

# REPRESENTAÇÃO E PROPAGAÇÃO DE INCERTEZAS EM DADOS DE SOLOS. I - ATRIBUTOS CATEGÓRICOS<sup>(1)</sup>

S. BÖNISCH<sup>(2)</sup>, M. L. LOPES ASSAD<sup>(3)</sup>,  
G. CÂMARA<sup>(4)</sup> & A. M. V. MONTEIRO<sup>(4)</sup>

## RESUMO

Este trabalho teve por objetivos utilizar krigagem por indicação para espacializar propriedades de solos expressas por atributos categóricos, gerar uma representação acompanhada de medida espacial de incerteza e modelar a propagação de incerteza pela álgebra de mapas por meio de procedimentos *booleanos*. Foram estudados os atributos: teores de potássio (K) e de alumínio trocáveis, saturação por bases (V), soma de bases (S), capacidade de troca catiônica (CTC), textura (Tx) e classes de relevo (CR), de profundidade efetiva do solo, de drenagem interna e de pedregosidade e, ou, rochosoidade, extraídos de 222 perfis pedológicos e de 219 amostras extras, referentes a solos do estado de Santa Catarina. A espacialização das incertezas evidenciou a variabilidade espacial dos dados a qual foi relacionada com a origem das amostras e com o comportamento do atributo. Os atributos S, V, K e CR apresentaram grau de incerteza maior do que o de Tx e CTC e houve aumento da incerteza quando representações categóricas foram integradas.

**Termos de indexação:** krigagem por indicação, análise espacial, geoestatística, SIG.

---

<sup>(1)</sup> Extraído da Tese de Mestrado apresentada, pelo primeiro autor, ao Programa de Pós-Graduação em Sensoriamento Remoto do Instituto Nacional de Pesquisas Espaciais – INPE. Trabalho realizado com Bolsa da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Recebido para publicação em junho de 2002 e aprovado em outubro de 2003.

<sup>(2)</sup> Mestre em Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais – INPE. Av. dos Astronautas 1758, CEP 12227-010 São José dos Campos (SP). E-mail: sbonisch@cwmatrix.com.br

<sup>(3)</sup> Professora do DRNPA/CCA, Universidade Federal de São Carlos – UFSCar. CEP 13600-970 Araras (SP). E-mail: assad@cca.ufscar.br

<sup>(4)</sup> Pesquisador do Instituto Nacional de Pesquisas Espaciais – INPE. CEP 12227-010 São José dos Campos (SP). E-mail: gilberto@dti.inpe.br; miguel@dpi.inpe.br

## SUMMARY: REPRESENTATION AND PROPAGATION OF SOIL DATA UNCERTAINTIES. I - CATEGORIC ATRIBUTES

*The objectives of this work were to generate a representation of soil properties expressed by categorical attributes by kriging indicator; to assess the uncertainty in estimates, and to model the uncertainty propagation of map algebra by means of boolean procedures. The studied attributes were potassium (K) and aluminum (Al) exchangeable contents, sum of bases (SB), cationic exchange capacity (CEC), base saturation (V), texture (Tx) and relief classes (RC), effective soil depth, internal drainage, and stoniness and/or rockiness, extracted from 222 pedologic profiles and 219 extra samples from soils of the State of Santa Catarina, Brazil. The uncertainties evidenced the spatial variability of the data related to the samples' origin and attribute behavior. Attributes SB, V, K, and RC presented higher uncertainties than Tx and CEC, and there was an increase of uncertainty when categorical representations were integrated.*

*Index terms: kriging indicator, spatial analysis, geostatistic, GIS.*

## INTRODUÇÃO

A informação pedológica é de grande importância nos estudos ambientais e, no Brasil, os dados disponíveis sobre os solos encontram-se representados em diferentes escalas, muitas vezes incompatíveis com a demanda dos usuários dessas informações. A extensão continental do País e os elevados custos de levantamentos de solos impõem a adoção de procedimentos inferenciais que permitam: (a) otimizar o uso dos mapeamentos já realizados e (b) melhorar a qualidade – precisão e confiabilidade – dos mapeamentos a serem realizados.

O geoprocessamento permite explorar diferentes representações espaciais para as propriedades do ambiente e realizar múltiplas operações de análise espacial. No mapeamento do mundo geográfico, a natureza empírica da realidade é representada por entidades geográficas que podem ser modeladas conceitualmente e classificadas em geocampos e geoobjetos, definidos, respectivamente, como variáveis espacialmente contínuas e variáveis espacialmente descontínuas e individualizáveis (Goodchild, 1992). Os geocampos podem ser categóricos ou numéricos. Em um categórico, o geocampo associa um valor de um conjunto nominal (uma classe de um mapa, por exemplo) a cada ponto do espaço, de determinada região geográfica e, em um numérico, o geocampo associa a cada ponto um valor do conjunto dos inteiros ou dos reais (Câmara, 1995).

Freqüentemente, os solos são representados espacialmente como geocampos categóricos, apoiando-se na irreal suposição de que propriedades físicas e químicas são constantes dentro de regiões homogêneas com bordas bem delineadas e definidas (Burrough & McDonnell, 1998). Esta premissa leva a dois tipos de incerteza: (a) um simples atributo não

pode ser associado com 100 % de certeza a uma determinada localização (ambigüidade do atributo); (b) as bordas entre entidades não estão localizadas com 100 % de certeza (confusão espacial) (Brown, 1998). A incerteza é a diferença entre a realidade e sua representação, incluindo erros, falhas e o conceito estatístico de variação (Heuvelink, 1998), sendo, portanto, influenciada pela variabilidade espacial das entidades geográficas e pelo procedimento de mapeamento usado.

A modelagem da incerteza de atributos espaciais pode ser feita por krigagem por indicação que constitui uma forma de krigagem não-paramétrica, na qual a distribuição de probabilidade associada ao atributo é estimada experimentalmente, sem uma suposição de normalidade da distribuição (Goovaerts, 1997). O procedimento de modelagem por krigagem por indicação tem vantagens sobre a modelagem estocástica linear (ou krigagem tradicional): (a) a modelagem é não-paramétrica; (b) a modelagem da incerteza é independente de um estimador; (c) é possível modelar dados de natureza categórica, além de dados de natureza numérica (Felgueiras, 1999).

Utiliza-se uma aproximação discretizada da função de distribuição acumulada, condicionada aos dados disponíveis. Esta aproximação é construída a partir das diferentes classes do atributo categórico; para cada uma das  $k$  classes, aplica-se uma transformação binária não-linear (chamada transformação por indicação), do tipo:

$$I(u; z_k) = 1 \text{ se } Z(u) = z_k$$

$$I(u; z_k) = 0 \text{ se } Z(u) \neq z_k$$

em que  $Z(u)$  é a variável aleatória categórica que para  $u = u_\alpha$  corresponde ao conjunto amostral do atributo categórico;  $u$  é um vetor posição dentro de uma região  $A$  da superfície terrestre;  $u_\alpha$  é um vetor

posição amostrado dentro de A;  $z_k$  é o valor (z) da classe (k) e  $I(u; z_k)$  é um novo conjunto amostral, binário, que constitui uma variável aleatória.

