

Uso da inferência bayesiana para obtenção de mapas de probabilidades para o Trimestre Chuvoso no Estado da Paraíba

Elias Silva de Medeiros^{a,1}, Paulo Justiniano Ribeiro Jr.^a

^a*Departamento de Ciências Exatas, ESALQ/USP, e-mail: eliasestatistica@gmail.com¹*

Resumo

O Estado da Paraíba possui um clima caracterizado pela irregularidade espacial e temporal da precipitação, o que ocasiona desastres climáticos. Em um estudo no qual se pretende analisar a dinâmica das variáveis regionalizadas, existem duas técnicas fundamentais para uma análise geoestatística, o semivariograma e a krigagem. Os semivariogramas são preferidos para caracterizar a estrutura de continuidade espacial da variável observado, por exigirem hipóteses de estacionariedade menos restritas (Hipótese Intrínseca). A técnica da Krigagem estima valores de variáveis para locais em que as mesmas não foram medidas a partir de valores adjacentes (vizinhos) interdependentes. A análise bayesiana considera a incerteza sobre todos os parâmetros existentes no modelo. Assim, pode-se realizar uma análise descritiva sobre cada um deles, sendo possível construir intervalos de credibilidades, considerando a incerteza sobre cada um do modelo. O presente trabalho foi desenvolvido com o objetivo principal de apresentar e implementar uma abordagem bayesiana no estudo geoestatístico aos dados de precipitação. De uso da inferência bayesiana foram estimados os parâmetros do modelo e em seguida a convergência dos mesmos foi verificada. Por meio do mapa de probabilidade, observou-se que as microrregiões Cariri/Curimataú se encontram dentro da área de ocorrência de chuvas mensal abaixo de 100 mm durante o Trimestre Chuvoso.

Palavras-chave: *Semivariogramas, Convergência, Intervalos de credibilidades.*

1. Introdução

O Estado da Paraíba, localizado no setor leste do Nordeste do Brasil, tem um clima que se caracteriza pela irregularidade espacial e temporal da precipitação, o que ocasiona desastres climáticos tais como secas, veranicos, enchentes, etc. Tal fato tem motivado o estudo dessas características sob os mais variados enfoques, já que ele serve de base à obtenção de informações essenciais a diversas atividades, tais como: dimensionamento de obras hidráulicas, irrigação, prevenção e controle de enchentes, etc.

Para o desenvolvimento deste artigo, as séries de dados diários de precipitação foram obtidas na Agência Executiva de Gestão das Águas do Estado da Paraíba (AESAs) responsável pelas informações da precipitação pluvial na Paraíba. Foram registradas as medidas das precipitações diárias, mensais, e anuais para 253 postos pluviométricos distribuídos nos 223 municípios paraibanos, no período de 1994 a 2011.

De acordo com Silva (2007) ficou caracterizada a divisão do Estado em duas grandes áreas com períodos chuvosos diferentes. Uma delas é ao Leste da Serra da Borborema representada pelas microrregiões do Litoral, Brejo e Agreste, e outra a Oeste com as microrregiões do

Cariri/Curimataú, Sertão e Alto Sertão. Verificou-se que o trimestre mais chuvoso nas microrregiões do Cariri/Curimataú, Sertão e o Alto Sertão é caracterizado pelos meses fevereiro, março e abril. Essas microrregiões são caracterizadas pela pouca quantidade de precipitação pluvial durante todo o ano.

O presente trabalho foi desenvolvido com o objetivo principal de apresentar e implementar uma abordagem bayesiana no estudo geostatístico aos dados de precipitação no Estado da Paraíba. Para isso, foram analisados os totais mensais durante 1994 e 2011 e em seguida foram extraídos os valores medianos para o período de maior ocorrência de chuvas (fevereiro, março e abril - Trimestre Chuvoso), nas microrregiões, Cariri/Curimataú, Sertão e Alto Sertão, que constituem a maior parte do Estado.

