

O USO DE TÉCNICAS GEOESTATÍSTICAS NA ESPACIALIZAÇÃO DA ARGILA DO SOLO

Elisângela Aparecida de Oliveira¹

Resumo: *O presente trabalho teve como objetivo utilizar técnicas geoestatísticas para proceder à espacialização da média da argila do solo medidas às profundidades de 0-5, 5-20 e 80-100 cm. A área do estudo foi na Estação Ecológica dos Caetetus (EEC), que possui área contínua de 2178,84 ha, situada nos municípios de Gália e Alvilândia, estado de São Paulo, entre as coordenadas geográficas (22° 41' e 22° 46' S e 49° 10' e 49° 16' W), na bacia hidrográfica do Médio Paranapanema. As etapas do trabalho envolveram a coleta de amostras de solo, análises químicas, análises exploratórias dos dados, criação, ajuste do modelo experimental, krigagem ordinária e criação do mapa temático do atributo em estudo. Na etapa inicial, foram coletadas amostras de solo num total de 256 amostras. As análises exploratórias dos dados procederam-se, inicialmente, como estatísticas descritivas e espaciais. Como resultado da krigagem espacial dos dados observou-se que a área apresentou teores de argila baixos, médios e altos, mostrando a existência da variação da argila no solo, no lote em estudo.*

1. INTRODUÇÃO

O termo “estatística espacial” é utilizado para descrever uma ampla gama de modelos e métodos estatísticos destinados à análise de dados espacialmente referenciados. Dentro da estatística espacial, o termo geoestatística se refere a modelos e métodos para os dados com as seguintes características: primeiro, os valores de Y_i : $i = 1, \dots, n$, são observados de um conjunto discreto de locais de amostragem x_i dentro de alguma região espacial A . Segundo, cada valor observado Y_i seja uma medida direta, ou seja estatisticamente relacionada, do valor de um fenômeno espacial subjacente contínuo, $S(x)$, no local de amostragem correspondente x_i (Diggle e Ribeiro Jr., 2006).

Segundo Isaaks e Srivastava, 1989, citado por Vasconcelo e Leite, 2007, a teoria fundamental da geoestatística é a esperança de que, na média, as amostras próximas no tempo e no espaço sejam mais similares entre si do que as que estiverem distantes.

As análises geoestatísticas, em geral, compreendem as seguintes etapas: análise exploratória dos dados, modelagem do semivariograma experimental, ajuste de um modelo teórico e interpolação por krigagem.

¹ Departamento de Ciências Exatas (LCE), Esalq/USP, elisangelao@rocketmail.com

Para a utilização das técnicas geoestatísticas, necessita-se, inicialmente, obter as informações da área em estudo. Os atributos a serem analisados devem ser conhecidos e localizados através de um processo de amostragem dentro de uma região geográfica definida. Esse procedimento é viabilizado pela implantação de um sistema de coordenadas locais ou geográficas, onde cada característica do solo estudada terá suas informações quantitativas ou qualitativas determinadas. É assim, por exemplo, com o número de decíduos, área basal (m²) dos decíduos, número de perenifólos, argila do solo, PH do solo entre outros, que são atributos em estudo sempre associados a uma posição no espaço. Estas técnicas têm bastantes aplicações em estudos florestais e agronômicos.

Com relação à dinâmica de florestas, poucas mudanças são esperadas em um curto período de tempo. No entanto, padrões climáticos atípicos estão sendo observados, gerando alterações nos regimes de precipitação e formação de tornados ou fortes ventos, possivelmente em conseqüências das mudanças climáticas globais. Estes eventos podem gerar mudanças imediatas na dinâmica de florestas (Barreto, 2009).

Os chamados solos argilosos contêm mais de 30% de argila. A argila é formada por grãos menores que os da areia. Além disso, esses grãos estão bem ligados entre si, retendo água e sais minerais em quantidade necessária para a fertilidade do solo e o crescimento das plantas. Mas se o solo tiver muita argila, pode ficar encharcado, cheio de poças após a chuva. A água em excesso nos poros do solo compromete a circulação de ar, e o desenvolvimento das plantas fica prejudicado. Quando está seco e compacto, sua porosidade diminui ainda mais, tornando-o duro e ainda menos arejado.

