

PEDOMETRIA APOIADA EM ATRIBUTOS TOPOGRÁFICOS, ÍNDICES ESPECTRAIS E GEOLOGIA COM USO DE REDES NEURAIIS ARTIFICIAIS

Angelo Evaristo SIRTOLI¹, Claudinei Taborda da SILVEIRA², Mosar Faria BOTELHO³,
Luiz Eduardo MANTOVANI⁴

- (1) Universidade Federal do Paraná – UFPR, Setor de Ciências Agrárias, Departamento de Solos.
Campus das Agrárias, rua dos Funcionários 1540, CEP 80035 050. Curitiba – PR. Endereço eletrônico: asirtoli@ufpr.br
(2) Universidade Federal do Paraná – UFPR, Setor de Ciências da Terra, Departamento de Geografia. Campus Centro Politécnico,
rua Francisco H. dos Santos 100, CEP 81531-980. Curitiba – PR. Endereço eletrônico: claudineits@ufpr.br
(3) Instituto Federal Sul de Minas – IFSULDEMINAS. Endereço eletrônico: mfbotelho@yahoo.com.br
(4) Universidade Federal do Paraná – UFPR, Setor de Ciências da Terra Departamento de Geologia.
Campus Centro Politécnico, rua Francisco H. dos Santos 100, CEP 81531-980. Curitiba – PR. Endereço eletrônico: lem@ufpr.br

Introdução
Área de estudo
Método e materiais
Geração das variáveis
Levantamento de dados de amostragem e verificação
Treinamento
Arquitetura da RNA
Teste de validação cruzada
Generalização e avaliação
Resultados e discussão
Resultado do treinamento e classificação da RNA
Análise da acurácia da classificação
Conclusões
Referências bibliográficas

RESUMO - Os levantamentos de solos e os mapas resultantes são produzidos, em sua grande maioria, por métodos convencionais. A delimitação das unidades de mapeamento dos solos, nestes mapas, depende grandemente da subjetividade e experiência do analista. Visando reduzir a subjetividade e integrar dados de diferentes origens para auxiliar os procedimentos básicos de delimitação de unidades de mapeamento preliminares de solos, foi utilizada a técnica de redes neurais artificiais (RNA). Foram integrados atributos do terreno, geologia e índices espectrais através de uma rede *perceptron* multicamadas com alimentação adiante e algoritmo de aprendizado de retropropagação de erro otimizado por *resilient*. Na RNA, foram utilizadas 12 variáveis ambientais combinadas em diferentes conjuntos. As variáveis testadas foram: geologia, declividade, orientação das vertentes, elevação, perfil e plano de curvatura, radiação solar anual incidente, índice de umidade, de capacidade de transporte de sedimentos, de corrente de máximo fluxo, NDVI e *iron oxide*. As melhores classificações executadas pela rede, quando comparadas à área de verificação, resultaram em coeficientes de correlação acima de 0,8. No entanto o delimitamento das unidades de solos com a utilização de todas as variáveis, ou somente com atributos secundários do relevo, geologia e índices espectrais foram semelhantes quando comparadas a 95% de probabilidade. Os resultados alcançados permitem afirmar que, a RNA possibilita a integração de diferentes variáveis para o processo de delimitamento de unidades preliminares em mapas de solos. Nos mapas obtidos pela RNA, referentes a combinações dos diferentes conjuntos testados, os maiores valores de índice kappa e exatidão global apresentaram também, nas unidades de mapeamento, maiores índices de exatidão do ponto de vista do usuário. Todos os mapas resultantes da classificação por RNA apresentaram mais detalhes espaciais do que o mapa de solos convencional e, em todos os conjuntos testados, a supressão da variável perfil da curvatura ou da variável secundária corrente de máximo fluxo, derivada do perfil da curvatura, melhorou o resultado da classificação.

Palavras – chave: RNA, atributos do relevo, índices espectrais, levantamento de solos.

ABSTRACT - The soil survey and the resultant soil maps are produced, in his majority, by conventional approaches. The mapping units delimitation in these maps, depends on the technical experience of the person that will be carrying out the task. Aiming to reduce the subjectivity in that task and integrate facts of different origins to define preliminary mapping units of soils, it was utilized the technics of artificial neural network (ANN). In the ANN here applied 12 environmental variables combined were utilized in different ways. The selected variables were: geology, slope, aspect, elevation, profile and plan of curvature, annual solar radiation, and some secondary topographic index: wetness, sediments transport capacity, the stream power index and the spectral indexes ndvi and iron oxide. The results are going to affirm that, to ANN enables the integration of different variables for the trial delimitation of preliminary units in soils maps . In the maps obtained by the ANN, regarding the combination of the different selected sets, the largest values of index kappa and global accuracy present also, in the units of mapping, larger indices of accuracy of the viewpoint of the user. All the resultant maps from the classification by RNA present more spatial details than the conventional soil map.

Key Words: ANN, topographic attributes, spectral indexes, soil survey.

INTRODUÇÃO

A tarefa de execução de trabalhos de mapeamento de solos, seja por meio de imagens aéreas ou orbitais, requer um conjunto de julgamentos subjetivos, conforme é destacado por Way (1978). Portanto, a delimitação de unidades de solos depende do conhecimento tácito adquirido através da experiência do pesquisador.

Exemplo prático dessa subjetividade foi comprovado por Bie & Beckett (1973), quando compararam quatro mapas de solos produzidos a partir de fotointerpretação por quatro mapeadores de solos que trabalharam independentemente e sem nenhuma restrição estabelecida. Os resultados demonstraram que os mapas produzidos eram consideravelmente diferentes entre si, concluindo que a porcentagem de pureza e o grau da variabilidade das propriedades dos solos, dentro das unidades de mapeamento, foram menores do que o da paisagem como um todo.

Essa experiência aponta para a necessidade de introduzir na pesquisa cartográfica dos solos, ferramentas que possam auxiliar os procedimentos básicos de delimitação em modelos preditivos amparados em técnicas pedométricas, permitindo a redução da subjetividade. A Pedometria é definida por Hengl (2003) como a modelagem quantitativa e digital dos solos

O desenvolvimento de técnicas de mapeamento digital para delimitar unidades preliminares de mapeamento dos solos, potencialmente empregada no suporte aos trabalhos de levantamentos e apoio à cartografia pedológica, é de elevada importância e converge para a necessidade de introduzir na pesquisa cartográfica dos solos elementos que possam auxiliar nos procedimentos de delimitação das unidades.

Com o avanço da tecnologia na informática e o surgimento dos sistemas de Sensoriamento Remoto de alta resolução espacial, tornou-se possível a análise mais detalhada de imagens, onde objetos menores estão mais evidenciados, fato que vem provocando uma evolução nos processos de geração e manutenção de bases cartográficas, como também nos processos de classificação digital de imagens. Porém, em uma abordagem de classificação semi-

automática, o uso de imagens espectrais orbitais, muitas vezes não é suficiente para distinguir as feições do ambiente relacionadas aos solos, assim a inserção de informações auxiliares, como a modelagem digital do relevo, torna-se necessária.

A modelagem digital do relevo é uma das técnicas quantitativas melhor desenvolvidas para prever atributos e classes de solos (McKenzie *et al.*, 2000). Ela se utiliza da parametrização do relevo, geomorfometria, (Wood, 1996), onde são obtidos atributos topográficos de um Modelo Digital do Terreno (MDT), cujas primeiras derivadas são os atributos primários e de segunda derivada os secundários (Moore *et al.*, 1993; Wilson e Gallant, 2000). A aplicação dos atributos topográficos apresenta grande potencial na distinção de unidades de solos.

O emprego de técnicas de pedometria tem sido combinado com Sistemas de Informações Geográficas (SIG) e técnicas de inteligência artificial (ex. as Redes Neurais Artificiais - RNA), e árvores de decisão (Mendonça-Santos e Santos, 2003), no intuito de descrever, classificar e estudar os padrões de variação espacial dos solos.

Dentre esses métodos de integração de variáveis na pedometria, a RNA emerge como uma alternativa vantajosa aos classificadores de imagens baseados em conceitos estatísticos, uma vez que nenhuma hipótese prévia sobre a distribuição dos dados, a serem classificados, é exigida. Outra vantagem é que os dados de entrada podem ter escalas diferentes.

A RNA é um processo matemático para simular, em parte, o funcionamento do cérebro humano por meio de uma rede de elementos de processamento unitários que juntos, e funcionando em paralelo, simulariam o fluxo de “informações” do cérebro humano (Mcbratney *et al.* 2003).