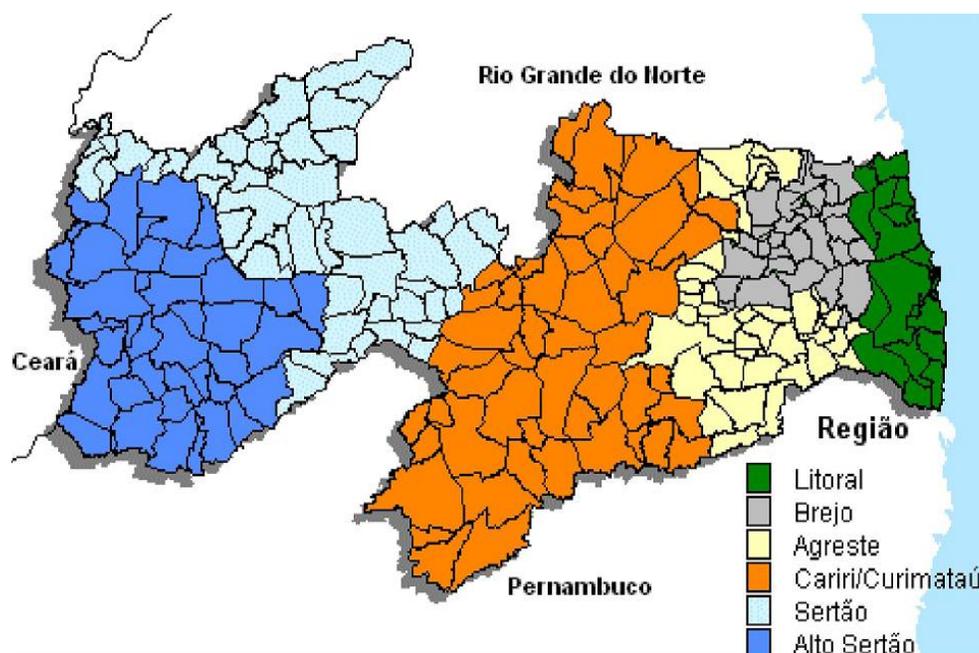


Figura 1: Localização geográfica das microrregiões do Estado da Paraíba. Fonte: Silva (2007)

2. Metodologia

Krige (1951)¹, trabalhando com dados de concentração de ouro, concluiu que somente a informação dada pela variância seria insuficiente para explicar o fenômeno em estudo. Para tal, seria necessário levar em consideração a distância entre as observações. A partir daí surge o conceito da *geoestatística*, que leva em consideração a *localização geográfica* e a *dependência espacial* (Camargo, 1997).

Matheron (1963) baseado nas observações de Krige, desenvolveu a teoria das variáveis regionalizadas, a partir dos fundamentos da geoestatística. Esta teoria vem sendo aplicada

¹Krige, D. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Chemical, Metallurgical and Mining Society of South Africa**, v.52, p.119-139, 1951.

em diversas áreas de pesquisa, tais como, ecologia, geologia, hidrologia, entre muitas outras, desde que as variáveis estudadas apresentem dependência espacial.

Seja $Z(s_i)$ definida como uma variável aleatória que assume diferentes valores Z em função da posição s_i dentro de uma região de estudo S , e representa pares de coordenadas (x,y) . Em um estudo no qual se pretende analisar a dinâmica das variáveis regionalizadas, existem duas técnicas fundamentais para uma análise de geoestatística, o *semivariograma* e a *krigagem*.

2.1. Semivariograma

Na geoestatística considera-se que o valor de cada ponto observado está relacionado com valores obtidos a partir de pontos localizados a uma certa distância, sendo assim, verifica-se que há uma grande relação entre os pontos tanto quanto menor for a distância entre eles. Para mensurar esta relação é definido o vetor de distâncias $\Delta \vec{h}$. A covariância pode ser uma das medidas que explica o grau de relação entre pontos numa certa direção e, embora a covariância exista entre todas as distâncias possíveis ao longo de \mathbf{h} , pode ser estipulado que somente sejam considerados valores entre pontos regularmente espaçados por múltiplos inteiros de $\Delta \vec{h}$.

A covariância entre valores determinados nessas distâncias separadas por $\Delta \vec{h}$ ao longo de \mathbf{h} é

$$C(\mathbf{h}) = E \{ [Z(s_i) - \mu] [Z(s_i + \mathbf{h}) - \mu] \} \quad (1)$$

em que μ é a média da variável regionalizada $Z(s_i)$.

Como a covariância depende do tamanho do vetor \mathbf{h} , se $\mathbf{h} = 0$, $C(\mathbf{h})$ será dita a variância. A expressão para variância neste caso, será:

$$Var [Z(s_i)] = E [Z(s_i) - \mu]^2 = C(0). \quad (2)$$

Os semivariogramas são preferidos para caracterizar a estrutura de continuidade espacial da característica avaliada, por exigirem hipóteses de estacionariedade menos restritas (Hipótese Intrínseca). Esta hipótese não é mais definida por meio das distribuições das variáveis aleatórias pontuais e sim por meio das diferenças entre elas (Ribeiro Junior, 1995).

A hipótese intrínseca é definida pelas seguintes condições:

- o incremento entre dois pontos tem esperança nula,

$$E [Z(s_{i+\mathbf{h}}) - Z(s_i)] = 0 \quad (3)$$

- para qualquer vetor \mathbf{h} o incremento tem uma variância finita independente da posição s_i .

$$Var [Z(s_i + \mathbf{h}) - Z(s_i)] = E [Z(s_i + \mathbf{h}) - Z(s_i)]^2 = 2\gamma(\mathbf{h}) \quad (4)$$

em que, $2\gamma(\mathbf{h})$ a definição de variograma.

Dividindo a Eq. 4 por dois, tem-se então a função de semivariograma $\gamma(\mathbf{h})$, a qual depende do vetor de distâncias $\Delta \vec{h}$. Espera-se que $\gamma(\mathbf{h})$ aumente com a distância \mathbf{h} até um valor máximo no qual a semivariância estabiliza.

Para construir um semivariograma é necessário dispor de um conjunto de valores obtidos em intervalos regulares dentro de um mesmo suporte geométrico. Sendo $s_1, s_2, \dots, s_i, \dots, s_n$, valores de uma variável regionalizada a seguinte expressão fornece a estimativa não tendenciosa da semivariância:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2} \frac{\sum_{i=1}^n [Z(s_i) - Z(s_i + \mathbf{h})]^2}{n(\mathbf{h})} \quad (5)$$

em que $n(\mathbf{h})$ é o número de pares de valores medidos $Z(s_i)$, $Z(s_i + \mathbf{h})$ separados por \mathbf{h} .

