

Atributos topográficos e dados do Landsat7 no mapeamento digital de solos com uso de redes neurais

César da Silva Chagas⁽¹⁾, Elpídio Inácio Fernandes Filho⁽²⁾, Carlos Antônio Oliveira Vieira⁽³⁾, Carlos Ernesto Gonçalves Reynaud Schaefer⁽²⁾ e Waldir de Carvalho Júnior⁽¹⁾

⁽¹⁾Embrapa Solos, Rua Jardim Botânico 1.024, Jardim Botânico, CEP 22460-000 Rio Janeiro, RJ. E-mail: chagas.rj@gmail.com, waldircj@cnpq.embrapa.br ⁽²⁾Universidade Federal de Viçosa (UFV), Departamento de Solos, Campus Universitário, CEP 36570-000 Viçosa, MG. E-mail: elpidio@ufv.br, carlos.schaefer@ufv.br ⁽³⁾UFV, Departamento de Engenharia Civil, Campus Universitário, CEP 36570-000 Viçosa, MG. E-mail: carlos.vieira@cfn.ufsc.br

Resumo – O objetivo deste trabalho foi avaliar variáveis discriminantes no mapeamento digital de solos com uso de redes neurais artificiais. Os atributos topográficos elevação, declividade, aspecto, plano de curvatura e índice topográfico, derivados de um modelo digital de elevação, e os índices de minerais de argila, óxido de ferro e vegetação por diferença normalizada, derivados de uma imagem do Landsat7, foram combinados e avaliados quanto à capacidade de discriminação dos solos de uma área no noroeste do Estado do Rio de Janeiro. Foram utilizados o simulador de redes neurais Java e o algoritmo de aprendizado “backpropagation”. Os mapas gerados por cada um dos seis conjuntos de variáveis testados foram comparados com pontos de referência, para a determinação da exatidão das classificações. Esta comparação mostrou que o mapa produzido com a utilização de todas as variáveis obteve um desempenho superior (73,81% de concordância) ao de mapas produzidos pelos demais conjuntos de variáveis. Possíveis fontes de erro na utilização dessa abordagem estão relacionadas, principalmente, à grande heterogeneidade litológica da área, que dificultou o estabelecimento de um modelo de correlação ambiental mais realista. A abordagem utilizada pode contribuir para tornar o levantamento de solos no Brasil mais rápido e menos subjetivo.

Termos para indexação: atributos do terreno, classificação de solos, modelo digital de elevação, redes neurais artificiais.

Topographic attributes and Landsat7 data in the digital soil mapping using neural networks

Abstract – The objective of this study was to evaluate discriminant variables in digital soil mapping using artificial neural networks. The topographic attributes elevation, slope, aspect, plan curvature and topographic index, derived from a digital elevation model, and the indexes of clay minerals, iron oxide and normalized difference vegetation, derived from a Landsat7 image, were combined and evaluated for their ability to discriminate soils of an area at the northwest of Rio de Janeiro State. The Java neural simulator and the backpropagation learning algorithm were used. The maps generated by each of the six tested sets of variables were compared with reference points for determining the rating accuracy. This comparison showed that the map produced with the use of all the variables reached a performance (73.81% of agreement) superior to maps produced by other sets of variables. Possible sources of error in the use of this approach are mainly related to the great lithological heterogeneity of the area, which hindered the establishment of a more realistic model of environmental correlation. The approach can help make the soil survey in Brazil faster and less subjective.

Index terms: terrain attributes, classification of soils, digital elevation model, artificial neural networks.

Introdução

A demanda por dados de levantamento de solos para subsidiar o planejamento e a modelagem ambiental tem crescido mundialmente. No entanto, o método tradicional de levantamento tem se mostrado caro, demorado e, algumas vezes, ineficiente (McBratney

et al., 2003). Os recursos financeiros destinados aos levantamentos de solos no Brasil têm diminuído muito, embora poucas áreas do território brasileiro apresentem levantamentos em média escala, entre 1:100.000 e 1:50.000. Além disso, o número de profissionais capacitados a realizar esses levantamentos vem sendo reduzido a cada ano.

Para superar essas dificuldades, novos métodos e estratégias vêm sendo desenvolvidos para a obtenção de informações de solo com uso de modelagem. Métodos de mineração de dados, tais como redes neurais artificiais (RNA), árvores de decisão, árvores de classificação, entre outros, podem fornecer soluções que auxiliem a extração automática de informações a partir de um conjunto de dados existentes (Behrens et al., 2005).

As RNA estão sendo amplamente utilizadas na ciência do solo, principalmente para estimativa de propriedades do solo. A sua aplicação como função de pedotransferência para estimativa de propriedades hidráulicas tem sido a mais comum (Minasny & McBratney, 2000; Minasny et al., 2004; Merdun et al., 2006; Elshorbagy & Parasuraman, 2008). A aplicação no mapeamento de classes de solos é mais rara e foi relatada em poucos estudos. Zhu (2000) utilizou RNA para alimentar um modelo de similaridade construído para representar o solo como um contínuo espacial. Nesse estudo, um conjunto de variáveis de fatores de formação do solo foi usado como dados de entrada na rede, e a informação espacial derivada do uso dessa abordagem revelou detalhes espaciais maiores e com maior qualidade do que os derivados do mapa de solo convencional. Uma abordagem para o mapeamento digital de classes de solos baseada em RNA foi desenvolvida por Behrens et al. (2005), com uso de 69 atributos do terreno, 53 unidades geológicas e 3 tipos de uso da terra, extraídos de mapas e bancos de dados de solos existentes. Em geral, a abordagem mostrou-se bastante satisfatória, economizou tempo, reduziu custos financeiros e produziu resultados confiáveis. Boruvka & Penizek (2007) utilizaram RNA no mapeamento de classes de solos, com dados de pH, teor de argila e gradiente textural de levantamentos de solos pré-existent, além de dados de elevação, aspecto e declividade do terreno, como variáveis ambientais discriminantes. Os resultados mostraram a existência de diferenças no sucesso da predição entre as classes de solos avaliadas, atribuídas à heterogeneidade de cada classe.

A complexidade da composição do solo e a especificidade da assinatura espectral de reflectância de cada um de seus componentes, além da resolução espacial e espectral das imagens, exigem esforços de pesquisa para se definir o que pode ser obtido a partir dessas imagens, de modo a se estabelecerem as relações entre o sinal contido ou derivadas delas e as informações do solo (Madeira Netto et al., 2007). Assim, Dobos et al. (2001) sugeriram que os dados

de sensores remotos fossem complementados com informações do terreno, para compensar as distorções que surgem das variações topográficas da paisagem e para fornecer dados adicionais para modelagem solo-paisagem.

