

INTEGRAÇÃO DE GEOESTATÍSTICA E SISTEMAS DE INFORMAÇÃO GEOGRÁFICA: UMA NECESSIDADE

EDUARDO CELSO GERBI CAMARGO

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE
DEPARTAMENTO DE PROCESSAMENTO DE IMAGENS - DPI
AV. DOS ASTRONAUTAS, 1758 CEP: 12201-970
SÃO JOSÉ DOS CAMPOS - SÃO PAULO - BRASIL
FONE: (012) 3456524 FAX: (012) 3456468
E-MAIL: eduardo@dpi.inpe.br

ANTONIO MIGUEL VIEIRA MONTEIRO

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE
DEPARTAMENTO DE PROCESSAMENTO DE IMAGENS - DPI
AV. DOS ASTRONAUTAS, 1758 CEP: 12201-970
SÃO JOSÉ DOS CAMPOS - SÃO PAULO - BRASIL
FONE: (012) 3456474 FAX: (012) 3456468
E-MAIL: miguel@dpi.inpe.br

CARLOS ALBERTO FELGUEIRAS

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE
DEPARTAMENTO DE PROCESSAMENTO DE IMAGENS - DPI
AV. DOS ASTRONAUTAS, 1758 CEP: 12201-970
SÃO JOSÉ DOS CAMPOS - SÃO PAULO - BRASIL
FONE: (012) 3456519 FAX: (012) 3456468
E-MAIL: carlos@dpi.inpe.br

SUZANA DRUCK FUKS

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA
CENTRO NACIONAL DE PESQUISAS EM SOLOS - CNPS
RUA JARDIM BOTÂNICO, 1024 CEP: 22460-000
RIO DE JANEIRO - RIO DE JANEIRO - BRASIL
FONE: (021) 2744999 FAX: (021) 2745291
E-MAIL: suzana@cnps.embrapa.br

RESUMO

Este trabalho tem como objetivo enfatizar a importância da integração Geoestatística e Sistemas de Informação Geográfica (SIG). A plataforma de integração é o Sistema de Processamento de Informações Georreferenciadas (SPRING), de propriedade do Instituto Nacional de Pesquisas Espaciais, desenvolvido na Divisão de Processamento de Imagens (DPI). A base conceitual deste trabalho está fundamentada em técnicas geoestatísticas, as quais são derivadas da Teoria das Variáveis Regionalizadas e formalizadas por Matheron (1971). Um exemplo é apresentado. Através de procedimentos geoestatísticos uma análise espacial do teor de argila é realizada, nos solos da Fazenda Canchim em São Carlos – SP, e os resultados são apresentados e discutidos.

ABSTRACT

This work emphasizes the importance of integrating Geostatistics and Geographic Information Systems (GIS). The Georeferenced Information Processing System (SPRING), developed at the Image Processing Division (DPI) of the National Institute for Space Research, is used to test such an integration. The conceptual basis of this study finds its foundations on geostatistics techniques. These techniques are derived from Matheron's theory on regionalized variables (Matheron, 1971). A case study is presented, which applies geostatistics procedures to spatially analyze the amount of clay on the Canchim Farm soils in São Carlos, SP. Results are then described and discussed.

INTRODUÇÃO

Os Sistemas de Informação Geográfica (SIGs), devido a sua crescente utilização, vem incorporando novos princípios e procedimentos de análise espacial. Esse esforço deve-se ao consenso, hoje existente, de que o futuro sucesso da tecnologia dos SIGs depende de sua capacidade de gerar cenários confiáveis, através de um conjunto de procedimentos e ferramentas para a modelagem espacial, que possam representar adequadamente, em meio digital, as propriedades, os processos e suas imprecisões.

Alguns processos espaciais, principalmente aqueles observados em aplicações ambientais, apresentam indexação no espaço e trazem como característica comum a continuidade, observando que seus valores variam de forma gradual numa determinada vizinhança. A representação destas propriedades nos SIGs, em geral, são oriundas de modelos limitados e, como consequência, o produto final é uma simplificação porque não trabalha a variabilidade espacial das propriedades e fenômenos ambientais em estudo.

A associação de procedimentos de estatísticas espaciais aos SIGs é importante porque permite representações com hipóteses mais próximas do contínuo dos fenômenos ambientais e também porque quantifica as imprecisões associadas aos produtos trabalhados.

Neste contexto, este trabalho apresenta o componente de análise espacial, baseado nas técnicas geoestatísticas, integrado ao Sistema de Processamento de Informação Georreferenciada - SPRING. O componente consiste de um conjunto de procedimentos integrados ao sistema, que possibilita novos modelos inferenciais e utiliza-se da biblioteca geoestatística, de domínio público, GSLIB (Deutsch e Journel, 1992).

Um exemplo é apresentado. Partindo de um conjunto de amostras de solos, foram aplicadas técnicas geoestatísticas para se obter a espacialização da variação do teor de argila nos solos da Fazenda Canchim, em São Carlos, SP. Uma análise comparativa, da representação espacial oriunda de procedimentos geoestatísticos, com outros métodos disponíveis no sistema SPRING foi realizada e os resultados são apresentados e discutidos.

