Sjournals
Scientific Journals
Agricultural Advances (2014) 3(4) 111-123
ISSN 2251-7820
doi: 10.14196/aa.v3i4.1267

Contents lists available at Sjournals
Agricultural Advances
Journal homepage: www.Sjournals.com

Original article

Mathematical model of soil cation exchange capacity using GMDH-type neural network and genetic algorithm

A. Bazrafshan[a,*], M. Shabanpour[a], M. Norouzi[a], S. Fallahi[b]
[a]*Department of Soil Science, Faculty of Agriculture, University of Guilan, Rasht, Iran.*
[b]*Department of Mathematics, Faculty of Sciences, University of Guilan, Rasht, Iran.*

*Corresponding author; Department of Soil Science, Faculty of Agriculture, University of Guilan, Rasht, Iran.

ARTICLE INFO

Article history,
Received 13 March 2014
Accepted 14 April 2014
Available online April 2014

Keywords,
Clay
GMDH model
Multivariate regression
Pedotransfer functions

ABSTRACT

Measuring the cation exchange capacity (CEC) for all horizons of every map unit component in a survey area is very time consuming and costly. This study was conducted (i) to evaluate the group method of data handling (GMDH) neural network (NN) and genetic algorithm model and (ii) to compare GMDH-type NN with other artificial neural networks such as the multilayer perceptron (MLP), radial basis function (RBF) and regression-based models for predicting CEC in soils of Lahijan, north of Iran. In this study, the proposed model was trained before requested predictions. The data set was divided into two parts: 70% were used as data for training (110 soil samples), and 30% (40 soil samples) were used as a test set, which were randomly extracted from the database. In order to evaluate the models, coefficient of determination (R^2), mean square error (MSE), root mean square error (RMSE) and mean absolute deviation (MAD) were used. Results showed that the GMDH-type and MLP-NN models had larger R^2 values than the multiple regression and RBF models. The results of GMDH model were very encouraging and congruent with the experimental results. In general, the GMDH-type-NNs models provided more reliable predictions than the artificial neural networks (ANNs) and regression-based models.

1. Introduction

CEC is the amount of negative charge in soil that is available to bind positively charged ions (cations). Cation exchange capacity is used as a measure of fertility, nutrient retention capacity and the capacity to protect groundwater from cation contamination (Sarmadian and Taghizadeh Mehrjardi, 2008).

Although CEC can be measured directly, its measurement is especially difficult and expensive in the soils. Pedotransfer functions (PTFs) provide an alternative by estimating CEC from more readily available soil data.in recent years, various PTFs have been developed to estimate CEC from basic physical and chemical soil properties (Amini et al., 2005; Seybold et al., 2005 and Sarmadian and Taghizadeh Mehrjardi, 2008). Multiple linear regression analysis is generally used to find the relevant coefficients in the model equations. However, models developed for one region may not give adequate estimates for a different region (Wagner et al., 2001). Several researchers have attempted to predict CEC from clay and organic carbon contents alone, using multiple regression (Drake and Motto, 1982; Sahrawat, 1983; Bell and Van Keulen, 1995). A neural network is an attempt to build a mathematical model that supposedly works in an analogous way to the 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 (Minasny and McBratney, 2002). A NN is an adaptable non-linear data transfer structure that can learn the relations between input and output data while being insensitive to measurement noise (Hecht-Nielsen, 1990).

The group method of data handling (GMDH) was used to develop PTFs in some previous soil studies (Pachepsky and Rawls, 1999). Grouping improved the accuracy of PTFs in most cases, probably because of similarities in PTF relations within groups (Pachepsky and Rawls, 1999). The GMDH is aimed at identifying the functional structure of a model hidden in the empirical data (Ivakhnenko, 1971). The main idea of the GMDH is the use of feed-forward networks based on short-term polynomial transfer functions whose coefficients are obtained using regression combined with emulation of the self-organizing activity behind NN structural learning (Farlow, 1984). The GMDH was developed in complex systems for the modeling, predicting, identifying, and approximating. It has been shown that, the GMDH is the best optimal simplified model for inaccurate, noisy, or small data sets, with a higher accuracy and a simpler structure than typical full physical models (Ghanadzadeh et al., 2012). To avoid the limitations of ANNs, a vapor–liquid equilibrium (VLE) prediction method was developed using the GMDH algorithm (Nariman-Zadeh, 2007; Ketabchi et al., 2010).

The general purpose of this research is predicting of CEC using the GMDH model and genetic algorithm. This model was compared with other ANNs such as the MLP and RBF networks and regression-based models.

2. Materials and methods

2.1. Study area

This study was conducted in Guilan province in north of Iran, as a part of Gorgan-Rasht tectonic zone. The geology of the area is still poorly known, because of its location in rainy forest and dense topography. Geologically, the region is composed of Quaternary Caspian deposits and mainly of Jurassic and Cretaceous volcanic rocks (Anells et al., 1975). The study area located between 37° 5′ to 37° 15′ northern latitude and 50° 0′ to 50° 10′ eastern longitude (Figure 1).

2.2. Soil sampling and laboratory analysis

In the present study, 150 soil samples were collected from different horizons of 48 soil profiles located in Lahijan, Guilan Province, in the north of Iran. Soil textural distribution is presented in Figure (2). After removing large stones, the soil samples were air-dried and passed through a 2 mm (10 mesh) sieve. Particle size distribution was determined by the hydrometer method (Gee and Or, 2002). Soil pH was measured in 0.01 M $CaCl_2$ in 1:2 soil: solution (Thomas, 1996). Organic carbone (OC) was determinate according to the Walkley–Black wet oxidation method (Nelson and Sommers, 1996). CEC was determined by saturation with 1 M ammonium acetate (NH4OAc) at pH=7.0 (Sumner & Miller, 1996).

