



## Modeling of soil penetration resistance using statistical analyses and artificial neural networks

Fábio Lúcio Santos<sup>1\*</sup>, Valquíria Aparecida Mendes de Jesus<sup>2</sup> and Domingos Sárvio Magalhães Valente<sup>1</sup>

<sup>1</sup>Departamento de Engenharia Agrícola, Universidade Federal de Viçosa, Av. P. H. Rolfs, s/n, 36570-000, Viçosa, Minas Gerais, Brazil.

<sup>2</sup>Programa de Pós-graduação em Fitotecnia, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. \*Author for correspondence. E-mail: fabio.ls@ufv.br

**ABSTRACT.** An important factor for the evaluation of an agricultural system's sustainability is the monitoring of soil quality via its physical attributes. The physical attributes of soil, such as soil penetration resistance, can be used to monitor and evaluate the soil's quality. Artificial Neural Networks (ANN) have been employed to solve many problems in agriculture, and the use of this technique can be considered an alternative approach for predicting the penetration resistance produced by the soil's basic properties, such as bulk density and water content. The aim of this work is to perform an analysis of the soil penetration resistance behavior measured from the cone index under different levels of bulk density and water content using statistical analyses, specifically regression analysis and ANN modeling. Both techniques show that soil penetration resistance is associated with soil bulk density and water content. The regression analysis presented a determination coefficient of 0.92 and an RMSE of 0.951, and the ANN modeling presented a determination coefficient of 0.98 and an RMSE of 0.084. The results show that the ANN modeling presented better results than the mathematical model obtained from regression analysis.

**Keywords:** modeling, soil physical properties, neural networks.

### Modelagem da resistência à penetração do solo usando análises estatísticas e redes neurais artificiais

**RESUMO.** Um importante fator para a avaliação da sustentabilidade de sistemas agrícolas é o monitoramento da qualidade do solo por meio de seus atributos físicos. Logo, atributos físicos do solo, como resistência à penetração, podem ser empregados no monitoramento e na avaliação da qualidade do solo. Redes Neurais Artificiais (RNA) tem sido empregadas na solução de vários problemas na agricultura, neste contexto, o uso desta técnica pode ser considerada uma abordagem alternativa para se prever a resistência à penetração do solo a partir de suas propriedades básicas como densidade e teor de água. Portanto, o objetivo desse trabalho foi desenvolver um estudo do comportamento da resistência à penetração do solo, medida a partir do índice de cone, empregando análise de regressão e modelagem por RNA. Ambas as técnicas mostraram que a resistência à penetração do solo está associada com a densidade e o teor de água do solo. A análise de regressão apresentou coeficiente de regressão de 0,92 e REMQ igual a 0,951 enquanto a modelagem por RNA apresentou coeficiente de determinação de 0,98 e REMQ igual a 0.084. Os resultados indicaram que a modelagem por RNA apresentou melhores resultados do que o modelo matemático obtido a partir da análise de regressão.

**Palavras-chave:** modelagem, propriedades físicas do solo, redes neurais.

### Introduction

An important factor for the evaluation of an agricultural system's sustainability is the monitoring of soil quality via its physical attributes. The monitoring of these attributes can result in better quality agricultural products, the promotion of more efficient mechanization processes and the establishment of the reasonable use of raw materials and natural resources (BEUTLER et al., 2001). Penetration resistance is a physical attribute of soil

that can be used to monitor and evaluate soil quality (ISLAM; WEIL, 2000).

Penetration resistance influences the growth of roots, and it can be used as a parameter for evaluating the effects of tillage systems on the roots' environment, the detection of compacted layers, the prediction of the traction force needed to perform mechanized processes and the prevention of the appearance of a physical barrier that can reduce the development of the plants

(CAMPANHARO et al., 2009; CUNHA et al., 2002).

The determination of the soil penetration resistance is performed a device called penetrometer, which allows the soil resistance to be measured quickly (TAVARES FILHO; RIBON, 2008). According to Dexter et al. (2007), the resistance to penetration is governed by fundamental properties of the soil, such as shear strength, compressibility and the friction force from the soil-metal interaction during the trial using the penetrometer. Hence, soil penetration resistance can be estimated as a quantity called cone index. This quantity can be expressed as the ratio of force per unit area of the base of the cone at a determined depth (CAMPANHARO et al., 2009; CUNHA et al., 2002).