Este trabalho descreve uma metodologia envolvendo as etapas e resultados de análises geoestatísticas para determinação dos valores da argila do solo de uma Estação Ecológica.

2. MATERIAL E MÉTODOS

2.1 Localização e caracterização da área de estudo

Este estudo foi desenvolvido na Estação Ecológica dos Caetetus (EEC), que possui área contínua de 2178,84 ha, situada nos municípios de Gália e Alvilândia, estado de São Paulo, entre as coordenadas geográficas (22° 41'e 22° 46'S e 49° 10'e 49° 16'W), na bacia hidrográfica do Médio Paranapanema. O clima local, segundo a classificação de Köppen, é Cwa, mesotérmico de inverno seco, com temperaturas inferiores a 18°C no inverno e superiores a 22°C no verão. No mês mais seco o total da precipitação não passa de 30 mm,

atingindo a precipitação anual de 1.100 a 1.700 mm. A altitude varia entre 520 e 680m (Tabanez *et al.*2005, citado por Barreto, 2009).

A EEC caracteriza-se como um grande remanescente de floresta estacional semidecidual do Planalto Ocidental do estado de São Paulo. Esta formação florestal revestia originalmente parte do Planalto Paulista, a Depressão Periférica, a Cuesta Basáltica e parte do Planalto Ocidental do interior paulista, restando apenas 6% da sua constituição original. Hoje esta formação florestal é uma das mais ameaçada do estado de São Paulo, devido à sua fragmentação como consequência de alterações antrópicas (Tabanez *et al.*2005, citado por Barreto, 2009).

O acompanhamento temporal da dinâmica florestal em larga escala, realizado por meio de parcelas permanentes, tem se mostrado muito eficiente e promissor para o entendimento das possíveis causas da mortalidade, recrutamento e do crescimento das árvores. O entendimento do funcionamento das florestas tropicais, associando estes aos componentes do ecossistema ou às alterações causadas nele pela ação humana, é um passo importante a ser dado pelos estudiosos de ecologia de todo o mundo (Parker & Picket 1999, Lewis *et al.* 2004^a, , citado por Barreto, 2009).

Esses estudos também são considerados essenciais para o manejo e conservação de espécies, restauração das formações florestais e para estabelecer indicadores de avaliação e monitoramento dos remanescentes florestais e das áreas restauradas (Whitmore 1989, Condit 1995, Tomás 1996, Rodrigues *et al.* 2009, citado por Barreto, 2009).

2.2 Análise exploratória dos dados

A análise exploratória dos dados é parte integrante da prática da estatística moderna, e a geoestatística não é exceção. No cenário da geoestatística, a análise exploratória é naturalmente dirigida para a investigação preliminar dos aspectos espaciais dos dados que são relevantes para verificar se as suposições feitas por qualquer modelo provisório são aproximadamente satisfeitas. No entanto, aspectos não espaciais podem e devem também ser investigados (Diggle e Ribeiro Jr., 2006). A partir dessas análises é possível, por exemplo, proceder a transformações dos valores observados e a remoção de tendência, antes das análises geoestatísticas.

2.3 A função Semivariograma

O formato básico para dados geoestatísticos univariados é:

$$(x_i, y_i) : i = 1, \dots, n,$$

em que x_i identifica uma localização espacial (tipicamente no espaço bidimensional, embora exemplos unidimensionais e tridimensionais também possam ocorrer) e y_i é um valor escalar associado ao local x_i . Chamamos y a variável medida ou resposta. A característica definidora da geoestatística é que a variável de medição é, pelo menos em princípio, definida ao longo de uma região de estudo contínua A . Além disso, supõe-se que o plano de amostragem para os locais x_i ou é determinístico (por exemplo, os x_i podem formar uma grade sobre a região de estudo), ou estocasticamente independente do processo que gera as medições y_i . Cada y_i é uma realização de uma variável aleatória Y_i cuja distribuição é dependente do valor no local de um x_i subjacente espacialmente contínuo no processo estocástico $S(x)$ que não é diretamente observável (Diggle e Ribeiro Jr., 2006).