FUNDAMENTOS DE GEOESTATÍSTICA

A geoestatística envolve análise e inferência de fenômeno espacial e ou temporal, tal como concentração de poluentes, variação do teor de zinco no solo, preço do petróleo no tempo, etc. A geoestatística teve sua origem na área de mineração (Krige, 1951). Atualmente, geoestatística é justamente um nome associado com uma classe de técnicas usada para analisar e inferir valores de uma variável distribuída no espaço e ou no tempo. Tais valores são implicitamente assumidos ser correlacionados com outros, e o estudo de tal correlação é denominada de análise estrutural ou modelagem do variograma. Depois da análise estrutural, inferências em localizações não amostradas são realizadas usando "Krigagem" ou podem ser *Simuladas* usando simulações condicionais. Resumidamente, os passos num estudo empregando técnicas geoestatísticas inclui: (a) análise exploratória dos dados, (b) análise estrutural (cálculo e modelagem do variograma) e (c) realização de inferências (Krigagem ou Simulação). Este trabalho, limita-se à realizar inferências utilizando as técnicas de Krigagem Simples e Ordinária.

O sucesso e as suposições das técnicas geoestatísticas estão fundamentadas na *Teoria das Variáveis Regionalizadas*, desenvolvida por Matheron (1963, 1971). A variável regionalizada é uma variável distribuída no espaço (ou tempo). A teoria diz que alguma medida pode ser vista como uma realização de uma função aleatória (ou processo aleatório, ou campo aleatório, ou processo estocástico). Esta teoria consolida o alicerce da geoestatística.

Segundo Burrough (1987), a variação espacial de uma variável regionalizada pode ser expressa pela soma de três componentes: a) uma componente estrutural, associada a um valor médio constante ou a uma tendência constante; b) uma componente aleatória, espacialmente correlacionada; e c) um ruído aleatório ou erro residual.

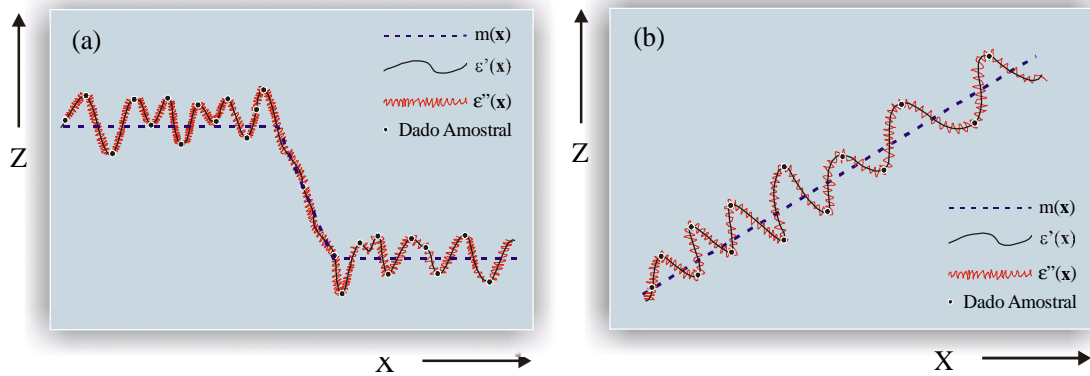
Se \mathbf{x} representa uma posição em uma, duas ou três dimensões, então o valor da variável Z , em \mathbf{x} , é dada por:

$$Z(\mathbf{x}) = m(\mathbf{x}) + \varepsilon'(\mathbf{x}) + \varepsilon'' \quad (1)$$

onde:

- $m(\mathbf{x})$ é uma função determinística que descreve a componente estrutural de Z em \mathbf{x} ;
- $\varepsilon'(\mathbf{x})$ é um termo estocástico, que varia localmente e depende espacialmente de $m(\mathbf{x})$;
- ε'' é um ruído aleatório não correlacionado, com distribuição normal com média zero e variância σ^2 .

A Figura 1(a) e 1(b) ilustram as três componentes principais da variação espacial. A Figura 1(a) apresenta uma componente determinística que varia abruptamente, enquanto a componente determinística, na Figura 1(b), apresenta uma tendência constante.



Figs. 1(a) e 1(b) - Principais componentes da variação espacial.

FONTE: Modificada de Burrough (1987), p. 155.

Diferente de outras técnicas de inferência, a Krigagem Simples e ou Ordinária requer algumas hipóteses (Burrough, 1987 e David, 1977). No caso mais simples, admite-se que a componente determinística, $m(\mathbf{x})$, é constante (não há tendências na região). Então, $m(\mathbf{x})$ é igual ao valor esperado da variável aleatória Z na posição \mathbf{x} , e a diferença média entre os valores observados em, \mathbf{x} e $\mathbf{x}+\mathbf{h}$, separados por um vetor de distância \mathbf{h} (módulo e direção) é nula.

$$E[Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})] = 0 \quad \text{ou} \quad E[Z(\mathbf{x})] = E[Z(\mathbf{x}+\mathbf{h})] = m(\mathbf{x}) = m \quad (2)$$

Admite-se também a estacionariedade da covariância, isto é, a covariância entre dois pares quaisquer $Z(\mathbf{x})$ e $Z(\mathbf{x}+\mathbf{h})$, separados por um vetor distância \mathbf{h} , existe e depende somente de \mathbf{h} . Então:

$$C(\mathbf{h}) = \text{Cov} [Z(\mathbf{x}), Z(\mathbf{x}+\mathbf{h})] = E[Z(\mathbf{x}).Z(\mathbf{x}+\mathbf{h})]-m^2, \quad \forall \mathbf{x}; \quad (3)$$

A estacionariedade da covariância também implica na estacionariedade do variograma, definido por:

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

As considerações acima resumem a hipótese de estacionariedade de 2ª ordem. Porém, na geoestatística adota-se uma hipótese de estacionariedade menos restritiva denominada *Intrínseca*. Neste caso, admite-se que $E[Z(\mathbf{x})] = m(\mathbf{x}) = m, \forall \mathbf{x}$ e que a variância das diferenças depende somente do vetor distância \mathbf{h} , isto é:

$$\text{Var}[Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h})] = E\{[Z(\mathbf{x})-Z(\mathbf{x}+\mathbf{h})]^2\} = 2\gamma(\mathbf{h}) \quad (5)$$

onde $2\gamma(\mathbf{h})$ é o variograma, conforme definido anteriormante.