Studies have been carried out to evaluate the influence of water content on the behavior of soil penetration resistance (CUNHA et al., 2002). Mathematical models have also been developed to predict the penetration resistance from basic soil properties, such as soil composition, bulk density and water content (CUNHA et al., 2002; DEXTER et al., 2007; SINGH; KAY, 1997). In this context, the use of Artificial Neural Networks (ANN) can be considered an alternative approach for predicting soil penetration resistance from soil bulk density and water content.

ANN have been employed to solve many problems in agriculture (ERZIN et al., 2008, 2010; KIM; GILLEY, 2008). Varella et al. (2002) used ANN for the determination of land cover from digital images. Khazaei and Daneshmandi (2007) used ANN to model the drying kinetics of sesame seeds. They concluded that the ANN technique presented better results than traditional mathematical modeling. Sarmadian et al. (2009) used ANN to model soil properties, and the results were better than the multivariate regression analysis, showing the effectiveness of the ANN technique.

The objective of this work is to determine the effect of the soil bulk density and water content on soil penetration resistance behavior measured from the cone index, using statistical analyses, specifically regression analysis, and ANN modeling.

## Material and methods

### Experimental procedure

The study was conducted at Cidade Gaucha, which is located in northwest Paraná State, Brazil. Samples were collected at a location where the predominant soil is classified as Rhodic Acrustox (EMBRAPA, 1999). The samples were collected at a depth of 0.10 m using steel cylinders in three areas

that had different levels of management and, therefore, presented different levels of soil compaction. Thus, 12 samples were collected at each point, totaling 36 samples.

First, preliminary tests were conducted in an oven for the samples related to each area to establish the average bulk density of the soil, the moisture saturation and the water loss gradient. Three replicates were used for the preliminary tests for each sampled area, totaling nine samples. The samples were saturated for 48 hours and then were placed in an oven at 105°C. The samples were removed from the oven at intervals of 30 minutes to check the weight loss.

A penetrometer was used to determine the soil penetration resistance. The penetrometer had a 4 mm diameter rod and a load cell of 200 kgf. Readings were taken at intervals of 1 second during the probe penetration into the soil sample. The penetration resistance was obtained from the average of the points obtained during the test. At the end of the tests, the soil densities of the samples were determined.

From the experimental procedure described, a dataset was obtained considering three replications for three average soil bulk densities (1.75, 1.90 and 2.05 kg dm<sup>-3</sup>) and three average water content levels (0.04, 0.08 and 0.12 kg kg<sup>-1</sup>).

### Statistical analyses

From the experimental results, a model was chosen considering the soil penetration resistance as a function of the soil bulk density and water content. A completely randomized factorial design with three densities was chosen, in which the evaluated factors were composed of three levels of bulk density and three levels of water content. The soil penetration resistance data were submitted to an analysis of variance at a 5% significance level. The effects of the soil bulk density and water content were studied by regression analysis.

All statistical analyses were performed using the SAS program, version 8.0. The model was chosen based on the coefficient of determination, the significance of regression coefficients and the lack of adjustment of the model.

### Artificial neural network modeling

According to Haykin (1999), Artificial Neural Networks (ANN) are massively parallel networks, are self-adaptive and are interconnected by basic structures called neurons. Neurons are processing units with limited learning capacity; however, their interactions allow the ANN to learn from a determined set of input data and their output

patterns. Figure 1 illustrates an ANN architecture, which is composed of an input layer, a processing layer (also known as a hidden layer) and an output layer. This type of architecture is called a "Multilayer Perceptron Network".

