International Scholars Journals

African Journal of Ecology and Ecosystems ISSN: 9428-167X Vol. 6 (3), pp. 001-013, March, 2019. Available online at www.internationalscholarsjournals.org © International Scholars Journals

Full Length Research Paper

Development of pedo transfer functions (PTFs) to predict soil physico-chemical and hydrological characteristics in southern coastal zones of the Caspian Sea

Hadi Rahimi Lake[1], Ali Akbarzadeh[2] and Ruhollah Taghizadeh Mehrjardi[2*]

[1]Department of Soil Science, Faculty of Agriculture, University of Guilan, Rasht, 41635- 1314, Iran. [2]Department of Soil Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, 31587-77871, Iran.

Accepted 17 January, 2019

The research presented in this paper attempts to develop a more realistic model using multi-layer perceptron (MLP), a feed forward artificial neural network (ANN), instead of traditional models like multiple linear regression (MLR) for predicting some soil physico-chemical and hydrological properties. The study area (Guilan Province) which is located in northern Iran bordering to south side of Caspian Sea in a coastal zone has udic and thermic soil moisture and temperature regimes respectively. The estimated soil parameters were CEC, EC, ESP, MWD and final steady- state infiltration rate (IR). Although these parameters can be measured directly, their measurement is difficult and expensive, so pedotransfer functions (PTFs) provide an alternative by estimating these parameters from more readily available soil data. In order to predict the mentioned parameters, soil sampling was conducted at 500 points in the region. Measured soil variables included texture, O.C, porosity, EC, CEC, SAR, ESP, MWD, soluble cations and anions and IR. Then, ANN and MLR models were tested. The data set was divided into two subsets for calibration (80%) and testing (20%) of the models and their normality were tested by Minitab software and Kolmogrov-Smirnov method. In order to evaluate the models, root mean square error (RSME) was used. The comparison of RSME for two mentioned models showed that the ANN model gives better estimates rather than the MLR model. So that the levels of RMSE and R^2 derived by ANNs models for EC, CEC, ESP, MWD and IR were 0.24, 0.96; 1.25, 0.90; 0.18, 0.94; 0.04, 0.84 and; 1.55, 0.92 respectively while these parameters for MLR models were 1.98, 0.73;, 7.92, 0.60; 1.13, 0.66; 0.187, 0.51 and; 9.45, 0.57 respectively. The superiority of ANN models compared with MLR models was probably due to a nonlinear relationship between the dependent and independent variables. Furthermore, results indicated that training is very important in increasing model accuracy for one region.

Key words: Artificial neural networks, multiple linear regressions, soil properties.

INTRODUCTION

Measurement of some soil physico-chemical and hydrological properties both at field and laboratory conditions are cumbersome, expensive, time-consuming, labour-intensive and they give only local scale results. In the resent 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). PTFs can be categorized into three main groups namely class PTFs,

*Corresponding author. E-mail: Rh_Taghizade@yahoo.com.

continuous PTFs and neural networks. Class PTFs calculate soil properties (e.g. soil hydraulic properties) for a textural class (e.g. sand) by assuming that similar soils have similar properties; continuous PTFs on the other hand, use measured percentages of clay, silt, sand and organic matter content to provide continuously varying soil properties across the textural triangle (Manyame et al., 2007) . Multiple linear regression (MLR) models commonly are classified into these two groups of PTFs. MLR analysis is generally used to find the relevant coefficients in the model equations. Often, however, models developed for one region may not give adequate estimates for a different region (Wagner et al., 2001).

A recent approach to model PTFs is the use of artificial neural networks (ANNs) (Schaap et al., 1998). 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.

The major advantage of neural networks over the two groups of PTFs described earlier is that they do not require a-priori concept of the relations between input and output data (Schaap and Leij, 1998). However, because of their greater feasibility, ANN models are generally expected to be superior to MLR models (Sarmadian et al., 2009; Amini et al., 2005; Minasny et al., 1999).

Many studies related to modeling various soil parameters using different types of PTFs has been conducted yet. For example, Vos et al. (2005) used 12 PTFs and Brazilian's database for prediction of bulk density. Their results showed that the separation of subsoil data from topsoil data did not increase the accuracy of prediction. Similarly, Heusher et al. (2005) and Kaur et al. (2002) reported that the soil texture and organic matter content were the main parameters for estimating of bulk density. Schaap et al. (1998) developed some functions for estimation of the different parameters of Vangenokhten, Vangenokhten-moalem, and Gardner equations by means of ANNs. Their results showed that with increasing the number of input data, the accuracy of functions would enhance. Najafi and Givi (2006) used the ANNs and PTFs methods for prediction of soil bulk density. They pointed out that the ANNs are able to predict the soil bulk density better than the PTFs. Amini et al. (2005) estimated the cation exchange capacity in the central of Iran using soil organic matter and clay contents. They used the ANN and five experimental models that were on the basis of regression methods for their predictions. They showed that a neural network PTF with eight hidden neurons was able to predict CEC better than the regression PTFs. Also the ANN model significantly improved the accuracy of the prediction by up to 25%. They concluded that network models are in general more suitable for capturing the non-linearity of the relationship between variables. Jain and Kumar (2006) indicated that the ANN technique can be successfully employed for the purpose of calibration of infiltration equations. They had also found that the ANNs are capable of performing very well in situations of limited data availability. Jiang and Cotton (2004) reported that the ANN model has show good performance when trained and tested with a spatially distributed dataset for estimation of soil moisture. They also derived correlation coefficient of 0.95 and 0.99 for the relationship between the model estimate and the soil moisture for training and testing respectively. In contrast Merdun et al. (2006) pointed out that although the differences between regression and ANN models were not statistically significant, regression predicted point and parametric variables of soil hydraulic parameters better than ANN.

Despite progress made in PTF development in general, little evaluation of PTFs has been done for the soils of humid regions of northern Iran (Guilan Province). Hence the present study was carried out with objective to comparison the efficiency of ANNs, multivariable regression, and three PTFs for estimation of some soil physico-chemical and hydraulic properties using some easily measurable soil parameters.

MATERIALS AND METHODS

Study area

The study area is located in northern Iran bordering to Caspian Sea in Guilan province of Iran lying between 36° 00´ northern latitude and 51°00´ eastern longitude (Figure 1). The climate of the region is humid with the mean annual precipitation of 1,250 mm. The mean annual temperature of the region is 15.5°C. The mean humidity is 75% and the annual evapotranspiration is 850 mm. The soil moisture and temperature regimes of the region by means of Newhall software are udic and thermic, respectively. The major geological formations are composed of thick sedimentary and metamorphic rocks of Tertiary and Quaternary periods. The coastal plain lying between Alborz mountain ranges and Caspian Sea is composed of marine, river and aeolian deposits of varying thicknesses. The physiographical units of the region from south to north direction are river alluvial plains, river bank, low lands, coastal lands and sand duns respectively. According to soil taxonomy system (USDA 2006) the soil of the region are classified in three orders of Alfisols, Inceptisols and Entisols.

