



# Utilização de redes neurais artificiais para avaliação de produtividade do solo, visando classificação de terras para irrigação<sup>1</sup>



Luciana C. Bucene<sup>1</sup> & Luiz H. A. Rodrigues<sup>2</sup>

<sup>1</sup> FEAGRI/UNICAMP, CP 6011, CEP 13083-970, Campinas, SP. E-mail: [luciana.bucene@agr.unicamp.br](mailto:luciana.bucene@agr.unicamp.br) (Foto)

<sup>2</sup> FEAGRI/UNICAMP, Campinas, SP, Fone (19) 37881062, E-mail: [lique@agr.unicamp.br](mailto:lique@agr.unicamp.br)

Protocolo 145 - 7/10/2002 - Aprovado em 1/3/2004

**Resumo:** Objetivando classificar terras para irrigação, faz-se necessário analisar e determinar alguns parâmetros, entre eles a produtividade do solo. A classificação de produtividade (comumente chamada fertilidade aparente) é delimitada em cinco classes: muito alta, alta, média, baixa e muito baixa, e em cada classe é preciso avaliar certos atributos do solo, como pH, CTC (capacidade de troca de cátions), V% (índice de saturação por bases), P (fósforo), Mg (magnésio) e K (potássio). Neste trabalho, objetivou-se identificar a produtividade na qual atributos do solo, da parte inicial da microbacia hidrográfica do Rio Pardo, localizada em Pardinho, SP, foram analisados e classificados nas classes que a delimitam, através de Redes Neurais Artificiais (RNAs) utilizando-se Perceptron Múltiplas Camadas (Multilayers Perceptrons - MLP) com o algoritmo de treinamento "backpropagation" – classificador de padrões, obtendo-se um número ótimo de camadas intermediárias e de neurônios; resultando na classificação de produtividade, a situação ótima da rede obteve 78% dos resultados iguais aos desejados, com duas camadas de neurônios, uma das quais intermediária, com 5 neurônios, e uma camada de saída.

**Palavras-chave:** inteligência artificial, reconhecimento de padrões, perceptron múltipla camada

## Use of artificial neural networks for evaluation of apparent fertility and classification of land for irrigation

**Abstract:** Productivity data (commonly known as apparent fertility) of the initial part of the river Pardo-SP watershed was analyzed and classified with Artificial Neural Networks (ANNs), in order to classify lands for irrigation. Soil attributes as pH, CEC (cation exchange capacity), V% (base saturation index), P (phosphorus), Mg (magnesium) and K (potassium) were defined in five classes: very high, high, medium, low and very low. Apparent fertility classification taking into account the five classes was performed by using Multiple Layers Perceptron (MLP). Backpropagation algorithm was performed with the training set. One hidden layer with 5 neurons was the situation that best performed.

**Keywords:** artificial intelligence; pattern recognition; multilayer perceptron

## INTRODUÇÃO

Trabalhos de pesquisa em microbacias hidrográficas têm sido desenvolvidos desde o início do século. Estudos comparativos são realizados com a finalidade principal de se conhecer as diferenciações existentes na terra, os quais levantam todos os problemas da bacia, analisam os conflitos, classificam parâmetros e indicam soluções em todos os níveis, integrando conclusões e recomendações ao meio ambiente (Rocha, 1997).

Técnicos do "Bureau of Reclamation", do Departamento de Interior dos Estados Unidos elaboraram, para o Brasil, em

1993, uma série de manuais sobre irrigação, entre eles "Classificação de Terras para Irrigação", tendo como base o Boletim Americano de 1982 (Carter, 1993). Foi declarado um dos sistemas de classificação de terras para irrigação mais difundido no mundo, definido pelos atributos físicos e químicos da terra. A classificação de terra é feita primordialmente para se obter informações necessárias à delimitação das terras em áreas de classes aptas para irrigação, eliminando as áreas inaptas nas condições econômicas prevaletentes (Carter, 1993). Essa classificação exige análise e avaliação das características de topografia, drenagem e solo, entre as quais a análise de fertilidade aparente do solo. Inicialmente, a seleção de terras

para irrigação deve identificar as áreas de terras que têm produtividade adequada para garantir segurança daquela área para o desenvolvimento da irrigação e, ainda, compreende uma análise criteriosa dos fatores químicos do solo (Hagan et al., 1967).

