

PROPAGAÇÃO DE INCERTEZAS NO MAPEAMENTO CATEGÓRICO

Simone Bönisch
Paulo Justiniano Ribeiro Junior
Antônio Miguel Vieira Monteiro
Gilberto Câmara

1. Introdução

Propriedades físicas e químicas ambientais, como solos, vegetação, uso e cobertura da terra, tradicionalmente são representadas por variáveis temáticas. Com a evolução do conhecimento sobre a natureza, tem se tornado claro que as bases de representações destas variáveis temáticas são freqüentemente insuficientes para comunicar a complexidade das entidades envolvidas nos processos ambientais (Burrough, 1996).

Esta realidade deve ser uma preocupação dos responsáveis pela geração das bases de dados de recursos naturais e de seus usuários, principalmente quando estes dados são operados em meio digital, onde não há a intervenção humana capaz de lidar com parte das incertezas.

Estas incertezas, aqui atribuídas às diferenças entre a realidade e suas representações da realidade, são afetadas pela variabilidade espacial das entidades geográficas e pelo procedimento de mapeamento usado (Heuvelink, 1998).

Nos procedimentos de mapeamento de variáveis temáticas é feita uma pobre e irreal suposição, a de que as propriedades e recursos naturais são constantes dentro de regiões homogêneas, com bordas bem delineadas e definidas (Burrough e McDonnell, 1998).

Nos processos de tomadas de decisões, onde a informação espacial é fator determinante, há uma tendência em se acreditar fielmente nos mapas, esquecendo-se de que são representações aproximadas da verdade geográfica e têm incertezas associadas. É imprescindível que estas incertezas sejam estimadas e, em processos de análise espacial, sejam propagadas em acordo com as técnicas de interação aplicadas.

É neste contexto que se encontra este artigo. Ele propõe uma forma de propagação de incertezas espaciais associadas às representações de variáveis temáticas, quando são integradas por operadores lógicos da Álgebra Booleana. As incertezas associadas às representações são expressas por valores de probabilidade de não classificação das classes temáticas.

O estudo de caso trata da estimativa da fertilidade do solo para o cultivo da soja no Estado de Santa Catarina, tendo como base informações coletadas das análises de perfis de solos.

2. Objetivo

O trabalho tem por objetivo propor um método de propagação de incertezas espaciais associadas às representações de variáveis temáticas, quando estas são integradas por técnicas da análise espacial e aquelas estão relacionadas à confiabilidade da classificação temática, ou seja, com qual probabilidade as classes temáticas foram atribuídas às áreas mapeadas.

Para tanto, será definido a seguir o método proposto de propagação de incertezas para, no item seguinte, ser apresentado o estudo de caso com os resultados obtidos e, finalmente, serem efetuados a avaliação destes resultados e os comentários finais.

3. Material e Métodos

Para a integração de representações temáticas, as técnicas de análise espacial se restringem aos operadores lógicos da Álgebra Booleana, a menos que as classes temáticas sejam ponderadas e, então, sejam aplicados operadores aritméticos para integrá-las.

Considerando que em cada ponto da área mapeada, para cada representação temática a ser integrada, é atribuída uma certa classe temática e um valor de incerteza expresso pela probabilidade desta classe não ser atribuída ao ponto. O método de propagação estimará a incerteza associada às classes obtidas da integração espacial dependendo do operador lógico escolhido e das regras de interação. Para isto é necessário prever as possíveis combinações entre as classes a serem integradas.

As incertezas das classes obtidas na representação de saída do processo de análise espacial correspondem às incertezas destas classes não ocorrerem em função da probabilidade da classificação incorreta das representações de entrada. Ou então, correspondem à soma das probabilidades de ocorrerem todas as outras combinações entre as classes, com exceção das combinações consideradas pelas regras de integração.