O semivariograma é uma função matemática, definida para representar o nível de dependência entre duas variáveis regionalizadas. Ele é usado para modelar valores correlacionados no espaço ou no tempo e é pré-requisito para os métodos de interpolação de Krigagem (Carvalho, 2002, citado por Vasconcelo, 2007). Muitas vezes é uma ferramenta de análise exploratória dos dados, indicando se existe dependência espacial nos dados estudados, e até mesmo se existe estacionariedade.

Um gráfico do semivariograma está ilustrado na **Figura 1**.

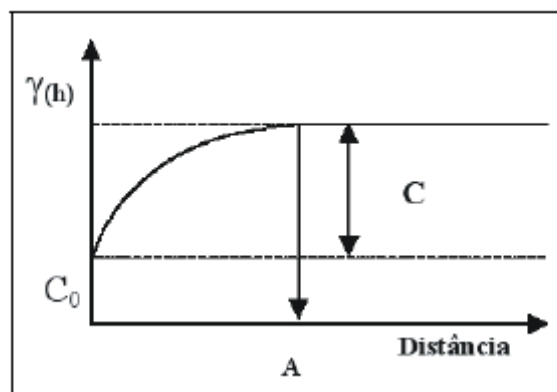


Figura 1 – Semivariograma típico

Na Figura anterior: $\gamma(h)$ - é a semivariância, C_0 - é o efeito pepita, C - é o patamar (parcial); $C + C_0$ - patamar total e A - Alcance.

2.4 Estimativa de Máxima Verossimilhança

Segundo Diggle e Ribeiro Jr. (2006), a estimativa de máxima verossimilhança é um método amplamente aceito na estatística, com propriedades de otimização bem conhecidas para amostras grandes. Sob condições de regularidade, o estimador de máxima

verossimilhança é assintoticamente normalmente distribuído, imparcial e totalmente eficiente. Dentro do contexto geoestatístico, a implementação da estimativa de máxima verossimilhança é simples quando os dados são gerados por um modelo Gaussiano. No entanto, e com a flexibilidade adicional fornecida por transformações marginais da variável resposta Y , este modelo é útil para muitas aplicações de geoestatística na qual Y é uma quantidade de valor contínuo.

2.5 Previsão Espacial - Krigagem

A previsão subjacente utilizada foi a krigagem simples. A krigagem simples é a menor média de erros quadrados de previsão sob o modelo Gaussiano estacionário, mas ignorando a incerteza de parâmetros, isto é, as estimativas de todos os parâmetros do modelo estão conectadas na equação de predição como se fossem os verdadeiros valores dos parâmetros.

O preditor do erro quadrado médio mínimo, digamos $\hat{S}(x)$, de $S(x)$ em um local arbitrário x é a função dos dados, $y = (y_1, \dots, y_n)$, que minimiza a quantidade $E[(\hat{S}(x) - S(x))^2]$. Um resultado normal, é que $\hat{S}(x) = E[S(x) | y]$. Para o processo Gaussiano estacionário, essa esperança condicional é uma função linear do y_i , ou seja,

$$\hat{S}(x) = \mu + \sum_{i=1}^n w_i(x)(y_i - \mu)$$

em que os $w_i(x)$ são funções explícitas dos parâmetros de covariância: variância (σ^2), efeito-pepita (τ^2) e alcance (ϕ).

Logo a krigagem é um processo de estimativa de valores de variáveis distribuídas no espaço, e/ou no tempo, a partir de valores adjacentes enquanto considerados como interdependentes por um modelo. Trata-se, em último caso, de um método de estimativa por médias móveis.

O termo, tradução do francês *krigeage*, e do inglês *kriging*, foi cunhado pela escola francesa de geoestatística em homenagem ao engenheiro de minas sul-africano e pioneiro na aplicação de técnicas estatísticas em avaliação mineira, Daniel G. Krige.

A krigagem pode ser usada, como algoritmo estimador, para:

a) previsão do valor pontual de uma variável regionalizada em um determinado local dentro do campo geométrico; é um procedimento de interpolação exato que leva em consideração todos os valores observados, o qual pode ser a base para cartografia automática por computador quando se dispõe de valores de uma variável regionalizada dispostos por uma determinada área;

b) cálculo médio de uma variável regionalizada para um volume maior que o suporte geométrico como, por exemplo, no cálculo do teor médio de uma jazida a partir de informações obtidas de testemunhas de sondagens.