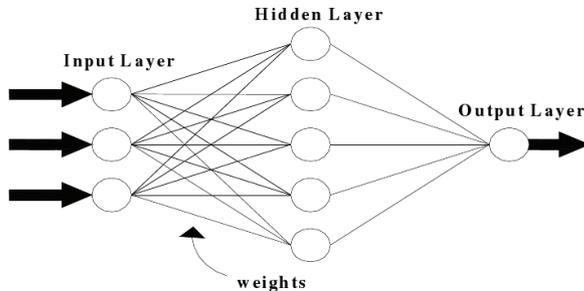


Figure 1. Multilayer Perceptron Network.

In this work, a model was developed based on the ANN technique to predict the soil penetration resistance using the soil bulk density and water content as the input data. The ANN modeling was composed of two stages: training and validation. In the training stage, architectures were considered that consisted of 2-n1-n2-1, where there were 2 elements in input vector and n1 and n2 represented the number of neurons in each hidden layer, with just one neuron in the output layer. Several configurations were tested in the ANN hidden layer, where the number of neurons in first hidden layer (n1) ranged from 1 to 15 and the number of neurons in second layer ranged (n2) from 0 to 15, totaling 240 architectures analyzed. Out of these architectures, 15 were composed of one hidden layer ANN when the number of neurons in layer n1 ranged from 1 to 15 and the number of neurons was equal to zero in layer n2.

An error backpropagation algorithm was used in the training stage. The data set was divided into a training set and a validation set. The training set consisted of 20 input and output patterns, where the input vector was composed of values of the soil bulk density and water content and the output consisted of the soil penetration resistance. To improve the ANN generalization capability, the output data were normalized, which allowed output values ranging from 0 to 1, according to Equation 1.

$$PR_N(y) = 1 - \left( \frac{PR(y)}{PR_{max}} \right) \quad (1)$$

where:

$PR_N(y)$  = normalized penetration resistance;  
 $PR(y)$  = penetration resistance to be normalized;  
 $PR_{max}$  = maximum value of the soil penetration resistance.

In the first step of the training stage, the ANN architectures with the best performance were determined during the training process. Thus, only architectures that reached a root mean square error (RMSE) of 0.001 were selected. However, to avoid over training, ANN models with minimal dimensions were selected. In the second step of the training stage, a study was developed to determine ANN parameters such as learning rate and momentum. The networks were trained so that these parameters could be determined properly. In this step, the RMSE and the number of training epochs were considered for the selection of the ANN architectures.

Once a given ANN was trained using the training data set, its performance must be evaluated using a validation set of data. The validation stage is essential to avoid ANN over-training. Thus, the performance of the ANNs selected were tested and compared using the determination coefficient ( $R^2$ ) and the RMSE. The final ANN selection considered the lowest errors presented in the training and validation stages.

All programs used in the ANN training and validation stages were developed in the C programming language using a gcc-gnu compiler and the FANN library (Fast Artificial Neural Network Library) for the Linux operating system UBUNTU.

## Results and discussion

### Statistical Analyses

Table 1 presents the results of the analyses of variance obtained from the experimental data of the soil penetration resistance determined from the cone index while considering the different levels of soil bulk density and water content. The interaction between the soil bulk density and water content factors was significant at a 5% probability.

Table 1. Analyses of variance for soil mechanical penetration resistance considering different levels of density and water content.

SV	DF	SS	MS	F	P-value
Bulk Density ( <i>D</i> )	2	306511347.90	153255673.90	--	--
Water Content ( <i>WC</i> )	2	67906589.20	33953294.60	--	--
<i>D</i> x <i>WC</i>	4	9337733.30	2334433.30	692.45*	< 0.001
Residue	18	60682.70	3371.30		
Total	26	383816353.10			

\*significant at 5% probability.

