



Universidade Estadual da Paraíba  
Centro de Ciências e Tecnologia  
Departamento de Estatística

Isaac Correia Marinheiro

# **Inferência espacial para a determinação da distribuição de nutrientes no solo em sistemas agroflorestais.**

Campina Grande - PB  
Dezembro de 2013

Isaac Correia Marinheiro

# **Inferência espacial para a determinação da distribuição de nutrientes no solo em sistemas agroflorestais.**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de Bacharel em Estatística.

Orientador:

Dr. Ricardo Alves de Olinda

Campina Grande - PB

Dezembro de 2013

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL – UEPB

M338i Marinheiro, Isaac Correia.  
Inferência espacial para a determinação da distribuição de nutrientes no solo em sistemas agroflorestais [manuscrito] / Isaac Correia Marinheiro. – 2013.  
29 f. : il. color.

Trabalho de Conclusão de Curso (Graduação em Estatística)  
– Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2013.

“Orientação: Prof. Dr. Ricardo Alves de Olinda, Departamento de Estatística”.

1. Dependência espacial. 2. Geoestatística. 3. Predições. I. Título.

21. ed. CDD 310

Isaac Correia Marinheiro

## Inferência espacial para a determinação da distribuição de nutrientes no solo em sistemas agroflorestais.

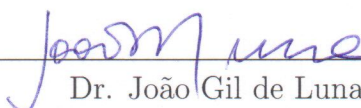
Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de Bacharel em Estatística.

Aprovado em: 16/12/2013

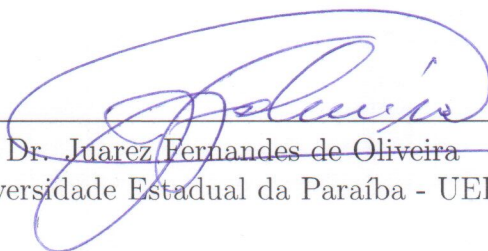
### Banca Examinadora:



Dr. Ricardo Alves de Olinda  
Orientador



Dr. João Gil de Luna  
Universidade Estadual da Paraíba - UEPB



Dr. Juarez Fernandes de Oliveira  
Universidade Estadual da Paraíba - UEPB

# Dedicatória

Dedico este trabalho aos meus pais, Antônio Marinheiro e Analíce Correia Marinheiro e a minha esposa Núbia Nogueira Barbosa Marinheiro e ao meu filho Iarley Miguel que estar para nascer. Por me incentivarem a superar todas as dificuldades.

# Agradecimentos

A Deus, por ter me dado força e discernimento para superar os momentos difíceis.

Ao professor Dr. Ricardo Alves de Olinda, pela confiança e orientação e também pela amizade durante o período em que estivemos juntos na graduação.

Ao meu irmãos Tony, Richardson e Chistiana, por terem dado apoio e incentivo a conclusão do curso.

A minha mãe Analíce e meu pai Antônio, por sempre ter me apoiado nas minhas escolhas e decisões.

# Resumo

O presente trabalho utiliza técnicas geoestatísticas com o objetivo de identificar a dependência espacial de nutrientes do solo em sistema agroflorestais. A estrutura espacial foi explorada pelo uso do semivariograma e as interpolações por meio de krigagens. Observou-se a existência da dependência espacial através do gráfico de envelope simulado, o índice de dependência espacial foi verificado utilizando-se a relação entre a variação estruturada e o patamar. Entre as diferentes funções de correlação estudadas, o modelo cúbico foi o modelo que melhor se ajustou aos dados, tomando-se como critério de seleção o método de Akaike(AIC). As previsões espaciais foram calculadas em uma grid de 81 localizações espaciais, em que, as precisões das krigagens ficaram em um patamar aceitável, tendo-se boas informações sobre a região em estudo.

**Palavras-chave:** Dependência espacial, geoestatística, previsões.

# Abstract

This study uses geostatistical techniques in order to identify the spatial dependence of soil nutrients in agroforestry system. The spatial structure was explored by using the semivariogram and interpolations by kriging. Observed the existence of spatial dependence across the graph simulated envelope, the index of spatial dependence was checked using the relation between the variation and structured level. Among the different correlation functions studied, circular model was the model that best fit the data, taking as a criterion for selecting the method of Akaike (AIC). the spatial predictions were calculated in a grid of 81 locations space, wherein the precision of kriging were on a porch acceptable, having good information about the area under study.

**Keywords:** Spatial dependence, geostatistics, predicts



# Sumário

**Lista de Figuras**

**Lista de Tabelas**

<b>1</b>	<b>Introdução</b>	p. 11
<b>2</b>	<b>Fundamentação Teórica</b>	p. 12
2.1	Estacionaridade . . . . .	p. 12
2.1.1	Hipótese de estacionaridade de primeira ordem . . . . .	p. 13
2.1.2	Hipótese de Estacionariedade de Segunda Ordem . . . . .	p. 13
2.2	Isotropia e anisotropia . . . . .	p. 14
2.3	O semivariograma . . . . .	p. 14
2.3.1	Funções de correlação para o semivariograma . . . . .	p. 15
2.3.2	Critério de avaliação de modelos . . . . .	p. 16
2.4	Grau de dependência espacial . . . . .	p. 17
2.4.1	Krigagem . . . . .	p. 17
2.4.2	Metodologia . . . . .	p. 18
<b>3</b>	<b>Resultados e Discussão</b>	p. 19
<b>4</b>	<b>Conclusão</b>	p. 27
	<b>Referências</b>	p. 28

# Lista de Figuras

1	Análise descritiva espacial do teor de nitrogênio no solo sob a árvore juazeiro. . . . .	p. 19
2	Gráfico de possíveis transformações da variável por meio da transformação ótima de Box-Cox(Esquerda), observando-se a dependência espacial por meio do envelope simulado(Direita). . . . .	p. 20
3	Modelo ajustado ao semivariograma conforme o método de máxima verossimilhança para a variável teor de nitrogênio sob a árvore juazeiro. . . . .	p. 21
4	Krigagem e probabilidades condicionais dos valores do teor de nitrogênio no solo. . . . .	p. 22
5	Análise descritiva espacial do teor de fósforo no solo sob a árvore juazeiro. . . . .	p. 23
6	Análise descritiva espacial do teor de fósforo no solo sob a árvore juazeiro com os dados transformados. . . . .	p. 24
7	Gráfico de possíveis transformações de variáveis por meio da transformação ótima de Box-Cox(Esquerda), observando a dependência espacial por meio do envelope simulado(Direita). . . . .	p. 24
8	Modelo ajustado ao semivariograma conforme o método de máxima verossimilhança para a variável teor de fósforo sob a árvore juazeiro. . . . .	p. 25
9	Krigagem e probabilidades condicionais dos valores do teor de fósforo no solo. . . . .	p. 26

# Lista de Tabelas

- 1 Estimativas dos parâmetros associados ao modelo por meio da máxima verossimilhança, assumindo-se a média da variável teor de nitrogênio, constante sobre a região de estudo sob árvore juazeiro. . . . . p. 21
- 2 Estimativas dos parâmetros associados ao modelo por meio da máxima verossimilhança, assumindo-se a média da variável teor de fósforo, constante sobre a região de estudo da árvore juazeiro. . . . . p. 25

# 1 Introdução

O estudo da variabilidade espacial de nutrientes do solo em sistema agroflorestais tem grande importância não só na escolha de uma área experimental, como também na locação das unidades experimentais, coletas de amostras e interpretação de resultados, levantamento e classificação de solo e nos esquemas de uso racional de fertilizantes. A variação das características do solo está relacionada com fatores de sua formação (JENNY, 1941).

