

Érika Fialho Morais Xavier

**Abordagem bayesiana para o processo
espaço-temporal log gaussiano de Cox com
aplicação no setor florestal**

Recife, PE

Abril de 2013



Universidade Federal Rural de Pernambuco
Pró-Reitoria de Pesquisa e Pós-Graduação
Programa de Pós-Graduação em Biometria e Estatística Aplicada

Érika Fialho Morais Xavier

**Abordagem bayesiana para o processo
espaço-temporal log gaussiano de Cox com
aplicação no setor florestal**

Recife, PE
Abril de 2013.

Érika Fialho Morais Xavier

Abordagem bayesiana para o processo espaço-temporal log gaussiano de Cox com aplicação no setor florestal

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada, em cumprimento das exigências legais para obtenção do título de Mestre.

Área de concentração: Biometria e Estatística Aplicada

Orientador:

Prof. Dr. Eufrázio de Souza Santos

Co-orientador:

Prof. Dr. Moacyr Cunha Filho

Recife, PE

Abril de 2013.

Ficha catalográfica

X3a

Xavier, Érika Fialho Morais.

Abordagem bayesiana para o processo espaço-temporal log gaussiano de Cox com aplicação no setor florestal / Érika Fialho Morais Xavier. - Recife, 2013.

60 f.: il.

Orientador: Eufrázio de Souza Santos.

Dissertação (Mestrado em Biometria e Estatística Aplicada) - Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática, Recife, 2013.

Referências.

1. LGCP 2. MCMC 3. Pontos de queimadas 4. Amazônia
I. Santos, Eufrázio de Souza, orientador II. Título

CDD 310

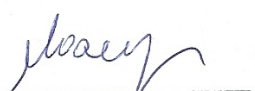
Érika Fialho Morais Xavier

Abordagem bayesiana para o processo espaço-temporal log gaussiano de Cox com aplicação no setor florestal

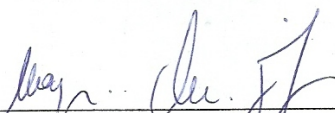
Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 10/04/2013 pela Comissão Examinadora.

Aprovado em: 10 / 04 / 2013

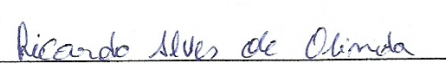
Comissão Examinadora:



Prof. Dr. Eufrázio de Souza Santos
Orientador
DEINFO - UFRPE



Prof. Dr. Moacyr Cunha Filho
Co-orientador
DEINFO - UFRPE



Prof. Dr. Ricardo Alves de Olinda
DE - UEPB



Prof. Dr. Humber Agrelli de Andrade
DEPAq - UFRPE

Dedico a Pedro (*in memoriam*) e Salete,
meus pais. E a Sílvio, meu esposo.

Agradecimentos

Agradeço primeiramente a Deus. Em Sua infinita misericórdia Ele conduziu os meus caminhos, me deu sabedoria e paciência, me fez aprender muito sobre a vida durante este tempo de trabalho e dedicação, esteve comigo em todos os momentos, me deu proteção e amparo. Sem o meu Deus eu nada seria, não poderia dar um passo sequer, não seria capaz de sair de minha cidade de origem para estudar em Recife e, principalmente, não conseguiria concluir este trabalho. A Ele toda honra e toda glória eternamente.

Agradeço aos amigos que conquistei durante o mestrado, em especial Neto, Priscilla, Renata, Mary, Danila, Clara, Gabriela e, sempre comigo, minha irmã Nyedja. Agradeço a estes por cada dia de estudo, por cada sorriso, por cada lágrima, pelo companheirismo e acima de tudo pela amizade construída.

Agradeço ao apoio dado por meus familiares de Recife, em especial a minhas tias Teca (*in memoriam*) e Penha, minha cunhada Fernanda, minha sogra Ana e meu sogro Sílvio. A presença de cada um destes foi fundamental para a minha conquista.

Agradeço aos meus familiares de Campina Grande, meu irmão Pedro e minha mãe Salete, que estão sempre ao meu lado, incondicionalmente. Agradeço por cada lágrima de saudade e cada sorriso no abraço dos reencontros.

Agradeço de maneira especial e com muito carinho ao meu esposo Sílvio que esteve comigo durante todo o meu percurso, que me apoiou, que acreditou em mim, que me animou e nunca me deixou desistir, e que muitas vezes estudou comigo e me ajudou didaticamente também. Agradeço a Deus por tê-lo ao meu lado.

Agradeço aos professores e funcionários do DEINFO, em especial aos professores Eufrázio e Moacyr, também ao sempre prestativo Marco.

Agradeço a CAPES pelo apoio financeiro.

“Porque o Senhor dá a sabedoria, da
Sua boca é que vem o conhecimento e
o entendimento.”

Provérbios 2:6

Resumo

Através da análise de Processos de Poisson tem sido possível realizar de forma satisfatória diversos estudos a partir de processos pontuais, com dados provenientes de contagem. Entretanto, estes processos limitam-se ao estudo de situações com padrões homogêneos, dificilmente encontrados em dados reais. Este trabalho propôs o estudo dos Processos Log Gaussianos de Cox (LGCP), processo que torna possível o estudo de dados com padrões pontuais heterogeneos a partir de uma generalização do processo de Poisson, baseado na realização de um campo aleatório Gaussiano. Foram realizadas duas aplicações para o processo, a primeira em dados simulados de focos de incêndio em Castilla-La Mancha, Reino da Espanha, com a finalidade de explorar as propriedades grafico-computacionais do LGCP, bem como a heterogeneidade proposta pelo processo. A segunda em dados reais de focos de calor e precipitação média de chuva no Bioma Amazônia, Brasil, detectados pelo satélite NOAA 15, entre os anos de 2007 e 2011. A inferência para esses processos é realizada sob a abordagem Bayesiana, utilizando o método de Monte Carlo em Cadeias de Markov (MCMC). Os objetivos propostos neste trabalho foram cumpridos de forma satisfatória, possibilitando previsões futuras a respeito dos dados em estudo.

Palavras-chave: LGCP, MCMC, Pontos de Queimadas, Amazônia.

Abstract

Through the analysis of Poisson processes has been possible to perform satisfactorily some studies with data point processes counting. However, these processes are limited to the study of situations with homogeneous patterns, hardly found in actual data. This research has proposed the study of Log Gaussian Cox Processes, process that makes possible the study of patterns points heterogeneous data, with a based from Poisson process with on the realization of a Gaussian random field. We did two applications for the process, the first with simulated data of outbreaks of fire in Castilla-La Mancha, Kingdom of Spain, in order to explore the properties of the graph and computational of LGCP, and study the heterogeneity proposed by the process. The second focuses on real data of fire points and average rainfall in the Amazon Biome, Brazil, detected by satellite NOAA 15, between the years 2007 and 2011. The Inference for these processes are carried out under the Bayesian approach, using the Monte Carlo Markov Chain (MCMC). The proposed objectives of this work were completed satisfactorily, enabling future predictions about the data in the study.

Keywords: LGCP, MCMC, Fire Points, Amazon.

Sumário

Lista de Figuras	p. 11
Lista de Tabelas	p. 13
1 Introdução	p. 14
2 Revisão de Literatura	p. 16
2.1 Teorema de Bayes	p. 19
2.2 Alguns métodos de análise Bayesiana	p. 19
2.2.1 Prioris conjugadas	p. 21
2.2.2 Priori não informativa de Bayes e Laplace	p. 21
2.2.3 Priori não informativa de Jeffreys	p. 22
2.2.4 Distribuição a Posteriori	p. 23
2.2.5 Estimadores de Bayes	p. 23
2.3 Métodos de Aproximação Bayesiana	p. 24
2.3.1 Método de Monte Carlo	p. 24
2.3.2 Monte Carlo via Função de Importância	p. 25
2.3.3 Método de Reamostragem Ponderada	p. 26
2.3.4 Monte Carlo via Cadeias de Markov	p. 27
2.3.4.1 Cadeias de Markov	p. 27
2.3.4.2 Algoritmo de Metropolis-Hastings	p. 28
2.3.4.3 Amostrador de Gibbs	p. 29
2.3.4.4 MCMC com Saltos Reversíveis (RJMCMC)	p. 30

3	Metodologia	p. 31
3.1	Processos Log Gaussianos de Cox	p. 32
3.1.1	Abordagem LGCP assumindo o tempo contínuo	p. 34
3.1.2	Abordagem LGCP assumindo o tempo discreto	p. 35
3.2	Análise Bayesiana	p. 36
3.2.1	Priori e posteriori para a abordagem no tempo contínuo	p. 37
3.2.2	Priori e posteriori para a abordagem no tempo discreto	p. 38
4	Resultados e discussão	p. 39
4.1	Aplicação em dados simulados	p. 40
4.2	Aplicação em dados reais	p. 44
4.2.1	Diagnóstico MCMC para os LGCP	p. 51
5	Conclusão	p. 55
	Referências	p. 57

Lista de Figuras

1	Thomas Bayes (1702 - 1761)	p. 16
2	Representação Gráfica de um Processo Estocástico como uma família de Variáveis Aleatórias. Fonte: (MORETTIN; TOLOI, 2004)	p. 27
3	Dados simulados referentes a Incêndios Florestais em Castilla-La Mancha, Reino da Espanha. (A) Padrões Pontuais, (B) Densidade estimada.	p. 40
4	Estimativa para a Densidade de $\Lambda(s)$ em conjunto com os pontos agregados.	p. 40
5	Padrões pontuais de acordo com as causas de cada incêndio em pontos agregados (A) e suavização para o processo (B).	p. 41
6	Estimativa para a média do processo $\mu = E[y(t)]$ (Número médio de focos de incêndio em relação ao tempo).	p. 42
7	Gráficos da Função de Covariância Exponencial para as estimativas $\sigma^2 = 1.3$ (em (A)) e $\phi = 2.3$ (em (B)).	p. 42
8	Estimativa para o parâmetro de correlação temporal $\theta = 3.4$	p. 43
9	Mapa brasileiro com as divisões dos Biomas existentes no país - Fonte: Portal Brasil, Ministério do Meio Ambiente.	p. 44
10	Densidade para os dados de quantidade de focos de calor e a precipitação média no Bioma Amazônia.	p. 45
11	Dados de focos de calor em padrões pontuais (A) e com a suavização do processo (B).	p. 45
12	Suavização do processo em conjunto com os pontos agregados de focos de calor e precipitação média.	p. 46
13	Estimativa para a média do processo $\mu = E[y(t)]$	p. 46
14	Gráficos da Função de Covariância Exponencial para as estimativas $\sigma^2 = 2.4$ (A) e $\phi = 6.0$ (B)	p. 47

15	Estimativa para o parâmetro $\theta = 2,1$	p. 47
16	Padrões pontuais e densidade estimada com um grid de dados 128×128	p. 48
17	Gráficos da Função de Covariância Exponencial para as estimativas $\sigma^2 = 1,5$ (A) e $\phi = 3,6$ (B).	p. 49
18	Estimativa para o parâmetro $\theta = 3,8$	p. 49
19	Gráfico para o risco relativo (A), os erros associados ao risco relativo (B) e a intensidade média de Poisson (C).	p. 50
20	Série temporal para a representação da Cadeia de Markov. (A) Série para os 20000 valores gerados pela cadeia, (B) série para uma amostra de cadeia com 1000 valores simulados.	p. 51
21	Gráficos para a função de autocorrelação. (A) Autocorrelação para os 20000 valores gerados pela cadeia, (B) autocorrelação para uma amostra de cadeia com 1000 valores simulados.	p. 52
22	Representação gráfica para a evolução da mediana no decorrer das iterações.	p. 52
23	Gráfico da estatística de Geweke para um dos conjuntos de cadeias.	p. 53
24	Gráfico da estatística de Gelman-Rubin para o conjunto de cadeias.	p. 53
25	Histograma para a quantidade de focos de calor no bioma Amazônia.	p. 54

Lista de Tabelas

1	Parâmetros estimatimados a partir do LGCP aplicado aos a pontos de incêndios em Castilla-La Mancha.	p. 43
2	Parâmetros Estimativados a partir do LGCP aplicado aos pontos de focos de calor no Bioma Amazônia.	p. 48
3	Novos parâmetros estimatimados para LGCP.	p. 50
4	Comparação entre os valores observados e previstos para a quantidade de focos de calor no Bioma Amazônia.	p. 50

1 Introdução

Processos de contagem são frequentemente encontrados na natureza e intuitivamente associados a distribuição de Poisson e ao processo de Poisson. Entretanto, um processo simples de Poisson limita-se ao estudo de processos pontuais homogêneos, dificilmente identificados em dados reais. Logo, naturalmente surge a necessidade de um estudo mais amplo e representativo de processos reais.

Moller *et al.* (1998), Diggle (2003), entre outros autores, propõem uma generalização para este processo: o processo de Cox, que nada mais é do que uma função de intensidade de um processo de Poisson como realização de um campo aleatório. Moller *et al.* (1998) acrescenta ainda que o processo de Cox é “duplamente estocástico” devido ao seu surgimento dado em função do processo de Poisson e da medida de intensidade aleatória. Estes processos são amplamente utilizados em modelagens de padrões pontuais por sua capacidade de reflexão e análise a respeito da heterogeneidade em dados ambientais reais.

Assumindo-se que o logaritmo da superfície de intensidade segue um processo gaussiano, é possível fazer uma aplicação dos processos de Cox a partir da distribuição Gaussiana da superfície de intensidade, obtendo-se, assim, uma rica modelagem de padrões pontuais. Com esta motivação Moller *et al.* (1996) levanta ideias e apresenta resultados de forma ainda mais definida, a partir de uma aplicação em um exemplo do setor florestal, propondo o modelo Log Gaussiano em um processo de Cox, ao qual propomos estudo.

O objetivo principal deste trabalho será analisar a aplicabilidade dos LGCP em dados provenientes de contagens, bem como estudar o processo fazendo-se uma abordagem a partir de métodos Bayesianos. Para isto, serão realizadas aplicações do Processo Log Gaussiano de Cox em dados de focos de calor e queimadas, usados muitas vezes como uma ferramenta na agricultura, porém responsáveis por alguns prejuízos como o empobrecimento do solo, a destruição da vegetação, interferências diretas em alterações na atmosfera e na saúde dos animais.

Organização dos Capítulos

Os Capítulos 1 e 2 correspondem a parte introdutória do trabalho. No Capítulo 1 discutiremos a proposta de estudo e no Capítulo 2 apresentaremos uma breve revisão literária, onde mostraremos alguns trabalhos já realizados juntamente com um breve resumo de suas propostas de estudo, bem como uma contextualização a respeito dos métodos Bayesianos de análise estatística. Discussões a respeito dos *Log Gaussian Cox Processes* serão vistas no Capítulo 3.

No Capítulo 4 serão apresentadas duas aplicações para os LGCP: Na Seção 4.1 o Processo Log Gaussiano de Cox será aplicado a dados simulados correspondentes a Focos de Incêndio em Castilla-La Mancha, Reino da Espanha. Esta aplicação nos mostrará as propriedades gráficas de análise do LGCP, onde será possível visualizar a heterogeneidade dos dados, que é o principal motivo de surgimento do processo.

A análise Bayesiana será feita com maior ênfase na Seção 4.2, onde o LGCP será aplicado em dados reais, provenientes do cruzamento entre as variáveis Focos de Calor e Precipitação Média no Bioma Amazônia, Brasil. Nesta aplicação utilizaremos o método MCMC tendo por finalidade a obtenção de boas estimativas para o processo. Por fim, o Capítulo 5 apresentará uma breve conclusão a respeito dos resultados obtidos, bem como a sugestão de estudos futuros.