No delineamento de classes de uso da terra, a abordagem a partir de Redes Neurais Artificiais (RNAs) tem apresentado expressivos resultados, quando comparados com os métodos tradicionais de classificação (Venturieri & Santos, 1998). Kim et al. (1995), ressaltaram a robustez do sistema de inteligência artificial, que oferece grande potencial na discriminação das classes de cobertura da terra, em comparação com os sistemas de classificação convencionais.

RNAs são modelos computacionais inspirados no sistema nervoso biológico, cujo funcionamento é semelhante a alguns procedimentos humanos, ou seja, aprendem pela experiência, generalizam exemplos através de outros e abstraem características (Wasserman, 1989). A capacidade de aprender através de exemplos e de generalizar a informação aprendida, é o atrativo principal da solução de problemas através de RNAs. A generalização, que está associada à capacidade da rede aprender através de um conjunto reduzido de exemplos e, posteriormente, dar respostas coerentes para dados não conhecidos, é uma demonstração de que a capacidade das RNAs vai muito além do que simplesmente mapear relações de entrada e saída. As RNAs são capazes de extrair informações não apresentadas de forma explícitas através de exemplos.

Segundo Haykin (2001) a utilização de um RNA na solução de uma tarefa passa, antes de tudo, por uma fase de aprendizagem, quando a rede extrai informações relevantes de padrões de informação apresentados para ela, criando assim uma representação própria para o problema. Ainda segundo o mesmo autor, a etapa de aprendizagem consiste em um processo iterativo de ajuste de parâmetros da rede, os pesos das conexões entre as unidades de processamento, que guardam, ao final do processo, o conhecimento que a rede adquiriu do ambiente em que está operando.

As redes são organizadas de acordo com o número de camadas que compõem a sua estrutura tipológica. Assim sendo, existem as redes neurais de camadas simples, constituídas por um grupo de neurônios arranjados em apenas uma camada, e as redes multicamadas, formadas pela combinação de várias redes de camada simples (Venturieri & Santos, 1998).

Um aspecto a considerar é que a Rede Neural pode aceitar diferentes dados de entrada. Assim, dados coletados em campo, tais como condições e medidas topográficas, parâmetros edáficos, valores de fitomassa, estágio de desenvolvimento

de culturas agrícolas etc., também podem ser utilizados como fonte de definição e de contexto de determinado alvo, importante na tarefa de memorização de característica temática. Com isto, aumenta-se a precisão da classificação de classes agrícolas e florestais (Venturieri & Santos, 1998).

Neste sentido e com o objetivo de classificar a fertilidade aparente do solo, utilizando-se as RNAs, avaliaram-se alguns dados analisando-se os atributos do solo, como pH, CTC, V%, P, Mg e K, permitindo a diferenciação das terras para fins de irrigação.

## MATERIAL E MÉTODOS

O presente estudo baseia-se em dados químicos da parte inicial da microbacia do Rio Pardo, localizada no município de Pardinho, SP, tendo como coordenadas 23°05' de latitude Sul e 48°14' de longitude WGr (IGC, 1978 e IBGE, 1982) possuindo o divisor de águas entre as bacias dos Rios Tietê e Paranapanema, fazendo parte da bacia deste último.

Os parâmetros químicos do solo, como pH, CTC, V%, P, Mg e K, foram avaliados para classificar a fertilidade aparente. Até cinco níveis de fertilidade podem ser distinguidos e classificados em: muito alta, alta, média, baixa e muito baixa.

A Tabela 1 mostra como as classes foram definidas de acordo com os parâmetros avaliados, tendo-se como resultado os diferentes níveis da fertilidade aparente.

