

# UTILIZAÇÃO DE GEOESTATÍSTICA PARA O ESTUDO DO ATRIBUTO COBRE NO SOLO

*Gabriel Tambarussi Avancini<sup>1</sup>*

## RESUMO

Este trabalho teve como objetivo estudar a espacialização do atributo cobre, Cu, do solo utilizando técnicas geoestatísticas. A área do estudo é uma área experimental localizada na Estação Exeperimental Camantulul, município de Santa Lúcia Cotzumalguapa, Guatemala. As etapas deste estudo consiste em fazer uma análise exploratória dos dados, ajustar um modelo, fazer o variograma, a krigagem, um mapa do atributo em estudo, um mapa de erro padrão, testar anisotropia e comparar o mapa de predição com um construído a partir de uma abordagem bayesiana. As amostras coletadas do solo consiste em uma malha regular, cobrindo toda a região de estudo. As análises exploratórias dos dados consistiu em fazer análises descritivas e posteriormente foi feita a análise geoestatística que observou um mapa de predição com grande variação do Cu no solo.

**Palavras-Chave:** geoestatística, krigagem ordinária, cobre

## 1 Introdução

Tendo em vista a grande concorrência dos mercados voltados para agricultura, a questão da produtividade é algo de extrema importância, e um ponto muito relevante para o aumento da produtividade é o conhecimento a respeito do solo no qual será plantado as devidas culturas. Geralmente as aplicações na agricultura são baseadas em teores médios dos atributos do solo, o que pode causar, em alguns locais, uma subestimação ou uma superestimação destes atributos no solo, causando assim um déficit ou excesso de nutrientes para a cultura em determinados locais, (Souza et al, 2004).

Assim, utilizaremos a análise geoestatística, que é considerada uma técnica difundida para estes estudos, com o objetivo de fazer um mapa de predição para o teor de cobre, através da Krigagem que atribui valores mais parecidos a locais mais próximos e portanto, estamos trabalhando com a suposição de dependência que é incorporada na análise através das covariâncias

---

<sup>1</sup>Mestrando no Programa de Pós-Graduação em Estatística e Experimentação Agronômica - ESALQ/USP

e semi-variâncias, este fato é confirmado em (Ribeiro Junior, 1995):

“(...) cada observação é descrita não apenas pelo seu valor, mas também por informações de sua posição, expressa por um sistema de coordenadas. Considerando que observações mais próximas geograficamente tendem a ter valores mais similares e que tal fato pode ser avaliado por medidas de associação, fica claro que em geoestatística não se pode trabalhar com a suposição de independência.”

## 2 Fonte dos dados

Os dados são provenientes de uma área experimental localizada na Estação Experimental Camantulul, município de Santa Lúcia Cotzumalguapa, Guatemala na qual foram retiradas amostras do solo, igualmente espaçadas, e observado 22 atributos. O objeto de estudo deste trabalho será o cobre, Cu, pois este teve correlação espacial enquanto em outros atributos não foi detectada esta correlação.

A Figura 1 representa os locais nos quais as amostras foram retiradas e o tamanho dos pontos representa a quantidade do atributo Cu, e as cores representam os quintis.

## 3 Metodologia

Utilizaremos uma metodologia geoestatística pois os valores considerados neste estudo possuem uma componente aleatória e um certa estrutura de dependência espacial, e portanto a geoestatística é a metodologia mais indicada para a análise destes dados.

Denotaremos por  $Y_i$  o atributo cobre, Cu, medido no local  $x_i$ , que será o vetor de coordenadas do local. Defini-se um processo espacial Gaussiano estacionário  $\{Y(x) : x \in \mathbb{R}^2\}$  como sendo:

- $E[Y(x)] = \mu, \quad \forall x;$
- $Var[Y(x)] = \sigma^2;$
- $\rho^e(u) = \text{Corr}[Y(x_i), Y(x_j)],$  a função de correlação para  $i \neq j.$