Para Souza *et al.* (2006), a propriedade mais importante da RNA é a habilidade de aprender seu ambiente e com isso melhorar seu desempenho. Este aprendizado é realizado por meio de processos iterativos de ajustes aplicados a pesos, o chamado treinamento. O aprendizado só ocorre quando a RNA atinge uma solução generalizada para uma classe de

problemas. A maioria dos modelos de RNA possui uma regra de treinamento, na qual os pesos de suas conexões são ajustados de acordo com os padrões apresentados, isso afirma que elas aprendem por meio de exemplos. A RNA passa pelo processo de treinamento a partir de um conjunto de dados reais, diferindo da computação programada, a qual necessita de um conjunto de regras pré-fixadas e algoritmos.

Devido a grande maioria dos problemas práticos não serem linearmente separáveis, a arquitetura de rede perceptron multicamadas gera interesse pela sua habilidade em aprender e reconhecer padrões dessa natureza.

Alguns exemplos de experiências que utilizaram RNA na cartografia de solos, são Chagas (2006), que considerou combinações entre as variáveis discriminantes: geologia, elevação, declividade, orientação das vertentes, plano de curvatura, índice de umidade e três índices derivados de uma imagem do sensor ETM+ do Landsat-7. Resultados satisfatórios foram alcançados apenas para o primeiro nível categórico, no SiBCS (EMBRAPA, 2006), nos demais níveis a concordância foi inferior a 50%, evidenciando, ainda mais, as discrepâncias, principalmente no 4º nível categórico, na qual a concordância foi de apenas 23,27%.

Silveira (2010) utilizou RNA para integrar oito variáveis, sete atributos topográficos (hipsometria, perfil e plano de curvatura, declividade, índices de corrente de máximo fluxo, umidade e de capacidade de transporte de sedimentos) e a geologia, para delimitar unidades preliminares de mapeamento de solos na bacia do arroio do Corvo/PR em nível de Grande Grupo. Na validação as unidades

mapeadas foram comparadas com mapas convencionais de solos pré existentes na área de estudo. Os melhores resultados foram com o conjunto das oito variáveis na RNA com um mapa pré existente na escala 1:10.000 cujo valor Kappa foi de 0,56 e a Exatidão Global de 0,72.

Aplicações de RNA para delimitar unidades de solos foram realizadas por Sirtoli (2008), onde empregou-se como variáveis um conjunto de atributos topográficos primários e secundários, índices espectrais extraídos do sensor orbital Ikonos II e geologia que foram integrados por RNA para mapear unidades de solos.

Nesse contexto, visando contribuir com o desenvolvimento de técnicas de pedometria para apoiar a identificação de unidade de solos, o presente trabalho teve por objetivo utilizar uma RNA perceptron multicamadas, alimentado adiante, utilizando o algoritmo de aprendizado retro-propagação para delineamento de unidades de mapeamento preliminares de solos, empregando-a na integração de dados de índices espectrais derivados da imagem espectral Ikonos, geologia e atributos topográficos obtidos de um MDT.

Área de estudo

A área de estudo representa uma quadrícula de 1 km², situada na bacia do rio Canguiri, afluente do rio Iraí, pertencente ao sistema de drenagem do alto curso do rio Iguaçu, no município de Pinhais, estado do Paraná (Fig. 1). Localiza-se entre as coordenadas 687974m e 689000m, 7189643m e 7190606m, do sistema de projeção UTM, no fuso -22, utilizando do Datum Horizontal SAD69.

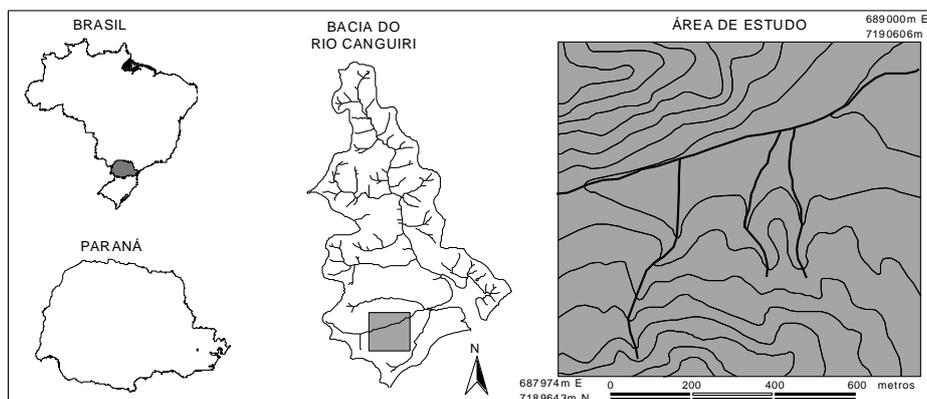


Figura 1. Localização da área de estudo, no detalhe curvas de nível e drenagem.

Sob o aspecto geológico a área está situada sobre a bacia sedimentar de Curitiba, em maior parte sobre a Formação Guabirota e na planície do canal de drenagem principal sobre Sedimentos Holocênicos. A Formação Guabirota foi depositada após sucessivas fases erosivas que desenvolveram a superfície do Alto Iguaçu, seguida por uma fase erosiva de clima úmido que dissecou esta superfície. Posteriormente, iniciou-se a deposição em ambiente semiárido, com chuvas torrenciais formando depósitos tipo *playa-lake*, constituídos por argilitos, arcósios, depósitos rudáceos e margas, com espessura variando de 10 a 80 m, sendo maior na região central (Minerpar, 2001). Esses depósitos assentam-se discordantemente sobre o embasamento, sendo compostos predominantemente por depósitos argilosos e silticos esverdeados que gradam para termos mais grossos de forma interdigitada ou abrupta, dependendo das características estratigráficas e estruturais locais, com intercalações de areias arcólicas e, mais raramente, de cascalheiras com predominância de seixos de quartzo (Salamuni, *et al.*, 1999). Os Sedimentos Holocênicos são oriundos de erosão e deposição das litologias mais antigas intemperizadas. O processo formador é hidráulico-deposicional, condicionado às calhas de drenagem dos rios e planícies de inundação. A formação destes depósitos continua

ocorrendo devido ao avanço do intemperismo, erosão e retrabalhamento dos sedimentos (Minerpar, 2001).

Sob o aspecto climático, prevalece na região a unidade climática Cfb, segundo a classificação de Köppen, que se caracteriza como quase sempre úmido, com características pluviais de clima quente-temperado. No mês mais quente podem ser registradas temperaturas médias inferiores a 22° C e o restante do ano com médias maiores que 10° C. Possui uma precipitação regular ao longo do ano, formação de geadas durante o inverno, raramente neva e pouco se percebe uma estação tipicamente seca (Maack, 1981).

Sob o aspecto da geomorfologia está inserida no Primeiro Planalto Paranaense, predominando na subunidade morfoescultural Planalto de Curitiba, segundo mapa geomorfológico do estado do Paraná (Santos *et al.*, 2006). Esse compartimento geomorfológico caracteriza-se por apresentar dissecação moderada composta por topos alongados e aplainados, vertentes convexas e formas de relevo predominantemente onduladas a suave onduladas. Secundariamente, nas áreas planas, ocorre a subunidade morfoescultural denominada de Planícies Fluviais sobre Sedimentos Inconsolidados do período Quaternário, cujos limites estão restritos ao leito do rio.

MÉTODO E MATERIAIS

Geração das variáveis

No delineamento das unidades de mapeamento preliminares de solos foram utilizadas doze variáveis, apresentadas na tabela 1. O objetivo da seleção dessas variáveis foi fundamentado no conceito solo-paisagem desenvolvido por Jenny (1941), segundo o qual o solo é resultado das interações entre os fatores de formação ao longo do tempo.

Nem todas as variáveis contempladas estão relacionadas com os fatores de formação dos solos, algumas estão relacionadas com os processos pedogenéticos (ex: adições coluvionares, perdas por erosão). Desse modo, as variáveis refletem ou fornecem indicativos das características dos solos nas diferentes unidades de paisagem.

Na RNA empregada a variável geologia (ge) foi estabelecida através do mapa geológico de escala 1:20.000, adaptado de Salamuni *et al.* (1999), uma vez que o material de origem é um dos condicionantes no processo de formação e na variabilidade espacial dos atributos do solo.

A partir dos dados vetoriais planialtimétricos, obtidas das folhas topográficas na escala 1:10.000 da COMEC (1976), foi produzido um MDT com resolução espacial de 5 metros. As operações foram executadas a partir da grade regular, nela foi executada uma varredura com janela móvel de nove nodos no qual foram computadas as variáveis (tabela 1).