Data collection and soil sample analysis

After preliminary studies of geological (1:100000, 1:250000) and topographic maps, using GPS, studying locations were appointed. 500 soil samples were collected from different horizons of 193 soil profiles located in Guilan Province. Particle-size distribution was determined after dissolution of $CaCO_3$ with 2 N HCl and decomposition of organic matter with 30% H_2O_2. After repeated washing to remove salts, samples were dispersed using sodium hexametaphosphate for determination of sand, silt and clay fractions by the pipette method (Day, 1965). The silt+fine sand content was determined by Kittrick and Hope, (1963) method. MWD (mean weight diameter in mm) was measured by dry sieving method (Klute, 1986). Organic carbon (O.C) was determined by Walkley-Black method (Nelson and Sommers, 1982); CEC (cation exchange capacity in cmolc kg^{-1} soil) by the method of Chapman (1965); EC (electrical conductivity in dS m^{-1}) of the saturated extract by the Bower and Wilcox (1965) method. Also, soluble (meq l^{-1}) cations and anions (such as Na^+ and Cl^-), SAR (sodium adsorption ratio) and ESP (exchangeable sodium percentage) were measured with respect to standard methods (Page, 1986). The total porosity was calculated using the following equation:

Percentage of total porosity = (1 - bulk density) × 100 (1)

The bulk density was measured by undisturbed sampling using cylinder shape vessels (Klute, 1986). The particle density was measured using the pycnometer method (Blake and Hartge, 1986). Also, the final steady-state infiltration rate was measured by double-ring variable-water level infiltrometers (Doaei et al., 2005). The internal diameter was 30 cm for inner and 60 cm for outer ring. At the final steady-state infiltration rate which is unique for each soil type, the velocity variations of percolating water into the soil with the time become constant. Before analysis, all data were normalized to have zero mean and unit variance; then the results were converted to the original scale. The normality of dataset was tested by Minitab software and Kolmogrov-Smirnov method (Ghorbani-Dashtaki and Homaei, 2002).

Methods to fit PTFs

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

$Y = aX_1 + bX_2 + cX_3 + \ldots$ (2)

Where Y denotes depended variable, Xn is independent variable and $a, b, \ldots$ are coefficients.

Artificial neural network: An artificial neural network is a highly interconnected network of many simple processing units called neurons, which are analogous to the biological neurons in the human brain. Neurons having similar characteristics in an ANN are arranged in groups called layers. The neurons in one layer are connected to those in the adjacent layers, but not to those in the same layer. The strength of connection between the two neurons in adjacent layers is represented by what is known as a 'connection strength' or 'weight'. An ANN normally consists of three layers, an input layer, a hidden layer, and an output layer. In a feed forward network, the weighted connections feed activations only in the forward direction from an input layer to the output layer. On the other hand, in a recurrent network additional weighted connections are used to feed previous activations back into the network. The structure of a feed-forward ANN is shown in Figure 2. This ANN is a popular neural network which known as the back propagation algorithm introduced by Karaca and Ozkaya (2006). This ANN had k input and one output parameters. They used this ANN for accurate modeling of the leachate flow-rate. They also reported that the input parameters, number of neurons at the hidden and output layer should be determined according to currently gathered data. Moreover, an important step in developing an ANN model is the training of its weight matrix. The weights are initialized randomly between suitable ranges, and then updated using certain training mechanism (Minasny and McBratney, 2002; Pachepsky et al., 1996; Schaap et al., 1998).

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.

Evaluation criteria

Accuracy of the regression equations for derivation of PTFs was evaluated using coefficient of determination (R^2) and root mean square error (RMSE) expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(Z_S - Z_O\right)^2} \quad (3)$$

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

RESULTS

Data summary information

Data summary of different soil properties used to testing and training are presented in Tables 1 and 2. Data subdivided in two sets: 20% of the data for testing and the remaining (80%) of the data for training or calibrating. For achievement to this subset, we selected the data in the manner in which some of their statistical criteria such as mean and standard deviation (Std.) were similar to each other as much as possible. The values of mean and standard deviation of training and testing data for

