

MAPAS SÃO DADOS, NÃO DESENHOS!! (PARTE 2)

Na primeira parte deste artigo (publicada na revista InfoGeo no 5), analisamos o problema de integração de dados em estudos ambientais, e mostramos que as técnicas de Geoprocessamento permitem gerar diferentes cenários, e que o uso de inferência baseada em processos de classificação contínua ("fuzzy") permite resultados melhores que técnicas booleanas. Isto ocorre porque a classificação contínua captura melhor a variabilidade das propriedades ambientais no espaço geográfico. No artigo a seguir, estaremos estendendo esta noção, procurando responder a uma pergunta: como capturar e representar matematicamente as relações entre os componentes do espaço ?

Para ilustrar nossos argumentos, consideramos um problema de formulação simples, que certamente muitos dos leitores já enfrentaram: como interpolar uma superfície a partir de um conjunto de amostras de forma a obter uma melhor aproximação da realidade ? Este conjunto de amostras pode representar dados de altimetria, perfis de solo, dados geofísicos (como anomalias do campo magnético da terra), geológicos (amostras de jazidas) ou ambientais (concentração de poluentes num lago). Assim, trata-se de um tema geral o suficiente para merecer nossa atenção. A figura 1 mostra um caso típico, em que temos amostras de teor de argila, obtidas numa fazenda experimental da EMBRAPA.

Ao resolver este problema, precisamos considerar aquilo que Waldo Tobler chama de "a primeira lei da geografia": *"todas as coisas são parecidas, mas coisas mais próximas se parecem mais que coisas mais distantes"*. O que quer nos dizer este princípio ? que nada na natureza (como na vida real) acontece por acaso. Se encontramos poluição num trecho de um lago, é provável que locais próximos a esta amostra também estejam poluídos. A grande dificuldade está em medir objetivamente este relacionamento.

Voltemos ao nosso problema de interpolação. A técnica mais comum para resolve-lo é simplesmente ajustar uma superfície baseado em interpoladores locais, como o inverso do quadrado da distância: para cada localização do mapa de saída, consideram-se as amostras mais próximas e o novo ponto será obtido a partir da média destas amostras, ponderado pelo inverso da distância. O resultado, apesar de representar uma primeira aproximação ao que queremos, ainda possui muitas limitações. A variabilidade espacial é ignorada, e numa área com poucas amostras, poderemos estar utilizando para interpolação amostras que já não tem nenhuma influência sobre o ponto considerado. Estes métodos desconsideram ainda o fato que o fenômeno pode variar de forma diferente em direções distintas do espaço (a *anisotropia*).

Uma abordagem mais adequada para este problema é a chamada *geo-estatística*. Não por acaso, a geo-estatística nasceu na África do Sul, onde um geólogo chamado Daniel Krige buscava melhorar o mapeamento de minas

de diamantes (donde o termo *krigeagem*, aplicado ao processo de interpolação associado). Krige teve a idéia de modelar a distribuição do "parentesco" entre as amostras. Sua idéia foi mais tarde formalizada pelo matemático francês Matheron, que estabeleceu algumas hipóteses de base para o correto funcionamento desta técnica: as amostras devem ser parte de um conjunto homogêneo, com uma média constante e com uma variância que só depende da distância entre elas (em termos matemáticos, trata-se de uma distribuição *estacionária em segunda ordem*). Para os não familiarizados com os termos da estatística, considerem que a variância entre duas amostras mede a "influência mútua". Por exemplo, quando duas amostras apresentam variância conjunta positiva, tanto mais o aumento da primeira implica no crescimento da segunda.

Para o caso dos conjuntos de dados que satisfaçam as hipóteses da geoestatística (e que são surpreendentemente muitos) uma boa notícia: é possível modelar estatisticamente a relação entre estes dados e obter uma superfície interpolada muito mais próxima da realidade. A modelagem estatística é obtida através de um processo chamado *variografia*, que indica precisamente o grau de influência mútua entre as amostras. A partir desta modelagem, é possível definir um interpolador (a *krigeagem*), bem mais acurado que o método empírico de inverso do quadrado da distância.