Em todas essas situações o método fornece, além dos valores estimados, o “erro” associado a tal estimativa, o que o distingue dos demais algoritmos à disposição (Landim, 2006).

3. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

As análises foram feitas utilizando o software livre R com a contribuição do pacote geoR.

A **Figura 2** a seguir ilustra a distribuição espacial dos valores da média da variável da argila do solo nos pontos de coleta. Este é um gráfico de círculo dos dados em que cada dado (x_i, y_i) é representado por um círculo com centro em x_i e raio proporcional ao y_i .

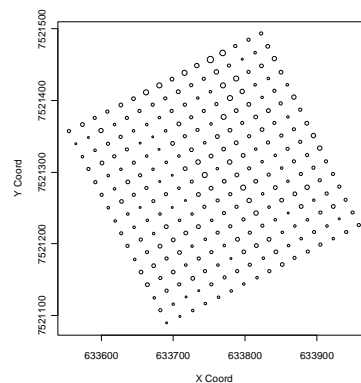


Figura 2 – Gráfico de círculos dos valores da argila do solo nos pontos amostrais.

Os dados foram amostrados em 256 localidades e a distância mínima entre quaisquer dois locais é cerca de 20 metros e a máxima é 424. Os valores verificados para argila no lote deste estudo apresentaram valores mínimo e máximo de 0,0001 a 12,3300, respectivamente, com média 4,1510, mediana 4 e desvio padrão de 5,3330.

O histograma dos 256 valores (**Figura 3**) indica uma assimetria dos dados mostrando que uma transformação da variável resposta deve ser utilizada a fim de tornar o modelo linear Gaussiano aplicável.

Na **Figura 4** é apresentado o gráfico exploratório dos dados transformados. Nota-se que aparentemente existe dependência espacial, mas este fato pode ser confirmado através do variograma.

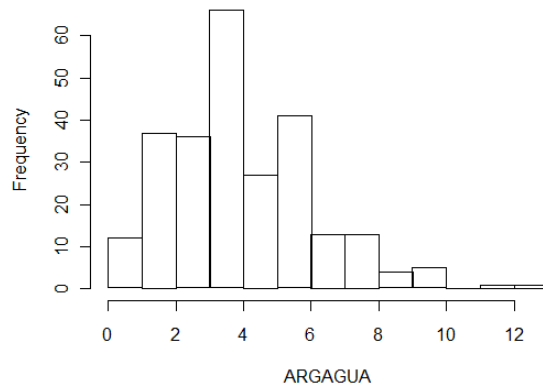


Figura 3 – Histograma de frequência dos valores das amostras de argila.