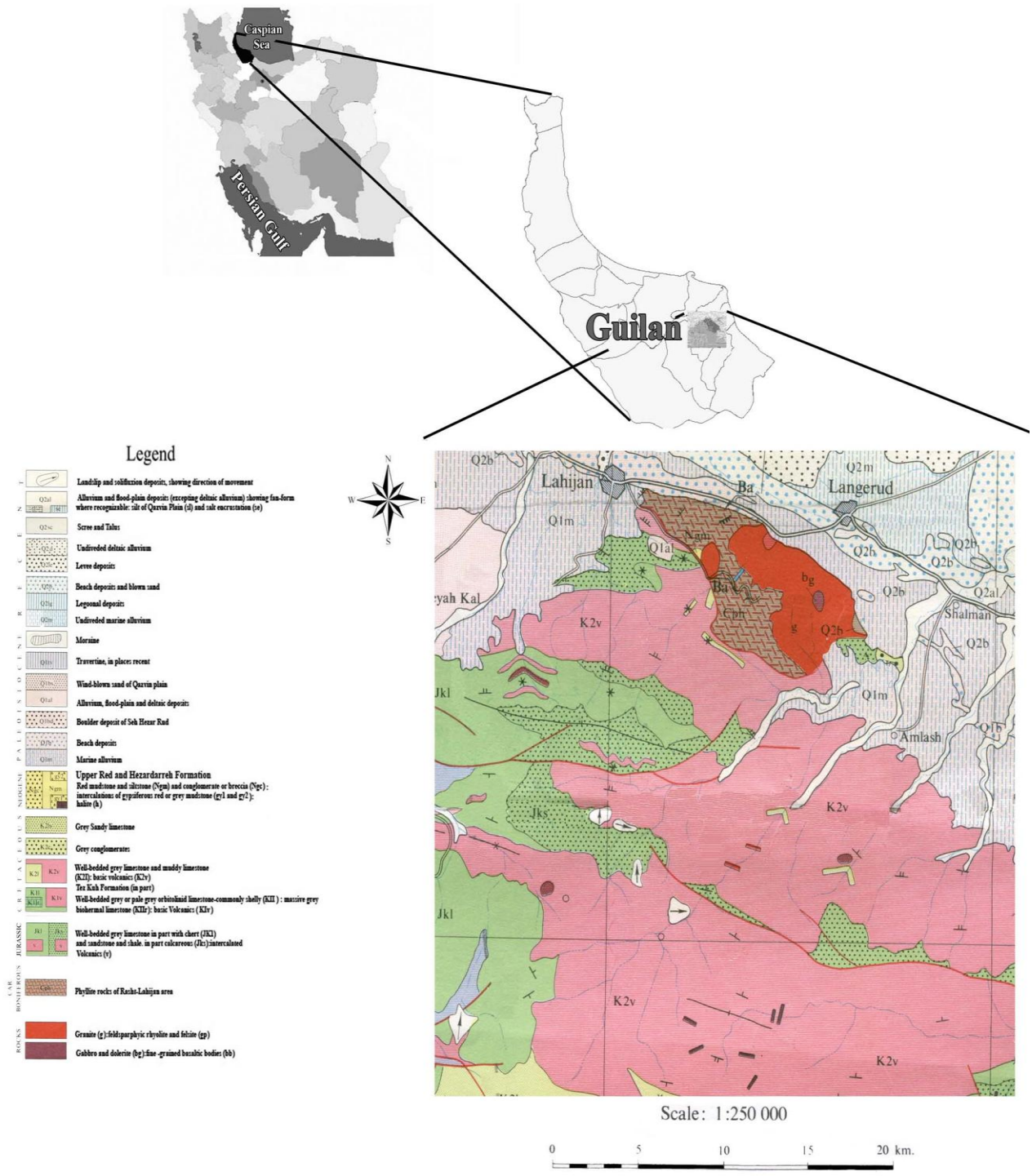

Fig. 1. Position of sampling area in Lahijan, Guilan province.