O cálculo dos valores do atributo para as posições não amostradas baseia-se na relação entre o valor esperado da variável aleatória por indicação  $I(u; z_k)$  e a distribuição de probabilidade do atributo, e é dado por:

$$E\{I(u; z_k) | (n)\} = 1 * \text{Prob}\{I(u; z_k) = 1 | (n)\} + 0 * \text{Prob}\{I(u; z_k) = 0 | (n)\} = 1 * \text{Prob}\{I(u; z_k) = 1 | (n)\} = F(u; z_k) | (n)$$

A esperança de  $I(u; z_k)$  condicionada a  $n$  valores vizinhos fornece, para cada classe  $z_k$ , uma estimativa de  $F(u; z_k | (n))$ , que é a função de distribuição de probabilidade acumulada condicionada (fdpac) da variável aleatória  $Z(u)$ . A esperança condicional  $E\{I(u; z_k) | (n)\}$  pode ser estimada por algoritmos de krigagem. Para cada posição  $u$ , podem-se estimar a esperança condicional e os valores da fdpac de  $Z(u)$  para as diferentes classes do atributo categórico.

O processo de modelagem do atributo categórico por procedimentos da geoestatística por indicação pode ser feito da seguinte forma: (a) para cada ponto do espaço, estima-se a probabilidade de cada uma das classes; (b) atribui-se a cada ponto a classe de maior probabilidade; e (c) a incerteza de classificação de cada ponto é estimada como a diferença entre a probabilidade da classe escolhida e a soma das probabilidades das demais classes.

Assim, a estimativa do valor não amostrado da variável aleatória categórica  $Z(u)$  pode ser feita por meio do estimador de moda, o qual atribui para cada posição  $u$  a classe de maior probabilidade (Soares, 1992), ou seja:

$$Z(u) = z_k \text{ se } p_k(u) > p_i(u) \quad \forall i = 1.., K \text{ e } k \neq i$$

A incerteza associada a esse valor estimado pode ser inferida por uma medida chamada de incerteza pela moda, que corresponde à probabilidade de não ser atribuída, para a posição  $u$ , a classe de maior probabilidade  $p_k(u)$  (Felgueiras, 1999), e é estimada por:

$$\text{Inc}(u) = 1 - p_k(u)$$

A propagação de incertezas em geoprocessamento corresponde ao tratamento das incertezas associadas aos dados de entrada, quando estes são submetidos a operações de álgebra de mapas. Uma representação de saída  $U(\cdot)$  originada de um processo algébrico de mapas corresponde às representações de entrada  $b_i(\cdot)$  e suas incertezas  $V_i(\cdot)$ , manipuladas pelas operações  $g(\cdot)$  (Heuvelink, 1998), ou seja:

$$U(\cdot) = g([b_1(\cdot) + V_1(\cdot)], \dots, [b_m(\cdot) + V_m(\cdot)])$$

As funções de álgebra de mapas, aplicadas às representações categóricas e, particularmente, às

funções de integração, restringem-se aos operadores lógicos da álgebra *booleana*, como os operadores “AND”, “OR” e “XOR”.

Os procedimentos da krigagem por indicação para atributos categóricos permitem estimar as medidas pontuais de incerteza, que correspondem à probabilidade de classificação incorreta. A aplicação da teoria de probabilidade depende do operador lógico usado na integração. Se for usado o operador “AND”, que corresponde a uma operação de interseção de dois eventos, a incerteza propagada é estimada pela probabilidade de não ocorrer esta interseção, ou seja:

$$\text{Inc} = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - (P(A) \cdot P(B/A))$$

em que  $\text{Inc}$  é a incerteza;  $P(A)$  é a probabilidade de ocorrer o evento A; e  $P(B/A)$  é a probabilidade condicional de ocorrer B, dado que A está ocorrendo.

Se os eventos A e B forem independentes, o fato de ocorrer A não indica nada a respeito da ocorrência de B, então  $P(B/A) = P(B)$  (Meyer, 1983). A incerteza propagada passa a ser estimada por:

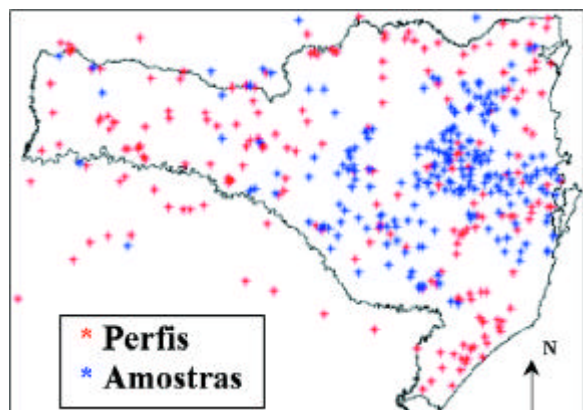
$$\text{Inc} = P(\overline{A \cap B}) = 1 - P(A \cap B) = P(\overline{A}) \cdot P(\overline{B})$$

Este trabalho teve por objetivos espacializar propriedades de solos expressas por atributos categóricos, utilizando krigagem por indicação; gerar, para cada atributo de solo estudado, uma representação acompanhada de medida espacial de incerteza e modelar a propagação de incerteza gerada em procedimentos *booleanos* de álgebra de mapas de atributos categóricos. Todo o procedimento se apóia em três premissas: (a) os dados de solos utilizados representam o universo de amostras disponíveis e possíveis de ser manipuladas, não sendo objeto deste trabalho adquirir novos dados; (b) os dados de solos utilizados no trabalho são confiáveis e não foram submetidos à crítica e (c) as classes de fertilidade e de limitação em fertilidade, estabelecidas em EMBRAPA (1997), utilizadas para expressar propriedades de solo na forma categórica e para testar funções de álgebra de mapas, estão corretas e representam adequadamente os fenômenos pedológicos considerados.

## MATERIAL E MÉTODOS

Os dados de solos foram extraídos de 222 perfis pedológicos e 219 amostras extras (Figura 1), obtidos em diferentes fontes (UFSM, 1972; IBGE, 1986; Almeida, 1988; Bognola, 1995; Almeida et al., 1997; EMBRAPA, 1997; EMBRAPA, 1998; Gmach, 1999; IBGE, 1999; Corrêa, 2001).

As variáveis consideradas foram: teor de potássio trocável (K), teor de alumínio trocável (Al), saturação



**Figura 1. Distribuição dos perfis completos e amostras complementares utilizados no estudo.**

por bases (V), soma de bases (S), capacidade de troca catiônica (CTC), classes de relevo (CR), profundidade efetiva do solo (PES), classe de drenagem interna (DR), pedregosidade e, ou, rochosidade (PR) e textura (TX), que serviram de base para o zoneamento pedoclimático para a cultura de soja no estado de Santa Catarina (EMBRAPA, 1997). As variáveis K, Al, V, S e CTC foram determinadas para cada horizonte ou camada de solo e CR, PES, DR e PR referiam-se à área amostrada. Quanto à TX, foi definida a partir de informações de perfis de solo e de amostras extras.