2 Revisão de Literatura

A análise de Inferência Estatística pode ser realizada através de dois tipos diferentes de abordagem: a Inferência clássica (ou frequentista), que consiste em obter tais conclusões com base apenas na amostra aleatória, bem como a Inferência Bayesiana, que é fundamentada em uma “complementação” desta amostra, a partir de informações obtidas em experimentos passados. De maneira geral, segundo Pollard (1986), a aproximação bayesiana pode ser vista como uma extensão da clássica.

A origem dos estudos que envolvem a inferência Bayesiana é relacionada ao Reverendo Thomas Bayes (Figura 1).



Figura 1: Thomas Bayes (1702 - 1761)

Bayes nasceu por volta de 1702, em Londres. Foi um dos primeiros seis ministros “*Nonconformist*” ordenados na Inglaterra, pastor da Igreja Cristã Presbiteriana e matemático.

Em 1719, ingressou na Universidade de Edimburgo para estudar teologia e lógica, entretanto seus estudos tiveram continuidade em uma universidade escocesa em virtude de sua formação religiosa, pois, na época, os ministros dissidentes foram proibidos de frequentar as universidades de Oxford e Cambridge. Em 1742, foi eleito membro da *Royal Society*.

Bayes faleceu em 17 de abril de 1761 em Tunbridge Wells, Kent e foi sepultado em Londres.

Apesar de não ter publicado obras matemáticas, o Reverendo tornou-se conhecido por ter utilizado os conceitos de probabilidade de forma intuitiva, estabelecendo bases sólidas para o estudo da inferência estatística Bayesiana por meio da formulação do Teorema de Bayes. Pouco tempo após sua morte, em 1763, o amigo pessoal de Thomas Bayes, Richard Price, publicou a obra do Reverendo intitulada “*An essay Towards Solving a problem in the Doctrine of Chances*”, Bayes (1763), fazendo com que todos os estudos realizados até os dias de hoje a respeito da Inferência Bayesiana fossem possíveis.

Entre alguns estudos que utilizam a metodologia Bayesiana, podemos destacar, por exemplo, Bartolucci (2006). Em um estudo de regressão logística para o número de sobreviventes associado a um tratamento recebido contra uma certa doença e aos fatores externos, ele utilizou 5 possíveis modelos (M_1, M_2, M_3, M_4 e M_5) e concluiu que o modelo mais completo (M_5) possuía a forma de $y_{ij} \sim \text{bin}(n_{ij}, p_{ij})$, com y_{ij}, n_{ij} e p_{ij} o número de sobreviventes, o número total de pacientes e a probabilidade de sobreviventes para o paciente com condição i e que recebeu o tratamento j , respectivamente. Para os parâmetros do modelo μ, μ_2^A, μ_2 e μ_{22}^{AB} ele propôs uma Distribuição a Priori Normal $N(0, 8)$, com a suposição de que os parâmetros do modelo são independentes.

Outro estudo importante foi Queiroz (2004), que produziu uma análise sobre os fatores de risco para a contaminação com o HIV em doadores de sangue. Para isto foi utilizado um modelo de Densidade Bernoulli para os dados coletados e como análise bayesiana a Priori não-informativa $N(0, 0.001)$ foi estimada.

Souza (2003) em seu texto abordou vários exemplos de aplicação da inferência bayesiana, entre eles o autor estimou a probabilidade de que o sétimo filho de um casal tivesse o sexo masculino dado que este casal concebeu seis filhos, dentre os quais a sexta filha possui sexo feminino e os demais são do sexo masculino, utilizando uma densidade a Priori Beta com 3 valores diferentes para os parâmetros α e β , para, então, mostrar que se a Informação a priori é vaga, então a Informação a posteriori também o será, já que esta é baseada na função de verossimilhança. Souza (2003) concluiu ainda que, a medida em que se aumenta a informação a priori, mais peso esta informação terá na distribuição a posteriori e menor será a dependência do conjunto de dados.

Freire (2001) utilizou o método bayesiano e estimou para seus dados epidemiológicos uma Distribuição Poisson e uma Distribuição a Priori Uniforme para o parâmetro da Poisson. Com isso fez uma revalidação de alguns métodos criados para corrigir o viés

provocado pelo problema da má classificação muito comum em estudos epidemiológicos.

Alexander (2000) propôs Distribuições a Priori Gama, Normal e não-informativa de Jeffreys em seu estudo sobre um parasita causador da doença filariose linfática, utilizando dados provenientes de uma Distribuição Binomial Negativa.

Moller *et al.* (1996) levanta ideia a respeito dos LGCP e mostra que estes processos fornecem uma classe rica de modelos de padrões com pontos agrupados. Moller *et al.* (1998) aborda as propriedades dos processos de Cox, dirigidos por um processo de intensidade Log Gaussian, fazendo uma aplicação no setor florestal com dados referentes a 126 mudas de Pinheiro Silvestre em um lote de $10 \times 10 m^2$, e conclui que este ajuste é o melhor na adequação aos dados com heterogeneidade estrutural. Moller *et al.* (1998) faz ainda uma análise Bayesiana empírica com respeito a heterogeneidade ambiental dos dados utilizados.

Paez (2004) propõe uma análise de modelos para a estimação e previsão de processos espaço-temporais, a partir de observações que apresentam estrutura de correlação espaço-temporal. Para isto, Paez (2004) utilizou alguns modelos hierárquicos que incorporavam esta estrutura, entre eles o modelo LGCP, baseando-se no artigo de Brix e Diggle (2001), com uma abordagem bayesiana pelo método de MCMC (*Markov Chain Monte Carlo*)

Taylor *et al.* (2011) apresenta um pacote do *Software R* para a previsão dos processos espaço-temporais Log Gaussian Cox fazendo uma abordagem ao método MCMC e, com isso, fornecendo um conjunto de funções para a implementação de algoritmos MCMC para os processos Log Gaussian Cox.

Simpson *et al.* (2011) apresenta uma abordagem inédita para a realização de inferências computacionais em processos Log Gaussian Cox, onde propõe estudar diretamente a probabilidade do processo pontual e não apenas uma partição do domínio, permitindo assim o estudo de modelos mais complexos.

Taylor e Diggle (2012) Estudaram a realização da inferência bayesiana nos Processos espaciais Log Gaussianos de Cox, assumindo um campo aleatório contínuo no espaço, de duas formas diferentes: pelo método de *Monte Carlo Markov Chain* (MCMC) e através da aproximação *Integrated Nested Laplace Approximation* (INLA), com uma aplicação em métodos da estatística computacional por meio de simulações. O estudo teve por finalidade verificar e comparar o desempenho e a precisão preditiva dos métodos estudados, verificou-se que, para o estudo proposto, o método INLA mostrou-se mais adequado.

2.1 Teorema de Bayes

Considerando-se o parâmetro θ desconhecido. Ao se observar uma quantidade aleatória X pode-se ter maiores informações a respeito do parâmetro, esta relação pode ser bem definida pela distribuição amostral $p(x|\theta)$. Segundo Ehlers (2007), a idéia de que após observar a quantidade $X = x$ é aumentada a informação sobre o parâmetro θ pode ser claramente vista através do Teorema de Bayes, que é a regra de atualização utilizada para quantificar este aumento da informação:

$$p(\theta|x) = \frac{p(x, \theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(\theta, x)d\theta} \quad (2.1)$$

Observa-se que em (2.1) o termo independente de θ $1/p(x)$ funciona como uma constante normalizadora de $p(\theta|x)$, além disso também pode-se perceber que para cada valor fixo de x a função $l(\theta; x) = p(x|\theta)$ fornecerá a verossimilhança de cada um dos possíveis valores do parâmetro θ . Intuitivamente, combinando-se a verossimilhança e a priori de θ , $p(\theta)$, chega-se à distribuição à posteriori do parâmetro, que é a forma usual do Teorema de Bayes, ou seja,

$$p(\theta|x) \propto l(\theta; x)p(\theta) \quad (2.2)$$

Pode-se observar que a proporcionalidade em (2.2) é uma forma reduzida para a Equação (2.1) quando o termo $p(x)$ é omitido, ou seja, $p(x)$ é uma constante normalizadora e (2.2) é bastante útil em problemas que envolvem estimação de parâmetros. A partir desta proporcionalidade a constante normalizadora pode ser recuperada como

$$p(x) = \int p(x, \theta)d\theta = \int p(x|\theta)p(\theta)d\theta = E_{\theta}[p(X|\theta)],$$

onde esta é a distribuição esperada para a observação de x dado θ , conhecida como distribuição preditiva, logo, antes de observar a variável X pode-se verificar a adequação da distribuição à priori por meio de predições através de $p(x)$.

2.2 Alguns métodos de análise Bayesiana

Em qualquer estudo estatístico a informação que é tida a respeito de um parâmetro θ é fundamental para sua concretização. Na maior parte dos casos θ é desconhecido, logo a informação inicial sobre o parâmetro θ é mínima, daí surge a necessidade de um $\hat{\theta}$ (θ

estimado) preciso e suficientemente representativo para o parâmetro populacional.

Na análise *Bayesiana* essa incerteza sobre θ é representada por meio dos modelos probabilísticos para o parâmetro, assim, de acordo com o grau de incerteza do pesquisador, podem surgir modelos probabilísticos distintos a respeito de θ , por isso quantidades observáveis e os parâmetros de um modelo são ambos considerados quantidades aleatórias, o que faz com que estudos com base nas observações ou no parâmetro sejam adequados a estudos científicos.

Desde a discussão inicial de Bayes, a inferência Bayesiana tem sido motivo de diversos estudos e debates, principalmente após o avanço da computação que, aliada a estudos científicos, impulsionou o interesse, a possibilidade e o avanço de técnicas e aplicações bayesianas.

Pollard (1986) define a aproximação Bayesiana como um método que busca informações a priori tanto quanto informações da amostra e especifica como a informação a priori pode ser corrigida a partir dos dados da amostra. Além disso, ressalta que a ideia de “ignorância” ou “nenhum conhecimento” que é conduzida a uma priori não-informativa deve ser utilizada com cautela pois entra em contradição com sua finalidade, assim a expressão de conhecimento a priori é melhor vista como uma aproximação ou representação de conhecimento vago do que total ignorância a priori.

Segundo Paulino *et al.* (2003) distribuições a priori não-informativas podem desempenhar um papel importante nos estudos bayesianos como gerar informações a posteriori, principalmente quando se parte de um estudo com pouco conhecimento sobre o parâmetro, permitir a comparação de resultados bayesianos com resultados obtidos pela inferência clássica, que baseia-se apenas no estudo da amostra, ou averiguar a influência da distribuição a priori subjetiva, que descreve a informação verdadeira sobre o parâmetro.

De acordo com Saraiva (2007), ao longo dos anos e com a investigação de várias técnicas, pode-se observar que as ideias e métodos fundamentais foram originalizados com Jeffreys, ainda segundo ela vale salientar que o ponto de vista de Jeffreys evoluiu através da escolha de prioris por convenção mais do que pela representação única de ignorância a priori. A seguir serão mostrados alguns dos principais métodos de estudo bayesianos para obtenção de prioris.

2.2.1 Prioris conjugadas

A partir do conhecimento que se tem sobre a quantidade θ pode-se definir uma família paramétrica de distribuições probabilisticamente representativas para θ , com parâmetros especificados de acordo com este conhecimento prévio (chamados de hiperparâmetros para que sejam diferenciados dos parâmetros de interesse θ).

Segundo Ehlers (2007), no caso de prioris conjugadas, a ideia geral é que as distribuições a priori e posteriori pertençam a mesma classe de distribuições, fazendo com que a atualização do conhecimento sobre θ envolva apenas uma mudança nos hiperparâmetros, assim, a aplicação do método Bayesiano pode ser dada definindo-se apenas a regra de atualização dos hiperparâmetros.

Definição 2.1 *Se $F = p(x|\theta), \theta \in \Theta$ é uma classe de distribuições amostrais então uma classe de distribuições P é conjugada a F se*

$$\forall p(x|\theta) \in F \quad e \quad p(\theta) \in P \Rightarrow p(\theta|x) \in P.$$

Se existe uma constante k tal que $k^{-1} = \int l(\theta; x) d\theta < \infty$ e todo $p \in P$ é definido como $p(\theta) = kl(\theta; x)$ então P é a família conjugada natural ao modelo amostral gerador de $l(\theta; x)$.

Alguns termos como “conhecimento vago”, “não informação” ou “ignorância a priori” são bastante comuns na estatística Bayesiana e dizem respeito a especificação de distribuições a priori onde é esperado que a informação obtida com o conjunto de dados seja dominante, no sentido em que a priori é “vaga”. Alguns tipos de prioris não informativas são dados a seguir.

2.2.2 Priori não informativa de Bayes e Laplace

Prioris não informativas são frequentemente utilizadas quando pouco ou nenhuma informação está disponível. Uma ideia bastante intuitiva para prioris não informativas seria supor uma distribuição a priori constante no espaço paramétrico, ou seja, uma Distribuição Uniforme. Esta ideia motivou Bayes (1763) a sugerir $p(\theta) \propto k$ para θ variando em um subconjunto da reta, fazendo com que os possíveis valores de θ fossem equiprováveis. Entretanto esta escolha de priori pode ocasionar algumas dificuldades técnicas, Ehlers (2007) apresenta algumas dessas dificuldades:

(i) Se o intervalo de variação θ for ilimitado, então a distribuição a priori é imprópria

$$\int p(\theta)d\theta = \infty.$$

(ii) Se $\phi = g(\theta)$ é uma reparametrização não linear monótona de θ , então $p(\Phi)$ é não uniforme, já que pelo teorema de transformação de variáveis

$$p(\Phi) = p(\theta(\Phi)) \left| \frac{d\theta}{d\Phi} \right| \propto \left| \frac{d\theta}{d\Phi} \right|.$$

Entretanto, nem sempre uma priori Uniforme é uma “boa referência” como distribuição por ser invariante sobre reparametrização. Uma outra Priori não informativa frequentemente utilizada foi proposta por Jeffreys (1961) e será discutida a seguir.

2.2.3 Priori não informativa de Jeffreys

Para encontrar uma classe de Prioris, Jeffreys utilizou o conceito de medida da Informação de Fischer.

Definição 2.2 Informação de Fischer

Considere uma única observação X com função densidade de probabilidade $p(x|\theta)$. A medida de informação esperada de Fischer de θ através de X é definida como

$$I(\theta) = E \left[-\frac{\partial^2 \log p(x|\theta)}{\partial \theta^2} \right].$$

Se θ for um vetor paramétrico define-se então a matriz de informação esperada de Fischer de θ através de X como

$$\mathbf{I}(\theta) = E \left[-\frac{\partial^2 \log p(x|\theta)}{\partial \theta \partial \theta'} \right].$$

Este conceito de informação pode ser associado a uma espécie de curvatura média da função de verossimilhança, onde, quanto maior é esta curvatura, mais precisa é esta informação contida na verossimilhança, ou ainda, maior o valor da Informação de Fischer. Além disso, a informação $I(\theta)$ pode ser considerada como uma medida de informação global por estar associada a esperança matemática, tomada em relação a distribuição amostral $p(x|\theta)$. Uma medida de Informação local pode ser obtida se não tomarmos o valor esperado na definição acima.

Com base nestes conceitos, Jeffreys propôs a seguinte definição para prioris:

Definição 2.3 *Seja uma observação X com função densidade de probabilidade $p(x|\theta)$. A Priori não informativa de Jeffreys tem função de densidade definida por*

$$p(\theta) \propto [I(\theta)]^{1/2}.$$

Se θ for um vetor paramétrico então

$$p(\theta) \propto |\det \mathbf{I}(\theta)|^{1/2}.$$

De acordo com Ehlers (2007), de modo geral a Priori não informativa de Jeffreys é obtida fazendo-se o parâmetro de escala da distribuição conjugada tender a zero e fixando-se os demais parâmetros convenientemente.