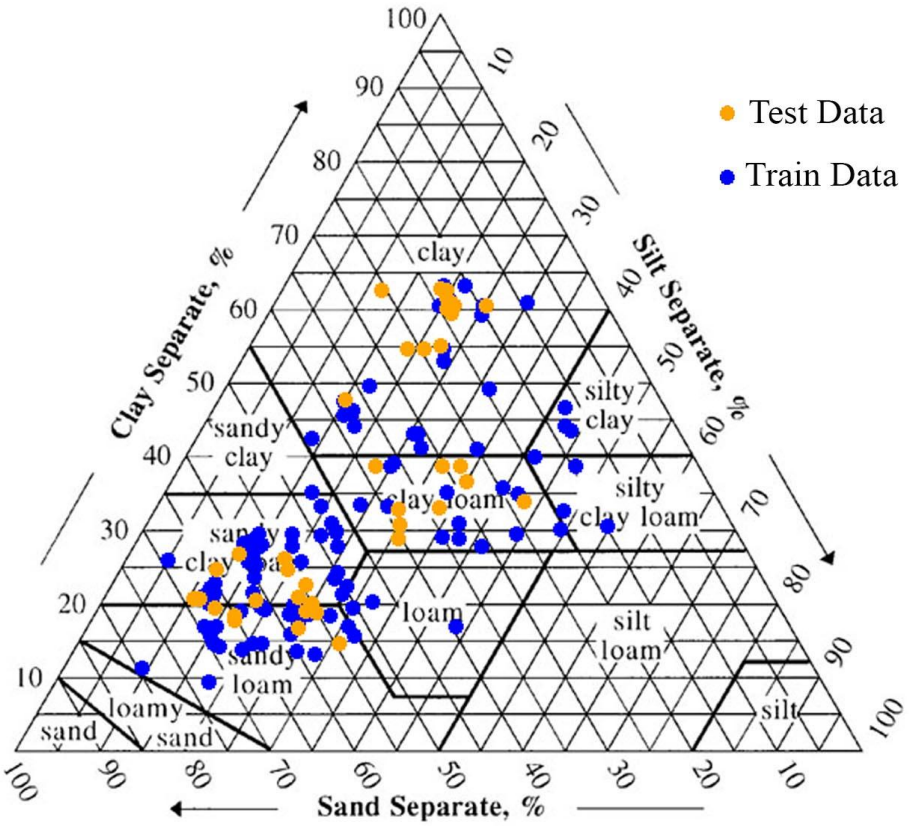

Fig. 2. Particle-size distribution of the 150 soil samples (train and test data).

2.3. Multivariate regression

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

Y = aX1 + bX2 + cX3 + ... (1)

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

2.4. Artificial neural network

An ANN, by means of its architecture, attempts to simulate the biological structure of the human brain and nervous system (Erzin et al., 2008). Artificial neural networks are the preferred tool for many predictive data mining applications because of their power, flexibility, and ease of use. ANNs used in predictive applications, such as the multilayer perceptron (MLP) and radial basis function (RBF) networks, are supervised in the sense that the model-predicted results can be compared against known values of the target variables (Ripley,1996).

2.5. Group method of data handling (GMDH)

Using the GMDH algorithm, a model can be represented as a set of neurons in which different pairs of them in each layer are connected through a quadratic polynomial and, therefore, produce new neurons in the next layer. Such representation can be used in modeling to map inputs to outputs. The formal definition of the identification of problem is to find a function, $\hat{f}$, that can be approximately used instead of the actual one, f, in order to predict output $\hat{y}$ for a given input vector $X=(x_1,x_2,x_3,\cdots,x_n)$ as close as possible to its actual output y. Therefore, given number of observations (M) of multi-input, single output data pairs so that

$$y_i = f(x_{i1},x_{i2},x_{i3},\cdots,x_{in})(i=1,2,3,\cdots,M) \quad (2)$$

It is now possible to train a GMDH-type-NN to predict the output values $\hat{y}_i$ for any given input vector $X=(x_{i1},x_{i2},x_{i3},\cdots,x_{in})$, that is

$$\hat{y}_i=\hat{f}(x_{i1},x_{i2},x_{i3},\cdots,x_{in})(i=1,2,3,\cdots,M) \quad (3)$$

In order to determine a GMDH type-NN, the square of the differences between the actual output and the predicted one is minimized, that is

$$\sum_{i=1}^{M}\left[\hat{f}(x_{i1},x_{i2},\ldots,x_i)-y_i\right]^2\rightarrow\min \quad (4)$$

The general connection between the inputs and the output variables can be expressed by a complicated discrete form of the Volterra functional series (Ivakhnenko, 1971) in the form of

$$y=a_o+\sum_{i=1}^{n}a_ix_i+\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j+\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}a_{ijk}x_ix_jx_k+\cdots \quad (5)$$

Where is known as the Kolmogorov-Gabor polynomial (Ivakhnenko, 1971). The general form of mathematical description can be represented by a system of partial quadratic polynomials consisting of only two variables (neurons) in the form of equation (6).

$$\hat{y}=G(x_i,x_j)=a_o+a_1x_i+a_2x_j+a_3x_ix_j+a_4x_i^4+a_5x_j^2\cdots \quad (6)$$

In this way, such partial quadratic description is recursively used in a network of connected neurons to build the general mathematical relation of the inputs and output variables given in equation (4). The coefficients a_i in equation (5) are calculated using regression techniques. It can be seen that a tree of polynomials is constructed using the quadratic form given in equation (5). In this way, the coefficients of each quadratic function G_i are obtained to fit optimally the output in the whole set of input–output data pairs, that is

$$E=\frac{\sum_{i=1}^{M}(y_i-G_i)^2}{M}\rightarrow\min \quad (7)$$

In the basic form of the GMDH algorithm, all the possibilities of two independent variables out of the total n input variables are taken in order to construct the regression polynomial in the form of equation (5) that best fits the dependent observations $(y_i,i=1,2,\ldots,M)$ in a least squares sense. Using the quadratic sub-expression in the form of equation (5) for each row of M data triples, the following matrix equation can be readily obtained as

$$Aa=Y \quad (8)$$

Where a is the vector of unknown coefficients of the quadratic polynomial in equation (5).

$$a=\{a_o,a_1,a_2,a_3,a_4,a_5\} \quad (9)$$

and

$$Y=\{y_1,y_2,y_3\ldots,y_M\}^T \quad (10)$$

Here Y is the vector of the output's value from observation. It can be readily seen that

$$A=\begin{bmatrix}1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2\\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & x_{Mp}^2 & x_{Mq}^2\end{bmatrix} \quad (11)$$

The least squares technique from multiple regression analysis leads to the solution of the normal equations in the form of

$$a=(A^TA)^{-1}A^TY \quad (12)$$

The multivariate regression (for PTF model), MLP, RBF and feed-forward GMDH-type NN for CEC was constructed using an experimental data set. A random sample selection of database is given in Table 1. The data was divided into two parts: 70% (110 points) was used as training data, and 30% (40 points) was used as test data.

