



Variáveis preditivas para padrões distintos de relevo em mapeamento pedológico digital ⁽¹⁾

Laura Milani da Silva Dias⁽²⁾; Ricardo Marques Coelho⁽³⁾; Gustavo Souza Valladares⁽⁴⁾.

⁽¹⁾ Com apoio Petrobrás (Rede de Sistemas de Produção – PBIO)

⁽²⁾ Mestre em Agricultura Tropical e Subtropical, Instituto Agronômico, Campinas, São Paulo; laurads5@yahoo.com.br; ⁽³⁾ Pesquisador; Instituto Agronômico; ⁽⁴⁾ Professor Doutor; Universidade Federal do Piauí.

RESUMO: O relevo é elemento determinante para a variabilidade dos solos na paisagem. Tipo e intensidade dos processos pedogenéticos variam com o relevo, o que pode influir no poder preditivo de variáveis usadas em mapeamento pedológico digital. Neste trabalho foram avaliadas variáveis preditivas de relevo, geologia e sensor remoto orbital em situações distintas de relevo em área da Bacia Sedimentar do São Francisco. A distribuição de atributos do relevo, litologia e índices de sensor remoto determinou a localização dos pontos de observação de solos. Combinações destes atributos compuseram diferentes matrizes de dados de atributos preditivos com classes de solo em dois níveis taxonômicos. A predição das classes de solos por algoritmos de mineração de dados foi validada em campo dividindo-se o mapeamento em áreas planas e não planas. Os resultados mostraram superioridade das variáveis de relevo na predição das classes de solo em áreas não planas. Com o uso em conjunto do relevo e da litologia o resultado se inverte, com as observações localizadas em áreas aplanadas preditas com maior acerto. Quando reunidas todas as variáveis preditivas, a predição nas áreas de maior ondulação volta a ter maior acerto. Os resultados se repetiram nos dois níveis taxonômicos. Isso enfatizou a importância da geologia e do relevo como elementos preditivos da distribuição dos solos, mas não deu suporte ao uso dos índices de sensor remoto. O baixo poder dos índices de sensoriamento em conjunto deve-se à inclusão do índice de vegetação, sensível à diversificada cobertura vegetal e baixa proporção de solo exposto.

Termos de indexação: mapeamento digital de solos, predição de classes de solos, NDVI.

INTRODUÇÃO

O relevo é tido como fator que modifica a distribuição dos processos hidrológicos e erosivos influenciando na formação dos horizontes e dos atributos do solo. Por isso, diferentes tipos de solos se formam em diferentes posições do relevo e em distintos segmentos da vertente (Moore et al., 1991; Moore et al., 1993).

O mapeamento digital de solos usa sistemas de informação espacial e modelos numéricos aplicados a informações de solo pré-existent e variáveis ambientais correlacionadas para inferir variações espaciais de classes e propriedades de solos (Lagacherie & McBratney, 2007). Dentre as variáveis ambientais, as do relevo têm tido o uso mais frequente nos trabalhos de predição de classes de solo em mapeamentos digitais no País (Ten Caten et al., 2012), provavelmente facilitado pelo acesso a modelos digitais de elevação (MDE) e à quantidade de variáveis geomorfológicas que podem ser derivadas em sistemas de informações geográficas (SIG).

No entanto, em áreas aplanadas em que variáveis do relevo têm pequena diferenciação, variáveis derivadas de sensores remotos podem ser mais eficientes na predição de classes de solo (Kingham and McKane, 1998). Dadas as características locais, em que 86% da área total do estudo foi classificada como de relevo plano e suave ondulado sugerindo pequena diferenciação dos solos, presumiu-se que os atributos derivados de razões de bandas de uma imagem de sensor remoto orbital do satélite Landsat 8 pudessem destacar informações sobre a resposta espectral das diferentes classes de solo e por isso fossem mais eficientes que outros na predição nestas condições.

O objetivo do trabalho foi validar mapas pedológicos digitais elaborados por predições de algoritmos de mineração de dados, a fim de verificar quais atributos utilizados têm maior capacidade para discriminar e prever as classes de solo em duas condições distintas de relevo.

MATERIAL E MÉTODOS

A área estudada localiza-se no norte do Estado de Minas Gerais, mesorregião de Montes Claros, e se estende por 110.289 ha, com seu centro no município de Capitão Enéas, MG. O clima é do tipo Aw, tropical de savana com estação seca de inverno. Encontra-se na Bacia Sedimentar do Rio São Francisco, onde há predomínio de superfícies aplanadas (Leite e Brito, 2012) e a litologia, compilada do mapa geológico em escala 1.100.000



(CPRM, 2011) e do mapa metalogenético em escala 1:250.000 (CPRM, 1985), ambos da folha Montes Claros, mostra predomínio de rochas pelíticas e metapelíticas.