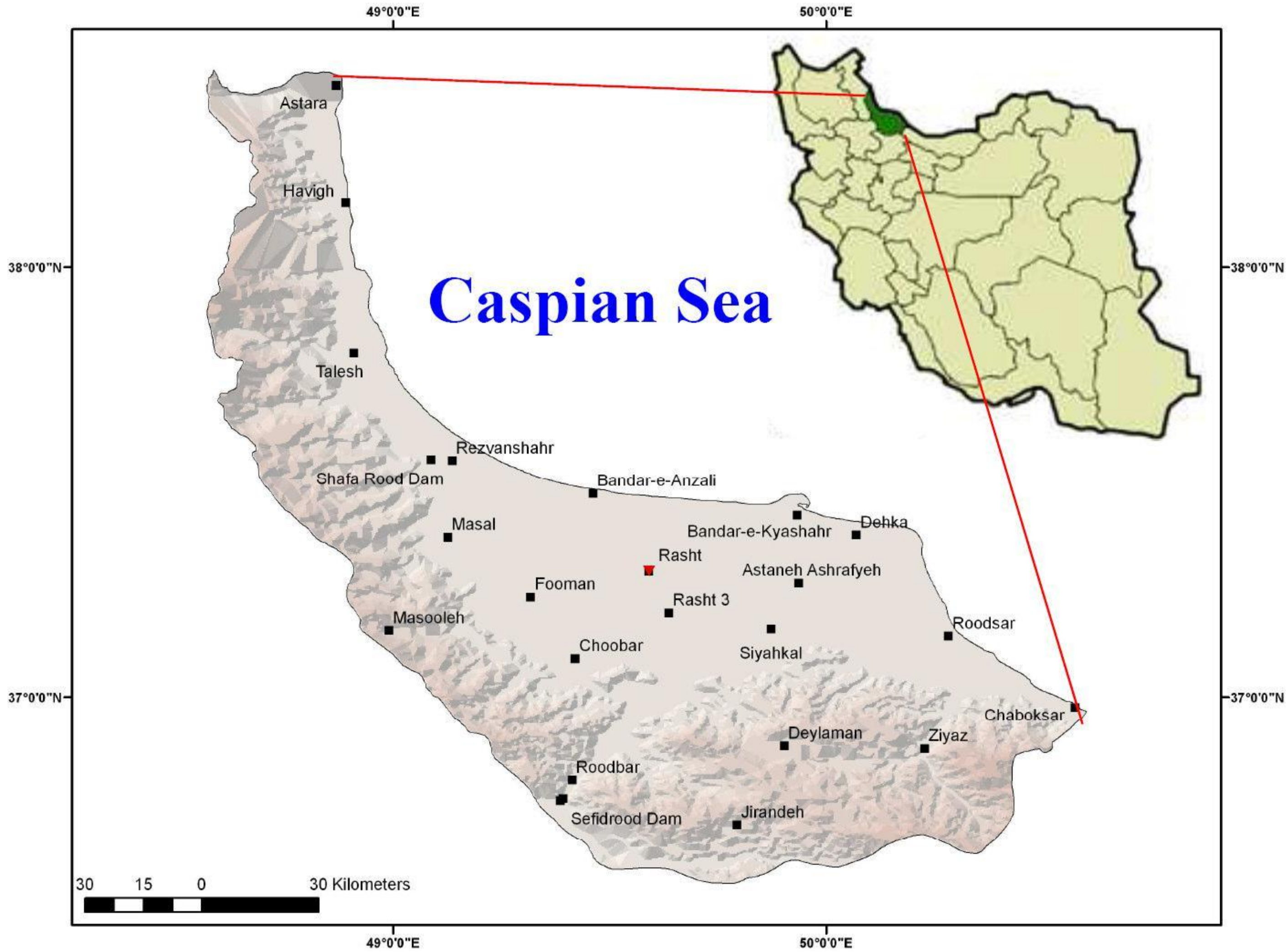

Figure 1. Location of Guilan province in northern Iran.

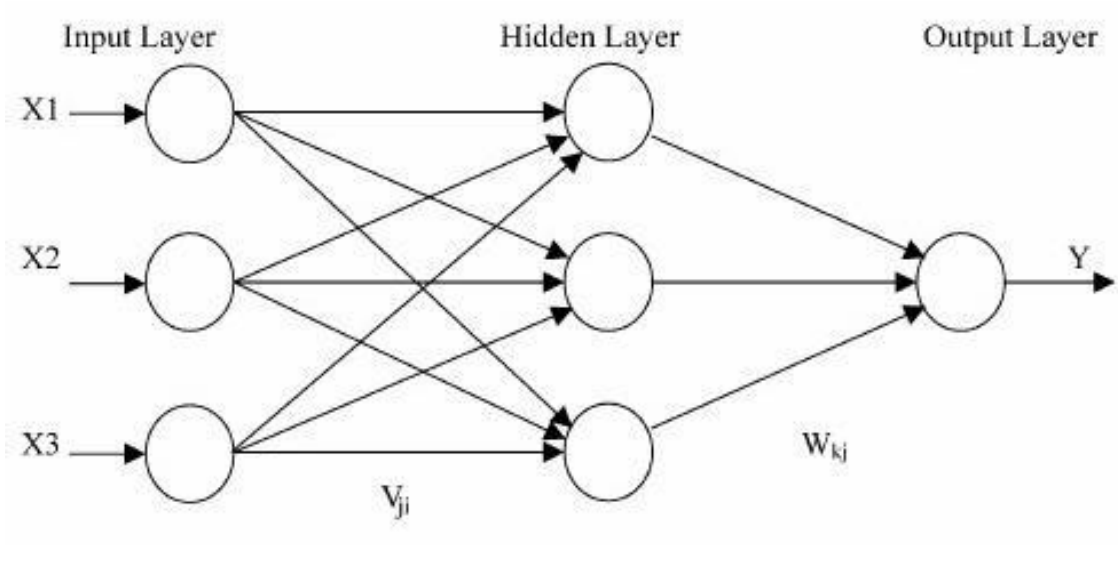

Figure 2. Architecture of an artificial neural network.

Table 1. Statistics of training and test data sets for some soil physical and hydrological properties.

		Clay (%)	Silt+fine sand (%)	Infiltration rate (mmh^{-1})	Porosity (%)	MWD (mm)
Training set	Min	1.02	13.60	5.08	39.55	0.20
	Max	67.30	40.69	88.13	76.94	0.91
	Mean	34.66	27.47	43.78	57.52	0.58
	Std.	14.49	8.71	6.34	17.48	0.57
Test set	Min	7.43	16.61	6.02	42	0.27
	Max	56.11	44.21	82.25	72.72	0.79
	Mean	32.20	29.58	37.70	54.05	0.51
	Std.	13.90	10.74	4.13	18.92	0.40

Table 2. Statistics of training and test data sets for some soil chemical properties.

		O.C (%)	CEC ($cmol_c\ kg^{-1}$)	EC_e ($dS\ m^{-1}$)	Na^+_s ($meq\ l^{-1}$)	Cl^-_s ($meq\ l^{-1}$)	ESP (%)	SAR
Training set	Min	0.02	5.42	0.73	4.09	0.75	0.27	0.29
	Max	5.89	33.10	8.32	12.29	3.10	6.98	7.3
	Mean	1.19	20.70	4.46	7.22	1.93	3.09	3.94
	Std.	0.98	4.55	0.34	5.91	3.94	1.09	4.36
Test set	Min	0.09	12.81	0.86	6.21	0.88	0.39	0.18
	Max	11.30	32.02	8.91	10.10	2.14	6.13	6.33
	Mean	2.08	23.96	4.18	8.43	1.52	3.14	3.16
	Std.	2.89	3.54	0.32	6.05	2.02	1.07	5.55

different soil parameters are also presented in these tables. As shown in Table 2, the O.C content of the soils in the region is usually very high, ranging from 0.02 - 5.89%, with an average of 1.19% in this study. The reason of this high content of O.C is due to that a large area of the region is located in forest area. Although in the some strip zones such as near to the beach and sand duns in northern parts and also, near to the mountainous areas in southern parts of the study area the O.C content into the soil was negligible. Hence, the heterogeneity of the soil O.C and other soil parameters is very high in the region.

In general the clay content of the most of the region except those areas explained for organic carbon is rather high. Therefore, because of dependency of CEC to O.C and clay content and their positive relationships, the high level of CEC in the region is expected. In addition with due attention to Table 1, it is obvious that the level of infiltration rate, MWD and soil porosity is quite high because; in the forest areas of the region due to presence of aggregates having large volume into the soil, the mean weight diameter of soil particles is naturally more than those areas located in the beach and mountainous zones. Consequently, because of existing direct relationships between soil porosity and infiltration rate with MWD; these soil characteristics had large levels. Furthermore, other soil chemical properties including soluble Na^+ and Cl^- due to high precipitation (1250 mm) and leaching of soluble cations and anions is relatively low. Thus, the soils electrical conductivity has not large value in the region same to the Na^+ and Cl^-. Nevertheless in some parts near to the Caspian Sea due to high soil salinity, the EC shows large values. Besides, dependency of SAR and ESP to basic cations such as exchangeable (Na^+_e) will cause these parameters show low levels in the region except areas near to the sea side (Table 2).