Diversos estudos de características química e físicas do solo e de plantas demonstram que a variabilidade não ocorre ao acaso, mas apresenta correlação ou dependência espacial (OLIVEIRA et al., 1999; SOUZA et al., 1998). Segundo Balastreire (1998), a agricultura de precisão é definida como um conjunto de técnicas que permite o gerenciamento localizado de culturas e que se fundamenta na percepção da variabilidade espacial da produtividade e de fatos a ela relacionados. A erosão por sua vez é um dos fenômenos bastante destrutivos, pois proporciona a distribuição da matéria orgânica do solo em pastagens (MELLO et al., 2006), alterando-se a variabilidade espacial. Sendo assim, os métodos geoestatísticos estão sendo utilizados para analisar tanto a dependência espacial como para interpolar atributos de solo por meio da Krigagem (SOUZA et al., 2004).

O semivariograma é a principal ferramenta geoestatística para diagnosticar a presença da correlação entre as unidades amostrais. A krigagem também fornece estimadores exatos com propriedades de não tendenciosidade e eficiência (SOUZA et al., 2004). Os métodos de krigagem usam a dependência espacial entre amostras vizinhas, expressas no semivariograma, para estimar valores em qualquer posição dentro da região de estudo, sem tendência e com variância mínima.

Diante do exposto o presente trabalho tem por objetivo modelar a dependência espacial do teor de nitrogênio e fósforo no solo em sistemas agroflorestais, com ênfase na análise do semivariograma como ferramenta de determinação da dependência espacial e realizar interpolações utilizando-se a metodologia da krigagem.

## 2 Fundamentação Teórica

A geoestatística é um conjunto de métodos utilizados para a compreensão e modelagem da variabilidade espacial inerente em um processo de interesse. Embora ela tenha sua origem na mineração, a geoestatística é uma parte básica de muitas disciplinas científicas incluindo-se as ciências do solo, hidrologia e engenharia ambiental (GOTWAY, C. A., 1996). A parte central da geoestatística é a ideia de que medidas mais próximas tendem a serem mais parecidas do que valores observados em locais distantes.

Segundo Andriotti (2003), entende-se como variáveis regionalizadas uma função que varia de um lugar para outro no espaço com certa aparência de continuidade. Para Guerra (1988), estas variáveis apresentam-se por uma certa quantidade de dados numéricos brutos disponíveis, onde são obtidas informações sobre as características do fenômeno natural em estudo.

Estas características são: i) Localização: mesmo que a variável regionalizada seja contínua no espaço, geralmente não se conhece seu valor em todos os pontos, mas apenas nos que foram obtidos a partir da amostragem; ii) Anisotropia ou Zonalidade: ocorre quando a variabilidade espacial é mais intensa numa direção do que em outra, tal característica é mais frequente em fenômenos naturais (DRUCK et al., 2004); iii) Continuidade: elementos próximos espacialmente tendem a ter valores mais semelhantes do que os elementos que estão distantes entre si.

### 2.1 Estacionaridade

Para estudos de geoestatística necessita-se, como restrição máxima, que o primeiro e o segundo momento em relação à origem sejam constantes, ou seja, exige-se no máximo a estacionaridade de segunda ordem. Se a esperança matemática de uma variável aleatória é constante, independentemente da origem que se toma no espaço ou no tempo, pode-se dizer que a variável é estacionária de primeira ordem e, portanto, a média será a mesma para todo o processo.

### 2.1.1 Hipótese de estacionaridade de primeira ordem

Segundo Genú (2004), a hipótese de estacionaridade de primeira ordem é definida como sendo o momento de primeira ordem da distribuição da função aleatória  $Z(s_i)$ , que por sua vez, é constante em toda a região de estudos, ou seja,

$$E[Z(s_i)] = E[Z(s_i) + \mathbf{h}] = \mu(s_i), \quad (2.1)$$

em que  $s_i$  é a média dos valores amostrais.

De acordo com a Equação 2.1, se for tomado um vetor  $\mathbf{h}$  de distâncias euclidianas entre dois pontos, o qual apresenta módulo e direção, para qualquer  $\mathbf{h}$  tem-se que

$$E[Z(s_i) - Z(s_i + \mathbf{h})] = 0. \quad (2.2)$$

Considerando-se que a diferença entre as duas variáveis aleatórias  $Z(s_i) - Z(s_i + \mathbf{h})$  é uma variável aleatória, isto corresponde a afirmar que o primeiro momento desta variável aleatória é igual a zero.

### 2.1.2 Hipótese de Estacionariedade de Segunda Ordem

De acordo com Guimarães (2004), um processo é estacionário de segunda ordem se o segundo momento em relação à origem é constante, tem-se então que a variância é constante independente da origem no espaço ou no tempo, isto é,

$$E[Z^2(s_i)] = \mu(s_i);$$

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

Define-se a covariância como sendo, a esperança do produto que ocorre em  $s_i$  e  $s_j$ , com  $\mathbf{h} = \|s_j - s_i\|$ , definida com

$$C(s_i, s_j) = E[Z(s_i)Z(s_j)] - \mu^2(s_i). \quad (2.3)$$

A estacionaridade de segunda ordem é definida quando, além de atender à estacionaridade de primeira ordem, a função aleatória apresenta a característica de, para cada par de valores  $Z(s_i)$ ,  $Z(s_i + \mathbf{h})$ , a covariância existe e depende apenas da distância euclidiana  $\mathbf{h}$ .

## 2.2 Isotropia e anisotropia

Segundo Vieira (1995), quando o semivariograma é idêntico para qualquer direção de  $\mathbf{h}$ , ele é chamado de isotrópico e quando o semivariograma apresenta os parâmetros  $\hat{\sigma}^2$ ,  $\hat{\tau}^2$ , diferenciado dependendo da direção de  $\mathbf{h}$ , ele é chamado anisotrópico (pode-se classificar a anisotropia em anisotropia geométrica ou anisotropia zonal). Se o semivariograma é anisotrópico ele deve sofrer transformações antes de ser usado. Vieira (1995), alega que, em geral, a precisão da interpolação ou o tipo de hipótese satisfeita, não são afetados se, ao invés de se preocupar com a escolha de método de transformação de anisotropia, apenas limitar a faixa de distância na qual se utiliza o semivariograma.