A partir de dados de hipsometria SRTM com resolução espacial de 90 metros foi gerado um modelo digital de elevação (MDE) e derivadas as variáveis de relevo altitude, declividade, curvatura planar, curvatura em perfil, orientação das vertentes, distância da drenagem e índice topográfico de umidade (Wilson e Gallant, 2000).

A metodologia desenvolvida para o sensor TM do Landsat (Sabins Junior, 1977) foi adaptada para dados do sensor OLI do Landsat 8 para derivar os índices *Clay minerals* (CMI) e *Iron Oxide* (IOI). A eles foi acrescentado o *Normalized Difference Vegetation Index* (NDVI).

Para obtenção das informações dos solos, as variáveis geradas foram inseridas no programa de amostragem aleatório estratificada (cLHS) como condicionantes à seleção de 300 pontos (Wyss & Jorgensen, 1998). Nos locais selecionados, o solo foi caracterizado morfologicamente e amostras coletadas para realização de análises químicas de rotina pedológica e granulométricas. O solo foi classificado no Sistema Brasileiro de Classificação de Solos até 4º nível categórico mais agrupamento textural (EMBRAPA, 2006), utilizando-se também a generalização desta para o 2º nível categórico.

Diferentes combinações entre as variáveis do relevo, os índices derivados da imagem Landsat 8 e a litologia compilada dos mapas geológico e metalogenético foram utilizadas como preditivas das classes de solo nos dois níveis categóricos e compuseram as matrizes de dados (Tabelas 1 e 2). As matrizes foram submetidas a procedimentos de pré-processamento dos dados e a uma modelagem com algoritmos de mineração de dados de árvores de decisão e de redes neurais (*Random Forest*, MLP e J48) no WEKA (Witten & Frank, 2005). Para cada matriz avaliada foi eleito e aplicado o modelo de maior acurácia para predição das classes de solo faltantes e posteriormente foram gerados e validados os mapas pedológicos digitais.

A validação de cada um dos mapas foi realizada por meio de 62 pontos classificados em campo, divididos em *Áreas Planas* (28 pontos), com menos de 3% de declividade, e *Áreas Não Planas* (34 pontos), com 3% ou mais de declividade. Os dados foram confrontados com os mapas preditos por meio de matrizes de erro.

RESULTADOS E DISCUSSÃO

Relevo plano e suave ondulado ocupa 86% da área de estudo. Isto sugere pequena diferenciação

dos solos e permite presumir que os atributos derivados de razões de bandas de imagens do sensor remoto orbital Landsat 8 possam destacar a resposta espectral das diferentes classes de solo e sejam mais preditivas que outras variáveis nestas condições.

Para as classes de solo no 4º nível categórico mais textura, os atributos de relevo promoveram maior acurácia na predição em Áreas Não Planas que em Áreas Planas. O mesmo deu-se na predição de classes de solo no 2º nível categórico, mostrando a sensibilidade destes atributos aumenta com a diferenciação do relevo, quando são considerados exclusivamente estes atributos preditivos (Tabela 1).

Incluída a litologia como variável preditiva nas matrizes com classe de solo no 4º nível, esta se mostrou determinante em áreas mais suavizadas ao produzir maior acurácia nas Áreas Planas, mas com pequeno ganho de acurácia nas áreas Não Planas (Tabela 1). De fato, os mapas base de geologia para a área indicam a ocorrência de distintas litologias nas áreas aplanadas: (i) depósitos aluviais, (ii) depósitos colúvio-eluviais, (iii) metassiltitos e folhelhos, o que contribuiu para a discriminação das classes de solo com uso da base de dados de geologia.

Quando considerado o conjunto completo de variáveis (relevo, geologia, índices de sensores), com pequena diferença (2,3%), as "Áreas Não Planas" voltam a apresentar melhor desempenho (55,9%) (Tabela 1) contrapondo-se à expectativa de aumento da acurácia nas áreas planas com o uso dos índices de sensores.

No 4º nível categórico, a maior acurácia de predição em Áreas Planas se deu com o uso conjunto das variáveis geomorfométricas e litologia (57,1%), mostrando que os índices derivados de sensor remoto foram pouco preditivos nestas condições de relevo (Tabela 1). Por outro lado, para as áreas de maior ondulação, o conjunto completo de variáveis apresentou o melhor desempenho (55,9%), sustentando o uso de todas elas na predição das classes de solo em relevos mais declivosos (Tabela 1).

