



Modelagem de Atributos Espaciais por Procedimentos Geoestatísticos por Indicação

Carlos Alberto Felgueiras



Organização da Apresentação

- 1. Considerações Iniciais***
- 2. A Krigagem por Indicação***
- 3. A Simulação por Indicação***
- 4. Considerações finais***
- 5. Aplicações em Geoprocessamento***



1. Considerações Iniciais

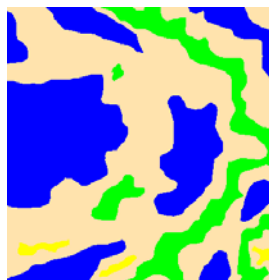


1. Motivação

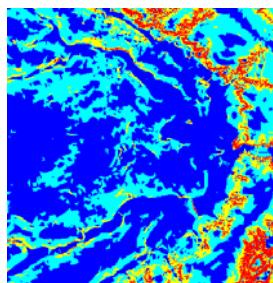
- Modelos de atributos espaciais são usados em estudos, simulação, análises e previsões sobre fenômenos geográficos (ambientais, por exemplo)
- Ferramentas de análise espacial integram representações de atributos espaciais em ambiente de SIGs
- Exemplos: Determinação de áreas de risco para moradias, geração de mapas de aptidão agrícola, distribuição de escolas e hospitais dentro de uma região, etc..

- *Análise espacial em SIG*

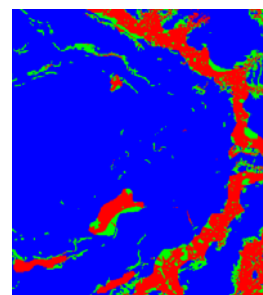
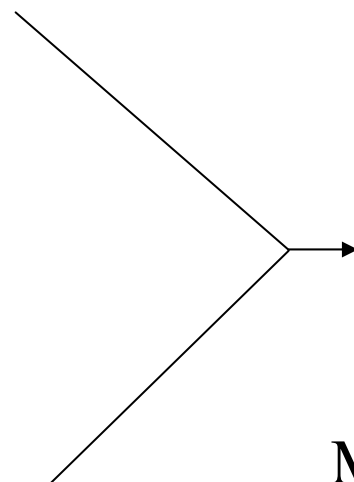
- Integram-se representações de atributos para se obter respostas, relatórios ou cenários que servirão de apoio a decisões sobre uma área em estudo



Mapa de Solos

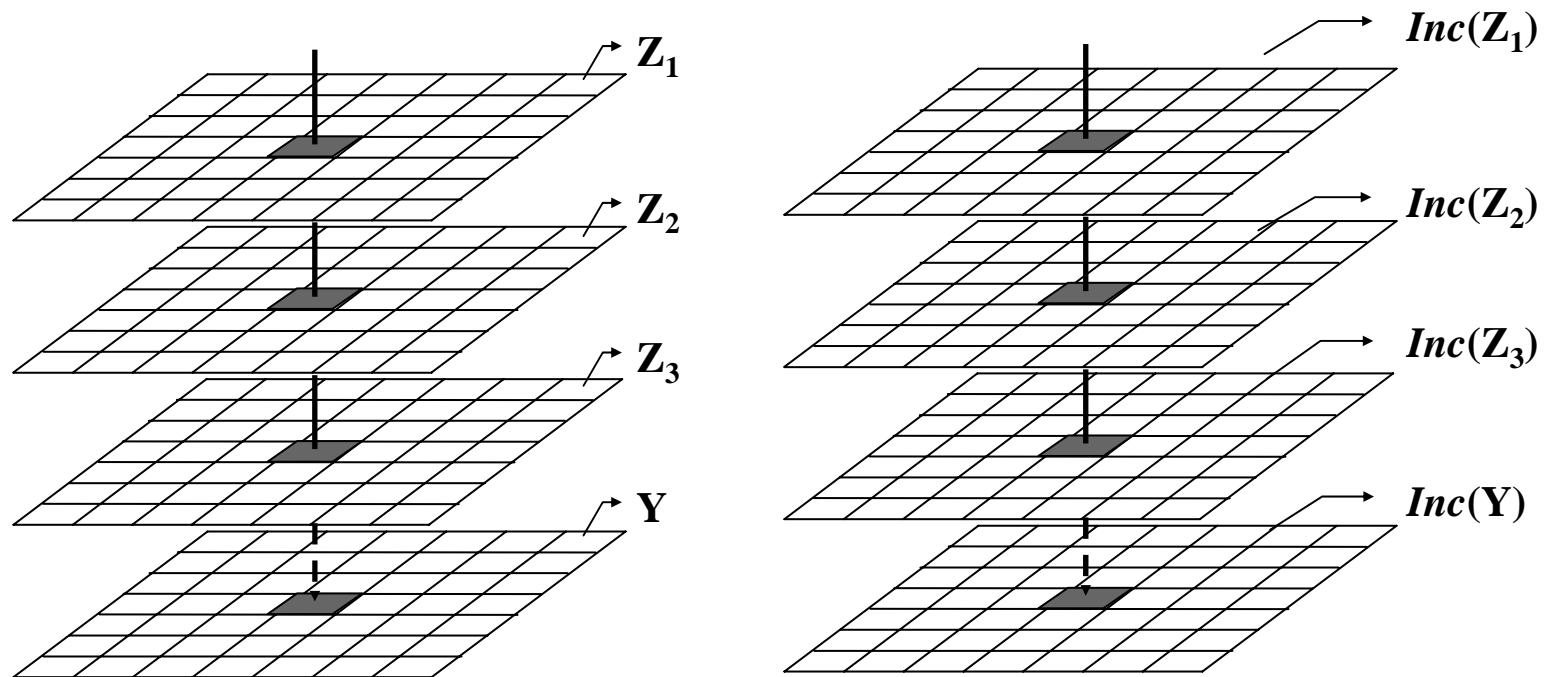


Mapa de Declividade



Mapa de Aptidão

- *Análise espacial em SIG*



Modelagem Espacial: $Y(\mathbf{u}) = g(Z_1(\mathbf{u}), \dots, Z_n(\mathbf{u}))$ para n entradas

As incertezas $Inc(Z_i)$ se propagam para a saída $Inc(Y)$



ESTIMADOR DE KRIGEAGEM LINEAR

(Eduardo C.G.Camargo)

- Segundo Journal (1988): $\mathbf{K} \cdot \boldsymbol{\lambda} = \mathbf{k} \Rightarrow \boldsymbol{\lambda} = \mathbf{K}^{-1} \mathbf{k}$

$$\begin{array}{c} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{array} = \begin{array}{c} C_{11} \ C_{12} \ \dots \ C_{1n} \ 1 \\ C_{21} \ C_{22} \ \dots \ C_{2n} \ 1 \\ \vdots \ \vdots \ \vdots \ \vdots \\ C_{n1} \ C_{n2} \ \dots \ C_{nn} \ 1 \\ 1 \ 1 \ \dots \ 1 \ 0 \end{array}^{-1} \begin{array}{c} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{array}$$

- Os elementos das matrizes de covariâncias são calculados da seguinte forma (Journal, 1988):

$$C_{ij} = C_0 + C_1 - \gamma(\mathbf{h})$$

- Substituindo os valores de C_{ij} nas matrizes encontram-se os pesos $\lambda_1, \lambda_2, \dots, \text{ e } \lambda_n$.

- Estimador de Krigagem (Journal, 1988): $Z_{\mathbf{x}_0}^* = \sum_{i=1}^n \lambda_i Z(\mathbf{x}_i)$

- Variância de Krigagem (Journal, 1988): $\sigma_{k_0}^2 = (C_0 + C_1) - \boldsymbol{\lambda}^T \mathbf{k}$

- Variância de krigagem como medida de incerteza

$$\sigma^2(\mathbf{u}) = C(0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \cdot \lambda_j \cdot C_{ij} - 2 \sum_{i=1}^n \lambda_i \cdot C_{iu}$$

- $C(0)$ é a variância das amostras. Quão errático é o atributo.
- C_{ij} é a covariância entre as amostras i e j . Considera aglomerados. Amostras próximas aumentam a variância final.
- C_{iu} é a covariância entre a amostra i e a posição \mathbf{u} do ponto a ser estimado. Considera a proximidade entre a amostra i e a posição \mathbf{u} . Amostras próximas de \mathbf{u} diminuem a variância final (sinal -).

Importante: Variâncias de krigagem, sendo independentes dos dados, fornecem uma comparação entre diferentes configurações geométricas de dados (Deutsch and Journel, 1998)

- *A medida de variância de uma função aleatória deve ser obtida diretamente de sua fdac multivariada.*
- O estimador de krigagem possibilita a determinação da fdac multivariada de uma FA por 2 métodos distintos: o *método multigaussiano* e o *método de krigagem por indicação*.
- ***Método Multigaussiano (paramétrico)***: Para modelos de FA FZ(u) gaussianos multivariados, a estimativa de krigagem simples e a sua variância determinam os parâmetros média e a variância da fdac Gaussiana. Este resultado é a base para os algoritmos de krigagem e simulação multigaussianas.

$$\mathbf{F}(z) = P(Z < z) = \int_{-\infty}^z \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz$$



2. A Krigagem por Indicação

- *Método da krigagem por indicação* (não paramétrico)

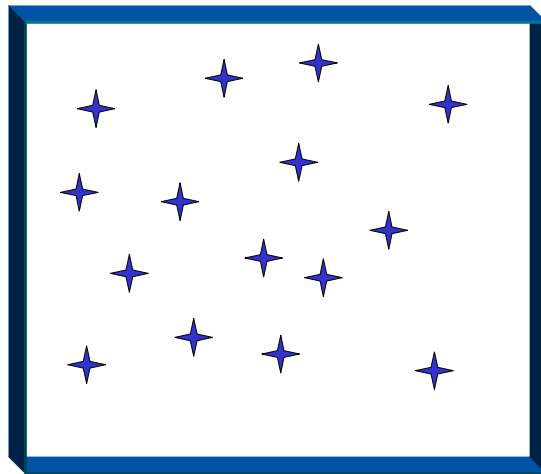
- Transformação por indicação (domínio probabilístico)

