Environmental Research Journal 3 (2): 19-24, 2009
ISSN: 1994-5396

Application of Artificial Intelligence in Modeling of Soil Properties (Case Study: Roodbar Region, North of Iran)

[1]A. Akbarzadeh, [1]R. Taghizadeh Mehrjardi, [2]H. Rahimi Lake and [2]H. Ramezanpour,
[1]Department of Soil Science, College of Agriculture and Natural Resources,
University of Tehran, Tehran, Iran
[2]Department of Soil Science, University of Guilan, Iran

Abstract: Investigation of soil properties like Cation Exchange Capacity (CEC) plays important roles in study of environmental researches as the spatial and temporal variability of this property have been led to development of indirect methods in estimation of this soil characteristic. Therefore, in this study indirect methods have been used to estimate cation exchange capacity. Eighty soil samples were collected from different horizons of 26 soil profiles located in the Roodbar region, Guilan Province, North of Iran. Measured soil variables included texture, organic carbon and cation exchange capacity. Then, multiple linear regression, Neuro-Fuzzy and feed-forward back-propagation network were employed to develop a pedotransfer function for predicting soil parameter using easily measurable characteristics of clay and organic carbon. Results showed that Neuro-Fuzzy was superior to artificial neural network and MR in predicting soil property.

Key words: Pedotransfer functions, artificial neural network, Nero-Fuzzy, CEC

INTRODUCTION

Cation Exchange Capacity (CEC) is the amount of negative charge in soil that is available to bind positively charged ions (cations). Essential plant nutrients, K^+, Ca^{2+}, Mg^{2+} and NH^{4+} and detrimental elements, Na^+, H^+ and Al^{+3} are cations. Cation exchange capacity is used as a measure of fertility, nutrient retention capacity and the capacity to protect groundwater from cation contamination. Cation exchange capacity buffers fluctuations in nutrient availability and soil pH. Soil components known to contribute to CEC are clay and organic matter and to a lesser extent, silt (Seybold *et al*., 2005).

A neural network is an attempt to build a mathematical model that supposedly works in an analogous way to human brain. A network consists of many elements or neurons that are connected by communication channels or connectors. These connectors carry numeric data arranged by a variety of means and organized into layers. The neural networks can perform a particular function when certain values are assigned to the connections or weights between elements. To describe a system, there is no assumed structure of the model, instead the networks are adjusted or trained so that a particular input leads to a specific target output (Minasny and McBratney, 2002).

The fuzzy logic approach is based on the linguistic uncertain expression rather than numerical uncertainty. It is an artificial intelligence technique that has been used currently in hydrological processes. 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). Nonetheless, the main problem with this approach is that there is no systematic procedure for a design of fuzzy controller. However, a neural network system has the ability to learn its structure from the input-output sets and adapt itself in an interactive manner. So, many researchers were proposed the usage of the ANFIS, consisted of the combination of the ANN and the fuzzy logic, to organize network structure itself and to adapt the parameters of fuzzy system to solve many engineering problems such as the estimating a controlled reservoir water level.

Tamari *et al*. (1996) gave a review on ANN and their application in predicting soil hydraulic properties. Most

Corresponding Author: R. Taghizadeh Mehrjardi, Department of Soil Science, College of Agriculture and Natural Resources, University of Tehran, Tehran, Iran

researchers have found that ANN performs better than multiple regressions. Amini *et al*. (2005) tested several published PTFs and developed two neural network algorithms using multilayer perceptron and general regression neural networks based on a set of 170 soil samples for predicting of Cation exchange capacity in central Iran. They found that the neural network-based models provided more reliable predictions than the regression-based PTFs. Minasny and McBratney (2002) claimed that an advantage of using the neural network approach is that no relationships need to be assumed beforehand. Schaap *et al*. (1998) used ANNs for predicting of some soil hydraulic properties. They also confirmed applicability of ANNs and concluded that accuracy of these models depend on number of inputs.

The objective of this study, is to evaluate the general applicability of artificial neural network, Nero-Fuzzy and multivariate regression in estimating cation exchange capacity in the soils of Iran.

MATERIALS AND METHODS

Data collection and soil sample analysis: The study area is located in the south direction of Roodbar city and by the side of Shahrood River in Guilan province of Iran. This study carried out in an area including 2200 ha between 36° 41' 11″-36° 37' 52″ in northern latitude and 49° 27' 20″-49° 31' 3″ eastern longitude. After interpretation of aerial photographs the digging site of soil profile was identified. Twenty six pedons was selected and then 80 soil samples were collected from different horizons of these profiles. Figure 1 illustrates the location of the study area. Soil moisture and temperature regimes of the region by means of Newhall software were arid and thermic, respectively. The soils were classified in 2 orders of Aridisols and Entisols on the basis of soil taxonomy system (2006). Measured soil factors included texture (determined by Bouyoucos hydrometer method), Organic carbon (determined Using Walkely and Black rapid titration) and CEC.