Aplicou-se o Perceptron Múltipla Camada (MLP) através do seu treinamento de forma supervisionada com o algoritmo de retropropagação de erro ("backpropagation"). O algoritmo "backpropagation" é baseado na regra delta proposta por Widrow & Hoff, razão por que é também chamado de regra delta generalizada e consiste na apresentação de um conjunto de amostras, o qual é comparado com um resultado de saída desejado (Sclunzen et al., 1993, citado por Venturieri & Santos, 1998). Desta forma, o algoritmo necessita de um conjunto de amostras com pares de entradas e saídas, ou padrões corretos, a partir do qual o algoritmo constrói uma configuração capaz de realizar o mapeamento desejado (Liporace, 1994, citado por Venturieri & Santos, 1998). Este algoritmo se baseia na regra de aprendizagem por correção de erro. Basicamente, a aprendizagem por retropropagação de erro consiste de dois passos através das diferentes camadas da rede: um passo para frente, a propagação, e um passo para trás, a retropropagação. No passo para frente, um padrão de atividade (vetor de entrada) é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada. Finalmente, um conjunto de saídas é produzido como a resposta real da rede. Durante o

Tabela 1. Padrões de classificação de fertilidade aparente

	Classe de Fertilidade Aparente				
	Muito alta	Alta	Média	Baixa	Muito baixa
pH (CaCl <sub>2</sub> )	>6,0	5,6 – 6,0	5,1 – 5,5	4,4 – 5,0	≤ 4,3
CTC (cmol <sub>c</sub> dm <sup>-3</sup> )		>10,0	5,0 – 10,0	<5,0	
V (%)	>90	71 – 90	51 – 70	26 – 50	≤ 25
P (mg dm <sup>-3</sup> )	>80	41 – 80	16 – 40	7 – 15	≤ 6
Mg (cmol <sub>c</sub> dm <sup>-3</sup> )		>0,8	0,5 – 0,8	<0,4	
K (cmol <sub>c</sub> dm <sup>-3</sup> )	>0,60	0,31 – 0,60	0,16 – 0,30	0,08 – 0,15	≤ 0,07

passo de propagação, os pesos sinópticos da rede são todos fixos. Durante o passo para trás, por outro lado, os pesos sinópticos são todos ajustados de acordo com uma regra de correção do erro. Especificamente, a resposta real da rede é subtraída de uma resposta desejada (alvo) para produzir um sinal de erro. Este sinal de erro é então propagado para trás através da rede, contra a direção das conexões sinópticas. Os pesos sinópticos são ajustados para fazer com que a resposta real da rede se mova para mais perto da resposta desejada. Genericamente, na aprendizagem por retropropagação começa-se tipicamente com uma amostra de treinamento e se usa o algoritmo de retropropagação para calcular os pesos sinópticos de um MLP carregando (codificando) tantos exemplos de treinamento quanto possível para dentro da rede (Haykin 2001). Espera-se que a rede neural assim projetada seja capaz de generalizar. Diz-se que uma rede está generalizada quando o mapeamento de entrada-saída computado pela rede for correto (ou aproximadamente correto) para dados de teste não-utilizados durante a criação ou treinamento da rede. Aqui se assume que os dados de teste foram retirados da mesma população usada para gerar os dados de treinamento.

As definições das classes estão apresentadas na Tabela 2 e os padrões de treinamento na Tabela 3. Executou-se o treinamento da rede utilizando-se o algoritmo supervisionado regra delta generalizada, para a classificação de 5 classes.

Tabela 2. Definições das classes

Classes	Saída (d)
Solo com fertilidade aparente muito alta	111
Solo com fertilidade aparente alta	101
Solo com fertilidade aparente média	11
Solo com fertilidade aparente baixa	1
Solo com fertilidade aparente muito baixa	0