O objetivo deste trabalho foi avaliar a eficiência de atributos do terreno e de dados do sensor ETM⁺ do Landsat7, no mapeamento digital de solos, com uso de redes neurais artificiais, tendo-se em vista a possibilidade de sua utilização como ferramenta auxiliar na elaboração de levantamentos de solos de média escala.

Material e Métodos

O estudo foi desenvolvido em uma área piloto, com 9.953 ha, na bacia hidrográfica do rio São Domingos, afluente do rio Muriaé, no noroeste do Estado do Rio de Janeiro, situada entre as coordenadas UTM 7.629.223 e 7.638.239 m N e 186.146 e 197.180 m E, zona 24S (Figura 1). De acordo com a classificação de Köppen, a área apresenta clima do tipo Aw (tropical, com inverno seco), com estação chuvosa no verão (novembro a abril) e seca no inverno (maio a outubro).

Dois sistemas de relevo são encontrados na área: o domínio colinoso de mar de morros, caracterizado por apresentar amplitude topográfica inferior a 100 m; e os alinhamentos serranos e degraus estruturais (NE-SW), caracterizados por elevações de grande amplitude topográfica. A maior parte da área apresenta altitudes inferiores a 300 m (86,36%), e a média estimada de declividade é de 26,63%. Os granulitos noríticos da Unidade São José de Ubá (63,45%), os migmatitos das Unidades Catalunha, São João do Paraíso e Vista Alegre (21,35%), e os sedimentos aluvionares do Quaternário (15,17%) são as litologias predominantes (Rio de Janeiro, 1980).

Foram utilizadas oito variáveis discriminantes no mapeamento digital de solos da área. Os atributos do terreno elevação, declividade, aspecto, plano de curvatura e índice topográfico combinado (CTI) foram usados para caracterizar as condições topográficas. Esses atributos foram derivados de um modelo digital de elevação (MDE), com 30 m de resolução espacial, obtido com a opção “topo to raster” do programa ArcGIS 9.2 (Environmental Systems Research Institute, 2004), a partir das curvas de nível equidistantes de 20 m, hidrografia e pontos cotados contidos nas cartas topográficas do IBGE, na escala de 1:50.000. Dados do

sensor ETM⁺ do Landsat7, de agosto de 1999, foram utilizados para derivar os índices minerais de argila (“clay mineral index”, CMI), óxido de ferro (“iron oxide index”, IOI) e vegetação por diferença normalizada (“normalized difference vegetation index”, NDVI), todos obtidos com a utilização do ERDAS Imagine 8.5 (ERDAS, 2001).

O CMI foi obtido com divisão da banda 5 (1,55–1,75 μm) pela banda 7 (2,08–2,35 μm) e o IOI com divisão da banda 3 (0,63–0,69 μm) pela banda 1 (0,45–0,52 μm). Embora os índices CMI, IOI e NDVI não apresentem relação direta com os fatores de formação do solo, eles foram utilizados como fonte de dados auxiliares e avaliados quanto à capacidade de melhorar o desempenho da classificação. Os índices de vegetação foram criados para tentar diminuir o trabalho de análise dos dados, pela maximização de informações espectrais da vegetação no menor número de bandas de operação dos sensores (Moreira, 2003). Estudos sobre

as variações espacial e temporal do índice NDVI têm mostrado a ligação desse índice com as propriedades físicas do solo (Lozano-Garcia et al., 1991; Narasimha Rao et al., 1993; Yang et al., 1997). Os índices CMI e IOI são empregados no sensoriamento remoto geológico e podem ser utilizados para auxiliar na distinção de solos com características físicas e mineralógicas distintas (Sabins Junior, 1997).

Pelo fato de as variáveis discriminantes utilizadas serem de fontes distintas, buscou-se identificar a importância de cada um dos grupos (atributos do terreno e dados de sensores remotos) e de seus respectivos componentes pelo uso das combinações apresentadas nos seguintes conjuntos: 1, elevação, declividade, aspecto, curvatura, CTI, CMI, IOI e NDVI; 2, elevação, declividade, aspecto, curvatura, CTI e CMI; 3, elevação, declividade, aspecto, curvatura, CTI e IOI; 4, elevação, declividade, aspecto, curvatura, CTI e NDVI; 5, elevação, declividade, aspecto, curvatura e CTI; e 6, CMI, IOI e NDVI.

A utilização de uma RNA inicia-se com o processo de treinamento da rede (Tso & Mather, 2009). Inicialmente, todos os valores referentes às variáveis discriminantes foram reescalados para o intervalo entre 0 e 1, com o ArcGIS 9.2, no caso dos atributos do terreno, e o ERDAS IMAGINE 8.5, no caso dos índices derivados da imagem do Landsat7. Esse procedimento foi realizado para facilitar o processo de treinamento e evitar a saturação das RNA, pois valores muito grandes poderiam impedir a solução do problema (convergência da rede). Além disso, esse procedimento pode prevenir que grandes variações de uma variável pouco importante inibam pequenas variações em outras variáveis.

Além dos atributos do terreno e dos índices derivados da imagem do Landsat7, foi utilizado um mapa geológico da área, na escala de 1:50.000 (Rio de Janeiro, 1980), para se ter informações acerca do material de origem dos solos. Com base nesse mapa e nas observações de campo, a área foi subdividida em granulitos noríticos (área 1) e migmatitos heterogêneos (área 2), para facilitar o processo de treinamento das RNA, já que um número grande de classes de solos poderia levar à confusão entre as classes.

Para o entendimento das relações solo-paisagem, toda a área foi percorrida, conforme as normas estabelecidas para levantamentos de solos convencionais. Na ocasião, foram feitas 256 observações sem coleta de amostras