Em diferentes fenômenos que envolva geoestatística, há vários modelos de ajustes aplicáveis. Estes por sua vez são denominados, modelos teóricos, que tem o papel de fornecer soluções confiáveis para o estimador linear - *krigagem*. Existem dois tipos de modelos de variogramas teóricos: os modelos com e sem patamar e, estes podem, ou não, apresentar o efeito pepita. Dentre os que apresentam patamar os mais utilizados são: o modelo esférico (Matheron), o modelo exponencial (Formery) e o modelo gaussiano (Gauss). Entre os modelos sem patamar encontram-se o Linear e o Logarítmico (Wijs), (Oliveira, 2003).

O efeito pepita (τ^2) corresponde a cota do ponto onde o semivariograma corta o eixo das ordenadas, o alcance (ϕ) corresponde ao conceito de dependência espacial, marcando a distância a partir da qual as amostras se tornam independentes, o patamar ($\sigma^2 + \tau^2$) corresponde ao ponto onde toda semivariância da amostra é de influência aleatória, correspondendo a variância total obtida pela estatística clássica. Quando o efeito pepita (τ^2) for aproximadamente igual ao patamar ($\sigma^2 + \tau^2$), denomina-se efeito pepita puro, demonstrando que a amostra não recebe influência espacial Trangmar et al. (1985). A seguir, serão apresentadas as expressões dos modelos de correlações do semivariograma.

- Os modelos mais clássicos com patamar são,

1. Modelo Esférico ou Modelo de Matheron, é o modelo mais comum, podendo-se afirmar que equivale à função de distribuição normal da estatística clássica,

$$\rho(\mathbf{h}) = \begin{cases} \tau^2 + \sigma^2 \left[\frac{3}{2} \left(\frac{\mathbf{h}}{\phi} \right) - \frac{1}{2} \left(\frac{\mathbf{h}}{\phi} \right)^3 \right] & \text{se } 0 \leq \mathbf{h} \leq \phi \\ \tau^2 + \sigma^2 & \text{se } \mathbf{h} > \phi \end{cases} \quad (6)$$

2. Modelo cardinal-seno (modelo wave) - pertencente a família dos modelos “hole effect”, o mesmo apresenta um comportamento isotrópico,

$$\rho(\mathbf{h}) = \begin{cases} \tau^2 + \sigma^2 \left[1 - \frac{\phi}{\mathbf{h}} \sin \left(\frac{\mathbf{h}}{\phi} \right) \right] & \text{se } 0 \leq \mathbf{h} \leq \phi \\ \tau^2 + \sigma^2 & \text{se } \mathbf{h} > \phi \end{cases} \quad (7)$$

3. Modelo Exponencial ou de Formery, este apresenta um comportamento linear na origem, alcança seu patamar somente assintoticamente e é representado por:

$$\rho(\mathbf{h}) = \begin{cases} \tau^2 + \sigma^2 \left[1 - e^{-\frac{\mathbf{h}}{\phi}} \right] & \text{se } 0 \leq \mathbf{h} \leq \phi \\ \tau^2 + \sigma^2 & \text{se } \mathbf{h} > \phi \end{cases} \quad (8)$$

4. O Modelo Gaussiano ou Parabólico, este apresenta um comportamento parabólico na origem e tem como equação o modelo a seguir

$$\rho(\mathbf{h}) = \begin{cases} \tau^2 + \sigma^2 \left[1 - e^{-\frac{\mathbf{h}^2}{\phi^2}} \right] & \text{se } 0 \leq \mathbf{h} \leq \phi \\ \tau^2 + \sigma^2 & \text{se } \mathbf{h} > \phi \end{cases} \quad (9)$$

- Os modelos sem patamar são,

1. O modelo linear, o qual é definido para o intervalo $0 < \phi < 2$ e $\sigma^2 > 0$, e tem a seguinte expressão,

$$\gamma(h) = \sigma^2 \times \mathbf{h}^\phi \quad (10)$$

Há ainda outros modelos sem patamar, que são, o Modelo de Potência e o Logarítmico ou de Wijs.

- Outras metodologias que estudam a dependência espacial e a obtenção de semivariogramas ajustados foi proposto também por Diggle e Ribeiro Junior (2002), eles sugerem este estudo por meio da modelagem utilizando a distribuição Gaussiana. Comentam sobre as diferentes estruturas de covariância utilizadas no estudo espacial e sugerem uma nova estrutura da função de correlação denominada de família de *Matérn*. A função de correlação de Matérn é definida por:

$$\rho(\mathbf{h}; \phi, \nu) = \frac{2^{2-\nu}}{\Gamma(\nu)} \left(\frac{\mathbf{h}}{\phi} \right)^\nu K_\nu \left(\frac{\mathbf{h}}{\phi} \right), \quad \mathbf{h} > 0, \quad (11)$$

em que $K_\nu(\cdot)$ é a função *Bessel* modificada de terceiro grau de ordem $\nu > 0$ e $\Gamma(\cdot)$ é a função *Gamma*. O primeiro desses parâmetros (ϕ) está relacionado ao alcance das correlações, sendo que maiores valores indicam dependências espaciais de maior alcance. O segundo parâmetro (ν) está relacionado à suavidade do processo, de forma que, quanto maior ν , maior será a suavidade.