Para VA contínua

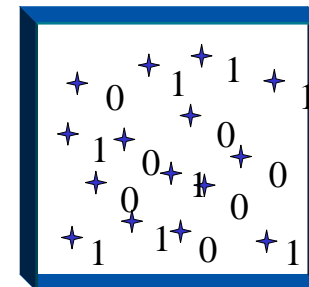
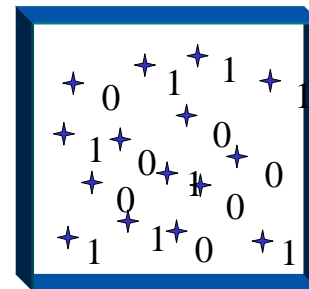
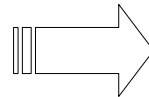
$$I(\mathbf{u}; z) = \begin{cases} 1, & \text{se } Z(\mathbf{u}) \leq z \\ 0, & \text{se } Z(\mathbf{u}) > z \end{cases}$$

Para VA categórica

$$I(\mathbf{u}; z) = \begin{cases} 1, & \text{se } Z(\mathbf{u}) = z \\ 0, & \text{se } Z(\mathbf{u}) \neq z \end{cases}$$

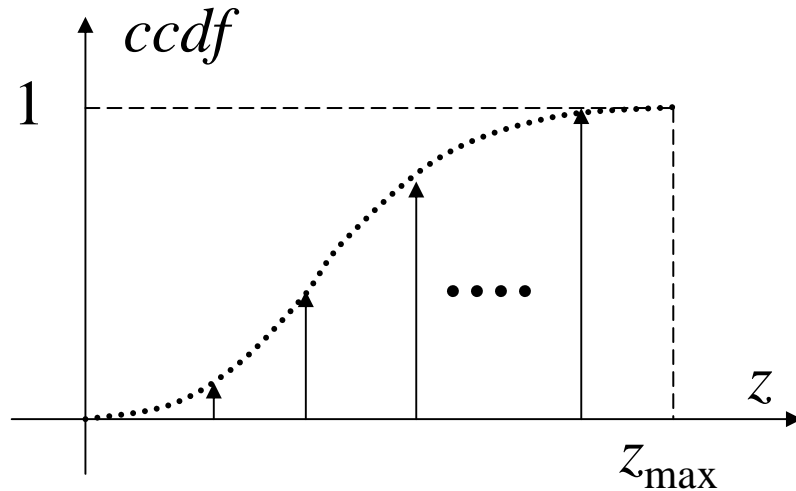


$I(\mathbf{u}_\alpha; z_k)$



A esperança condicional de $I(\mathbf{u}; z)$ é igual a fdac de $Z(\mathbf{u})$ como mostrado abaixo:

$$\begin{aligned} E\{I(\mathbf{u}; z) / (n)\} &= 1 \cdot \text{Prob}\{I(\mathbf{u}; z) = 1 / (n)\} + 0 \cdot \text{Prob}\{I(\mathbf{u}; z) = 0 / (n)\} \\ &= 1 \cdot \text{Prob}\{I(\mathbf{u}; z) = 1 / (n)\} = F(\mathbf{u}; z / (n)) \end{aligned}$$



O estimador de krigagem ordinária por indicação fornece

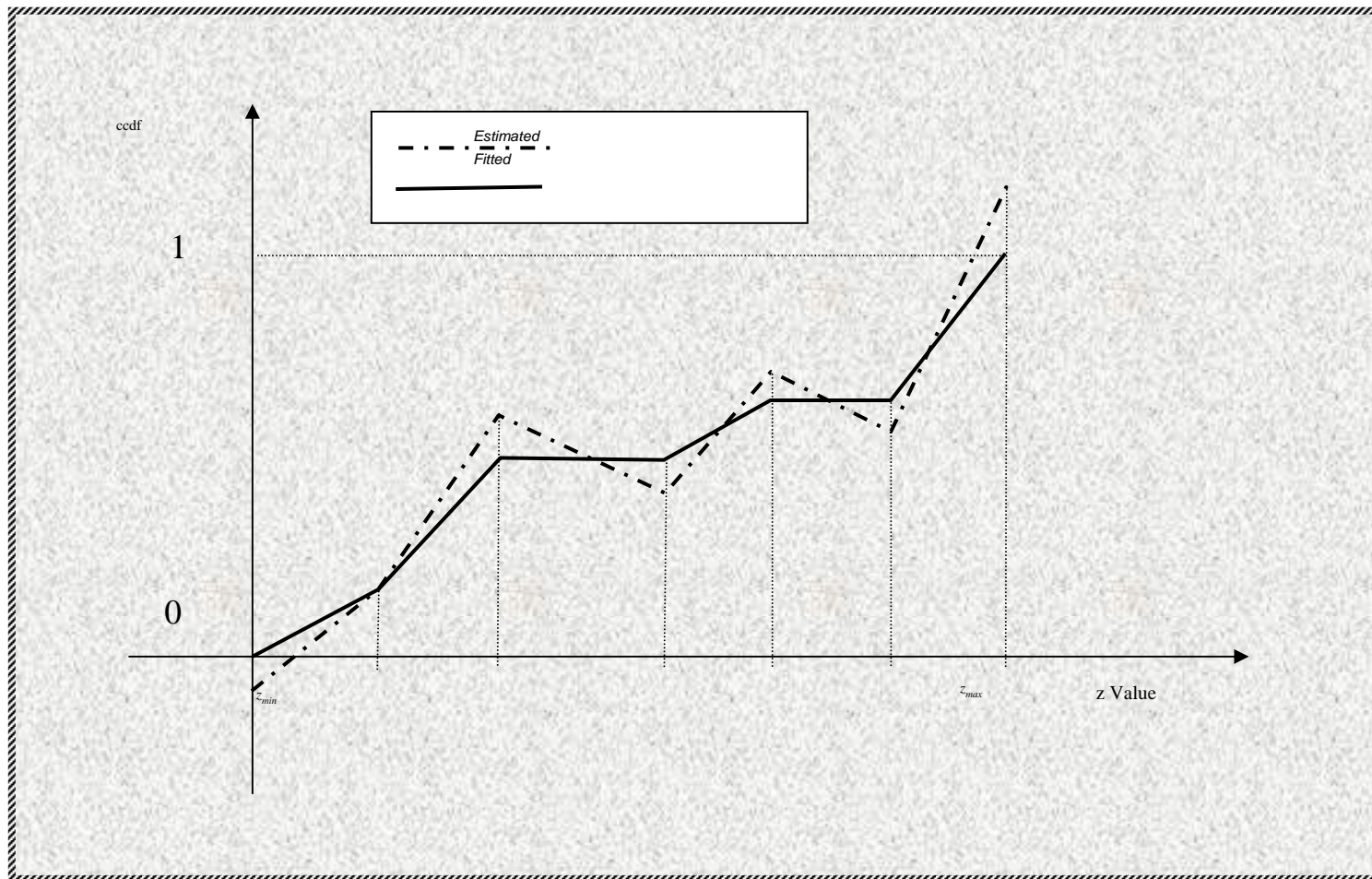
(para cada valor de corte $z=z_k$):

$$F_o^*(\mathbf{u}; z_k / (n)) = \sum_{\alpha=1}^n \lambda_o(\mathbf{u}; z_k) i(\mathbf{u}; z_k)$$

Os pesos λ_o são determinados pela solução do sistema :

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n \lambda_o(\mathbf{u}; z_k) C_I(\mathbf{h}_{\alpha\beta}; z_k) + \mu(\mathbf{u}; z_k) = C_I(\mathbf{h}_{\alpha}; z_k) \quad \forall \alpha = 1, 2, \dots, n \\ \sum_{\beta=1}^n \lambda_o(\mathbf{u}; z_k) = 1 \end{array} \right.$$

- Função de distribuição de probabilidade discretizada com correções de ordem:



- Pelo método da krigagem por indicação obtêm-se:
 - Para VA categórica \Rightarrow Valores de probabilidade de cada classe.
 - Para VA contínua \Rightarrow Valores de probabilidade acumulada em valores predefinidos da VA (“cutoffs”). Neste caso é necessário estimar-se a fdac em valores diferentes dos cutoffs.
 - O valor esperado e a variância de uma VA contínua são definidos como:

$$E[Z] = \int_{-\infty}^{\infty} z \cdot f(z) dz$$

$$Var[Z] = E[(Z - E(Z))^2] = \int_{-\infty}^{\infty} (z - E(Z))^2 \cdot f(z) dz$$

- Uma aproximação numérica para o valor esperado da fdac para K “cutoffs” é dado por (Journel, 1998):

$$[z(\mathbf{u})]_{\text{E}}^* = \int_{-\infty}^{\infty} z \cdot dF(\mathbf{u}; z / (n)) \approx \sum_{k=1}^{K+1} z'_k \cdot [F(\mathbf{u}; z_k / (n)) - F(\mathbf{u}; z_{k-1} / (n))]$$

- *Como fdac modela a incerteza E sobre os valores estimados ?*
- Para VA categórica: conhecido as probabilidades p_j de k classes

- Por *máxima probabilidade* de pertencer a uma categoria.

$$E(\mathbf{u}) = 1 - \text{Max}_{j=1}^k (p_j(\mathbf{u}))$$

- Por *entropia de Shannon*: medida de confusão

$$E(\mathbf{u}) = - \sum_{j=1}^k p_j(\mathbf{u}) \cdot \ln(p_j(\mathbf{u}))$$

- Outras medidas ??? (μ e σ fuzzy com pesos para as classes)