Table 1
Statistics of the training and test data sets.

	pH	OC	Clay	Silt	Sand	CEC
Train Data (110)						
Mean	4.22	1.48	29.31	23.10	47.36	20.70
Standard Error	0.043	0.140	1.265	0.905	1.686	0.942
Standard Deviation	0.447	1.473	13.272	9.492	17.680	9.883
Variance	0.200	2.171	176.144	90.102	312.569	97.669
Skewness	0.519	1.756	0.874	1.139	-0.485	1.306
Kurtosis	0.609	3.440	0.040	0.919	-0.903	1.584
Minimum	3.3	0.04	8.88	8.72	10.24	7
Maximum	5.6	6.7	62.72	54.1	80.56	52
Test Data (40)						
Mean	4.07	1.12	36.70	21.29	42.02	22.95
Standard Error	0.060	0.201	2.748	1.188	2.996	1.360
Standard Deviation	0.379	1.270	17.378	7.514	18.949	8.604
Variance	0.144	1.612	301.999	56.458	359.077	74.028
Skewness	1.652	1.409	0.362	0.597	0.105	0.232
Kurtosis	3.394	0.732	-1.537	0.2698	-1.650	-1.419
Minimum	3.58	0.11	14.1	10.1	15.3	10.8
Maximum	5.2	4.6	62.3	42.5	69.8	40.4

2.6. Assessing models accuracy and reliability

In order to investigate the reliability of the proposed method, the accuracy of the model was determined using coefficient of determination (R^2), mean square error (MSE), root mean square error (RMSE) and mean absolute deviation (MAD) which are defined as follows:

$$R^2 = \frac{\left[\mathrm{cov}(Y_{i(actual)}, Y_{i(\mathrm{mod}el)})\right]^2}{\mathrm{var}(Y_{i(actual)}).\mathrm{var}(Y_{i(\mathrm{mod}el)})} \quad (13)$$

$$RMSE = \left[\frac{\sum_{i=0}^{M}\left(Y_{i(\mathrm{mod}el)} - Y_{i(acutual)}\right)^2}{M}\right]^{1/2} \quad (14)$$

$$MSE = \frac{\sum_{i=0}^{M}\left(Y_{i(\mathrm{mod}el)} - Y_{i(acutual)}\right)^2}{M} \quad (15)$$

$$MAD = \frac{\sum_{i=0}^{M}\left|Y_{i(\mathrm{mod}el)} - Y_{i(acutual)}\right|}{M} \quad (16)$$

3. Results and discussion

In the present study, after the learning and training with the models, the output (CEC) is generated. The statistics of the complete dataset (train and test data) are given in Table 1. The CEC content was predicted using multivariate regression (for PTF model), MLP, RBF and feed-forward GMDH-type NN. In this way, genetic algorithm (GA) is arranged in a new approach to design the whole architecture of the GMDH-type-NNs. It provides the optimal number of neurons in each hidden layer and their connectivity configuration to find the optimal set of appropriate coefficients of quadratic expressions to model CEC. Scatter plot matrix between variables (studied soil

properties) are given in Figure (3). As expected, the correlations between the sand content and other soil properties were all negative.

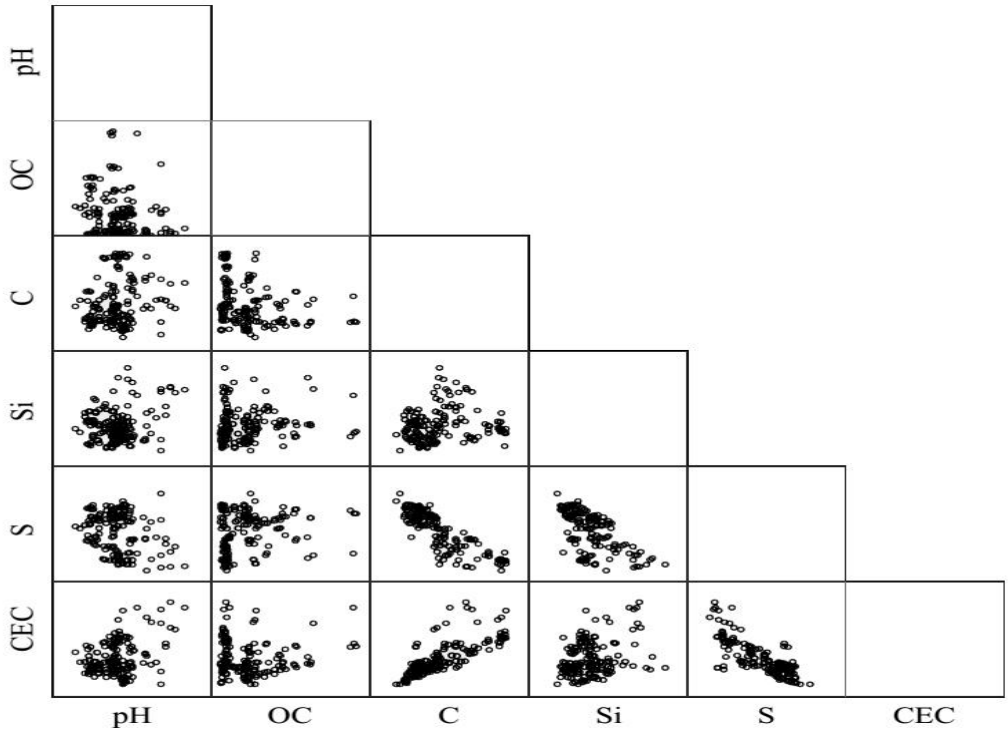

Fig. 3. Scatter plot matrix of studied soil properties.

