on binary regression inference. *Journal of Statistical Planning and Inference*, Elsevier, v. 33, n. 2, p. 213–231, 1992. Citado 2 vezes nas páginas 11 e 13.

GEISSER, S.; EDDY, W. F. A predictive approach to model selection. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 74, n. 365, p. 153–160, 1979. Citado na página 48.

GELFAND, A. E.; DEY, D. K. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, JSTOR, p. 501–514, 1994. Citado na página 48.

GELFAND, A. E.; SMITH, A. F. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 85, n. 410, p. 398–409, 1990. Citado na página 39.

GELFAND, A. E.; VOUNATSOU, P. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics*, Biometrika Trust, v. 4, n. 1, p. 11–15, 2003. Citado 2 vezes nas páginas 11 e 20.

GELMAN, A.; HWANG, J.; VEHTARI, A. Understanding predictive information criteria for bayesian models. *Statistics and Computing*, Springer, v. 24, n. 6, p. 997–1016, 2014. Citado 3 vezes nas páginas 46, 47 e 48.

GEWEKE, J. et al. *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. [S.l.]: Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991. v. 196. Citado na página 51.

GILKS, W. R.; BEST, N.; TAN, K. Adaptive rejection metropolis sampling within gibbs sampling. *Applied Statistics*, JSTOR, p. 455–472, 1995. Citado na página 35.

GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. Introducing markov chain monte carlo. *Markov chain Monte Carlo in Practice*, London: Chapman and Hall, v. 1, p. 19, 1996. Citado na página 38.

GILKS, W. R.; WILD, P. Adaptive rejection sampling for gibbs sampling. *Applied Statistics*, JSTOR, p. 337–348, 1992. Citado na página 41.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, Biometrika Trust, v. 57, n. 1, p. 97–109, 1970. Citado na página 40.

JIN, X.; CARLIN, B. P.; BANERJEE, S. Generalized hierarchical multivariate car models for areal data. *Biometrics*, Wiley Online Library, v. 61, n. 4, p. 950–961, 2005. Citado na página 11.

LI, K.-C.; DUAN, N. Regression analysis under link violation. *The Annals of Statistics*, JSTOR, p. 1009–1052, 1989. Citado 2 vezes nas páginas 11 e 13.

PRATES, M. O. *Link Specification and Spatial Dependence for Generalized Linear Mixed Models*. [S.l.]: University of Connecticut, 2011. Citado na página 27.

PRATES, M. O. et al. Transformed gaussian markov random fields and spatial modeling of species abundance. *Spatial Statistics*, Elsevier, 2015. Citado 5 vezes nas páginas 11, 12, 35, 55 e 59.

REICH, B. J.; HODGES, J. S.; ZADNIK, V. Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models. *Biometrics*, Wiley Online Library, v. 62, n. 4, p. 1197–1206, 2006. Citado na página 13.

ROBERTS, G. O.; ROSENTHAL, J. S. et al. Optimal scaling for various metropolis-hastings algorithms. *Statistical Science*, Institute of Mathematical Statistics, v. 16, n. 4, p. 351–367, 2001. Citado na página 40.

RODRIGUES, E. C. *Estruturas de Covariância em Modelos Espaciais Bayesianos*. [S.l.]: Universidade Federal de Minas Gerais, 2012. Citado 6 vezes nas páginas 7, 11, 12, 22, 25 e 31.

RUE, H.; HELD, L. *Gaussian Markov random fields: theory and applications*. [S.l.]: CRC Press, 2005. Citado 2 vezes nas páginas 25 e 27.

SAIN, S. R. et al. A spatial analysis of multivariate output from regional climate models. *The Annals of Applied Statistics*, Institute of Mathematical Statistics, v. 5, n. 1, p. 150–175, 2011. Citado 3 vezes nas páginas 11, 21 e 36.

SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002. Citado na página 46.

STROUP, W. W. *Generalized linear mixed models: modern concepts, methods and applications*. [S.l.]: CRC press, 2012. Citado na página 13.

WATANABE, S. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *The Journal of Machine Learning Research*, JMLR. org, v. 11, p. 3571–3594, 2010. Citado na página 47.