Methods to fit PTFs

Multivariate regression: The most common method used in estimation PTFs is to employ multiple linear regressions. For example:

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

where:

Y	=	Depended variable
Xn	=	Depended variable
a,b, ...	=	Are coefficients

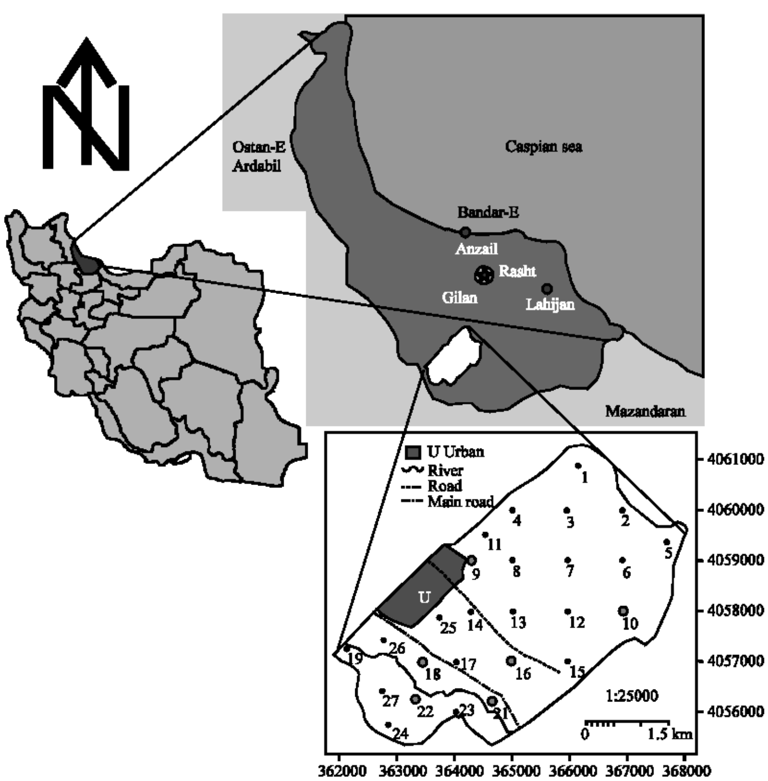

Fig. 1: Study area in North of Iran (Guilan Province) and sampling locations

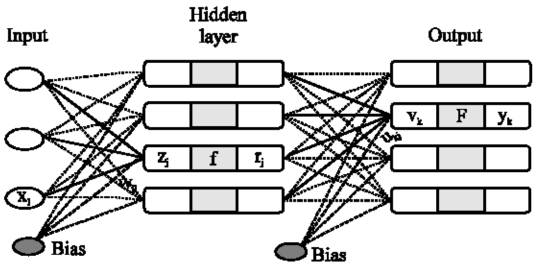

Fig. 2: Structure of feed-forward ANN

Feedforward neural networks: Artificial Neural Networks (ANNs) are universal estimators of multivariate non-linear mappings that are capable of learning and generalizing from examples (training data). The key to successfully training an Artificial Neural Network is choosing the right network architecture and training algorithm. A feedforward artificial neural network is used in this study to approximate the relation between hydraulic conductivity/Transmissivity values of the region in question resulting hydraulic conductivity values. Feedforward networks are a subclass of layered networks in which there no intra-layer connections are and whosemain feature is that connections are allowed from node 'i' only to nodes in layer iC1. Feedforward neural networks are among the most common neural networks in use (Luis and Shigidi, 2006). They were chosen for use in this study because they are simple, easily trained and can be readily inverted. The feedforward process from which the name was derived involves presenting an input pattern to input layer neurons that pass the input values into the first hidden layer. Each of the hidden layer nodes (neurons) computes a weighted sum of the inputs, passes the sum through the transfer (activation) function and presents the results to the next layer until the output layer is reached. Determining the architecture of a neural network involves determining the number of layers in the network as well as the number of nodes (neurons) in each layer (Mehrotra *et al.*, 1997). 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 Fig. 2.