An ANN, by means of its architecture, attempts to simulate the biological structure of the human brain and nervous system (Hagan, et al., 2002). The structure and architecture of MLP, RBF and GMDH-type-NN models are presented in Figure (4). The architecture tab is used to specify the structure of the network (Ripley, 1996). This pattern has proved to be useful when modelling input-output relations. Automatic architecture selection builds a network with one hidden layer. Automatic architecture selection uses the default activation functions for the hidden and output layers (Haykin, 1998). Specify the minimum and maximum number of units allowed in the hidden layer, and the automatic architecture selection in order to compute the best number of units in the hidden layer. The usage of a number of hidden layers in the ANN depends on the degree of complexity in the pattern recognition problem, and one or two hidden layers are found to be quite useful for most problems (Sonmez et al., 2005). The best structure in GMDH were reached with two hidden layers with 150 generations, cross over probability of 0.9 and mutation probability of 0.1, to model CEC. The developed GMDH NN was successfully used to obtain a model for calculating CEC (Table 2). The optimal structures of the developed NN with 2-hidden layers are shown in Figure (4). The “cbbcaedd” are corresponding genome representations for the CEC, in which a, b, c, d and e stand for pH, Organic carbon (OC), Clay, Silt and Sand respectively. All input variables were accepted by the models.

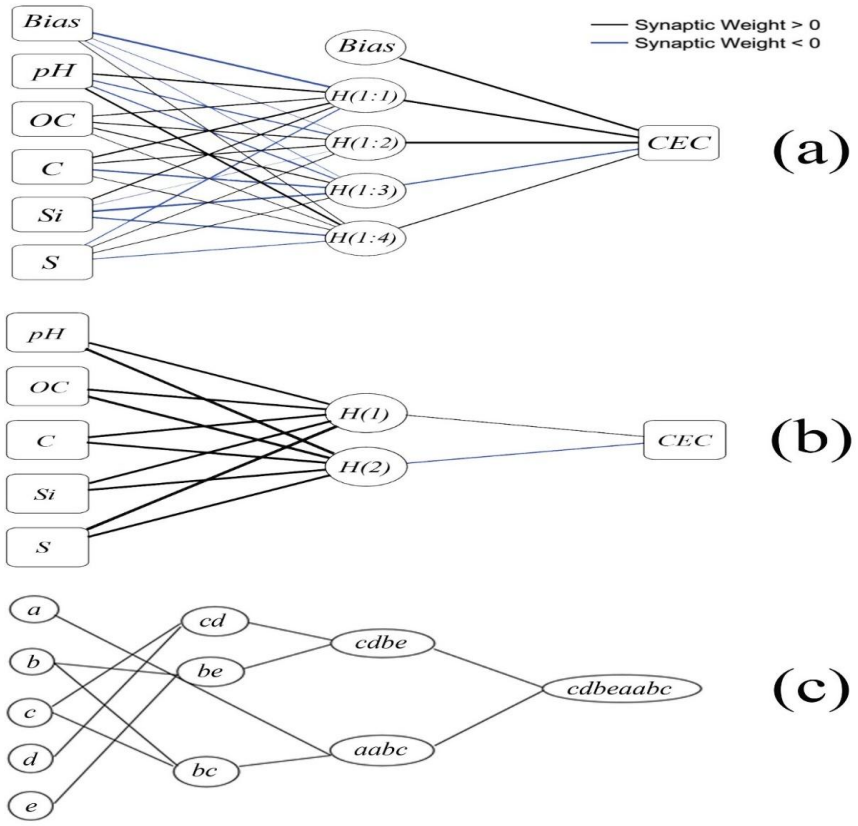

Fig. 4. The structure of (a) multilayer perceptron (MLP) neural network, (b) radial basis function (RBF) neural network and (c) developed structure of GMDH-type-NN model for the ternary system, a, b, c, d and e stand for pH, Organic carbon (OC), Clay, Silt and Sand respectively.

When using GMDH- NN and regression models to predict the CEC, the relations between characteristics need to be described by well-defined equations. Multiple linear regressions equation for PTF model and polynomial equations of the GMDH model for CEC prediction were presented in Table 2. It can be shown that these models are more reliable in comparison with ANN (MLP and RBF) models. The GMDH-type-NN provides an automated selection of essential input variables, and builds polynomial equations for the modeling. This polynomial equation shows the quantitative relationship between input and output variables (Table 2). Our proposed models behavior in prediction of CEC demonstrated in Figure (5). This Figure shows plot of the experimental data and GMDH-NN model. This diagram demonstrates that the predicted values are close to the experimental values, as many of the data points fall very close to the diagonal line.

The scatter plot of the measured against predicted CEC for the train and test data set is given in Figure (6) for the whole models. As depicted in Figure (6) RBF NN model cannot predict greater contents and most contents tended to CEC below 30 $Cmol_{(+)}\ Kg^{-1}$. The MLP NN model shows slightly better results as comparing to GMDH model in train data, but GMDH model had the best results in test data. The statistical results of the comparisons are given in Table 3, which shows that the GMDH-type and MLP-NN models had larger R^2 values than the multiple regression and RBF models. Furthermore, these models had lowest RMSE, MSE and MAD indexes in comparison with multiple regression and RBF models. The absolute value of MD should always be small, and indicate a high model performance and are a measure of the overall error of the estimation (Zacharias, 2007). Amini et al. (2005) reported that in prediction of CEC, the neural network-based models provided more reliable predictions than the regression-based PTFs.