Se as condições especificadas pela hipótese intrínseca são contempladas, então o semivariograma pode ser estimado a partir do dado amostral:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})]^2 \quad (6)$$

onde $N(\mathbf{h})$ é o número de pares de pontos amostrais separados pelo vetor distância \mathbf{h} .

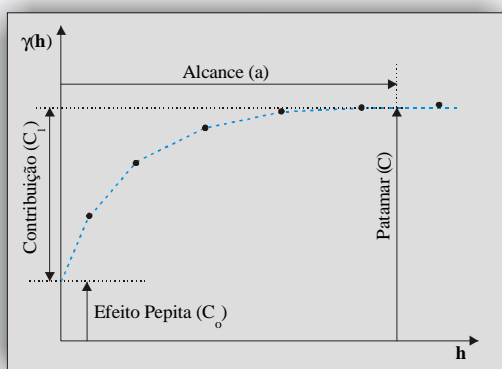


Fig. 2 – Semivariograma Experimental.

O gráfico $\gamma(\mathbf{h})$ versus \mathbf{h} é conhecido como semivariograma experimental. A Figura 2 ilustra um semivariograma experimental com características muito próximas do ideal. O seu padrão representa o que, intuitivamente, se espera de dados de campo, isto é, que as diferenças $\{Z(\mathbf{x}_i) - Z(\mathbf{x}_i + \mathbf{h})\}$ decresçam à medida que \mathbf{h} diminui. É esperado que observações mais próximas geograficamente tenham um comportamento mais semelhante entre si do que aquelas separadas por maiores distâncias. Desta maneira, é esperado que $\gamma(\mathbf{h})$ aumente com a distância \mathbf{h} .

O procedimento de ajuste, do semivariograma experimental, não é direto e automático, como no caso de uma regressão, por exemplo, mas sim interativo, pois nesse processo o intérprete faz um primeiro ajuste e verifica a

adequação do modelo teórico. Resumidamente, os modelos de ajuste estão divididos em dois tipos (Isaaks e Srivastava, 1989): modelos com patamar e modelos sem patamar. Modelos do primeiro tipo são referenciados na geoestatística como modelos transitivos. Exemplos de modelos transitivos são: Esférico, conforme definido na equação (7), Exponencial e Gaussiano.

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , |\mathbf{h}| = 0 \\ C_0 + C_1 \left[\frac{3}{2} \left(\frac{|\mathbf{h}|}{a} \right) - \frac{1}{2} \left(\frac{|\mathbf{h}|}{a} \right)^3 \right] = C_0 + C_1 [\text{Sph}(|\mathbf{h}|)] & , 0 < |\mathbf{h}| \leq a \\ C_0 + C_1 & , |\mathbf{h}| > a \end{cases} \quad (7)$$

onde:

C_0 é o efeito pepita, C_1 a contribuição do modelo teórico e a o alcance.

Modelos do segundo tipo não atingem o patamar, e continuam aumentando enquanto a distância aumenta. Tais modelos são utilizados para modelar fenômenos que possuem capacidade infinita de dispersão. Exemplo deste último podemos citar o modelo Potência.

Quanto aos métodos de inferência de Krigagem Simples e Ordinária, seus estimadores são conforme as equações (8) e (9) respectivamente (Journal, 1988):

$$Z_{\mathbf{x}_0}^* = m + \sum_{i=1}^n \lambda_i [Z(\mathbf{x}_i) - m] \quad (8)$$

$$Z_{\mathbf{x}_0}^* = \sum_{i=1}^n \lambda_i Z(\mathbf{x}_i) \quad , \quad \text{com} \quad \sum_{i=1}^n \lambda_i = 1 \quad (9)$$

Nas equações (8) e (9) o que difere o estimador de Krigagem Simples do estimador de Krigagem Ordinária, é que este último não requer o prévio conhecimento da média amostral m .

A grande diferença entre a Krigagem e outros métodos de inferência é a maneira como os pesos são atribuídos às diferentes amostras. No caso da interpolação por média simples, por exemplo, os pesos são todos iguais a $1/N$ (N = número de amostras); na interpolação baseada no inverso do quadrado das distâncias, os pesos são definidos como o inverso do quadrado da distância que separa o valor interpolado dos valores observados. Na Krigagem, o procedimento é semelhante ao de interpolação por média móvel ponderada, exceto que aqui os pesos são determinados a partir de uma análise espacial, baseada no semivariograma experimental.

Journal (1988) mostra que, o sistema de Krigagem simples e ordinária podem ser escritos em notação matricial como:

$$\mathbf{K} \cdot \boldsymbol{\lambda} = \mathbf{k} \Rightarrow \boldsymbol{\lambda} = \mathbf{K}^{-1} \cdot \mathbf{k} \quad (10)$$

onde, \mathbf{K} e \mathbf{k} são matrizes das covariâncias (ou variogramas) e $\boldsymbol{\lambda}$ o vetor dos pesos.