Adaptive Neuro-Fuzzy Inference System (ANFIS): Fuzzy inference system is a rule based system consists of three conceptual components. These are: a rule base, contains fuzzy if-then rules, a database, defines the membership function and an inference system, combines the fuzzy rules and produces the system results. First phase of fuzzy logic modeling is the determination of membership functions of input-output variables, second is the construction of fuzzy rules and the last is the determination of output characteristics, output membership function and system results. To determine the membership function of the input-output variables, two methods, named as backward propagation algorithm and hybrid-learning algorithm, provide learning of the ANFIS and construction of the rules, are used. A general structure of fuzzy system is demonstrated in Fig. 3.

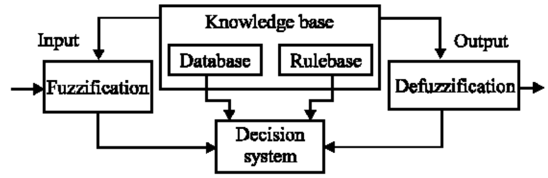

Fig. 3: The general structure of the fuzzy inference system

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. ANFIS possesses properties such as capability of learning, constructing, expensing and classifying. It has the advantage of allowing the extraction of fuzzy rules from numerical data or expert knowledge and adaptively constructs a rule base. Moreover, it can adapt the complicated conversion of human intelligence to fuzzy systems. The main difficulty of the ANFIS predicting model is the time required for training structure and determining parameters. In this study, the ANFIS method consisting of the combination of the artificial neural networks and fuzzy logic approach has been used to estimate the River flow. ANFIS uses the learning ability of the ANN to define the input-output relationship and construct the fuzzy rules by determining the input structure. The system results were obtained by thinking and reasoning capability of the fuzzy logic. The hybrid learning algorithm and subtractive function were used to determine the input structure. The detailed algorithm and mathematical background of the hybrid-learning algorithm can be obtained from the research of Jang *et al.* (1997). The consequence parameter in Sugeno inference system is a linear equation or constant coefficient. The linear equation is called zero-order Sugeno inference system and the constant type is called first-order Sugeno inference system. For simplicity, it was assumed that the fuzzy inference system had 2 inputs, x and y and one output, z (Jain and Kumar, 2006; Jang, 1993; Jang *et al.*, 1997;

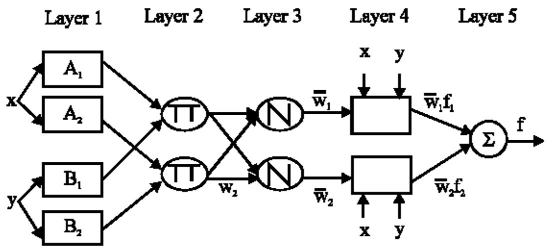

Fig. 4: Equivalent ANFIS architecture

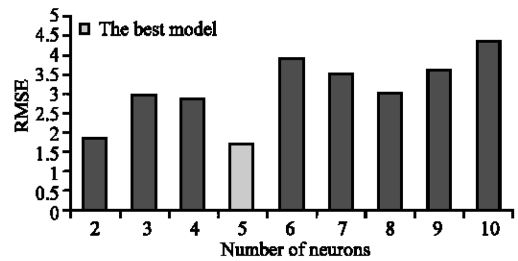

Fig. 5: RMSE value for 2-10 neurons (cation exchange capacity)

Firat and Gungor, 2007). The structure of this fuzzy inference system is shown in Fig. 4. For the first-order Sugeno inference system, typical 2 rules can be expressed:

Rule 1: If x is A_1, y is B_1 then $f_1 = P_1x+q_1y+r_1$
q_2y+r_2+ Rule 2: If x is A_2, y is B_2 then $f_2 = P_1x$

The resulting Sugeno fuzzy reasoning system is shown in Fig. 4. Here, the output z is the weighted average of the individual rules outputs and is itself a crisp value. The corresponding ANFIS architecture is shown in Fig. 3. Nodes at the same layer have similar functions. The output of the ith node in layer l is denoted as Ol, i.

Evaluation criteria: Accuracy of the regression equations for derivation of PTFs was evaluated using R^2 and RMSE between the measured and predicted values and expressed as:

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

Z_s = Observed value
Z_o = Predicted value
n = Number of samples

RESULTS AND DISCUSSION

Some soil parameters including: clay and organic carbon were input data for prediction of CEC. Amini *et al.* (2005) stated that CEC has high correlation with these inputs. He found that inputs like sand and silt can not improve accuracy of prediction of CEC. The RMSE of the different neurons in hidden layer is plotted in Fig. 5. This Fig. 5 illustrated that the best model obtained with 5 neurons for CEC. Correlation coefficient and RMSE have been obtained 0.89 and 1.7 for cation exchange capacity.