Os algoritmos de treinamento foram treinados com uma e duas camadas intermediárias. Todas as situações, num total de sete situações diferentes, como mostra a Tabela 4, tiveram 6 entradas (pH, CTC, K, P, Mg e V%) e três saídas (Y1, Y2 e Y3), cuja combinação corresponde a um nível de fertilidade aparente: muito alta, alta, média, baixa e muito baixa. Do total dos dados, 90% foram separados para o treinamento da rede neural, definidos como padrões de treinamento, escolhidos aleatoriamente, a fim de que a rede “aprenda” as regras e não “decore” exemplos. O restante dos dados, 10%, definidos como padrões de teste, somente foi apresentado à rede na fase de teste a fim de que ela possa “deduzir” corretamente o inter-relacionamento entre os dados. Todos os testes foram

Tabela 3. Padrões de treinamento\*

x1	x2	x3	x4	x5	x6	Y1	Y2	Y3
>6,0	>12,0	>90	>80	> 1,1	>0,60	1	1	1
5,6 – 6,0	9,1 – 12,0	71 - 90	41 - 80	0,9 – 1,1	0,31– 0,60	1	0	1
5,1 – 5,5	6,1 – 9,0	51 - 70	16 - 40	0,5 – 0,8	0,16-0,30	0	1	1
4,4 – 5,0	4,0 – 6,0	26 - 50	7 - 15	0,4 – 0,1	0,08– 0,15	0	0	1
≤4,3	<4,0	≤ 25	≤ 6	<0,1	≤ 0,07	0	0	0

\* x1 = pH; x2 = CTC; x3 = V%; x4 = P; x5 = Mg e x6 = K; Y1, Y2 e Y3 = Saídas; Se Y1 = 1, Y2 = 1, Y3 = 1, significa muito alta; Se Y1 = 1, Y2 = 0, Y3 = 1, significa alta; Y1 = 0, Y2 = 1, Y3 = 1, significa média; Y1 = 0, Y2 = 0, Y3 = 1, significa baixa; Y1 = 0, Y2 = 0, Y3 = 0, significa muito baixa

realizados no software Matlab (Demuth & Beale, 1998). A função de ativação foi do tipo tangente hiperbólica, definida por  $f(x) = \tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$ . O número máximo de épocas de treinamento, quantidade de vezes que apresentou o conjunto de treinamento à rede, de tal forma que ela alcançasse os valores desejados, foi de 1500. A precisão de treinamento foi de  $0.5 \cdot 10^{-6}$ .

## RESULTADOS E DISCUSSÃO

Todos os resultados mencionados abaixo foram realizados a partir dos 10% dos dados destinados à fase de teste, e estão apresentados na Tabela 4.

Na primeira situação, testando-se o algoritmo com apenas uma camada intermediária e com 5 neurônios, alcançou-se um resultado com 7 respostas iguais às desejadas. Numa segunda situação, com 10 neurônios na única camada intermediária, obtiveram-se 4 respostas iguais às respostas desejadas. Em um terceiro momento e se testando com 15 neurônios na camada intermediária, o resultado encontrado foi o mesmo da primeira situação, enquanto na quarta situação, quando testado com 20 neurônios na camada intermediária, o resultado foi também de 7 respostas iguais às desejadas.

Tabela 4. Resultados do treinamento, utilizando-se os padrões de teste

	Nº de camadas intermediárias							
	1				2			
Nº de neurônios	5	10	15	20	1ª camada	5	5	10
					2ª camada	5	10	10
Nº de acertos	7	4	7	7*		5	6	6*

\* A rede não convergiu, chegando à precisão de treinamento ( $0.5 \cdot 10^{-6}$ ) antes de se obter os resultados esperados

Analisando-se as situações acima mencionadas e se comparando os resultados obtidos, as redes com 5 e 15 neurônios apresentaram maior número de acertos que quando testadas com 10 neurônios; trabalhando com 10 neurônios, o resultado da rede discordou mais dos resultados desejados, tendo apenas 4 respostas iguais às respostas desejadas, sendo este o teste que mais se afastou das respostas desejadas; quando testada a rede com 20 neurônios, a rede não convergiu, não se terminando de testa-la, chegando à precisão de treinamento antes de se obter os resultados esperados.