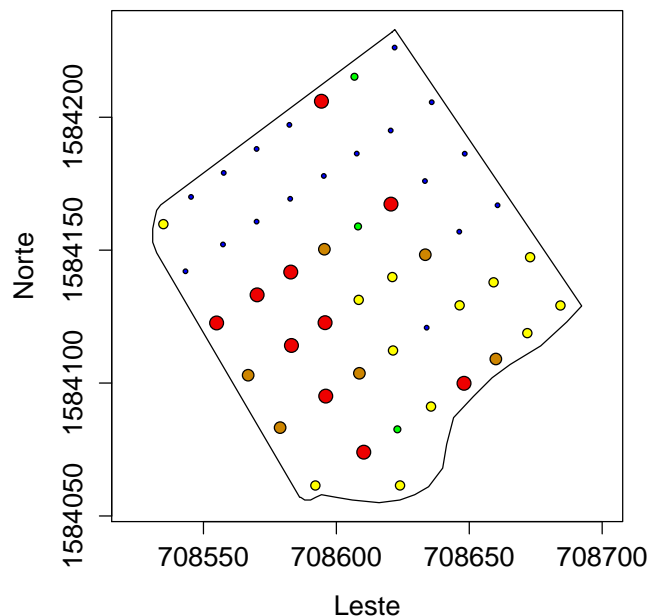


Figura 1: Planejamento do experimento

em que  $u = \|x_i - x_j\|$  correspondente a distância Euclidiana (distância usual) entre  $x_i$  e  $x_j$ , (Diggle e Ribeiro Jr., 2007).

Também podemos definir a covariância  $\text{Cov}[Y(x_i), Y(x_j)] = \gamma^e \|x_i - x_j\|$ . Assim, o processo sendo estacionário temos:  $\gamma^e(u) = \sigma^2 \rho^e(u)$ .

Considerando  $\text{Var}[Y(x_i), Y(x_j)] = 2V(x_i, x_j)$ , chamaremos  $2V(\cdot)$  de *função variograma* (Matheron, 1963), e chamaremos  $V(\cdot)$  de *função semi-variograma*.

## 4 Resultados e Discussões

### 4.1 Análise exploratória

Primeiramente foi feito um box plot dos dados verificando que a variabilidade dos não não é muito grande e que há um ponto atípico.

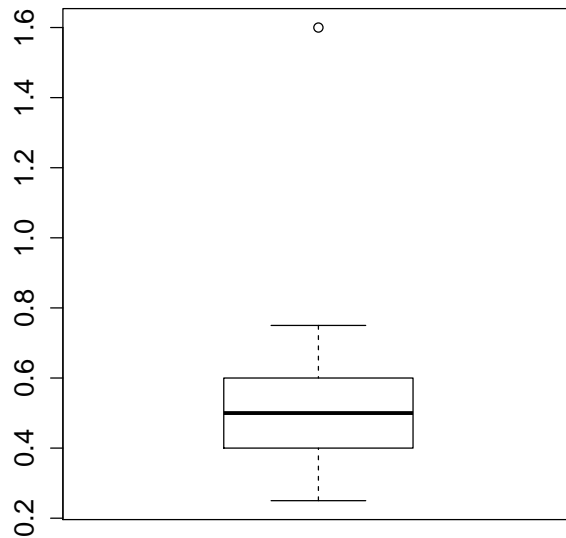


Figura 2: Box-plot dos dados

A Tabela 1 representa uma análise geral dos dados e assim podemos observar que a média do teor de Cu é de 0,5245, que é considerado um valor baixo, temos também que a menor distância entre duas observações é de 14,59 unidades, o que representa uma grande parte da maior distância que é de 167,42, o que não é aconselhável em um experimento, pois não poderemos estimar o variograma para pequenas distâncias.

Para evitar este tipo de problema, poderíamos ter colhido as observações não em uma malha regular, mas sim em um malha mista (regular com alguns pontos aleatórios) ou também em um malha aleatória.

Utilizamos o teste Box-Cox, para verificar se havia a necessidade de transformar os dados, a constatação foi positiva e utilizamos  $\lambda = -0,46$ .