Supondo a integração de duas representações temáticas “A” e “B”, cada qual com três classes temáticas (A₁, A₂, A₃) e (B₁, B₂, B₃), pelo o operador lógico “AND”, que corresponde a uma operação de interseção, as nove possíveis relações entre as classes são:

$$A_1 \cap B_1, A_1 \cap B_2, A_1 \cap B_3, A_2 \cap B_1, A_2 \cap B_2, A_2 \cap B_3, A_3 \cap B_1, A_3 \cap B_2 \text{ e } A_3 \cap B_3.$$

Para cada uma destas relações é atribuída uma classe que define uma das condições intituladas por boa, média ou ruim, que irão compor a representação final como mostra:

	A ₁	A ₂	A ₃		A ₁	A ₂	A ₃
B ₁	boa	média	ruim	B ₁	P(A ₁)*P(B ₁)	P(A ₂)*P(B ₁)	P(A ₃)*P(B ₁)
B ₂	média	média	ruim	B ₂	P(A ₁)*P(B ₂)	P(A ₂)*P(B ₂)	P(A ₃)*P(B ₂)
B ₃	ruim	ruim	ruim	B ₃	P(A ₁)*P(B ₃)	P(A ₂)*P(B ₃)	P(A ₃)*P(B ₃)

Considerando que cada classe tem uma probabilidade de classificação P(I_j) e uma incerteza igual a I(I_j) = 1-P(I_j) as incertezas das classes de saída são determinadas por:

$$I(\text{boa}) = 1 - (P(A_1)*P(B_1)) = (P(A_1)*P(B_2)) + (P(A_1)*P(B_3)) + (P(A_2)*P(B_1)) + (P(A_2)*P(B_2)) + (P(A_2)*P(B_3)) + (P(A_3)*P(B_1)) + (P(A_3)*P(B_2)) + (P(A_3)*P(B_3))$$

$$I(\text{média}) = 1 - [(P(A_1)*P(B_2)) + (P(A_2)*P(B_1)) + (P(A_2)*P(B_2))] = (P(A_1)*P(B_1)) + (P(A_1)*P(B_3)) + (P(A_2)*P(B_3)) + (P(A_3)*P(B_1)) + (P(A_3)*P(B_2)) + (P(A_3)*P(B_3))$$

$$I(\text{ruim}) = 1 - [(P(A_1)*P(B_3)) + (P(A_2)*P(B_3)) + (P(A_3)*P(B_1)) + (P(A_3)*P(B_2)) + (P(A_3)*P(B_3))] = (P(A_1)*P(B_1)) + (P(A_1)*P(B_2)) + (P(A_2)*P(B_1)) + (P(A_2)*P(B_2))$$

Efetuados os cálculos o resultado encontrado para cada ponto é um valor de incerteza associada à classe atribuída a este ponto segundo as regras de integração.

Para os demais operadores lógicos, o procedimento de estimativa da incerteza é o mesmo apresentado, bastando definir as possíveis relações e estabelecer as regras de interação.

4. Estudo de Caso

O estudo de caso trata da modelagem da fertilidade do solo para a cultura de soja no Estado de Santa Catarina. Baseia-se no projeto governamental Zoneamento Pedoclimático do Brasil (Embrapa, 1997), que aborda não só a fertilidade da soja, mas um conjunto de variáveis climáticas e pedológicas para estimativa de um zoneamento pedoclimático para diversas culturas permanentes e temporárias integradas à vocação ecológica de cada região.

Para a determinação da fertilidade, a metodologia considera cinco propriedades químicas: capacidade de troca de cátions, saturação por bases, soma de bases, alumínio trocável e potássio. Estas informações normalmente são encontradas nos relatórios de análise de perfis e amostras de solos.

Foram coletadas aproximadamente 220 análises entre perfis e amostras de solos. Os dados relativos à fertilidade foram classificados de acordo com os critérios apontados na metodologia do zoneamento pedoclimático, como mostra a tabela abaixo:

Tabela 1 – Critérios para a classificação das propriedades químicas segundo a fertilidade

Classes de Fertilidade	Capacidade de troca de cátions (cmolc/kg)	Saturação por bases (%)	Soma de bases (cmolc/kg)	Alumínio trocável (cmolc/kg)	Potássio (mg/kg)
Alta	≥ 8	≥ 50	≥ 4	$< 0,3$	≥ 80
Média	≥ 6 e < 8	≥ 35 e < 50	≥ 2 e < 4	$\geq 0,4$ e < 1	≥ 45 e < 80
Baixa	≥ 4 e < 6	≥ 10 e < 35	< 2	≥ 1 e < 4	< 45
Muito Baixa	< 4	< 10	< 2	≥ 4	< 45

Fonte: CEPA, 1985 citado por Embrapa 1997.

Depois de classificados os perfis e amostras de solo em classes de fertilidade, estas classes foram espacializadas por krigeagem por indicação, que é uma técnica da geoestatística (Felgueiras, 1999).