Por exemplo, no caso da Krigagem Ordinária as matrizes são assim definidas:

$$\mathbf{K} = \begin{vmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ : & : & & : & : \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{vmatrix}, \quad \boldsymbol{\lambda} = \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ : \\ \lambda_n \end{vmatrix} \quad \text{e} \quad \mathbf{k} = \begin{vmatrix} C_{10} \\ C_{20} \\ : \\ C_{n0} \\ 1 \end{vmatrix}$$

Determinado os pesos, equação (10), aplica-se o estimador da equação (8) ou (9). Além da inferência de valores, esta técnica fornece também uma medida de imprecisão denominada de Variância de Krigagem (Simples ou Ordinária), dada por (Journal, 1988) como: $\sigma_k^2 = C(0) - \boldsymbol{\lambda}^T \cdot \mathbf{k}$ (11)

O MÓDULO DE GEOESTATÍSTICA NO SPRING

O Módulo de procedimentos geoestatísticos, integrado ao sistema SPRING, tem como objetivo a análise em duas dimensões, para dados espacialmente distribuídos, no que diz respeito a inferência de superfícies geradas a partir de amostras georreferenciadas. A estrutura implementada utiliza a biblioteca de procedimentos geoestatísticos GSLIB e segue o modelo apresentado na Figura 3.

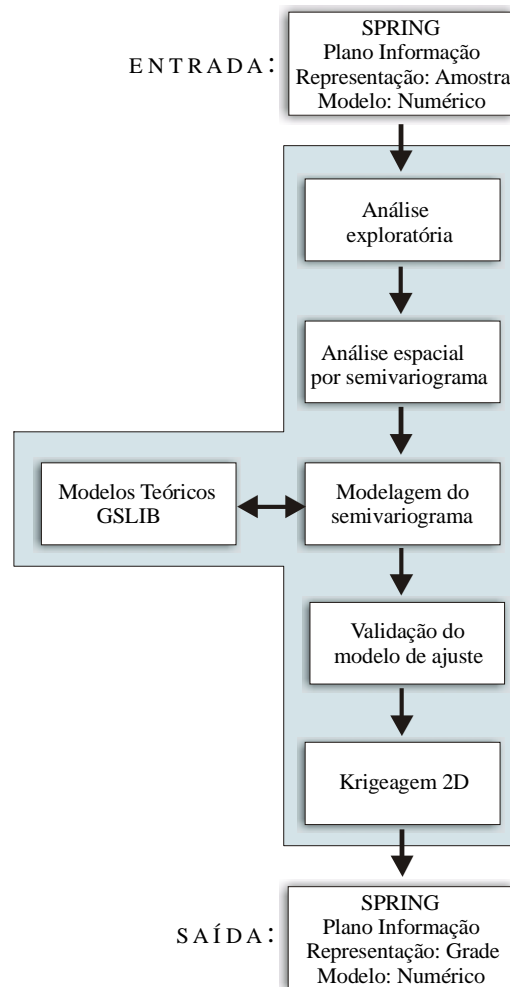


Fig. 3 – Estrutura do Módulo de Geoestatística no SPRING

Inicialmente sobre o dado de entrada, neste caso proveniente de um Plano de Informação (PI), é realizada uma análise exploratória através de estatísticas univariadas e bivariadas. As estatísticas univariadas fornecem um meio de organizar e sintetizar um conjunto de valores, que se realiza principalmente através do histograma. As estatísticas bivariadas fornecem meios de descrever o relacionamento entre duas variáveis, isto é, entre dois conjuntos de dados ou de duas distribuições. Esta relação pode ser visualizada através do diagrama de dispersão. Seguindo, uma análise de continuidade espacial baseada em semivariograma é realizada. A análise completa do semivariograma compreende os seguintes passos: (a) levantamento do semivariograma experimental, (b) ajuste a uma família de modelos de semivariogramas e (c) validação do modelo a ser utilizado nos procedimentos da Krigeagem. Para finalizar, a inferência dos dados é realizada utilizando as técnicas de Krigeagem e o resultado armazenado num Plano de Informação do sistema SPRING.

Os procedimentos gerais que compõem o módulo estão agrupados no menu principal do SPRING, conforme mostra a Figura 4, e possibilitam que as etapas fundamentais da geoestatística sejam executadas.