A transformação Box-Cox, segundo (BOX e COX, 1964), consiste em:

Tabela 1: Análise descritiva dos dados

<b>observações</b>	49
<b>min coord X</b>	708534.9
<b>max coord X</b>	708684.3
<b>min coord Y</b>	1584061
<b>max coord Y</b>	1584226
<b>min distância</b>	14.59948
<b>max distância</b>	167.42101
<b>Mínimo</b>	0.2500
<b>1º Quartil</b>	0.4
<b>Mediana</b>	0.5
<b>Média</b>	0.5245
<b>3º Quartil</b>	0.6
<b>Máximo</b>	1.6

$$\begin{cases} y_i^* = \frac{y_i^{\lambda^e - 1}}{\lambda^e} & \text{se } \lambda^e \neq 0 \\ y_i^* = \log(y_i) & \text{se } \lambda^e = 0 \end{cases} \quad (1)$$

## 4.2 Análise Geoestatística

Foram testadas todas as possibilidades de tendências até de segundo grau em relação as coordenadas *lat* e *long* porém estas tendências não tiveram um acréscimo significativo no modelo quando inseridas e portanto não foram utilizadas.

O primeiro gráfico da Figura 3 nos dá indícios da dependência espacial já que os atributos representados por símbolos iguais se encontram significativamente aglomerados e estes símbolos representam os quintis. Analisando o gráfico da direita inferior da Figura 3 consideramos que esta satisfeito o pressuposto da normalidade.

A dependência espacial foi verificada utilizando o semivariograma empírico e os envelope simulado, constando a dependência devido ao fato de termos pelo menos um ponto fora do envelope.

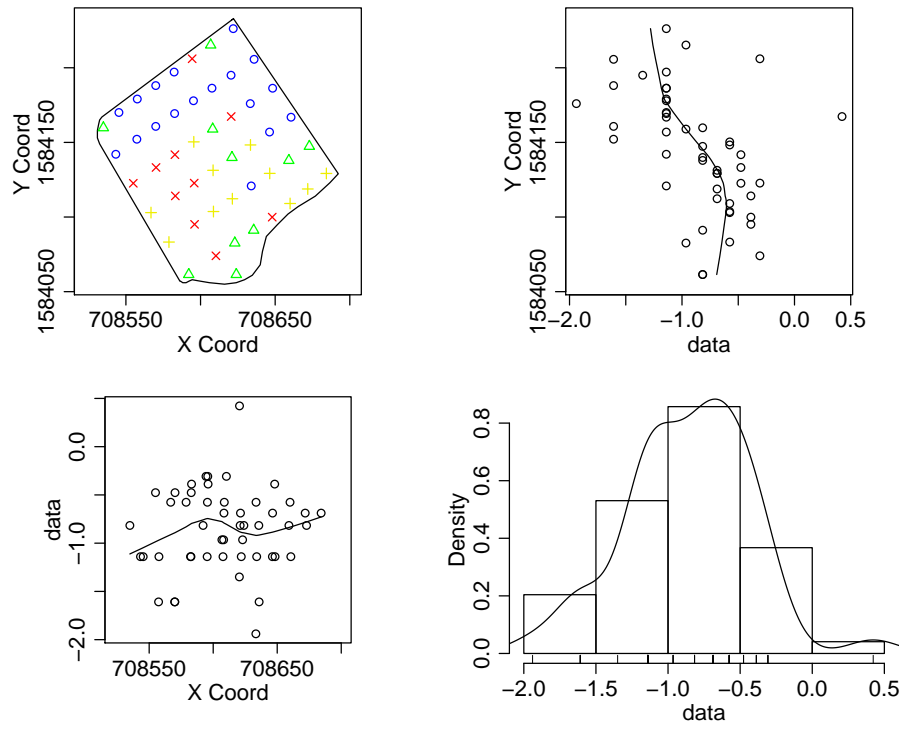


Figura 3: Gráficos descritivos

### 4.3 Ajuste do modelo teórico

O ajuste do modelo foi feito para 7 modelos: Cauchy, Circular, Cúbico, Exponencial, Gaussiano, Matern e Esférico (Cressie, 1993), (Oliveira, 2003). Sendo os valores iniciais dados pela função *eyefit* do pacote *geoR* do software R (RIBEIRO JR., P. J.; DIGGLE, P. J, 2001).

Para todos os modelos foi comparado os valores dos AIC's para o modelo com efeito espacial e sem efeito espacial, e como esperado em todos os modelos os AIC's com dependência espacial foram menores, indicando a a dependência espacial dos dados e portanto que a análise geoestatística está correta em ser aplicada.