As matrizes de erro com classes de solo no segundo nível categórico mantiveram as tendências da avaliação com as classes de solo no 4º nível (Tabela 2). A litologia aumenta o poder preditivo em relação ao uso unicamente de variáveis do relevo nas Áreas Planas, com pequena redução (3,5%) de acurácia com o uso dos índices do sensor nestas áreas. Nas Áreas Não Planas, com 61,8% de acurácia com uso exclusivo de atributos do relevo, o uso da litologia reduz a acurácia em 14,3%, acurácia recuperada em 11,3% a partir da inclusão dos índices de sensoriamento.



Dois resultados contrariam a expectativa inicial para as Áreas Não Planas, nos dois níveis taxonômicos: (i) a redução da acurácia quando da adição da litologia ao relevo como elemento preditivo e (ii) a elevação da acurácia dessa matriz (relevo + litologia) quando são incluídos os índices de sensores, ao contrário das Áreas Planas. Apesar da classe de relevo plano estar distribuída por toda a área mapeada, há predomínio destas áreas na porção norte da área do estudo. De forma semelhante, apesar da distribuição das diferentes litologias por toda a área do estudo, a base geológica mostra polígonos maiores de litologias com predomínio de relevo plano na porção norte, especialmente depósitos aluvionares e depósitos colúvio-eluviais, mas também metassiltitos e folhelhos, estes últimos com proporção um pouco maior de declividades superiores a 3%. Na porção sul da área mapeada, com menor proporção de relevo plano, há menos polígono de litologias distintas, o que reduz o poder discriminatório da variável litologia, e estes polígonos são menores, com possível redução na acurácia da predição por dificultar a precisão cartográfica. Também na porção sul, com menor proporção de relevo plano, pode-se esperar maior poder preditivo de variáveis derivadas do relevo.

Para o efeito positivo das variáveis derivadas do sensoriamento apenas nas Áreas Não Planas, não foi encontrada justificativa aparente. O baixo poder preditivo das variáveis derivadas do sensoriamento orbital nas áreas planas, contrariando a expectativa, pode estar relacionada à escolha dos índices de sensoriamento. A sensibilidade dos índices *clay minerals* e *iron oxides*, específicos para determinações em solos, pode ter sido neutralizada ou confundida pelo índice NDVI, sensível a variações na cobertura vegetal, mas nem sempre sensível a variação nas propriedades espectrais dos solos, especialmente quando há um mosaico diversificado de uso e cobertura do solo. Como a cobertura vegetal em toda a área é heterogênea, especialmente nas áreas mais planas ao norte, com maior aproveitamento agrícola e fragmentação de usos da terra, isso pode ter promovido variação no NDVI desvinculada e dispare das variações dos índices de solo.

CONCLUSÕES

Aspectos metodológicos influenciaram os resultados. A escolha dos índices de sensoriamento, misturando índices mais sensíveis às resposta espectrais dos solos com índice mais sensível à resposta espectral da vegetação, afetou

negativamente o poder preditivo dos índices derivados do sensoriamento remoto orbital.

O baixo poder preditivo da variável litologia em relação ao relevo nas áreas mais onduladas e com menor variabilidade litológica mapeada, mostra que, apesar de geologia ser determinante para a distribuição e variabilidade dos solos, seu uso deve estar associado à disponibilidade de mapas em escala compatível com o mapeamento pedológico, especialmente no caso de baixa variabilidade geológica mapeada.