Visando à obtenção de um único valor de cada propriedade por ponto do espaço, foi aplicada uma

média ponderada aos dados disponíveis para horizontes e camadas, onde o peso atribuído para cada uma dependeu da percentagem de raízes de soja normalmente presente na profundidade correspondente àquele horizonte ou àquela camada. Tomaram-se por referência a fase de floração e a formação das primeiras vagens da soja, na qual a distribuição do peso das raízes secas atingiu aproximadamente 71 % do peso total, entre zero e 7,5 cm de profundidade; 15 %, entre 7,5 e 15 cm; 4 %, entre 15 e 23 cm; e o restante (10 %), além dos 23 cm de profundidade (Miyasaka & Medina, 1981).

Nas amostras extras, para as variáveis K, Al, V, S e CTC, foram considerados apenas os dados referentes a uma profundidade mínima de 20 cm, já que a concentração das raízes até esta profundidade foi de aproximadamente 90 %. As amostras que não atingiram esta profundidade mínima foram desconsideradas. A variável PR foi considerada apenas para os dados de perfis, por não estar disponível nas informações referentes às amostras extras, enquanto DR e PES foram extraídas de perfis de solos e de algumas amostras extras.

O primeiro passo na representação e avaliação dos atributos categóricos foi a transformação das variáveis K, Al, V, S e CTC (numéricas) e TX, CR, PES, DR e PR (nominais) em atributos categóricos - classificação categórica - considerando as exigências da cultura de soja no manejo C, definido por Ramalho & Beek (1995), conforme os critérios (Quadro 1) adotados pelo Zoneamento Pedoclimático do Brasil (EMBRAPA, 1997).

**Quadro 1. Critérios para classificação dos atributos dos solos, em classes de fertilidade<sup>(1)</sup>, para as variáveis teor de potássio (K), teor de alumínio trocável (Al), soma de bases (S), capacidade de troca catiônica (CTC), saturação por bases (V), e em classes de aptidão agrícola<sup>(2)</sup>, para os atributos: textura (TX), classe de relevo (CR), profundidade efetiva do solo (PE), classe de drenagem (DR) e pedregosidade/rochosidade (PR), tomando por base as exigências da cultura de soja**

Classe de fertilidade <sup>(1)</sup>	Alta	Média	Baixa	Muito baixa
K, mg kg <sup>-1</sup>	≥ 80	≥ 45 e < 80	< 45	
Al, cmol <sub>c</sub> kg <sup>-1</sup>	< 0,3	≥ 0,4 e < 1	≥ 1 e < 4	≥ 4
S, cmol <sub>c</sub> kg <sup>-1</sup>	≥ 4	≥ 2 e < 4	< 2	
CTC, cmol <sub>c</sub> kg <sup>-1</sup>	≥ 50	≥ 35 e < 50	≥ 10 e < 35	< 10
V (%)	≥ 8	≥ 6 e < 8	≥ 4 e < 6	< 4
Classe de aptidão agrícola <sup>(2)</sup>	Boa	Regular	Restrita	Inapta
TX	med, arg, mtarg, med/arg e mtarg	aren/arg e med/argcasc	aren, medcasc, medcasc/argcasc	argcasc, mtargcasc e org
CR	pln e suaveond	ond	ond	ftond, mont e esc
PE, cm	PES > 100	50 < PES < 100	30 < PES < 50	PES < 30
DR	acentdren e bemdren	ftdren a modren	excdren e impdren	maldren
PR	ausente	ausente	pouca e moderada	abundante

<sup>(1)</sup> Fonte: CEPA (1985) citado por EMBRAPA (1997) e Ramalho Filho & Beek (1995). <sup>(2)</sup> Adaptado de EMBRAPA (1997). <sup>(3)</sup> Abreviações: med = média, arg = argilosa, mtarg = muito argilosa, aren = arenosa, argcasc = argilo-cascalhenta, medcasc = media cascalhenta, mtargcasc = muito argilosa e cascalhenta, org = solo orgânico, pln = plano, suaveond = suave ondulado, ond = ondulado, ftond = forte ondulado, mont = montanhoso, esc = escarpado, acentdren = acentuadamente drenado, bemdren = bem drenado, ftdren = fortemente drenado, modren = moderadamente drenado, excessivamente drenado, dren = bem drenado, impdren = imperfeitamente drenado e maldren = mal drenado.

Em seguida, foi feita a transformação por indicação dos conjuntos amostrais nominais e a análise dos modelos teóricos de semivariogramas ajustados aos modelos experimentais de cada classe (Quadros 2 e 3).

Após a definição dos modelos de semivariogramas teóricos, as classes dos atributos estudados foram espacializadas com o uso do interpolador de krigagem ordinária, sendo corrigidos os resultados da espacialização com eventual desvio de relação de ordem. Partindo das probabilidades  $p_k(u)$ , a estimativa do valor não amostrado de cada variável aleatória categórica considerada foi obtida pelo estimador de moda. A incerteza associada a esse valor estimado foi inferido pela medida de incerteza pela moda, que corresponde à probabilidade de não ser atribuída, para a posição  $u$ , a classe de maior probabilidade  $p_k(u)$ .

Visando avaliar a propagação de incertezas associadas às representações categóricas, os atributos categóricos K, Al, V, S e CTC foram integrados para espacialização da limitação das

terras em fertilidade. Foram aplicados procedimentos de álgebra *booleana* por meio do operador “AND”, prevalecendo sempre a classe de maior restrição.

Foi feita a verificação da independência entre os eventos, considerando um nível de confiança de 95 %, adotando-se um teste estatístico, cuja hipótese nula é a igualdade entre a probabilidade de interseção dos eventos e o produto das probabilidades individuais destes eventos. O teste partiu das probabilidades marginais dos eventos, as quais foram estimadas com as próprias amostras. Assim, no caso de independência entre A e B, as probabilidades marginais,  $p_{i*}$  e  $p_{*j}$ , foram estimadas por (Snedecor & Cochran, 1971):

$$\hat{p}_{i*} = \frac{n_{i*}}{n} \text{ e } \hat{p}_{*j} = \frac{n_{*j}}{n}$$

Destas probabilidades marginais, podem ser estimadas as probabilidades  $p_{ij}$ , e delas,  $E_{ij}$ :

$$\hat{p}_{ij} = \hat{p}_{i*} \cdot \hat{p}_{*j} \text{ e } \hat{E}_{ij} = \hat{p}_{ij} \cdot n$$

A estatística do teste correspondeu a:

