

Uma retrospectiva da modelagem em solos

Luciano Campos Cancian¹

¹Affiliation not available

September 14, 2022

Este texto é uma proposta incentivada na disciplina de Modelagem Geoestatística do Ambiente, ministrada no Programa de Pós Graduação em Ciência do Solo da Universidade Federal de Santa Maria. Ele é desenvolvido com base nos conhecimentos adquiridos a cada aula da disciplina, sendo movido e motivado por um trecho extraído do livro *Geostatistics for Natural Resources Evaluation*, de Pierre Goovaerts (1997, p.442):

“[...] beware that uncertainty is not intrinsic to the phenomenon under study: rather it arises from our imperfect knowledge of that phenomenon, it is data-dependent and most importantly model-dependent, that model specifying our prior concept (decisions) about the phenomenon. No model, hence no uncertainty measure, can ever be objective: the point is to accept that limitation and document clearly all aspects of the model.”

Isso nos permite refletir sobre os modelos usados para estudar os fenômenos espaciais de interesse na ciência do solo, considerando aspectos relacionados tanto ao surgimento como aos rumos tomados por esse campo ao longo da história.

Não é de hoje que sabemos da importância de obter informações sobre o solo. Contudo, a necessidade de que os usuários tenham uma forma mais dinâmica de visualizar tais informações instigou – e ainda instiga – muitos pesquisadores a desenvolver formas para que isso se tornasse possível. Uma das formas mais práticas e interativas é a criação de uma forma visual por meio da criação de um mapa, como apresentado na Figura 1, tornando a informação mais intuitiva ao usuário final, que tem nele a variação das informações na paisagem. O objetivo do mapeamento é permitir que possamos dizer algo preciso sobre partes específicas da área e que não poderíamos dizer sobre o todo (WEBSTER and BECKETT, 1968).

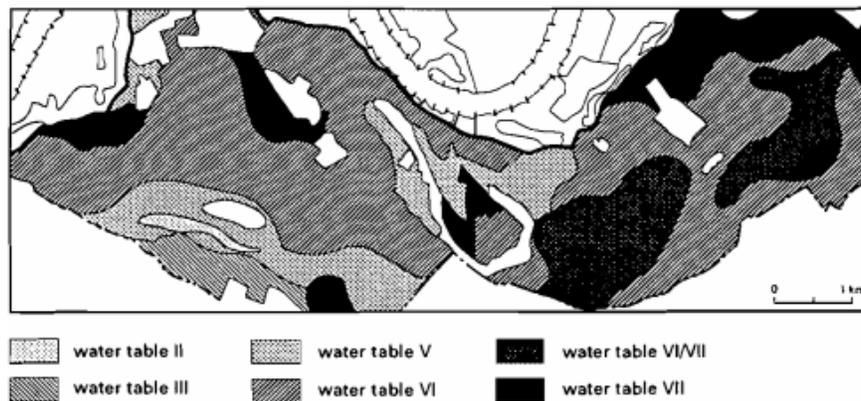


Figure 2. Map of water table classes in the Ooypolder.

Figure 1: Map of water table classes in the Ooypolder (HEUVELINK, 1996)

Não é de hoje também a intenção de avaliar a qualidade das informações contidas nos mapas. As informações contidas nestes mapas podem conter diversos erros embutidos, o que tornaria tais informações imprecisas – ou até inúteis. Conforme Webster e Beckett (1968), a avaliação é necessária para que possam ser feitas afirmações precisas sobre as informações do solo dentro das unidades mapeadas e que as informações ali presentes possam ser utilizadas representando a realidade do local com a maior acurácia possível.

Nos primeiros estudos abordando o mapeamento, as medidas sobre a qualidade dos mapas e os possíveis erros neles existentes eram muito mais subjetivas e intrínsecas ao pesquisador do que propriamente avaliadas por testes, ao passo que métodos estatísticos eram menos difundidos dentro das pesquisas em ciência do solo, principalmente na Pedologia (WEBSTER and BECKETT, 1968). A avaliação dos resultados também seguia o mesmo caminho: o usuário final precisava se basear na opinião do pesquisador sobre seu mapa, por mais “pobres” - em termos de informação - que esta fosse, devido a menor disponibilidade ou difusão de avaliações estatísticas. Mesmo com essa menor disponibilidade, o pesquisador buscava avaliar o mapa, mesmo que sem maior refinamento, por algum método estatístico que permitisse o permitisse afirmar que seus dados condiziam com a realidade e que representavam a variação local.