2.2.4 Distribuição a Posteriori

Definição 2.4 *Uma Densidade a Posteriori para θ é dada pela Densidade Condicional de Θ , dados os valores amostrais. Assim,*

$$f_{\theta|x}(\theta|x) = \frac{f_{x,\Theta}(x, \theta)}{f_x(x)}.$$

Em que $f_{x|\Theta}(x, \theta)$ é a função de densidade conjunta para x e Θ , dada pelo produto entre a função de verossimilhança $P(x|\theta)$ e a função de densidade a priori $f_{\theta}(\theta)$, e $f_x(x)$ é a distribuição marginal de x em relação a densidade conjunta.

Larson (1982) observa que enquanto a densidade a priori expressa nossa crença sobre os valores de θ antes da amostra, a densidade a posteriori expressa nosso grau de credibilidade sobre a situação de θ , dados os resultados da amostra.

2.2.5 Estimadores de Bayes

Em uma amostra aleatória X_1, \dots, X_n tomada de uma função densidade de probabilidade $p(x|\theta)$, com θ desconhecido, o valor de θ é estimado a partir dos dados observados na amostra. Se $\theta \in \Theta$, é intuitivo supor que os possíveis valores de um estimador $\delta(X)$ também devam pertencer ao espaço Θ .

Um bom estimador é tal que, com alta probabilidade, o erro $\delta(X) - \theta$ seja bastante próximo de zero. Associando a cada possível valor de θ e cada possível estimativa $a \in \Theta$ uma função $L(a, \theta)$ de perda, de modo que quanto maior a distância entre a e θ , maior o

valor da perda, podendo-se estimar uma perda a posteriori da seguinte forma

$$E [L(a, \theta)|x] = \int L(a, \theta)p(\theta|x)d\theta.$$

A regra de estimadores de Bayes consiste em escolher a estimativa que minimiza esta função de perda esperada. Alguns exemplos de funções de perda podem ser vistos em Bernardo e Smith (1994) e O'Hagan (1994).

2.3 Métodos de Aproximação Bayesiana

Supondo $g(\theta)$ como uma função particular do parâmetro θ , pode-se, convenientemente, fazer uma abordagem para a Distribuição a Posteriori, em termos da esperança de $g(\theta)$:

$$E[g(\theta)|\mathbf{x}] = \int g(\theta)p(\theta|\mathbf{x})d\theta,$$

ou, para o caso em que θ é multidimensional, esta abordagem pode ser feita a partir das distribuições a posteriori marginais de θ , por exemplo, para $\theta = (\theta_1, \theta_2)$, então

$$p(\theta_1|\mathbf{x}) = \int p(\theta|\mathbf{x})d\theta_2.$$

O problema da Inferência Bayesiana, de maneira geral, consiste em calcular tais valores esperados, segundo a distribuição a posteriori de θ . Entretanto, existem funções onde o cálculo da integral é bastante complexo, principalmente com a abordagem multidimensional, assim, uma Inferência exata e precisa será possível apenas se estas integrais puderem ser analiticamente calculadas. Para o caso onde este cálculo de integrais não é possível, deve-se usar, de maneira cautelosa, os métodos de Aproximação.

A seguir, serão apresentados, de maneira resumida, alguns dos métodos computacionais de aproximação. Mais detalhes sobre estes métodos podem ser obtidos em (ROBERT; CASELLA, 1999), (GAMERMAN, 1997), (GAMERMAN; LOPES, 2006).

2.3.1 Método de Monte Carlo

A ideia do método consiste em escrever em função de um valor esperado a integral que se deseja calcular. De modo geral, para o caso em que a integral é a esperança de uma função $g(\theta)$, onde θ tem função de densidade definida por $p(\theta)$,

$$I = \int_a^b g(\theta)p(\theta)d\theta = E[g(\theta)], \quad (2.3)$$

de forma que o estimador de Monte Carlo \hat{I} será definido por

$$\hat{I} = \bar{g} = \frac{1}{n} \sum_{i=1}^n g(\theta_i),$$

onde se tem cada θ_i independente.

Pela Lei Forte dos Grandes Números, para um n consideravelmente grande, \hat{I} converge quase certamente para I

$$\frac{1}{n} \sum_{i=1}^n g(\theta_i) \xrightarrow{q.c.} E[g(\theta)], \quad n \rightarrow \infty.$$

Pode-se, então, concluir que

$$E[g(\theta_i)] = E[g(\theta)] = I \quad \text{e} \quad Var[g(\theta_i)] = \sigma^2 = \frac{1}{n^2} \sum_{i=1}^n (g(\theta_i) - \bar{g})^2.$$

É importante lembrar que o tamanho de n não se refere ao tamanho amostral dos dados, e está dentro do controle do pesquisador. Portanto, a aproximação pode ser tão precisa quanto se deseja, a partir do aumento ou diminuição no tamanho de n .

2.3.2 Monte Carlo via Função de Importância

O procedimento de Monte Carlo via função de importância, ou, procedimento de amostragem por importância, é indicado para os casos em que a simulação de valores da distribuição a posteriori é bastante difícil, ou mesmo impossível.

Inicialmente, é preciso recorrermos a uma função $q(\theta)$, de fácil amostragem, chamada função de importância. Se a função $q(\theta)$ é uma função de densidade definida no mesmo espaço da variação de θ , então pode-se reescrever a integral 2.3 como

$$I = \int \frac{g(\theta)p(\theta)}{q(\theta)} q(\theta) dx = E \left[\frac{g(\theta)p(\theta)}{q(\theta)} \right].$$

Neste caso, a esperança diz respeito a distribuição de q , de maneira que, para uma amostra aleatória $\theta_1, \dots, \theta_n$ tomada da distribuição de q , o estimador de Monte Carlo será definido por

$$\hat{I} = \frac{1}{n} \sum_{i=1}^n \frac{g(\theta_i)p(\theta_i)}{q(\theta_i)},$$

com as mesmas propriedades do estimador de Monte Carlo simples.

Alguns cuidados devem ser tomados ao escolher a densidade q . A escolha ótima, no sentido de minimizar a variância do estimador, é tomar $q(\theta) \propto g(\theta)p(\theta)$.

2.3.3 Método de Reamostragem Ponderada

Os métodos de reamostragem devem ser aplicados nos casos onde é muito difícil, ou impossível, simular valores para as distribuições. Segundo Ehlers (2007), a ideia geral destes métodos é dada em duas etapas, onde, primeiramente, são gerados valores aleatórios de uma distribuição auxiliar conhecida. Esta distribuição auxiliar, na prática, geralmente é a distribuição a priori, conforme proposto por Ehlers (2007) *apud* (SMITH; GELFAND, 1992). Na segunda etapa, a proposta é a utilização de um mecanismo de correção para que os valores obtidos sejam representativos da distribuição a posteriori.

Entre os métodos de reamostragem, estudaremos a Reamostragem Ponderada. Esta técnica utiliza as duas etapas acima comentadas, entretanto, diferentemente de alguns outros métodos de reamostragem, não há necessidade de maximização da verossimilhança.

Suponha-se uma amostra $\theta_1, \dots, \theta_n$ da distribuição auxiliar q . A partir desta amostra, é possível construir os pesos

$$w_i = \frac{p(\theta_i|\mathbf{x})/q(\theta_i)}{\sum_{j=1}^n p(\theta_j|\mathbf{x})/q(\theta_j)}, \quad i = 1, \dots, n.$$

Tomando-se uma reamostra de tamanho m da distribuição discreta $\theta_1, \dots, \theta_n$, com probabilidades w_1, \dots, w_n , e assumindo a priori como densidade auxiliar, os pesos se simplificam para

$$w_i = \frac{p(\mathbf{x}|\theta_i)}{\sum_{j=1}^n p(\mathbf{x}|\theta_j)}, \quad i = 1, \dots, n.$$

Em ambos os métodos de simulação citados até aqui, métodos não iterativos, os valores de estimativa da distribuição a posteriori são gerados de forma independente e com a preocupação única voltada para o tamanho da amostra, facilitando, assim, a simulação. Entretanto, na prática, é difícil, e muitas vezes impossível, conseguir encontrar uma densidade de importância que seja ao mesmo tempo fácil de ser amostrada e uma boa aproximação da posteriori. Uma alternativa aos métodos não iterativos são os métodos de Monte Carlo via Cadeias de Markov, descritos a seguir.

2.3.4 Monte Carlo via Cadeias de Markov

Os métodos de Monte Carlo via Cadeias de Markov (MCMC) também tem como ideia geral a obtenção de uma amostra da distribuição a posteriori e o calculo das estimativas amostrais de características desta distribuição, no entanto, são usadas tecnicas iterativas de simulação com base nas cadeias de Markov, fazendo com que os valores gerados não sejam independentes. Segundo Ehlers (2007), os métodos MCMC simulam um passeio aleatório no espaço de θ , convergindo para uma distribuição estacionária, que é a distribuição de interesse no problema.

2.3.4.1 Cadeias de Markov

A definição de uma Cadeia de Markov está inteiramente ligada aos Processos de Markov, por isso, segue a definição.

Definição 2.5 (Processo Estocástico) *Seja T um conjunto arbitrário. Um processo estocástico é uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória.*

Em resumo, Processos Estocásticos são sequências geradas aleatoriamente de acordo com as leis probabilísticas, em que, se para cada $t \in T$, $Z(t)$ é uma variável aleatória definida sobre o espaço de probabilidades Ω , então pode-se afirmar que $Z(t)$ é uma função de dois argumentos, $Z(t, \omega)$, com $\omega \in \Omega$ e $t \in T$. Esta interpretação pode ser vista na Figura 2.

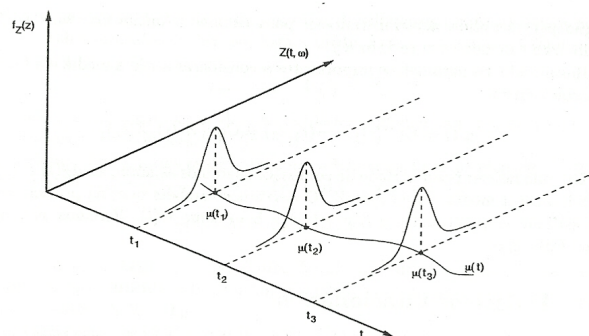


Figura 2: Representação Gráfica de um Processo Estocástico como uma família de Variáveis Aleatórias. Fonte: (MORETTIN; TOLOI, 2004)

Para cada $t \in T$ temos uma variável aleatória (v.a.) com uma distribuição de probabilidade. A função densidade de probabilidade no instante t_1 pode, ou não, ser a mesma

no instante t_i , com $i = 1, \dots, n$. Além disso, para cada $\omega \in \Omega$ fixado será definida uma função t , ou seja, uma trajetória do processo.

Uma Cadeia de Markov é um caso particular dos Processos Estocásticos, onde os estados anteriores do processo não são relevantes para as previsões dos estados seguintes, desde que o estado atual seja conhecido.

Definição 2.6 (Cadeia de Markov) *Seja o Processo Estocástico $\{X_0, X_1, \dots\}$, onde cada $X_i, i = 1, \dots, t$ representa uma variável aleatória, com possíveis valores de probabilidade em um conjunto arbitrário T . X_t é Cadeia de Markov se*

$$P(X_t \in T | X_0, \dots, X_{t-1}) = P(X_t \in T | X_{t-1}).$$

Ou seja, uma Cadeia de Markov é um processo estocástico $\{X_0, X_1, \dots\}$, tal que a distribuição de X_t , dados todos os valores anteriores X_0, \dots, X_{t-1} , depende exclusivamente de X_{t-1} .

O método de simulação MCMC requer ainda que a Cadeia de Markov seja homogênea (onde não há variação entre as probabilidades de transição de um estado para o outro), irreduzível (em que cada estado pode ser atingido, a partir de qualquer outro estado e em um número finito de iterações) e aperiódica (onde não hajam estados absorventes).

Supondo-se uma distribuição $\pi(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d$, com pelo menos uma constante multiplicativa conhecida, complexa o suficiente para que não seja possível a obtenção direta de uma amostra. Dadas as realizações $\{X^{(t)}, t = 0, 1, \dots\}$ de uma cadeia de Markov onde π é a distribuição de equilíbrio, então

$$X^{(t)} \xrightarrow{t \rightarrow \infty} \pi(x) \quad \text{e} \quad \frac{1}{n} \sum_{t=1}^n g(X_i^{(t)}) \xrightarrow{n \rightarrow \infty} E_\pi(g(X_i))$$

Assim, a média dos valores da cadeia é um estimador consistente para a média teórica. Conforme o número de iterações aumenta, a cadeia “esquece” os valores iniciais e, eventualmente, converge para uma distribuição de equilíbrio, fazendo-se com que, na prática, algumas iterações iniciais sejam descartadas.

2.3.4.2 Algoritmo de Metropolis-Hastings

A ideia principal dos Algoritmos de Metropolis-Hastings consiste em gerar valores de uma distribuição auxiliar, aceitando-os dadas suas probabilidades, bem como nos métodos

de rejeição. Este processo de reamostragem garante a convergência da cadeia para a distribuição de equilíbrio, que, neste caso, é a própria distribuição a posteriori.

Supondo que a cadeia esteja no estado θ , os valores θ' gerados a partir de uma distribuição proposta $q(\cdot|\theta)$ podem compor a amostra oriunda da distribuição a posteriori, dependendo da aceitação dada sua probabilidade, onde θ pode ser um parâmetro de interesse ou um vetor de parâmetros.

A partir de uma cadeia de Markov, é gerado um θ' para o próximo estado, dada a distribuição $q(\cdot|\theta)$ proposta. O valor θ' gerado depende dos estados anteriores da cadeia de modo que $\theta_{t-1} = \theta'_t$. Este valor será adicionado a cadeia e aceito com probabilidade de aceitação definida por

$$\alpha(\theta, \theta') = \min \left(1, \frac{\pi(\theta')q(\theta|\theta')}{\pi(\theta)q(\theta'|\theta)} \right), \quad (2.4)$$

em que π é a distribuição de interesse. A probabilidade 2.4 não se altera, por isso, precisamos conhecer π apenas parcialmente, logo, este processo nos dá a possibilidade de gerar uma amostra a partir de uma distribuição não conhecida.

2.3.4.3 Amostrador de Gibbs

Nos amostradores de Gibbs, ao contrário dos Algoritmos de Metropolis-Hastings, não existe um mecanismo de aceitação pois a cadeia sempre irá se mover para um novo valor. As transições entre os estados da cadeia serão feitas a partir de distribuições condicionais completas $\pi(\theta_i|\theta_{-i})$, onde $\theta_{-i} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_d)$, onde cada θ_i pode ser uni ou multidimensional.

A distribuição condicional completa é a distribuição da i -ésima componente de θ , condicionada em todas as outras componentes. Esta distribuição é obtida através da distribuição conjunta

$$\pi(\theta_i|\theta_{-i}) = \frac{\pi(\theta)}{\int \pi(\theta)d\theta_i}$$

Se as distribuições condicionais forem completamente conhecidas, então o algoritmo de Gibbs pode ser dado a partir da obtenção de $\theta^{(t)}$, por meio da especificação de θ^{t-1} , através da geração sucessiva de valores

$$\begin{aligned}
\theta_1^{(t)} &\approx \pi(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}) \\
\theta_2^{(t)} &\approx \pi(\theta_2 | \theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}) \\
&\vdots \\
\theta_d^{(t)} &\approx \pi(\theta_d | \theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{d-1}^{(t)})
\end{aligned}$$

Cada iteração se completa após d movimentos ao longo das componentes de θ . Após o final das iterações, onde ocorre a convergência, os valores resultantes irão compor a amostra de $\pi(\theta)$. O amostrador de Gibbs é, portanto, um caso especial do algoritmo de Metropolis-Hastings, onde os elementos de θ são atualizados com base na distribuição condicional completa, onde a probabilidade de aceitação é igual a 1.