- Para VA continua

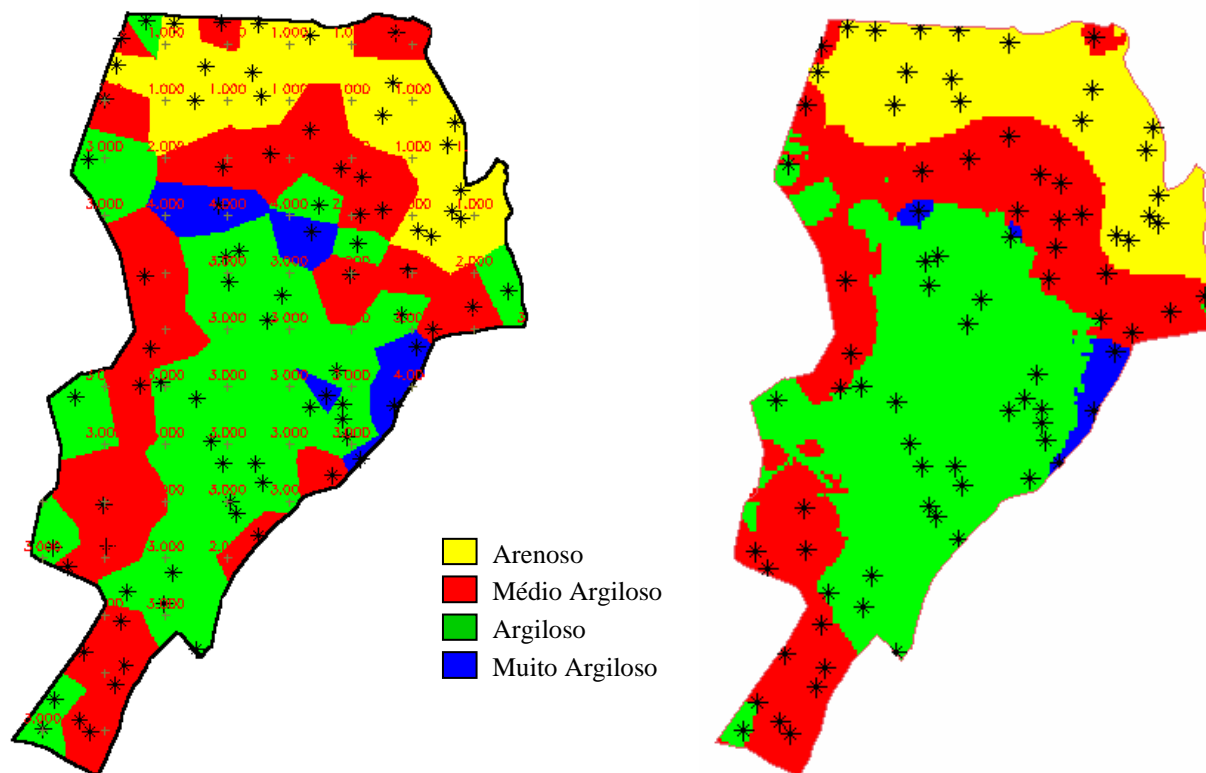
- A variância ou o desvio padrão é a medida de incerteza.

- Intervalos de confiança: $1\sigma = 68\%$; $2\sigma = 95\%$...

$$\text{Prob} \left\{ Z(\mathbf{u}) \in \left[z^*(\mathbf{u}) \pm 2\sigma(\mathbf{u}) \right] \right\} \cong 0.95$$

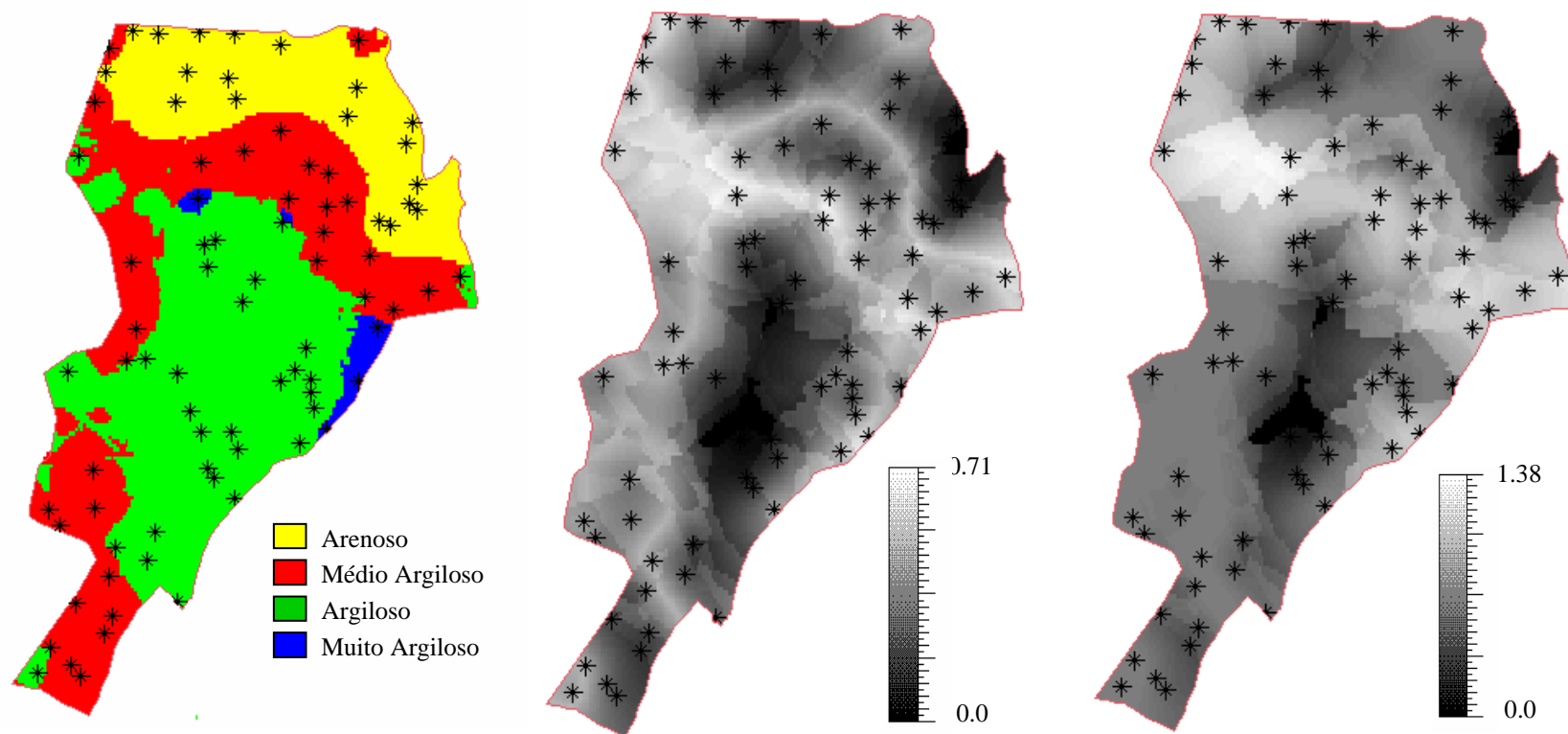
- Intervalos interquantis

- Exemplos - Modelagem de variáveis categóricas



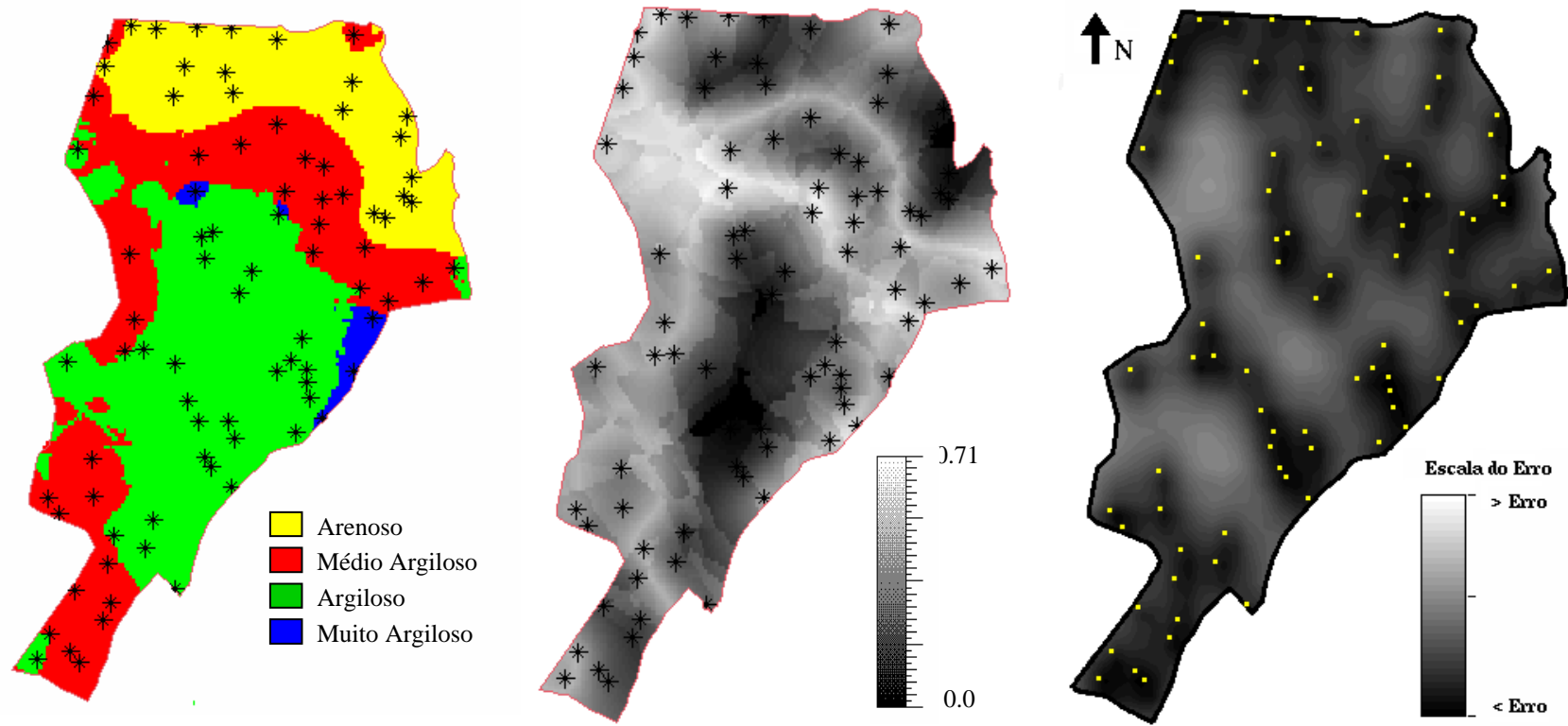
Figuras: Distribuição de um conjunto de amostras de textura de solo. Mapa com interpolação por vizinho mais próximo. Mapa classificado por máxima probabilidade (krigeagem por indicação).