Tabela 1. Variáveis utilizadas na RNA.

Variáveis	Descrição / significado	Referência
Declividade % (de)	$\beta = \arctan \left[\left(f_x^2 + f_y^2 \right)^{1/2} \right]$	Moore et al. (1993)
Orientação das vertentes (as)	$\psi = 180 - \arctan \left(\frac{f_y}{f_x} \right) + 90 \left(\frac{dx}{ f_x } \right)$	Moore et al. (1993)
Elevação (hi)	Elevação acima do nível do mar (m)	MDE de base cartográfica 1:10.000
Perfil da curvatura (pfc)	$Pfc = \frac{-2(f_{xx} f_x^2 + f_{xy} f_x f_y + f_{yy} f_y^2)}{f_x^2 + f_y^2}$	Moore et al. (1993)
Plano de curvatura (plc)	$Plc = \frac{2(f_{xx} f_y^2 - f_{xy} f_x f_y + f_{yy} f_x^2)}{f_x^2 + f_y^2}$	Moore et al. (1993)
Índice topográfico de umidade (iu)	$\omega = \ln \left(\frac{A_s}{\tan \beta} \right)$	Moore et al. (1993)
Radiação solar anual (rg)	Determinada em MJ cm ⁻² ano ⁻¹	McCune & Keon (2002).
Índice de capacidade de transporte de sedimentos (ts)	$\tau = \left(\frac{A_s}{22.13} \right)^m \left(\frac{\sin \beta}{0.0896} \right)^n$	Moore et al. (1993)
Índice de corrente de máximo fluxo (cmf)	$\Omega = A_s \tan \beta$	Moore et al. (1993)
Índice de vegetação (ndvi); "iron oxide" (io).	NDVI=1xb4-1xb3/1xb4+1xb3 IO=1xb3/1xb1	Campbell, (1987) Sabins (1997)
Geologia (ge)	Mapa geológico na escala 1:20 000 digitalizado	Salamuni et al. (1999)

Levantamento de dados de amostragem e verificação

Foram realizadas coletas de solo em campo com descrição morfológica, segundo procedimentos descritos em Santos et al., 2005, e utilizados dados secundários de amostragem

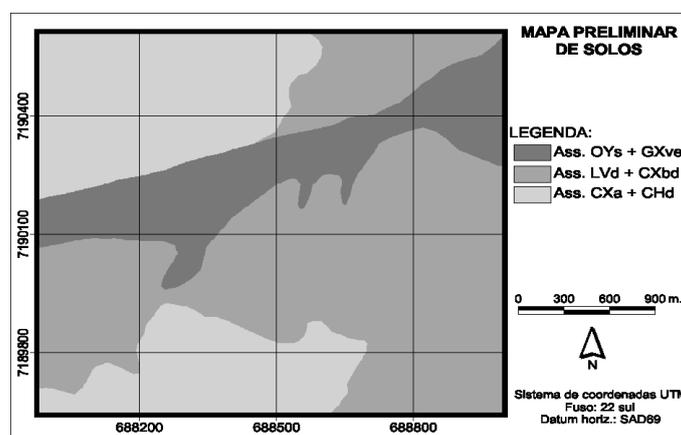
de seis perfis com caracterização química e física de Lima (2005). Com base nesses dados foram caracterizadas três classes de solos no quarto nível categórico (tabela 2), conforme o SiBCS de Embrapa (2006).

Tabela 2. Composição das unidades de mapeamento de solos (Um).

Unidade	Descrição
Um1	Associação Latossolo Vermelho Distrófico típico + Cambissolo Háptico Tb Distrófico latossólico (LVd+CXbd).
Um2	Associação Organosolo Mésico Sáprico típico + Gleissolo Háptico Ta Eutrófico típico (OYs+GXve).
Um3	Associação Cambissolo Háptico Alumínico típico + Cambissolo Húmico Distrófico típico (Cxa+CHd).

Como área de verificação, foi utilizado o mapa de solos (Fig. 2), previamente elaborado, por método convencional, com utilização de análise estereoscópica sobre fotografias aéreas pretas e brancas na escala 1:10.000 e

delineamento manual das unidades de solo-paisagem. Essa referência serviu para coleta de amostras de treinamento, para o processamento da classificação e a verificação das unidades de solos delineadas utilizando-se as RNA.

**Figura 2.** Unidades de mapeamento preliminares de solos.

Na imagem de verificação (Fig. 2), contendo as unidades de mapeamento de solos previamente delimitadas, foram definidos os elementos amostrais para treinamento da RNA e verificação. Para cada unidade de mapeamento foram selecionadas 200 amostras para o treinamento, sendo cada amostra composta por 3 subamostras de 3 *pixels* cada, buscando com isso aumentar a representatividade de cada subamostra. Tomou-se a precaução que as mesmas estivessem bem distribuídas na imagem e que fossem representativas.

Treinamento

As amostras de treinamento foram intercaladas, conforme quadro 1, para que o treinamento não se tornasse tendencioso. A quantidade de amostras utilizadas no treinamento foi definida a partir de testes preliminares. Nos testes com uma quantidade menor de amostras de treinamento por classe, não se atingiu um bom aprendizado da RNA, entretanto para um maior número, o aprendizado da tornou-se lento. Em cada unidade de mapeamento foram também coletadas mais 60 unidades amostrais estratificadas para a verificação.

Quadro 1. Amostra de dados para treinamento da RNA.

E	N	ov	de	ge	hi	pf	plc	rg	iu	ts	cmf	ndvi	io	um
688026	7190591	0	2	1	4	0,1	0	2	4,89108	0,40304	0,60857	0,58442	0,30573	1
688002	7190162	0	0	0	1	0,1	0	2	5,42684	0,20126	0,35618	0,41582	0,69565	3
688668	7190587	0	2	1	1	0	1	1	5,3567	0,22045	0,38206	0,18182	0,92093	2
688030	7190539	1	2	1	4	0,1	0	2	5,03491	0,33452	0,52706	0,58515	0,52198	3
688010	7190122	1	2	0	1	0	0	2	5,61385	0,15785	0,29543	0,53846	0,52941	2
688728	7190587	0	2	1	1	0,1	1	1	5,55202	0,17107	0,31429	0,13151	1,20532	1

Legenda: E=este, N=norte, ov=orientação das vertentes, de=declividade, ge=geologia, hi=hipsometria, pf=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, NDVI=índice de vegetação da diferença normalizada, io= iron oxide, um=unidade de mapeamento.

Arquitetura da RNA

Estruturar a arquitetura de uma RNA consiste em definir o número de camadas escondidas e a quantidade de neurônios por camada. Esta definição foi realizada por meio de uma série de experimentos, nos quais a arquitetura foi variada até encontrar a ideal, que garantisse a convergência em um número menor de iterações e com um resultado coerente com a área de verificação. A RNA foi treinada com duas camadas escondidas e, para todos os teste adotou-se um erro médio quadrático (rms) de 0,01, onde o tempo médio de processamento foi em torno de 30 segundos.

Neste estudo foi utilizada a arquitetura de multicamadas com alimentação adiante (*feed-forward*), com aprendizado de retro-propagação (*backpropagation*) e a função de transferência adotada foi a sigmóide. O método de aprendizado foi o *Backpropagation* otimizado por *Resilient*, desenvolvido por Reidmiller & Braun (1993), executado com o intuito de minimizar as oscilações nos ajustes dos pesos quando o treinamento estava sendo processado,

atingindo assim, o ponto de mínimo em uma quantidade menor de iterações com um erro médio quadrático menor (Ribeiro, 2003).

Depois de definida a quantidade de camadas escondidas foi necessário determinar, experimentalmente, o número de nós por camada. Para tanto, considerou-se a regra de Hilton (1989) que comenta que, no caso de duas camadas intermediárias, o número de nós da segunda não deve exceder 1/3 do conjunto de nós da primeira.

Para a predição das unidades de mapeamento dos solos por RNA, foi utilizado o programa MatLab na versão 7 (Matworks, 2007), que possui um módulo de redes neurais, denominado de *Neural network toolbox for use in MatLab*. A grande flexibilidade na manipulação e formatação de arquivos e diversidade de algoritmos implementados eficientemente (funções de ativação e transferência) justificou a escolha do programa.

Considerando como área de verificação o mapa de solos convencional (fig. 2), as variáveis utilizadas na classificação da RNA

foram combinadas em diferentes conjuntos e avaliadas quanto à capacidade de discriminação das unidades de solos.