**Quadro 2. Parâmetros dos semivariogramas teóricos obtidos para cada um dos atributos de fertilidade estudados**

Classe	Direção	Modelo	<sup>(1)</sup> C <sub>0</sub>	<sup>(1)</sup> C	<sup>(1)</sup> a (km)
Potássio, mg kg <sup>-1</sup>					
Alta	Isotrópico	Exponencial	0,211	0,031	83,88
Média	Isotrópico	Esférico	0,217	0,007	73,44
Baixa	Isotrópico	Exponencial	0,074	0,016	45,64
Muito baixa	Isotrópico	Exponencial	0,115	0,009	67,78
Alumínio trocável, cmol <sub>c</sub> kg <sup>-1</sup>					
Alta	100°	Exponencial	0,069	0,026	117,55
	10°	Exponencial	0,072	0,008	74,45
Média	Isotrópico	Esférico	0,074	0,012	81,53
	Baixa	100°	Exponencial	0,193	0,056
		10°	Exponencial	0,193	0,034
Muito baixa	Isotrópico	Gaussiano	0,175	0,080	63,29
Capacidade de troca de cátions, cmol <sub>c</sub> kg <sup>-1</sup>					
Alta	Isotrópico	Exponencial	0,046	0,055	151,50
Média	Isotrópico	Gaussiano	0,037	0,005	93,94
Baixa	Isotrópico	Exponencial	0,033	0,024	47,59
Muito baixa	Isotrópico	Esférico	0,005	0,001	24,42
Soma de bases, cmol <sub>c</sub> kg <sup>-1</sup>					
Alta	Isotrópico	Exponencial	0,145	0,054	40,36
Média	Isotrópico	Exponencial	0,166	0,017	150,95
Baixa	Isotrópico	Esférico	0,106	0,007	58,88
Muito baixa	Isotrópico	Gaussiano	0,176	0,059	42,88
Saturação por bases, %					
Alta	100°	Exponencial	0,108	0,047	178,85
	10°	Exponencial	0,127	0,018	88,76
Média	Isotrópico	Exponencial	0,072	0,010	93,93
Baixa	100°	Esférico	0,185	0,056	79,65
	10°	Esférico	0,202	0,030	45,11
Muito baixa	100°	Exponencial	0,179	0,084	173,96
	10°	Exponencial	0,167	0,061	103,32

<sup>(1)</sup> C<sub>0</sub>: efeito pepita; C<sub>1</sub>: variância estrutural; a: alcance (cf. definições propostas por Vieira, 2000).

$$Q = \sum_{i=1}^L \sum_{j=1}^C \left( \frac{n_{ij} - \hat{E}_{ij}}{\hat{E}_{ij}} \right)^2$$

que segue uma distribuição  $\chi^2$ , com  $((L - 1) * (C - 1))$  graus de liberdade.

As probabilidades marginais, necessárias para a estimativa da estatística do teste, foram determinadas com a operação denominada tabulação cruzada, a partir do cruzamento de cada duas representações de atributos do solo. Considerando os resultados obtidos (Quadro 4), a integração das representações de limitação em fertilidade foi feita em duas partes;

uma delas considerou as representações referentes às propriedades CTC e V, que se mostraram independentes entre si, e a outra considerou as representações referentes às propriedades S, Al e K, que também se mostraram independentes entre si.

As representações parciais foram, posteriormente, integradas para gerar uma única representação e, da mesma forma que nas integrações anteriores, foi aplicado o operador “AND” da álgebra *booleana*, prevalecendo a classe de maior restrição. O teste de independência entre as representações parciais apontou serem elas independentes para um nível de confiança de 95 %.

**Quadro 3. Parâmetros dos semivariogramas teóricos obtidos para cada um dos atributos de aptidão agrícola estudados**

Classe	Direção	Modelo	<sup>(1)</sup> C <sub>0</sub>	<sup>(1)</sup> C	<sup>(1)</sup> a (km)
Textura					
Boa	Isotrópico	Exponencial	0,043	0,025	55,65
Regular	Isotrópico	Esférico	0,002	0,009	11,82
Restrita	Isotrópico	Exponencial	0,014	0,008	66,76
Inapta	Isotrópico	Gaussiano	0,037	0,031	311,32
Classes de relevo					
Boa	Isotrópico	Exponencial	0,083	0,160	343,72
Regular	Isotrópico	Gaussiano	0,189	0,153	80,23
Restrita	Isotrópico	Exponencial	0,028	0,017	79,03
Inapta	Isotrópico	Gaussiano	0,207	0,035	71,45
Profundidade efetiva <sup>2</sup>					
Boa	Isotrópico	Exponencial	0,191	0,019	65,05
Regular	Isotrópico	Exponencial	0,136	0,020	74,51
Restrita	Isotrópico	Gaussiano	0,044	0,002	193,03
Inapta <sup>(2)</sup>	Isotrópico	Gaussiano	29,108	12,232	166,02
Drenagem <sup>(3)</sup>					
Boa	Isotrópico	Exponencial	0,156	0,053	38,62
Regular	Isotrópico	Exponencial	0,134	0,029	9,06
Restrita <sup>(3)</sup>	Isotrópico	Esférico	48,049	24,536	26,99
Inapta	Isotrópico	Exponencial	0,023	0,026	73,41
Pedregosidade e, ou, rochiosidade <sup>(4)</sup>					
Boa	106°	Esférico	0,083	0,101	214,63
	16°	Esférico	0,091	0,080	147,70
Restrita	106°	Esférico	0,083	0,101	214,63
	16°	Esférico	0,091	0,080	147,70

<sup>(1)</sup> C<sub>0</sub>: efeito pepita; C<sub>1</sub>: variância estrutural; escala ou contribuição; a: alcance (cf. definições propostas por Vieira, 2000). <sup>(2)</sup> O modelo de variograma experimental da classe inapta corresponde ao semivariograma relativo geral. <sup>(3)</sup> O modelo de variograma experimental da classe restrita corresponde ao semivariograma relativo geral. <sup>(4)</sup> Não foram identificadas áreas nas classes regular e inapta.

**Quadro 4. Resultados obtidos no teste de independência entre as representações categóricas utilizadas na avaliação da limitação em fertilidade dos solos para a cultura de soja no estado de Santa Catarina**

	Saturação por bases (V)	Soma de bases (S)	Alumínio trocável (Al)	Potássio trocável (K)
CTC	Independente	Dependente	Independente	Dependente
V		Independente	Dependente	Dependente
S			Independente	Independente
Al				Independente

Todos os procedimentos de espacialização usando krigagem por indicação, com ajustes de semivariogramas, propagação de incertezas e tabulação cruzada, foram feitos no SPRING® versão 3.4, desenvolvido pelo INPE – Instituto Nacional de Pesquisas Espaciais.

## RESULTADOS E DISCUSSÃO

Da análise dos semivariogramas gerados (Quadros 2 e 3), verificou-se que os atributos considerados foram ajustados principalmente por semivariogramas unidirecionais, indicando variabilidade espacial isotrópica. Os modelos de semivariogramas diferiram entre si para todos os atributos, com exceção de PR, no qual todas as classes foram ajustadas por semivariogramas esféricos anisotrópicos. As classes de fertilidade, segundo os atributos Al e V, também apresentaram anisotropia nos semivariogramas ajustados para duas e três classes, respectivamente. O caráter anisotrópico do semivariograma indica que a variabilidade espacial depende da direção, e este aspecto pode ser útil na definição de formas de amostragem desses atributos. Couto et al. (2000), estudando a variabilidade espacial do teor de argila e do potássio trocável em solos do Mato Grosso, constataram que esses atributos apresentaram moderada dependência espacial e mostraram anisotropia tanto nos horizontes alterados por atividades antrópicas quanto nos não alterados.