## 2.3 O semivariograma

De acordo com Diggle e Ribeiro Jr. (2000), o semivariograma é uma função utilizada para determinar os “pesos” de cada amostra coletada em uma determinada região de estudos. De forma intuitiva, pode-se imaginar pontos que se encontrem mais próximos dos pontos a serem estimados e que possuam peso maior do que amostras mais espaçadas do ponto que se deseja estimar, o semivariograma pode ser definido como

$$\gamma(\mathbf{h}) = \frac{1}{2} \text{Var}[Z(s_i) - Z(s_i + \mathbf{h})]. \quad (2.4)$$

Observa-se que a função de semivariância  $\gamma(\mathbf{h})$  é definida como sendo a esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados por um vetor de distância euclidiana ( $\mathbf{h}$ ), conforme a seguinte equação

$$\gamma(\mathbf{h}) = \frac{1}{2} E[Z(s_i) - Z(s_i + \mathbf{h})]^2 \quad (2.5)$$

e pode ser estimado por

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2} \frac{\sum_{i=1}^n [Z(s_i) - Z(s_i + \mathbf{h})]^2}{n(\mathbf{h})}, \quad (2.6)$$

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

### 2.3.1 Funções de correlação para o semivariograma

Adicionalmente, é útil incluir no modelo, alguma flexibilidade na forma geral da função de correlação. A seguir pode-se observar algumas das principais funções de correlação do semivariograma utilizadas na geoestatística.

#### Função de correlação gaussiana

$$\rho(\mathbf{h}) = \tau^2 + \sigma^2 \left[ 1 - \exp \left[ -3 \left( \frac{\mathbf{h}}{\phi} \right)^2 \right] \right] \quad 0 \leq \mathbf{h} \leq \pi.$$

#### Função de correlação exponencial

$$\rho(\mathbf{h}) = \tau^2 + \sigma^2 \left[ 1 - \exp \left[ -3 \left( \frac{\mathbf{h}}{\phi} \right) \right] \right] \quad 0 \leq \mathbf{h} \leq \pi.$$

A função de correlação exponencial frequentemente produz um ajuste qualitativamente razoável à estrutura de correlação dos dados, mas as previsões baseadas nesse modelo tendem a ser pouco robustas a pequenas variações sobre o modelo escolhido (DIGGLE e RIBEIRO Jr., 2000). Tanto na função exponencial quanto no modelo de gaussiano,  $\pi$  é a distância máxima na qual o semivariograma é definido e nestes modelos o patamar  $\sigma^2 + \tau^2$  é atingido apenas assintoticamente. O parâmetro  $\phi$  é determinado visualmente como a distância após a qual o semivariograma se estabiliza.

#### Função de correlação Esférica

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

A função de correlação esférica depende apenas do parâmetro  $\phi$ , não há flexibilidade para a forma. Adicionalmente, a lógica de sua aplicação é questionável, em problemas no espaço bidimensional (DIGGLE e RIBEIRO Jr., 2000).

#### Função de correlação Circular

$$\rho(\mathbf{h}) = \begin{cases} \frac{2}{\pi} (\text{sen}^{-1} (\mathbf{h} + \sqrt{1 - \mathbf{h}^2})) \\ \rho(\mathbf{h}) = 1 \end{cases}, \quad \mathbf{h} > 1$$



### Função de correlação Matérn

Segundo Kruger (2005), a função de correlação Matérn, proposta por Bertil Matérn em 1986 é provavelmente a melhor escolha para função de correlação de uso geral por sua flexibilidade e simplicidade (apenas dois parâmetros) e possui a seguinte forma

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

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 ( $\phi$ ) está relacionado ao alcance das correlações, sendo que maiores valores indicam dependências espaciais de maior alcance. O segundo parâmetro ( $\nu$ ) está relacionado à suavidade do processo, de forma que, quanto maior  $\nu$ , maior será a suavidade (DIGGLE e RIBEIRO Jr., 2007).

### Função de correlação cúbico

Conforme Olinda(2012), Outra função de correlação limitada e de alcance finito é o modelo Cúbico. Sua equação é definida por

$$\rho(\mathbf{h}; \phi) = \begin{cases} 7 \left(\frac{\mathbf{h}}{\phi}\right)^2 - 8,75 \left(\frac{\mathbf{h}}{\phi}\right)^3 + 3,5 \left(\frac{\mathbf{h}}{\phi}\right)^5 - 0,75 \left(\frac{\mathbf{h}}{\phi}\right)^7 & \text{para } \mathbf{h} \leq \phi \\ 0; & \text{para } \mathbf{h} > \phi. \end{cases}$$

O parâmetro  $\phi$  possui um intervalo finito, aproximando-se do seu limite mais lentamente do que no modelo *Spherical*(OLINDA, 2012).

## 2.3.2 Critério de avaliação de modelos

Nos modelos com estrutura de dependência espacial utilizam-se os critérios de informação onde a avaliação do ajuste é penalizada por uma função do número de parâmetros. Cabe observar que é necessário ainda levar em conta o número de parâmetros independentes ao se incluir funções espaciais nos modelos (CÂMARA et al., 2003). Existem vários métodos para verificar a qualidade do ajuste do semivariograma aos dados experimentais, entre eles pode-se citar: Método do Critério de Informação de Akaike (AIC), método do coeficiente de determinação ( $R^2$ ), Método do Critério de informação bayesiana (BIC). Neste trabalho adota-se o critério de informação de Akaike.

O critério de informação de Akaike é expresso por

$$AIC = -2 \times (L(\boldsymbol{\theta}; y)) + 2p,$$

em que,  $L(\boldsymbol{\theta}; y)$  é o logaritmo da função de máxima verossimilhança e  $p$  é o número de parâmetros. Segundo este critério, o melhor modelo é o que possui menor valor de *AIC*.

## 2.4 Grau de dependência espacial

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)

$$IDE = \frac{\tau^2}{\sigma^2 + \tau^2} \times 100. \quad (2.7)$$

Quanto ao grau de dependência espacial da variável em estudo, pode ser classificada como:

- i) **Variável com forte dependência espacial.** Se o efeito pepita for menor ou igual a 25% do patamar  $\left(\frac{\tau^2}{\sigma^2 + \tau^2} < 0,25\right)$ ;
- ii) **Variável com moderada dependência espacial.** Se o efeito pepita representar entre 25% e 75% do patamar  $\left(0,25 \leq \frac{\tau^2}{\sigma^2 + \tau^2} \leq 0,75\right)$ ;
- iii) **Variável com baixa dependência espacial.** Se a relação entre efeito pepita e patamar estiver entre 75% e 100%  $\left(0,75 \leq \frac{\tau^2}{\sigma^2 + \tau^2} \leq 1,00\right)$ ;
- iv) **Variável independente espacialmente.** Se a relação entre efeito pepita e patamar for igual a 100%, neste caso tem-se o semivariograma com efeito pepita puro  $\left(\frac{\tau^2}{\sigma^2 + \tau^2} = 1,00\right)$

### 2.4.1 Krigagem

O termo *krigagem* foi aplicado por Matheron em 1965, em homenagem ao Engenheiro de Minas Sul-Africano Daniel G. Krige, que primeiro formulou e implementou essa forma de interpolação, em 1951. Isaaks e Srivastava (1989), afirmam que, assim como os métodos tradicionais de interpolação de pontos (como a ponderação da distância inversa, a triangulação e a média das amostras locais), a Krigagem pode fornecer a estimativa para um local específico. Frequentemente, os métodos tradicionais podem ser tão acurados mas consomem muito mais tempo do que a Krigagem.