- Variáveis categóricas



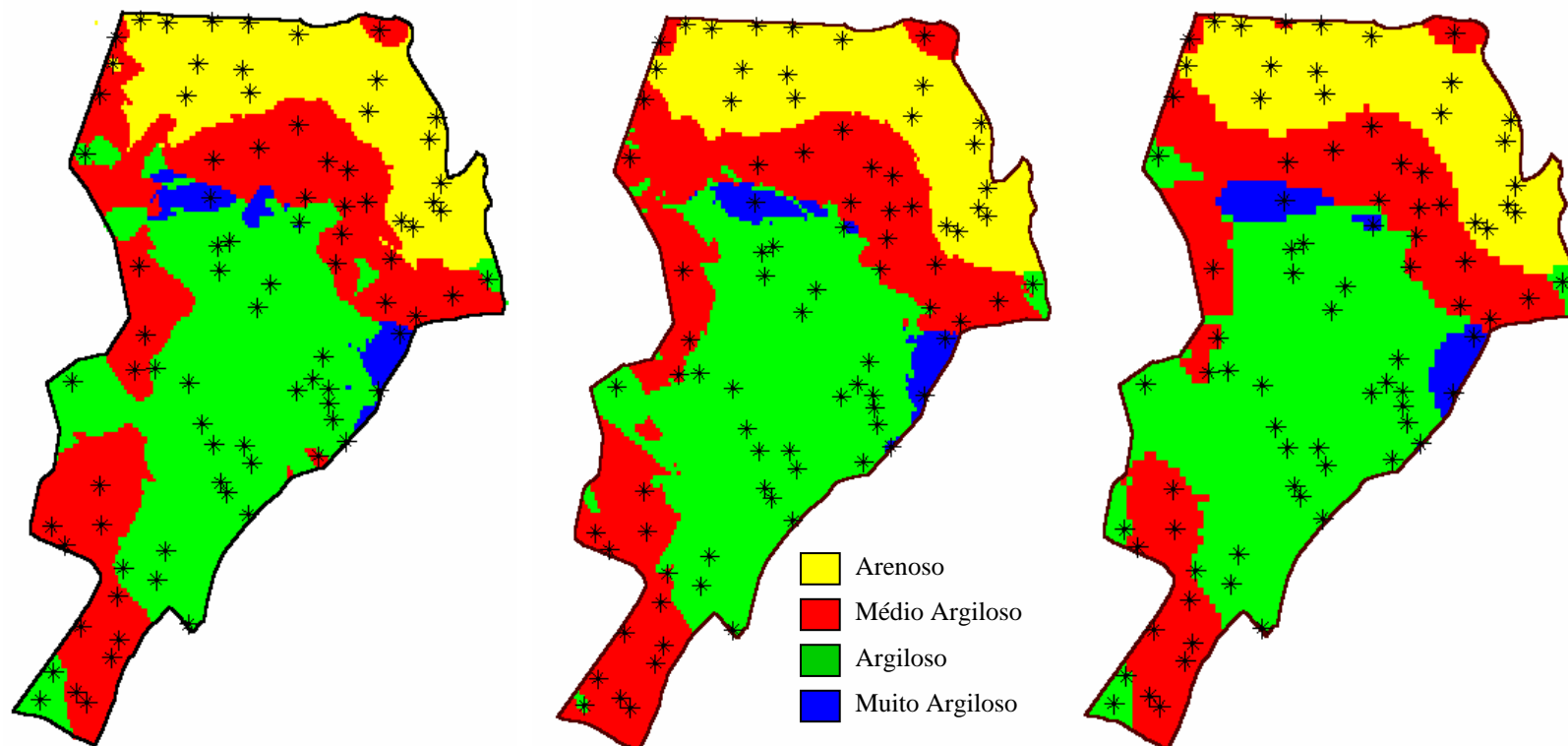
Figuras: Mapas de classe mais provável (esquerda) e de incerteza a partir das probabilidades definidas pela krigagem por indicação: por máxima probabilidade e por entropia de Shannon

- Variáveis categóricas



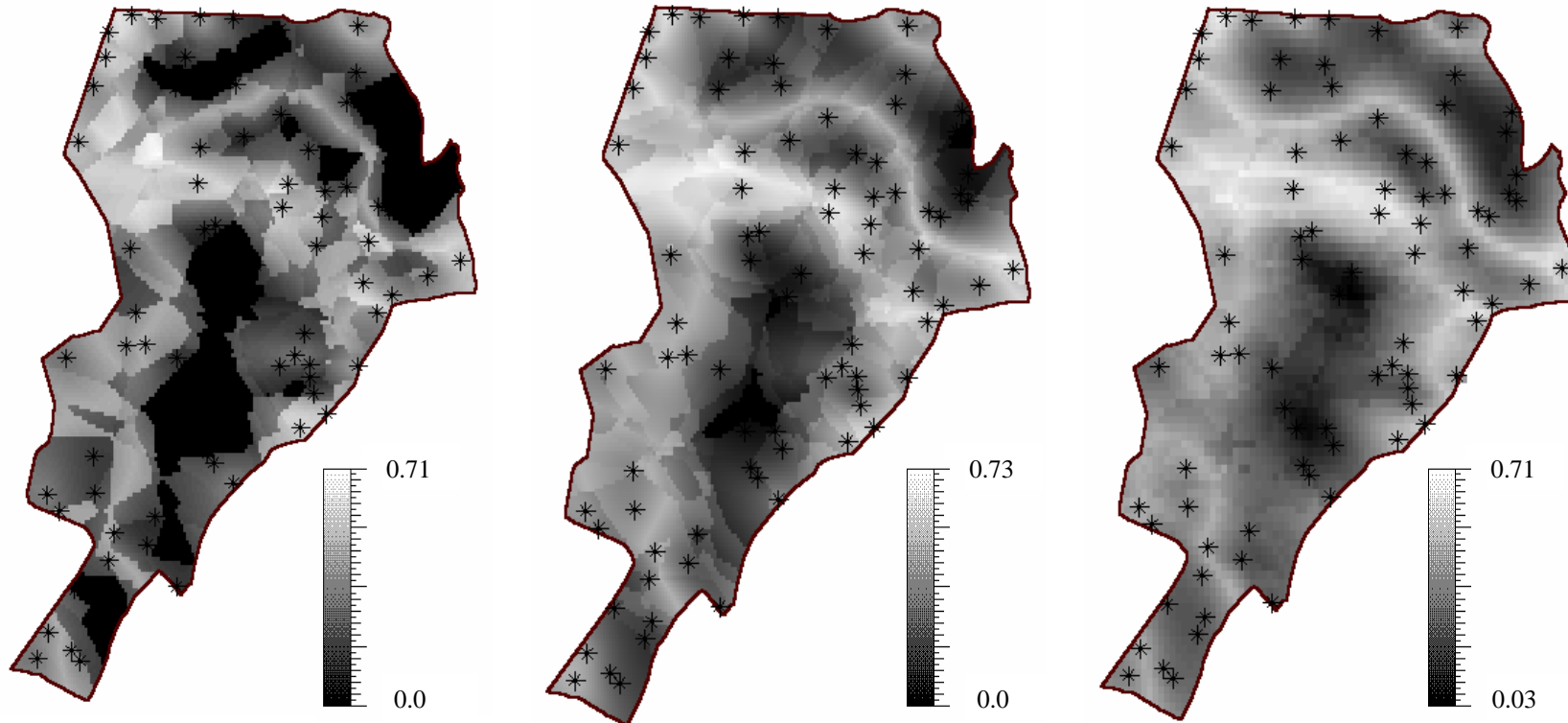
Figuras: Mapas de classe mais provável e de incerteza a partir das probabilidades definidas pela krigagem por indicação: por máxima probabilidade e por variância da krigagem

- Variáveis categóricas



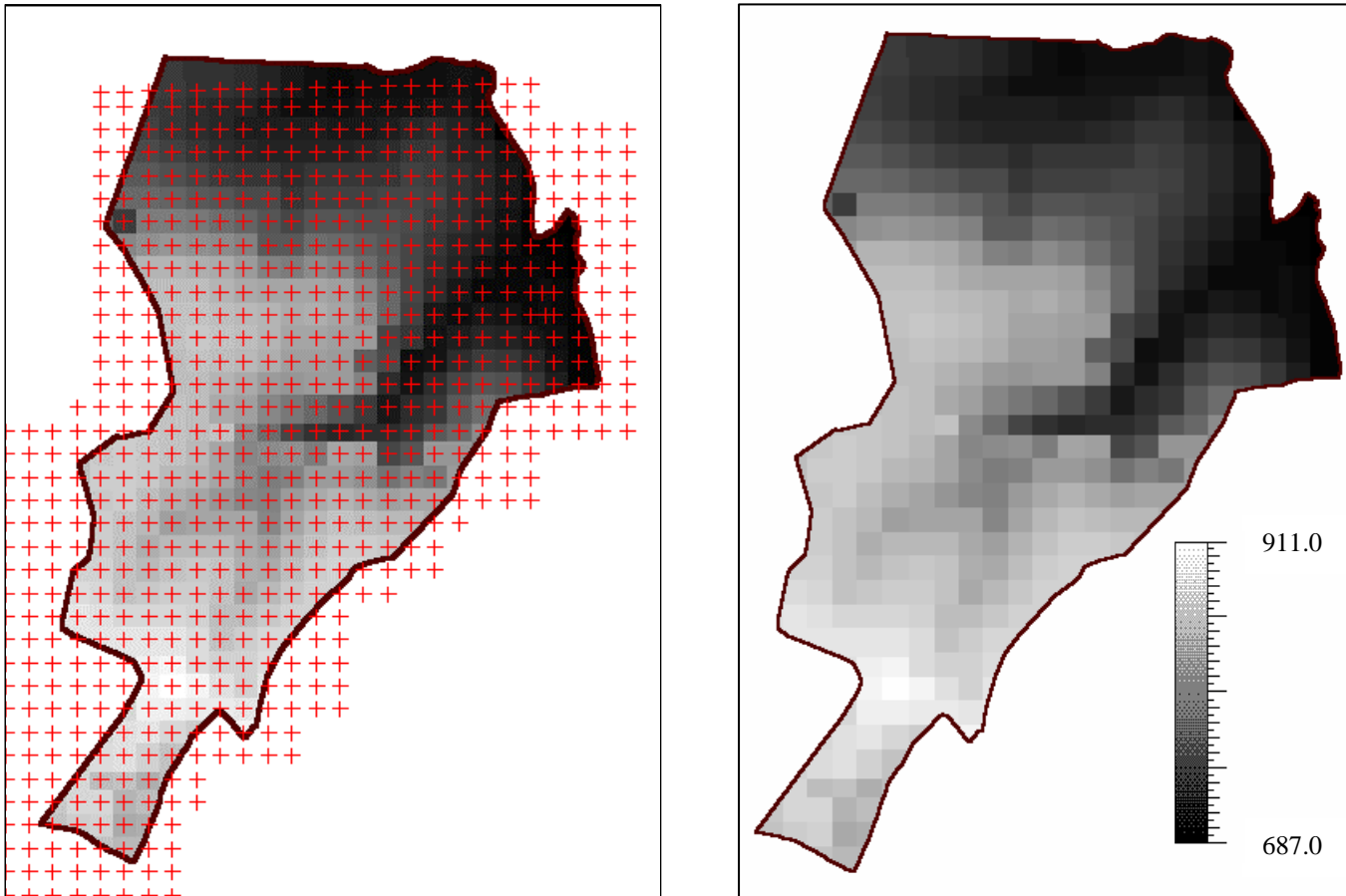
Figuras: Mapas de incerteza a partir das probabilidades definidas pela krigagem por indicação usando-se 4, 8 e 16 vizinhos mais próximos.