Também foi testado para todos os modelos encaixados o teste da razão de máxima verossimilhança para confirmar o fato de que as covariáveis não eram significativas ao serem acrescentados no modelo.

O modelo escolhido foi o modelo circular, já que apresentou o menor AIC, resultando no seguinte semi-variograma:

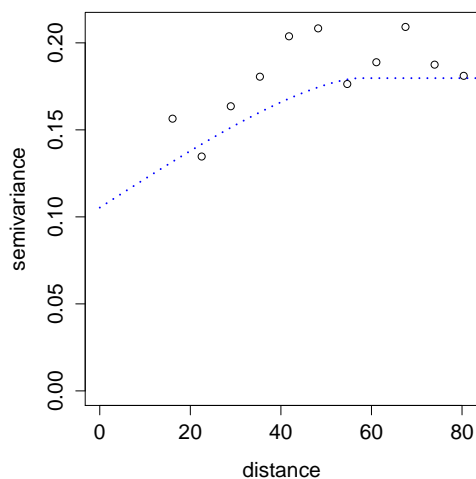


Figura 4: Semivariograma ajustado para os dados utilizando o modelo circular

#### 4.4 Mapa de predição (Krigagem)

Como apenas a informação dada pela variância dos dados não é suficiente para realizar a predição (Krige, 1951), utilizaremos também a distância entre as observações para realizar a predição dos valores do Cu nos locais que estes não foram amostrados. A interpolação foi feita por meio da Krigagem ordinária e também foi feito o mapa de erro padrão, que resultaram, respectivamente na Figura 5.

#### 4.5 Abordagem Bayesiana

Na abordagem bayesiana estamos considerando uma incerteza nos parâmetros do modelo. Precisamos definir as distribuições a priori dos parâmetros do modelo, porém esta escolha é não é tão simples e esta sujeita a muitas discussões. Foram escolhidas a distribuição Normal, para o parâmetro  $\beta$ , a distribuição Exponencial para o  $\phi$  e a distribuição *sc.inv.chisq* para o parâmetro  $\sigma$ .

A Figura 6 representa o mapa gerado pela abordagem bayesiana

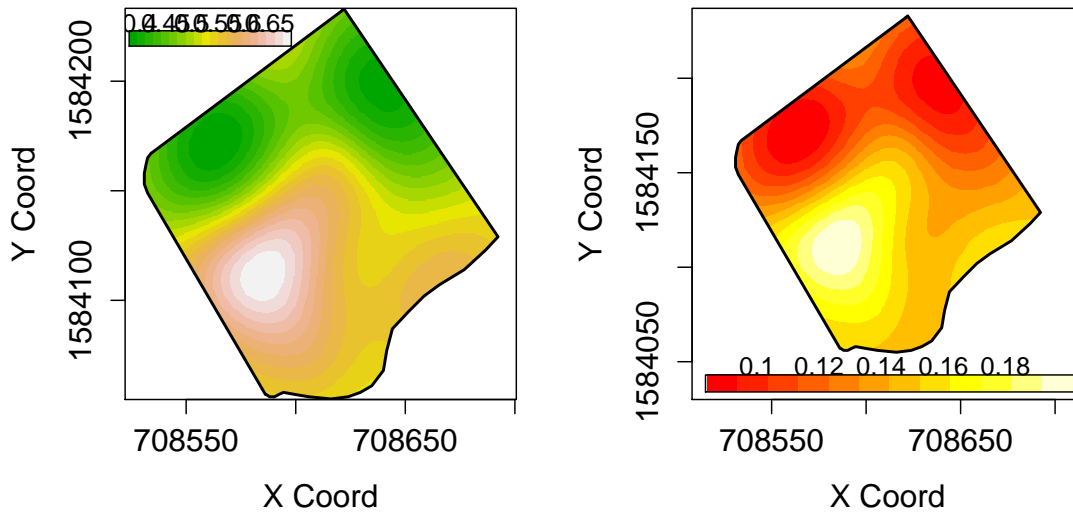


Figura 5: Mapa de krigagem (Esquerda) e mapa dos erros padrões (Direita)

## 5 Conclusão

Concluimos então que a utilização da geoestatística de fato foi apropriada devido as comprovações de dependência espacial pelo envelope simulado, já que obtivemos pontos fora do mesmo. O modelo que se ajustou melhor aos dados foi o modelo circular, contrariando o senso comum de utilizar o modelo exponencial ou o modelo esférico para o estudo de atributos no solo.