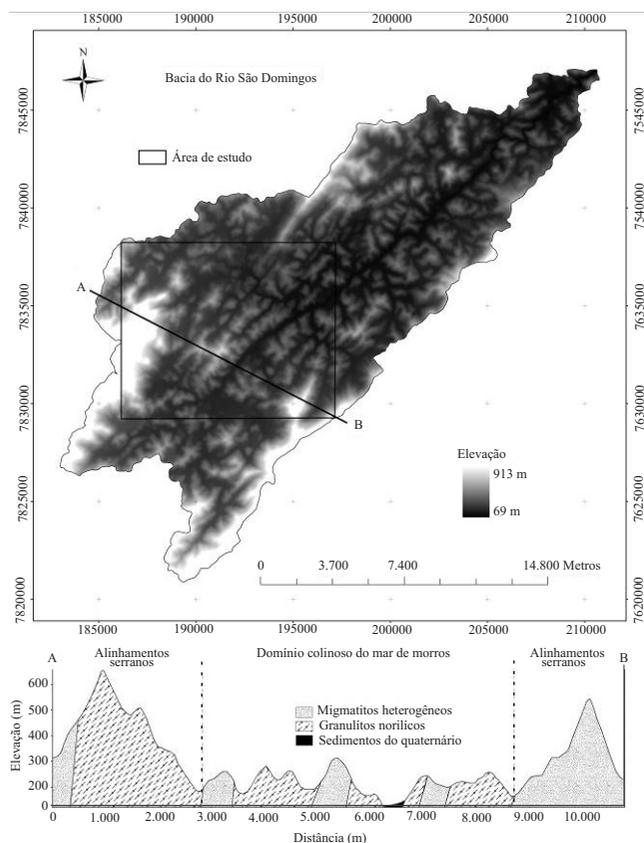


Figura 1. Localização da área de estudo na Bacia do Rio São Domingos, no noroeste do Rio de Janeiro, e perfil de elevação na linha diagonal A–B, com evidência dos domínios geomorfológicos e das principais litologias da área.

e, posteriormente, para a caracterização das classes de solos identificadas, foram utilizados 73 perfis de solos completos e 11 amostras extras, obtidos em estudos preexistentes ou coletados especificamente para este estudo. As classes de solos, de acordo com o Sistema Brasileiro de Classificação de Solos (Santos et al., 2006), estão apresentadas na Tabela 1.

O simulador de redes neurais utilizado foi o “Java Neural Network Simulator”, que é baseado no “Stuttgart Neural Network Simulator 4.2 Kernel” (Zell et al., 1996). Para cada conjunto de variáveis discriminantes, foram coletados, independentemente, dois conjuntos de amostras: um para o treinamento e outro para a validação das RNA, que também são independentes entre os conjuntos de variáveis discriminantes avaliados. Essas amostras (amostras estratificadas) foram coletadas de uma imagem que continha as variáveis discriminantes, com o auxílio do software ERDAS IMAGINE 8.5 e a opção “converter células para ASCII”, para capturar um conjunto que representasse o máximo possível as características de

cada classe de solo, quanto à elevação, declividade, aspecto, curvatura, índice CTI e aos índices CMI, IOI e NDVI. Em seguida, os arquivos obtidos no formato ASCII foram convertidos para o formato de trabalho do “Java Neural Network Simulator”. Assim, para cada conjunto testado, foi criado um arquivo de treinamento e outro de validação.

Embora Congalton & Green (1999) tenham desenvolvido procedimentos estatísticos para definir o número mínimo de amostras para classificadores, optou-se por seguir as recomendações de Zhu (2000) que, em seu estudo com RNA, indica que o número de amostras de treinamento deve estar relacionado ao número de classes de solos existentes em uma determinada área, pois quanto mais classes, mais complexo será o mapeamento da relação entrada (variáveis)/saída (classes de solos) e, por sua vez, mais amostras serão necessárias para treinar as redes. Assim, o tamanho da amostra de treinamento utilizado foi de aproximadamente 30 vezes o número de classes de solos (300 pixels por classe de solo), enquanto o

Tabela 1. Classes de solos identificadas na Bacia Hidrográfica do rio São Domingos, no Noroeste do Estado do Rio de Janeiro.

Símbolo	Número de perfis	Classe de solo
Granulitos noríticos (área 1)		
AR1	-	Afloramentos de rocha
RLe1	1	Neossolo Litólico eutrófico típico
CXbe	6	Cambissolo Háptico Tb eutrófico típico
PVe1	7	Argissolo Vermelho eutrófico cambissólico
PVe2	3	Argissolo Vermelho eutrófico típico
PVe3	7	Argissolo Vermelho eutrófico nitossólico
PVAe1	7	Argissolo Vermelho-Amarelo eutrófico típico
PVAe2	3	Argissolo Vermelho-Amarelo eutrófico típico ou cambissólico
PVAd	7	Argissolo Vermelho-Amarelo distrófico latossólico
LVAd	1	Latossolo Vermelho-Amarelo distrófico
GXbe	3	Gleissolo Háptico Tb eutrófico solódico ou típico + Cambissolo Háptico Tb eutrófico gleissólico
PAe1	1	Argissolo Amarelo eutrófico típico
Migmatitos heterogêneos (área 2)		
AR2	-	Afloramentos de rocha
RLe2	1	Neossolo Litólico eutrófico típico
CXve	10	Cambissolo Háptico Ta eutrófico típico
PVe4	2	Argissolo Vermelho eutrófico abrupto cambissólico
PVe5	7	Argissolo Vermelho eutrófico abrupto
PVe6	4	Argissolo Vermelho eutrófico abrupto latossólico
PVe7	3	Argissolo Vermelho eutrófico abrupto ou abrupto cambissólico + Argissolo Vermelho-Amarelo eutrófico abrupto ou abrupto cambissólico abrupto ou abrupto cambissólico
LVe	1	Latossolo Vermelho distrófico típico
GXve	9	Gleissolo Háptico Ta eutrófico solódico ou típico + Cambissolo Háptico Ta eutrófico gleissólico
PAe2	1	Argissolo Amarelo eutrófico abrupto ou abrupto cambissólico

conjunto de validação foi 50% do tamanho do conjunto de treinamento (150 pixels).

Em razão do número elevado de classes de solos e da grande complexidade envolvida no processo de treinamento e validação das redes, optou-se pelo desenvolvimento de dois conjuntos de RNA, um para cada uma das áreas apresentadas na Tabela 1. Diferentes arquiteturas foram testadas, tendo-se variado o número de neurônios na camada de entrada, correspondentes ao número de variáveis discriminantes utilizadas e variado o número de neurônios na camada interna. Na camada de saída todas tiveram o mesmo número de neurônios com 12 classes de solos para a área 1, e 10 para a área 2. No treinamento das RNA, foi utilizado o algoritmo de aprendizado “backpropagation”, com alocação aleatória dos pesos interneurônios entre -0,5 e 0,5 e uma taxa de aprendizado de 0,2, considerando 10 mil ciclos de aprendizagem.

A avaliação dos resultados foi realizada com medidas estatísticas como o índice Kappa e a exatidão global da classificação, derivados de uma matriz de confusão (Congalton & Green, 1999). Uma matriz de significância foi gerada com os resultados dos testes estatísticos realizados, tendo-se utilizado os valores de Kappa e da variância do Kappa entre as classificações. O teste estatístico Z verifica, inicialmente, se a classificação difere de uma classificação casual e, em segunda análise, se existe diferença significativa entre os valores de Kappa resultantes da avaliação dos diferentes conjuntos, aos pares.