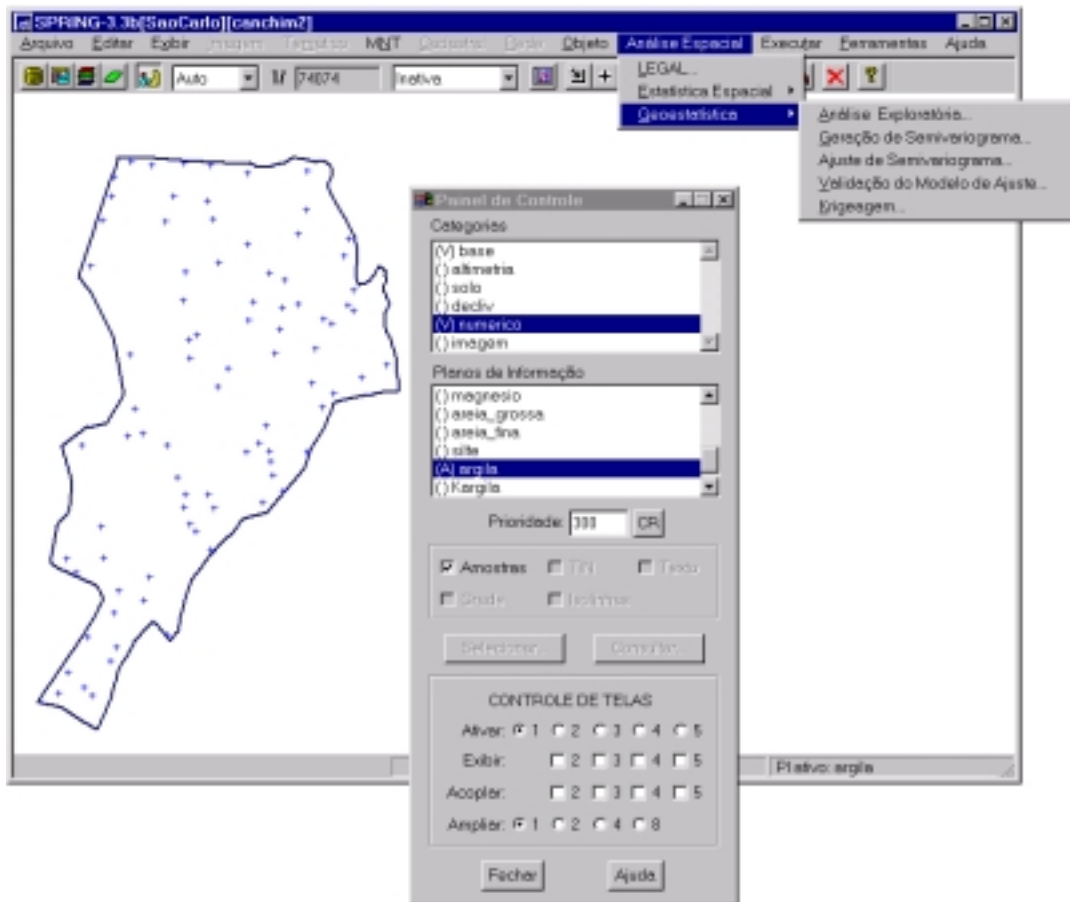


Fig. 4 – Tela Principal do sistema SPRING

EXEMPLO

Este exemplo tem como objetivo enfatizar, através de um experimento prático, a importância da integração Geoestatística - SIG na análise e inferência de dados espacialmente distribuídos.

Os dados utilizados, de propriedade do Centro Nacional de Pesquisas de Solos (CNPS - RJ), foram obtidos no levantamento dos solos da Fazenda Canchim. Estes se referem a uma amostragem de 85 observações georreferenciadas coletadas no horizonte Bw (camada do solo com profundidade média de 1m). Dentre as variáveis disponíveis, selecionou-se para estudo o teor de argila.

Considera-se o teor de argila ao longo do perfil, classificado do seguinte modo (Calderano Filho et al., 1996):

- MUITO ARGILOSO: solos que apresentam 59% ou mais de argila;
- ARGILOSO: solos que apresentam de 35% a 59% de argila;
- MÉDIO: solos que apresentam de 15% a 35% de argila;
- ARENOSO: solos que apresentam menos de 15% de argila.

Para tal experimento, são realizados os seguintes procedimentos: (a) análise exploratória (b) análise estrutural do teor de argila (modelagem variograma) (c) inferência de valores através de Krigagem Ordinária (d) análise e apresentação dos resultados obtidos e (e) comparação dos resultados obtidos com outros métodos convencionais de estimação.

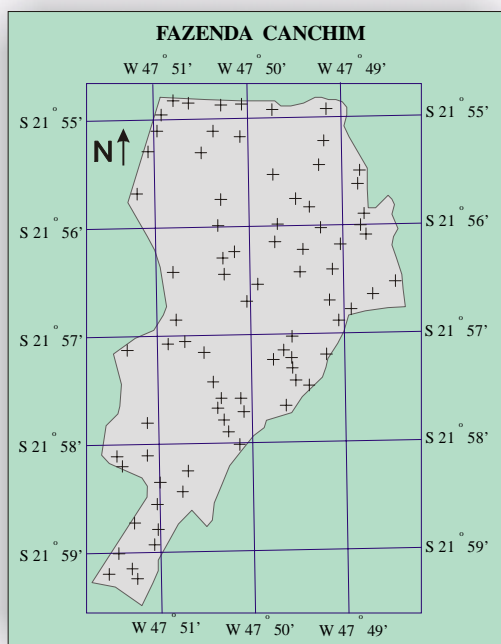


Fig. 5 – Mapa de localização da Fazenda Canchim.

A região de estudo refere-se à área da Fazenda Canchim, base física do Centro de Pesquisa Pecuária do Sudeste (CPPSE- EMBRAPA), no município de São Carlos, Estado de São Paulo, cobrindo uma área de 2.660 hectares, entre as coordenadas de 21°55'00" - 21°59'00" S e 47°48'00" e 47°52'00" W. O mapa de localização da região em estudo é apresentado na Figura 5.

A análise exploratória dos dados é uma etapa de fundamental importância e obrigatória dentro de qualquer estudo geoestatístico. É uma etapa precursora introduzida para averiguar as propriedades estatísticas e matemáticas dos dados (Burrough et al., 1996).

A partir da distribuição das observações, Figura 5, pode-se proceder a uma análise preliminar baseada em inspeção visual. Essa análise permite a identificação de alguns problemas, tais como escassez ou excesso de observações em determinado local, ocorrência de valores aberrantes, etc. Após a análise preliminar, foram obtidas algumas estatísticas da amostra, conforme ilustra a Tabela 1. Além das estatísticas referidas, utilizou-se também, para uma melhor caracterização, os recursos gráficos de histograma, conforme apresentado na Figura 6.

TABELA 1 – ESTATÍSTICAS DO TEOR DE ARGILA.

Número de observações	85
Média	33,035
Variância	288,034
Desvio Padrão	16,972
Coefficiente de variação	0,514
Coefficiente de assimetria	0,214
Coefficiente de curtose	2,344
Valor mínimo	4
Quartil Inferior	10
Mediana	33
Quartil superior	43
Valor máximo	73

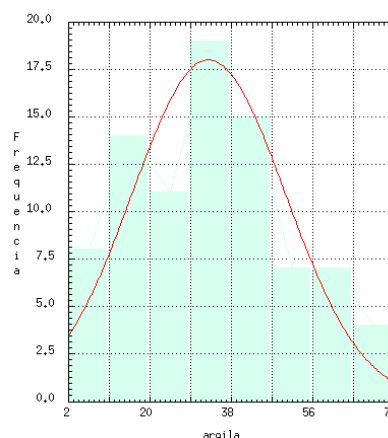


Fig. 6 – Histograma do teor de argila.