Para ilustrar, considere as figura 2 e 3, que apresentam uma visualização comparativa dos resultados obtidos pelo método do inverso do quadrado da distância (fig. 2) e pela *krigeagem* (fig. 3). A simples visualização dos dados mostra a existência de "artefatos" (pontos extremos falsos) na primeira interpolação e um comportamento mais próximo ao esperado no mapa produzido pelas técnicas de geoestatística.

Deste modo, a integração de ferramentas de geoestatística a sistemas de informação geográfica oferece um ambiente muito poderoso e adequado para a representação da realidade. Maiores detalhes poder ser encontrados na webpage "AI-GEOSTATS" (<http://curie.ei.jrc.it/ai-geostats.htm>), que apresenta um grande quantidade de informações, e na tese de Eduardo Camargo, do INPE (<http://www.dpi.inpe.br/teses/eduardo>), que descreve a integração de procedimentos geoestatísticos ao SPRING.

Em resumo, nesta série de artigos, estamos procurando convencer nossos leitores que o tratamento matemático-estatístico compõe um conjunto imprescindível de técnicas para lidar com a geo-informação e que os resultados compensam amplamente o esforço investido em seu aprendizado.

O autor agradece a Eduardo Camargo, Carlos Felgueiras e Antônio Miguel Monteiro (INPE) e Suzana Fuks (EMBRAPA), a cessão dos exemplos e as discussões sobre o tema. Para os leitores com conhecimentos de estatística, minhas desculpas: as simplificações do artigo são necessárias para um público mais amplo.

Gilberto Câmara é coordenador do programa de P&D de Geoprocessamento do INPE e foi um dos responsáveis pelo desenvolvimento do Spring (homepage: www.dpi.inpe.br/~gilberto).

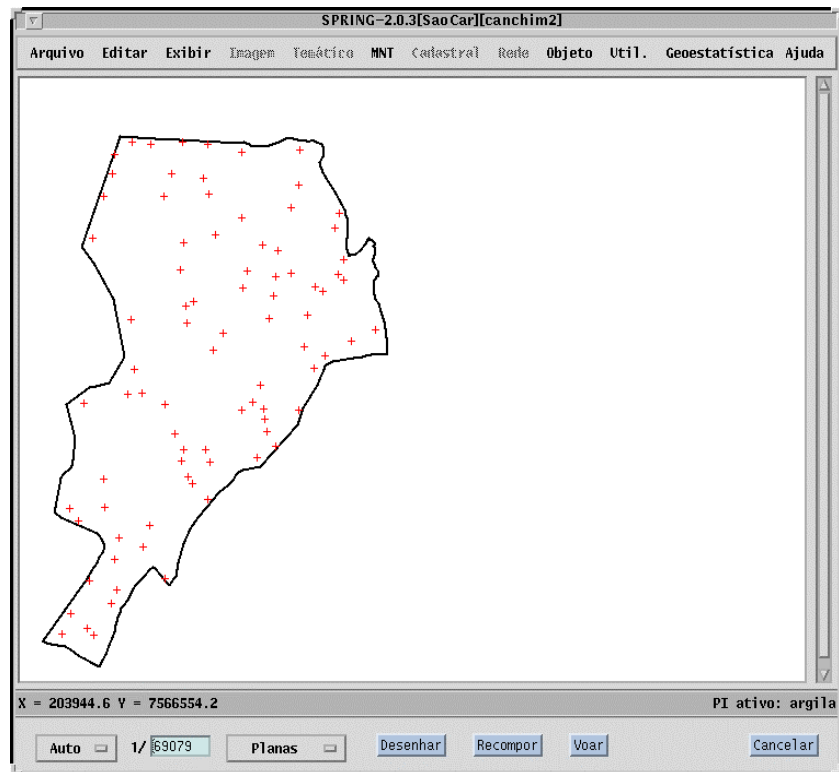


Figura 1 - Amostras de teor de argila (fonte: EMBRAPA-Solos).