Training and testing of PTFs

Correlation matrixes among various soil parameters were also calculated (Table 3). This correlation matrix will help us to distinguish those soil parameters having the most correlation with each other. As Table 3 illustrates correlations among O.C, clay and CEC and also, between EC, Na^+ and Cl^- were positive and highly significant. For example the correlation coefficients between CEC and

Table 3. Correlation matrix among different soil parameters.

	O.C	Clay	CEC	Silt + fine sand	I.R.	Porosity	MWD	EC	Na	Cl	ESP	SAR
OC	1											
Clay	0.31*	1										
CEC	0.62**	0.69**	1									
Silt + fine sand	0.37*	-0.22*	-0.39*	1								
I.R.	0.46**	-0.44**	0.31	-0.47**	1							
Porosity	0.57**	-0.49*	0.42	-0.61**	0.78**	1						
MWD	0.84**	-0.60**	0.34	-0.59**	0.46**	0.84**	1					
EC	0.33*	0.38	0.09	0.38*	0.45**	0.71**	0.39**	1				
Na	0.27*	0.34*	0.19	0.37*	-0.38**	-0.52*	-0.28**	0.74**	1			
Cl	0.19*	0.13*	0.03	0.19*	0.34*	0.27**	0.14*	0.60**	0.88**	1		
ESP	0.15*	0.14*	0.27	0.24	-0.41**	-0.74*	-0.40**	0.33*	0.91**	0.18	1	
SAR	0.28*	0.08*	0.20	0.22	-0.50**	-0.79**	-0.48**	0.36*	0.86**	0.13	0.84**	1

*p < 0/05, **p < 0/01.

Table 4. The results of linear regression and neural network-based pedo-transfer functions.

Models	Soil parameters	RMSE	R^2
Linear regression			
	EC	1.985	0.73
	CEC	7.923	0.60
	ESP	1.130	0.66
	MWD	0.187	0.51
	Infiltration rate	9.446	0.57
ANN	EC	0.242	0.96
	CEC	1.250	0.90
	ESP	0.184	0.94
	MWD	0.038	0.84
	Infiltration rate	1.549	0.92

O.C content (r = 0.62) is rather similar to the between CEC and clay content (r = 0.69) . Also, the correlation coefficient between EC and Na^+ (r = 0.74) is rather more than between EC and Cl^- (r = 0.60). However with regarding to these correlation coefficients and due to their high amount, both of them are suitable for developing PTFs for EC prediction in soils of northern Iran. Similarly these correlations between ESP and SAR (r = 0.84) and also, between infiltration rate and soil porosity (r = 0.78) were positive and significant. Although, the correlation between infiltration rate and silt+fine content (r = 0.47) and among infiltration rate and MWD (r = 0.46) were relatively high, but we did not enter the silt+fine content and MWD for building our equations. Because, these correlation coefficients were less than 0.50 and for enhancing the model accuracy we did not use them for development of equations.

In addition with regarding to this table it is clear that MWD is negatively correlated with SAR (r = 0.93), ESP (r = 0.85), clay (r = 0.60) and silt+fine sand (r = 0.59) while, it is positively correlated with O.C content (r = 0.84). Hence with respecting to Table 3, multivariable linear regression equations were developed for those studied parameters which had high significant correlation with each other using Minitab and SAS software. From the numerous available PTFs derived to predict soil physico-chemical and hydrological properties we selected only those regression models that had a coefficient of determination, R^2, greater than 0.5 (Amini et al., 2005). These equations were expressed as:

EC = 0.401 Na^+ + 0.213 Cl^- + 0.468…. (4)
CEC = 2.03 O.C + 0.109 Clay + 12.66…. (5)
ESP = 0.917 SAR + 0.224…….. (6)
MWD = 0.098 O.C - 0.002 Clay - 0.015 (silt+fine sand) - 0.003 ESP + 0.727……. (7)
Final steady-state infiltration rate = 0.613 Porosity + 4.17…… (8)

After determining of these equations, performance of multivariate linear regression was developed for test data set and then, correlation coefficient and RMSE have been obtained for EC, CEC, ESP, MWD and infiltration rate. The results of coefficient of determination and RMSE values related to studied soil parameters for multivariable linear regression method are presented in Table 4. With respecting to these results it is obvious that linear regression equations can estimate EC (R^2 = 0.73) and ESP (R^2 = 0.66) with more accuracy than other soil parameters in Southern costal zones of Caspian Sea. Other soil parameters including CEC (R^2 = 0.60), infiltration rate (R^2 = 0.51) and MWD (R^2 = 0.51) are in the subsequent orders respectively.

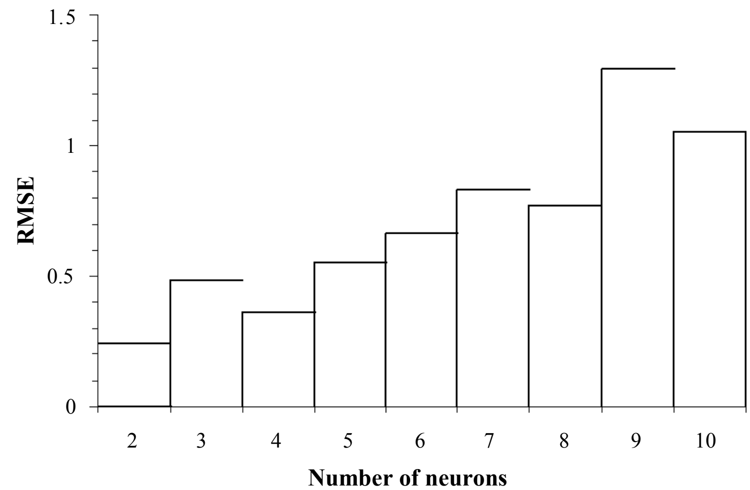

Figure 3. RMSE values for 2-10 neurons in hidden layer (electrical conductivity).