Para a aplicação da krigagem assume-se: que sejam conhecidas as realizações  $z(s_1)$ ,  $z(s_2)$ , ...,  $z(s_n)$  da variável  $Z(s)$ , nos locais  $s_1, s_2, \dots, s_n$ ; que o semivariograma da variável já tenha sido determinado; e que o interesse seja estimar um valor  $\hat{s}$  na posição  $s_0$ . A

Krigagem ordinária, que é a variação mais utilizada da Krigagem simples, como o valor interpolado  $\hat{Z}(s_0)$  de uma variável regionalizada  $Z$ , num local  $s_0$  pode ser determinada por

$$\hat{Z}(s_0) = \sum_{i=1}^n w_i Z(s_i),$$

em que:  $n$  é o número de amostras de  $Z(s_i)$  envolvidas na estimativa de  $\hat{Z}(s_0)$ , e  $w_i$  são os pesos associados a cada valor medido em  $Z(s_i)$ .

## 2.4.2 Metodologia

O estudo foi realizado em 1996, em uma fazenda de gado de corte pertencente à Agropecuária Jaçanã, situada na Cidade de Custódia, PE (8013.70 ao Sul e 37044,70 a Oeste). A precipitação média de chuva no local é 740 milímetros e a temperatura média anual é de 260 C (dados a partir dos registros da fazenda). A vegetação dentro da área da Agropecuária, inicialmente consistia em uma vegetação, de caatinga semidecídua seca, a qual foi desmatada por tratores em 1984, e em seguida foram retirados os entulhos como tocos e pedras que porventura ficaram no local dentre os 5.000 ha (MENEZES; SALCEDO, 1999).

A coleta de dados foi realizada por pesquisadores e estudantes do Departamento de Energia Nuclear da UFPE. As amostras de solo foram coletadas na profundidade de 0-15 cm, com as devidas coordenadas X e Y.

O trabalho tem por objetivo modelar a dependência espacial do teor de nitrogênio e fósforo no solo em sistemas agroflorestais, com ênfase na análise do semivariograma como ferramenta de determinação da dependência espacial e realizar interpolações utilizando-se a metodologia da krigagem para o teor de nitrogênio e fósforo no solo.

### 3 Resultados e Discussão

Segundo Correia (2012), na região de estudos, foi selecionada a árvore Juazeiro. Deu-se preferência por árvores maduras isoladas uma das outras. Para cada árvore, estabeleceram-se 81 pontos de amostragem regularmente distribuídas, onde a espécie arbórea sempre encontra-se no centro da região de estudos. As análises foram realizadas com o auxílio do pacote *geoR* do Software R, versão 3.0.2.

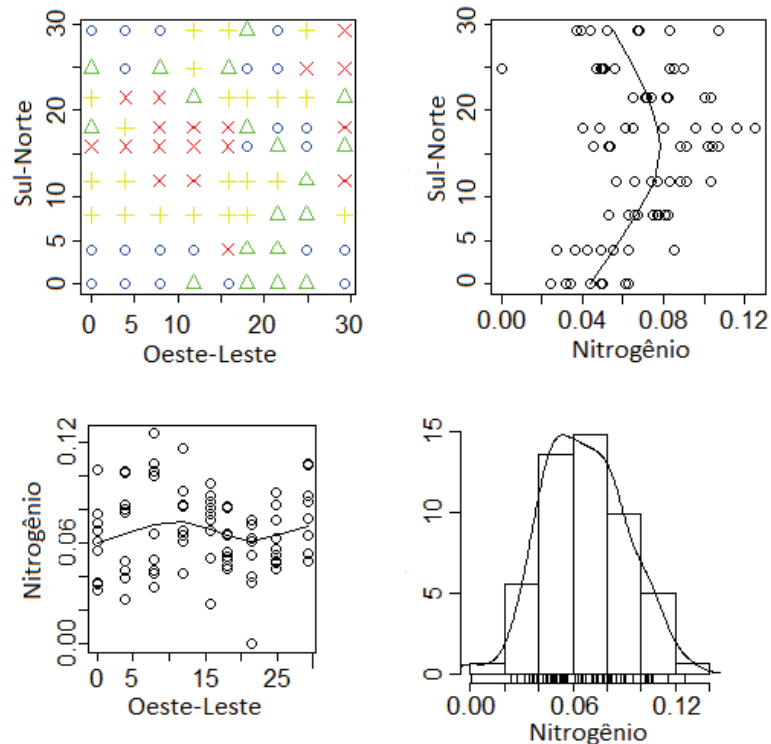


Figura 1: Análise descritiva espacial do teor de nitrogênio no solo sob a árvore juazeiro.

Inicialmente analisa-se a variável do teor de nitrogênio no solo. Diante do exposto, pode-se observar por meio da Figura 1, o gráfico do canto superior esquerdo os quartis da variável nitrogênio, isto é, “+” 1º quartil, “Δ” 2º quartil, “o” 3º quartil e “x” 4º quartil. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão da variável em torno da média geral. Com base no gráfico do canto inferior direito, pode-se observar que o pressuposto de normalidade dos resíduos sob a variável em estudo não está sendo violado, fato este corroborado pela imagem do lado esquerdo

da Figura 2.

A imagem ao lado direito da Figura 2 apresenta alguns pontos fora do intervalo, indicando assim uma dependência espacial do teor de nitrogênio no solo, vale salientar que foram utilizadas 1.000 simulações. Para que haja dependência espacial, deve haver ao menos um ponto do variograma fora do envelope simulado, deste modo, se isso ocorre rejeita-se a hipótese nula ( $H_0$ ), de que não há dependência espacial (DIGGLE e RIBEIRO, 2007).