Na krigeagem por indicação os elementos do conjunto amostral (perfis e amostras), para cada classe, são codificados para os valores 0 ou 1 e depois é feita uma análise de semivariograma, onde são determinados os modelos de semivariograma a serem utilizados e seus parâmetros.

Em seguida, com o modelo definido, cada classe é modelada por algoritmos de krigeagem, tendo como resultado, para cada ponto, a probabilidade de o ponto pertencer a cada classe.

A classe efetivamente atribuída para cada ponto é determinada pelo operador de moda, que determina qual é a classe cuja probabilidade de ocorrência é maior. A incerteza associada é determinada pela probabilidade de não ocorrer tal classe, ou a soma das probabilidades de ocorrer todas as outras classes.

Todo este procedimento de modelagem foi aplicado para as cinco propriedades químicas e o resultado é apresentado na Figura 1.

Em seguida, estas representações foram integradas pelo operador “AND” da Álgebra Booleana, segundo a regra determinada pela metodologia do zoneamento pedoclimático que para cada possível relação de classes, prevalece a classe de maior restrição e para esta, ainda é associada uma classe de aptidão pedológica.

Esta nova classificação respeita o critério de que o manejo da soja no Brasil está sob o nível C, que prevê uso de corretivos e fertilizantes e, portanto, a fertilidade não representa um entrave para o desenvolvimento das culturas (Embrapa, 1997). Desta forma, as classes de fertilidade obtidas da integração: “alta”, “média” e “baixa”, são associadas à classe de aptidão pedológica “boa” e para a classe “muito baixa”, à classe “regular”.

Os quadros abaixo ilustram as regras de integração e reclassificação:

Classes Fertilidade	Alta	Média	Baixa	Muito Baixa
Alta	Alta	Média	Baixa	Muito Baixa
Média	Média	Média	Baixa	Muito Baixa
Baixa	Baixa	Baixa	Baixa	Muito Baixa
Muito Baixa	Muito Baixa	Muito Baixa	Muito Baixa	Muito Baixa

Quadro 1: Regra de integração.

Classes Fertilidade	Alta	Média	Baixa	Muito Baixa
Alta	Boa	Boa	Boa	Regular
Média	Boa	Boa	Boa	Regular
Baixa	Boa	Boa	Boa	Regular
Muito Baixa	Regular	Regular	Regular	Regular

Quadro 2: Classificação para aptidão pedológica.

Para a propagação das incertezas foi aplicado o procedimento descrito no item 3, constando de que as incertezas associadas à classe “boa” são determinadas por 1 – a soma das probabilidades de ocorrer a classe “boa” ou pela soma das probabilidades de ocorrer a classe “regular” e as incertezas associadas à classe “regular” são determinadas pela forma inversa.

Os resultados obtidos para a integração e propagação das incertezas são mostrados na Figura 2.

5. Avaliação dos resultados e conclusão

Analisando os resultados é possível observar um comportamento semelhante em todas as representações de incerteza, o de que os maiores valores se concentram nas zonas de transição das classes temáticas.

Este comportamento é decorrente da modelagem categórica dos atributos, que ignora a variabilidade espacial das variáveis quando supõe a homogeneidade dentro das classes temáticas. É o caso das variáveis relacionadas à fertilidade e também das variáveis relacionadas à maioria dos processos da natureza.

As estimativas das incertezas são essenciais na modelagem dos recursos naturais, não só como indicativo da qualidade e limitações das representações geradas, mas também permitem uma avaliação do conjunto amostral e da metodologia aplicada (Bönisch, 2003).

Com a estimativa das incertezas, em processos de análise espacial de integração de dados, a propagação destas incertezas torna-se imprescindível. O método proposto atende perfeitamente esta determinação, permitindo a estimativa espacial da incerteza associada à representação final, ou seja, um valor de incerteza para cada ponto da área mapeada.

As limitações deste método estão relacionadas ao esforço demandado para estabelecer as possíveis relações entre as classes e a correspondência destas relações com as classes da representação final. Elas aumentam proporcionalmente com o aumento do número de representações consideradas no processo de integração e/ou com o aumento do número de classes temáticas.

Atualmente, não se deveria mais elaborar metodologias de geoprocessamento, principalmente aplicadas aos estudos de propriedades e processos naturais, desconsiderando os procedimentos de avaliação dos produtos gerados, pois uma informação espacial da qualidade agregada às representações espaciais é essencial nos processos de tomada de decisões mais precisas.