Ao término do processo de treinamento e validação das RNA, foi selecionada para o mapeamento digital das classes de solos da área a arquitetura de rede que obteve

o melhor resultado para o índice Kappa, para cada um dos conjuntos de variáveis discriminantes utilizados.

O único levantamento de solos existente na área está na escala de 1:500.000, escala incompatível com o nível de detalhe exigido (entre 1:100.000 e 1:50.000 – média escala), para comparação com os mapas digitais produzidos pelos diferentes conjuntos. Assim, foram utilizados 126 pontos de referência (perfis de solo, amostras extras e pontos de observação), não utilizados anteriormente no processo de treinamento das RNA, para determinar a percentagem de locais classificados corretamente pelos diferentes mapas produzidos, conforme procedimento utilizado por Zhu (2000).

Resultados e Discussão

Os resultados obtidos para o índice Kappa, quando se utilizaram as amostras de validação, em ambas as áreas, variaram consideravelmente em razão do número de neurônios na camada interna de todos os conjuntos testados. Dessa forma, na Tabela 2 estão apresentados somente os resultados obtidos pela RNA que teve o melhor desempenho no treinamento, em cada um dos conjuntos testados.

A combinação dos atributos do terreno com todos os três índices derivados da imagem do Landsat7 (conjunto 1) foi a que apresentou o melhor desempenho para ambas as áreas avaliadas, com valores de índice Kappa de 0,908 (oito neurônios na camada interna) e de 0,893 (cinco neurônios na camada interna), respectivamente para área 1 e 2, os quais devem ser considerados muito bom e excelente (Landis & Koch, 1977). No entanto, o pior desempenho para as duas

Tabela 2. Testes estatísticos com os valores de Kappa e variância para a avaliação do treinamento dos conjuntos testados⁽¹⁾.

Conjunto	Granulitos noríticos (área 1)						Migmatitos heterogêneos (área 2)					
	1 (8)	2 (15)	3 (10)	4 (5)	5 (6)	6 (3)	1 (5)	2 (15)	3 (20)	4 (5)	5 (20)	6 (3)
Exatidão global	91,6	89,4	85,8	85,8	87,3	25,1	90,4	86,9	84,8	83,9	86,2	22,3
Kappa	0,908	0,885	0,845	0,845	0,861	0,182	0,893	0,854	0,831	0,821	0,847	0,137
Variância ⁽²⁾	0,051	0,062	0,080	0,081	0,073	0,116	0,071	0,093	0,105	0,110	0,098	0,126
1	127,2						106,0					
2	2,2*	112,4					3,1*	88,6				
3	5,5*	3,4*	94,5				4,7*	1,6	81,1			
4	5,5*	3,4*	0,0	93,9			5,4*	2,3*	0,7	78,3		
5	4,2*	2,1*	1,3	1,3	100,8		3,5*	0,5	1,1	1,8	85,6	
6	56,2*	52,7*	47,4*	47,3*	49,4*	16,9	53,9*	48,5*	45,7*	44,5*	47,4*	12,2*

⁽¹⁾Conjunto: 1, atributos do terreno + índices de mineral de argila (CMI), de óxido de ferro (IOI) e de vegetação por diferença normalizada (NDVI); 2, atributos do terreno + CMI; 3, atributos do terreno + IOI; 4, atributos do terreno + NDVI; 5, atributos do terreno; 6, índices CMI, IOI e NDVI. Valores entre parênteses definem o número de neurônios na camada interna de cada conjunto. ⁽²⁾Valores multiplicados por 1000. *Significativo a 5% de probabilidade.

áreas foi obtido pelo conjunto 6, em que se utilizaram somente os índices derivados da imagem do Landsat7. De modo geral, os resultados obtidos foram satisfatórios para todos os conjuntos ($Kappa > 0,8$), exceto para o conjunto 6.

O melhor desempenho obtido pelo conjunto 1 difere de todos os demais conjuntos, em ambas as áreas. Entre os índices, a maior contribuição para a discriminação dos solos, tanto na área 1 quanto na área 2, foi dada pelo índice CMI, visto que a combinação deste com os atributos do terreno (conjunto 2) obteve resultados estatisticamente superiores aos alcançados pelos índices IOI (conjunto 3) e NDVI (conjunto 4). O índice IOI, quando associado com os atributos do terreno, produziu resultado igual (área 1) ou ligeiramente superior (área 2) ao obtido com o índice NDVI (conjunto 4) (Tabela 2).

Os índices derivados da imagem do Landsat7 foram utilizados como variáveis adicionais para melhorar a

discriminação das classes de solos da área e, embora não apresentem relação direta com os fatores de formação do solo, sua utilização com os atributos do terreno teve efeito significativo sobre o desempenho do treinamento. Na Figura 2 estão apresentados os valores médios das amostras de treinamento das diferentes variáveis discriminantes entre as classes de solos. Nas duas áreas, observa-se que o índice NDVI apresenta uma distribuição muito próxima ao índice CMI. Neste caso, quando associadas aos atributos do terreno no conjunto 1, a RNA mostrou habilidade para estimar a importância de cada uma e para ignorar possíveis redundâncias entre essas variáveis.

Por sua vez, o desempenho inferior do índice NDVI pode estar relacionado à baixa diferenciação desse índice entre as classes de solos (Figura 2), pelo fato de que praticamente toda a vegetação original da área foi substituída por pastagens, principalmente de capim-colômbio. O baixo desempenho do treinamento,