Cambardella et al. (1994), propuseram os seguintes intervalos para avaliar a % da semivariância do Efeito Pepita: $\leq 25\%$ - **forte dependência espacial**; entre 25% e 75% - **moderada dependência espacial** e $\geq 75\%$ - **fraca dependência espacial**, denominado de **IDE** (Índice de Dependência Espacial) e pode ser determinado de acordo com a seguinte expressão:

$$IDE = \frac{\tau^2}{\sigma^2 + \tau^2} 100 \quad (12)$$

2.2. Seleção dos modelos

A estimação dos parâmetros de modelos baseados em verossimilhança produz de maneira geral estimadores não viesados e eficientes quando aplicados para grandes amostras. Diggle e Ribeiro Junior (2002), afirmam ainda que ao utilizar estes métodos deve-se considerar a pressuposição Gaussiana. Ilustram este estudo com um modelo linear Gaussiano em que a média é definida por um modelo de regressão linear.

O teste de razão de verossimilhanças compara os modelos com e sem a estrutura de dependência espacial, em que a razão de verossimilhanças somente é válida para modelos sequencialmente reduzidos (Rao, 1973). A estatística do teste da razão de verossimilhanças é definida como:

$$LRT = -2\ln\left(\frac{ML2}{ML1}\right) = 2[\ln(ML1) - \ln(ML2)] \quad (13)$$

em que $ML1$ é a máxima verossimilhança no modelo sem o componente espacial e $ML2$ é a verossimilhança no modelo com o componente espacial. Se o valor LRT da Eq. 13 for menor do que o valor do χ^2 tabelado, então há fortes evidência de que o modelo com menor número de parâmetro é o mais adequado.

Para a avaliação da adequação dos modelos espaciais foi utilizado o critério de informação de Akaike - AIC. Esse critério, em geral, tem sido utilizado para comparação de modelos com diferentes estruturas de covariâncias. O AIC é baseado na teoria de decisão e é definido como a quantidade:

$$AIC = -2l + 2p \quad (14)$$

em que, l é o logaritmo natural da função de máxima verossimilhança e p é o número de parâmetros do modelo. De acordo com este critério o melhor modelo ajustado será aquele que possuir o menor valor de AIC .

2.3. Krigagem

A técnica da Krigagem estima valores de variáveis para locais em que as mesmas não foram medidas a partir de valores adjacentes (vizinhos) interdependentes. Para a aplicação da krigagem, suponha que sejam conhecidas as realizações $z(s_1), z(s_2), \dots, z(s_k)$; e o semivariograma já tenha sido determinado, sendo assim, o interesse está em estimar \hat{Z} na posição s_0 (Maestre et al., 2007).

A equação de estimação da krigagem é dada por:

$$\hat{Z}(s_0) = \sum_{i=1}^k \lambda_i z(s_i) \quad (15)$$

em que, \hat{Z} é o estimador para o ponto (s_0), k é o número de vizinhos utilizados na estimação, λ_i é o peso atribuído a cada vizinho e, $z(s_i)$ é o valor observado em cada vizinho.

A krigagem, ao contrário de outros métodos de interpolação, estima uma matriz de covariância espacial que permite determinar pesos atribuídos às diferentes amostras, a redundância dos dados, a vizinhança a ser considerada no procedimento inferencial e o erro associado ao valor estimado (Goovaerts, 1997).

Existem várias técnicas de estimativas para a krigagem e a mais usual é a krigagem ordinária. Esta é “linear” porque suas estimativas são combinações lineares ponderadas dos dados disponíveis; é “não-viciada” porque busca o valor de erro ou resíduo médio igual a 0; e é “melhor” porque minimiza a variância dos erros (Isaaks e Srivastava, 1989).

E para obter a variância mínima sob a condição de $\sum_{i=1}^k \lambda_i = 1$, introduz-se o multiplicador de Lagrange para a dedução das equações e o sistema de krigagem resultante é:

$$\sum_{i=1}^k \lambda_i \gamma(s_i, s_0) + \mu = \gamma(s_i, s_0) \quad (16)$$

sendo μ é o multiplicador de Lagrange. O estimador da variância é da forma,

$$\hat{\sigma}^2 = \mu + \sum_{i=1}^k \lambda_i \gamma(s_i, s_0) \quad (17)$$

2.4. Abordagem Bayesiana

Segundo Diggle e Ribeiro Junior (2002) um modelo geoestatístico é especificado conjuntamente para um processo espacial contínuo não observado $\{S(x) : x \in \mathfrak{R}^2\}$ e para um conjunto de dados observados $Y(x)$ nas localizações x , condicionado ao processo espacial com efeito nas mesmas localizações. Sendo $\boldsymbol{\theta} = \{\theta_1; \dots; \theta_p \in \mathfrak{R}\}$ o conjunto de parâmetros desconhecidos e não observáveis no modelo, então:

$$P(Y(x); S(x) | \boldsymbol{\theta}) = \frac{P(Y(x); S(x); \boldsymbol{\theta})}{P(\boldsymbol{\theta})} = P(Y(x) | S(x); \boldsymbol{\theta}) P(S(x) | \boldsymbol{\theta}) \quad (18)$$