A dependência espacial observada para os valores de nitrogênio mostram que as análises da estatísticas clássica, em que as observações são consideradas independentes especialmente falando, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

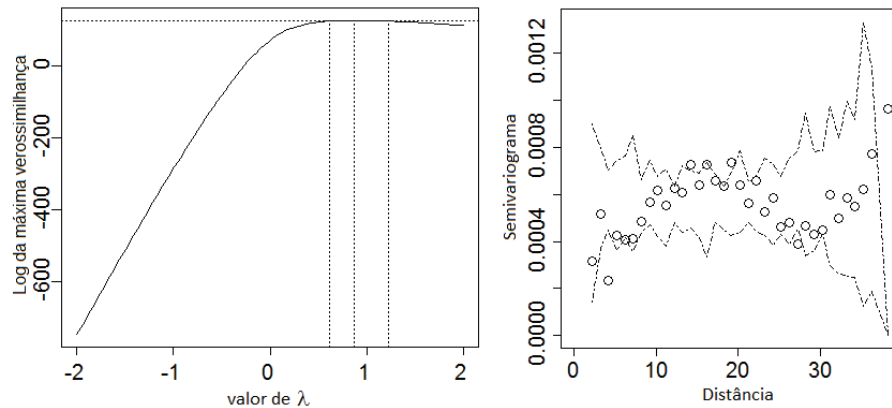


Figura 2: Gráfico de possíveis transformações da variável por meio da transformação ótima de Box-Cox (Esquerda), observando-se a dependência espacial por meio do envelope simulado (Direita).

Continuando-se as análises do teor de nitrogênio da árvore juazeiro, através da estatística espacial, observa-se por meio da Tabela 1 as estimativas associadas aos parâmetros dos modelos ajustados ao semivariograma por meio da função de máxima verossimilhança, considerando-se a média como polinômio de primeira ordem sobre as coordenadas  $(x, y)$  para a matriz de delineamentos  $\mathbf{X}$ .

Levando-se em consideração o polinômio de primeira ordem apresentado na Tabela 1, observa-se que as estimativas para o parâmetro  $\beta_1$  apresenta valores próximos para as diferentes funções de correlação, o parâmetro  $\sigma^2$  apresenta valores bem parecidos, no entanto, o parâmetro  $\phi$  apresenta valores discrepantes para as demais funções de correlação em estudo, ainda na Tabela 2, pode-se observar que o efeito pepita  $\tau^2$  apresenta valores próximos para as demais funções de correlação em estudo. Adotando-se o método de seleção AIC o modelo selecionado foi o cúbico com um valor de  $AIC = -401,4$ , com

alcance de aproximadamente  $24m$ , ou seja, amostras do teor de nitrogênio no solo selecionadas a distâncias inferiores são especialmente dependentes.

Tabela 1: Estimativas dos parâmetros associados ao modelo por meio da máxima verossimilhança, assumindo-se a média da variável teor de nitrogênio, constante sobre a região de estudo sob árvore juazeiro.

função de correlação	$\hat{\beta}_1$	$\hat{\tau}^2$	$\hat{\sigma}^2$	$\hat{\phi}$	AIC	IDE
matérn kappa = 0,5	0,0621	0,0001	0,0005	7,9543	-396,8	16,66%
matérn kappa = 1,0	0,0623	0,0002	0,0004	5,8638	-398,0	33,33%
esférico	0,0639	0,0002	0,0004	18,3229	-400,7	33,33%
gaussiano	0,0635	0,0003	0,0004	10,3288	-400,9	42,85%
circular	0,0639	0,0002	0,0004	17,3094	-401,2	33,33%
cúbico	0,0638	0,0002	0,0004	23,6324	-401,4	33,33%

Na sequência, pose-se observar por meio da Figura 3 o semivariograma ajustado conforme o método de máxima verossimilhança. Pode-se observar que todas as características estudadas nos procedimentos aqui descritos, apresentam estruturas espaciais bem definidas, isto é, existe uma função estrutural com semivariância de comportamento modelável.

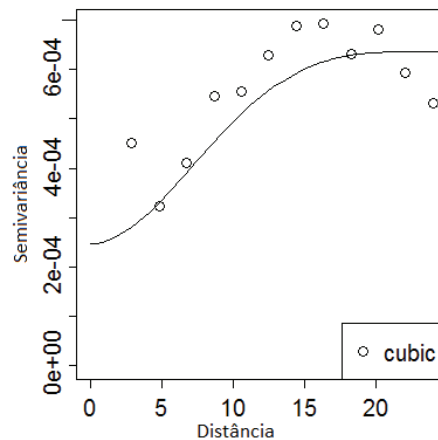


Figura 3: Modelo ajustado ao semivariograma conforme o método de máxima verossimilhança para a variável teor de nitrogênio sob a árvore juazeiro.

Para verificar o *IDE*, entre as características da árvore Juazeiro do teor de nitrogênio foi utilizada a relação entre a variação estruturada e o patamar, observou-se que o  $IDE = 33,33\%$ , classificando-se segundo Cambardella et al. (1994), como uma moderada dependência espacial.

Conforme visto anteriormente, a krigagem é uma metodologia utilizada para estimar o valor de uma variável de interesse em um local onde não foram coletadas observações. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para o índice

de fertilidade do solo em uma malha de 81 localizações espaciais, sendo que as estimativas paramétricas foram substituídas na forma de krigagem ordinária.

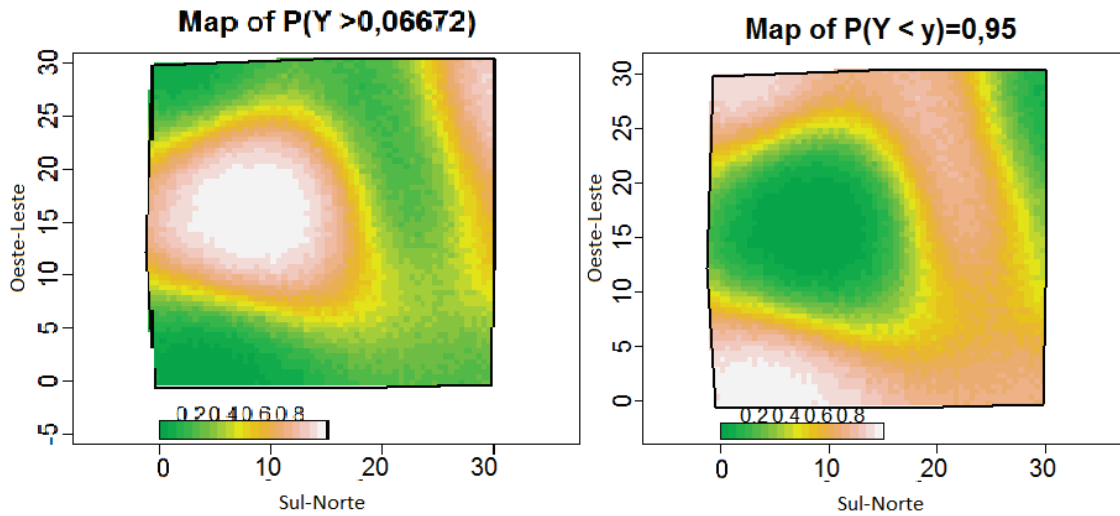


Figura 4: Krigagem e probabilidades condicionais dos valores do teor de nitrogênio no solo.