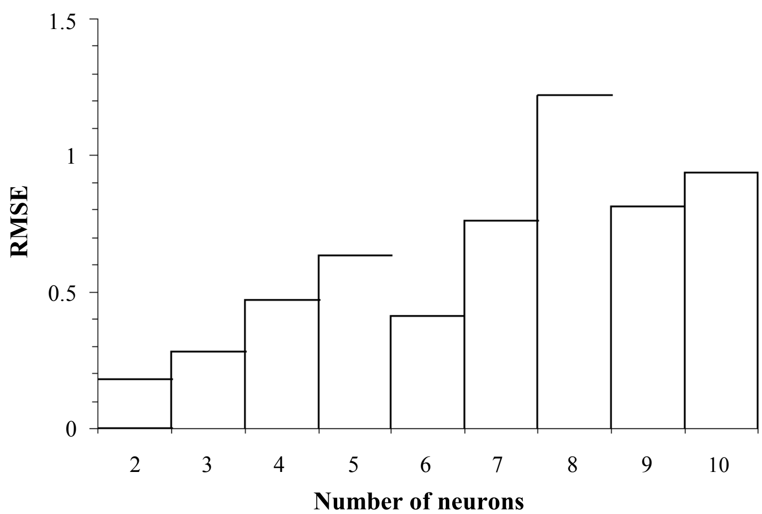

Figure 4. RMSE values for 2-10 neurons in hidden layer (exchangeable sodium percentage).

Neural network model

For predicting the soil physico-chemical and hydrological properties by means of artificial neural networks, the input data were similar to those used for multivariate linear regression. In the present study for predicting soil properties we did not increase the input date for constructing artificial neural network. Because according to findings of Amini et al. (2005) increasing the number of inputs will decrease the accuracy of the estimations. For example for predicting a soil characteristics if just one types of the input data have low correlation coefficients with output data, the accuracy of the model will automatically decrease. Therefore, in constructing of neural networks we used from those soil parameters that had the most correlation coefficient with each other (Table 3). The input data in these models were consisted of the percentages of clay and organic carbon for CEC, concentration of soluble Na^+ and Cl^- for EC, the level of SAR for ESP, percentage of organic carbon, clay, silt+fine sand and ESP level for MWD and percentage of total porosity for infiltration rate.

After determination the complexes of training and testing data, in the next step the various models of neural network having one hidden layer and 2 - 10 neurons in this layer were made. Then, the optimum structures of network by means of correlation coefficient and RMSE criteria were determined. The RMSE values for various numbers of neurons related to studied soil parameters are presented in the Figures 3, 4, 5, 6 and 7. As shown in this Figures 3 and 4, the minimum level of RMSE for EC and ESP are related to the network having two neurons

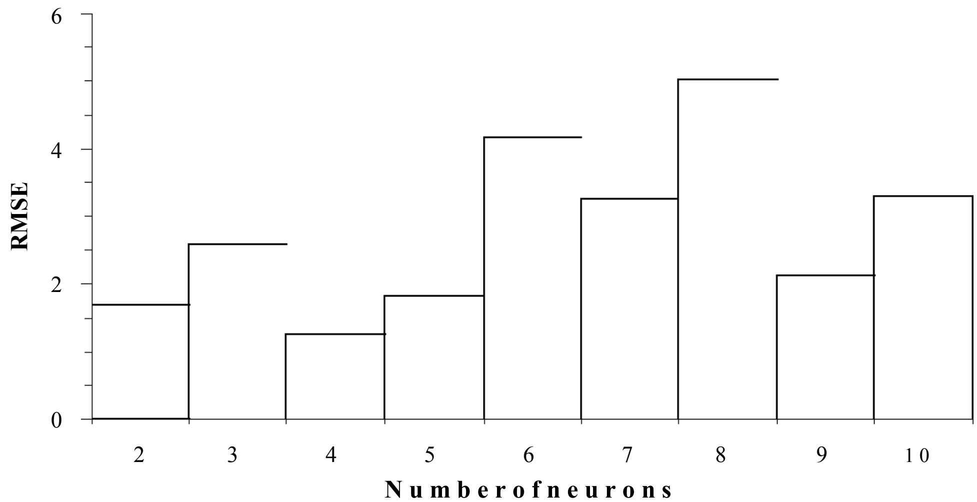

Figure 5. RMSE values for 2-10 neurons in hidden layer (cation exchange capacity).

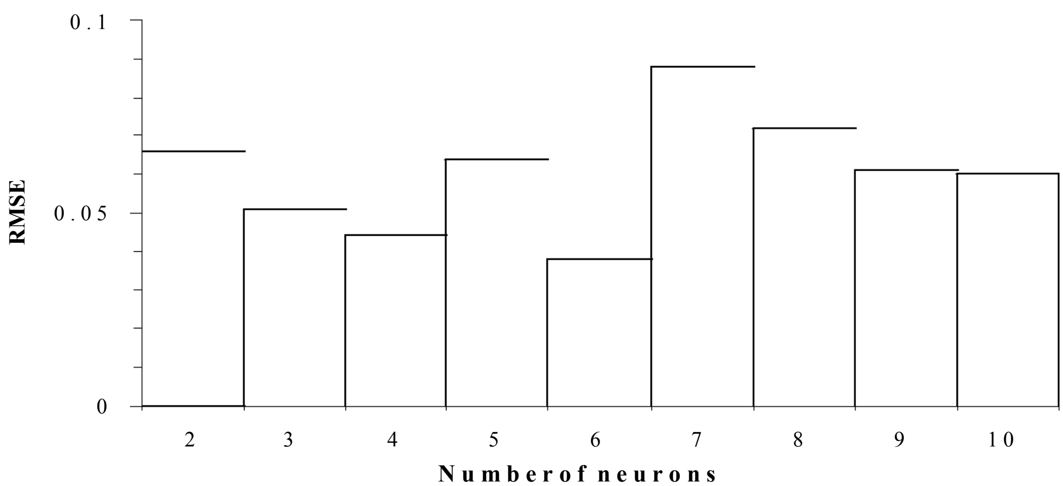

Figure 6. RMSE values for 2-10 neurons in hidden layer (mean weight diameter of aggregates).

