Development of Pedotransfer Functions to Predict Soil Hydraulic Properties in Golestan Province, Iran

F. Sarmadian and R. Taghizadeh Mehrjardi

Department of Soil Science, College of Agriculture & Natural Resources, University of Tehran, Karaj, 31587-77871, Iran

Abstract
The research presented in this paper attempts to develop a more realistic model using multi-layer perceptron (MLP) and Adaptive neuro-fuzzy inference system (ANFIS) instead of traditional models like multiple linear regression (MLR) for predicting of infiltration rate and deep percolation. Soil samples were collected from different horizons of profiles located in the Gorgan Province, North of Iran. Measured soil variables included texture, organic carbon, water saturation percentage, bulk density, infiltration rate and deep percolation. Then, MLR, ANFIS and ANN models were employed to develop a pedotransfer function for predicting soil parameters using easily measurable characteristics of clay, silt, SP, Bd and organic carbon. The performance of models was evaluated using RMSE. Results showed that the neuro-fuzzy model gives better estimation than the other techniques for all characteristics. After neuro-fuzzy model, artificial neural network had better accuracy than multivariate regression.

Key Words
Infiltration rate, deep percolation, pedotransfer function.

Introduction
In the recent years, the development of prediction methods that use cheap secondary information to spatially extend sparse and expensive soil measurements has been a sharpening focus of research (Bishop and McBratney 2001). Several attempts have been made to estimate indirectly soil properties from more easily measurable and more readily available soil properties such as particle size distribution (sand, silt and clay content), organic matter or organic carbon content, bulk density, porosity, etc. Such relationships are referred to as pedo-transfer functions (PTFs) (Mermoud and Xu 2006).

A recent approach to model PTFs is the use of artificial neural networks (ANNs) (Schaap *et al.* 1998) ANFIS (Wieland and Wilfried Mirschel 2008). ANN offers a fundamentally different approach for modeling soil behavior. ANN is an oversimplified simulation of the human brain and composed of simple processing units referred to as neurons. It is able to learn and generalize from experimental data even if they are noisy and imperfect. This ability allows this computational system to learn constitutive relationships of materials directly from the result of experiments. Unlike conventional models, it needs no prior knowledge, or any constants and/or assumptions about the deformation characteristics of the geomaterials. Other powerful attributes of ANN models are their flexibility and adaptivity, which play an important role in material modeling. When a new set of experimental results cannot be reproduced by conventional models, a new constitutive model or a set of new constitutive equations needs to be developed. However, trained ANN models can be further trained with the new data set to gain the required additional information needed to reproduce the new experimental results. These features ascertain the ANN model to be an objective model that can truly represent natural neural connections among variables, rather than a subjective model, which assumes variables obeying a set of predefined relations (Banimahd *et al.* 2005). In brief, a neural network consists of an input, a hidden, and an output layer all containing “nodes”. The number of nodes in input (e.g. soil bulk density, soil particle size data and etc) and output (different soil properties) layers corresponds to the number of input and output variables of the model (Manyame *et al.* 2007). A type of ANN known as multilayer perceptron (MLP), which uses a back-propagation training algorithm, is usually used for generating PTFs (Schaap *et al.* 1998; Minasny *et al.* 1999; Minasny and McBratney 2002; Amini *et al.* 2005). This network uses neurons whose output is a function of a weighted sum of the inputs.

Since Zadeh (1965) proposed the fuzzy logic approach to describe complicated systems, it has become popular and been successfully used in various engineering problems, especially on control processes (Barreto-Neto and Filho 2008). Ralf Wieland and Wilfried Mirschel (2008) a feed forward neural network (NN), a radial basis function network (RBF) and a trained fuzzy algorithm compared for regional yield estimation of agricultural crops (winter rye, winter barley).

Despite progress made in PTF development in general, little evaluation of PTFs has been done for the soils of humid regions of northern Iran (Golestan Province). Hence the present study was carried out with objective to comparison the efficiency of ANN, ANFIS and MLR for estimation of some soil hydraulic properties using some easily measurable soil parameters.

Materials and methods

Data collection and soil sample analysis

Soil samples were collected from different horizons soil profiles located in the Gorgan Province, North of Iran. Measured soil factors included texture, Organic carbon, infiltration rate and deep percolation. The clod method was used to determine bulk density (Saprks *et al.* 1996).

Methods to fit PTFs

Multivariate regression

The most common method used in estimation PTFs is to employ multiple linear regressions (Minasny and McBratney 2002). For example:

$$Y = aX_1 + bX_2 + cX_3 + ...$$

Where Y is depended variable, Xn is in depended variable and $a,b,$ …. are coefficients.

Artificial Neural Network

ANN is a popular neural network which known as the backpropagation algorithm introduced by Karaca and and Ozkaya (2006). In this study, the training process was performed by the commercial package MATLAB, which includes a number of training algorithms including the back propagation training algorithm. This is a gradient descent algorithm that has been used successfully and extensively in training feed forward neural networks.

Adaptive neuro-fuzzy inference system (ANFIS)

The ANFIS is a multilayer feed-forward network uses ANN learning algorithms and fuzzy reasoning to characterize an input space to an output space. It has been shown to be powerful in modeling numerous processes such as wind speed time series and real-time reservoir operation (Mahmut Firat and Mahmud Gungor 2007).

Evaluation criteria

We used root mean square error (RMSE) to comparison of the efficiency of models that expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(Z_S - Z_O)^2} \qquad (1)$$

Z_s is observed value, Z_o is predicted value, n is number of samples.