Analisando-se a Figura 4 percebe-se que as amostras ficaram condizentes com os dados observados, ou seja, localizações com  $P(Y > 0,06672)$  possuem entre 80-100% do teor de nitrogênio no solo (parte clara da imagem esquerda). Localização com  $P(Y < y)$  menor que a média possui uma concentração do teor de nitrogênio no solo de 40-70% (parte laranja da imagem esquerda). Localização com  $P(Y < y)$  menor que a média possui uma concentração do teor de nitrogênio no solo de 10-30% (parte verde da imagem esquerda).

Neste momento, dando sequência as análises, trabalha-se com a variável do teor de fósforo. Pode-se observar por meio da Figura 5 com os dados originais o gráfico do canto superior esquerdo os quartis da variável fósforo. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão da variável em torno da média geral. Com base no gráfico do canto inferior direito, observou-se a análise residual sob o pressuposto de gaussianidade da variável em estudo.

A transformação ótima de Box-Cox, por meio do logaritmo da função de máxima verossimilhança, apresenta qual o valor do parâmetro que mais aproxima os resíduos à distribuição normal. Tendo em vista que o valor 1 (um) não está contido no intervalo com 95% de confiança, é preciso utilizar uma transformação na variável em estudo. Neste caso, para a variável fósforo foi encontrado o valor de  $\lambda = -0,39$ . Após a transformação dos dados, percebe-se por meio da Figura 6 que o gráfico do canto inferior direito condiz com a hipótese de que os resíduos se aproximam à distribuição normal

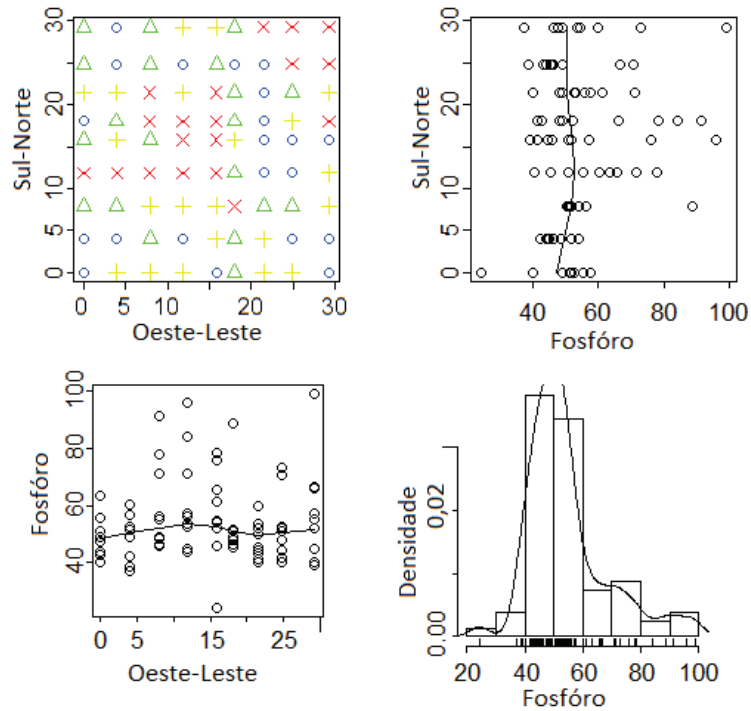


Figura 5: Análise descritiva espacial do teor de fósforo no solo sob a árvore juazeiro.

A existência de dependência espacial pode ser verificada por meio do diagnóstico gráfico utilizando-se do envelope simulado. A imagem ao lado direito da Figura 5 apresenta alguns pontos fora do intervalo, indicando assim uma dependência espacial do teor de fósforo no solo, vale salientar que foram utilizados 1.000 simulações. Para que haja dependência espacial, deve haver ao menos um ponto do variograma fora do envelope simulado, deste modo, se isso ocorre rejeita-se a hipótese nula ( $H_0$ ), de que não há dependência espacial (DIGGLE e RIBEIRO Jr., 2007).

A dependência espacial observadas para os valores do teor de fósforo mostram que as análises da estatísticas clássica, em que as observações são consideradas aleatórias e independentes especialmente, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

Continuando-se as análises através da estatística espacial, observa-se por meio da Tabela 2, as estimativas de máxima verossimilhança para os parâmetros dos modelos univariados ajustados, considerando-se a média como polinômio de primeira ordem sobre as coordenadas  $(x, y)$  para a matriz de delineamentos  $\mathbf{X}$ . Levando-se em consideração o polinômio de primeira ordem apresentado na Tabela 1, observa-se que as estimativas para o parâmetro  $\beta_1$  apresenta valores bem próximos para as diferentes funções de correlação,



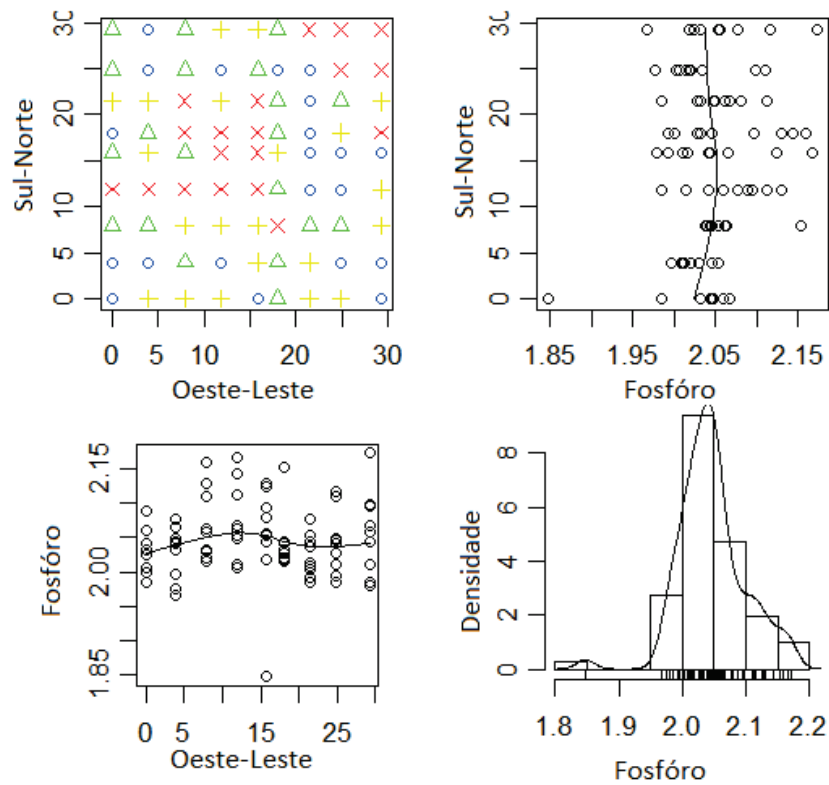


Figura 6: Análise descritiva espacial do teor de fósforo no solo sob a árvore juazeiro com os dados transformados.