O histograma apresentado na Figura 6 mostra que a distribuição do teor de argila é levemente alongada à direita. Neste caso, a distribuição é dita ser positivamente assimétrica (Costa Neto, 1977, p. 30), com coeficiente de assimetria, expresso na Tabela 1, igual a 0,214. Quanto ao grau de achatamento, o coeficiente de curtose (2,344) indica que a distribuição é ligeiramente platicúrtica (Costa Neto, 1977, p. 32). Dentre outros valores apresentados na Tabela 1, nota-se que a média e a mediana, medidas que procuram caracterizar o centro da mesma distribuição de frequências, possuem valores próximos (33,035 e 33,0), respectivamente. Assim sendo, embora a distribuição da variável em estudo seja aproximadamente normal, supôs-se a hipótese de normalidade dos dados.

A análise estrutural do teor de argila é realizada com o auxílio do semivariograma. Esta é uma das etapas mais importantes, pois o modelo de semivariograma escolhido representa a estrutura de correlação espacial a ser utilizada nos procedimentos inferenciais de Krigagem Ordinária. Os resultados obtidos da análise dos semivariogramas direcionais estão apresentados na Figura 7.

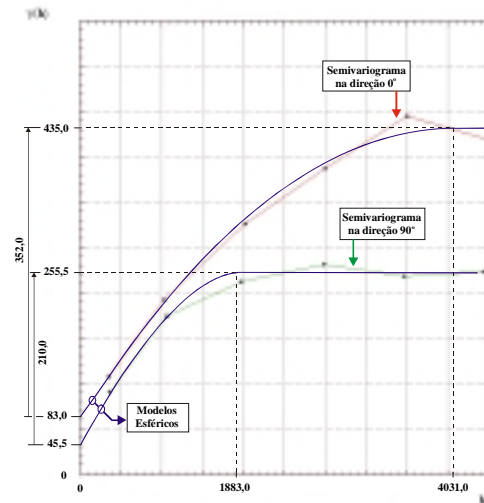


Fig. 7 - Semivariogramas direcionais e omnidirecional do teor de argila.

Conforme ilustra a Figura 7, os semivariogramas que apresentaram maior e menor alcance, foram detectados nas direções 0° e 90°, respectivamente. O semivariograma relativo à direção 0° possui um alcance próximo a 4000m, enquanto o semivariograma relativo à direção 90° possui um alcance próximo a 2000m. O semivariograma com maior patamar foi ajustado com um efeito pepita de (83,0), uma estrutura esférica com contribuição (352,0) e alcance (4.031,0). De maneira análoga, o semivariograma na direção 90° foi ajustado com efeito pepita (45,5), uma estrutura esférica com contribuição (210,0) e alcance (1.883,0).

A anisotropia presente é do tipo zonal combinada, esta deve ser considerada e modelada, a fim de retratar a verdadeira continuidade espacial do fenômeno. Um método para combinar os modelos direcionais num único modelo consistente para todas as direções é apresentado por Isaaks e Srivastava (1989, p. 377-386). O modelo único, é expresso através da seguinte equação:

$$\begin{aligned} \gamma(\mathbf{h}) = & 45,5 + 210,0 \left[\text{Sph} \left(\sqrt{\left(\frac{h_0^0}{4031,0} \right)^2 + \left(\frac{h_{90}^0}{1883,0} \right)^2} \right) \right] + 142,0 \left[\text{Sph} \left(\sqrt{\left(\frac{h_0^0}{4031,0} \right)^2 + \left(\frac{h_{90}^0}{\infty} \right)^2} \right) \right] + \\ & + 37,5 \left[\text{Sph} \left(\sqrt{\left(\frac{h_0^0}{\epsilon} \right)^2 + \left(\frac{h_{90}^0}{\infty} \right)^2} \right) \right] . \end{aligned} \quad (12)$$

O modelo completo, apresentado na equação (12), é composto de quatro termos, a saber:

- O primeiro termo referente ao efeito pepita (C_0), igual a 45,5.
- O segundo termo referente à 1ª componente estrutural, um modelo esférico com contribuição 210, alcance 4031,0 na direção Norte (0°) e alcance 1883,0 na direção Leste (90°).
- O terceiro termo referente à 2ª componente estrutural, um modelo esférico com contribuição 142,0 e alcance 4031,0 relativo à direção Norte (0°).
- Por último, o quarto termo referente à 3ª componente estrutural, é um modelo esférico com contribuição 37,5 e alcance muito pequeno, ϵ , relativo à direção Norte (0°). Ainda com relação ao quarto termo da equação (12), este é um artifício empregado para representar o segundo efeito pepita.

Uma vez definido o modelo que descreve a continuidade espacial do teor de argila., torna-se necessário realizar o processo de validação do modelo proposto. Esta é uma etapa que precede a Krigagem, e seu principal objetivo é avaliar a adequação do modelo sugerido no processo que envolve a re-estimação dos valores amostrais conhecidos. Os resultados obtidos da validação como: estatísticas e histograma do erro, diagrama espacial do erro e gráfico de valores verdadeiros versus estimados, foram cuidadosamente analisados e o modelo proposto na equação (12) foi dado como aceito.