Multi regression was computed for 3 soil train data set by MINITAB software. These equations were expressed as:

$$\text{CEC} = 2.31 + 0.408\text{Clay} + 2.86\text{OC} \tag{2}$$

After determining of these equations, performance of multivariate regression was developed for test data set. Correlation coefficient and RMSE have been obtained 0.72 and 5.3 for cation exchange capacity.

Results showed that artificial neural network with 5 neurons in hidden layer had better performance in predicting CEC than multivariate regression which is in line with the work done by Amini *et al.* (2005), Tamari *et al.* (1996), Minasny and McBratney (2002) and Schaap *et al.* (1998). Amini *et al.* (2005) found that the neural network-based models provided more reliable predictions than the regression-based PTFs. Schaap *et al.* (1998) confirmed applicability of ANNs and concluded that accuracy of these models depend on number of inputs. Koekkoek and Booltink (1999) found that ANN performed slightly better, but the differences were not significant. One of the advantages of neural networks compared to traditional regression PTFs is that they do not require a priori regression model, which relates input and output data and in general is difficult because these models are not known Schaap and Leij (1998).

In the main, each fuzzy system consists of three main sections, Fuzzifier, Fuzzy data base and Defuzzifier. At first, input information is made as fuzzy data after bypassing the fuzzifier section. This operation is done by membership functions, in which the precise amount value becomes as fuzzy values by membership functions. Later then, fuzzy parameters are entered to the fuzzy data base. Fuzzy data base includes two main sections, Fuzzy rule base and inference engine. In fuzzy rule base, rules related to fuzzy propositions are described. Thereafter, analysis operation is applied by fuzzy inference engine. 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 parameter. The best structure of Neuro-fuzzy model obtained according to less RMSE. The characteristic and structure of this model showed in Table 1.

Table 1: Model structures for estimation of CEC

3	Number of Mfs
254	Epoch
trimf	Mf type
Back propegation	Optimum method
Wtaver	Defuzz method

Table 2: Comparison among different models for predicting CEC

Models	RMSE	R^2
ANFIS	0.87	0.97
ANN	1.70	0.89
MR	5.30	0.72

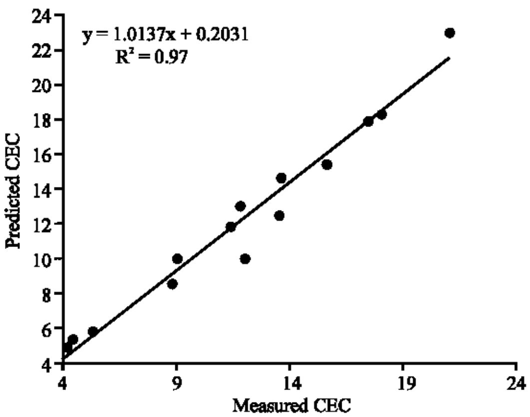

Fig. 6: The scatter plot of the measured versus predicted CEC

Result of ANFIS, ANN and MR showed in Table 2. As this table demonstrates ANFIS had the accurate for predicting the parameter. The value of RMSE for estimation of CEC was 0.87.

Comparing the 3 estimation models, it can be seen that the values of the RMSE of the ANFIS model is much lower than ANN and MR model. It appears that the RMSE of the ANFIS model is lower compared to the ANN and the MR during testing. The RMSE value of the ANFIS model was also lower than ANN model and MR. It may be noted that a trial and error procedure has to be performed for ANN model to develop the best network structure, while such a procedure is not required in developing an ANFIS model. Moreover, in the current study, the ANFIS model was trained by using just 254 epochs, while the ANN model took 850 epochs. The results suggest that the ANFIS method is superior to the ANN method in the modeling and forecasting of CEC.

The scatter plot of the measured against predicted CEC for the test data set is given in Fig. 6 for the ANFIS model, which we identified as being the best model for predicting soil parameter. As this figure showed that ANFIS model predicted mentioned property with very high accuracy which this point demonstrate applicability and performance of ANFIS and fuzzy logic for prediction of CEC.

CONCLUSION

At present research, we compare applicability and accuracy of 3 models for prediction of CEC. Results revealed that the neuro-fuzzy model gives better estimates than the other techniques. After neuro-fuzzy model, artificial neural network had better accuracy than multivariate regression for prediction of CEC. It was founded that ANFIS and ANNs had high accuracy for prediction of mentioned parameter but the application of artificial neural networks and fuzzy systems to real problems should be done with care. A good software basis with an integrated graphical analysis can relax the situation significantly. There still remain, however, many traps for the modeler.