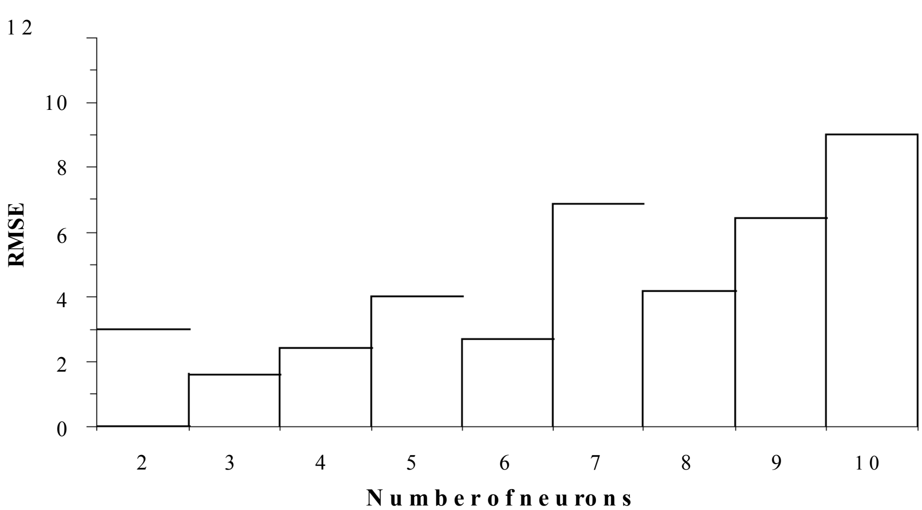

Figure 7. RMSE values for 2-10 neurons in hidden layer (infiltration rate).

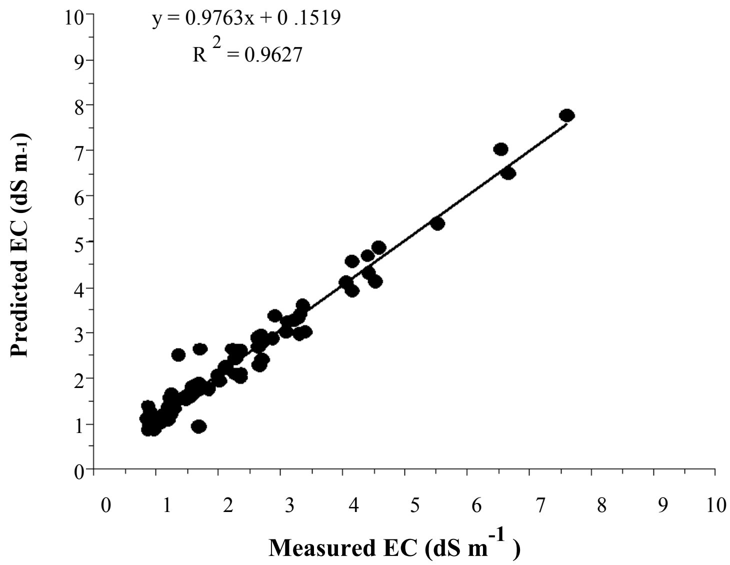

Figure 8. The scatter plot of the measured versus predicted EC.

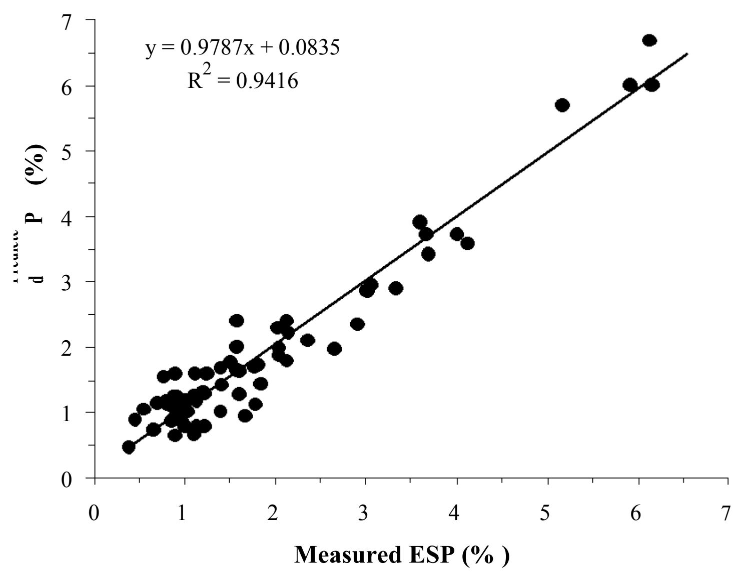

Figure 9. The scatter plot of the measured versus predicted ESP

in the hidden layer. Also, with regarding to these figures can be realize that with increasing the number of neurons, the efficiency of models will decrease and hence, the best efficiency is related to the networks having low numbers of neurons. However, the similar results were observed for other studied soil characteristics such as infiltration rate, CEC and MWD. So that as shown in the tables, the least levels of RMSE for infiltration rate, CEC and MWD were related to three, four and six neurons in hidden layer, respectively. The results of coefficient of determination and the least values of RMSE related to studied soil parameters for artificial neural network method are presented in Table 4. As shown in this table the levels of RMSE and R^2 for EC, CEC, ESP, MWD and final infiltration rate were 0.242,0.96, 1.250,0.90, 0.184,0.94, 0.038,0.84 and 1.549,0.92 respectively. In addition, on the basis of this able the levels of correlation coefficient and RMSE derived by artificial neural network for all studied soil parameters had higher values than those derived by multivariate linear regression. The scatter plots of the measured against predicted EC, ESP, CEC, MWD and infiltration rate for the test data set are given in Figures 8, 9, 10, 11 and 12 for the ANN model, which we identified as being the best model for predicting

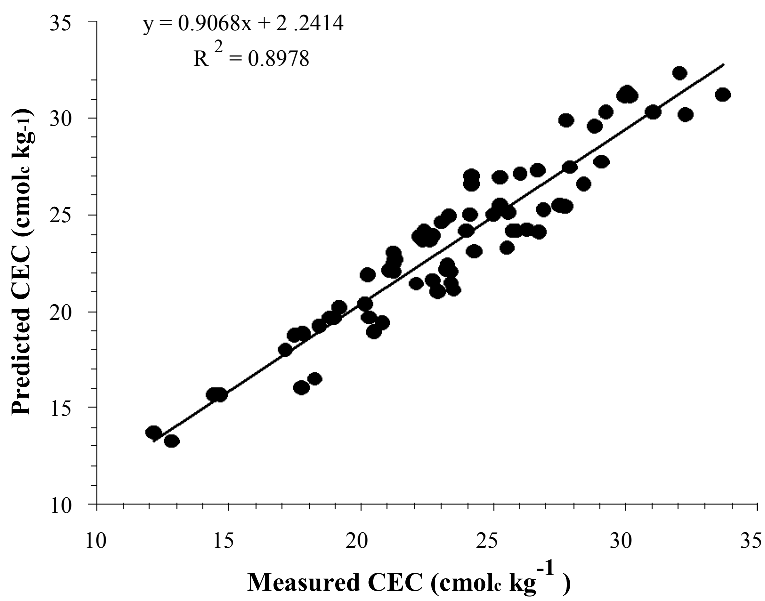

Figure 10. The scatter plot of the measured versus predicted CEC.