O patamar é o valor da semivariância em que a curva se estabiliza num valor constante correspondente ao alcance e constitui a soma da variância estrutural com o efeito pepita. O alcance é, portanto, a distância da origem a partir da qual as amostras não são correlacionadas (Goovaerts, 1997). Desta forma, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras ( $\text{Var}[Z(x) - Z(x + h)]$ ) torna-se invariante com a distância. O valor do efeito pepita (Quadros 2 e 3) indica a descontinuidade dos dados para distâncias menores do que a menor distância entre as amostras. Parte desta descontinuidade pode ser também devida a erros de medição, mas é impossível quantificar se a maior contribuição provém dos erros de medição ou da variabilidade de pequena escala não captada pela amostragem (Isaaks & Srivastava, 1989). Em outros termos, o efeito pepita é o valor da semivariância para a distância zero e representa o componente da variabilidade espacial que não pode ser relacionado com uma causa específica (variabilidade ao acaso). Quando comparados com os valores de variância estrutural (Quadros 2 e 3), o efeito pepita foi muito maior, principalmente para os atributos S, V, Al, K e PES, confirmando o forte componente ao acaso da variabilidade espacial desses atributos, seja pelo histórico de manejo (todos dependentes das condições

de uso), seja pelas diferentes origens dos dados. Assim, pode-se inferir que o número de dados utilizados neste estudo para o estado de Santa Catarina não foi suficiente para explicar a variabilidade espacial dos atributos considerados.

As representações resultantes da krigagem de atributos categóricos dos solos (Figuras 2 e 3) mostraram a probabilidade de que a classe do atributo considerada exceda a classe definida. Assim, as representações a, b, c, d e e (Figuras 2 e 3) foram obtidas pelo operador de moda, atribuindo para cada posição geográfica a classe de maior probabilidade do atributo considerado; enquanto as representações f, g, h, i e j mostram a incerteza associada dada por “1 – a probabilidade da classe atribuída”. Neste caso, a legenda do mapa de incertezas, em escala de cinza, indica a *probabilidade de que a classe do atributo considerado exceda a classe definida como mais provável*.

Em certas posições, as probabilidades de serem classificadas pelas  $n$  classes consideradas foram grandes e isto se refletiu nas incertezas estimadas. Isto porque, em cada localização  $u$ , foi inferido um valor de probabilidade  $p_k(u)$  para cada uma das classes  $z_k$ . Partindo das probabilidades  $p_k(u)$ , a estimativa do valor não amostrado da variável aleatória categórica  $Z(u)$  foi feita por meio do estimador de moda. Este atribuiu, para cada posição  $u$ , a classe de maior probabilidade e, por vezes, a probabilidade de não ser classificada por esta classe também foi grande. A aplicação de uma métrica de incerteza baseada no uso da entropia de Shannon poderia melhorar esta relação (Felgueiras, 1999), mas não modificá-la.

De modo geral, as representações das incertezas estimadas apresentaram valores relativamente elevados tanto nas zonas de transição das classes definidas, quanto no interior dessas classes (Figuras 2 e 3). Isto indica uma elevada variabilidade espacial e confirma o quanto é fraca a suposição feita na modelagem categórica de propriedades de solos de que estas são constantes e homogêneas para áreas que teriam limites bem delineados e bem definidos.

Os atributos utilizados no presente estudo puderam ser agrupados em dois conjuntos de estruturas distintas: (a) atributos que resultam de uma classificação de resultados quantitativos obtidos em análises de laboratório, realizadas em amostras coletadas em pontos do espaço (perfis e amostras extras), e (b) atributos que resultam de uma apreciação qualitativa feita por profissionais, no campo, referente a uma área. No primeiro grupo, estão os atributos CTC, V, S, Al e K e, no segundo, estão os atributos CR, DR, PES e PR. Assim, poder-se-ia esperar que a incerteza associada aos atributos CR, DR, PES e PR fosse maior do que a dos demais, pois esses dados foram gerados com finalidades variadas por profissionais com diferentes graus de experiência, de instituições distintas e em diferentes épocas.

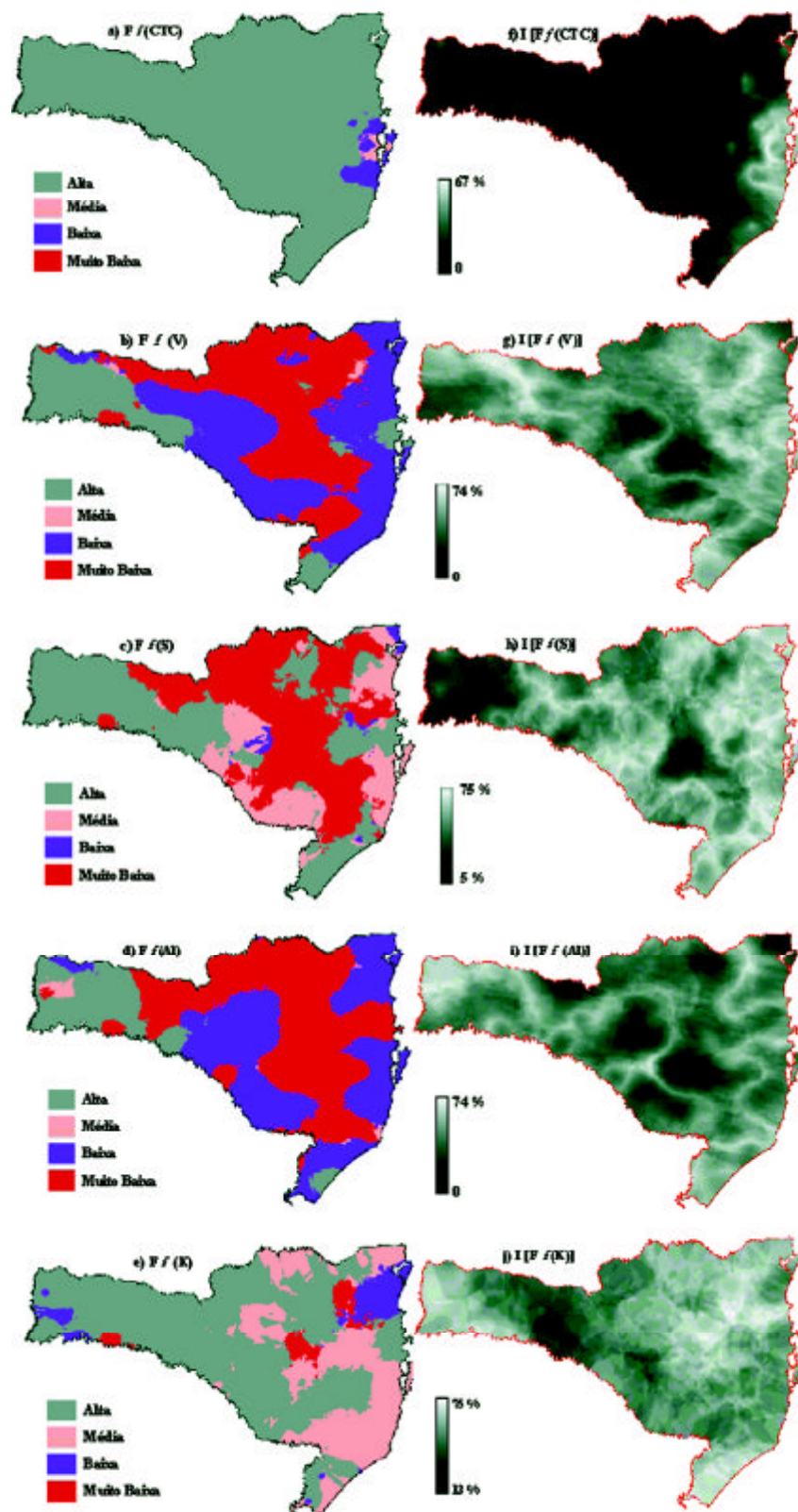


Figura 2. Espacialização das variáveis de solos: capacidade de troca catiônica, saturação por bases, soma de bases e teores de alumínio trocável e de potássio trocável, expressas por atributos categóricos de fertilidade [Ff(CTC), Ff(V), Ff(S), Ff(Al) e Ff(K)], e incertezas associadas [I [Ff(CTC)], I [Ff(V)], I [Ff(S)], I [Ff(Al)] e I [Ff(K)]], com a legenda do mapa de incertezas, em escala de cinza, indicando a probabilidade (em porcentagem) de que a classe do atributo considerado exceda a classe definida como mais provável.