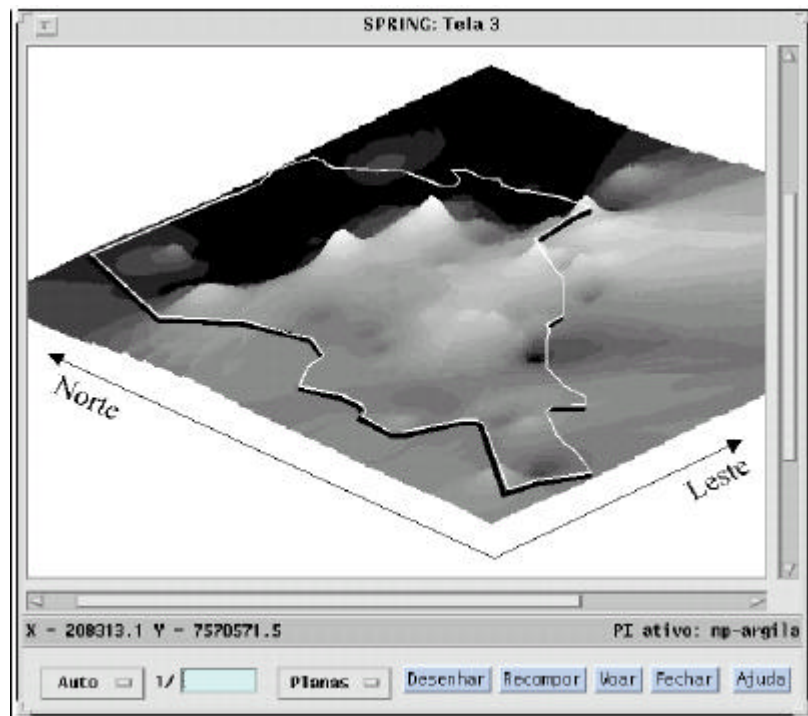


Figura 2 - Superfície interpolada a partir das amostras pelo método do inverso do quadrado da distância (fonte: Eduardo Camargo, INPE)

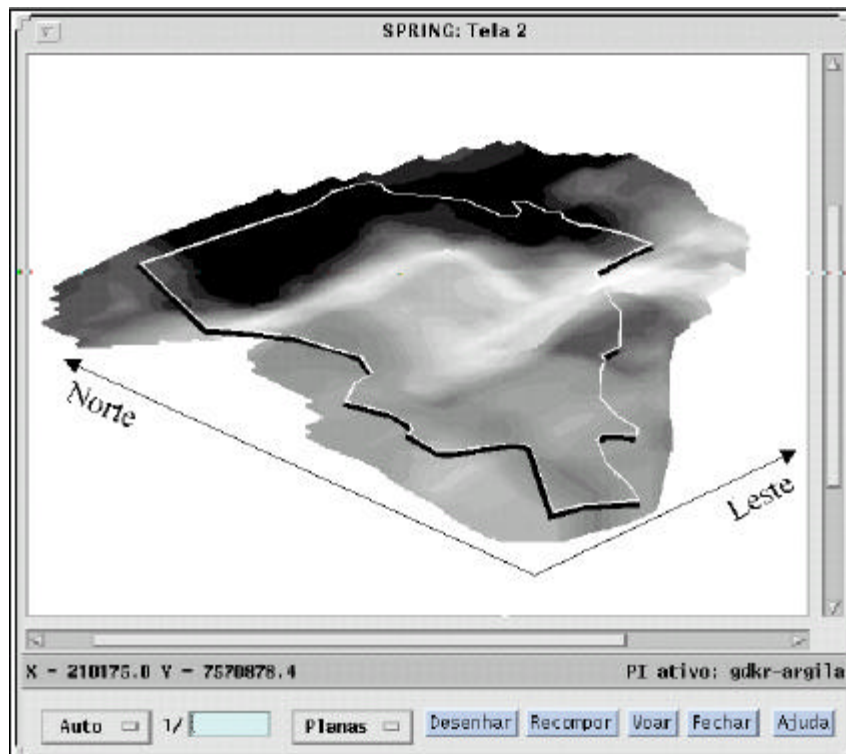


Figura 3 - Superfície interpolada a partir das amostras pela técnica de geoestatística (fonte: Eduardo Camargo, INPE).