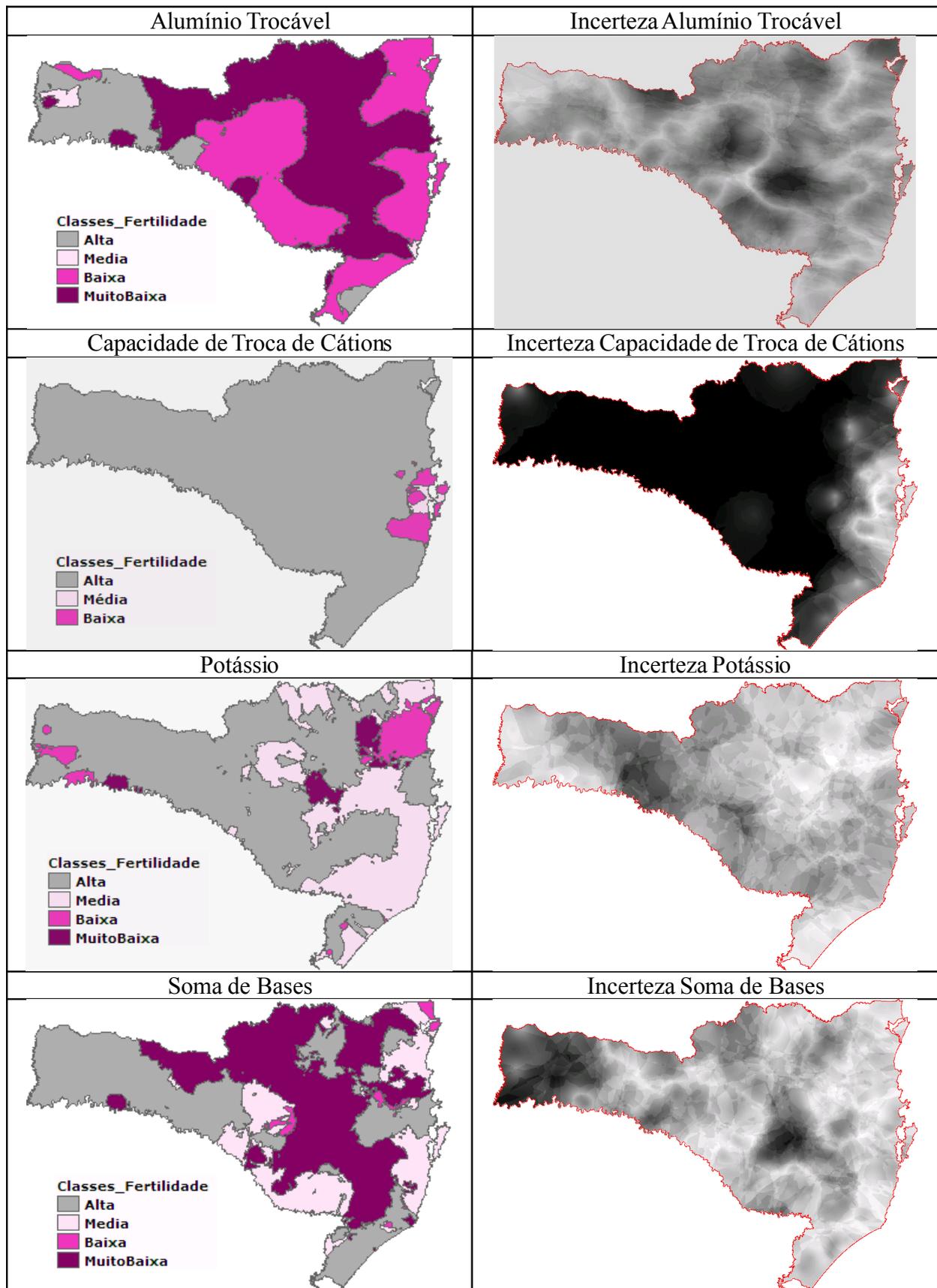


Figura 1 – Espacialização das propriedades químicas que definem a fertilidade e incerteza associada.