2.3.4.4 MCMC com Saltos Reversíveis (RJMCMC)

O método MCMC com Saltos Reversíveis pode ser visto como uma extensão dos algoritmos de Metropolis-Hastings, onde propõe-se novos valores para a cadeia de Markov, de acordo com a probabilidade de aceitação estimada. Entretanto, no caso do método RJMCMC, esta proposição de novos valores poderá ocorrer entre espaços com diferentes dimensões. O processo é executado a partir dos métodos usuais MCMC, onde, em cada iteração realizada, ocorre a atualização dos parâmetros, de acordo com o modelo em estudo.

Inicialmente, supondo-se que o estado atual da cadeia é (k, θ) , ou seja, a cadeia está no modelo k com parâmetros θ . Nestas condições, é proposto, com probabilidade $r_{k,k'}$, um novo modelo k' com parâmetros θ' . Assumindo-se que o modelo proposto possui dimensão maior que o modelo atual, ou seja, $n_{k'} > n_k$, e que o novo vetor paramétrico $\theta' = g(\theta, \mathbf{u})$, em que g é uma função determinística e \mathbf{u} é um vetor aleatório com dimensão $n_{k'} - n_k$ e distribuição $\mathbf{u} \sim q(\mathbf{u})$.

A probabilidade de aceitação de (k', θ') será definida por

$$\min \left(1, \frac{\pi(k', \theta')}{\pi(k, \theta)} \times \frac{r_{k',k}}{r_{k,k'} q(\mathbf{u})} \left| \frac{\partial g(\theta, \mathbf{u})}{\partial(\theta, \mathbf{u})} \right| \right)$$

3 Metodologia

Com a classe dos Modelos Log Gaussianos no Processo de Cox, é possível realizar um estudo amplo na análise dos processos temporais de padrões pontuais de Cox. Por exemplo, estudos relacionados a incidência de doenças, ao número de imigrantes de uma localidade, ou ainda às sementes de uma planta em uma determinada região, enfim, qualquer evento relacionado a um processo de contagem.

Definição 3.1 *Processo de Contagem:*

Um Processo $\{N(t), t \geq 0\}$ é dito ser um Processo de Contagem se $N(t)$ representa o número de eventos que ocorrem em um intervalo $(0, t)$ de tempo.

Estudos dentro deste contexto podem ser analisados, segundo Paez (2004), por meio das próprias observações dos eventos e suas contagens em intervalos de tempo ou espaço, entretanto, os modelos LGCP fornecem uma rica modelagem, o que torna o estudo mais abrangente e mais cheio de informações. Através dos LGCP é possível explorar diferentes níveis de agregação da variável de interesse principal e, conseqüentemente, diferentes quantidades de informação fornecidas pelo conjunto de dados, tornando possíveis considerações a respeito do ajuste do modelo proposto.

Ao se pensar em processos de contagem, pode-se fazer uma breve associação entre estes processos e a distribuição de Poisson, bem como aos Processos de Poisson, intuitivamente. Logo, segue a definição de um Processo de Poisson.

Definição 3.2 *Processo de Poisson:*

Um processo de contagem $\{N(t), t \geq 0\}$ é dito ser um Processo de Poisson com uma taxa $\lambda, \lambda > 0$, se:

i) $N(0) = 0$.

ii) O Processo tem incrementos independentes, ou seja, o número de eventos que ocorrem em intervalos de tempo disjuntos são independentes.

iii) O número de eventos em qualquer intervalo de tempo com comprimento t segue uma distribuição de Poisson com parâmetro λt , isto é, para $s, t \geq 0$,

$$P[N(t+s) - N(s) = n] = \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots$$

Em estudos relacionados a processos pontuais, os processos de Poisson desempenham um papel importante, entretanto, não muito representativo de processos reais em virtude da homogeneidade nos processos. Uma generalização dos processos de Poisson que considere a função de intensidade do processo como uma realização de um campo aleatório pode representar melhor a heterogeneidade geralmente encontrada nos processos reais. A esta generalização dar-se o nome de Processo de Cox.

Os processos de Cox são dados em função de um processo não homogêneo de Poisson, juntamente com uma medida de intensidade aleatória, chamada por alguns autores de processo de intensidade, ou de superfície. Moller *et al.* (1998) define um processo de Cox com uma função de intensidade aleatória como

Definição 3.3 *Seja um subconjunto X , definido no plano \mathbb{R}^d . X será um processo de Cox com processo aleatório de intensidade $\Lambda = \{\Lambda(s) : s \in \mathbb{R}^d\}$, se a distribuição condicional de X dado Λ é um processo de Poisson com função de intensidade $\Lambda(\cdot)$.*

Assim, pode-se assumir que X será um processo de Cox com função de intensidade Λ quando, em conjuntos limitados de Borel $\mathcal{B} \subset \mathbb{R}^d$, existir a distribuição condicional de $X|\Lambda$, onde $(X \cap \mathcal{B})$ possui distribuição de Poisson com parâmetro $\int_{\mathcal{B}} \Lambda(s) ds$ finito e não negativo.

Abordaremos, em especial, o caso em que Λ , e portanto X , é fixo e onde a distribuição de Λ é invariante sob transformações e, possivelmente, rotações em \mathbb{R}^d .

3.1 Processos Log Gaussianos de Cox

Em um processo Log Gaussiano de Cox a heterogeneidade dos dados é modelada por

$$\Lambda(s) = \exp\{Y(s)\},$$

onde Y representa um processo gaussiano, em que a distribuição conjunta de qualquer vetor finito $(Y_{(s_1)}, \dots, Y_{(s_n)})$ é Gaussiana (MOLLER *et al.*, 1997). Além disso, a distribuição de $X|\Lambda$, é um processo de Poisson, em que Λ é a função de intensidade do processo.

A medida aleatória $v(\mathcal{B}) = \int_{\mathcal{B}} \Lambda(s) ds$, para os conjuntos limitados de Borel $\mathcal{B} \subset \mathbb{R}^2$, precisa ser unicamente determinada pela distribuição de Y e estar bem definida. Para isto, será necessário impor a Y algumas condições.

Primeiramente, as realizações de Λ precisam ser integráveis. Além disso, impondo a condição de que v é dada em função de uma alteração contínua de Y , v será unicamente determinada por Y pois todas as alterações contínuas são indistinguíveis, ou seja, suas realizações são idênticas e com probabilidade um. Segue também que $v(\mathcal{B}) < \infty$.

Por hipótese de estacionariedade e isotropia, a distribuição de Y (e, portanto, a de X) é especificada pela média $\mu = E[Y(s)]$, a variância $\sigma^2 = var[Y(s)]$ e função de covariância $c(r) = Cov(Y_{(s_1)}, Y_{(s_2)})$, em que $r = ||s_1 - s_2||$ é a distância euclidiana entre os pontos s_1 e s_2 .

Discussões mais detalhadas a respeito dos Processos Log Gaussianos de Cox são apresentadas em (MOLLER *et al.*, 1996), (MOLLER *et al.*, 1997) e (MOLLER *et al.*, 1998).

Neste trabalho segue-se o procedimento abordado por Paez (2004) para estimar o modelo matemático LGCP. Assumindo-se um processo definido no tempo, com $s \in \mathbb{R}^d$ e definindo-se a função de intensidade Λ como o produto entre a intensidade populacional ρ e uma função de risco Π .

Para a abordagem do LGCP, propõe-se para Π um modelo log linear bem como a implementação de covariáveis que tratarão da variação temporal não explicada por ρ . Dessa forma, tem-se a função de intensidade do processo definida por

$$\Lambda(t) = \rho(t)\Pi(t), \quad (3.1)$$

em que $\rho(t)$ é conhecido, pois se trata da intensidade populacional, $\Pi(t) = \exp[\gamma(t) + \beta\mathbf{X}(t)]$, γ é um processo temporal, $\mathbf{X}(t)$ é um vetor de variáveis explicativas no tempo t e β é um vetor de coeficientes associados a \mathbf{X} .

De acordo com Paez (2004), o processo temporal γ é especificado por uma estrutura auto-regressiva. Fixando-se $E[\exp(\gamma(t) + \beta\mathbf{X}(t))] = 1$, faz-se com que $\exp(\gamma(t))$ e $\beta\mathbf{X}(t)$ assumam o papel de modelar desvios da média do processo de $\rho(t)$. Deste modo, pode-se impor a restrição $E[\gamma(t)] = 0, 5\sigma^2$ que, ainda segundo Paez (2004), assegura o valor 1 para a esperança fixada.

Com isto, a estrutura auto-regressiva de γ será definida pela Equação 3.2

$$[\gamma(t_j) + 0, 5\sigma^2] = \varphi_{t_i, t_j} [\gamma(t_j) + 0, 5\sigma^2] + \epsilon(t_i, t_j), \quad (3.2)$$

com $\epsilon(t_i, t_j) \sim N(0, \tau_{t_i, t_j}^2)$, em que $\tau_{t_i, t_j}^2 = \sigma^2(1 - \varphi_{t_i, t_j}^2)$, para $t_i < t_j \in \mathbb{R}^d, d = 1, 2, \dots$, e φ_{t_i, t_j} é a correlação entre $\gamma(t_i)$ e $\gamma(t_j)$.

Paez (2004) afirma ainda que a correlação φ é definida por uma função exponencial com parâmetro de decaimento ϕ , proporcional à distância temporal entre dois eventos, assim

$$\varphi_{t_i, t_j} = \exp\left[-\frac{t_i - t_j}{\phi}\right], \quad \phi > 0 \quad \text{e} \quad t_i < t_j \in \mathbb{R}^d, d = 1, 2, \dots \quad (3.3)$$

Os LGCP tratam-se de processos espaço-temporais, em virtude do espaço amostral em que é possível realizar aplicações deste processo. Aqui, concentra-se à realização e ao estudo deste processo em decorrência apenas do tempo, que poderá ser feita em duas abordagens diferentes: assumindo-se o tempo contínuo ou discreto. Estas duas abordagens serão discutidas a seguir.

3.1.1 Abordagem LGCP assumindo o tempo contínuo

Para a abordagem do LGCP considerando-se o tempo como uma variável aleatória contínua, suponha-se o estudo de um experimento aleatório onde z representa o conjunto de tempos $z = (z(1), \dots, z(T))$ de ocorrência de um evento de interesse.

Assumindo-se z como uma realização do processo Log Gaussiano de Cox Z , com função de intensidade definida por 3.1, logo

$$\Lambda(z(t)) = \rho(z(t))\Pi(z(t)), \quad \text{com} \quad t = 1, \dots, T \quad \text{e} \quad z(t) \in \mathbb{R}^d,$$

em que $\rho(z(t))$ é a intensidade populacional e $\Pi(z(t)) = \exp[\gamma(z(t)) + \beta\mathbf{X}(z(t))]$.

A função de distribuição acumulada de ocorrência do evento $z(t)$, dado Λ e o evento anterior a $z(t)$, ou seja, $z(t-1)$, será definida por (3.4).

$$F(z(t)|z(t-1), \Lambda) = \exp\left(-\int_{z(t-1)}^{z(t)} \Lambda(s)ds\right). \quad (3.4)$$

E a dependência de γ no tempo será definida assim como em (3.2) por:

$$[\gamma(z(t)) + 0, 5\sigma^2] = \varphi_{z(t-1), z(t)}[\gamma(z(t-1)) + 0, 5\sigma^2] + \epsilon(z(t-1), z(t)),$$

em que $z(t-1), z(t)$ tem distribuição $\text{Normal}(0, \tau_{z(t-1), z(t)}^2), t = 2, \dots, T$, $\varphi_{z(t-1), z(t)}^2$ é definido por $\exp[-\frac{z(t)-z(t-1)}{\phi}]$, $\tau_{z(t-1), z(t)}^2$ é definido por $\sigma^2(1 - \varphi_{z(t-1), z(t)}^2)$ e $\gamma(z(1)) \sim N(-0, 5\sigma^2; \sigma^2)$. Assim, dados ϕ e σ^2 , $z(t)$ somente dependerá das observações passadas por meio de $z(t-1)$.

3.1.2 Abordagem LGCP assumindo o tempo discreto

Considerando-se ainda o mesmo processo de Cox Z definido na seção anterior, com o estudo de um experimento aleatório em que z também representa o conjunto de tempos $z = (z(1), \dots, z(T))$ de ocorrência de um evento de interesse. Suponha-se que não há acesso à informação sobre o momento exato do tempo de ocorrência de cada dado coletado na amostra.

Sendo assim, é mais interessante que os dados sejam observados por meio de agregações em intervalos de tempo, de forma que o processo resposta será dado em função do número de casos reportados durante cada um destes intervalos. Logo, seja $y(i)$ o número de casos observados do evento de interesse no intervalo $i = 1, 2, \dots, T$. A série temporal $y = y(1), \dots, y(T)$ é uma realização do processo de Poisson Z onde

$$y(t) \sim \text{Poisson} \left(\int_{t-1}^t \Lambda(s) ds \right), \quad t = 1, \dots, T, \quad (3.5)$$

com função de intensidade

$$\int_{t-1}^t \Lambda(s) ds \quad \text{ou} \quad \int_{t-1}^t \rho(s) \Pi(s) ds, \quad t = 1, \dots, T.$$

Considerando-se o tempo de maneira discreta, pode-se assumir, por consequência, que o modelo matemático não fará distinção entre $\Lambda(t_i)$ e $\Lambda(t_j)$ para cada t_i e t_j pertencentes ao mesmo intervalo de tempo. Deste modo, pode-se fazer a aproximação

$$\int_{t-1}^t \Lambda(s) ds \approx \Lambda_{apr}(t-1).$$

Logo, pode-se escrever a Equação (3.5) como

$$y(t) \sim \text{Poisson}(\Lambda_{apr}(t)), \quad t = 1, 2, \dots, T \quad (3.6)$$

em que a intensidade Λ medida em $\Lambda_{apr}(t)$, definida por 3.6, não se refere à intensidade no início do intervalo de tempo t , e sim a intensidade média do intervalo de tempo.

O processo temporal γ também será aproximado por intervalo, em que

$$\int_{t-1}^t \gamma(s) ds \approx \gamma(t-1).$$

A dependência de γ no tempo é explicada pela mesma estrutura temporal auto-regressiva de (3.2), com adaptação ao tempo discreto

$$[\gamma_{apr}(t) + 0.5\sigma^2] = \varphi [\gamma_{apr}(t-1) + 0.5\sigma^2] + \epsilon(t), \quad t = 2, \dots, T, \quad (3.7)$$

em que $\epsilon(t)$ tem distribuição Normal com média 0 e variância τ^2 e $\varphi = \exp\left(-\frac{1}{\phi}\right)$ é a correlação entre $\gamma_{apr}(t)$ e $\gamma_{apr}(t-1)$, constantes para todo $t = 2, \dots, T$ e $\tau^2 = \sigma^2(1 - \varphi^2)$.