Tendo-se as variáveis pré-processadas e utilizando-se como nível de abstração o *pixel*, coletou-se as amostras de treinamento e treinou-se a RNA, executando-se testes para verificar quais combinações de variáveis ofereceram melhor resposta.

Teste de validação cruzada

O método de validação cruzada múltiplo conhecido como deixo um de fora (Haykin,

2001) foi utilizado. Neste caso, um conjunto de variáveis (cjn) é usado para treinar o modelo (Quad. 2) e efetuar a primeira modelagem. Em seguida, uma variável é substituída pela imagem zero, representada pelo sinal negativo em frente a mesma. Na próxima modelagem, a variável anterior é reinsertada e uma segunda é substituída pela imagem zero, e assim sucessivamente para todas as variáveis. Este procedimento buscou avaliar a importância da variável deixada de fora na modelagem em questão.

Quadro 2. Exemplo do método de validação cruzada múltiplo.

Conjunto/variáveis	Variáveis				Simulações
cjn	iu	ts	cmf	ge	1ª modelagem
- iu	iu	ts	cmf	zero	2ª modelagem
- ts	iu	ts	zero	ge	3ª modelagem
-cmf	iu	zero	cmf	ge	4ª modelagem
-ge	zero	ts	cmf	ge	5ª modelagem

Legenda: iu=índice de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=índice de corrente de máximo fluxo, ge=geologia.

Generalização e avaliação

Depois do treinamento foi efetuada a generalização que consiste na aplicação do modelo matemático a todos os dados originais, não apenas aqueles usados no treinamento. Para cada conjunto de variáveis foi executada a generalização, a qual gerou uma imagem classificada. Esta imagem foi submetida a um filtro não linear de mediana, para retirada de ruídos. A escolha desse filtro deu-se por que as bordas não são degradadas em extremo, e os valores originais são preservados.

A avaliação dos resultados da classificação constou da determinação do nível de exatidão através do coeficiente Kappa, da exatidão

global e acurácia do ponto de vista do usuário, através da análise da matriz de confusão.

Também foi utilizado um índice de qualidade universal de imagem (índice q), proposto por Wang & Bovik (2002). Esse índice usa uma combinação de três fatores: perda da correlação, da distorção da luminância e da distorção do contraste. Neste caso, por tratar-se de imagem temática, somente foi considerada a correlação. Assim, este índice foi utilizado para medir o grau de afinidade ou similaridade entre a imagem classificada e a imagem de verificação, apresentando uma faixa de variação de 0 a 1. O valor 0 indica nenhuma correlação, enquanto que o valor 1, perfeita correlação.

RESULTADOS E DISCUSSÃO

Resultado do treinamento e classificação da RNA

Dentre as inúmeras combinações testadas, as dos conjuntos ilustrados no quadro 3, apresentaram melhor potencial na predição das

unidades de mapeamento de solos, pois possuíram porcentagem de acertos superior a 90% das amostras de treinamento e menor erro médio quadrático (rms).

Quadro 3. Conjuntos utilizados e respectivas variáveis ambientais.

Conjuntos	Variáveis ambientais
cj1	ov, de, hi, pfc, plc, rg, iu, ts, cmf, ge, ndvi, io
cj2	ov, de, hi, pfc, plc, rg, ge, ndvi, io
cj3	iu, ts, cmf, ge, ndvi, io
cj4	ov, de, hi, pfc, plc, rg, iu, ts, cmf, ge

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, NDVI=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia.

Os melhores resultados de arquitetura (Tab. 3) foram com a camada de entrada em função das variáveis de entrada, duas camadas

escondidas e a camada de saída em função das unidades de mapeamento de solos consideradas.

Tabela 3. Identificação dos melhores conjuntos de variáveis e respectiva arquitetura da RNA

Conjuntos	Neurônios	Iterações	% de acertos no	
	camadas escondidas		treinamento	rms
cj1	100, 35	47	93,89	0,0539
cj2	60, 18	55	96,12	0,0609
cj3	29, 9	148	90	0,0812
cj4	81, 20	56	93,89	0,0604

Nas tabelas 4 e 5 verificou-se que ao utilizar todas as variáveis do cj1 e cj2, como camada de entrada da rede, e na substituição de cada uma delas pela imagem zero, as iterações variaram de 36 a 119 no primeiro conjunto e, de 53 a 455 no segundo conjunto, demonstrando que a RNA convergiu rapidamente.

Ao se analisar a taxa de *pixels* corretamente classificada no treinamento (Tab. 4) verificou-se que esta variou de 86 a 96 nas substituições respectivas de hi e de no cj1. No cj2 (Tab. 5), a taxa variou de 83 na substituição de ge a 96 na substituição de pfc.

Foi realizada uma inferência estatística a partir da qual se verificou que, em todos os conjuntos avaliados, os valores de rms estão dentro de um intervalo de confiança de 95%. Os valores de rms no cj1 (Tab. 4) foram muito semelhantes com a utilização de todas as variáveis ou na substituição de cada uma pela imagem zero. A exceção ocorreu na substituição da variável hi, onde o valor de rms subiu para 0,1377. Este resultado influenciou o índice q, o qual resultou no valor mais baixo desse conjunto.

Sobre os valores dos índices q, em todos os conjuntos estudados e, na substituição de suas variáveis, foi realizada uma inferência estatística, a partir da qual também se verificou que todos os valores apresentados estão dentro de um intervalo de confiança de 95%.

Ao se substituir a variável pfc no primeiro conjunto (Tab. 4), os valores do índice q resultaram em 0,8837. Na substituição da variável hi o índice q resultou em seu menor valor, 0,7194.

No segundo conjunto (Tab. 5), os valores do índice q variaram de 0,8821 a 0,7349 nas substituições respectivas de pfc e rg. Assim, pode-se afirmar que os resultados, obtidos nesses dois conjuntos, refletem a influência negativa que o perfil da curvatura (pfc) exerce na delimitação das unidades preliminares de mapeamento de solos. A sua substituição, nos dois conjuntos, permitiu que houvesse uma melhora na predição das unidades de solos. Esse melhor índice q ocorreu por que, no delineamento convencional das unidades de solos, o mapeador tem dificuldade em considerar este aspecto, mesmo se utilizasse uma imagem de alta resolução associada a um MDT onde é possível visualizar detalhes da encosta. Por outro lado, num levantamento utilizando produtos em escala maior, com imagens de alta resolução, associado um MDT equivalente, essa variável ambiental pode ser de grande importância. Nesse sentido Pennock et al. (1987) afirma que a configuração da encosta de acordo com a curvatura pode ser usada para identificar áreas de diferentes regimes de umidade, permitindo uma avaliação mais acurada da distribuição dos solos.

Tabela 4. Experimentos realizados com o (cj1) e substituindo suas variáveis.

Conjunto/variáveis	Iterações	% acertos treinamento	rms	Índice q
cj1	50	93,34	0,0539	0,7566
- io	47	95	0,0554	0,842
- ndvi	55	93,89	0,0623	0,8193
- cmf	36	94,45	0,0524	0,809
- ts	48	93,89	0,0564	0,7565
- iu	50	91,67	0,0468	0,7557
- rg	63	91,67	0,0675	0,7832
- plc	43	95	0,0479	0,7504
- pfc	50	95	0,0649	0,8837
- hi	119	86,12	0,1377	0,7194
- ge	70	87,68	0,0993	0,7585
- de	44	96,67	0,0496	0,7632
- ov	70	93,94	0,0514	0,7557

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ndvi=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia, um=unidade de mapeamento.

No cj2 (Tab. 5) o rms resultou em valores muito semelhantes entre si, ocorrendo valores díspares somente quando da substituição das variáveis hi e ge. Porém o índice q, resultante da substituição dessas duas variáveis, não foi menor como seria o esperado. Esses resultados indicam a necessidade de uma segunda avaliação das imagens classificadas através do

coeficiente kappa, exatidão global e exatidão do ponto de vista do usuário. Embora o rms seja utilizado como uma medida do desempenho das redes neurais, nesse conjunto, não se mostrou adequado para a escolha do melhor conjunto de variáveis para a predição das unidades de mapeamento preliminares de solos.

Tabela 5. Experimentos realizados com o (cj2) e substituindo suas variáveis.