Muitos métodos tiveram de esperar o desenvolvimento de computadores para serem utilizados intensamente (Heuvelink and Webster, 2001). Com o desenvolvimento e avanço de ferramentas computacionais, tornou-se possível a programação e automação de métodos, permitindo um maior avanço na disponibilização e criação de métodos de avaliação. Mesmo neste cenário, alguns destes métodos não se mostravam tão usuais e recomendáveis pela grande demanda de processamento de algumas das técnicas, principalmente em bancos de dados com maior número de informações. Isso demonstrava que ainda era grande o caminho para que estes métodos fossem explorados com eficiência. Mesmo assim, o avanço trazido pela popularização de sistemas computacionais facilitou a análise e levou, cada vez mais, os métodos para longe da subjetividade, permitindo assim uma avaliação mais criteriosa da sua qualidade.

Com o foco voltado ao desenvolvimento não só de computadores com maior capacidade de processamento, mas também a criação e ao aprimoramento de softwares e modelos que processassem e gerassem dados com maior velocidade, visualizava-se um número cada vez maior de adeptos e entusiastas com interesse nesta linha de pesquisa. A modelagem em solos pode ser definida como o “desenvolvimento de modelos matemáticos que tentam simular, da melhor forma possível, processos que ocorrem no solo (Vereecken et al., 2016), considerando suas variações a partir da utilização de dados obtidos do solo.

Sobre a modelagem de processos do solo, cabe ressaltar que, pelo fato de não podermos medir o solo em todos os locais (o que seria um censo), qualquer informação obtida em locais não amostrados envolve previsão, visto que a variação no solo complexa ao ponto de nenhuma descrição dele ser completa (Heuvelink and Webster, 2001). A criação de modelos buscando a geração de mapas com dados mais exatos não só possibilitou que os erros fossem avaliados de melhor forma, mas também que permitisse ao pesquisador verificar o efeito que diferentes modelos com novas abordagens possuíam em variações – por exemplo - amostrais permitindo avaliar se um maior número de amostras implicaria em uma melhor predição e redução de erros. Isso possibilita que nós pesquisadores, baseados em experiências documentadas, tomemos certos caminhos ao seguir linhas de pesquisa que utilizem modelos eficientes ou rumar para outras que tragam melhores resultados para nossos propósitos científicos.

Um passo importante que devemos considerar é a introdução da geoestatística para modelar a variação espacial. Ela assume implicitamente que a semivariância em uma área depende apenas da distância de separação entre dois pontos, e não das suas posições. De modo geral, assume-se que os pontos próximos são mais parecidos uns com os outros em média do que os mais distantes (Heuvelink and Webster, 2001). Outro ponto que considero importante na evolução dos modelos de predição é a busca por novos modelos e alternativas – que muitas vezes são derivações de modelos já utilizados – que consigam se adaptar melhor também aos dados disponíveis. Um exemplo disso é o modelo apresentado por HEUVELINK (1996), onde a união de um modelo discreto à outro contínuo, proporcionou ao modelo misto melhores estimativas e menores valores de erro. Embora insira melhorias nas previsões sobre o solo, é importante termos em mente que a mistura entre esses métodos, pela maior “flexibilidade”, pode fazer com que perca capacidade preditiva

em certas situações em detrimento de ganho em outras. O fato de que diferentes fontes de dados permitem que modelos transmitam maiores ou menores erros ao resultado final força essa busca, fazendo com que nós pesquisadores optemos por modelos que, além de gerar bons resultados, também tragam valores de erro menores.