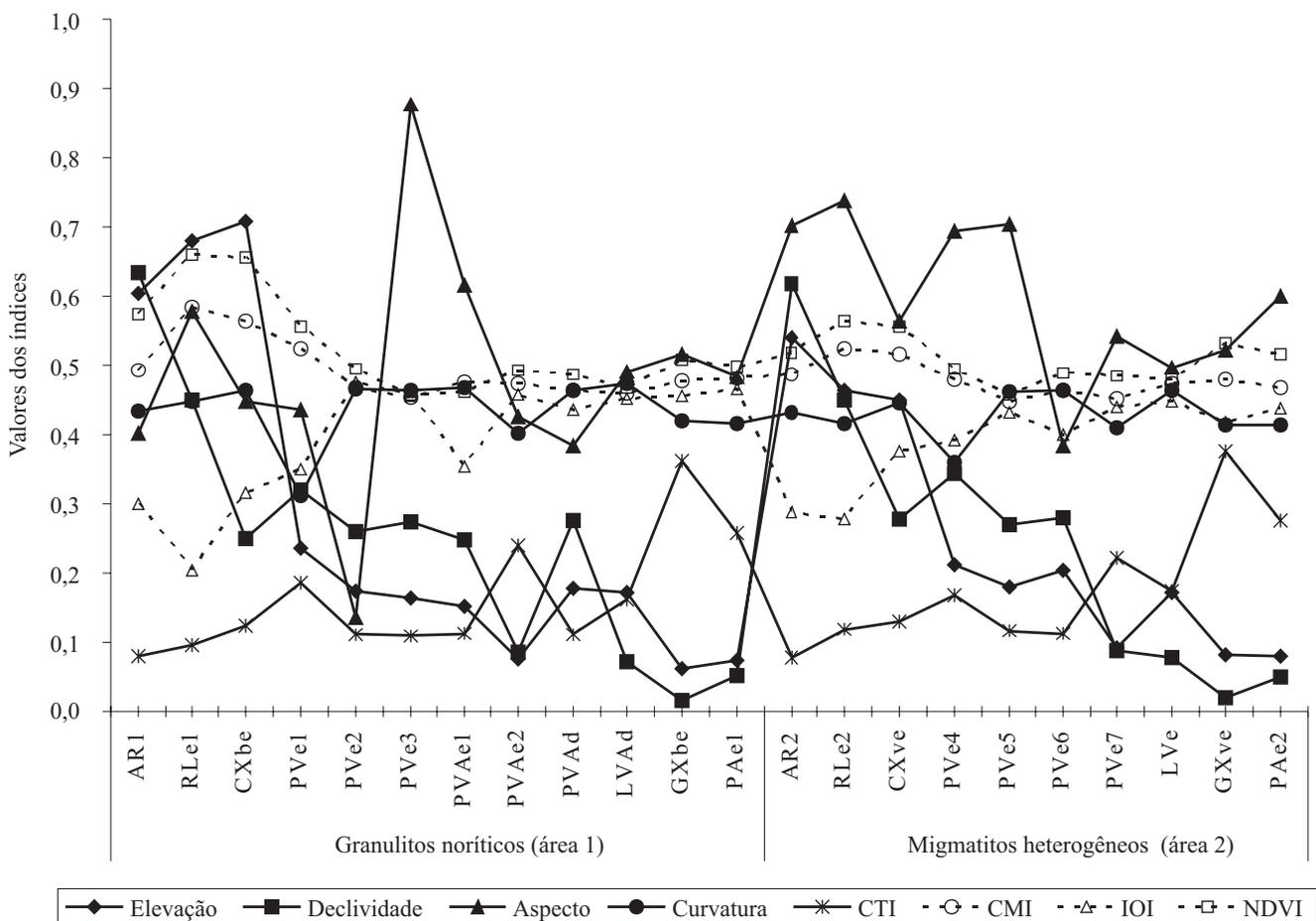


Figura 2. Distribuição das variáveis discriminantes entre as classes de solos descritas na Tabela 1. Valores dos índices reescalados entre 0 e 1 para permitir a apresentação em conjunto. CTI, índice topográfico combinado; CMI, índice de mineral de argila; IOI, índice de óxido de ferro; NDVI, índice de vegetação por diferença normalizada.

quando se utilizaram somente os três índices derivados da imagem do Landsat7 (Tabela 2), corroborou a afirmativa de Dobos et al. (2001).

De modo geral, as diferenças nos atributos do terreno entre as classes de solos são mais marcantes do que as verificadas nos índices derivados da imagem do Landsat7, o que contribuiu com maior peso para a discriminação dos solos, conforme comprovado pelos valores do índice Kappa obtidos quando foram utilizados somente os atributos do terreno em ambas as áreas (Tabela 2).

A forte influência dos atributos do terreno sobre a distribuição espacial dos solos é um princípio bem conhecido (Park & Vlek, 2002), que foi confirmado neste estudo. No entanto, apesar da maior importância dos atributos do terreno, a utilização deles separadamente apresentou um desempenho estatisticamente inferior ao obtido quando os índices derivados da imagem do Landsat7 foram adicionados. Resultados semelhantes foram obtidos por Cole & Boettinger (2007) com atributos do terreno e dados do sensor Landsat5 no mapeamento de classes de solos.

Entre os atributos, apenas o plano de curvatura mostrou pouca diferenciação entre as classes (Figura 2), o que, entretanto, foi muito importante para a separação de PVe1 e PVe4 + CXve das demais classes (curvatura côncava). A forma da curvatura de uma encosta pode influenciar muito a distribuição lateral dos processos pedológicos, hidrológicos e geomórficos e, por conseguinte, os solos que resultam das interações entre esses processos (MacMillan et al., 2000). Essa influência tem sido relacionada, principalmente, ao controle que as formas côncava e convexa exercem sobre a distribuição de água e materiais solúveis das partes mais elevadas para as mais baixas. Na área estudada, os solos mais profundos e evoluídos ocupam encostas convexas, enquanto os mais rasos e menos evoluídos tendem a ocorrer em encostas côncavas, em consequência do maior fluxo convergente de água nessas áreas, o que favorece a morfogênese e leva esses solos a um constante rejuvenescimento por erosão.

O índice CTI apresentou valores médios muito similares entre as classes que ocorrem no terço médio ou superior de encosta (AR1 a LVAd e AR2 a LVe) que, por sua vez, foi distinto dos observados para as classes localizadas no terço inferior (GXbe, GXve, PAe1 e PAe2). Assim, o índice foi importante para a

identificação, principalmente dos solos hidromórficos que ocorrem na área.

Entre os atributos do terreno, a declividade e a elevação foram os que apresentaram o comportamento mais distinto entre as classes de solos, tanto na área 1 quanto na área 2 (Figura 2) e, certamente, contribuíram para melhorar o desempenho do treinamento. A forte influência que esses atributos têm sobre os processos pedogenéticos e, conseqüentemente, sobre a distribuição dos solos na paisagem, foi ressaltada por Gallant & Wilson (2000) e Thompson et al. (2001). Crivelenti et al. (2009), ao utilizar uma técnica de mineração de dados (árvore de decisão) em um mapeamento digital de solos no Estado de São Paulo, verificaram que os atributos do terreno que mais influenciaram a determinação das unidades de mapeamento foram a distância diagonal da drenagem, a declividade e a curvatura em perfil. Esses autores associaram a importância desses atributos ao acúmulo de água e aos fluxos hídricos, que possuem grande influência na formação dos solos.