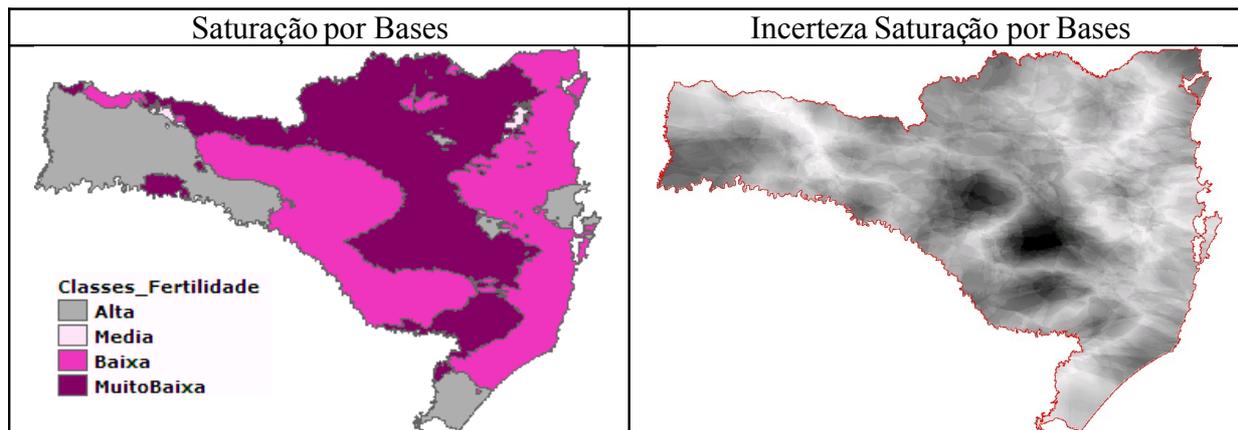


Figura 1 – Continuação

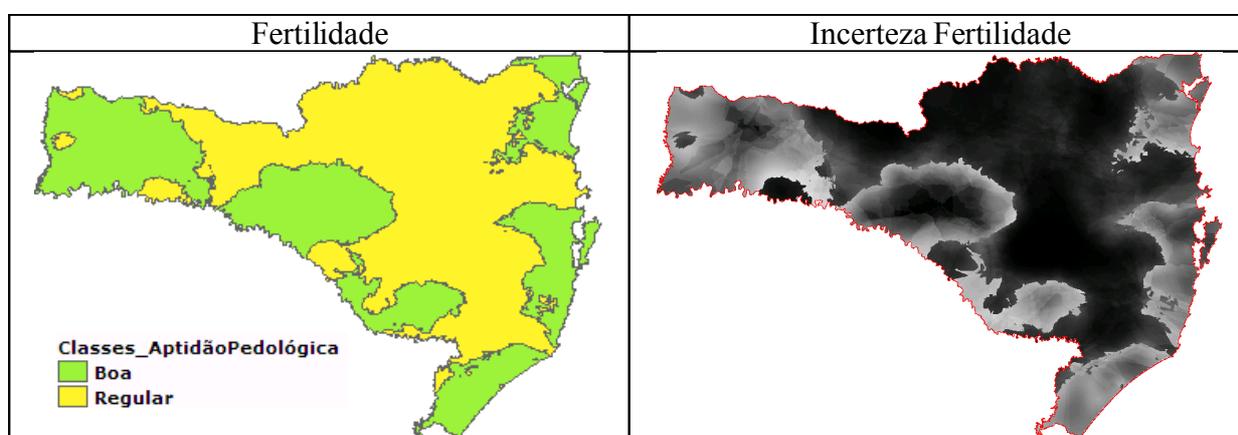


Figura 2 – Fertilidade representada por classes de aptidão pedológica e incerteza associada.

6. Referências bibliográficas

Bönisch, S. **Geoprocessamento ambiental com tratamento de incerteza: o caso do zoneamento pedoclimático para soja no estado de Santa Catarina**. São José dos Campos. 249p. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, 2003.

Burrough, P. A. Natural objects with boundaries. In: Burrough, P. A.; Frank, A. U. ed. **Geographic objects with indeterminate boundaries**. London: Taylor & Francis, 1996, p.3-28.

Burrough, P. A.; McDonnell, R. A. **Principles of geographical information systems**. New York: Oxford University Press, 1998. 344p.

Empresa Brasileira de Pesquisa Agropecuária (Embrapa). **Zoneamento pedoclimático do estado de Santa Catarina**. Rio de Janeiro: Embrapa, 1997. 59p.

Felgueiras, C. A. **Modelagem ambiental com tratamento de incertezas em sistemas de informações geográficas: paradigma geoestatístico por indicação**. São José dos Campos. 181p. Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, 1999.

Heuvelink, G. B. M. **Error propagation in environmental modeling with GIS**. London: Taylor & Francis, 1998. 127p.