Conjunto/variáveis	Iterações	% acertos no treinamento	rms	Índice q
cj2	55	96,12	0,0609	0,7561
- io	84	92,78	0,0675	0,8112
- ndvi	74	94,45	0,0660	0,8029
- rg	83	91,12	0,0619	0,7349
- plc	61	93,34	0,0541	0,7557
- pfc	53	96,12	0,0540	0,8821
- hi	455	80,56	0,1528	0,7877
- ge	115	83,89	0,1473	0,741
- de	94	90,56	0,0934	0,7563
- ov	86	93,34	0,0599	0,7781

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ndvi=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia, um=unidade de mapeamento.

As imagens apresentadas na figura 3 são resultantes da integração das variáveis do primeiro conjunto. Neste, foram avaliadas 12 variáveis: atributos primários e secundários do relevo, geologia e índices espectrais derivados

de imagens Ikonos. No segundo conjunto (Fig. 4), as imagens são resultantes da integração de 9 variáveis ambientais entre atributos primários do relevo, geologia e os índices espectrais.

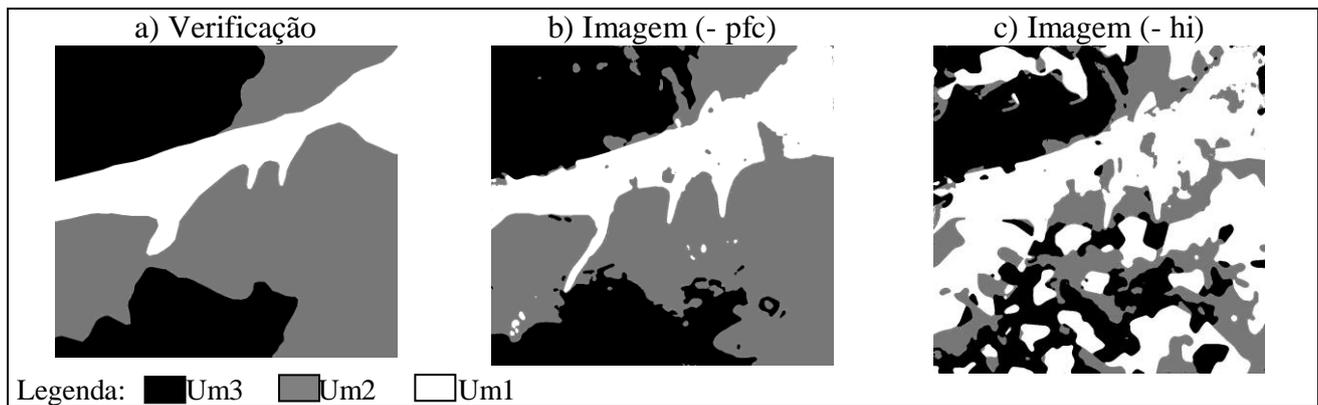


Figura 3. imagem de verificação (a), imagem simulada do cj1 com maior índice q (b) e imagem com menor índice q (c).

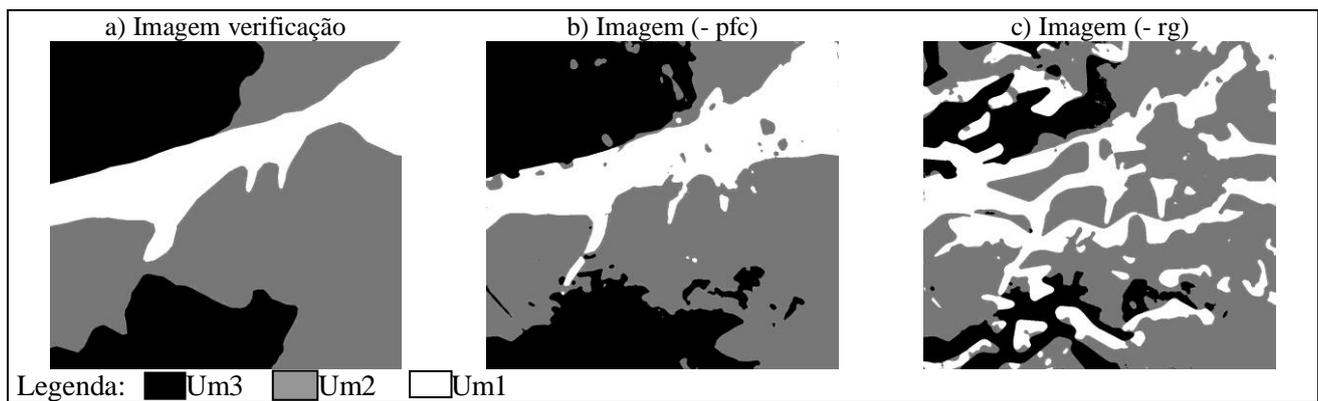


Figura 4. Imagem de verificação (a), imagem simulada do cj2 com maior índice q (b) e imagem com menor índice q (c).

Avaliando os melhores resultados obtidos nos dois primeiros conjuntos, tanto através de uma análise visual (Fig. 3 e 4), como considerando o índice q (Tab. 4 e 5), conclui-se que as imagens que apresentaram maiores valores desse índice são muito semelhantes entre si e apresentam respectivamente índice q 0,8837 e 0,8821. Isso mostra que, apesar do aumento da dimensionalidade dos dados, no conjunto 1, não houve significativa melhora da qualidade da imagem classificada. Isso pode ser creditado ao fenômeno de Hughes que ocorre devido ao número limitado de amostras frente à inclusão de mais variáveis.

Nas Tabelas 6 e 7 verificou-se que, colocando todas as variáveis utilizadas nos conjuntos 3 e 4 como camadas de entrada da RNA e, na substituição de cada uma delas pela imagem zero, as iterações variaram no cj3 de 148 a 561 e no cj4 de 56 a 678. Esses resultados demonstram que a rede convergiu rapidamente, conforme observado também nos conjuntos 1 e 2.

Nesses dois conjuntos (cj3 e cj4) a menor taxa de *pixels* corretamente classificada no

treinamento, 75 no primeiro e 80 no segundo, ocorreu quando foi substituída a geologia (ge), demonstrando a importância desta variável, na classificação por RNA, quando da combinação dos atributos desses dois conjuntos.

Quando foram realizados testes com atributos secundários do relevo associados à geologia e índices espectrais derivados de imagens Ikonos, a maior quantidade de acertos no treinamento (Tab. 6) ocorreu na substituição da imagem referente ao índice de vegetação (*ndvi*). Este resultado demonstra que, o índice espectral (*ndvi*) influencia negativamente na classificação por RNA neste conjunto de dados, pois na sua substituição o índice q resultante foi bom, apesar de não ser estatisticamente diferente do melhor resultado a 95% de probabilidade. Nesse sentido, esse resultado concorda com Dobos et al. (2001), que afirmam serem os dados de sensores remotos grandemente influenciados pela variabilidade do terreno, porém, ainda não conseguem identificar com precisão toda a variabilidade do solo que ocorre na paisagem.

Tabela 6. Experimentos realizados com o (cj3) e substituindo suas variáveis.

Conjunto/variáveis	Iterações	% acertos no treinamento	rms	Índice q
cj3	148	90	0,0893	0,8446
- io	561	83,89	0,1220	0,8212
- ndvi	312	92,78	0,0778	0,8441
- cmf	225	91,12	0,0720	0,8529
- ts	191	92,23	0,0757	0,8427
- iu	174	88,89	0,1339	0,8464
- ge	200	75	0,4609	0,7611

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ndvi=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia, um=unidade de mapeamento.

No cj4, maior porcentagem de acertos no treinamento ocorreu na substituição do as pela imagem zero. Porém, na análise do índice q, este resultou em pior resultado embora que,

estatisticamente a 95% de probabilidade, não difere do melhor, quando foi substituída a variável pfc por zero (Tab. 7).

Tabela 7. Experimentos realizados com o (cj4) e substituindo suas variáveis.

Conjunto/variáveis	Iterações	% acertos no treinamento	rms	Índice q
cj4	56	93,89	0,0604	0,822
- cmf	65	92,23	0,0695	0,7576
- ts	86	92,78	0,0530	0,7147
- iu	77	93,34	0,0651	0,714
- rg	100	92,78	0,0696	0,7241
- plc	61	93,89	0,0496	0,7302
- pfc	61	92,23	0,0568	0,8774
- hi	678	83,34	0,1595	0,7495
- ge	213	80,56	0,1612	0,802
- de	74	90,56	0,0776	0,7203
- as	129	96,12	0,0468	0,7119

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ndvi=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia, um=unidade de mapeamento.