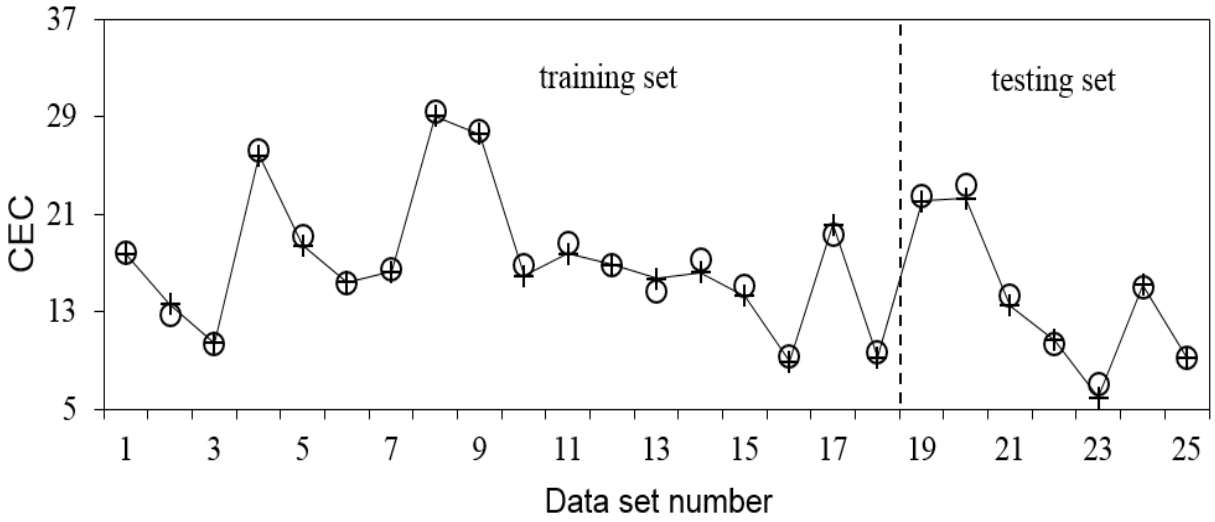

Fig. 5. Plot of CEC against data set number to illustrate the prediction of the experimental data using the GMDH model; (○) experimental points; (+) calculated points.

Table 2
Pedotransfer function (PTF) and polynomial equations of the GMDH model for Cation exchange capacity (CEC) in this study.

Model	Equation
PTF	CEC = -8.988-0.15 (Sand) + 0.408 (Clay) + 2.143 (OC) + 5.131 (pH)
GMDH	CEC = -2.8699 + 0.9501 Y_4 + 0.2811 Y_5 - 0.1014 Y_4^2 - 0.0770 Y_5^2 + 0.1772 Y_4 Y_5
Y1 = 8.5426 + 0.1779 (Clay) – 0.1291 (Silt) - 0.0016 $(Clay)^2$ - 0.0034 $(Silt)^2$ + 0.0199 (Clay) (Silt)	
Y2 = 46.1905 + 0.3956 (OC) - 0.7762 (Sand) + 0.4278 $(OC)^2$ + 0.0041 $(Sand)^2$ - 0.0195 (OC) (Sand)	
Y3 = 1.7759 - 2.8270 (OC) + 0.7143 (Clay) + 0.4937 $(OC)^2$ - 0.0034 $(Clay)^2$ + 0.0975 (OC) (Clay)	
Y4 = 1.0723 + 0.4940Y_1 + 0.3398 Y^2 + 0.0296 Y_1^2 + 0.0282 Y_2^2 - 0.0538Y_1 Y_2	
Y5 = 160.6425 - 71.2619 Y_3 - 0.8948 (pH) +7.9456 Y_3^2 + 0.0046 $(pH)^2$ + 0.3895 Y_3 (pH)	