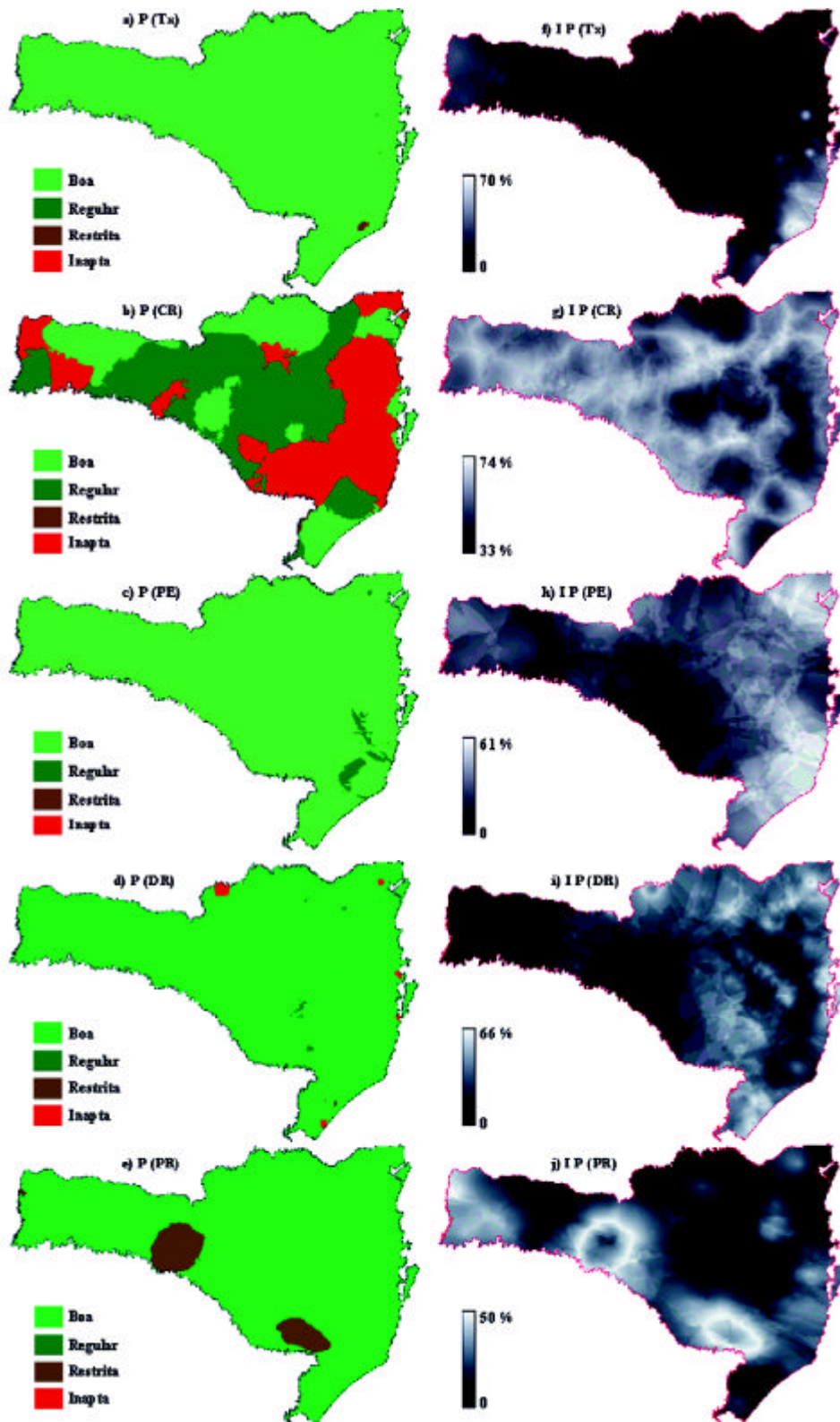


Figura 3. Espacialização das variáveis de solos: classe de textura, classe de relevo, profundidade efetiva, classe de drenagem e classe de pedregosidade/rochosidade, expressas por atributos categóricos de aptidão pedológica [Pf(Tx), Pf(CR), Pf(PE), Pf(DR) e Pf(PR)], e incertezas associadas [I[Pf(Tx)], I[Pf(CR)], I[Pf(PE)], I[Pf(DR)] e I[Pf(PR)]], com a legenda do mapa de incertezas, em escala de cinza, indicando a probabilidade (em percentagem) de que a classe do atributo considerado exceda a classe definida como mais provável.

Entretanto, comparando, por exemplo, a representação da incerteza de K (Figura 2e) com a de PR (Figura 3e), verificou-se que, na melhor situação (nível escuro), existe uma probabilidade de 87 % (Figura 2j) da informação apresentada na representação categórica de fertilidade segundo o atributo K estar correta (isto é,  $p_{\text{correto}} = 1 - 0,129$ ), enquanto essa probabilidade, para o atributo PR (Figura 3j), pode ser de 100 %. Na pior situação (níveis claros), a incerteza da representação categórica, ou seja, a probabilidade de não ser o que está representado, é de 75 % para K e de 50 % para PR. Ou seja, os dados obtidos a partir de uma apreciação qualitativa pareceram menos incertos do que os decorrentes de determinação quantitativa.

Na espacialização de atributos categóricos, o dado qualitativo estimado pelo pedólogo representou melhor a variabilidade espacial do que o dado numérico obtido em pontos do espaço. Isto é explicado pelo levantamento tradicional de solos que, conforme descrito por Wilding (1985), consiste num exercício de modelagem que envolve métodos científicos e elementos de arte. Neste sentido, a apreciação de campo feita pelo pedólogo abrange o desenvolvimento de um modelo mental que relaciona o solo com a paisagem. Este modelo envolve aproximações semelhantes ao que um pintor faz reconstruindo uma imagem numa tela. Nestas aproximações, estariam embutidas as possíveis incertezas dos dados.

Verificou-se que as representações de fertilidade, segundo os atributos V, S, K e CR, apresentaram variabilidade dos valores das incertezas associadas às representações categóricas maiores do que as relativas aos atributos TX e CTC (Figuras 2 e 3). Esta diferença pode estar relacionada com as características do processo representado pelo atributo. A capacidade de troca catiônica (CTC) é a quantidade total de cátions trocáveis adsorvidos pelo complexo coloidal e é condicionada pela mineralogia dos argilominerais e pela estrutura química da matéria orgânica.