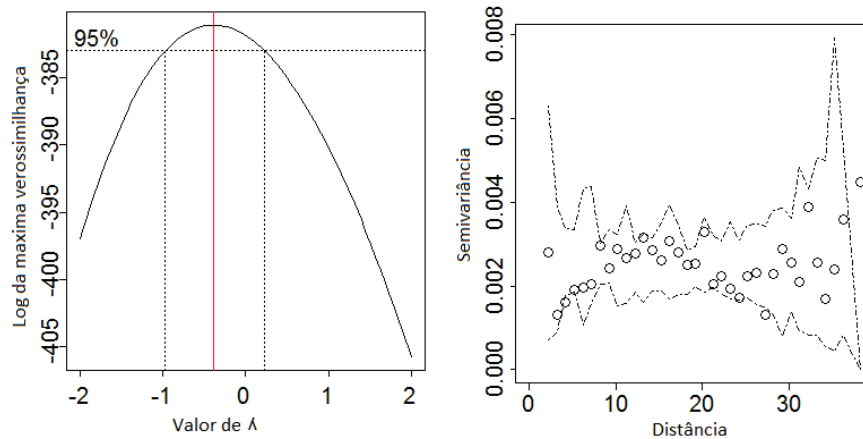


Figura 7: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de Box-Cox(Esquerda), observando a dependência espacial por meio do envelope simulado(Direita).

o parâmetro  $\hat{\sigma}^2$  apresenta valores bem parecidos, no entanto, o parâmetro  $\phi$  apresenta um valor bem discrepantes para as demais funções de correlação em estudo, ainda na Tabela 2, pode-se observar que o efeito pepita  $\hat{\tau}^2$  apresenta valores bem próximos para as demais funções de correlação em estudo. O próximo passo é a seleção do modelo que melhor se ajusta aos dados de fósforo. Adotando-se o método de seleção de AIC, o qual seleciona

como melhor modelo aquele que apresentar o menor valor para a estatística AIC, o modelo selecionado foi o cúbico com um valor de  $AIC = 626,7$ .

Tabela 2: Estimativas dos parâmetros associados ao modelo por meio da máxima verossimilhança, assumindo-se a média da variável teor de fósforo, constante sobre a região de estudo da árvore juazeiro.

função de correlação	$\hat{\beta}_i$	$\hat{\tau}^2$	$\hat{\sigma}^2$	$\hat{\phi}$	AIC	IDE
matérn kappa = 0,5	2,0435	0,0008	0,0020	5,2328	629,8	28,57%
matérn kappa = 1,0	2,0439	0,0012	0,0016	3,9100	628,9	42,85%
esférico	2,0452	0,0012	0,0016	14,2118	627,4	42,85%
gaussiano	2,0453	0,0014	0,0014	6,6532	640,0	50,00%
circular	2,0457	0,0010	0,0016	9,0547	626,8	38,46%
cubic	2,0455	0,0014	0,0014	15,8143	626,7	50,00%

Na Figura 8 encontram-se os semivariogramas ajustados conforme o método de máxima verossimilhança para uma distância máxima de 2.000 metros, pode-se, observar que todas as características estudadas nos procedimentos aqui descritos, apresentam estruturas espaciais bem definidas, isto é, existe uma função estrutural com semivariância de comportamento modelável. Para todas as características avaliadas, o modelo que melhor se ajustou ao semivariograma experimental foi o modelo cubic apresentando-se o menor valor de AIC. Isto significa que as estimativas com o modelo cubic, considerando-se a componente espacial, independente do procedimento de amostragem, trará otimização nas estimativas.

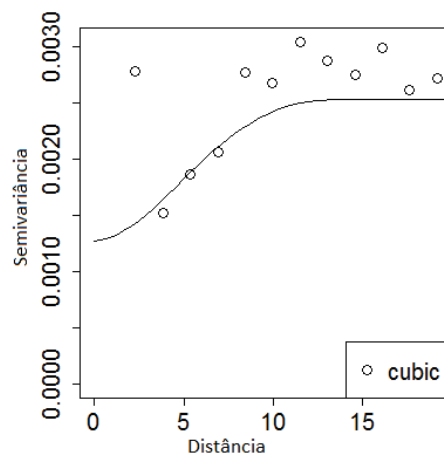


Figura 8: Modelo ajustado ao semivariograma conforme o método de máxima verossimilhança para a variável teor de fósforo sob a árvore juazeiro.

Para verificar o  $IDE$ , entre as características da árvore Juazeiro da variável sob o teor de fósforo foi utilizada a relação entre a variação estruturada e o patamar, observou-se que o  $IDE = 50,00\%$ , classificando-se como uma moderada dependência espacial. Nota-se

que a variável fósforo apresenta dependência espacial, que pode ser descrita pelo modelo cubic com alcance de aproximadamente 15,4183m, ou seja, amostras de fósforo do solo selecionadas a distâncias inferiores são especialmente dependentes.

Conforme visto anteriormente, a Krigagem é uma metodologia utilizada para estimar o valor de uma variável de interesse em um local onde não foi coletadas observações. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para o índice de fertilidade do solo em uma malha de 81 localizações espaciais, sendo que as estimativas paramétricas foram substituídas na forma de krigagem. Uma análise importante que se dispõe com a utilização dos mapas preditivos de krigagem são os de probabilidade condicionais marginais, pois eles podem prever a partir de um determinado valor numérico as chances de ocorrência de um evento na região de estudos.

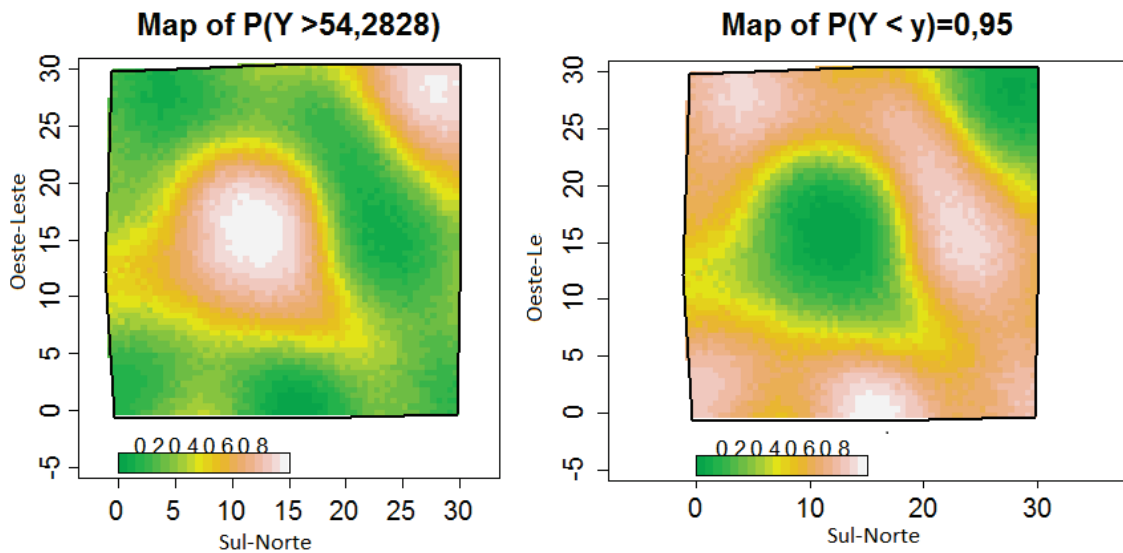


Figura 9: Krigagem e probabilidades condicionais dos valores do teor de fósforo no solo.