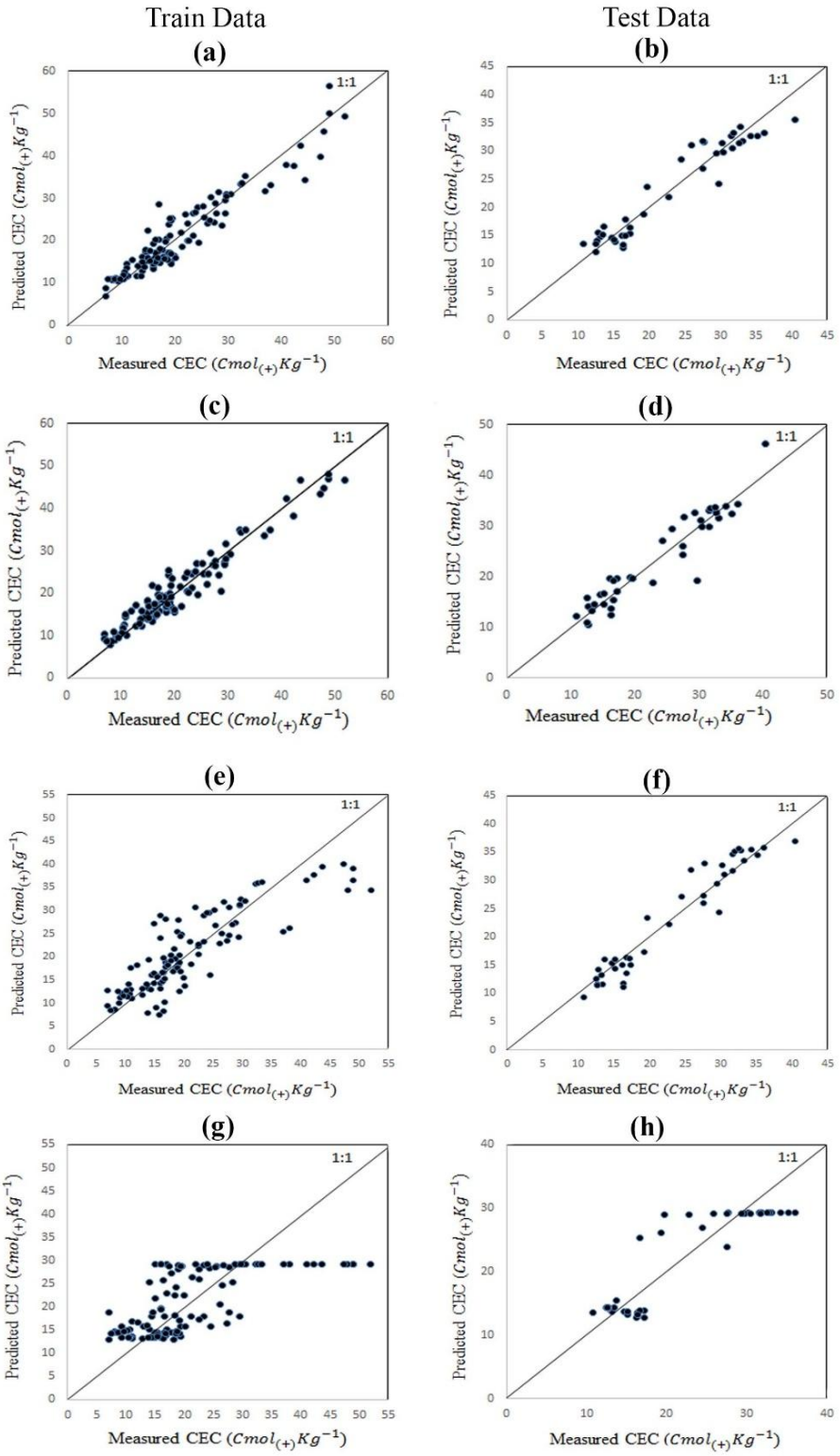

Fig. 6. Performance of models for the train and test dataset: measured vs. predicted CEC ($Cmol_{(+)}\ Kg^{-1}$) with reference to the 1:1 line. (a) and (b) for GMDH neural network model, (c) and (d) for MLP neural network, (e) and (f) for Pedotransfer function (PTF) model, (g) and (h) for RBF neural network model.

Table 3
Evaluation indices for the Group Method of Data Handling (GMDH), multilayer perceptron (MLP), Pedotransfer function (PTF) and radial basis function (RBF) models.

		Data	R^2	RMSE	MSE	MAD
1	GMDH	Training	0.90	3.13	9.83	2.48
		Testing	0.92	2.47	6.08	2.01
2	MLP	Training	0.93	2.54	6.46	2.05
		Testing	0.90	2.84	8.07	2.12
3	PTF	Training	0.73	5.13	26.29	3.72
		Testing	0.90	2.57	6.58	1.96
4	RBF	Training	0.43	7.40	54.72	5.73
		Testing	0.77	4.11	16.89	3.29

Clay and OC are main factors for predict of CEC. Figure (7) shows a 3D scatter plot of OC, Clay, and CEC from study soil samples. The results of Sahrawat (1983) in Philippine soils and Bell and van Keulen (1995) in Mexico soils showed that greater than 50% of variation in CEC could be explained by the variation in clay and OC content. Amini et al., (2005) reported that the relationship between CEC and clay and organic matter appeared to be dominantly linear.

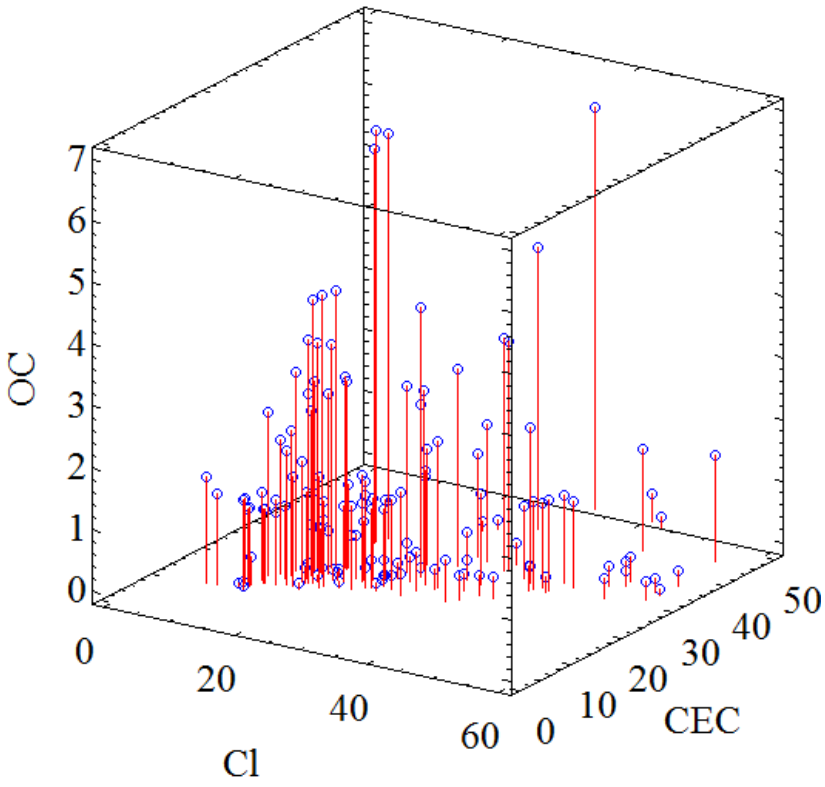

Fig. 7. A 3D scatter plot showing Organic carbon (OC), Clay (Cl), and Cation exchange capacity (CEC) from study soil samples.