Equation 2 represents the selected model from the regression analyses. The model was chosen considering the determination coefficient ( $R^2$ ), the significance of regression coefficients and the lack of adjustment of the model. It can be observed that the

model presented a determination coefficient of 0.92. The selected model also presented an RMSE of 0.951.

$$(PR = -41821.28 + 26701.59D - 40100.79WC \ R^2 = 0.92) \quad (2)$$

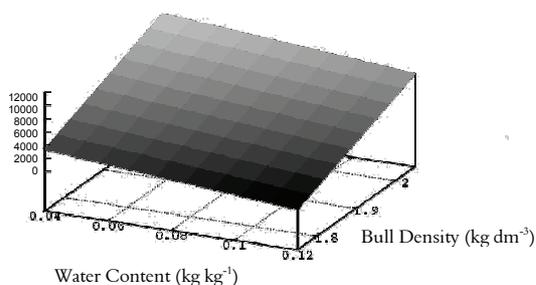
where:

$PR$  = soil penetration resistance, kPa;

$D$  = soil bulk density,  $\text{kg dm}^{-3}$ ;

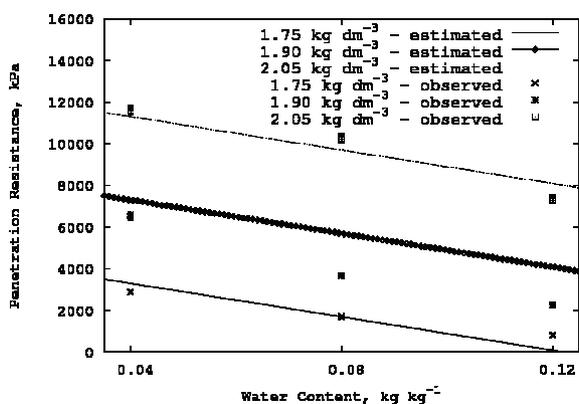
$WC$  = soil water content,  $\text{kg kg}^{-1}$ .

Figure 2 presents the surface response relating the soil penetration resistance to the soil bulk density and water content. The model analysis shows that the highest penetration resistance tended to occur at a higher soil density and lower water content, as reported in the literature (CUNHA et al., 2002; DEXTER et al. 2007).



**Figure 2.** Surface response for penetration resistance ( $PR$ ) related to the density ( $D$ ) and water content ( $WC$ ) of the soil.

To improve the visualization of these factors' influence on soil penetration resistance, the response surface is presented in Figure 3 for the different levels of soil bulk density and water content.

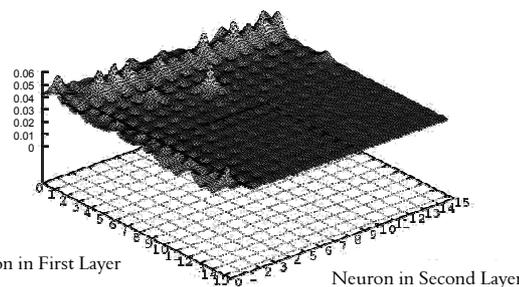


**Figure 3.** Response surface for the soil penetration resistance obtained from the different levels of soil bulk density and water content.

Figure 3 shows that the soil penetration resistance tended to decrease with higher values of soil water content, which can be explained by the reduction of the cohesion forces and the internal friction (KLEIN et al., 1998). Moreover, the soil penetration resistance tended to increase with higher values of soil bulk density. This effect can be explained by the reduction of the soil pore spaces, which resulted in an increase of the penetration resistance (CUNHA et al. 2002).

**Artificial neural network modeling**

In this study, an ANN model was employed to predict the soil penetration resistance from the soil bulk density and water content as input data. Figure 4 verifies the result of the study performed during the training stage, in which ANN architectures were evaluated. The study considered one and two hidden layer ANNs.



**Figure 4.** Artificial Neural Network learning capacity considering the number of neurons in the first hidden layers ( $n_1$ ) and the second hidden layer ( $n_2$ ).

Figure 4 shows that the learning capacity of the two hidden layers was significantly higher than one hidden layer. This feature indicates that when increasing the number of hidden layers, the ANN learning capacity increases. However, the number of neurons in each hidden layer can vary according to the complexity of the problem (ERZIN et al., 2010; KHAZAEI; DANESHMANDI, 2007; KIM; GILLEY, 2008).

Considering that several evaluated architectures reached the RMSE established during the training stage, for the validation stage, only the ANN architectures that presented minimal dimensions were selected. Among these, the ANN architecture composed by the 2-2-2-1 configuration presented the best results. The ANN results are presented in Table 2.