3.2 Análise Bayesiana

O ponto inicial de toda análise Bayesiana se dá por meio da estimativa de distribuições a priori para os parâmetros de interesse do modelo. Em geral, quando não há nenhuma informação prévia a respeito destas distribuições a priori, esta estimativa pode ser feita de diversas formas, como foi discutido no Capítulo 2.

Particularmente neste trabalho, existe a informação prévia abordada na tese de doutorado de Paez (2004) a respeito de distribuições a priori para o parâmetro β , o processo γ , e os hiperparâmetros ϕ e σ^2 .

A seguir serão apresentadas as prioris obtidas por Paez (2004) para a abordagem considerando-se o tempo contínuo e discreto, bem como serão tecidas as análises Bayesianas adequadas ao modelo em estudo.

3.2.1 Priori e posteriori para a abordagem no tempo contínuo

Partindo-se da informação inicial de que as distribuições a priori para β e ϕ são, respectivamente

$$\begin{aligned}\beta_i &\sim N(\mu_b, \sigma_b^2) \\ \phi &\sim \text{Uniforme}(a_\phi, b_\phi).\end{aligned}$$

Ao invés de especificar uma priori para σ^2 , Paez (2004) utiliza a prática comum de especificar uma priori para a precisão σ^{-2} . Dessa forma, adota-se a distribuição a priori

$$\sigma^{-2} \sim \text{Gama}(a_\sigma, b_\sigma).$$

A função de distribuição acumulada de z , descrita em (3.4), pode ser descrita na forma da função de distribuição

$$f(z|\beta, \phi, \sigma^2, \gamma) = \prod_{t=2}^T f(z(t)|z(t-1), \beta, \phi, \sigma^2, \gamma) f(z(1)|\beta, \phi, \sigma^2, \gamma) \quad (3.8)$$

Com relação ao processo γ , vamos agora supor que $\gamma^* = [\gamma(z(1)), \gamma(z(2)), \dots, \gamma(z(T))]$ seja o conjunto de valores de γ associados aos tempos de observação dos eventos. Dessa forma, pode-se considerar $\gamma^*(t) = \gamma(z(t))$ e γ^{-*} como todos os valores de γ não incluídos em γ^* , assim, $\gamma^{-*} = \{\gamma - \gamma^*\}$, tal que $\gamma = (\gamma^*, \gamma^{-*})$.

A função de verossimilhança descrita em (3.8) irá depender de γ por meio de γ^* . De modo geral, pode-se escrever a distribuição a posteriori também em função de γ^* , isto é,

$$f(\beta, \phi, \sigma^2, \gamma|z) \propto f(\beta, \phi, \sigma^2, \gamma^*|z) f(\gamma^{-*}|\phi, \sigma^2, \gamma^*). \quad (3.9)$$

Combinando-se as prioris descritas anteriormente com a verossimilhança descrita em 3.8, pode-se definir a função $f(\beta, \phi, \sigma^2, \gamma^*|z)$ como

$$f(\beta, \phi, \sigma^2, \gamma^*|z) \propto f(z|\beta, \phi, \sigma^2, \gamma^*) f(\gamma^*|\phi, \sigma^2) f(\phi) f(\sigma^2) f(\beta), \quad (3.10)$$

além disso, dados ϕ e σ^2 , escreve-se $f(\gamma^*|\phi, \sigma^2)$ como

$$f(\gamma^*|\phi, \sigma^2) \propto f(\gamma^*(T)|\gamma^*(T-1), \phi, \sigma^2) f(\gamma^*(T-1)|\gamma^*(T-2), \phi, \sigma^2) \dots f(\gamma^*(1)|\sigma^2) \quad (3.11)$$

Não existe uma forma fechada para a posteriori definida por (3.9), logo, para realizar-se inferências a respeito dos parâmetros, precisaremos utilizar técnicas computacionais de

reamostragem.

3.2.2 Priori e posteriori para a abordagem no tempo discreto

Os parâmetros de interesse descritos na Seção 3.1.2 são $\beta, \gamma_{apr}, \phi$ e σ^2 , com distribuições a priori novamente estimadas por

$$\begin{aligned}\beta_i &\sim N(\mu_b, \sigma_b^2), \\ \phi &\sim Uniforme(a_\phi, b_\phi) \quad \text{e} \\ \sigma^{-2} &\sim Gama(a_\sigma, b_\sigma)\end{aligned}$$

Ainda na Seção 3.1.2, vimos que y , dados $\beta, \gamma_{apr}, \phi$ e σ^2 , tem distribuição poisson. A função de verossimilhança de y pode ser definida por

$$f(y|\beta, \gamma_{apr}, \phi, \sigma^2) \propto \prod_{i=1}^T (\rho(i)e^{[\gamma_{apr}(i)+\beta X]})^{y_i} e^{-[\rho(i)e^{[\gamma_{apr}(i)+\beta X]}]} \quad (3.12)$$

em que γ_{apr} é novamente o conjunto de valores associados a classes de tempo $\gamma_{apr} = (\gamma_{apr}(1), \gamma_{apr}(2), \dots, \gamma_{apr}(T))$, com uma distribuição semelhante a 3.10 em uma versão regularmente espaçada em virtude das classes de tempo, definida pela Equação (3.13)

$$f(\gamma_{apr}|\phi, \sigma^2) \propto f(\gamma_{apr}(T)|\gamma_{apr}(T-1), \phi, \sigma^2) f(\gamma_{apr}(T-1)|\gamma_{apr}(T-2), \phi, \sigma^2) \dots f(\gamma_{apr}(1)|\sigma^2), \quad (3.13)$$

em que $f(\gamma_{apr}(T)|\gamma_{apr}(T-1), \phi, \sigma^2)$ é definido pela Equação (3.7)

Uma estimativa para a distribuição a posteriori pode ser dada por meio do produto entre a função de verossimilhança (3.12) e as prioris definidas anteriormente. A distribuição a posteriori para $\beta, \gamma_{apr}, \phi$ e σ^2 é definida por

$$f(\beta, \gamma_{apr}, \phi, \sigma^2) \propto f(y|\beta, \gamma_{apr}, \phi, \sigma^2) f(\gamma_{apr}|\phi, \sigma^2) f(\sigma^2) f(\phi) f(\beta) \quad (3.14)$$

Assim como na Seção 3.2.1, em que a distribuição a posteriori não possui forma fechada, no caso em que o tempo é considerado discreto a distribuição a posteriori também não possui. Deste modo, as técnicas de reamostragem computacional também serão aplicadas aqui com o objetivo da realização de inferências a respeito dos parâmetros.

4 Resultados e discussão

Os incêndios e as queimadas florestais são usados muitas vezes como uma ferramenta na agricultura, entretanto, são também responsáveis por alguns prejuízos como, por exemplo, o empobrecimento do solo, a destruição da vegetação, além de interferirem diretamente em alterações na composição química da atmosfera (CRUTZEN; ANDREAE, 1990) e causarem efeito deletério sobre a saúde humana (RIBEIRO; ASSUNÇÃO, 2002). Neste capítulo serão apresentadas duas aplicações para o LGCP, ambas com dados referentes a queimadas e pontos de incêndio.

Na seção 4.1 discute-se a respeito de um banco de dados simulados de pontos de incêndios florestais em Castilla-La Mancha, Reino da Espanha. Nesta aplicação aborda-se o tempo de maneira contínua, considerando-se as observações diárias dentro do intervalo contínuo de horas, minutos... Além disso, será possível visualizar a heterogeneidade dos dados através da proveniência de cada observação (incêndios florestais por meio de relâmpagos, ocasionados de forma acidental ou intencional e gerados por meio de causas desconhecidas).

Apresenta-se na seção 4.2 uma aplicação dos LGCP em dados reais referentes a quantidade diária de focos de calor detectados no Bioma Amazônia através do satélite NOAA-15, considerando-se um campo aleatório gaussiano formado por dados de precipitação média diária de chuva. Nesta aplicação o tempo será considerado de forma discreta, contado em dias desde o dia 1 (referente ao primeiro dia de observação) até o dia 1048 (referente ao último dia de observação no intervalo de tempo amostrado). Todas as causas dos pontos de queimadas são desconhecidas, portanto, nesta aplicação, a heterogeneidade dos dados será desconsiderada.

As análises apresentadas a seguir serão feitas a partir do pacote *lgcp* (TAYLOR *et al.*, 2012), criado para a análise dos LGCP por meio do Software R (TEAM, 2011). Mais detalhes sobre este pacote podem ser vistos em Taylor *et al.* (2011) e Taylor e Diggle (2012).

4.1 Aplicação em dados simulados

Inicialmente realiza-se a aplicação do LGCP em dados simulados referentes a incêndios florestais em Castilla-La Mancha, Reino da Espanha, gerados a partir do Software R (TEAM, 2011).

Pode-se observar por meio da Figura 3 o comportamento dos dados a partir de padrões pontuais (*planar point pattern*). Na Figura 3(A) tem-se uma ideia da localização de cada foco de incêndio, de acordo com a latitude e longitude das observações. Com a Figura 3(B) é possível ver um mapa de suavização do processo. Pode-se notar que em 3(B) há uma forte tendência de cor exatamente onde 3(A) nos mostra a maior quantidade de pontos agregados.

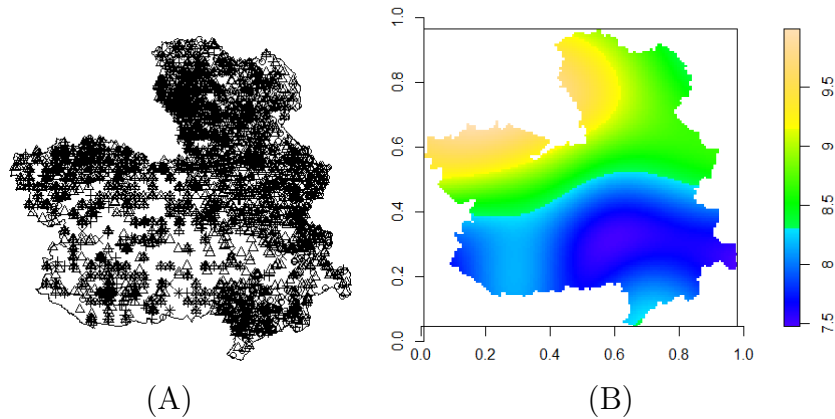


Figura 3: Dados simulados referentes a Incêndios Florestais em Castilla-La Mancha, Reino da Espanha. (A) Padrões Pontuais, (B) Densidade estimada.

Pode-se observar por meio da Figura 4 a suavização do processo em conjunto com os pontos agrupados. Esta suavização, segundo Taylor *et al.* (2011), pode ser interpretada como uma estimativa para a densidade $\Lambda(s)$ do processo.

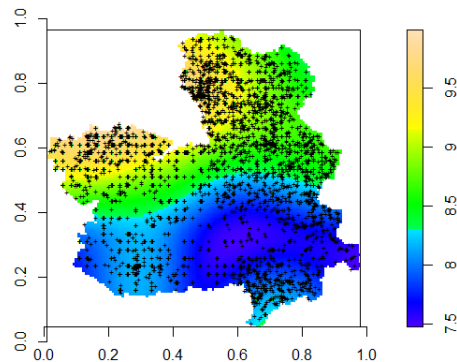


Figura 4: Estimativa para a Densidade de $\Lambda(s)$ em conjunto com os pontos agrupados.

As Figuras 3 e 4 apresentaram a densidade e os pontos agregados de incêndio sem que haja distinção entre as causas de cada um dos focos de incêndio. Entretanto, como visto em seções anteriores, a classe dos LGCP nos permite a análise de conjuntos de dados com padrões de pontos agregados provenientes de uma heterogeneidade natural. A Figura 5 nos mostra em (A) os pontos de focos de incêndio e em (B) a densidade estimada de acordo com cada uma das causas observadas para a ocorrência dos incêndios (incêndios florestais causados por meio de relâmpagos, de forma acidental ou intencional e com causas desconhecidas).

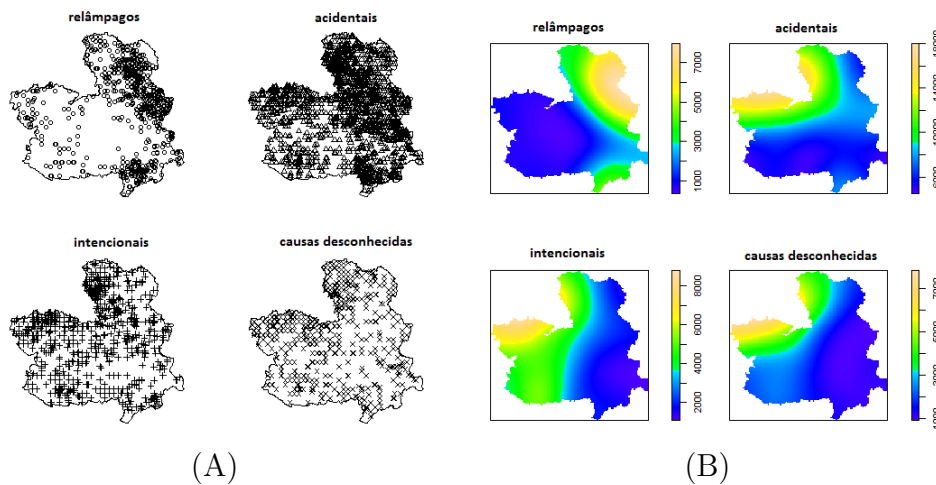


Figura 5: Padrões pontuais de acordo com as causas de cada incêndio em pontos agregados (A) e suavização para o processo (B).

Nesta aplicação, tem-se que y representa o índice de quantidade de focos de incêndio, portanto $\mu = E[y(t)]$ é a média temporal de focos de incêndio, bem como $\sigma^2 = var[y(t)]$ representa a variância temporal de focos de incêndio. A especificação do processo será feita, portanto, através da especificação de estimativas para a média $\mu = E[y(t)]$, a variância $\sigma^2 = var[y(t)]$ e função de covariância $c(r) = Cov(Y_{(t_1)}, Y_{(t_2)})$, que será indicada pelo parâmetro θ .

Inicialmente, gera-se uma estimativa para $\mu = E[y(t)]$ a partir de métodos computacionais, obtendo-se o valor de 0,845 para o número médio de focos de Incêndio em Castilla-La Mancha. O gráfico desta estimativa pode ser visto na Figura 6.

Para a obtenção de estimativas a respeito de σ^2 e ϕ analisa-se os gráficos da função de covariância exponencial. Estes gráficos geram possíveis valores para as estimativas a medida em que, com o uso do Software R, “move-se” as curvas intituladas “empírica” e “teórica”. A escolha de uma boa estimativa é, segundo Taylor e Diggle (2012), aquela que permita a melhor aproximação entre as curvas “empírica” e “teórica” nos gráficos da

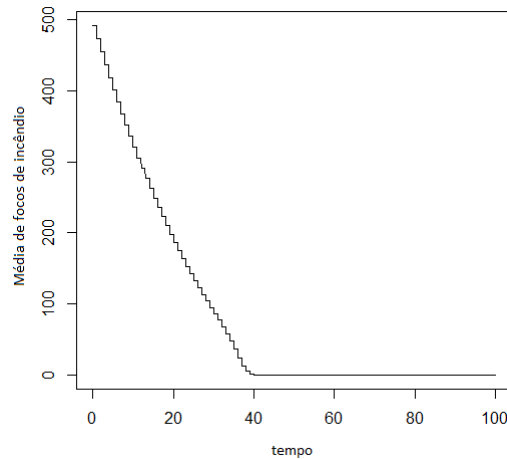


Figura 6: Estimativa para a média do processo $\mu = E[y(t)]$ (Número médio de focos de incêndio em relação ao tempo).

função de covariância exponencial.