O índice q variou de 0,8529 a 0,7611 no cj3, nas substituições respectivas de cmf e ge. No cj4 variou de 0,8774 a 0,7119 nas substituições de pfc e as. As melhores imagens resultantes apresentaram resultados visuais semelhantes aos observados sobre os cj1 e cj2. Pelo fato das imagens simuladas não serem estatisticamente diferentes a 95% de probabilidade, demonstra novamente a ocorrência do fenômeno de hughes. Pois, apesar do aumento da

dimensionalidade dos dados do cj 4 em relação ao cj 3, não houve melhora significativa na predição das unidades de solos.

Independente do índice q, as imagens produzidas pela RNA, tanto as ilustradas nas figuras 3 e 4, como nas 5 e 6, apresentaram muito mais detalhes espaciais - o que concorda com a realidade de campo - do que o mapa de solos convencional.

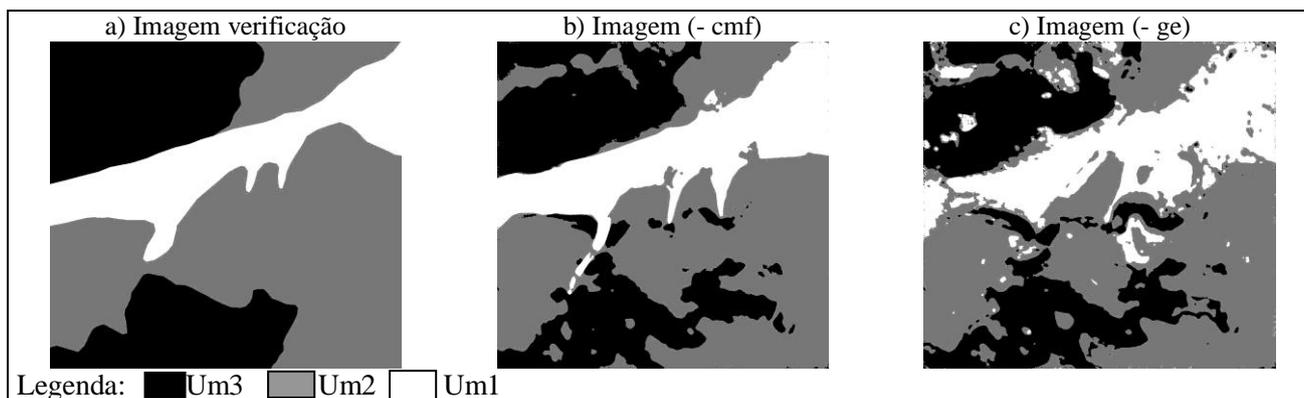


Figura 5. Imagem de verificação (a), imagem simulada do cj3 com maior índice q (b) e imagem com menor índice q (c).

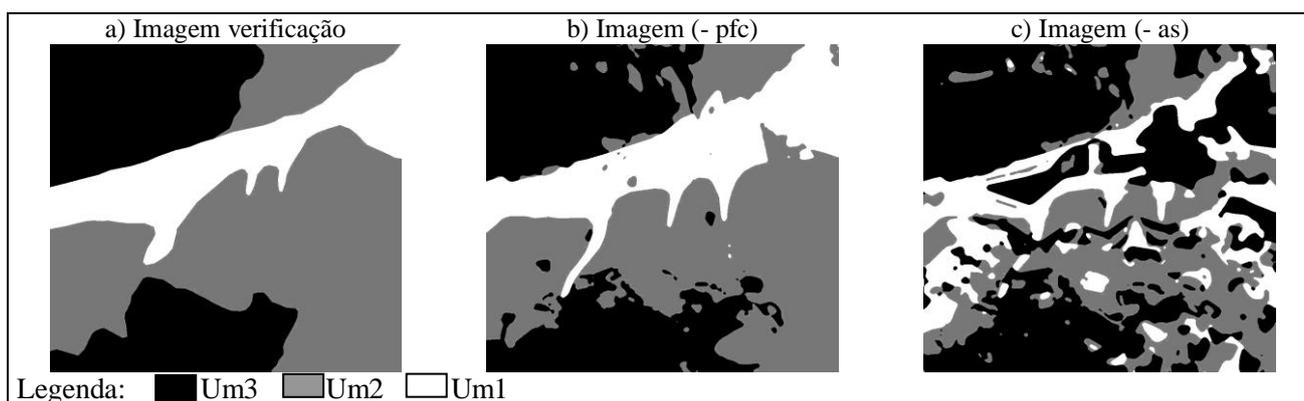


Figura 6. imagem de verificação (a), imagem simulada do cj4 com maior índice q (b) e imagem com menor índice q (c).

Esse detalhamento se deve ao fato que a variação ambiental pôde ser caracterizada quantitativamente e com maior precisão devido à escala das informações básicas e, por se tratar de interpretação não sujeita aos efeitos da subjetividade e generalização de detalhes pelo analista. Assim, com a integração das informações através de RNA é possível identificar pequenos componentes na paisagem que são difíceis de serem identificados no mapeamento convencional, mesmo utilizando produtos com alta resolução espacial onde os detalhes são visíveis, mas dependentes de conhecimento e interpretação do analista.

A imagem classificada (Fig. 5) que apresentou menor valor do índice q, corresponde ao maior rms apresentado no conjunto no cj3 (Tab. 6). Por outro lado, no cj4 o menor índice q (Fig. 6) não correspondeu ao maior rms (Tab. 7). Esses resultados também evidenciam a necessidade da utilização de outras formas de avaliação das imagens classificadas por RNA.

Análise da acurácia da classificação

A acurácia dos mapas de solos gerados pelas RNA, também foi avaliada através do coeficiente Kappa, da acurácia global e da acurácia do ponto de vista do usuário.

O desempenho da classificação medido pelo coeficiente kappa para o cj1 e cj2 (Tab. 8 e 9), oscilou em função das variáveis utilizadas. Os valores apresentados variaram desde razoáveis até muito bons, segundo classificação proposta por Landis & Koch (1977). O melhor desempenho ocorreu quando da substituição da variável pfc nos dois conjuntos. O pior desempenho aconteceu quando da substituição da variável as no cj1 e ge no cj2.

Os resultados, referentes ao cj1, ilustrados na tabela 8, apontam que, na combinação de variáveis onde foi substituída a imagem referente ao perfil de curvatura (pfc) pela imagem nula, os resultados da exatidão global foram maiores. No entanto, essa exatidão foi menor quando da substituição da variável elevação (hi). Concorda, este resultado, com o maior rms calculado nesse conjunto.

Tabela 8. Resultado da classificação por RNA para o (cj1) nas diferentes combinações.

Conjunto/variáveis	Kappa	Exatidão global	Usuário		
			Um1	Um2	Um3
cj1	0,3701	59,09	94,16	42,69	41,63
- io	0,6649	79,5	66,36	83,68	91,69
- ndvi	0,5842	75,59	75,82	86,89	42,41
- cmf	0,5187	72,22	65,3	87,02	42,11
- ts	0,364	58,5	94,99	40,77	42,29
- iu	0,3632	58,66	94,56	41,76	41,1
- rg	0,4764	67,65	92,66	60,36	42,44
- plc	0,4689	66,85	89,3	60,7	43,07
- pfc	0,8172	88,56	94,47	84,22	90,18
- hi	0,3722	56,09	67,87	39,39	82,68
- ge	0,3787	60,25	93,67	45,31	41,6
- de	0,3713	59,21	94,58	42,42	42,36
- ov	0,3567	57,96	93,83	40,23	42,9

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ndvi=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia, um=unidade de mapeamento.

A acurácia do ponto de vista do usuário em relação a cada uma das unidades de mapeamento consideradas foi diferente, principalmente na (um1) apresentada na tabela 8. Nesta os resultados foram muito semelhantes na utilização de todas as variáveis e nas substituições de: ts, iu, pfc e de. Esses resultados demonstram que, o peso dessas variáveis é semelhante na delimitação desta unidade de solos que se encontra em relevo mais suavizado.

Na (um2) o melhor resultado aconteceu na retirada do atributo topográfico secundário cmf (Tab. 8). Embora seja bem conhecida a forte influência dos atributos do terreno sobre a variação espacial dos solos (Park & Vlek, 2002), neste conjunto a sua presença influenciou negativamente a modelagem, provavelmente pelo excessivo peso dado a essa variável pela rede.

A (um3) apresentou seu melhor resultado na substituição de io. A influência negativa dessa variável pode estar relacionada ao fato que, os índices espectrais que se referem às propriedades dos solos perdem significado quando o mesmo não se encontra descoberto.