Apesar de relacionada com a matéria orgânica, que é um constituinte do solo de instabilidade estrutural muito maior do que os argilominerais, a CTC é um atributo relativamente estável no tempo, não sendo sujeito a variações muito grandes, a menos que se faça um aporte muito expressivo de matéria orgânica ao solo e, principalmente, se o solo apresenta elevado conteúdo de areia. Neste sentido, a CTC tem comportamento semelhante ao da textura, que também revela uma estabilidade muito grande.

O mesmo não se pode dizer dos atributos V, S e K. Estes podem ser modificados rapidamente, por aporte de adubos e corretivos. Segundo Berg & Klamt (1997), num estudo sobre a variabilidade espacial de solos do Rio Grande do Sul, atributos como S e V apresentaram elevada variabilidade espacial, mesmo quando a distância entre as

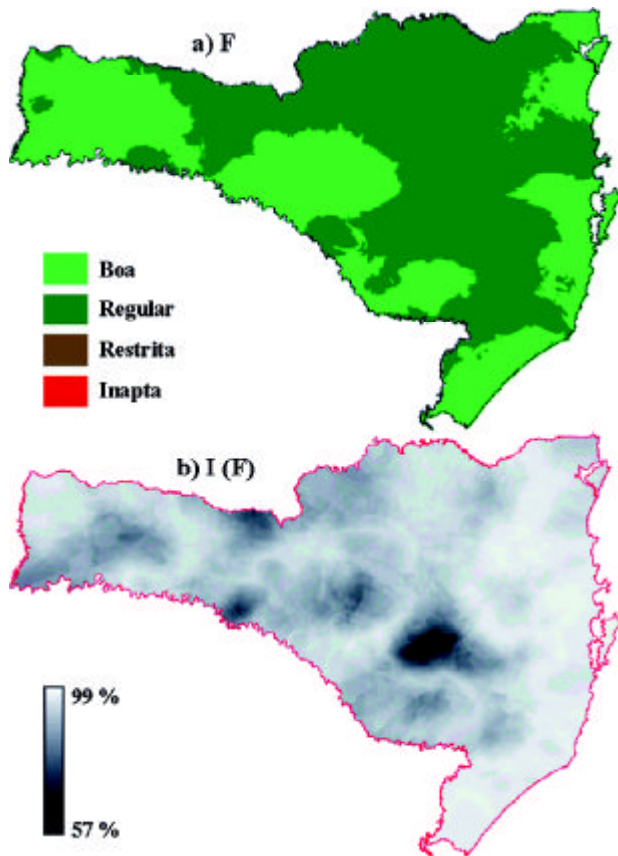
amostras era pequena ( $< 5$  m). Por outro lado, esses autores indicaram que parâmetros, como TX e CTC, apresentaram um padrão de variabilidade espacial mais estável, com maiores variações apenas em grandes distâncias entre amostras ( $> 1$  km). Para esses autores, o uso do solo é o principal responsável por variações de atributos químicos na superfície, enquanto o material de origem é o principal responsável por variações de textura a grandes distâncias.

Os dados utilizados neste trabalho, conforme indicado, provêm de diferentes fontes e foram obtidos em diferentes épocas. Como a krigagem por indicação constitui um procedimento de interpolação espacial, a inferência de um atributo para uma localização  $u$  não conhecida depende do valor do atributo na localização  $u$  conhecida. Quando os dados são coletados em épocas diferentes, é razoável esperar que atributos que podem ser modificados no tempo pela ação do homem possam apresentar um grau de incerteza maior do que aqueles menos modificáveis, mesmo quando situados relativamente próximos.

Quanto à classe de relevo (Figura 3b), a incerteza elevada da aptidão pedológica, segundo o atributo categórico, foi maior nas zonas de transição de classes do que no interior de classes. Neste caso, trata-se de um atributo dificilmente modificado por ação do homem e a incerteza nas bordas das classes definidas constitui uma confusão espacial (Brown, 1998) que reflete a mudança de classe. O atributo Al, que pode ser modificado pelos processos de correção do solo, também apresentou zonas de maior incerteza associadas, principalmente, às bordas das classes definidas.

No teste de funções de álgebra de mapas aplicados a atributos categóricos (classes de fertilidade), definidos a partir dos critérios estabelecidos no Zoneamento Pedoclimático do estado de Santa Catarina para a cultura de soja, três aspectos devem ser destacados. O primeiro refere-se aos resultados obtidos no teste de independência de variáveis (Quadro 4). Os valores referentes aos atributos S, V, K e Al indicaram independência por vezes inconsistente com os processos aos quais estão relacionados, provavelmente por causa da origem dos dados.

Outro aspecto é que a espacialização das incertezas da limitação em fertilidade (Figura 4) também indicou elevada incerteza tanto no interior das classes definidas quanto nas zonas de transição entre classes. Segundo o Zoneamento Pedoclimático do Brasil para a cultura da soja (EMBRAPA, 1997), enquanto as classes de fertilidade alta, média e baixa correspondem a uma aptidão boa, a classe de fertilidade muito baixa corresponde a uma aptidão regular. Estes critérios consideram que, no nível de manejo C, usam-se corretivos e fertilizantes e, portanto, a fertilidade não representa um entrave para o desenvolvimento das culturas.



**Figura 4. Espacialização de limitação em fertilidade e da incerteza associada, obtida a partir da integração dos atributos categóricos K, Al, V, S e CTC por meio de álgebra de mapas por meio de procedimentos *booleanos*.**

Como exemplo, a incerteza associada ao atributo S (Figura 2c) varia entre [0,05 - 0,749] e a incerteza associada à limitação em fertilidade (Figura 4) varia entre [0,566 - 0,992]. Estes resultados indicam que, na melhor situação, há 57 % de probabilidade de a informação apresentada pela representação categórica de limitação em fertilidade não estar correta, enquanto essa probabilidade, para o atributo S foi de apenas 5 %. Por outro lado, constatou-se que, na representação da incerteza associada à limitação em fertilidade (Figura 4b), houve forte predominância de áreas claras, isto é, de áreas com incerteza mais alta, e este aspecto foi mais expressivo do que nas representações das incertezas associadas aos atributos categóricos K, Al, V, S e CTC (Figura 3). Portanto, a incerteza aumentou quando diferentes representações foram integradas.

Um terceiro e importante aspecto diz respeito à escala de amostragem. Conforme destacam Burrough & McDonnell (1998), uma premissa para computar um semivariograma é assumir que todos os dados estão localizados numa mesma classe ou domínio e o semivariograma constitui um modelo

global que define a interpolação em toda área. No entanto, isso não é necessariamente verdadeiro, pois diferentes partes de área de estudo podem ter importantes diferenças na cobertura do solo, no tipo de rocha, no tipo de solo, cada um dos quais com sua estrutura ou padrão de correlação espacial própria, única (Burrough & McDonnell, 1998). Este aspecto é particularmente válido para os dados aqui utilizados que cobrem uma área de 96.442,9 km<sup>2</sup> (conforme dados obtidos em [www.ibge.com.br](http://www.ibge.com.br)), o que representa uma média de uma amostragem a cada 227 km<sup>2</sup> (considerando-se o conjunto de perfis e amostras de solos).