Após a análise de diversas possibilidades para boas estimativas, decidiu-se optar por $\sigma^2 = 1.3$ e $\phi = 2.3$, seguindo a proposta de Taylor e Diggle (2012). O gráfico da função de covariância exponencial para estas estimativas pode ser visto na Figura 7, em que considerou-se o comportamento das curvas “empírica” e “teórica” o mais aproximado possível.

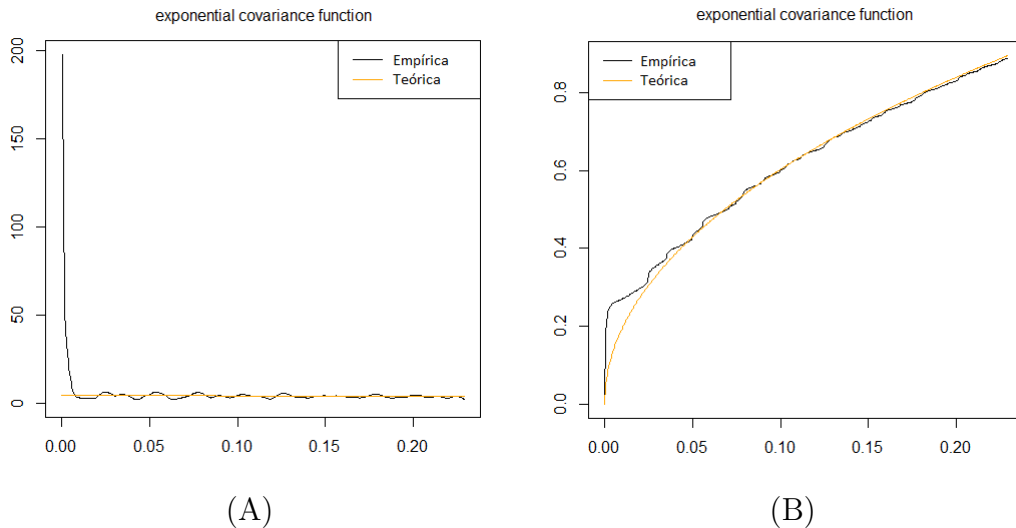


Figura 7: Gráficos da Função de Covariância Exponencial para as estimativas $\sigma^2 = 1.3$ (em (A)) e $\phi = 2.3$ (em (B)).

Especifica-se ainda uma estimativa para o parâmetro de correlação temporal θ através de forma semelhante à análise utilizada para estimar os valores de σ^2 e ϕ . Foram analisados gráficos de Autocovariância do processo na busca de uma melhor aproximação entre as

curvas “empírica” e “teórica”, levando-nos a crêr que o melhor valor estimado para θ é 3,4. Com a Figura 8 pode-se ver, através do gráfico de Autocovariância do processo, uma estimativa possível para θ , dados os valores estimados de σ^2 e ϕ .

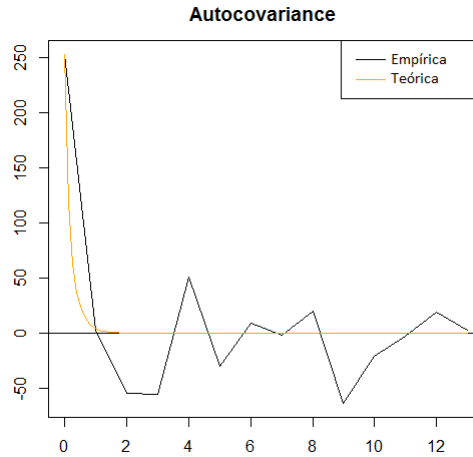


Figura 8: Estimativa para o parâmetro de correlação temporal $\theta = 3.4$.

Os valores estimados por meio da aplicação dos LGCP em dados simulados para pontos de incêndio em Castilla-La Mancha, Reino da Espanha, podem ser vistos na Tabela 1.

Parâmetro	Estimativa
σ^2	1,300
ϕ	2,300
μ	0,845
θ	3,400

Tabela 1: Parâmetros estimatimados a partir do LGCP aplicado aos a pontos de incêndios em Castilla-La Mancha.

Segundo Paez (2004), não existe ainda uma forma matemática fechada para os LGCP, assim, para a realização de previsões a respeito dos dados é preciso utilizar métodos computacionais. Para a realização destas previsões, deve-se utilizar as estimativas dadas na Tabela 1, que definem o processo. Entretanto, neste trabalho aborda-se o LGCP em dados simulados, como visto, e em dados reais. A Seção 4.2 irá realizar a segunda abordagem deste trabalho, de forma que, realizar previsões a respeito de dados reais é mais interessante do que previsões sobre dados simulados, uma vez que dados simulados apresentam um comportamento diferente do real. Por isso, possíveis previsões utilizando as estimativas dos parâmetros serão apresentadas apenas na Seção a seguir.

4.2 Aplicação em dados reais

Existem no Brasil seis biomas diferentes: a Amazônia, o Cerrado, a Mata Atlântica, a Caatinga, o Pampa e o Pantanal. Pode-se observar por meio da Figura 9 o mapa brasileiro com as divisões aproximadas de cada um destes biomas (MMA, 2013b).

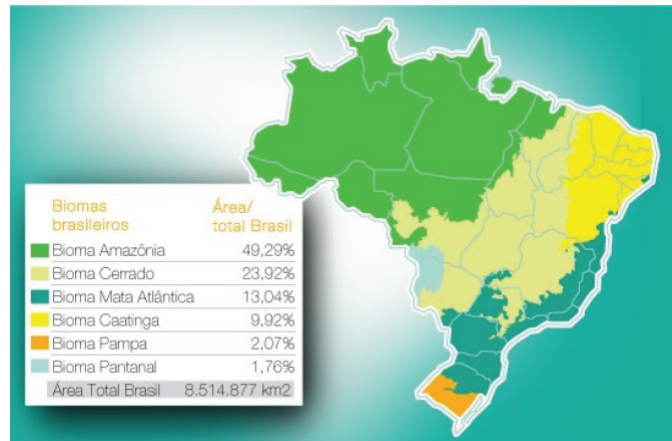


Figura 9: Mapa brasileiro com as divisões dos Biomas existentes no país - Fonte: Portal Brasil, Ministério do Meio Ambiente.

Com uma extensão aproximada de 4.196.943 quilômetros quadrados, a Amazônia é o bioma que ocupa a maior parte do território brasileiro, ocupando totalmente os estados Acre, Amapá, Amazonas, Pará e Roraima, bem como 98.8% de Rondônia, 54% do Mato Grosso, 34% do Maranhão e 9% de Tocantins. Apresenta rios com fluxo intenso, florestas densas e com vegetação alta, chuvas torrenciais com uma precipitação média anual entre 1500mm e 1700mm aproximadamente (ANJOS *et al.*, 2012). Além disso caracteriza-se pelo clima equatorial, quente e úmido, com uma temperatura média de 25°C.

O bioma Amazônia é a maior reserva de biodiversidade do mundo. O Brasil possui, no bioma Amazônia, cerca de 60% da Bacia Amazônica, que é a maior bacia hidrográfica do mundo (cobre aproximadamente 6 milhões de Km², tem 1.100 afluentes e é responsável por escoar aproximadamente 20% do volume mundial de água doce). Estima-se que este bioma abrigue mais de um terço das espécies que vivem sobre a Terra (MMA, 2013a).

Na Amazônia, segundo Ignotti *et al.* (2007), a queima de biomassa decorrente do desmatamento e dos demais fatores é bastante intensa. Na agricultura brasileira o uso do fogo é utilizado como uma tecnologia agrícola nos sistemas de produção, desde o preparo das terras até a pré e pós colheita (MACHADO; ALVES, 2011) *apud* (MIRANDA *et al.*, 2006), entretanto, queimadas e desmatamentos estão entre os principais fatores que causam impactos sobre o clima e a biodiversidade (JUSTINO; ANDRADE, 2000).

A seguir será apresentada a aplicação em dados reais de focos de calor e precipitação média de chuva, detectados diariamente pelo satélite NOAA-15 (*National Oceanic and Atmospheric Administration*), entre o período de agosto de 2007 até agosto de 2011. Os dados foram fornecidos pelo Centro de Previsão do Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais - INPE.

Pode-se observar na Figura 10 a densidade dos pontos agregados para as variáveis quantidade diária de focos de calor no Bioma Amazônia em conjunto com a precipitação média diária. Em seguida, a Figura 11 apresenta em (A) os pontos agregados de focos de queimadas e precipitação média e em (B) a suavização do processo. Note que na primeira aplicação todas as análises gráficas são apresentadas a partir do mapa de Castilla-la Mancha, dadas as latitudes e longitudes, enquanto nesta segunda aplicação as análises gráficas são apresentadas de forma mais simples por não haver acesso a latitude e longitude dos focos de queimadas em nosso banco de dados, portanto as figuras serão apresentadas a partir das unidades de medida escalar densidade estimada (eixo “y”) versus quantidade de pontos agregados (eixo “x”).

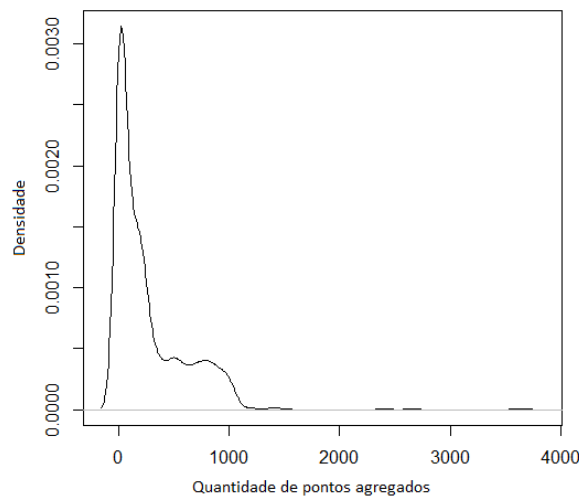


Figura 10: Densidade para os dados de quantidade de focos de calor e a precipitação média no Bioma Amazônia.

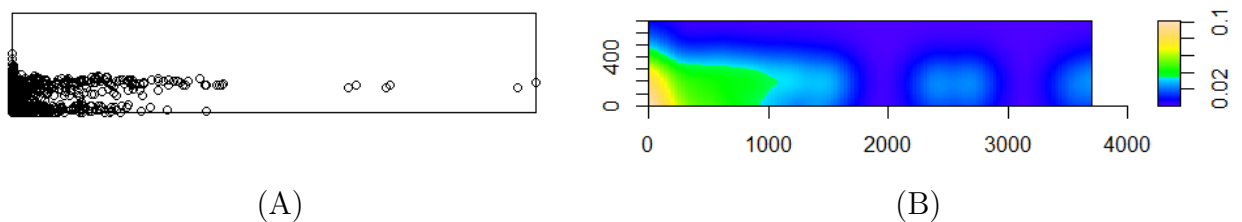


Figura 11: Dados de focos de calor em padrões pontuais (A) e com a suavização do processo (B).

A Figura 10 apresenta indícios de existência de uma grande quantidade de pontos agregados entre as variáveis em estudo, em virtude do “pico” na densidade entre o intervalo de 0 a 1000 pontos agregados. Esta ideia pode ser reforçada por meio da Figura 11, onde pode-se observar a grande massa de pontos agregados entre as variáveis focos de calor e precipitação média por meio dos pontos agrupados mostrados em 11(A) e da cor intensa em 11(B). Isso nos dá indícios de acreditar em uma boa associação entre estas duas variáveis, ou seja, acreditar que elas estão diretamente relacionadas. Assim, o ajuste dos dados por meio do LGCP parece ser indicado para estudo e análise do processo. A Figura 12 mostra o agrupamento entre a densidade e a os pontos de focos de calor.

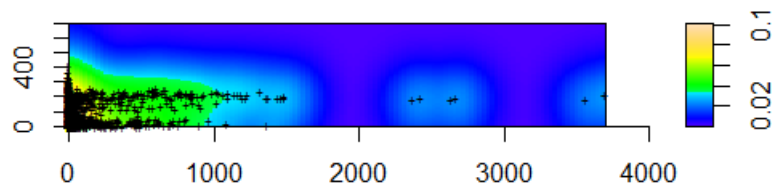


Figura 12: Suavização do processo em conjunto com os pontos agregados de focos de calor e precipitação média.

Inicialmente, tem-se a estimativa de $\mu = E[y(t)]$ conforme apresentado na Figura 13, em que $y(t)$ representará agora a quantidade de focos de calor, que é nossa variável resposta.

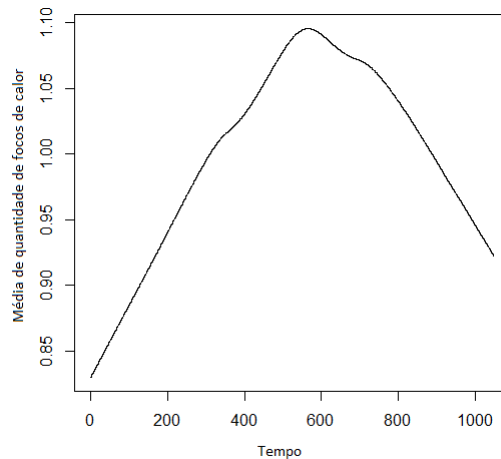


Figura 13: Estimativa para a média do processo $\mu = E[y(t)]$.

Assim como na primeira aplicação deste trabalho, estimam-se valores para os parâmetros por meio de uma análise em diversos gráficos para a função de covariância exponencial. A Figura 14 apresenta o comportamento das curvas “empírica” e “teórica” no gráfico da

função de covariância exponencial para σ^2 e ϕ estimados pelos valores 2.4 e 6.0, respectivamente.

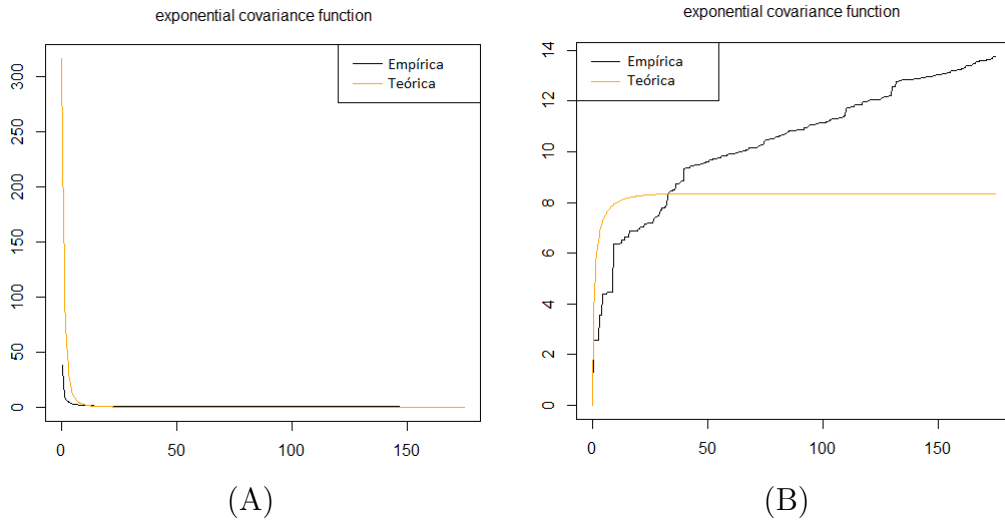


Figura 14: Gráficos da Função de Covariância Exponencial para as estimativas $\sigma^2 = 2.4$ (A) e $\phi = 6.0$ (B)

A estimativa para o parâmetro de correlação temporal θ também será definida de forma semelhante à análise utilizada para estimar os valores de σ^2 e ϕ . Foram analisados gráficos de autocovariância do processo na busca de uma melhor aproximação entre as curvas “empírica” e “teórica”, levando-nos a crêr que o melhor valor estimado para θ é 2.1. A Figura 15 apresenta o gráfico de autocovariância do processo para a melhor estimativa encontrada para θ , dados os valores estimados de σ^2 e ϕ .