A distribuição preditiva do processo $S(x)$ é definida condicionalmente às observações $Y(x)$ como:

$$P(S(x) | Y(x)) = \frac{P(S(x); Y(x))}{P(Y(x))} \quad (19)$$

Para se obter a distribuição conjunta de $(S(x); Y(x))$ na Equação 19, integra-se a distribuição de $(S(x); Y(x); \boldsymbol{\theta})$ sobre o espaço paramétrico, ou seja:

$$P(S(x); Y(x)) = \int_{\boldsymbol{\theta}} P(S(x); Y(x); \boldsymbol{\theta}) P(\boldsymbol{\theta}; S(x)) d\boldsymbol{\theta}$$

que, substituída em 19 resulta:

$$P(S(x) | Y(x)) = \int_{\boldsymbol{\theta}} P(S(x) | Y(x); \boldsymbol{\theta}) P(\boldsymbol{\theta} | Y(x)) d\boldsymbol{\theta} \quad (20)$$

Aplicando-se o Teorema de Bayes, a distribuição a posteriori dos parâmetros é dada por:

$$P(\boldsymbol{\theta} | Y(x)) = \frac{P(\boldsymbol{\theta}; Y(x))}{P(Y(x))} = \frac{P(Y(x) | \boldsymbol{\theta}) P(\boldsymbol{\theta})}{P(Y(x))}. \quad (21)$$

O termo $P(Y(x) | \boldsymbol{\theta})$ é a função de verossimilhança de $Y(x) | \boldsymbol{\theta}$ com distribuição gaussiana multivariada e $P(\boldsymbol{\theta})$ é a distribuição a priori de $\boldsymbol{\theta}$ que expressa o conhecimento prévio acerca da distribuição de probabilidades dos parâmetros.

Na prática, a escolha da distribuição a priori dos parâmetros do modelo é um assunto delicado na inferência bayesiana pois ela se dá, ora por conhecimento (objetivo ou subjetivo) da sua distribuição, ora por uma conveniência que resulte em uma distribuição a posteriori com solução analítica. Distribuições a priori que resultem em uma posteriori da mesma família, são chamadas de prioris conjugadas e sua escolha se dá devido à tratabilidade analítica decorrente e a uma conveniência computacional (Silva, 2008).

A análise bayesiana considera a incerteza sobre todos os parâmetros existentes no modelo. Assim pode-se realizar uma análise descritiva sobre cada parâmetro, sendo possível construir intervalos de credibilidades, considerando a incerteza sobre cada parâmetro do modelo.

Na metodologia bayesiana associa-se uma distribuição a priori para os parâmetros $(\beta, \sigma^2, \phi, \tau^2)$. Em geral, para o parâmetro ϕ utiliza-se como priori uma Distribuição Gamma, para os parâmetros σ^2 e τ^2 , uma Distribuição Gamma inversa, e para β uma Distribuição Normal.

3. Resultados e Discussão

Foi construído o gráfico que descreve o valor estimado para o parâmetro da transformação Box Cox, Figura 2(a) e, pode ser observado que o número 1 não faz parte do intervalo sugerindo então que seja realizada uma transformação no atributo. A Figura 2(b) apresenta os gráficos de dispersão e do histograma da variável precipitação. O histograma da variável após a transformação Box Cox apresentou pouca assimetria.

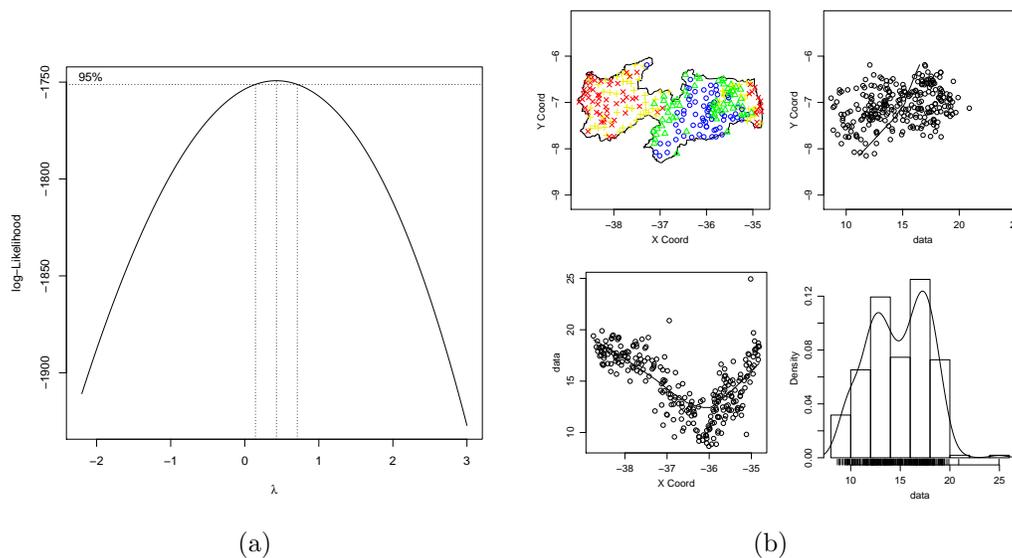


Figura 2: Gráfico para a estimação do parâmetro da transformação Box Cox (a) e, gráfico de dispersão e histograma após a transformação Box Cox(b)

Considerando que as observações medidas apresentam dependência espacial, os dados foram analisados sugerindo alguns modelos que incluíram a estrutura de dependência $S(x)$. O modelo selecionado por meio do teste da razão de verossimilhanças é descrito a seguir.

$$Y_i = \beta_0 + \beta_1 f_1(s_i) + \beta_2 f_2(s_i) + S(x) + \varepsilon_i \quad (22)$$

sendo f_1 o efeito da latitude e, f_2 o efeito quadrático da longitude.