A utilização de modelos de variação espacial mais elaborados pode ser recomendável em situações particulares, embora se deva notar que modelos mais complexos contêm mais parâmetros que, de algum modo, precisam ser estimados a partir dos dados obtidos, ficando assim sujeitos a mais erros (HEUVELINK, 1996). Isso nos permite afirmar que apesar de ser mais complexo, podendo resultar em melhores estimativas, nem sempre essa maior complexidade trará benefícios aos dados de interesse. Contudo, o surgimento de métodos como o modelo linear misto de variação espacial, embora “mais complexo” por possuir mais dados de entrada, proporcionou formas de integrar a modelagem estocástica com a compreensão dos processos do solo, favorecendo a modelagem do solo. A grande vantagem desta formulação é que foi possível propor estimadores baseados em probabilidade para os parâmetros dos efeitos aleatórios (Lark, 2012), tornando-se apropriado por minimizar o viés nas estimativas devido à incerteza nos coeficientes de efeitos fixos estimados serem conhecidos. Incorporar dados preditores como efeitos fixos permite que a variação restante nas propriedades de interesse do solo seja modelada como efeitos aleatórios espacialmente dependentes, assimilando processos do solo na forma de dados.

Porém, apesar de todo o desenvolvimento que observamos nos últimos anos, quando consideramos a avaliação de erros – e de formas para sua identificação e quantificação, acredito que esta área ainda seja “nebulosa” para muitos pesquisadores, inclusive para mim. Os resultados de uma análise de erros e sua propagação nos modelos só serão realmente úteis quando os erros embutidos nos dados forem identificados corretamente. Entretanto, esta é uma condição difícil de ser atendida, visto que na maioria dos casos há apenas estimativas brutas de erro dos dados. Essa falta de informação – e porque não de conhecimento de nós pesquisadores? - é possivelmente a principal razão de as análises de propagação de erros ainda serem uma exceção e não uma regra nos trabalhos de modelagem e mapeamento. De qualquer forma, devemos ter em mente o exposto por Heuvelink and Webster (2001): “As pesquisas passadas, presentes e futuras foram, são e serão direcionadas ao desenvolvimento de modelos que diminuam os erros”. Esse é o caminho para tornarmos a pesquisa em Pedometria e Mapeamento Digital de Solos mais exata.

References

- HEUVELINK. Identification of field attribute error under different models of spatial variation. *International Journal of Geographical Information Systems*, 10, 1996.
- Heuvelink and Webster. Modelling soil variation: past, present, and future. *Geoderma*, 100, 2001.
- R.M. Lark. Towards soil geostatistics. *Spatial Statistics*, 1:92–99, may 2012. doi: 10.1016/j.spasta.2012.02.001. URL <https://doi.org/10.1016%2Fj.spasta.2012.02.001>.
- H. Vereecken, A. Schnepf, J.W. Hopmans, M. Javaux, D. Or, T. Roose, J. Vanderborght, M.H. Young, W. Amelung, M. Aitkenhead, S.D. Allison, S. Assouline, P. Baveye, M. Berli, N. Brüggemann, P. Finke, M. Flury, T. Gaiser, G. Govers, T. Ghezzehei, P. Hallett, H.J. Hendricks Franssen, J. Heppell, R. Horn, J.A. Huisman, D. Jacques, F. Jonard, S. Kollet, F. Lafolie, K. Lamorski, D. Leitner, A. McBratney, B. Minasny, C. Montzka, W. Nowak, Y. Pachepsky, J. Padarian, N. Romano, K. Roth, Y. Rothfuss, E.C. Rowe, A. Schwen, J. Šimůnek, A. Tiktak, J. Van Dam, S.E.A.T.M. van der Zee, H.J. Vogel, J.A. Vrugt, T. Wöhling, and I.M. Young. Modeling Soil Processes: Review Key Challenges, and New Perspectives. *Vadose Zone Journal*, 15(5):0, 2016. doi: 10.2136/vzj2015.09.0131. URL <https://doi.org/10.2136%2Fvzj2015.09.0131>.
- WEBSTER and BECKETT. Is soil variation random? *Geoderma*, 97, 1968.