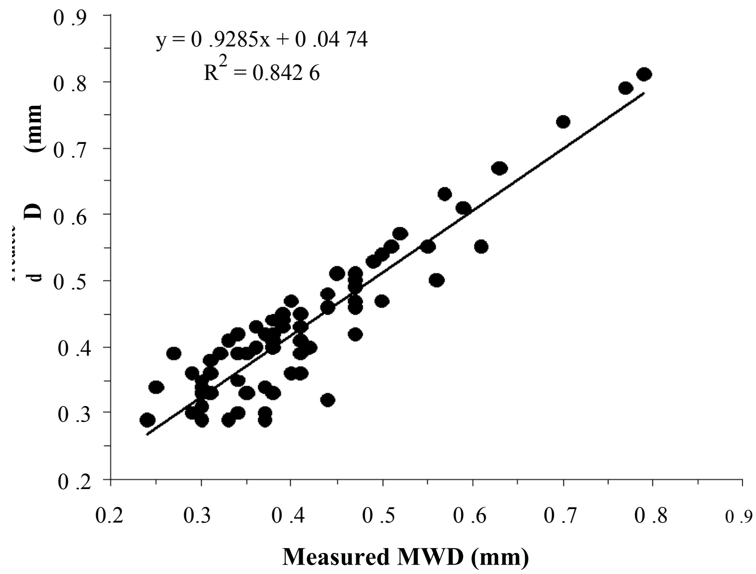

Figure 11. The scatter plot of the measured versus predicted MWD.

soil parameters. So that according to all of these diagrams, the best fitted line has the angle of near to 45° that shows the high accuracy of estimation by the neural network model.

DISCUSSION

The correlations between CEC and soil O.C (r = 0.62) and between CEC and clay content (r = 0.69) in the present study were both similar to respective values for aridisols of Isfahan in central Iran reported by Amini et al., (2005) (r = 0.65 for CEC and organic matter, and r = 0.66 for CEC and clay). In contrast, Seybold et al. (2005) believe that the prediction equations in aggregate provide a reasonable estimate of CEC for most soils of the United States. The results also indicated that ESP and SAR were positively correlated (r = 0.84) with each other and related linear equation for soils of northern Iran was ESP = 0.917 SAR + 0.224 with R^2 = 0.66. Seilsepour et al.

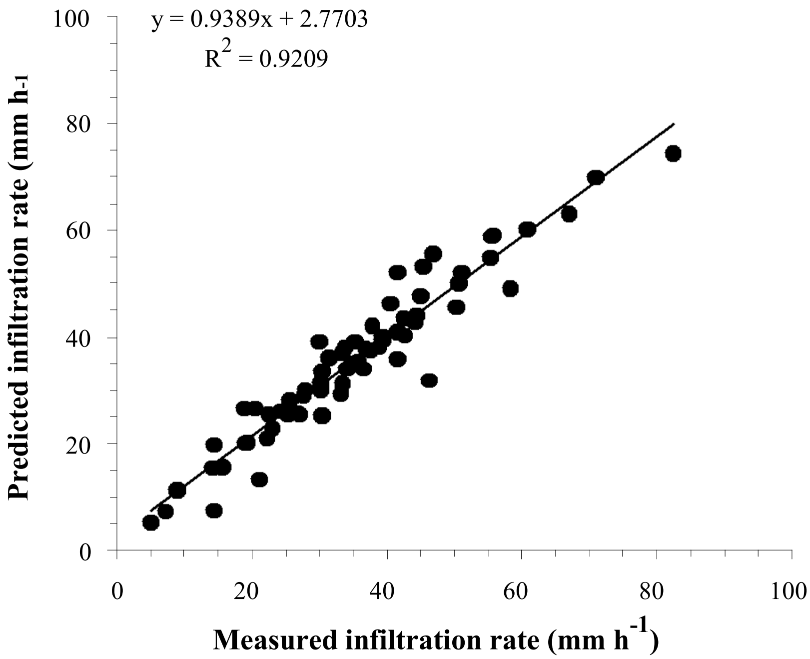

Figure 12. The scatter plot of the measured versus predicted infiltration rate.

(2009) recommended the similar model for predicting soil ESP in an arid region (Namely as Varamin region) of central Iran (ESP = 1.95 + 1.03 SAR with R^2 = 0.92).

The results also indicated that infiltration rate and porosity had high correlation (r = 0.78) with each other. Similarly, Ersahin (2003) reported that bulk density was significantly related to infiltration rate on an 8.5 ha alluvial field (loamy mesic Ustifluvent) located in Central Anatolia of Turkey. Hence with regarding to direct relationships among bulk density and soil porosity (equation 1), we can easily attribute the variations of infiltration rate to the porosity. Moreover, the high correlations between EC and soluble Na^+ (r = 0.74) and among EC and concentration of Cl^- (r = 0.60) in the present study were similar to findings reported by Taghizadeh Mehrjardi et al. (2008). They concluded from studding water samples provided from 625 wells in Azarbayjan province of Iran that the correlation coefficients between EC and Na^+ (r = 0.80) and between EC and Cl^- (r = 0.76) were significantly high. On the basis the results (Table 4), the comparing of RMSE and coefficient of determination related to all studied soil parameters for both of two mentioned types of pedo-transfer functions revealed that artificial neural network had better performance in predicting all soil properties (EC, ESP, CEC, MWD and final steady-state infiltration rate) than multivariate regression which is in line with the work done by Amini et al. (2005); Minasny and McBratney (2002); Najafi and ghivi (2006) and Schaap et al. (1998). The reason of this superior efficiency of ANNs models compare with the basic regression equations is probably because; the PTFs that have derived from various areas have the different efficiencies. On the other hand, according to the hypothesis of Schaap et al. (1998), for designing of a neural network we do not need to a special equation. They also believe that with creation a suitable equation between input and output data we are able to achieve to the best results. Also, due to the occurring of nonlinear equations between dependent variables and predicting variables, the neural network have the better efficiency compared with the basic regression equations. Koekkoek and Booltink (1999) found that ANN performed slightly better, but the differences were not significant. They also reported that network models for three parameters were more suitable for capturing the non-linearity of the relationship between variables. Pachepsky et al. (1996) investigated the accuracy of artificial neural network and analyzed the regression method using correlation co-efficient and the root of mean square error. They reported that the neural network is able to predict the easily measurable soil parameters with more accuracy and less error. The similar results have reported by the Tamari et al. (1996) as well. They found that using artificial neural network leads to less RMSE values than the multivariable linear regression. They also reported that the neural network has not better efficiency than linear regression models in occasion of high stability of data. However, the high accuracy of data leads to more efficiency of neural network and also, shows the proper selection of testing and training data. Analysis of the ANN parameters suggested that more input variables were necessary to improve the prediction of soil parameters (Tamari et al., 1996; Mermoud and Xu, 2006).