No entanto, em função da maior exatidão do ponto de vista do usuário para todas as unidades de mapeamento, de forma global, os melhores resultados ocorreram quando da substituição da variável perfil da curvatura (pfc).

Para o cj2 (Tab. 9), a exatidão global apresentou maior valor quando da substituição da variável pfc e menor quando foi substituída a variável ge. Essa menor exatidão demonstra que, no conjunto avaliado, o material de origem tem grande peso na predição das unidades de solos.

Tabela 9. Resultado da classificação por RNA para o (cj2) nas diferentes combinações.

Conjunto/variáveis	Kappa	Exatidão global	Usuário		
			Um1	Um2	Um3
cj2	0,3456	57,32	93,05	39,5	41,72
- io	0,5263	72,82	64,01	89,14	41,82
- ndvi	0,5174	67,82	94,61	42,33	92,04
- rg	0,3647	60,62	45,86	73,54	50,47
- plc	0,3562	58,14	94,29	40,91	41,08
- pfc	0,8072	87,93	93,78	83,4	90,17
- hi	0,5139	70,55	45,71	78,83	92,5
- ge	0,3331	56,44	92,23	39,43	39,4
- de	0,5288	70,19	79,02	60,97	80,51
- ov	0,4914	69,8	50,14	82,1	70,53

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ndvi=índice de vegetação da diferença normalizada, io= iron oxide, ge=geologia, um=unidade de mapeamento.

No entanto, para a unidade 1 o melhor resultado foi na substituição do *ndvi*. Esse índice, embora muito utilizado como variável auxiliar para a predição de características ou classes de solos (Zhu et al., 2001), quando associado às outras as variáveis do conjunto produziu menor desempenho para esta unidade. Esse resultado pode estar relacionado com a baixa diferenciação deste índice entre as unidades de mapeamento consideradas, pois a vegetação local encontra-se bastante alterada. O pior resultado, nessa unidade, ocorreu na substituição de *hi*. Esse resultado pode estar relacionado a baixa amplitude altimétrica encontrada nessa unidade de mapeamento e, por consequência sua pouca importância na predição da unidade de solos.

Para a unidade 2 o melhor resultado foi na substituição de *io*, demonstrando também que

este índice espectral perde importância quando o solo não se encontra descoberto. A pior exatidão do ponto de vista do usuário ocorreu na utilização de todas as variáveis do conjunto.

Na unidade 3 o melhor resultado ocorreu na substituição da variável *hi* e o pior na substituição da *ge*. Isso demonstra que, nesse conjunto, a presença da elevação influencia negativamente a predição desta unidade de mapeamento. Ao mesmo tempo, esse conjunto é dependente da geologia, pois a sua retirada da rede acarreta na redução da exatidão da imagem classificada.

Os melhores resultados de coeficiente kappa aconteceram na substituição das variáveis *cmf* e *pfc* respectivamente nos *cj3* e *cj4* (Tab. 10 e 11). Os piores coeficientes foram medidos nas substituições de *ge* no *cj3* e *iu* no *cj4*.

Tabela 10. Resultado da classificação por RNA para o (*cj3*) nas diferentes combinações.

Conjunto/variáveis	Kappa	Exatidão global	Usuário		
			Um1	Um2	Um3
<i>cj3</i>	0,7075	81,68	82,9	77,52	91,41
- <i>io</i>	0,6737	79,65	75,92	77,77	91,97
- <i>ndvi</i>	0,7144	82,09	84,07	77,54	91,59
- <i>cmf</i>	0,7222	82,73	80,68	80,97	91,63
- <i>ts</i>	0,7082	81,8	80,41	79,08	91,49
- <i>iu</i>	0,7136	82,23	78,81	81,29	91,22
- <i>ge</i>	0,614	76,01	72,07	79,97	80,46

Legenda: *cj*=conjunto, *iu*=índice topográfico de umidade, *ts*=índice de capacidade de transporte de sedimentos, *cmf*=corrente de máximo fluxo, *ndvi*=índice de vegetação da diferença normalizada, *io*= *iron oxide*, *ge*=geologia, *um*=unidade de mapeamento.

A estimativa de qualidade, medida pela exatidão global (Tab. 10 e 11), apresentou maiores valores quando da substituição da variável *cmf* no *cj3* e *pfc* no *cj4*. As menores exatidões ocorreram na substituição das variáveis *ge* e *iu* respectivamente nos dois conjuntos. Para o usuário, os dois piores resultados para as (*um1* e *um2*) acompanharam os resultados apresentados no coeficiente kappa e exatidão global, ou seja, na substituição do índice topográfico secundário de umidade *iu*. Na (*um3*), o pior resultado aconteceu na substituição da variável *cmf*.

Ao ser considerada a exatidão das unidades de mapeamento no *cj3*, do ponto de vista do usuário, a (*um1*) apresentou maior exatidão na substituição da variável *ndvi*. Para a (*um2*) o melhor resultado ocorreu na substituição da variável *iu* e para a (*um3*) a substituição da variável *io* proporcionou maior exatidão.

Para o *cj4* (Tab. 11) a maior exatidão do ponto de vista do usuário, ao ser considerada cada uma das três unidades de mapeamento individualmente, ocorreu na omissão de diferentes variáveis, *cmf* na unidade 1, utilizando todas as variáveis do conjunto na unidade 2 e na substituição de *iu* na unidade 3.

Tabela 11. Resultado da classificação por RNA para o (cj4) nas diferentes combinações.

Conjunto/variáveis	Kappa	Exatidão global	Usuário		
			Um1	Um2	Um3
cj4	0,5197	72,44	61,78	88,94	44,35
- cmf	0,3605	58,33	94,61	40,87	41,75
- ts	0,3793	57,31	62,98	41,76	91,87
- iu	0,3338	53,08	52,06	40,05	92,74
- rg	0,3588	58,92	91,02	44,07	42,52
- plc	0,441	62,69	48,99	61,27	92,15
- pfc	0,7998	87,46	93,18	82,73	90,56
- hi	0,6188	75,82	77,42	70,06	89,55
- ge	0,6513	77,47	89,3	68,12	82,64
- de	0,3991	60,81	49,68	61,64	78,99
- ov	0,4141	61,8	89,25	50,84	42,7

Legenda: cj=conjunto, ov=orientação das vertentes, de=declividade, hi=hipsometria, pfc=perfil da curvatura, plc=plano de curvatura, rg=radiação solar, iu=índice topográfico de umidade, ts=índice de capacidade de transporte de sedimentos, cmf=corrente de máximo fluxo, ge=geologia, um=unidade de mapeamento.

CONCLUSÕES

A integração de dados de diferentes origens é pertinente na predição de unidades preliminares de mapeamento de solos, já que a automação do processo de separação de unidades via RNA, visa simular o comportamento humano nesta tarefa.

De acordo com a qualidade dos resultados pode-se afirmar que a RNA permite a integração de diferentes variáveis para o processo de predição de unidades preliminares de mapeamento de solos. Fato que não é possível com classificadores estatísticos convencionais.

A definição da arquitetura da RNA é atividade mais trabalhosa, pelo razão de ser empírica, porém uma vez a rede treinada pode-se variar as variáveis de entrada, e verificar sua importância. A RNA necessita ser treinada, igualmente o intérprete necessita, além do treino, conhecimento e elementos que auxiliem a interpretação da cena. A vantagem da RNA é que a mesma consegue integrar grande quantidade de variáveis e, o analista tem limitações para isso.

Nos diferentes conjuntos testados, as unidades de mapeamento que apresentaram melhores índices de exatidão do ponto de vista do usuário, apresentaram também mais detalhes

espaciais do que o mapa de solos convencional. Isso ocorreu porque a variação ambiental pôde ser melhor quantificada com as variáveis aqui utilizadas. Por outro lado, a disponibilidade de dados mais detalhados permitiu a RNA no processo de treinamento e classificação detectar componentes na paisagem que são difíceis de serem identificados pelo mapeamento convencional devido a subjetividade do intérprete.

Nos conjuntos 1, 2 e 4 o maior coeficiente kappa e exatidão global aconteceram na substituição da variável pfc. No conjunto 3, o kappa e exatidão global foram maiores na substituição de cmf. No entanto, as maiores exatidões do ponto de vista do usuário foram diferentes para cada unidade de mapeamento e, seus resultados dependeram do conjunto de variáveis consideradas e de suas combinações.