Tabela 1: Estimativas de máxima verossimilhança dos parâmetros associados ao modelo e o valor do AIC

Função de correlação	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\tau}^2$	$\hat{\sigma}^2$	$\hat{\phi}$	AIC
Matérn (kappa = 1,0)	2008,20	109,15	1,49	1,26	3,16	0,33	2383,19
Gaussian	1939,08	105,77	1,45	1,39	2,59	0,37	2383,21
Exponencial	2044,96	110,89	1,51	1,05	3,78	0,74	2383,10
Esférico	2115,20	114,66	1,56	1,16	4,53	1,79	2381,07
Wave	2184,71	118,29	1,61	2,06	2,89	0,44	2414,09

A Tabela 1 apresenta os resultados das estimativas dos parâmetros e os valores AIC para o modelo 22. Pelo Critério de Informação Akaike (AIC), o modelo esférico foi o que melhor se ajustou ao semivariograma experimental.

Uma forma não-paramétrica de identificar a dependência espacial para o modelo 22 foi utilizando do envelope simulado, realizando 1000 simulações, Figura 3(a). Por meio desta figura verifica-se que há uma quantidade considerável de pontos do variograma fora do envelope simulado, isto mostra que há fortes evidências para uma dependência espacial. O modelo ajustado apresentou um $IDE = 20\%$), o que representa uma forte dependência espacial.

A Figura 3(b) mostra o mapa dos valores ajustados pelo modelo para os valores medianos da precipitação pluviométrica no Trimestre Chuvoso. Esse mapa indica as regiões em que ocorrem as maiores e as menores estimativas pelo método da krigagem ordinária. Como pode ser observado no mapa, as microrregiões do Cariri/Curimatá apresentaram baixos índices pluviométrico, já no Sertão e no alto-Sertão foram verificadas as maiores médias mensais da precipitação pluviométrica, que pode estar diretamente associada com o período chuvoso da região.

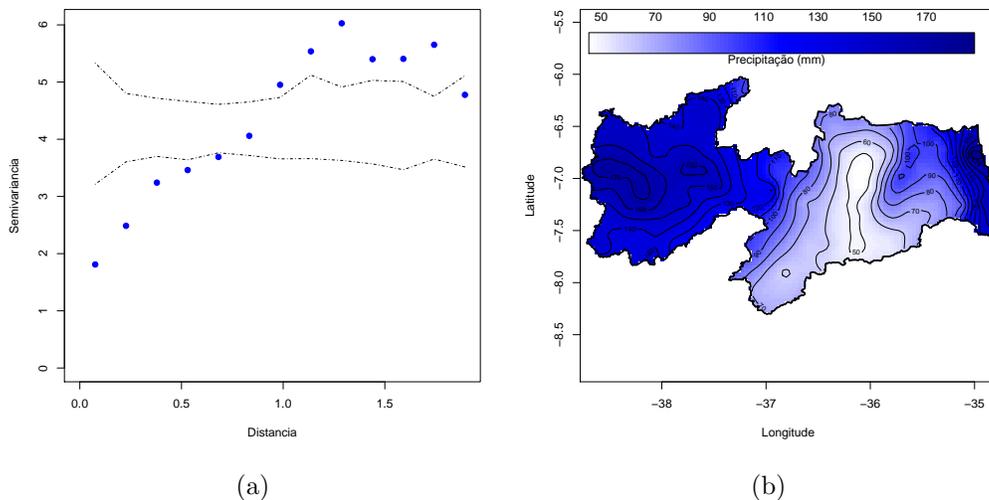


Figura 3: Envelope simulado e variograma empírico (a), e interpolação por krigagem ordinária dos valores preditos pelo modelo 22 para o Trimestre Chuvoso (b)

Na Tabela 2 é apresentado um resumo da distribuição a posteriori de todos os parâmetros do modelo 22. Observa-se que as estimativas para os parâmetros pelo método Bayesiano foram próximas das estimativas realizadas na abordagem clássica, com maior discrepância para σ^2 .

Tabela 2: Resumo da distribuição a posteriori dos parâmetros do modelo 22

Parâmetros	Média	2,5%	Mediana	97,5%
β_0	2078,68	1273,84	2074,29	2896,69
β_1	112,83	69,10	112,64	157,16
β_2	1,54	0,95	1,54	2,14
σ^2	1,68	1,35	1,67	2,02
ϕ	1,63	1,20	1,60	2,20
τ^2	1,70	1,42	1,69	2,03

Os seguintes procedimentos foram utilizados para a obtenção de alguns resultados por meio de uma abordagem bayesiana. Diante dos algoritmos MCMC, foram executadas 10 mil simulações e, o período de aquecimento foi de 1000 simulações.