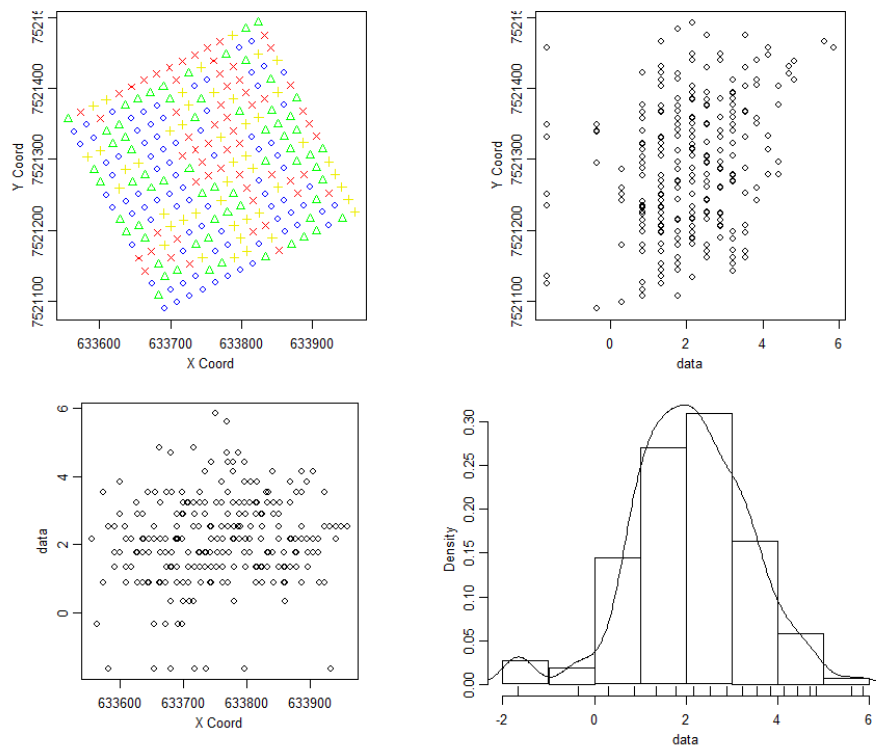


Figura 4 – Locais dos pontos (superior esquerdo), valores dos dados contra coordenadas (canto superior direito e inferior esquerdo) e histograma (canto inferior direito) das medidas transformadas de argila.

O variograma apresentado na **Figura 5** confirma a correlação espacial dos dados estudados, indicados na **Figura 4**.

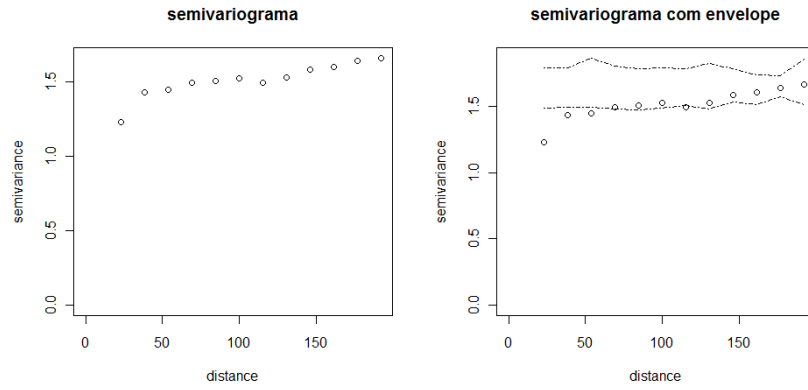


Figura 5 – Variogramas empíricos para os dados transformados a esquerda e variograma dos dados transformados com envelope a direita.

Obtivemos então as estimativas de máxima verossimilhança do modelo Gaussiano cujos parâmetros estimados foram: $\hat{\beta}_0 = 2,1737$, $\hat{\tau} = 0,1181$; $\hat{\sigma}^2 = 1,5779$ e $\hat{\phi} = 20,4385$. Na **Figura 6** observa-se de uma forma mais ampla, a distribuição espacial através do mapa de krigagem.

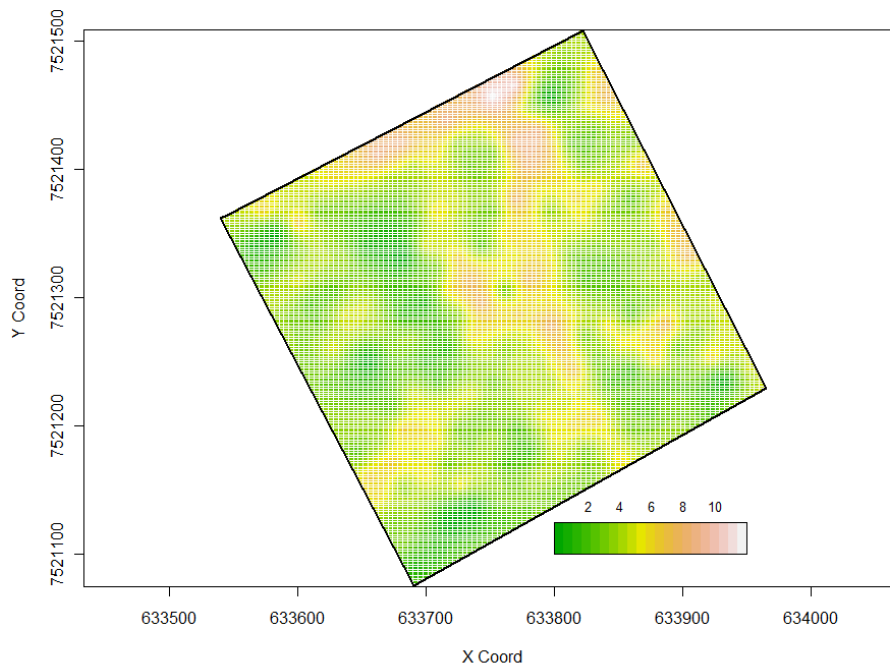


Figura 6 – Krigagem simples de previsões para os dados de argila da superfície.