Analisando-se a Figura 9 percebe-se que as amostras ficaram condizentes com os dados observados, ou seja, localizações com  $P(Y > 54,2828)$  possuem entre 80-100% do teor de fósforo no solo (parte clara da imagem esquerda). Localização com  $P(Y < y)$  menor que a média possui uma concentração do teor de fósforo no solo de 40-70% (parte laranja da imagem esquerda). Localização com  $P(Y < y)$  menor que a média possui uma concentração do teor de fósforo no solo de 10-30% (parte verde da imagem esquerda).

## 4 Conclusão

Feita a análise exploratória dos dados, verificou-se que houve uma assimetria a direita com os dados originais. Realizou-se uma transformação ótima de Box-Cox com intuito de normalizar os dados, após a transformação os dados se tornaram simétricos conforme observa-se na Figura 8.

O modelo que melhor se ajustou aos dados foi o modelo cubic, o qual foi selecionado pelo método de seleção de *AIC*, método este que seleciona como melhor modelo, aquele que apresentar o menor valor para a estatística de *AIC*. Observou-se que houve uma forte dependência espacial, o que justifica-se também pelo baixo valor do efeito pepita.

A precisão da krigagem ficou em um patamar aceitável o que implica em boas informações sobre a região em estudo. Observa-se através do gráfico de predição espacial para a probabilidade condicional, pois ele prediz a partir um determinado valor numérico as chances de ocorrência de um evento na região de estudos.

Visando quantificar as relações solo-planta em sistema agroflorestal, vários estudos tem demonstrado que a diminuição da fertilidade do solo no semiárido pode ser revertida e ou prevenida através da introdução e/ou preservação de espécies arbóreas. esse mesmo estudo realizado em épocas diferentes com Menezes e Salcedo (1999) e Menezes et al., (2002) avaliaram a fertilidade do solo, e encontraram “ilhas de fertilidades” nas áreas sob a influência das copas de árvores presentes em pastagens ou seja observou-se que com a preservação do juazeiro (*ziziphus joazeiro*) nas pastagens aumentou os níveis de materia orgânica e nutrientes na camada superficial do solo.

# Referências

- ANDRIOTTI, J. L. S. **Fundamentos de Estatística e Geoestatística**. São Leopoldo: Unisinos, 2003. 165p.
- BALASTREIRE, L. A. **Agricultura de Precisão**, FEALQ: Piracicaba, 1998. 150p.
- BERTEL, M. **Spatial Variation**, Springer: Berlin, 1986. 365p.
- CÂMARA, G.; MONTEIRO, A. M.; FUCKS S. D.; CARVALHO M. S. **Introdução à ciência da geoinformação**. INPE: São José dos Campos, 2003. 186p.
- CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in Central Iowa soils. **Soil Science Society of America Journal**, v.58, p.1501-1511, 1994.
- CORRÊA J. B.; REIS R. A.; CARVALHO J. G.; GUIMARÃES, P. T. G. Avaliação da fertilidade do solo e do estado nutricional de cafeeiros do Sul de Minas Gerais. **Ciência e Agrotecnologia**, v.25 p.1279-1286, 2012.
- DIGGLE, P. J.; RIBEIRO Jr, P. J. **Model Based Geostatistics**, Caxambu: São Paulo, Associação Brasileira de Estatística. 2000. 140p.
- DIGGLE, P.J.; RIBEIRO Jr., P. J. **Model-Based geostatistics**. New York: Springer, 2007. 230p.
- DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. Análise Espacial de Dados Geográficos. Brasília: **EMBRAPA**, v.85, p1-19, 2004.
- MELLO, J. M.; OLIVEIRA, M. S.; BATISTA, J. L. F.; RIBEIRO JR, P. J.; KANEGAE JR, H. Uso do estimador geoestatístico para predição volumétrica por talhão. **Floresta**, Curitiba, v.36, p. 251-260, 2006.
- FOOK. K. D. **Integração da Estatística Espacial em Ambientes Gis**, 2005. 86p.
- GENÚ. A. M.; **Geoestatística Multivariada**, 2004. 280p.
- GOODCHILD, M. F.; HAINING, R. P. GIS and spatial data analysis: covering perspectives. **Papers in Regional Science**, v.83, p.363-385, 2004.
- GOTWAY, C. A., HARTFORD, A. H. Geostatistical methods for incorporating auxiliary information in the prediction of spatial variables. *J. Agri., Biol. Environ. Statistics*, v.1, p.17-39, 1996.
- GUERRA, P. A. G. **Geoestatística operacional**, Brasília: MME/DNPM, 1988. 145p.

GUIMARÃES, E. C. **Geoestatística Básica e Aplicada**, Universidade Federal de Uberlândia Faculdade de Matemática, Núcleo de Estudos Estatísticos e Biométricos: Minas Gerais, 2004. 92p.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An Introduction to Applied Geostatistics**. New York: Oxford University Press, 1989. 561p.

JENNY.H. **Courier Dover Publications Science**: New Work. 1941. 281p.

MENEZES, R.; SALCEDO, I.; ELLIOTT, E. Microclimate and nutrient dynamics in a silvopastoral system of semiarid northeastern brazil. **Agroforestry Systems**, v.30, p.27-38, 2002.

MENEZES, R. S. C.; SALCEDO, I. H. Inuence of tree species on the herbaceous understory and soil chemical characteristics in a silvopastoral system in semi-arid northeastern brazil. **Revista brasileira de Ciências do Solo**, p.817-826, 1999.

Olinda) OLINDA, Ricardo Alves de. **Modelagem estatística de extremos espaciais com base em processos max-stable aplicados a dados meteorológicos no estado do Paraná**. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2012. Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11134/tde-17092012-103936/j>. Acesso em: 2013-12-17. 2012. 164p.

OLIVEIRA, S. **Tratado de metodologia científica**. São Paulo: Pioneira. 1999. 156p.

KRIGE, D. G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Chemistry, Metal. and Mining Society of South Africa**, v.52, p.119-139, 1951.

KRUGER, M. C. **Análise geoestatística de dados meteorológicos do estado do paraná utilizando um software livre**, 2005. 96p.

SOUZA, L. C. **Variabilidade espacial da salinidade de um solo aluvial no semi-árido paraibano**. Campina Grande, 1999. 77p. Dissertação (Mestrado) - Universidade Federal da Paraíba.

VIEIRA, S. R.; NOVAIS, R. F.; ALVARES, V. H.; SCHAEFFER, C. E. G. R. Geoestatística em estudos de variabilidade espacial do solo. Tópicos em ciência do solo. Viçosa: **Sociedade Brasileira de Ciência do Solo**, p.1-54, 2000.