**Table 2.** Parameters and results of ANN architecture 2-2-2-1 selected after the training and validation stages.

ANN parameters			Transfer function	Validation stage results			
architecture	learning rate	momentum		training RMSE	validation RMSE	R <sup>2</sup>	global mean error (%)
2-2-2-1	0.4	0.5	Sigmoid	0.0320	0.0840	0.98	6.74

Table 2 shows that, during the training and validation, ANN architecture 2-2-2-1 presented a RMSE equal to 0.032 and 0.084, respectively. Moreover, it can be observed that there is a significant difference between the results obtained from the ANN modeling (RMSE equal to 0.084 and  $R^2$  equal to 0.98) and the results obtained from the mathematical model performed by regression analysis (RMSE equal to 0.951 and  $R^2$  equal to 0.92).

Figure 5 presents the results for the ANN validation stage (architecture 2-2-2-1) using a validation set composed of 9 patterns. In general, the global mean error between the ANN estimated values and the observed values was below 6.74%, which confirms the great prediction capability of the ANN for the proposed problem. Additionally, there are several works in the literature demonstrating ANN' capabilities when applied to systems modeling (AHAMAD et al., 2007; GUNAYDIN et al., 2010; KHAZAEI; DANESHMANDI, 2007; SANTOS et al., 2009; SARMADIAN et al., 2009; TURK et al., 2001).

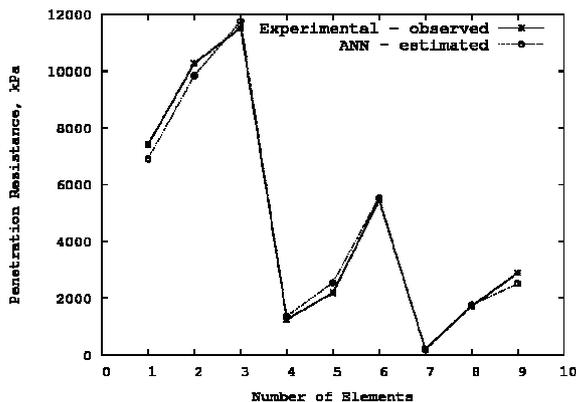


Figure 5. Validation stage results for 2-2-2-1 ANN architecture.

## Conclusion

Penetration resistance is associated with the soil bulk density and water content. The highest penetration resistance values tended to occur at a higher density and lower water content, whereas the lowest penetration resistance values tended to occur at lower soil bulk density and higher water content.

The ANN trained by the backpropagation algorithm was able to learn the correlation between the penetration resistance with the soil bulk density and the water content. ANN modeling can be used to predict the soil penetration resistance from soil bulk density and water content as the input data.

ANN architecture 2-2-2-1 presented an RMSE less than 0.085, an  $R^2$  equal to 0.98 and a global mean error of approximately 6.75%, whereas the

model obtained from statistical analyses presented an RMSE of 0.951 and an  $R^2$  of 0.92. These results show that the ANN model presented better results than the statistical model obtained from regression analysis.