4. Conclusions

In this study, PTF, ANNs (MLP and RBF) and feed-forward GMDH-type NN models developed using experimental CEC data. Then the results were compared with the experimental data. Despite the complexity of the

system studied, the GMDH model permits a good prediction of CEC. Thus, the GMDH model is suitable for predicting the CEC. The agreements between the experimental and calculated data were found to be excellent. This paper concluded GMDH model as being more effective and accurate in predicting of CEC than other ANNs and regression-based models.

References

Amini, M., Abbaspour, K.C., Khademi, H., Fathipour, N., Afyuni, M., Schulin, R., 2005. Neural network models to predict cation exchange capacity in arid regions of Iran. Eur. J. Soil Sci., 56, 551-559.

Anells, R.N., Arthurton, R.S., Bazley, R.A., Davis, R.G., 1975. Explanatory text of the Qazvin and Rasht Quadrangle map, 1:250,000. Geological survey of Iran, E3 and E4, p. 94.

Bell, M.A., van Keulen, J., 1995. Soil pedotransfer functions for four Mexican soils. Soil Sci. Soc. Am. J., 59, 865–871.

Drake, E.H., Motto, H.L., 1982. An analysis of the effect of clay and organic matter content on the cation exchange capacity of New Jersey soils. Soil Sci., 133, 281–288.

Erzin, Y., Hanumantha, Rao, B., Singh, D.N., 2008. Artificial neural network models for predicting soil thermal resistivity, Int. J. Therm. Sci., 47, 1347–1358.

Farlow, S.J., 1984. Self-Organizing Method in Modelling: GMDH-Type Algorithm, Marcel Dekker, New York.

Gee, G.W., Or, D., 2002. Particle-size analysis, In: Dane J.H., Topp G.C., (eds.). Methods of soil analysis, Part 4-Physical methods. Agronomy Monograph, vol. 9. ASA and SSSA, Madison, WI. USA. p. 255-293.

Ghanadzadeh, H., Ganji, M., Fallahi, S., 2012. Mathematical model of liquid–liquid equilibrium for a ternary system using the GMDH-type neural network and genetic algorithm. Appl. Math. Model., 36, 4096–4105.

Hagan, M.T., Demuth, H.B., Beale, M., 2002. Neural Network Design, Thomson Learning Pub., Singapore. p. 736.

Haykin, S., 1998. Neural Networks: A Comprehensive Foundation, 2nd edition. New York: Macmillan College Publishing.

Hecht-Nielsen, R., 1990. Neurocomputing. Addison-Wesley pub., Harlow. p. 433.

Ivakhnenko, A.G., 1971. IEEE Trans. Syst. man Cybernetics, 1, 364-78.

Ketabchi, S., Ghanadzadeh, H., Ghanadzadeh, A., Fallahi, S., Ganji, M., 2010. Estimation of VLE of binary systems (tert-butanol + 2-ethyl-1-hexanol) and (n-butanol + 2-ethyl-1-hexanol) using GMDH-type neural network. J. Chem. Thermodyn., 42(11), 1352–1355.

Minasny, B., McBratney, A.B., 2002. The neuro-m method for fitting neural network parametric pedotransfer functions. Soil Sci. Soc. Am. J., 66, 352–361.

Nariman-Zadeh, N., Jamali, A., 2007. Pareto Genetic Design of GMDH-type Neural Networks for Nonlinear Systems. In: Proc. of the International Workshop on InductiVe Modelling, Drchal J, Koutnik J, (eds.). Czech Technical University: Prague, Czech Republic. p. 96-103.

Nelson, D.W., Sommers, L.E., 1996. Total carbone, organic carbone, and organic matter, In: Sparks DL, (ed.), Methods of soil analysis, Part 3- chemical methods. Agronomy Monograph, vol. 9. ASA and SSSA, Madison, WI. USA. p. 961-1010.

Pachepsky, Y.A., Rawls, W.J., 1999. Accuracy and Reliability of Pedotransfer Functions as Affected by Grouping Soils. Soil Sci. Soc. Am. J., 63, 1748–1757.

Ripley, B.D., 1996. Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press.

Sahrawat, K.L., 1983. An analysis of the contribution of organic matter and clay to cation exchange capacity of some Philippine soils. Commu. Soil Sci. Plant Anal., 14, 803–809.

Sarmadian, F., Taghizadeh Mehrjardi, R., 2008. Modeling of some soil properties using artificial neural network and multivariate regression in Gorgan province, north of Iran. Global J. Environ. Res., 2(1), 30-35.

Seybold, C.A., Grossman, R.B., Reinsch, T.G., 2005. Predicting cation exchange capacity for soil survey using linear models, Soil Sci. Soc. Am. J., 69, 856–863.

Sonmez, H., Gokceoglu, C., Nefeslioglu, H.A., Kayabasi, A., 2005. Estimation of rock modulus: For intact rocks with an artificial neural network and for rock masses with a new empirical equation, Int. J. Rock Mech. Mining Sci., 43(2), 224–235.

Sumner, M.E., Miller, W.P., 1996. Cations exchange capacity and Exchange Coefficients, In: Sparks DL, (ed.). Methods of soil analysis, Part 3- chemical methods. Agronomy Monograph, vol. 9. ASA and SSSA, Madison, WI., USA. p. 1201-1230.

Thomas, G.W., 1996. Soil pH and soil acidity, In: Sparks DL, (ed.). Methods of soil analysis, Part 3- chemical methods. Agronomy Monograph, vol. 9. ASA and SSSA, Madison, WI. USA. p. 475-490.

Wagner, B., Tarnawski, V.R., 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.

Zacharias, S., 2007. Excluding organic matter content from pedotransfer predictors of soil water retention. Soil Sci. Soc. Am. J., 71, 43-50.