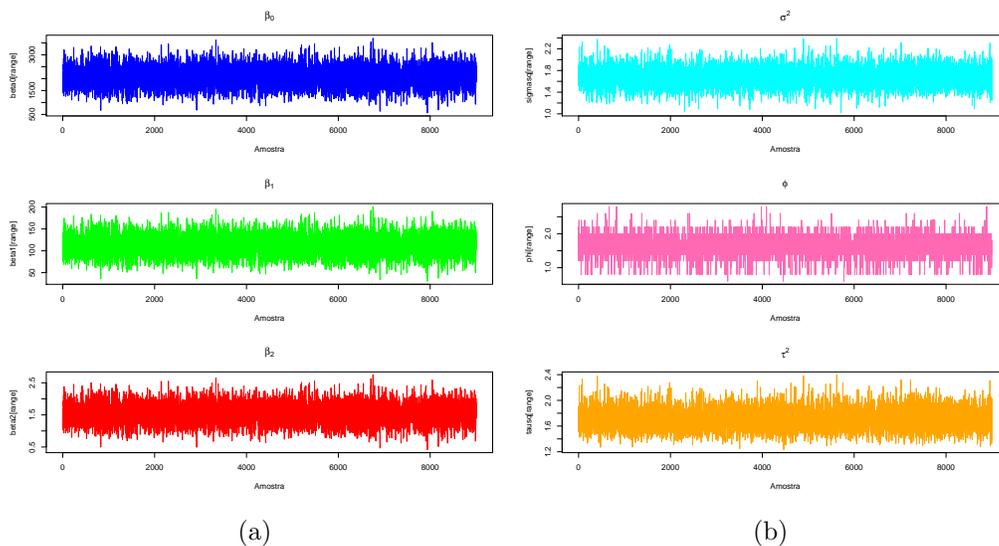


Figura 4: Traço de uma cadeia a posteriori dos parâmetros

A Figura 4 apresenta os gráficos dos traços de uma cadeia a posteriori para os parâmetros do efeito fixo (a) e dos efeitos aleatórios (b) do modelo 22 ajustado. Logo, observo-se que a linha foi aleatória e não ficou fixa em nenhuma região ao longo das iterações, permitindo verificar que existe fortes evidências para a convergência da cadeia.

Os resultados obtidos da predição e variância da predição com a estrutura de covariância Esférica, utilizando a inferência bayesiana, encontram-se na Figura 5.

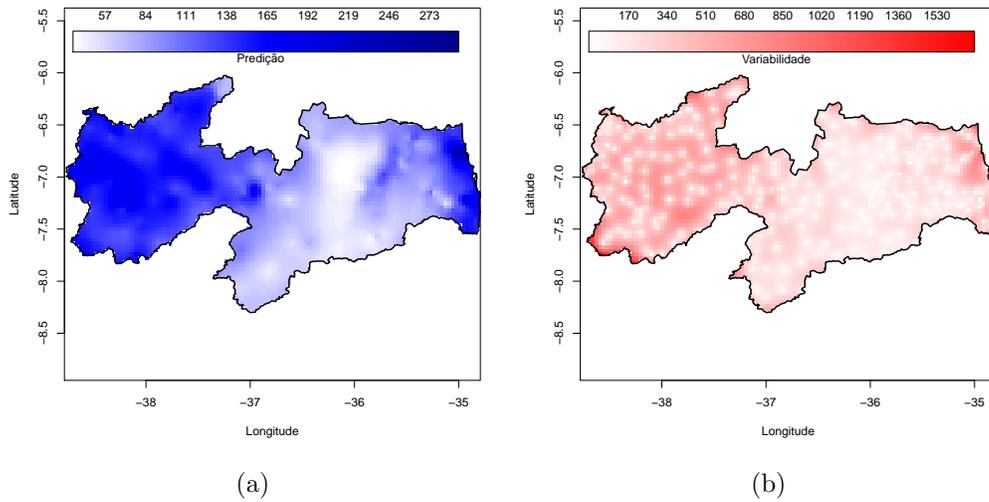


Figura 5: Mapa da predição (a) e da variância (b) da predição por inferência bayesiana

Observa-se que as variâncias estão bem próximas e que nos pontos mais claros estas são menores (Figura 5(b)).

Silva (2007) realizou um estudo sobre a precipitação no Estado da Paraíba, neste ela analisou o comportamento pluviométrico ocorrido entre os anos de 1975 a 2005, e verificou que a precipitação pluvial no Trimestre Chuvoso, em média, foi superior a 100 mm. De base nesta informação encontra-se na Figura 6 um mapa de probabilidade de ocorrência de chuvas inferior a 100 mm. Verifica-se, em especial, nas microrregiões do Cariri/Curimataú um forte agravante, pois a probabilidade é bastante alta para ocorrências da precipitação mensal ser inferior a 100 mm.