Foi mapeado os valores esperados do atributo Cu em toda a região experimental. O mapa poderia ser mais real se o experimento fosse feito em uma malha que proporcionasse conclusões melhores a respeito de observações mais próximas e portanto se não fosse desenvolvido em uma malha regular e sim uma mista (regular e aleatória) ou em uma malha aleatória.

A utilização da abordagem bayesiana modificou bastante o mapa de predição, o que pode ter ocorrido pelo fato de que as prioris definidas não foram escolhidas da melhor maneira possível.



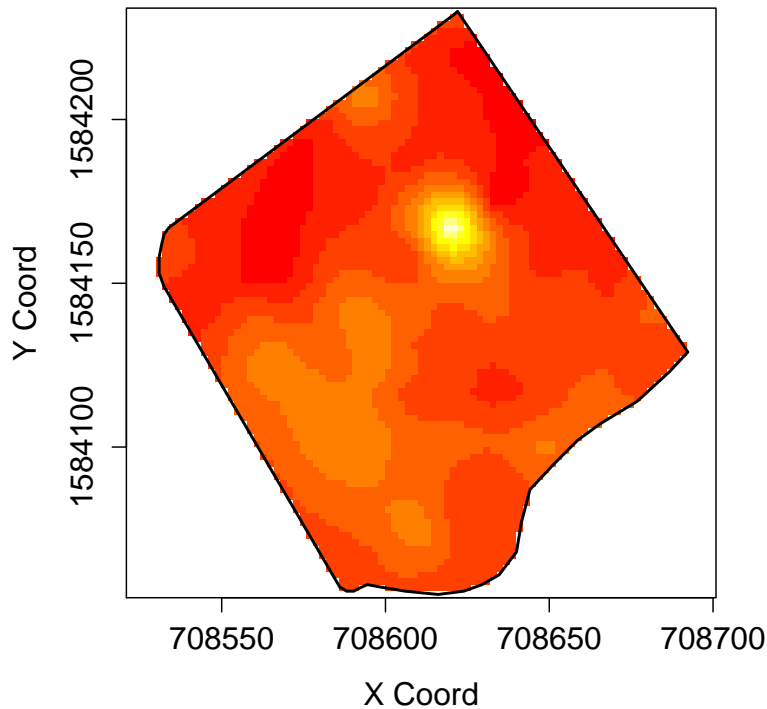


Figura 6: Mapa de krigagem realizado com a abordagem bayesiana

## 6 Referência Bibliográfica

BOX, G. E. P.; COX, D. R. **An Analysis of transformations**. Journal of the Royal Statistical Society, v.26, n.2, p.211-252, 1964.

CRESSIE, N.A.C. **Statistics for spatial data**. New York: John Wiley & Sons 900p, 1993.

DIGGLE, P. J.; RIBEIRO JR, P. J. **Model-based geostatistics**. New York: Springer, 2007.

Krige, 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.

MATHERON, G. **Principles of geostatistics**. Economic Geology, v.58, p.1246-1266, 1963.

Oliveira, M. C. N. **Métodos de estimação de parâmetros em modelos geoestatísticos**

**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, 2003.

RIBEIRO JUNIOR, P.J. **Métodos geostatísticos no estudo da variabilidade espacial de parâmetros do solo.** Dissertação de Mestrado, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1995.

RIBEIRO JR., P. J.; DIGGLE, P. J. **geoR: a package for geostatistical analysis.** R.News, v.1, n.2, p.15-18, 2001.

SOUZA, Z. M.; MARQUES JÚNIOR, J.; PEREIRA, G. T.; MOREIRA, L. F. **Variabilidade espacial do pH, Ca, Mg e V% do solo em diferentes formas do relevo sob cultivo de cana-de-açúcar.** Ciência Rural, Santa Maria, v.34, n.6, p.1763-1771, 2004.