A partir da quinta situação, testou-se o algoritmo com duas camadas intermediárias. Com 5 neurônios na primeira camada intermediária e 5 neurônios na segunda, o resultado foi de 5

respostas iguais às respostas desejadas; numa próxima situação, testando-se com 5 neurônios na primeira camada intermediária e 10 neurônios na segunda, o resultado foi de 6 respostas semelhantes às respostas desejadas; por fim, na última situação, com 10 neurônios na primeira e na segunda camada intermediária, o resultado foi também de 6 respostas iguais às respostas desejadas.

Analisando-se agora as situações com duas camadas intermediárias, os resultados encontrados não foram satisfatórios, apresentando números de acertos menores do que quando testada a rede com apenas uma camada intermediária; quando, porém, testada com 10 neurônios na primeira e na segunda camadas, observou-se novamente que a rede não convergiu, como aconteceu quando testada com 20 neurônios com uma camada intermediária.

Para treinar a rede, foram utilizados dados que representavam todas as classes existentes, porém numa proporção bem desigual, apresentando poucas situações caracterizando as classes muito altas, altas e médias, sendo que a maioria dos dados classificava em baixa e muito baixa, podendo ter influenciado no aprendizado da rede e contribuído para que o número de acertos não fosse maior que o apresentado. Observou-se, também, que as respostas encontradas que discordavam das respostas desejadas eram as mesmas em todas as situações, tanto com uma ou duas camadas intermediárias.

## CONCLUSÕES

1. O melhor resultado dentre as redes testadas, foi com 5 e 15 neurônios em apenas uma camada intermediária, obtendo-se 78% dos resultados iguais aos desejados, sendo que a indicada para trabalhar neste caso é a rede com 5 neurônios, pois já satisfaz com um número menor de neurônios.

2. As redes neurais artificiais apresentaram condições suficientes ao desenvolvimento do trabalho, sendo eficiente na classificação de fertilidade aparente para diferenciação de terras para fins de irrigação, da parte inicial da microbacia do Rio Pardo, o que permite o uso para casos semelhantes.

3. Sugerem investigações através dos quais se busque identificar parâmetros adicionais, não considerados no presente trabalho, mas que possam ter influenciado no treinamento.

## LITERATURA CITADA

- Carter, V.H. Classificação de terras para irrigação. Brasília, Ministério da Integração Regional/Secretaria de Irrigação, 1993. 203p.
- Demuth, H; Beale, M. Neural network toolbox. For use with MATLAB. Computation, visualization, programming. User's guide. Version 3. The Math Works. 1998. Available at <http://www.mathworks.com/access/helpdesk/help/toolbox/nnet/>
- Hagan, R.M.; Haise, H.R; Edminister, T.W. Irrigation of agricultural lands. Colorado; American Society Agronomy, 1967. p.27-173,
- Haykin, S. Redes neurais, princípios e prática. 2 edição. Porto Alegre: Bookman, 2001. 900p.
- IBGE - Instituto Brasileiro de Geografia e Estatística, Carta do Brasil - Quadricula de Pardino. Rio de Janeiro, Escala 1:50.000, 1982.
- IGC. Carta séries SF-22-Z-B-VI-3-SO-F a SF-22-Z-D-III-1-NE-E. Rio de Janeiro, Escala 1:10.000, 1978.
- Kim, K.; Yang, Y.; Lee, J.; Choi, K.; Kim, T. Classification of multispectral image using neural network. In: Procudings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS) '95, IEEE catalog #95CH35770, vol. 2. Firenze, Italy, 1995, p.446-448.
- Rocha, J.S.M. Manual de projetos ambientais. Santa Maria: Imprensa Universitária, 1997. 423p.
- Venturieri, A.; Santos, J.R. Técnicas de classificação de imagens para análise de cobertura vegetal. In: Assad, E. D; Sano, E.E. Sistema de informações geográficas. Aplicações na agricultura. 2 ed., Brasília: Embrapa – SPI/ Embrapa – CPAC, 1998. p.351-371.
- Wasserman, P.D. Neural computing: theory and practice. New York: van Nostrand Reinhold. 1989.