Como resultado final da integração Geoestatística – SIG, as Figuras 8(a) e 8(b) apresentam os resultados obtidos relativos à técnica de Krigagem Ordinária. A figura 8(a) refere-se à continuidade espacial do teor de argila dentro dos limites da Fazenda Canchim. A Figura 8(b) refere-se à Variância de Krigagem, a qual pode ser utilizada como um indicador para as regiões em que a amostragem deve ser melhorada. Observe na Figura 8(b) que próximo dos pontos amostrais o valor da Variância de Krigagem é baixo e aumenta à medida em que se afasta dos mesmos.

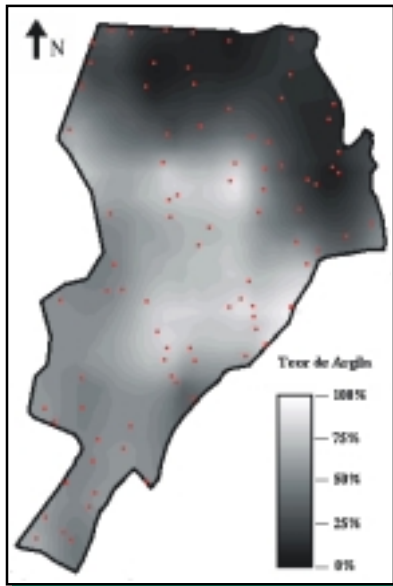


Fig. 8(a)- Resultado da Krigagem Ordinária

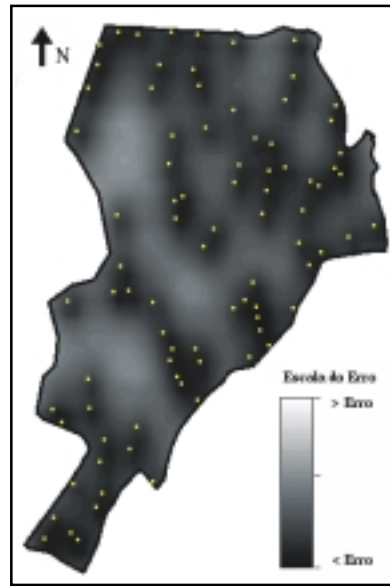


Fig. 8(b)- Resultado da Variância de Krigagem

O resultado obtido é expressivo e coerente. Isto pode ser verificado comparando a variabilidade espacial do teor de argila, mostrada na Figura 8(a), com o mapa geológico da região de estudo conforme Figura 9.



Fig. 9 – Mapa de Geologia da Fazenda Canchim

Dentro dos limites da Fazenda Canchim, afloram as seguintes litologias:

- **Arenito Superficial** predomina sobre a região norte-nordeste. Neste tipo de litogia encontram-se os solos com baixo teor de argila e os solos arenosos. Observe então a variabilidade espacial do teor de argila expresso na Figura 8(a) e compare-a com o mapa geológico.
- **Formação Serra Geral (Diabásio)** está localizado na região central da Fazenda Canchim. Neste tipo de litologia encontram-se os solos muito argilosos. De maneira análoga à anterior, observe a variabilidade espacial expressa na Figura 8(a) e compare-a com a Figura 9.
- **Arenito Botucatu** predomina sobre as demais regiões, conforme ilustra a Figura 9. Aqui encontram-se os solos com teor médio de argila.

Seguindo apresenta-se, de forma comparativa, os resultados obtidos da krigagem com outros métodos de estimação, como: média ponderada pelo inverso do quadrado da distância, média simples e vizinho mais próximo. Não é

o objetivo aqui detalhar cada um destes métodos, pois os mesmos são descritos na literatura, como, por exemplo, em Burrough (1987).

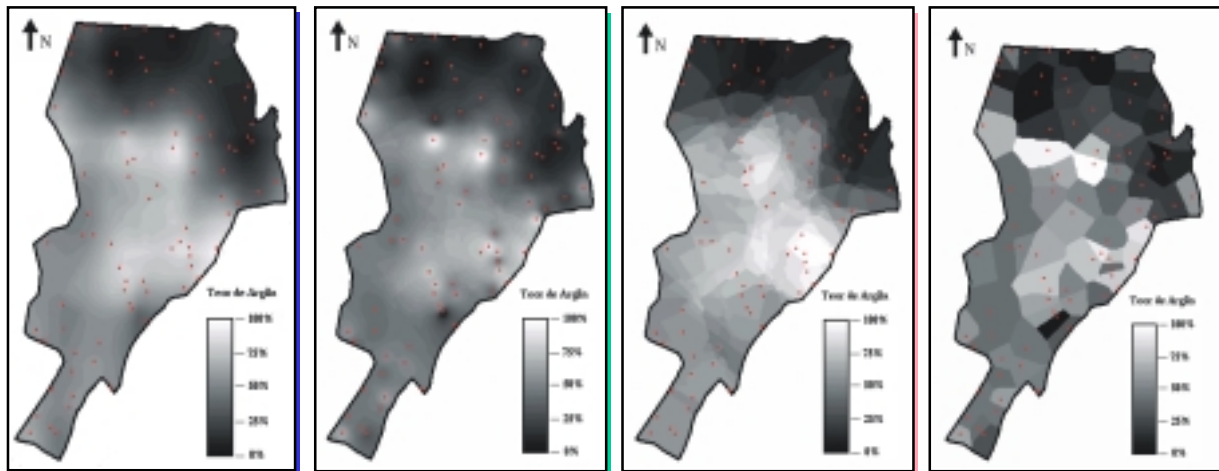


Fig.10- Krigagem Fig.11- Inverso Distância Fig.12- Média Simples Fig.13 – Vizinho + Próximo

Através das imagens apresentadas nas Figuras 10 a 13, pode-se constatar alguns fatos. Com algumas ressalvas, o método da Média Ponderada pelo Inverso do Quadrado da Distância, Figura 11, produz resultado que se assemelha ao resultado da Krigagem, Figura 10. O ponto crítico, porém, ocorre em regiões onde há superestimação e subestimação do teor de argila, o que difere do resultado oriundo da krigagem, onde as mudanças de valores do teor de argila são graduais. Os resultados produzidos pelos métodos Média Simples, Figura 12, e Vizinho mais Próximo, Figura 13, são menos expressivos com relação aos demais. O método da Média Simples produz resultado que apresenta imbricação, principalmente na região central da área de estudo. Já o método de inferência relativo ao Vizinho mais Próximo, embora sendo o que pior expressa a variabilidade espacial do fenômeno estudado, revela a área de influência de cada ponto de observação. Tal informação é de grande valia, como, por exemplo, numa análise preliminar para detecção de valores amostrais suspeitos.