REFERÊNCIAS

- KINGHAM, R.; McKANE, D. Angledool radiometric study. In: Bui, E. (ed.) A Soil Information Strategy for the Murray-Darling Basin (MDBSIS). CSIRO Land and Water. Final Report 1998. p.109-115. (<http://catalogue.nla.gov.au/Record/3082617>). (Acesso em 05/06/2015)
- EMBRAPA. Sistema Brasileiro de Classificação de Solos. 2.ed. Brasília: Embrapa Produção de Informação; Rio de Janeiro: Embrapa Solos, 2006. 312p
- LAGACHERIE, P.; MCBRATNEY, A. B. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: LAGACHERIE, P. et al. Digital soil mapping: an introductory perspective. Amsterdam: Elsevier, 2007, Cap. 1, p. 3-22
- LANDIS, J. R. & KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, v.33, p.159-174, 1977.
- LEITE, M.R.; BRITO, J.L.S. Mapeamento morfoestrutural e morfoescultural na região de cerrado no Norte de Minas Gerais. *Soc. & Nat.*, Uberlândia, ano 24 n. 1, 115-126, jan/abr. 2012
- MAPA GEOLÓGICO. Folha Montes Claros - SE-23-X-A-VI. Escala 1.100.000. CPRM, 2011
- MAPA METALOGÊNICO. Folha Montes Claros - SE-23-X-A. Escala 1: 250.00. CPRM, 1985.
- MOORE, I. D., GRAYSON, R. B., LADSON, A. R. Digital terrain modeling – a review of hydrological, geomorphological, and biological applications. *Hydrology Process*. v. 5, n.1, p. 3 - 30, 1991.
- MOORE, I. D.; GESSLER, P. E.; NIELSEN, G. A.; PETERSON G. A. Soil attribute prediction using terrain analysis. *Soil Science Society American Journal*, Madison, v. 57, p. 443 - 452, 1993.
- SABINS, F. F. Remote Sensing: Principles and Interpretation. 3rd ed. New York: W. H. Freeman and Company, 1997. 432p.



TEN CATEN, A; DALMOLIN, R. S. D; MENDONÇA-SANTOS, M. L; GIASSON, E. Mapeamento digital de classes de solos: características da abordagem brasileira. *Ciência Rural*, v.43, n.11, p. 1989 -1997, nov, 2012.

WILSON, J. P.; GALLANT, J. C. (Ed.) *Terrain Analysis: Principles and Applications*. New York, John Wiley & Sons, Inc., 2000, 479 pp.

WITTEN, I.H., FRANK, E. *Data mining: practical machine learning tools and techniques*. 2nd edition. San Francisco: Morgan Kaufmann; 2005.

WYSS, G.D. & JORGENSEN, K.H. *A user's guide to LHS: Sandia's Latin Hypercube Sampling Software*. Albuquerque, Sandia National Laboratories, 1998. 138p. (SAND98-0210)

Tabela 1. Acurácia (%) das matrizes em duas situações de relevo e solo no segundo nível categórico.

Matrizes	Áreas Planas	Áreas Não Planas
Relevo	42,8	47,0
Relevo e litologia	57,1	50,0
Relevo, litologia e índices de sensor remoto	53,6	55,9

Tabela 2. Acurácia (%) das matrizes em duas situações de relevo e solo no segundo nível categórico.

Matrizes	Áreas Planas	Áreas Não Planas
Relevo	50,0	61,8
Relevo e litologia	60,7	47,5
Relevo, litologia e índices de sensor remoto	57,1	58,8

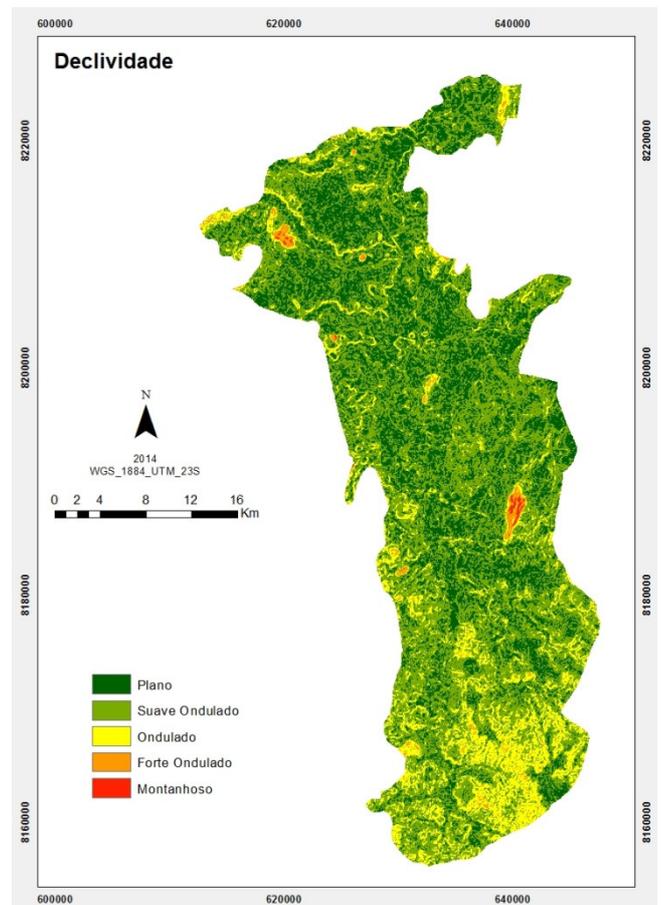


Figura 1. Classes de declividade da área de estudo