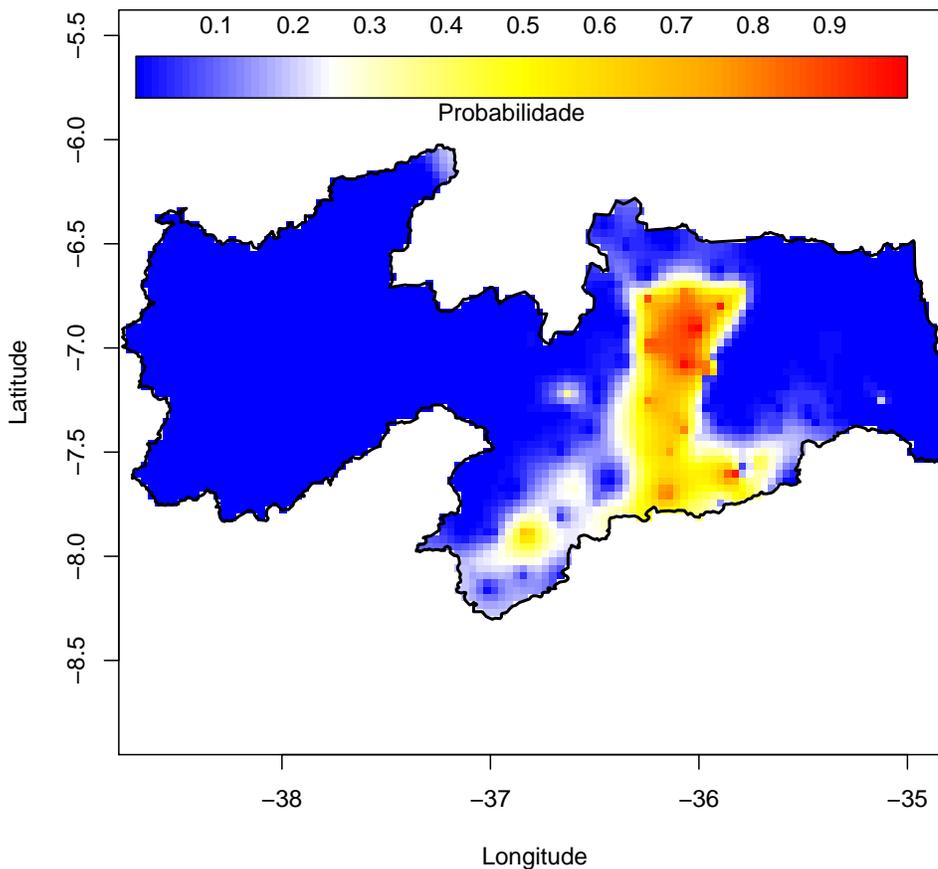


Figura 6: Mapa de probabilidade da precipitação mediana abaixo de 100 mm no Trimestre Chuvoso por inferência bayesiana

4. Conclusões

A metodologia desenvolvida foi baseada na teoria de verossimilhança e, principalmente na inferência bayesiana, uma vez que pode-se considerar nas predições a incerteza associada à estimação dos parâmetros. O modelo que melhor se ajustou aos dados foi o Esférico. Verificou-se que houve uma forte dependência espacial. De uso da inferência bayesiana foram estimados os parâmetros do modelo e em seguida a convergência dos mesmos foi verificada. Por meio do mapa de probabilidade, observou-se que as microrregiões, Cariri/Curimataú, se encontram dentro da área de ocorrência de chuvas mensal abaixo de 100 mm durante o Trimestre Chuvoso.

Referências

Camargo, E. C. G., 'Junho 1997. **Desenvolvimento, implementação e teste de procedimentos geoestatísticos (krigeagem) no sistema de processamento de informações**

- georreferenciadas (Spring)**. Dissertação (mestrado), Instituto Nacional de Pesquisas Espaciais - INPE, São Jos dos Campos.
- Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., Konopka, A. E., 1994. Field-scale variability of soil properties in central iowa soils. **Soil Sci. Soc. Am. J.** 58, 1501 – 1511.
- Diggle, P. J., Ribeiro Junior, P. J., 2002. Bayesian inference in gaussian model-based geostatistics. **Geographical and Environmental modelling** 6, 131–148.
- Goovaerts, P., 1997. **Geostatistics for natural resources evaluation**. New York: Oxford University Press.
- Isaaks, E. H., Srivastava, R. M., 1989. **An Introduction to Applied Geostatistics**. New York: Oxford University Press.
- Maestre, M. R., França, E. J., Araújo, A. L. L., Sartório, S. D., Barbin, D., 2007. **Concentrações de potássio em parte da parcela permanente do parque estadual Carlos Botelho situado na cidade Sete Barras - SP**. In: Anais da 52a RBras e 12o SEAGRO. Universidade Federal de Santa Maria.
- Matheron, G., 1963. Principles of geostatistics.principles of geostatistics. **Economic Geology**. 58, 1246–1266.
- Oliveira, M. C. N., 2003. **Métodos de estimação de parâmetros em modelos geostatísticos com diferentes estruturas de covariâncias: uma aplicação ao teor de cálcio no solo**. Tese de Doutorado, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba.
- Rao, C. R., 1973. **Linear statistical inference and its applications**. New York: John Wiley & Sons.
- Ribeiro Junior, P. J., 1995. **Métodos geoestatísticos no estudo da varabilidade espacial de parâmetros do solo**. Dissertação, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo., Piracicaba.
- Silva, E. A. A., 2008. **Aplicação de métodos geoestatísticos multivariados em problemas demapeamento de variáveis do sistema solo-planta**. Tese de Doutorado, Universidade Federal do Paraná.
- Silva, L. L., 2007. **Precipitações pluviais da pré-estação chuvosa no período chuvoso e suas influências na produtividade agrícola da Paraíba**. Dissertação, Universidade Federal da Paraíba, Piracicaba.
- Trangmar, B. B., Yost, R. S., Uehara, G., 1985. Application of geostatistics to spatial studies of soil properties. **Advances in Agronomy**. 38, 45–94.