A utilização de atributos topográficos, geologia e dados de sensores remotos, integrados em uma RNA podem contribuir com atividades de levantamento de solos, no entanto, cabe destacar que a experiência e o conhecimento técnico sempre serão fundamentais em qualquer trabalho de mapeamento.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BIE, S. W. & BECKETT, P. H. T. Comparison of four independent soil surveys by air-photointerpretation, Paphos area (Cyprus). *Photogrammetria*, Amsterdam, v. 29, p. 189 - 202, 1973.
2. CAMPBELL, J. B. **Introduction to remote sensing**. New York: The Guilford Press, 1987. 551p.
3. CHAGAS, C. S. **Mapeamento digital de solos por correlação ambiental e redes neurais em uma bacia**

- hidrográfica no domínio de mar de morros.** Viçosa, 2006. 223p. (Doutorado -Universidade Federal de Viçosa – UFV).
4. COMEC - COORDENAÇÃO DA REGIÃO METROPOLITANA DE CURITIBA. Cartas Topográficas, folhas: **SG-22-X-D-I-4-SO-B** e **SG-22-X-D-I-4-SE-A**, escala 1:10.000, 1976.
 5. DOBOS, E. MICHELI, E. BAUMGARDNER, M. F. A regional scale soil mapping approach using integrated AVHRR and DEM data. **International Journal of Applied Earth Observation and Geoinformation - JAG**, v.3, p.30-42, 2001.
 6. EMBRAPA. Centro Nacional de Pesquisa em Solos (Rio de Janeiro, RJ). **Sistema Brasileiro de Classificação de Solos**. Brasília: Embrapa produção de informação; Rio de Janeiro: Embrapa Solos. 2 ed. 306p, 2006.
 7. HAYKIN, S. **Redes neurais. Princípios e prática**. Porto Alegre: Bookman, 900 p., 2001.
 8. HENGL, T. **Pedometric mapping: bridging the gaps between conventional and pedometric approaches**. Wageningen University, Enschede, 2003. 233p. PhD thesis - Wageningen University. Disponível em: http://spatial-analyst.net/wiki/index.php?title=Pedometric_mapping:_PhD_thesis.
 9. HILTON, G. E. Connections learning procedures. **Artificial Intelligence**, v. 40, n. 1-3, p. 185-234, 1989.
 10. JENNY, H. **Factors of soil formation; a system of quantitative pedology**. New York: McGraw-Hill, 1941. 281p.
 11. LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n.1, p. 159-174, 1977.
 12. LIMA, M.R. **Atributos de solos e macrófitas aquáticas flutuantes: uma contribuição a sustentabilidade agrícola e ambiental na bacia do rio Iraí (PR)**. – Universidade Federal do Paraná. Curitiba, 2005. 111p. Tese (Doutorado em Agronomia – Produção vegetal) – Setor de Ciências Agrárias – Universidade Federal do Paraná.
 13. MAACK, R. **Geografia física do Paraná**. 2ª edição, José Olympio. Rio de Janeiro, 1981. 450p.
 14. MATWORKS. **User's Guide – Neural Network Toolbox**, MATLAB. version 7.2. Math Works, 2007.
 15. McBRATNEY, A. B.; MENDONÇA SANTOS, M. L.; MINASNY, B. On digital soil mapping. **Geoderma**, v. 117, p. 3-52, 2003.
 16. McCUNE, B. KEON, D. Equations for potential annual direct incident radiation and heat load. **Journal of Vegetation Science**. v. 13, p. 603-606, 2002.
 17. McKENZIE, N.J.; GESSLER, P.E.; RYAN, P.J.; O'CONNEL, D. The role of terrain analysis in soil mapping. In: Wilson, J.P., Gallant, J.C. (Eds.), **Terrain Analysis - Principles and Applications**. Wiley, New York, p. 245-265, 2000.
 18. MENDONÇA-SANTOS, M. L.; SANTOS, H. G. **Mapeamento digital de classes e atributos de solos – métodos, paradigmas e novas técnicas**. Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA, Rio de Janeiro, Documentos 55, 19p. 2003.
 19. MINEROPAR – Minerais do Paraná SA. **Atlas Geológico do Estado do Paraná**. Curitiba: Mineropar, 2001.
 20. MOORE, I.D., GRAYSON, R.B., LADSON, A. R. Digital terrain modeling: A review of Hydrological, geomorphological and biological applications. **Hydrological Processes**, v. 5, p. 3-30, 1991.
 21. MOORE, I.D.; GESSLER, P. E.; NIELSEN, G.A.; PETERSON, G.A. Soil attribute prediction using terrain analysis. **Soil Science Society American Journal**. v. 57, p. 443-452, 1993.
 22. 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.
 23. PENNOCK, D. J.; ZEBARTH, B. J. & DE JONG, E. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. **Geoderma**, v.40, p.297-315, 1987.
 24. REIDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning. The RPROP algorithm. In: INTERNATIONAL CONFERENCE ON NEURAL NETWORK (ICNN), 1993. San Francisco, 1993. p. 586-591.
 25. RIBEIRO, S. R. A. **Integração de imagens landsat TM e dados auxiliares para a delimitação automática de unidades fisionômicas (paisagem) e geológicas usando redes neurais artificiais**. Curitiba, 2003. 169 p. Tese (Doutorado em Ciências Geodésicas) – Setor de Ciências da Terra, Universidade Federal do Paraná.
 26. SABINS, F. F. **Remote Sensing: Principles and Interpretation**. 3rd ed. New York: W. H. Freeman and Company, 1997. 432p.
 27. SALAMUNI, E.; SALAMUNI, R.; EBERT, H.D. Contribuição à geologia da bacia sedimentar de Curitiba (PR). **Boletim Paranaense de Geociências**. n. 47, p. 123-142. 1999.
 28. SANTOS, L. J. C.; OKA-FIORI, C.; CANALI, N. E.; FIORI, A. P.; SILVEIRA, C. T.; SILVA, J. M. F.; ROSS, J. L. S. Mapeamento Geomorfológico do Estado do Paraná. **Revista Brasileira de Geomorfologia**, v.7, n.2, p. 3-11, 2006.
 29. SANTOS, R.D.; LEMOS, R.C.; SANTOS, H.G.; KER, J.C.; ANJOS, L.H.C. **Manual de descrição e coleta de solo no campo**. 5. ed. Viçosa: SBCS, 2005. 92 p.
 30. SOUZA, E. C. B. DE; RIBEIRO, S. R. A.; BOTELHO, M. F.; KRUGER, C. P.; CENTENO, J.A.S. Geração de isolinhas, com dados obtidos por levantamento PS/L₁L₂, mediante técnicas de Redes Neurais Artificiais. **Acta scientiarum technology**. v.28, n.2, 2006.
 31. SILVEIRA, C.T. **análise digital do relevo na predição de unidades preliminares de mapeamento de solos: integração de atributos topográficos em Sistemas de Informações Geográficas e Redes Neurais Artificiais**. Curitiba, 2010. 152p. Tese (Doutorado em Geografia) – Universidade Federal do Paraná, Setor de Ciências da Terra.
 32. SIRTOLI, A. E.; **Mapeamento de solos com auxílio de atributos do terreno, índices espectrais e geologia integrados por redes neurais artificiais**. Curitiba, 2008. 102p. Tese (Doutorado em Geologia) – Universidade Federal do Paraná, Setor de Ciências da Terra.
 33. WANG, Z.; BOVIK, A. C. **A universal image quality index**. IEEE Signal Processing Letters, v. 9, n. 3, p. 81-84, 2002.
 34. WAY, D. S. **Terrain Analysis**. 2. ed. Stroudsburg : Dowden, Hutchinson & Ross, 1978. 438 p.
 35. WILSON, J. P.; GALLANT, J. C. Digital terrain analysis, chapter 1. In: WILSON, J. P.; GALLANT, J. C. (eds.). **Terrain analysis: principles and applications**. New York: John Wiley & Sons, p.1-27, 2000.
 36. WOOD, J. **The geomorphological characterisation of digital elevation models**. Leicester, UK, 1996. 185p. PhD Thesis - University of Leicester. Disponível em: <http://www.soi.city.ac.uk/~jwo/phd>.
 37. ZHU, A.X ; HUDSON, B; BURT, J.; LUBICH, K.; SIMONSON, D. Soil mapping using GIS, expert knowledge, and fuzzy logic. **Soil Science Society of American Journal**. v.65, p.885-894, sep/oct 2001.

*Manuscrito recebido em: 20 de Setembro de 2011
Revisado e Aceito em: 04 de Outubro de 2013*