Result and Discussion

Data summary of test and train are presented in Table 1 Data subdivided in two sets: 20% of the data for testing and the remaining 80% of the data were used for training.

Table 1. Statistics of the training and test data sets of Infiltration Rate and Deep percolation

		Clay	Silt	BD	SP	OC	I	P
Training set	Min	15.00	19.00	1.30	38.30	0.34	0.25	0.09
	Max	54.00	73.00	1.65	84.00	8.80	6.50	8.70
	Mean	34.30	43.11	1.48	51.99	2.05	1.51	2.55
	Std	10.92	12.37	0.09	12.40	1.89	1.89	2.87
Test set	Min	26.00	30.00	1.30	60.00	1.22	0.40	0.40
	Max	47.00	46.00	1.55	72.60	10.25	4.70	5.50
	Mean	33.60	36.14	1.39	67.80	5.26	1.75	3.11
	Std	7.82	5.46	0.08	4.31	3.38	1.58	2.01

Some soil parameters including: clay, silt, Bulk density, water saturation percentage and organic carbon were input data for prediction of Infiltration rate and Deep percolation. Firs step was to evaluate accuracy of ANN for predicting known data. So we modeled the ANN for predicting of training data. Results revealed that high accuracy of ANN. After confirming of performances of ANN, different neurons were examined for achieving the best neuron for predicting of soil properties. In this stage we used RMSE criteria for determine the best model. Results showed that for infiltration rate five neurons and for deep percolation two neurons had the lowest RMSE.

Then, MLR was computed for soil training data set by MINITAB software. These equations were expressed as:

$$I = 12.7 - 0.188\ Clay - 0.053\ silt - 10.3\ BD + 0.187\ SP - 0.199\ OC \quad (1)$$

$$P = 37.3 - 0.289\ Clay - 0.176\ silt - 17.4\ BD + 0.130\ SP - 0.488\ OC \quad (2)$$

After determining of these equations, performance of MLR was developed for test data set. Results showed that ANN had better performance in predicting all soil properties than MLR which is in line with the work done by Amini *et al.* 2005, Tamari and Wösten (1996), Minasny and McBratney (2002) and Schaap *et al*. (1998).

After MLR and ANN, ANFIS model was computed. There are several fuzzy inference engines which can be utilized for this purpose, which Sugeno and Mamdani are of the most important ones. At this stage, we compute neuro-fuzzy model for predicting mentioned parameters. The best structure of Neuro-fuzzy model obtained according to less RMSE. Result of ANFIS, ANN and MR showed in Table 2. As this table demonstrates ANFIS had the highest accurate for predicting the both parameters.

Table 2. The results of linear regression and neural network-based pedo-transfer functions.

Models	Soil parameters	RMSE
Linear regression	IR	9.44
	DP	7.92
ANN	IR	7.24
	DP	6.54
ANFIS	IR	5.78
	DP	5.55

Conclusion

At present research, we compare applicability and accuracy of three models for prediction of IR and DP. Results revealed that the ANFIS model gives better estimates than the other techniques. After this model, ANN had better accuracy than MLR for prediction of mentioned parameters. It is founded that ANFIS and ANNs had high accuracy for prediction of mentioned parameters but the application of artificial neural networks and fuzzy systems to real problems should be done with care.

References

Amini M, Abbaspour KC, Khademi H, Fathianpour N, Afyuni M, Schulin R (2005) Neural network models to predict cation exchange capacity in arid regions of Iran. *European Journal of Soil Science* **56,** 551-559.

Banimahd M, Yasrobi SS, Woodward PK (2005) Artificial neural network for stress–strain behavior of sandy soils: Knowledge based verification. *Comput. Geotech*. **32,** 377–386.

Barreto-Neto AA, Filho CRDS (2008) Application of fuzzy logic to the evaluation of runoff in a tropical watershed. *Environmental Modelling & Software* **23**, 244-253.

Bishop TFA, McBratney AB (2001) A comparison of prediction methods for the creation of field-extent soil property maps. *Geoderma* **103,** 149–160.

Karaca F, Ozkaya B (2006) NN-LEAP: A neural network-based model for controlling leachate flow-rate in a municipal solid waste landfill site. *Environmental Modelling & Software* **21,** 1190-1197.

Mahmut F, Mahmud G (2007) River flow estimation using adaptive neuro fuzzy inference system. *Mathematics and Computers in Simulation* **75,** 87–96.

Manyame C, Morgan CL, Heilman JL, Fatondji D, Gerard B, Payne WA (2007). Modeling hydraulic properties of sandy soils of Niger using pedotransfer functions. *Geoderma* **141,** 407–415.

Merdun H, Cınar O, Meral R, Apan M (2006). Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil &*

Tillage Research **90**, 108–116.
Minasny B, McBratney AB, Bristow KL (1999) Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma* **93**, 225–253.
Minasny B, McBratney AB (2002) The neuro-m methods for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal* **66,** 352–361.
Schaap MG, Leij FJ, van Genuchten MTh (1998) Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal* **62,** 847–855.
Schaap MG, Leij FJ (1998) Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil & Tillage Research* **47,** 37–42.
Sparks DL, Page AL, Helmke PA, Leoppert RH, Soltanpour PN, Tabatabai MA, Johnston GT, Summer ME (1996) Methods of soil analysis. (Soil Sci. Soc. Of Am., Madison, Wisconsin).
Tamari S, Wosten JHM, Ruiz-Suarez JC (1996) Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Science Society of America Journal* **60**, 1732–1741.
Wieland R, Mirschel W (2008) Adaptive fuzzy modeling versus artificial neural networks. *Environmental Modelling & Software* **23,** 215-224.