ACKNOWLEDGEMENT

We wish to thank University of Guilan for supporting field studies and samplings. We would like to gratefully thank all the members of the soil science laboratory of faculty of agriculture, University of Guilan, for providing the facilities to carry out this study and for their suggestions.

REFERENCES

Amini, M., K.C. Abbaspour, H. Khademi, N. Fathianpour, M. Afyuni and R. Schulin, 2005. Neural network models to predict cation exchange capacity in arid regions of Iran. Eur. J. Soil Sci., 56: 551-559. 10.1111/j.1365-2389.2005.0698.x. http://www3.interscience.wiley.com/journal/118710376/abstract.

Barreto-Neto, A.A. and C.R.D.S. Filho, 2008. Application of fuzzy logic to the evaluation of runoff in a tropical watershed. Environ. Mod. Soft, 23: 244-253. 10.1016/j.envsoft.2007.07.006.http://www.sciencedirect.com/science?_ob=ArticleURLand_udi=B6VHC4PMJB84and_user=1400009and_coverDate=02%2F29%2F2008and_alid=842703388and_rdoc=1and_fmt.

Firat, M. and M. Gungor, 2007. River flow estimation using adaptive neuro fuzzy inference system. Math. Comput. Simulation, 75: 87-96. Portal.acm.org/citation.cfm?id=1265806anddl=ACMandcoll=portalandCFID=15151515andCFTOKEN=6184618.

Jain, A. and A. Kumar, 2006. An evaluation of artificial neural network technique for the determination of infiltration model parameters. Applied Soft Comp., 6: 272-282. 10.1016/j.asoc.2004.12.007.http://www.sciencedirect.com/science?_ob=ArticleURLand_udi=B6W864FX23CJ1and_user=1400009and_rdoc=1and_fmt=and_orig=searchand_sort=dandview=cand_acct=C000052577and_version=1and_urlVersion=0.

Jang, J.S.R., 1993. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans. Sys. Manage. Cybernetics, 23 (3): 665-685. 25872. https://eprints.kfupm.edu.sa/25872/.

Jang, J.S.R., C.T. Sun and E. Mizutani, 1997. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligance. Prentice Hall, Upper Saddle River, New Jersey, USA. portal.acm.org/citation.cfm?id=1290619.

Koekkoek, E.J.W. and H. Booltink, 1999. Neural network models to predict soil water retention. Eur. J. Soil Sci., 50: 489-495. www.blackwell-synergy.com/doi/abs/10.1046/j.1365-2389.1999.00247.x.

Luis, A.G. and A. Shigidi, 2006. Using neural networks for parameter estimation in ground water. J. Hydro., 318: 215-231. linkinghub.elsevier.com/retrieve/pii/S0022169405003057.

Mehrotra, K., C.K. Mohan and S. Ranka, 1997. Elements of Artificial Neural Networks. The MIT Press, Boston,pp: 344. 10.1006/jmps.1997.1244. 0262-133283. http://books.google.com/books?hl=enandlr=andid=6d68Y4Wq_R4Candoi=fndandpg=PR13anddq=%22Mehrotra%22+%22Elements+of+Artificial+Neural+Networks%22+andots=6pzdU1D1Aaandsig=CJXZUte7EB76HXMJB05lWgnhqkw#PPT1,M1.

Minasny, B. and A.B. McBratney, 2002. The neuro-m methods for fitting neural network parametric pedotransfer functions. Soil Sci. Soc. Am. J., 66: 352-361. 10.2335/sssaj2001.0018. linkinghub.elsevier.com/retrieve/pii/S0016706106001066.

Schaap, M.G., F.J. Leij and M.T. Van Genuchten, 1998. Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J., pp: 62. 847855.10.1340/sssaj1998.0408.soil.scijournals.org/cgi/content/abstract/62/4/847.

Schaap, M.G. and F.J. Leij, 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. Soil Till Res., 47: 37-42. 10.1103/sssaj1998.0009.linkinghub.elsevier.com/retrieve/pii/S0167198705002436.

Seybold, C.A., R.B. Grossman and T.G. Reinsch, 2005. Predicting cation exchange capacity for soil survey using linear models. Soil Sci. Soc. Am. J., 69: 856-886. 10.2136/sssaj2004.0026.http://soil.scijournals.org/cgi/content/abstract/69/3/856.

Tamari, S., J.H.M. Wosten and J.C. Ruiz-Suarez, 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. Soil Sci. Soc. Am. J., pp: 60. 17321741.10.2033/sssaj1996.0026.soil.scijournals.org/cgi/content/abstract/60/6/173.

Zadeh, L.A., 1965. Fuzzy sets. Inform. Cont., 8: 338-353. books.google.com/books?isbn=981022421.