- Variáveis categóricas



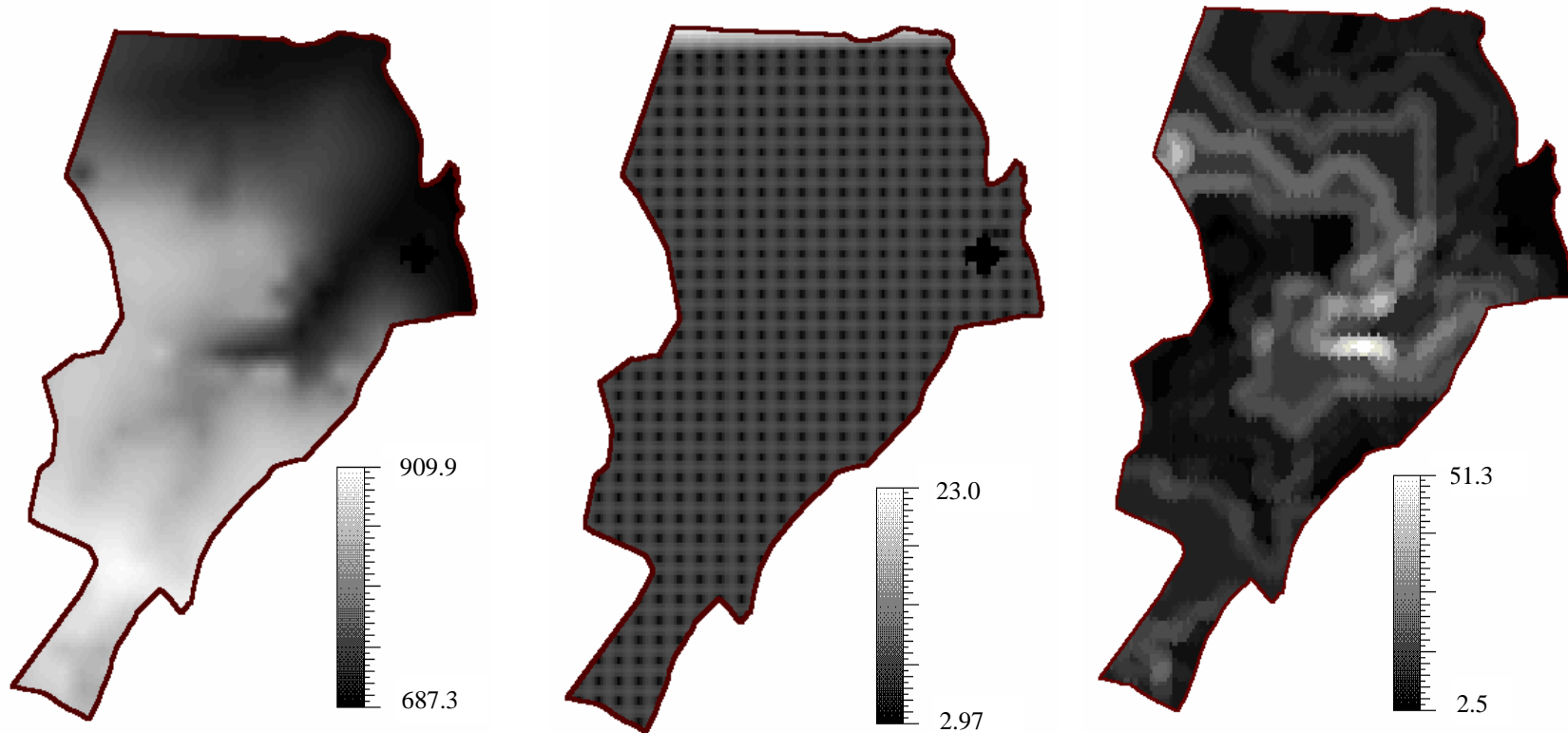
Figuras: Mapas de incerteza a partir das probabilidades definidas pela krigagem por indicação usando-se 4, 8 e 16 vizinhos mais próximos.

- Exemplos - Modelagem de variáveis contínuas



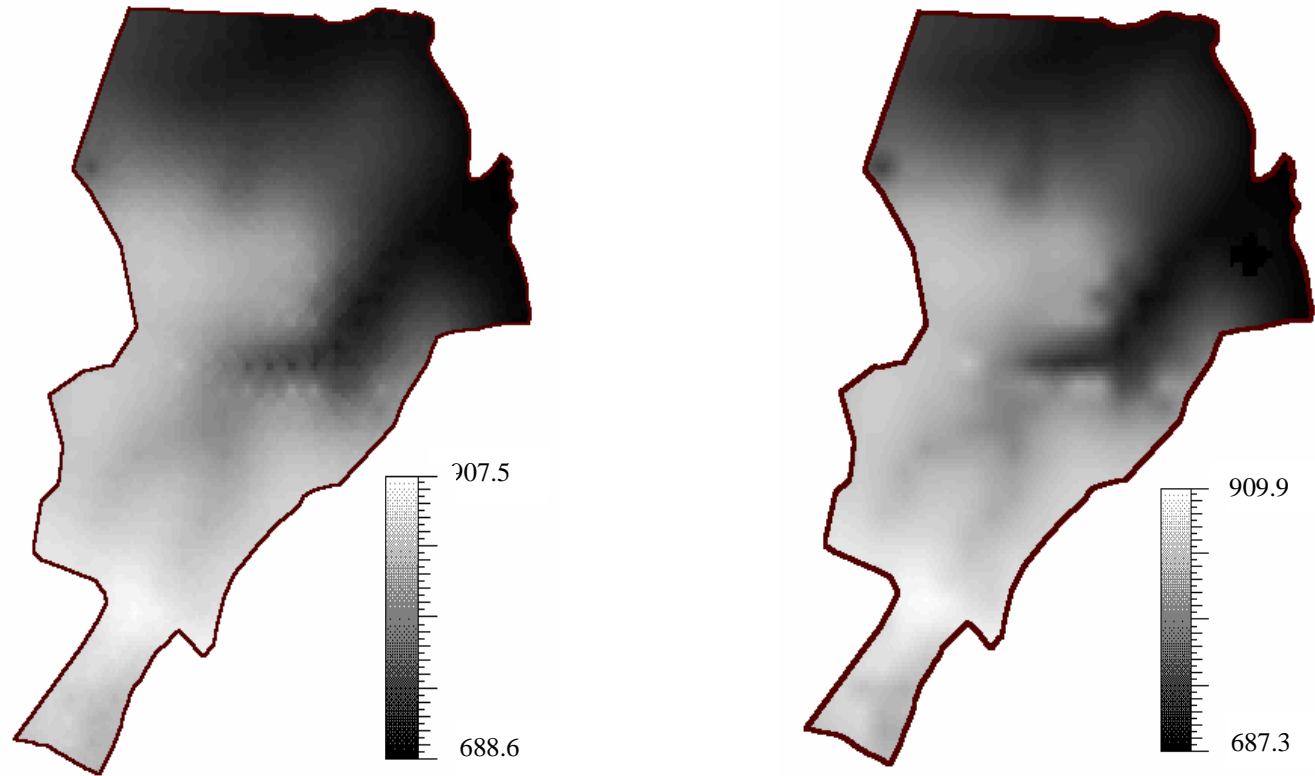
Figuras: Distribuição de amostras e interpolação por vizinho mais próximo.

- Variáveis contínuas



Figuras: Mapas de média (esquerda) e desvio padrão calculados a partir da variância de krigeagem ordinária e da fdac determinada pela krigeagem por indicação

- Variáveis contínuas



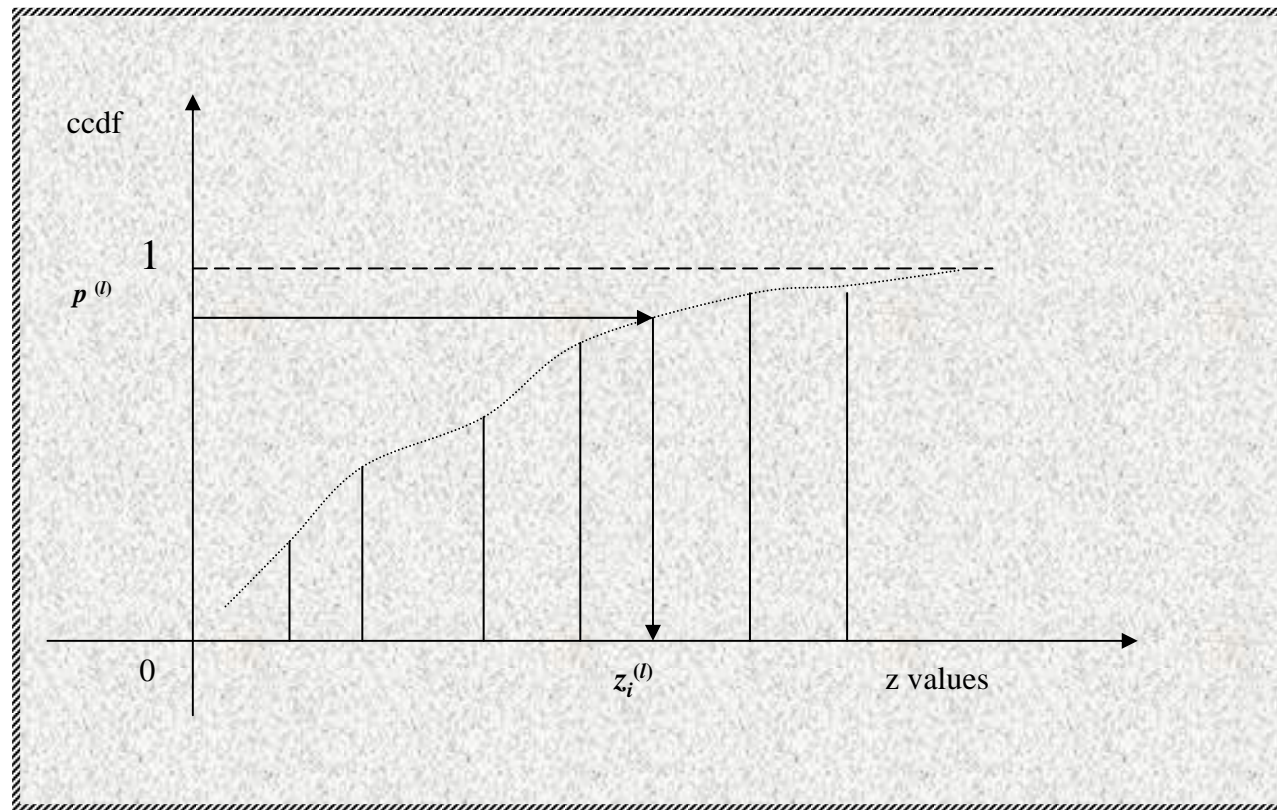
Figuras: Interpolação por média dos vizinhos por quadrante e por krigagem ordinária.