O atributo aspecto apresentou comportamento diferenciado entre as classes de solo, em ambas as áreas e, por afetar diretamente o regime de temperatura do ar e do solo, a evapotranspiração e a umidade do solo tiveram efeito significativo sobre a distribuição dos solos, sobretudo na diferenciação das classes PVe2 + PVe3 das classes PVAe1 + PVAe2 e das classes PVe5 da PVe6. Nestes casos, as variações microclimáticas, determinadas pelo aspecto nas encostas convexas, condicionaram a formação de solos mais intemperizados, o que é comprovado pelos resultados analíticos dos perfis de solos coletados, com horizonte B latossólico em profundidade, como nas classes PVAe2 e PVe6 nas encostas voltadas para o sul, que são relativamente mais frias e úmidas do que as encostas com orientação norte, conforme também verificado por Rech et al. (2001).

Para facilitar a avaliação dos mapas digitais e possibilitar a comparação com os pontos de referência, algumas classes de solos foram combinadas (Figura 3). A classificação, com uso do conjunto 1 (atributos do terreno + índices derivados da imagem do Landsat7), inferiu corretamente as classes de solo em 93 locais, com índice de exatidão global de 73,81%, e alcançou o melhor resultado entre todas as classificações realizadas (Tabela 3). O pior desempenho foi obtido pelo conjunto 6 (índices derivados da imagem do Landsat7),

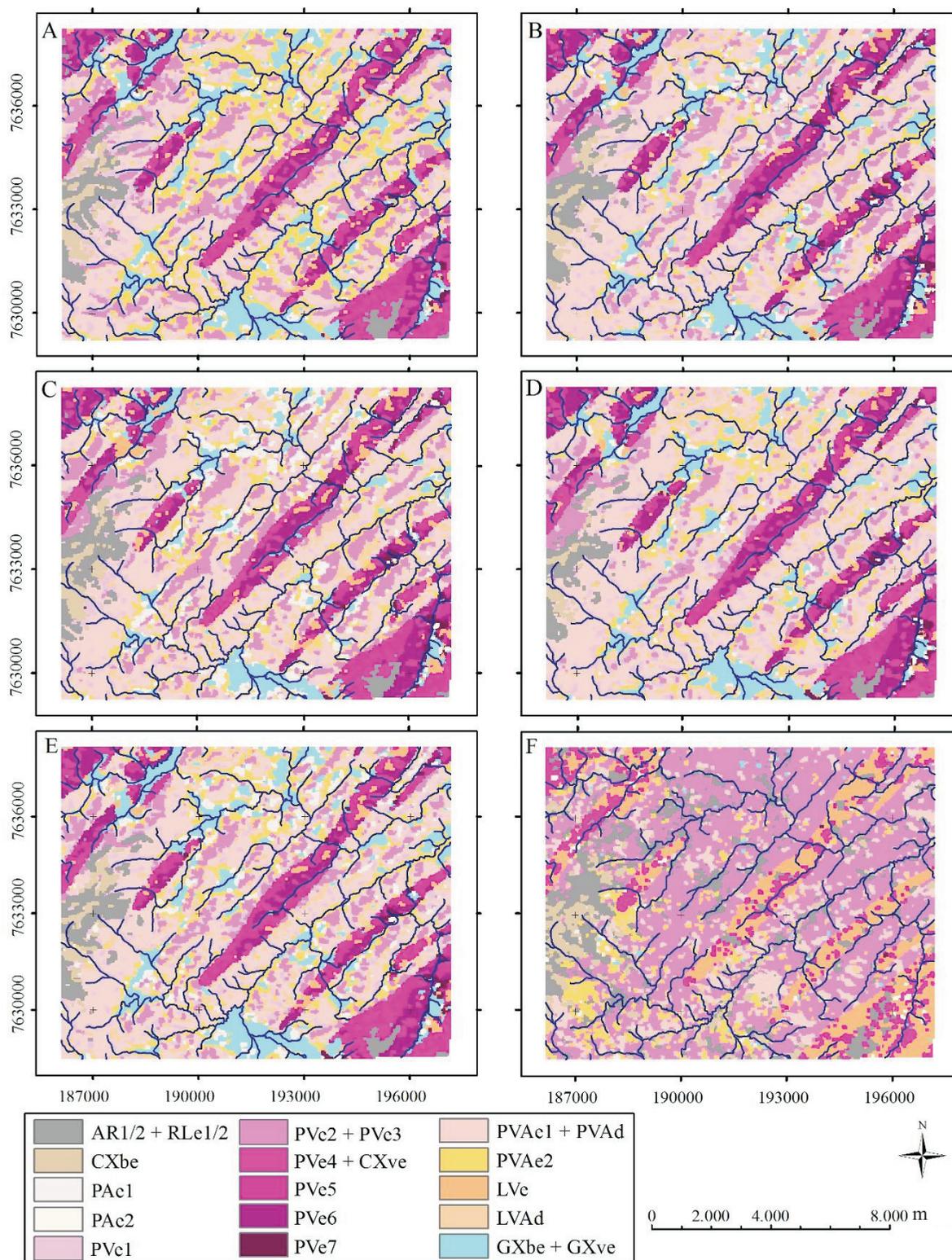


Figura 3. Mapas de classes de solos descritas na Tabela 1, produzidos pelos conjuntos de variáveis discriminantes. A, atributos do terreno + índices de mineral de argila (CMI), de óxido de ferro (IOI) e de vegetação por diferença normalizada (NDVI); B, atributos do terreno + CMI; C, atributos do terreno + IOI; D, atributos do terreno + NDVI; E, atributos do terreno; F, CMI, IOI e NDVI.

que obteve índice de exatidão global de apenas 16,67%. Os conjuntos 2, 3, 4 e 5, com desempenho inferior ao do conjunto 1, apresentaram resultados semelhantes entre si, e mostraram que os índices CMI, IOI e NDVI, quando utilizados isoladamente com os atributos do terreno, não contribuem para melhorar o desempenho da classificação.

As diferenças mais marcantes observadas entre as classificações apresentadas na Tabela 4 e visualizadas na Figura 3, exceto o conjunto 6 (pior desempenho), ocorreram entre as classes PVAe1 + PVAd e PVAe2. A utilização dos atributos do terreno com os índices derivados da imagem do Landsat7 (conjunto 1)

classificou 18,9% da área como pertencentes às classes PVAe1 + PVAd, enquanto que para os demais conjuntos os valores foram de 27,1, 28,8, 27,3 e 25,4%, respectivamente, conjuntos 2, 3, 4 e 5. Para a classe PVAe2, foi constatada uma relação inversa com o conjunto 1, que classificou mais áreas como pertencentes a essa classe (23,5%) do que os demais conjuntos, cujos valores obtidos foram de 9,8, 12,9, 10,9 e 13,84% da área total (conjuntos 2, 3, 4 e 5, respectivamente).