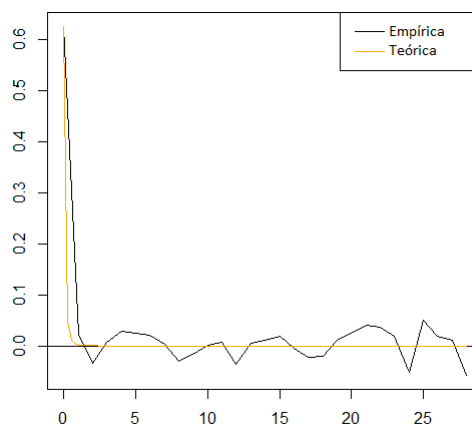


Figura 15: Estimativa para o parâmetro $\theta = 2.1$.

Pode-se observar por meio da Tabela 2 as estimativas para os parâmetros referentes a quantidade de focos de calor no Bioma Amazônia. Estas estimativas definem o LGCP, tornando possível a realizações de inferências a respeito das variáveis em estudo.

Parâmetro	Estimativa
σ^2	2,40
ϕ	6,00
μ	2,88
θ	2,10

Tabela 2: Parâmetros Estimativados a partir do LGCP aplicado aos pontos de focos de calor no Bioma Amazônia.

Estas estimativas definem o modelo do processo. Entretanto, como já visto anteriormente, não há ainda uma forma matemática fechada para o LGCP, por isto apenas as estimativas não são suficientes para realizar previsões sobre os dados, mas são fundamentais para a utilização dos métodos computacionais. Para a obtenção dos resultados a serem apresentados a seguir, foram realizadas 120000 iterações em um computador com processador Intel® Core™ i3-350M, 2.27GHz, com 4GB de memória RAM e HD de 500GB.

Para realizar previsões sobre os dados de focos de calor e precipitação média no Bioma Amazônia foi preciso, inicialmente, fazer uma minimização no grid destes dados. Anteriormente o grid era um retângulo com representatividade para a área em que houve queimadas no Bioma Amazônia, aproximadamente. Com esta extensão de grid não seria possível realizar uma simulação no Software R, seriam necessários muitos dias para que o Software realizasse todas as iterações. Faz-se então uma redução no grid, passando a ter aproximadamente o equivalente a uma área de $128Km \times 128Km$, com isso as iterações foram realizadas em aproximadamente $6h29min$. Obviamente, esta redução ocasionou também uma perda de informação no conjunto de dados, entretanto tornou-se possível realizar a aplicação do método computacional LGCP para a análise dos dados. O novo grid com padrões pontuais é apresentado na Figura 16.

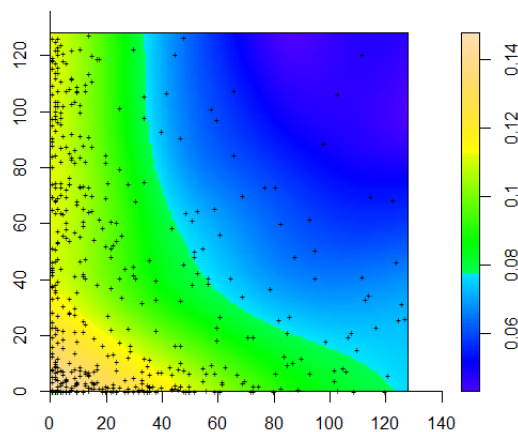


Figura 16: Padrões pontuais e densidade estimada com um grid de dados 128×128 .

Com o novo grid, os valores estimados anteriormente para os parâmetros do LGCP provavelmente não serão mais bem representativos para o modelo, em virtude dos dados “descartados”. Será feita uma nova estimativa com base no grid utilizado, a partir dos gráficos da função de covariância exponencial.

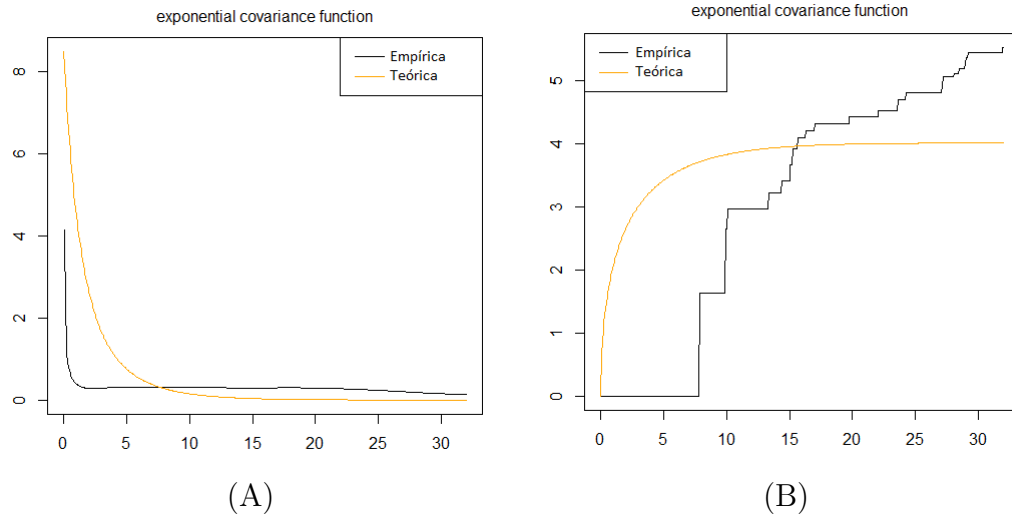


Figura 17: Gráficos da Função de Covariância Exponencial para as estimativas $\sigma^2 = 1, 5$ (A) e $\phi = 3, 6$ (B).

A Figura 17 apresenta os gráficos para a função de covariância exponencial, a partir do novo grid, para as estimativas $\sigma^2 = 1, 5$ e $\phi = 3, 6$. As novas estimativas obtidas para σ^2 e ϕ implicam em uma nova estimativa também para θ e μ . A Figura 18 apresenta o gráfico para a autocovariância do processo, a partir das novas estimativas.

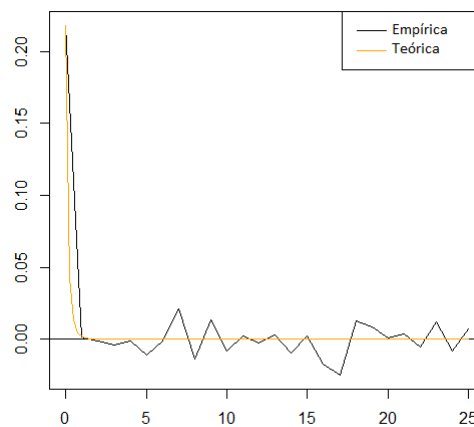


Figura 18: Estimativa para o parâmetro $\theta = 3, 8$.

A Tabela 3 apresenta as novas estimativas para os parâmetros referentes a quantidade de focos de calor no Bioma Amazônia. A partir destas estimativas serão feitas inferências a respeito dos dados.

Parâmetro	Estimativa
σ^2	1,500
ϕ	3,600
μ	1,125
θ	3,800

Tabela 3: Novos parâmetros estimatimados para LGCP.

Pode-se observar por meio da Tabela 4 uma comparação entre os valores obtidos com a previsão LGCP e os valores reais observados para a quantidade de focos de calor no Bioma Amazônia no período de tempo indicado.

Tempo		Quantidade de Focos de Calor	
Dia	Data	Observada	Reamostrada
817	23/03/2010	1	0
818	24/03/2010	1	0
819	25/01/2010	5	1
820	27/03/2010	1	2
821	29/03/2010	2	0
822	01/04/2010	4	2

Tabela 4: Comparação entre os valores observados e previstos para a quantidade de focos de calor no Bioma Amazônia.

A Figura 19 apresenta em (A) o risco relativo, em (B) os erros associados a este risco relativo e em (C) a intensidade média de Poisson.

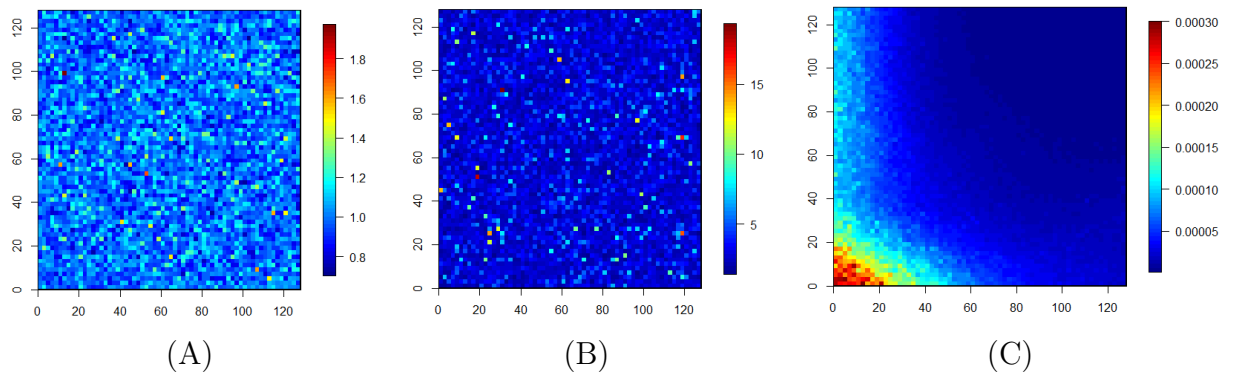


Figura 19: Gráfico para o risco relativo (A), os erros associados ao risco relativo (B) e a intensidade média de Poisson (C).

A partir da Figura 19 é possível verificar a qualidade no ajuste dos dados reamostrados. Nas Figuras 19(A) e (B) o risco relativo e os erros associados ao risco são dispostos de maneira aleatória, além disso, na Figura 19(C) a intensidade média de Poisson parece ser semelhante a densidade real dos dados amostrados vista na Figura 16, o que gera indícios para crer em um ajuste representativo para os dados reais.

4.2.1 Diagnóstico MCMC para os LGCP

Os métodos MCMC tem como ideia geral a obtenção de uma amostra da distribuição a posteriori e o cálculo das estimativas amostrais de características desta distribuição. Utiliza-se nesta Seção o método MCMC aplicado ao LGCP por meio de técnicas iterativas de simulação.

A Figura 20 apresenta a série temporal para a reamostra obtida por meio do método MCMC aplicado ao LGCP. Em 20(A) pode-se observar a série para os 20000 valores simulados a partir da cadeia. Em 20(B) é apresentada a série temporal para a representação de uma amostra da cadeia, onde pode-se observar a variação entre os valores da série do tempo para a amostra. Estes valores não permanecem perto de um mesmo ponto, ou seja, a série parece apresentar estacionariedade, gerando indícios de convergência.

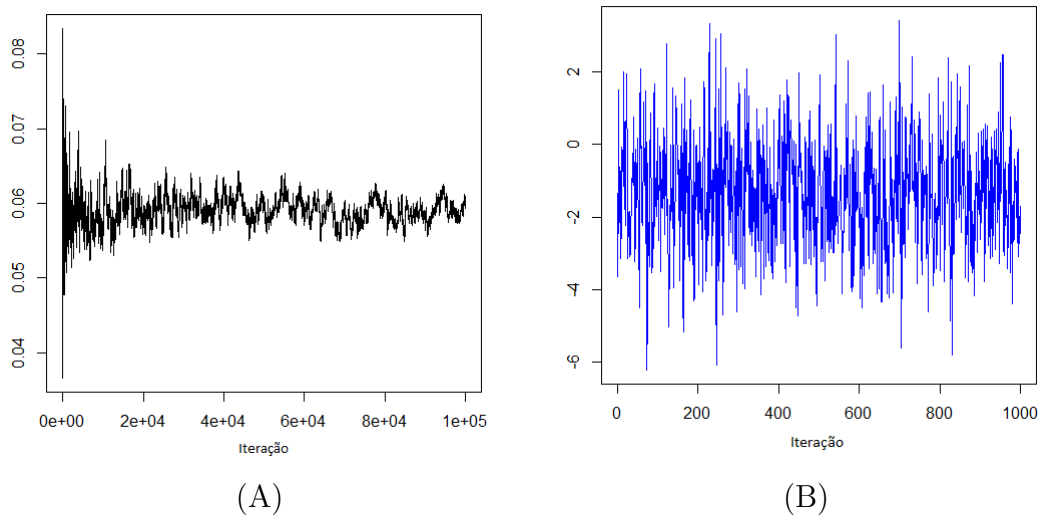


Figura 20: Série temporal para a representação da Cadeia de Markov. (A) Série para os 20000 valores gerados pela cadeia, (B) série para uma amostra de cadeia com 1000 valores simulados.

A Figura 21 apresenta o gráfico da função de autocorrelação para os valores gerados da cadeia. Os valores gerados por meio das iterações em uma cadeia de Markov são, por definição, correlacionados. Em algumas situações estes valores podem ser altamente correlacionados e em geral a autocorrelação será positiva, ocasionando pouco ganho em termos de informação no armazenamento de todos os valores simulados da cadeia, essa situação pode ser observada em 21(A). Torna-se, então, mais interessante realizar inferências a partir de uma amostra da cadeia, em busca de uma análise mais representativa para os dados reais. Em 21(B) pode-se observar a autocorrelação para uma amostra dos valores gerados.

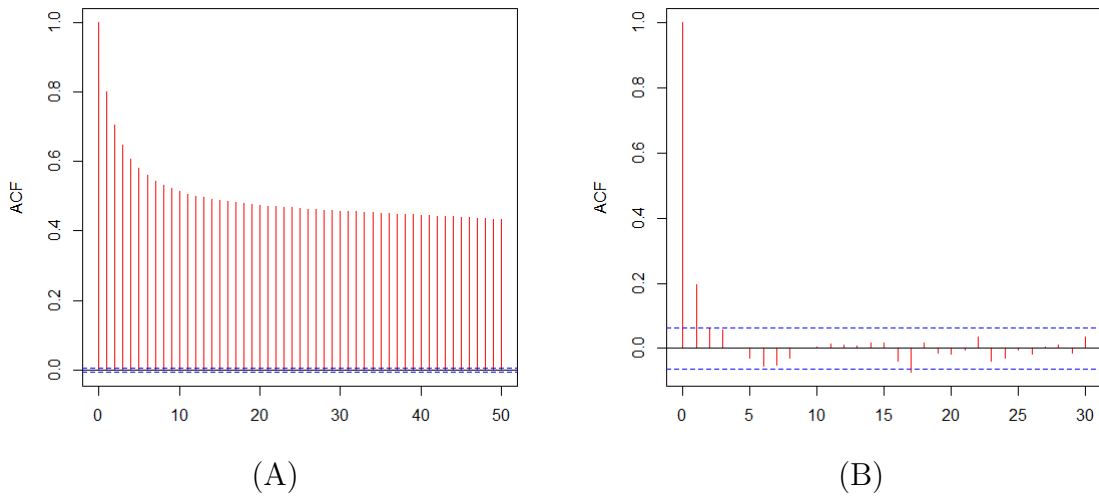


Figura 21: Gráficos para a função de autocorrelação. (A) Autocorrelação para os 20000 valores gerados pela cadeia, (B) autocorrelação para uma amostra de cadeia com 1000 valores simulados.