3. A Simulação por Indicação

Simulação estocástica

A partir da fdac e de um gerador de valores aleatórios (0 a 1) é possível se obter realizações da V. A.



Processo de obtenção de uma realização a partir da fdac univariada estimada

Simulação sequencial - Algoritmo

- Considerando-se a distribuição conjunta de N VAs Z_i condicionadas aos n dados originais, $|(n)$:

$$F_{(N)}(z_1, \dots, z_N | (n)) = \text{Prob}\{Z_i \leq z_i, i=1, \dots, N | (n)\}$$

- Obtenha um valor $z_1^{(l)}$ a partir da fdac univariada de Z_1 condicionada aos n dados originais: $\text{Prob}\{Z_1 \leq z_1 | (n)\}$. Atualize o conjunto de dados originais (n) para: $(n+1) = (n) \cup \{Z_1 = z_1^{(l)}\}$;
- Obtenha um valor $z_2^{(l)}$ a partir da fdac univariada de Z_2 condicionada aos $(n+1)$ dados: $\text{Prob}\{Z_2 \leq z_2 | (n+1)\}$. Atualize o conjunto de dados originais $(n+1)$ para: $(n+2) = (n+1) \cup \{Z_2 = z_2^{(l)}\}$;
- Sequencialmente considere todas as N VAs Z_i 's.
- Repita a sequência acima para novas realizações l 's

- *Simulação sequencial de variável categórica*

Importante: Deve-se definir uma ordenação qualquer das classes para construir a fdac

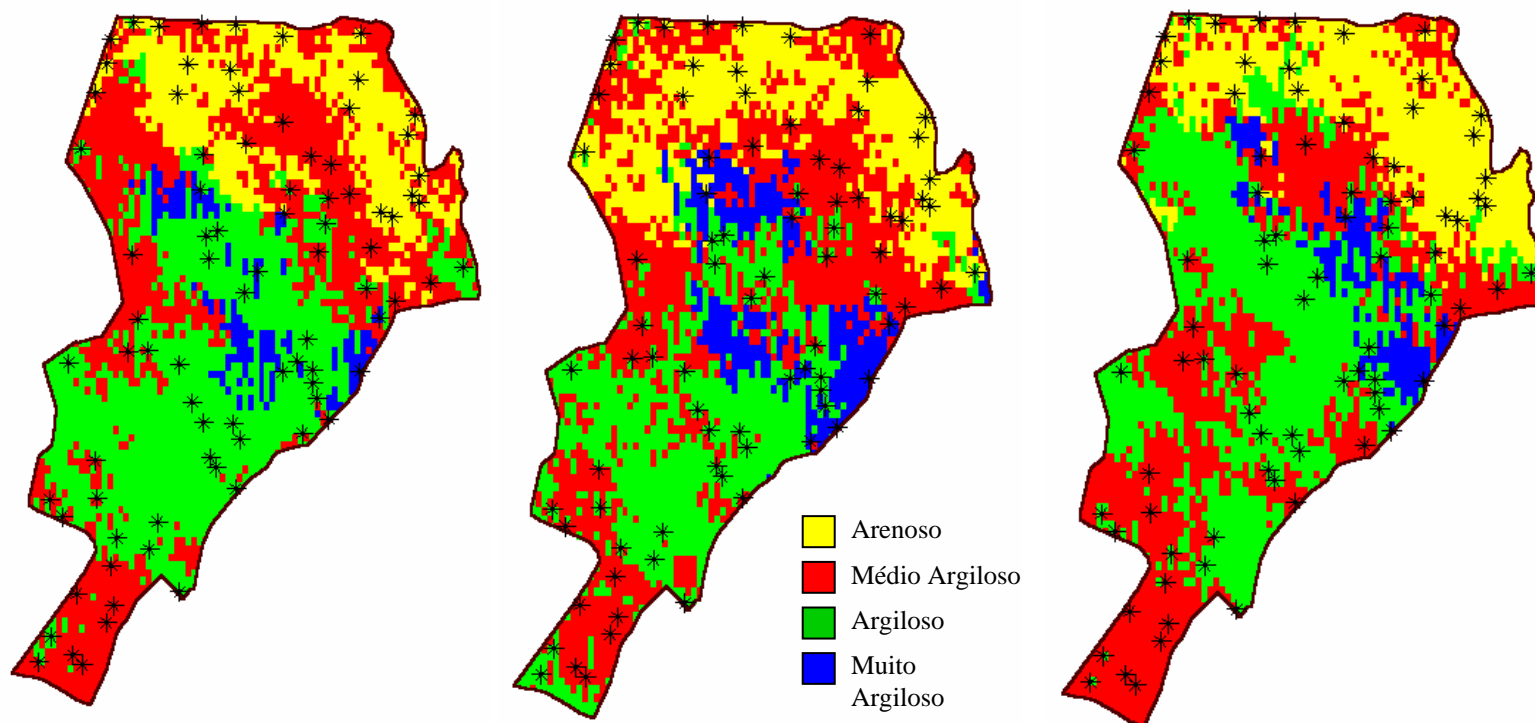


Figura: Exemplos de realizações da função aleatória definida para representar a variação anisotrópica de textura do solo da região de Canchim.

• *Simulação sequencial de variável categórica*

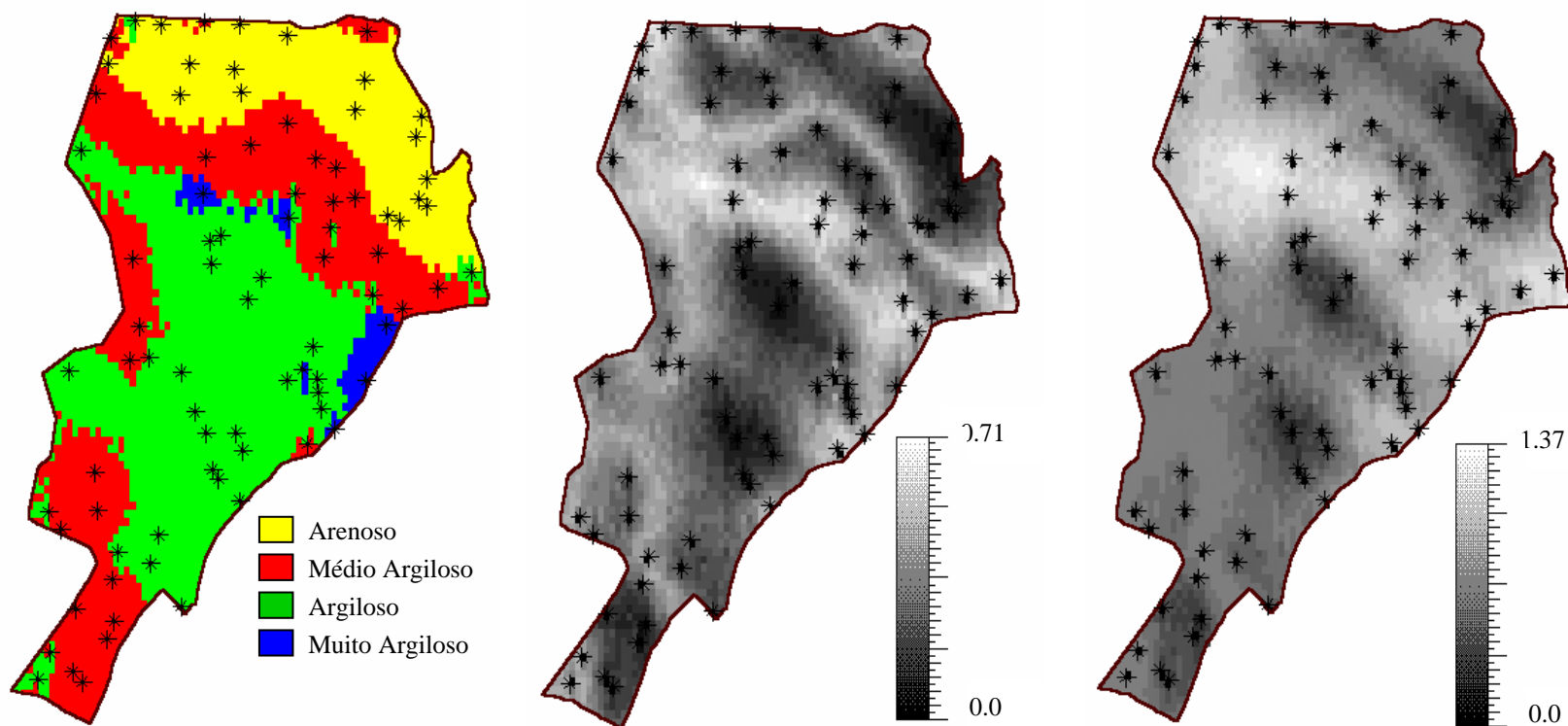


Figura: Mapas de classes de textura, de incerteza por máxima probabilidade e de incerteza pela entropia de Shannon gerados por simulação sequencial considerando comportamento anisotrópico da textura de solo