Conclusion

In this study the values of some soil physico-chemical and hydrological properties were estimated using models of ANNs and linear regression. A systematic approach is therefore presented to acquire and verify the stored knowledge of a general ANN (Perceptron neural network) based constitutive soil model. This network was consisted of one hidden layer, a sigmoid activation function in hidden layer, and a linear activation function in output layer. Sensitivities of the output to corresponding inputs are defined mathematically. A sensitivity analysis is then performed to extract the dominant rules of the proposed model, which compare favorably with experimental observations. For predicting the soil properties by means of PTFs, the input data were consisted of the percentages of clay and organic carbon for CEC, concentration of soluble Na^+ and Cl^- for EC, the level of SAR for ESP, percentage of organic carbon, clay, silt + fine sand and ESP level for MWD and percentage of total porosity for infiltration rate. With regarding to the evaluation criteria, the results of this study revealed that the artificial neural networks had superiority to the basic regression equations for prediction of mentioned soil parameters. This is a crucial result because, since ANN–PTFs formed from local data produce more accurate predictions than those built from data spread from a wider area, the concept of data conservation becomes a critical factor in ANN–PTF construction (Baker and Ellison, 2008). However, due to difficulties of direct measurement of soil parameters, we recommend using of neuro-fuzzy models in the future studies for obtaining the logical equations of other soil parameters, especially soil hydraulic properties, in each area.

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. Eur. J. Soil Sci. 53: 748-757.

Baker L, Ellison D (2008). Optimisation of pedotransfer functions using an artificial neural network ensemble method. Geoderma. 144: 212–224.

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.

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

Blake GR, Hartge KH (1986). Particle density. p. 377–382. In A. Klute (ed.) Methods of soil analysis. Part 1. Agron. Monogr. 9. ASA, Madison, WI.

Day PR (1965). Particle fractionation and particle-size analysis. In: Black, C.A. (Ed.), Methods of soil analysis: Part 2. Monogr. Ser., vol. 9. American Society of Agronomy, Madison WI, pp. 545-566.

Doaei M, Shabanpour M, Bagheri F, Navvabian M (2005). Modeling of hydraulic conductivity using artificial neural networks in soil of Guilan province. Guilan University MSc. Thesis. Rasht 128 pp. (in Persian)

Ghorbani-Dashtaki SH, Homaei M (2002). Parametric assessment of hydraulic functions of soil unsaturated zones using pedo-tranfer functions. J. Agric. Engin. Invest. 3(12): 3-15. (in Persian)

Heusher SA, Brandt CC, Jardin PM (2005). Using soil physical and chemical properties to estimate bulk density. Soil Sci. Soc. Am. J. 69: 51-56.

Jain A, Kumar A (2006). An evaluation of artificial neural network technique for the determination of infiltration model parameters. Appl. Soft Comput. 6: 272–282.

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

Kaur R, Kumar S, Gurung HP (2002). A pedotransfer function soil data and its comparison with existing PTFs. Aust. J. Soil Res. 40: 847-857.

Kittrick JA, Hope EW (1963). A procedure for the particle size separation of soils for X-ray diffraction analysis. Soil Sci. 96: 312-325.

Klute A (1986). Methods of soil analysis. Part 1: Physical and mi - neralogical methods. Soil sci. Soc. Am. J. Madison, Wisconsin, USA.

Koekkoek EJW, Booltink H (1999). Neural network models to predict soil water retention. Eur. J. Soil Sci. 50: 489-495.

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 Till. Res. 90: 108–116.

Mermoud A, Xu D (2006). Comparative analysis of three methods to generate soil hydraulic functions. Soil Till. Res. 87: 89–100.

Minasny B, McBratney AB (2002). The neuro-m methods for fitting neural network parametric pedotransfer functions. Soil Sci. Soc. Am. J. 66: 352–361.

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.

Najafi M, Givi J (2006). Evaluation of prediction of bulk density by artificial neural network and PTFs. 10th Iranian Soil Science Congress. Karaj. 680-681. (in Persian)

Nelson DW, Sommers LE (1982). Total carbon, organic carbon, and organic matter. In: Page, AL, Miller, RH, Keeney DR (Eds.), Methods of Soil Analysis. Part II, 2nd ed. American Society of Agronomy, Madison, WI, USA, pp. 539-580.

Pachepsky YA, Timlin D, Varallyay G (1996). Artificial neural networks to estimate soil water retention from easily measurable data. Soil Sci. Soc. Am. J. 60: 727–733.

Page, AL (1986). Methods of soil analysis. Part 2: Chemical and micro-biological properties. Soil Sci. Soc. Am. J. Madison, Wisconsin, USA.

Sarmadian F, Taghizadeh Mehrjardi R, Akbarzadeh A (2009). Modeling of some soil properties using artificial neural network and multivariate regression in Gorgan province, north of Iran. Australian J. of Basic and Applied Sci. 3(1): 323-329.

Schaap MG, Leij FJ (1998). Using neural networks to predict soil water retention and soil hydraulic conductivity. Soil Till. Res. 47: 37–42.

Schaap MG, Leij FJ, van Genuchten MTh (1998). Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J. 62: 847–855.

Seilsepour M, Rashidi M, Ghareei Khabbaz B (2009). Prediction of soil exchangeable sodium percentage based on soil sodium adsorption ratio. American-Eurasian J. Agric. Environ. Sci. 5(1): 01-04.

Seybold CA, Grossman RB, Reinsch TG (2005). Predicting cation exchange capacity for soil survey using linear models. Soil Sci. Soc. Am. J. 69: 856–863.

Taghizadeh Mehrjardi A, Akbarzadeh A, Mahmoodi Sh, Heidari A, Sarmadian F (2008). Application of geostastical methods for mapping groundwater quality in Azarbayjan province, Iran. American-Eurasian J. Agric. Environ. Sci. 3(5): 726-735.

Tamari S, Wosten JHM, Ruiz-Suarez JC (1996). Testing an artificial neural network for predicting soil hydraulic conductivity. Soil Sci. Soc. Am. J. 60: 1732-1741.

Vos BD, Meirvenne MV, Quataert P, Deckers J, Muys B (2005). Predictive quality of pedotransfer functions for estimating bulk density

of forest soils. Soil Sci. Soc. Am. J. 69: 500–510.
Wagner B, Tarnawski VR, Hennings V, Muller U, Wessolek G, Plagge R (2001). Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set. Geoderma. 102: 275–279.