## References

- AHAMAD, I.; NAGGAR, H. E.; KHAN, A. N. Artificial neural network application to estimate kinematic soil pile interaction response parameters. **Soil Dynamics and Earthquake Engineering**, v. 27, n. 9, p. 892-905, 2007.
- BEUTLER, A. N.; SILVA, M. L. N.; CURI, N.; FERREIRA, M. M.; CRUZ, J. C.; PEREIRA FILHO, I. A. Resistência à penetração e permeabilidade de latossolo vermelho distrófico típico sob sistemas de manejo na região dos cerrados. **Revista Brasileira de Ciência do Solo**, v. 25, n. 1, p. 167-177, 2001.
- CAMPANHARO, W. A.; SPERANDIO, H. V.; CECÍLIO, R. A.; HOLLANDA, M. P.; GUARIZ, H. R. Variabilidade espacial da resistência a penetração do solo a penetração em áreas sob cultivos puros e consorciados de café e eucalipto. **Revista Brasileira de Agroecologia**, v. 4, n. 2, p. 2721-2724, 2009.
- CUNHA, J. P. A. R.; VIEIRA, L. B.; MAGALHÃES, A. C. Resistência mecânica do solo à penetração sob diferentes densidades e teores de água. **Engenharia na Agricultura**, v. 10, n. 1, p. 1-7, 2002.
- DEXTER, A. R.; CZYŻ, E. A.; GATE, O. P. A method for prediction of soil penetration resistance. **Soil and Tillage Research**, v. 1, n. 93, p. 412-419, 2007.
- EMBRAPA-Empresa Brasileira de Pesquisa Agropecuária. **Sistema brasileiro de classificação de solos**. Rio de Janeiro: Centro Nacional de Pesquisa de Solos, 1999.
- ERZIN, Y.; RAO, B. H.; SINGH, D. N. Artificial neural network models for predicting soil resistivity. **International Journal of Thermal Sciences**, v. 47, n. 10, p. 1347-1358, 2008.
- ERZIN, Y.; RAO, B. H.; PATEL, A.; GUMASTE, S. D.; SINGH, D. N. Artificial neural network models for predicting electrical resistivity of soils from thermal resistivity. **International Journal of Thermal Sciences**, v. 49, n. 1, p. 118-130, 2010.
- GUNAYDIN, O.; GOKOGLU, A.; FENER, M. Prediction of artificial soil's unconfined compression strength test using statistical analyses and artificial neural networks. **Advances in Engineering Software**, v. 41, n. 9, p. 1115-1123, 2010.
- HAYKIN, S. **Neural networks: a comprehensive foundation**. New Jersey: Prentice Hall, 1999.
- ISLAM, K. R.; WEIL, R. R. Soil quality indicator properties in mid-Atlantic soils as influenced by conservation management. **Journal of Soil and Water Conservation**, v. 55, n. 1, p. 69-78, 2000.
- KHAZAEI, J.; DANESHMANDI, S. Modeling of thin-layer drying kinetics of sesame seeds: mathematical and neural networks modeling. **International Agrophysics**, v. 21, n. 1, p. 335-348, 2007.

- KIM, M.; GILLEY, J. E. Artificial neural network estimation of soil erosion and nutrient concentration in runoff from land application areas. **Computers and Electronics in Agriculture**, v. 64, n. 2, p. 268-275, 2008.
- KLEIN, V. A.; LIBARDI, P. L.; SILVA, A. P. Resistência mecânica do solo à penetração sob diferentes condições de densidade e teor de água. **Engenharia Agrícola**, v. 18, n. 2, p. 45-54, 1998.
- SANTOS, F. L.; DUARTE, M. L. M.; FARIA, M. T. C.; EDUARDO, A. C. Balancing of a rigid rotor using artificial neural network to predict the correction masses. **Acta Scientiarum. Technology**, v. 31, n. 2, p. 151-157, 2009.
- SARMADIAN, F.; MEHRJARDI, T.; AKBARZADEH, A. Modeling of some soil properties using artificial neural network and multivariate regression in Gorgan province, north of Iran. **Australian Journal of Basic and Applied Sciences**, v. 3, n. 1, p. 323-329, 2009.
- SINGH, A. P.; KAY, B. D. Estimating the least limiting water range of soils from properties and management. **Soil Science Society of America Journal**, v. 61, n. 1, p. 877-883, 1997.
- TAVARES FILHO, J.; RIBON, A. A. Resistência do solo à penetração em resposta ao número de amostras e tipo de amostragem. **Revista Brasileira de Ciência do Solo**, v. 32, n. 1, p. 487-494, 2008.
- TURK, G.; LOGAR, J.; MAJES, B. Modelling soil behaviour in uniaxial strain conditions by neural networks. **Advances in Engineering Software**, v. 32, n. 10, p. 805-812, 2001.
- VARELLA, C. A. A.; PINTO, F. A. C.; QUEIROZ, D. M.; SENA JÚNIOR, D. G. Determinação da cobertura do solo por análise de imagens digitais e redes neurais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, n. 2, p. 225-229, 2002.

*Received on November 5, 2010.*

*Accepted on May 16, 2011.*

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.