A classe PVAe2 ocorre basicamente em áreas com declives que variam de 8 a 20%, tanto em morrotes isolados quanto no terço inferior das encostas, e as classes PVAe1 + PVAd, por sua vez, ocorrem em declives superiores a 20% e, neste caso, os conjuntos 2, 3, 4 e 5 parecem subestimar a área da classe PVAe2 e superestimar a das classes PVAe1 + PVAd. A percentagem de área com declives entre 8 e 20%, obtida diretamente do mapa de declividade, é de aproximadamente 25% e corrobora o melhor resultado, obtido pelo conjunto 1.

Apesar da pouca diferença observada entre os conjuntos de variáveis para a classificação das classes GXbe + GXve, observa-se na Figura 3 que todos, em maior ou menor grau, subestimaram a área pertencente a estas classes. Parte da área que deveria ser atribuída

Tabela 3. Comparação entre os mapas digitais produzidos pelos diferentes conjuntos de variáveis discriminantes e os pontos de referência.

Avaliação	Conjuntos ⁽¹⁾					
	1	2	3	4	5	6
Total de pontos	126	126	126	126	126	126
Pontos corretamente classificados	93	71	64	71	73	21
Exatidão global	73,81	56,35	50,79	56,35	57,94	16,67

⁽¹⁾Conjuntos: 1, atributos do terreno + índices de mineral de argila (CMI), de óxido de ferro (IOI) e de vegetação por diferença normalizada (NDVI); 2, atributos do terreno + CMI; 3, atributos do terreno + IOI; 4, atributos do terreno + NDVI; 5, atributos do terreno; 6, CMI, IOI e NDVI.

Tabela 4. Percentagem de ocorrência das classes de solos obtidas pelos diferentes conjuntos de variáveis discriminantes⁽¹⁾.

Classe de solo ⁽²⁾	Área (%)					
	Conjunto 1	Conjunto 2	Conjunto 3	Conjunto 4	Conjunto 5	Conjunto 6
AR1+RLe1+AR2+RLe2	3,3	3,5	3,6	2,9	3,6	13,1
CXbe	1,6	1,6	1,5	0,9	1,4	6,2
PVe1	6,7	11,7	5,4	12,3	7,5	0,0
PVe2 + PVe3	11,9	10,3	11,2	6,6	9,2	40,8
PVAe1 + PVAd	18,9	27,1	28,7	27,3	25,4	10,1
PVAe2	23,5	9,8	12,9	10,9	13,8	6,2
LVAd	1,2	3,5	3,8	8,1	3,7	1,7
GXbe + GXve	12,1	9,5	6,3	9,0	10,6	0,4
PAe1	1,6	2,4	6,1	1,2	4,8	0,3
PVe4 + CXve	8,0	4,7	5,4	5,8	4,4	6,2
PVe5	3,3	4,2	3,9	4,4	7,0	1,1
PVe6	6,1	7,7	6,9	7,6	5,7	0,0
PVe7	0,6	1,5	1,4	1,2	1,1	0,0
LVe	1,3	2,0	2,6	1,5	1,7	13,3
PAe2	0,2	0,4	0,3	0,5	0,2	0,5
Total	100	100	100	100	100	100

⁽¹⁾Conjuntos: 1, atributos do terreno + índices de mineral de argila (CMI), de óxido de ferro (IOI) e de vegetação por diferença normalizada (NDVI); 2, atributos do terreno + CMI; 3, atributos do terreno + IOI; 4, atributos do terreno + NDVI; 5, atributos do terreno; 6, CMI, IOI e NDVI. ⁽²⁾Classes de solos descritas na Tabela 1.

a estas classes foi atribuída à classe PVAe2, o que pode estar relacionado ao tamanho da célula do grid utilizado (30 m), que não permitiu a identificação de pequenas, porém importantes áreas ocupadas por essas classes.

Embora diferenças estatísticas tenham sido verificadas entre as classificações, principalmente entre o conjunto 1 e os demais, há que se destacar que uma importante causa de discordância entre as classificações e os pontos de referência deve-se à elevada complexidade geológica da área, que não está representada adequadamente no mapa geológico disponível. Problemas relacionados à qualidade de mapas geológicos, utilizados em estudos de correlação ambiental, foram relatados por McKenzie & Ryan (1999) e Thomas et al. (1999). Além disso, em alguns locais, a profundidade de exposição das rochas tem importância maior para a distribuição dos solos, muitas vezes com influência mais decisiva do que a morfometria. Além disso, McKenzie & Ryan (1999) destacam que os modelos de correlação ambiental podem não utilizar toda a capacidade preditiva dos modelos mentais intuitivos, usados nos levantamentos convencionais. Finalmente, há circunstâncias em que a variação dos solos ocorre sem relação com variáveis ambientais facilmente observáveis pela falta de um preditor confiável.

Conclusões

1. A utilização dos atributos do terreno – elevação, declividade, curvatura, aspecto e índice topográfico combinado – e dos índices de mineral de argila, de óxido de ferro e vegetação por diferença normalizada, derivados de uma imagem do Landsat7 em uma abordagem por redes neurais artificiais, obteve resultados satisfatórios tanto no processo de treinamento das redes quanto na comparação com pontos de referência de perfis de solo.

2. Os atributos do terreno que mais contribuíram para melhorar o desempenho da classificação de classes de solos foram a declividade e a elevação e, entre os índices derivados da imagem do Landsat7, a maior contribuição foi dada pelo índice de mineral de argila, sendo que atributos e índices apresentam comportamento distinto entre as classes de solos avaliadas.

3. As principais causas de divergências entre os pontos de referência e os mapas de solos digitais estão

relacionadas à escala do mapa geológico utilizado, às deficiências no aprendizado dos critérios utilizados no treinamento das redes neurais artificiais e às deficiências das variáveis ambientais em representar todas as variações das características das classes de solos.