- *Simulação sequencial de variável contínua*

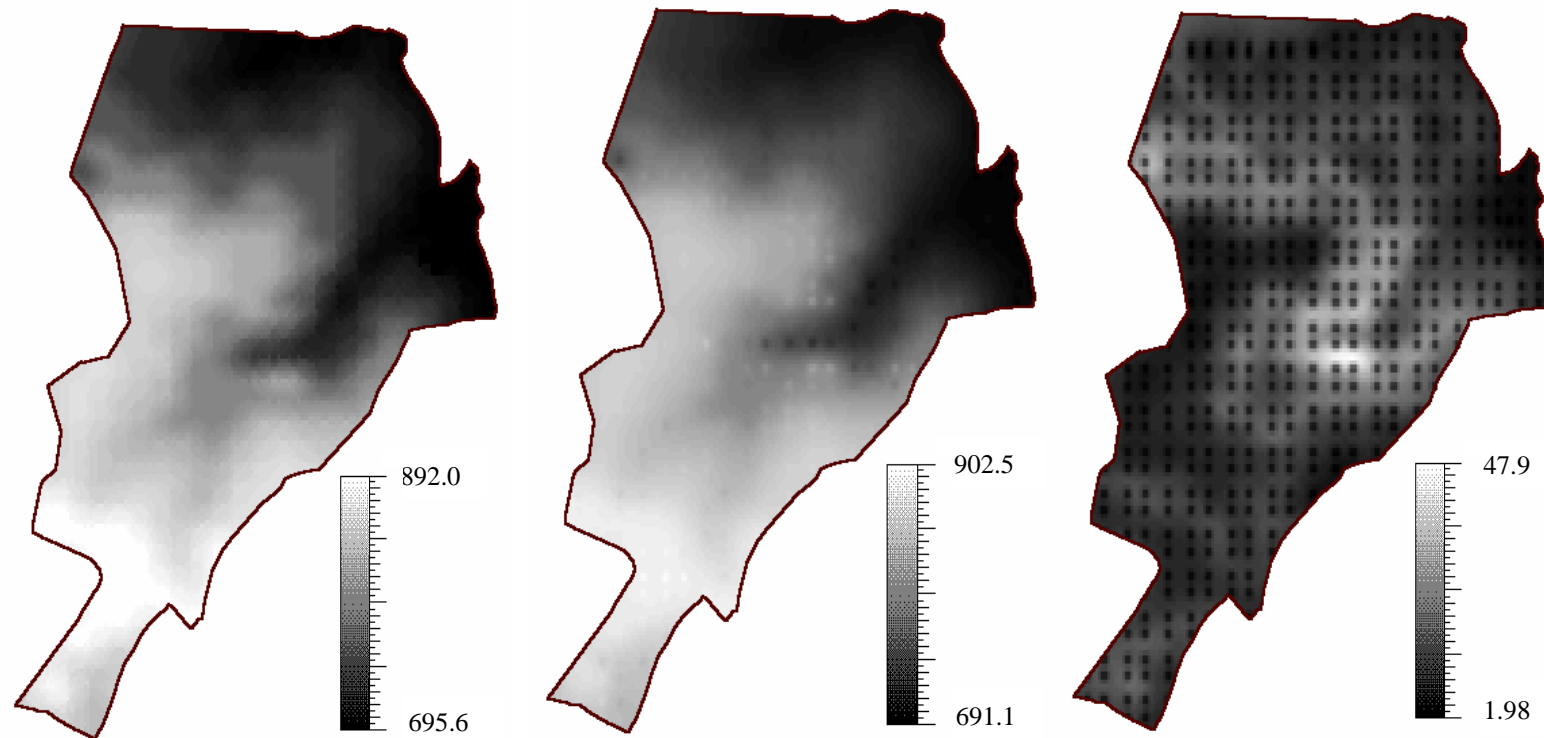


Figura:Imagens em nível de cinza referentes aos dados de altimetria da área de Canchim: (a) médias interpoladas por krigagem por indicação; (c) médias obtidas por simulação estocástica sequencial e; (d) desvios padrão das médias obtidas por simulação estocástica sequencial



4. Considerações Finais

- *Krigeagem x Simulação*

- Ambos baseiam-se no modelo conjunto de FA definido pelo estudo da variabilidade espacial do atributo (variografia). Também estão condicionados as amostras (estimadores exatos)
- A krigeagem cria uma realização em que a acurácia local, mínima variância de estimação, é o mais importante. Isto gera uma superfície suavizada, mascarando a variabilidade do dado.
- A simulação estocástica fornece representações globais alternativas $z^{(1)}(\mathbf{u})$, onde prevalece a representação de padrões de continuidade espacial.
- O resultado da média de um conjunto grande (>100) de imagens estocásticas é a realização de krigeagem do atributo.

• *Conclusões*

Procedimentos geoestatísticos por indicação:

- Desvantagem principal:

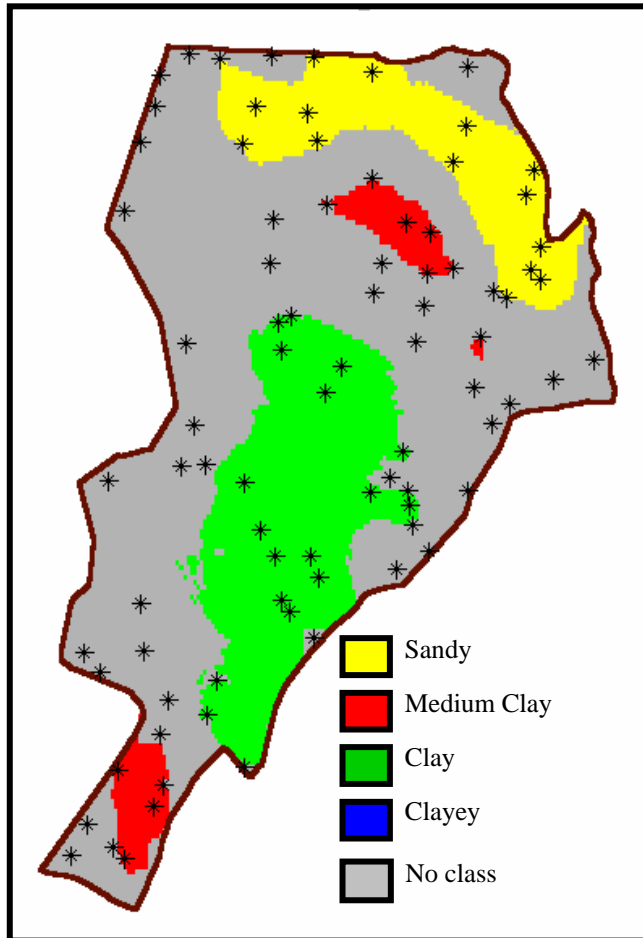
Modelagem de variogramas a priori para cada corte.

- Vantagens:
 - Possibilita usar diferentes estimadores e não só a média
 - Permite espacializar dados categóricos
 - Fornece informação de incerteza associada ao atributo. Essas incertezas podem ser propagadas para cenários de modelos espaciais.
 - Permite o uso de informação soft em conjunto com a informação hard
 - Incertezas são usadas para modelos de custos baseados em perdas de má-classificação. Estes são usados como apoio a decisões
- Pesquisas Atuais: uso de informação secundária, modelagem de custos (baseados em incertezas), validação dos modelos e modelos de propagação da incerteza

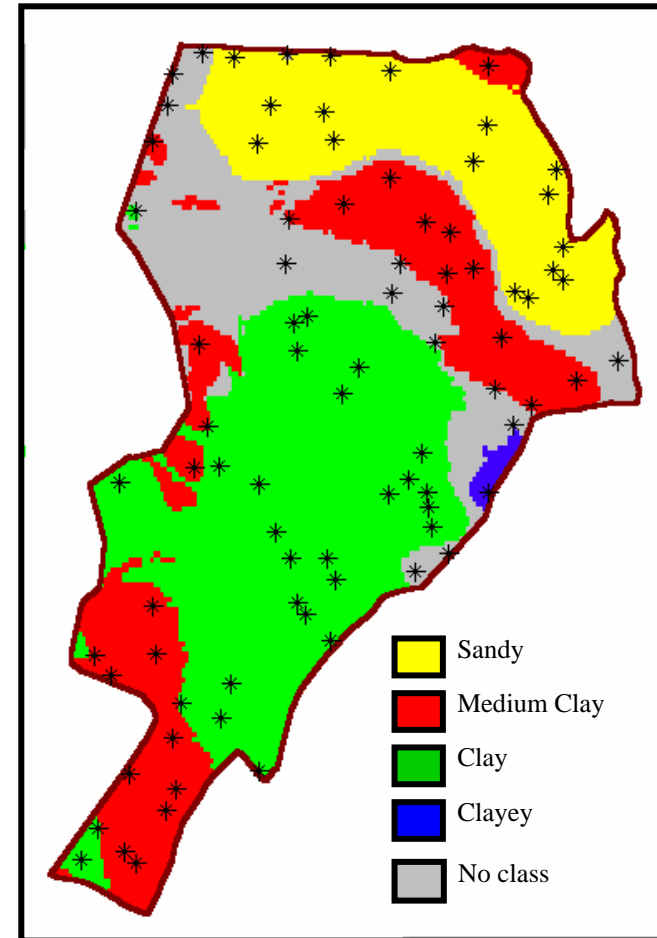


5. Exemplos de Aplicações em Geoprocessamento

Classificação com risco (a partir das funções de custo aplicadas sobre valores de incerteza)