Uma avaliação da estacionariedade da cadeia pode ser obtida por meio de uma análise gráfica para a evolução da mediana no decorrer das iterações. A Figura 22 apresenta uma representação gráfica para a evolução no 1º e 3º quartis ao longo das iterações. As linhas pontilhadas representam as medianas da cadeia, enquanto a linha contínua representa os valores gerados. Pode-se observar que há indícios de estacionariedade na cadeia, em virtude da não-correlação apresentada entre as medianas e os dados por ela gerados.

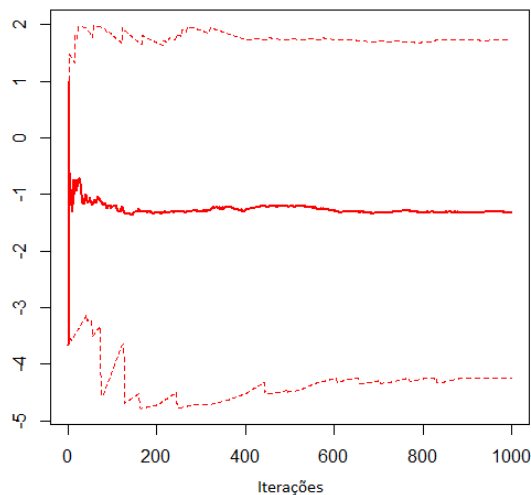


Figura 22: Representação gráfica para a evolução da mediana no decorrer das iterações.

O diagnóstico de Geweke considera uma cadeia de Markov de tamanho T para, utilizando técnicas de análise espectral, fornecer um diagnóstico para a ausência de convergência. Este diagnóstico baseia-se em um teste para a diferença de médias de primeira e última parte da cadeia de Markov, geralmente primeiros 10% e dos últimos 50%, onde

a hipótese nula testa se as duas partes da cadeia pertencem a uma mesma distribuição. Para isto, utiliza-se a estatística de teste Z – score normalizado. A Figura 23 apresenta a análise gráfica para a estatística de Geweke em uma amostra da cadeia, dando indícios de convergência em virtude do comportamento de cada Z – score dentro dos limites de $(-2, 2)$.

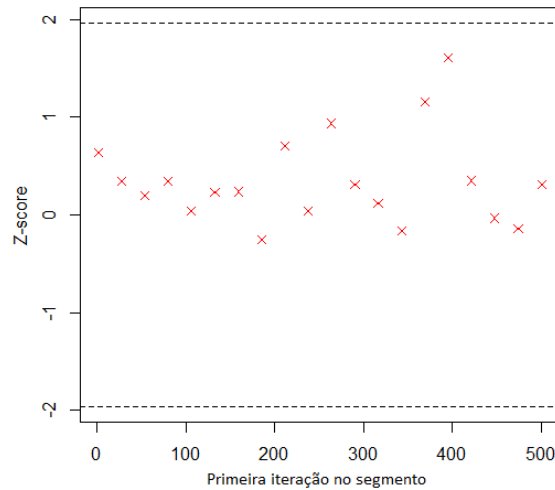


Figura 23: Gráfico da estatística de Geweke para um dos conjuntos de cadeias.

A Figura 24 apresenta o gráfico de diagnóstico Gelman-Rubin, que consiste em considerar M cadeias de Markov de tamanho T , geradas a partir de M pontos iniciais dispersos e comparar a variância das M cadeias com as médias das variâncias de cada uma das cadeias. As igualdades entre estas variâncias indicam a convergência da cadeia. A Partir da Figura 24 pode-se observar que há indícios de convergência para a cadeia de Markov em estudo.

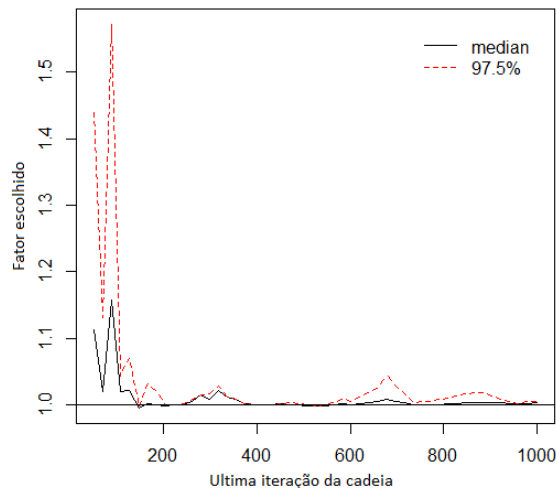


Figura 24: Gráfico da estatística de Gelman-Rubin para o conjunto de cadeias.

A Figura 25 apresenta o histograma para a quantidade de focos de calor no bioma Amazônia, a partir da reamostra obtida pelo MCMC aplicado ao LGCP.

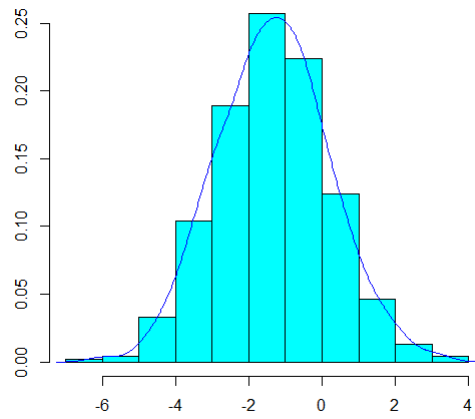


Figura 25: Histograma para a quantidade de focos de calor no bioma Amazônia.

5 Conclusão

A classe dos Modelos Log Gaussian no Processo de Cox torna possível a realização de um estudo amplo na análise de processos com padrões pontuais. Estudos dentro deste contexto podem ser analisados por meio de métodos já bastante conhecidos, entretanto, os modelos LGCP fornecem a possibilidade de um estudo mais abrangente e, consequentemente, mais cheio de informações em virtude da exploração dos diferentes níveis de agregação e diferentes informações fornecidas pelo conjunto de dados.

Através da aplicação dos LGCP em dados simulados, na Seção 4.1, foi possível visualizar claramente a heterogeneidade proposta pelo processo. Realizamos o estudo com base na latitude e longitude dos focos de incêndio, possibilitando uma boa análise visual dos mapas. Além disso, utilizamos dados provenientes de diferentes fontes (incêndios florestais causados por meio de relâmpagos, de forma acidental ou intencional e com causas desconhecidas) e com diferentes quantidades de informação fornecidas, mostrando assim a viabilidade na utilização do LGCP nas mais diversas áreas que utilizam processos de contagem como base de dados. Fizemos ainda a estimativa dos parâmetros do processo de forma empírica por meio do pacote *lgcp* (TAYLOR *et al.*, 2012), obtendo como estimativas $\sigma^2 = 1,3$, $\phi = 2,3$, $\mu = 0,845$ e $\theta = 3,4$.

Na Seção 4.2 o LGCP foi aplicado em dados reais das variáveis Quantidade de Focos de Queimadas e Precipitação Média de Chuva no Bioma Amazônia, mostrando a adequação do processo a dados provenientes de contagens como variável resposta (quantidade de focos de queimadas) em conjunto com um campo aleatório gaussiano (precipitação). Nesta aplicação exploramos os gráficos sem que houvesse a necessidade da latitude e longitude dos pontos, aplicando os padrões pontuais em uma janela com a área aproximada de pontos. Obtivemos as estimativas de forma satisfatória com $\sigma^2 = 1,5$, $\phi = 3,6$, $\mu = 1,125$ e $\theta = 3,8$. Além disso, utilizamos o método Bayesiano de reamostragem MCMC, possibilitando um bom diagnóstico com relação aos dados.

Os objetivos propostos neste trabalho foram cumpridos de forma satisfatória no estudo

sobre o LGCP, bem como nas aplicações realizadas em dados reais e em dados simulados, possibilitando previsões futuras para a incidência de pontos de incêndio no Bioma Amazônia e de quaisquer outros dados com padrões pontuais em que o LGCP possa ser aplicado. Logo, o trabalho teve resultados satisfatórios que contribuem positivamente para a tentativa de redução da incidência de Focos de Calor e pontos de Queimadas, minimizando, conseqüentemente, os impactos ambientais ocasionados por estes focos.

As maiores dificuldades encontradas na realização deste trabalho foram a análise Bayesiana do Processo, por não ter sido possível determinar uma forma fechada para a Distribuição a Posteriori dos parâmetros, assim como a análise computacional para o MCMC aplicado ao LGCP. Como estudos futuros, sugerimos métodos computacionais que tornem possível a análise dos processos com maior rapidez e sem que haja perda de informação proveniente dos dados.

Referências

- ALEXANDER, N. e. a. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics*, v. 1, n. 4, p. 453 – 463, 2000.
- ANJOS, P. S. dos; MORAIS, E. F.; NETO, J. G. C.; BARBOSA, N. F. M. Modelo de séries temporais para a predição de focos de calor no bioma amazônia. *Vigésimo Sinape, João Pessoa, Paraíba*, 2012.
- BARTOLUCCI, F. Efficient bayes factor estimation from the reversible jump output. *Biometrika*, v. 93, n. 1, p. 41 – 52, 2006.
- BAYES, T. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, v. 53, p. 370–418, 1763.
- BERNARDO, J. M.; SMITH, A. F. M. *Bayesian Theory*. New York: [s.n.], 1994.
- BRIX, A.; DIGGLE, P. J. Spatiotemporal prediction for log-gaussian cox processes. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, v. 63, p. 823–841, 2001.
- CRUTZEN, P. J.; ANDREAE, M. O. Biomass burning in the tropics: impact on atmospheric chemistry and biogeochemical cycles. *Science*, v. 250, p. 1669 – 1678, 1990.
- DIGGLE, J. P. *Statistical Analysis for Spatial Point Patterns*. 2. ed. London: Arnold.: [s.n.], 2003.
- EHLERS, R. S. *Inferência Bayesiana*. 2007. Disponível em: <<http://www.icmc.usp.br/~ehlers/bayes/>>.
- FREIRE, J. C. M. Métodos para estimar "odds ratio" através de abordagens clássica e bayesiana, em estudos epidemiológicos. *Dissertação de Mestrado em Biometria e Estatística Aplicada - Universidade Federal Rural de Pernambuco*, p. 62, 2001.
- GAMERMAN, D. Markov chain monte carlo: Stochastic simulation for bayesian inference. *Texts in Statistical Sciences. Chapman and Hall, London*, 1997.
- GAMERMAN, D.; LOPES, H. Markov chain monte carlo: Stochastic simulation for bayesian inference. *Texts in Statistical Sciences Series. CRC Press*, 2006.
- IGNOTTI, E.; HACON, S. S.; SILVA, A. M. C.; JUNGER, W. L.; CASTRO, H. Efeitos das queimadas na amazônia: método de seleção dos municípios segundo indicadores de saúde. *Revista Brasileira de Epidemiologia.*, p. 453 – 464, 2007.
- JEFFREYS, H. *Theory of Probability*. 3. ed. [S.l.: s.n.], 1961.

- JUSTINO, F.; ANDRADE, K. Programa de monitoramento de queimadas e prevenção de controle de incêndios florestais no arco do desflorestamento na amazônia (proarco). *XI Congresso Brasileiro de Meteorologia, SBMET*, p. 647 – 653, 2000.
- LARSON, H. *Introduction to Probability Theory and Statistical Inference*. [S.l.: s.n.], 1982.
- MACHADO, A. T.; ALVES, M. C. Análise e quantificação de focos de calor utilizando satélites noaa-12 e noaa-15 no estado de mato grosso. *Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR*, p. 7888 – 7894, 2011.
- MIRANDA, E. d.; MORAES, A. d.; OSHIRO, O. Queimada na amazônia brasileira em 2005. *Comunicado Técnico 18. EMBAPA, São Paulo*, 2006.
- MMA, G. d. B. *Ministerio do Meio Ambiente*. Fevereiro de 2013. Disponível em: <<http://www.mma.gov.br/biomas/amazônia>>.
- MMA, P. B. *Portal Brasil - Ministerio do Meio Ambiente*. Fevereiro de 2013. Disponível em: <<http://www.brasil.gov.br/sobre/meio-ambiente/geografia/biomas-brasileiros>>.
- MOLLER, J.; SYVERSVEEN, A.; WAAGEPETERSEN, R. Log gaussian cox processes. *Scand. J. Statist.*, v. 25, p. 451–482, 1998.
- MOLLER, J.; SYVERSVEEN, A. R.; WAAGEPETERSEN, R. P. Log gaussian cox processes. *Scandinavian Journal of Mathematics*, 1996.
- MOLLER, J.; SYVERSVEEN, A. R.; WAAGEPETERSEN, R. P. Log gaussian cox processes: A statistical model for analyzing stand structural heterogeneity in forestry. *In Proceedings First European Conference for Information Technology in Agriculture (H. Kure, I. Thysen and A. R. Kristensen, eds.) 339-342. Dept. Mathematics and Physics, Royal Veterinary and Agricultural Univ., Denmark.*, 1997.
- MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais*. São Paulo: Edgard Blucher, 2004.
- O'HAGAN, A. *Bayesian Inference*. Cambridge: [s.n.], 1994.
- PAEZ, M. S. Análise de modelos para a estimação e previsão de processos espaço-temporais. *Tese de doutorado - Universidade Federal do Rio de Janeiro, Instituto de Matemática*, 2004.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. Estatística bayesiana. *Fundação Calouste Gulbankian / Lisboa*, p. 446, 2003.
- POLLARD, W. E. Bayesian statistics for evaluation research an introduction. *Sage Publications New Delhi*, p. 241, 1986.
- QUEIROZ, N. Regressão logística - uma estimativa bayesiana aplicada na identificação de fatores de risco para hiv, em doadores de sangue. *Dissertação de Mestrado em Biometria e Estatística Aplicada - Universidade Federal Rural de Pernambuco*, p. 97, 2004.

- RIBEIRO, H.; ASSUNÇÃO, J. V. Efeitos das queimadas na saúde humana. *Estudos Avançados*, v. 16, n. 44, p. 125 – 148, 2002.
- ROBERT, C. P.; CASELLA, G. *Monte Carlo Statistical Methods*. New York: [s.n.], 1999.
- SARAIVA, C. A. Prioris para modelos probabilísticos discretos em ciências agrárias. *Dissertação de Mestrado em Biometria e Estatística Aplicada - Universidade Federal Rural de Pernambuco*, 2007.
- SIMPSON, D.; ILLIAN, J.; LINDGREN, F.; SORBYE, S. H.; RUE, H. Going off grid: Computationally efficient inference for log-gaussian cox processes. *Cornell University Library*, 2011.
- SMITH, A. F. M.; GELFAND, A. E. Bayesian statistics without tears: A sampling-resampling perspective. *The American Statistician*, v. 46, p. 84 – 88, 1992.
- SOUZA, A. D. P. Fundamentos de inferência bayesiana e aplicações. *Texto dos cursos de Matemática, Estatística e Ciência da Computação - FCT/UNESP*, p. 35, 2003.
- TAYLOR, B. M.; DAVIES, T. M.; ROWLINGSON, B. S.; DIGGLE, P. J. lgcp - an r package for inference with spatio-temporal log-gaussian cox processes. *Cornell University Library*, 2011.
- TAYLOR, B. M.; DAVIES, T. M.; ROWLINGSON, B. S.; DIGGLE, P. J. *lgcp: Log-Gaussian Cox Process*. [S.l.], 2012. R package version 1.0. Disponível em: <<http://CRAN.R-project.org/package=lgcp>>.
- TAYLOR, B. M.; DIGGLE, P. J. Inla or mcmc? a tutorial and comparative evaluation for spatial prediction in log-gaussian cox processes. *Cornell University Library*, 2012.
- TEAM, R. D. C. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2011. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org/>>.