Referências

- BEHRENS, T.; FORSTER, H.; SCHOLTEN, T.; STEINRÜCKEN, U.; SPIES, E.D.; GOLDSCHMITT, M. Digital soil mapping using artificial neural networks. **Journal of Plant Nutrition and Soil Science**, v.168, p.21-33, 2005.
- BORUVKA, L.; PENIZEK, V. A test of an artificial neural network allocation procedure using the Czech soil survey of agricultural land data. In: LAGACHERIE, P.; MCBRATNEY, A.B.; VOLTZ, M. (Ed.). **Digital soil mapping: an introductory perspective**. Amsterdam: Elsevier, 2007. p.415-424.
- COLE, N.J.; BOETTINGER, J.L. Pedogenic understanding raster classification methodology for mapping soils, Powder River Basin, Wyoming, USA. In: LAGACHERIE, P.; MCBRATNEY, A.B.; VOLTZ, M. (Ed.). **Digital soil mapping: an introductory perspectives**. Amsterdam: Elsevier, 2007. p.377-388.
- CONGALTON, R.G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. New York: Lewis Publishers, 1999. 137p.
- CRIVELENTI, R.C.; COELHO, R.M.; ADAMI, S.F.; OLIVEIRA, S.R. de M. Mineração de dados para inferência de relações solo-paisagem em mapeamentos digitais de solo. **Pesquisa Agropecuária Brasileira**, v.44, p.1707-1715, 2009.
- DOBOS, E.; MONTANARELLA, L.; NÈGRE, T.; MICHELI, E. A regional scale soil mapping approach using integrated AVHRR and DEM data. **International Journal of Applied Earth Observation and Geoinformation**, v.3, p.30-42, 2001.
- ELSHORBAGY, A.; PARASURAMAN, K. On the relevance of using artificial neural networks for estimating soil moisture content. **Journal of Hydrology**, v.362, p.1-18, 2008.
- ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. **ArcGIS 9.2: getting started with ArcGIS**. Redlands: ESRI, 2004. 272p.
- ERDAS. **ERDAS Imagine**. Version 8.5. Atlanta: Erdas, 2001.
- GALLANT, J.C.; WILSON, J.P. Primary topographic attributes. In: WILSON, J.P.; GALLANT, J.C. (Ed.). **Terrain analysis: principles and applications**. New York: John Wiley & Sons, 2000. p.51-85.
- LANDIS, J.R.; KOCH, G.G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, p.159-174, 1977.
- LOZANO-GARCIA, D.F.; FERNANDEZ, R.N.; JOHANNSEN, C.J. Assessment of regional biomass-soil relationships using vegetation indexes. **IEEE Transactions on Geoscience and Remote Sensing**, v.29, p.331-338, 1991.

- MACMILLAN, R.A.; PETTAPECE, W.W.; NOLAN, S.C.; GODDARD, T.W. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. **Fuzzy Sets and Systems**, v.113, p.81-109, 2000.
- MADEIRANETTO, J.S.; ROBBEZ-MASSON, J.M.; MARTINS, E. Visible-Nir hyperspectral imagery for discriminating soil types in the La Peyne watershed (France). In: LAGACHERIE, P.; MCBRATNEY, A.; VOLTZ, M. (Ed.). **Digital soil mapping: an introductory perspective**. Amsterdam: Elsevier, 2007. p.219-234.
- MCBRATNEY, A.B.; SANTOS, M.L.M.; MINASNY, B. On digital soil mapping. **Geoderma**, v.117, p.3-52, 2003.
- MCKENZIE, N.J.; RYAN, P.J. Spatial prediction of soil properties using environmental correlation. **Geoderma**, v.89, p.67-94, 1999.
- MERDUN, H.; ÇINAR, Ö.; MERAL, R.; APAN, M. Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. **Soil and Tillage Research**, v.90, p.108-116, 2006.
- MINASNY, B.; HOPMANS, J.W.; HARTER, T.; ECHING, S.O.; TULI, A.; DENTON, M.A. Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. **Soil Science Society of America Journal**, v.68, p.417-430, 2004.
- MINASNY, B.; MCBRATNEY, A. Evaluation and development of hydraulic conductivity pedotransfer functions for Australian soil. **Australian Journal of Soil Research**, v.38, p.905-926, 2000.
- MOREIRA, M.A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. 2.ed. Viçosa : UFV, 2003. 307p.
- NARASIMHA RAO, P.V.; VENKATARATNAM, L.; KRISHNA RAO, P.V.; RAMANA, K.V.; SINGARAO, M.N. Relation between root zone soil moisture and Normalized Difference Vegetation Index of vegetated fields. **International Journal of Remote Sensing**, v.14, p.441-449, 1993.
- PARK, S.J.; VLEK, P.L.G. Environmental correlation of three-dimensional soil spatial variability: a comparison of three adaptive techniques. **Geoderma**, v.109, p.117-140, 2002.
- RECH, J.A.; REEVES, R.W.; HENDRICKS, D.M. The influence of slope aspect on soil weathering processes in the Springerville volcanic field, Arizona. **Catena**, v.43, p.49-62, 2001.
- RIO DE JANEIRO. Secretaria de Estado de Indústria, Comércio e Turismo. Departamento de Recursos Minerais. **Projeto Carta Geológica do Estado do Rio de Janeiro**. Niterói: DRM, 1980. Escala 1:50.000. Folhas: Miracema e São João do Paraíso.
- SABINS JUNIOR, F.F. **Remote sensing: principles and interpretation**. 3rd ed. New York: W. H. Freeman and Company, 1997. 432p.
- SANTOS, H.G. dos; JACOMINE, P.K.T.; ANJOS, L.H.C. dos; OLIVEIRA, V.A. de; OLIVEIRA, J.B. de; COELHO, M.R.; LUMBRERAS, J.F.; CUNHA, T.J.F. (Ed.). **Sistema brasileiro de classificação de solos**. 2.ed. Rio de Janeiro: Embrapa Solos, 2006. 306p.
- THOMAS, A.L.; KING, D.; DAMBRINE, E.; COUTURIER, A. Predicting soil classes with parameters derived from relief and geologic materials in a sandstone region of the Vosges mountains (Northeastern France). **Geoderma**, v.90, p.291-305, 1999.
- THOMPSON, J.A.; BELL, J.C.; BUTLER, C.A. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. **Geoderma**, v.100, p.67-89, 2001.
- TSO, B.; MATHER, P.M. **Classification methods for remotely sensed data**. 2nd ed. Boca Raton: CRC Press, 2009. 356p.
- YANG, W.; YANG, L.; MERCHANT, J.W. An assessment of AVHRR/NDVI-ecoclimatological relations in Nebraska, USA. **International Journal of Remote Sensing**, v.18, p.2161-2180, 1997.
- ZELL, A.; MAMIER, G.; VOGT, M.; MACHE, N.; HUBNER, R.; DORING, S.; HERRMANN, K.; SOYEZ, T.; SCHMALZL, M.; SOMMER, T.; HATZIGEORGIOU, A.; POSSELT, D.; SCHREINER, T.; KETT, B.; CLEMENTE, G.; WIELAND, J.; GATTER, J. **Stuttgart Neural Network Simulator: user manual**. Version 4.2. Stuttgart: University of Stuttgart, 1996. 338p.
- ZHU, A.X. Mapping soil landscape as spatial continua: the neural network approach. **Water Resources Research**, v.36, p.663-677, 2000.

Recebido em 18 de agosto de 2009 e aprovado em 18 de abril de 2010