Para finalizar, a Tabela 2 apresenta de forma sintetizada as principais razões pelas quais os resultados obtidos pela krigagem produziram resultados mais significativos em relação aos outros métodos de inferência.

TABELA 2 - CARACTERÍSTICAS DOS MÉTODOS DE INFERÊNCIA.

Métodos Geoestatísticos	Métodos Convencionais
<ul style="list-style-type: none"> Os pesos são determinados a partir de uma análise de correlação espacial baseada no semivariograma. $\lambda_i = f[\gamma(\mathbf{h})]$ 	<ul style="list-style-type: none"> Os pesos são determinados meramente em função da distância. $\lambda_i = f(d_i)$
<ul style="list-style-type: none"> Área de influência na interpolação é indicada pelo alcance. 	<ul style="list-style-type: none"> Raio de busca é arbitrário.
<ul style="list-style-type: none"> Modela anisotropia, isto é, detecta as direções de maior e menor continuidade espacial do fenômeno. 	<ul style="list-style-type: none"> Anisotropia é ignorada.
<ul style="list-style-type: none"> Trata redundância (“Clusters”), isto é, atribui pesos adequados para agrupamentos de amostras. 	<ul style="list-style-type: none"> Redundância é ignorada. Neste caso, podem ocorrer superestimação ou subestimação de valores.

CONCLUSÃO

A qualidade dos mapas oriundos da krigagem (Figuras 10 e 11), revelou a boa adequação do modelo empregado. Os resultados foram de certa forma expressivos quando comparados ao mapa geológico da região e a outros métodos de inferência. O sucesso dos resultados obtidos se deve principalmente às características intrínsecas do interpolador de krigagem (Tabela 2). Ficou constatado, por exemplo, que o teor de argila varia mais intensamente numa direção do que em outra. Tal fato refere-se à anisotropia do fenômeno, a qual é considerada na krigagem e ignorada nos métodos tradicionais de inferência. O resultado disto é que modelos provenientes da krigagem proporcionam resultados mais próximos à realidade.

Desta forma, conclui-se que é possível melhorar a análise de dados espaciais significativamente quando a geoestatística é aplicada. Muitos aspectos particulares dos dados ficariam ocultos se não fosse o uso de semivariogramas, mostrando, por exemplo, a tendência da variabilidade espacial nos dados de teor de argila. Informações como estas não são apresentadas quando se usam apenas parâmetros clássicos como médias e variâncias.

Um outro aspecto que deve ser ressaltado é a utilização da potencialidade do SIG na apresentação e análise dos resultados. Neste sentido, a integração geoestatística e SIG mostrou-se extremamente conveniente e deve ser adotada sempre que possível como rotina em análises de dados espaciais, possibilitando métodos inferenciais com controle mais significativos sobre os parâmetros do interpolador usado.

REFERÊNCIAS BIBLIOGRÁFICAS

Burrough, P. A. **Principles of geographical information systems for land resources assessment**. Oxford, Clarendon Press, 1987. 193p.

Burrough, P. A.; Varekamp C.; Skidmore, A.K. Using Public Domain Geostatistical and GIS Software for Spatial Interpolation. **Photogrammetric Engineering & Remote Sensing**, 62 (7): 845-854, 1996.

Calderano Filho, B.; Fonseca, O. O. M.; Santos, H. G.; Lemos A. L. **Levantamento Semidetalhado dos Solos da Fazenda Canchim São Carlos - SP**. Rio de Janeiro, EMBRAPA- CNPS, 1996. 261p.

Costa Neto, P. L. O. **Estatística**. São Paulo, Edgard Blücher, 1977. 264p.

David, M. **Geostatistical ore reserve estimation**. New York, Elsevier Scientific, 1977. 364p.

Deutsch, C.V.; Journel, A. G. **GSLIB: Geostatistical Software Library and user's guide**. New York, Oxford University Press, 1992. 339p.

Instituto Nacional de Pesquisas Espaciais. Departamento de Processamento de Imagens. (INPE/DPI). **Sistema de Processamento de Informações Georreferenciadas (SPRING)**. <http://www.inpe.br/spring>. Março, 1977.

Isaaks, E. H.; Srivastava R. M. **An Introduction to Applied Geostatistics**. New York, Oxford University Press, 1989. 560p.

Journel, A.G. **Fundamentals of geostatistics in five lessons**. California, Stanford Center for Reservoir Forecasting Applied Earth Sciences Department, 1988.

Krige, D.G. A statistical approach to some basic mine evaluation problems on the Witwatersrand. **Johannesburg Chemistry Metallurgy Mining Society South African**, 52 (6): 119-139, 1951.

Matheron, G. **The theory of regionalized variables and its applications**. Paris, Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau, 1971. 211p.

Matheron, G. Principles of geostatistics. **Economic Geology**, 58 (8): 1246-1266, Dec. 1963.