4. CONCLUSÕES

Os procedimentos adotados no presente trabalho possibilitaram o estudo da dependência espacial da argila do solo, de forma a tornar viável a execução das análises geoestatísticas, envolvendo desde o estabelecimento da função semivariograma até a execução de interpolação espacial, por krigagem. Como resultado observou-se que a área apresentou teores de argila baixos, médios e altos, mostrando a existência da variação da argila no solo, no lote em estudo.

6. REFERÊNCIAS

BARRETO, Tiago E., **Dinâmica de 10,24ha de um trecho de floresta estacional semidecidual, Gália – Alvilândia, SP, Brasil**, 2009. Dissertação (Mestrado em Biologia Vegetal) - Universidade Estadual de Campinas, Campinas, SP, 2009.

DIGGLE, Peter J.; RIBEIRO JR., Paulo J., *Model-based Geostatistics*, Springer, 2007.

LANDIM P.M.B.. 2006. Sobre Geoestatística e mapas. *Terra Didática*, 2(1):19-33. <<http://www.ige.unicamp.br/terraedidatica/>>

RIBEIRO JR, Peter J.; DIGGLE, Paulo J.; **geoR: a package for geostatistical analysis**. In: R NEWS, 1(2):15-18. June, 2001.

VASCONCELO Thybério L. F. C.; LEITE, Eugênio P. F., **Espacialização do PH do solo através do uso de técnicas geoestatísticas**. In II Congresso de Pesquisa e Inovação da Rede Norte Nordeste de Educação Tecnológica João Pessoa, PB, 2007.

APÊNDICE

Procedimentos utilizados no software R, para análise descritiva dos dados, criação do modelo e krigagem.

```
# análise geoestatística - Relatório

rm(list=ls(all=T))

setwd("C:/Users/Usuario/Desktop/Disciplinas II semestre/Geoestatística/Analise/testes")
read.table("agua.txt", head=T)
dados=read.table("agua.txt", head=T)
str(dados)

require(geoR)

dec=as.geodata(dados,coords.col=c(1,2), data.col=3)

## visualizando os dados
x11()
points(dec)
summary(dec)
plot(dec, low=T)

# Histograma dos dados
with(dec, hist(data, main = "", xlab = "ARGAGUA"))

## definindo bordas
points(dec)
bor=(locator(type="l"))
#convertendo lista em matriz
bor=matrix(unlist(bor),ncol=2)
bor=rbind(bor,bor[1,])

####Transformação da variável
require(MASS)
boxcox(dec, lam=seq(0, 1, len=100))
plot(dec, lam=0.6)

## obtendo o variograma para a variável transformada
par(mfrow=c(1,2))
v <- variog(dec,lam=0.6,max.dist=200)
plot(v, main="semivariograma")
v.mc <- variog.mc.env(dec,obj.variog=v)
plot(v, env=v.mc, main="semivariograma com envelope")

## escolhendo função de variograma
x11()
ef <- eyefit(v)
ef

## definindo um "grid"de predição
gr=pred_grid(bor,by=1)

## fazendo a predição espacial (krigagem)
kc <- krige.conv(dec, loc=gr,borders=bor, krige=krige.control(obj.m=ef))

## visualizando os valores preditos na forma de um mapa
image(kc, col=terrain.colors(21))
points(dec, add=T)

## estimação de parâmetros e predição espacial
(ml1 <- likfit(dec, ini=c(2, 15), nug=1.5,lambda=0.6))

Kc1 <- krige.conv(dec, loc=gr, krige=krige.control(obj.model=ml1))

## mapa de predição por krigagem (E[S|y])
image(kc1, col=terrain.colors(21),
      x.leg=c(633800,633950), y.leg=c(7521100, 7521125))
```