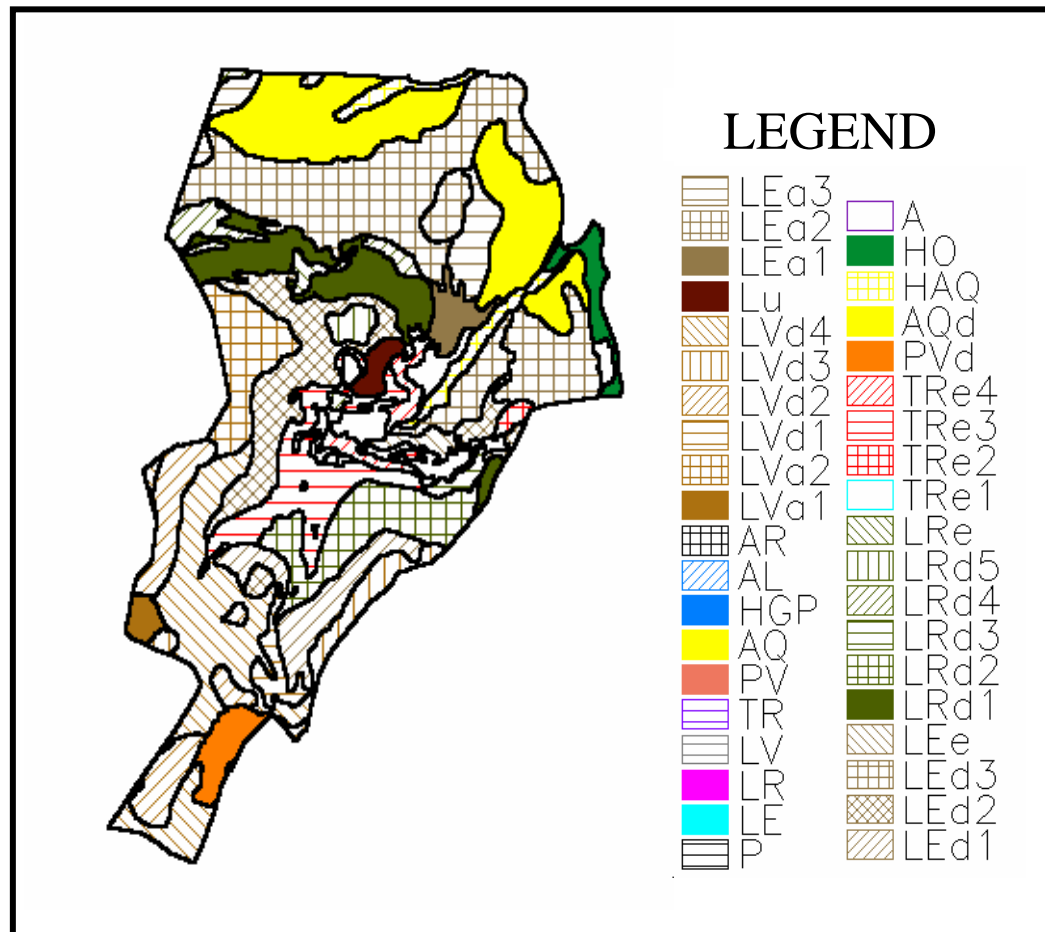


38 %

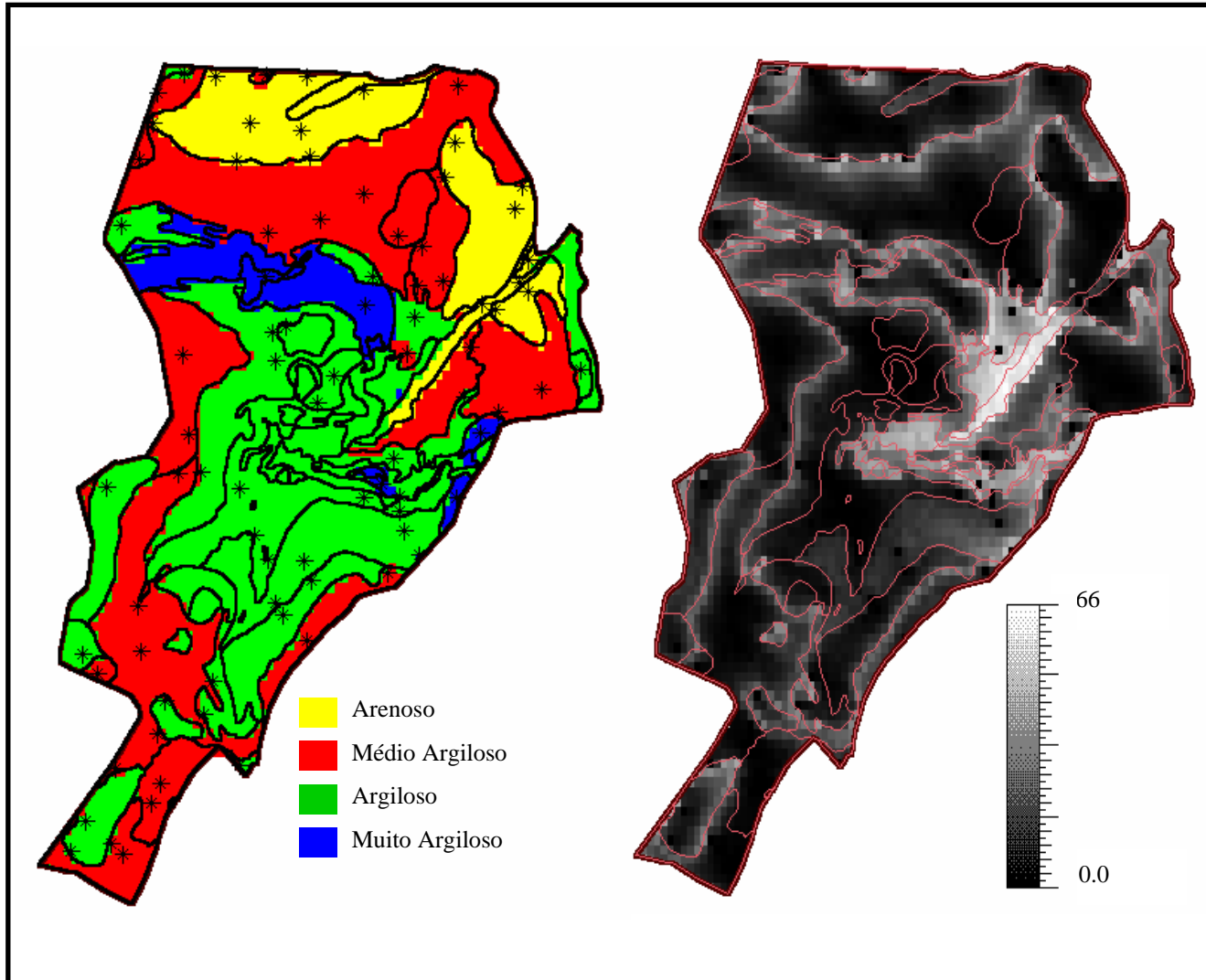


50 %

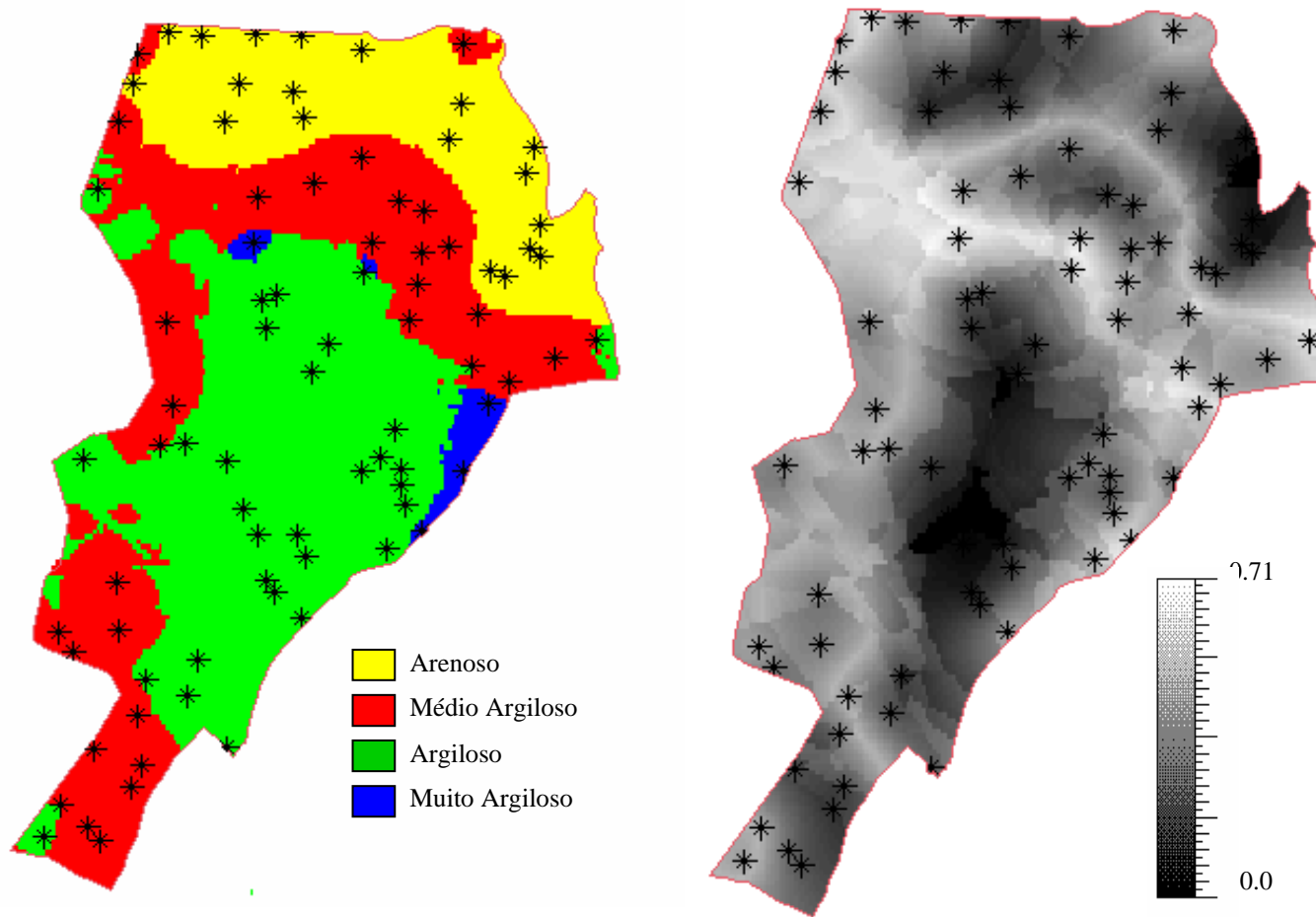
Uso de informação indireta (soft data)



Modelagem com Uso de informação indireta (soft data)



Modelagem sem informação indireta (krigeagem ou simulação estocástica)



•Referencias básicas

Burrough, P. A., 1986. Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press – Oxford – London.

Burrough, P. A.; McDonnell, R. A.,1998. *Principles of Geographical Information Systems*. Oxford University Press, Inc, New York, USA.

Bussab, W. O.; Morettin, P. A.. 1987. *Estatística Básica*. Atual Editora Ltda, Sao Paulo, Brasil.

Deutsch, C. V.; Journel, A. G., 1998. Geostatistical Software Library and User's Guide. Oxford University Press, New York, USA.

Heuvelink, G. B. M., 1998. *Error Propagation in Environmental Modelling with GIS*. Taylor and Francis Inc, Bristol, USA.

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