Conforme se pode constatar (Figura 1), as amostras utilizadas neste estudo não se encontravam uniformemente distribuídas na área de abrangência do estado de Santa Catarina, aspecto que pode ter interferido nas estimativas de incertezas. Portanto, este trabalho deve ser visto como uma primeira aproximação de estimativa de incertezas, quando da integração de propriedades naturais em modelos que buscam gerar cenários possíveis para intervenções específicas, a qual se mostrou útil para uma avaliação de todo o procedimento de análise espacial, que integra dados em meio digital.

## CONCLUSÕES

1. A krigagem por indicação permitiu espacializar atributos de solos como geocampos categóricos, gerar, para cada representação categórica, uma medida espacial de incerteza e modelar a propagação de incerteza gerada em procedimentos *booleanos* de álgebra de mapas de atributos categóricos.
2. As representações das incertezas estimadas para representações de atributos categóricos apresentaram valores, de modo geral, elevados tanto nas zonas de transição das classes definidas, quanto no seu interior, evidenciando elevada variabilidade espacial.
3. As probabilidades de estas posições serem classificadas pelas  $n$  classes consideradas foram grandes e, como o atributo foi classificado localmente pela classe de maior probabilidade, a probabilidade de ser classificado em outras classes também foi grande, o que refletiu nas altas incertezas estimadas.
4. A espacialização das incertezas evidenciou o grau de “confusão” das amostras, relacionado com a origem das amostras e, ou, com o comportamento do atributo considerado.
5. Os atributos saturação por bases, soma de bases, potássio trocável e classes de relevo apresentaram incertezas associadas às representações categóricas maiores do que os atributos textura e capacidade de troca catiônica.

6. Houve aumento da incerteza quando diferentes representações categóricas foram integradas por procedimentos *booleanos* de álgebra de mapas.

## LITERATURA CITADA

- ALMEIDA, J.A. Projeto pastagens nativas. Lages, Empresa Catarinense de Pesquisa Agropecuária e Universidade do Estado de Santa Catarina, 1988. Não publicado.
- ALMEIDA, J.A.; KÄMPF, N. & ALMEIDA, R. Caracterização mineralógica de Cambissolos originados de rochas pelíticas nos patamares do Alto Rio Itajaí e no Planalto de Lages (SC). R. Bras. Ci. Solo, 21:181-190, 1997.
- BOGNOLA, I.A. Caracterização química, física e mineralógica de solos intermediários entre Latossolos Brunos e Latossolos Roxos. Viçosa, Universidade Federal de Viçosa, 1995. 215p. (Tese de Mestrado)
- BROWN, D.G. Classification and boundary vagueness in mapping presettlement forest types. Int. J. Geog. Inf. Syst., 12:105-129, 1998.
- BURROUGH, P.A. & McDONNELL, R.A. Principles of geographic information systems. New York, Oxford University Press, 1998. 333p.
- CÂMARA, G. Modelos, linguagens e arquiteturas para bancos de dados geográficos. São José dos Campos, Instituto Nacional de Pesquisas Espaciais, 1995. 251p. (Tese de Doutorado)
- CORRÊA, J. Mineralogia e gênese das principais classes de solos de encostas basálticas no Estado de Santa Catarina. Lages, Universidade do Estado de Santa Catarina, 2001. 165p. (Tese de Mestrado)
- COUTO, E.G.; KLAMT, E. & STEIN, A. Estimativa do teor de argila e do potássio trocável em solos esparsamente amostrados no sul do estado do Mato Grosso. R. Bras. Ci. Solo, 24:129-140, 2000.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. Centro Nacional de Pesquisas de Solos. Levantamento de reconhecimento dos solos do estado de Santa Catarina. Rio de Janeiro, 1998. Não publicado.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. Zoneamento pedoclimático do estado de Santa Catarina. Rio de Janeiro, 1997. 59p.
- FELGUEIRAS, C.A. Modelagem ambiental com tratamento de incertezas em sistemas de informações geográficas: paradigma geoestatístico por indicação. São José dos Campos, Instituto Nacional de Pesquisas Espaciais, 1999. 181p. (Tese de Doutorado)
- GMACH, J. Evolução mineralógica dos solos e relações solo-superfícies geomórficas em áreas de encostas basálticas no Vale do Rio do Peixe, SC: topossequência de Ipira – Peritiba. Lages, Universidade do Estado de Santa Catarina, 1999. 142p. (Tese de Mestrado)
- GOODCHILD, M.F. Geographical data modeling. Comp. Geoc., 18: 401-408, 1992.
- GOOVAERTS, P. Geostatistical for natural resources evaluation. New York, Oxford University Press, 1997. 477p.
- HEUVELINK, G.B.M. Error propagation in environmental modelling with GIS. London, Taylor & Francis, 1998. 127p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Levantamento de recursos naturais. Rio de Janeiro, 1986. v.33. 796p.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Levantamento de recursos naturais. Rio de Janeiro, 1999. v.35. Não publicado.
- ISAAKS, E.H. & SRIVASTAVA, R.M. An introduction to applied geostatistics. New York, Oxford University Press, 1989. 561p.
- MEYER, P.L. Probabilidade: aplicações à estatística. Rio de Janeiro, Livros Técnicos e Científicos, 1983. 426p.
- MIYASAKA, S. & MEDINA, J.C. A soja no Brasil. Campinas, ITAL, 1981. 1062p.
- RAMALHO FILHO, A. & BEEK, K.J. Sistema de avaliação da aptidão agrícola das terras. Rio de Janeiro, Empresa Brasileira de Pesquisa Agropecuária, 1995. 65p.
- SNEDECOR, G.W. & COCHRAN, W.G. Méthodes statistiques. Paris, Association de Coordination Technique Agricole, 1971. 649p.
- SOARES, A. Geostatistical estimation of multi-phase structure. Mathem. Geol., 24:149-160, 1992.
- UNIVERSIDADE FEDERAL DE SANTA MARIA – UFSM. Levantamento de reconhecimento dos solos do Estado de Santa Catarina. R. Centro Ci. Rurais, 2:11-494, 1972.
- van den BERG, M. & KLAMT, E. Variabilidade espacial de características de solos na região do planalto médio, RS. II - Análise de semivariância e da variância. R. Bras. Ci. Solo, 21:401-408, 1997.
- VIEIRA, S.R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R.F.; ALVAREZ V, V.H. & SCHAEFER, C.E.G.R., eds. Tópicos em ciência do solo. Viçosa, Sociedade Brasileira de Ciência do Solo, 2000. v.1. p.1-54.
- WILDING, L.P. Spatial variability: its documentation, accommodation and implication to soil surveys. In: NIELSEN, D.R. & BOUMA, J., eds., Soil spatial variability. Int. Soc. Soil Sci. and Soil Sci. Soc. Amer. Workshop. Las Vegas, 1984. Proceedings. Wageningen, PUDOC publishers, Wageningen, 1985. p.166-194.