International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:4, No:2, 2010

Estimating Saturated Hydraulic Conductivity from Soil Physical Properties using Neural Networks Model

B. Ghanbarian-Alavijeh, A.M. Liaghat, and S. Sohrabi

***Abstract*—**Saturated hydraulic conductivity is one of the soil hydraulic properties which is widely used in environmental studies especially subsurface ground water. Since, its direct measurement is time consuming and therefore costly, indirect methods such as pedotransfer functions have been developed based on multiple linear regression equations and neural networks model in order to estimate saturated hydraulic conductivity from readily available soil properties e.g. sand, silt, and clay contents, bulk density, and organic matter. The objective of this study was to develop neural networks (NNs) model to estimate saturated hydraulic conductivity from available parameters such as sand and clay contents, bulk density, van Genuchten retention model parameters (i.e. θ_r, α, and *n*) as well as effective porosity. We used two methods to calculate effective porosity: (1) $\phi_{eff} = \theta_s - \theta_{FC}$, and (2) $\phi_{eff} = \theta_s - \theta_{\inf}$, in which θ_s is saturated water content, θ_{FC} is water content retained at -33 kPa matric potential, and $\theta_{\inf}$ is water content at the inflection point. Total of 311 soil samples from the UNSODA database was divided into three groups as 187 for the training, 62 for the validation (to avoid over training), and 62 for the test of NNs model. A commercial neural network toolbox of MATLAB software with a multi-layer perceptron model and back propagation algorithm were used for the training procedure. The statistical parameters such as correlation coefficient (R^2), and mean square error (MSE) were also used to evaluate the developed NNs model. The best number of neurons in the middle layer of NNs model for methods (1) and (2) were calculated 44 and 6, respectively. The R^2 and MSE values of the test phase were determined for method (1), 0.94 and 0.0016, and for method (2), 0.98 and 0.00065, respectively, which shows that method (2) estimates saturated hydraulic conductivity better than method (1).

***Keywords*—**Neural network, Saturated hydraulic conductivity, Soil physical properties.

I. INTRODUCTION

SEVERAL methods have been developed to estimate saturated hydraulic conductivity such as empirical models [1], semi empirical models [2-4], pedotransfer functions [5-11].

Semi empirical equation of Kozeny-Carman was proposed by Kozeny and later modified by Carman. The resulting equation is largely known as the Kozeny–Carman (KC) equation, although the two authors never published together. Limitations of Kozeny-Carman formula has been widely discussed in Carrier [12].

Ahuja et al. [13] modified Kozeny-Carman formula by defining an effective porosity as follows and showed that its applicability to a wide range of soils from the Southern Region of the USA, Hawaii, and Arizona:

$$K_s = B\phi_e^n \tag{1}$$

where K_s is saturated hydraulic conductivity, ϕ_e is the effective porosity, and *B* and *n* are the coefficient and exponent of model, respectively.

Ahuja et al. [14] and Franzmeier [15] indicated a degree of universality of Ahuja et al. [13] model for several soils from Korea and a variety of soils from Indiana, respectively.

Messing [16] presented data for some Norwegian soils where Ahuja et al. [13] model fitted the data for individual soils well, but the coefficients varied with soil type. Timlin et al. [17] reported a value of 0.0021 (m/s) and 3.29 for *B* and *n* parameters in Eq. (1), respectively.

Since different values of coefficient *B* and exponent *n* have been found for different data sets, Pachepsky et al. [18] used neural networks to find how these parameters are related to the parameters of Brooks and Corey [19] such as air entry value and pore-size distribution index. Their results showed that as the values of pore-size distribution index decreased from 1 to 0, values of *B* and *n* decreased.

Han et al. [20] developed a new model to estimate saturated hydraulic conductivity from soil structural properties derived from water retention curve. These authors showed that the inflection point of water retention curve includes useful information for modeling K_s.

Recently, neural networks have been widely used in simulation, classification and optimization of engineering sciences [21]. In the soil science, neural networks models have been applied in estimation of infiltration rate [22], cation exchange capacity [23], and saturated hydraulic conductivity [11, 24, 25].

Since there is no NNs model to estimate saturated hydraulic

B. Ghanbarian-Alavijeh is research assistant in University of Tehran, Karaj, 31587-77871 Iran (corresponding author to provide phone: +98-261-2241119; e-mail: b.ghanbarian@ gmail.com).
A. M. Liaghat, Professor in University of Tehran, Karaj, 31587-77871 IRAN. (e-mail: aliaghat@ut. ac.ir).
S. Sohrabi is a MSc student in Tabriz University, Tabriz, Iran.

conductivity from available data such as soil texture, soil water retention curve as well as bulk density, the objective of this study was to (1) develop a neural networks model for estimating saturated hydraulic conductivity from sand and clay contents, van Genuchten retention model parameters, and bulk density, and (2) compare two different definitions of effective porosity in estimation of saturated hydraulic conductivity using NNs model.

II. Materials and Methods

A. Data collection

In this study, 311 samples of the UNSODA database [26] with a wide range of soil texture were used. Fig. (1) gives the distribution of samples within the soil texture triangle based on the USDA classification.

van Genuchten [27] retention model parameters such as $\theta_s, \theta_r, \alpha$ and n were determined by direct fitting of van Genuchten [27] model to the measured data using RETC code [28]. The optimized vG model parameters were used to calculate water content retains at -33 kPa, θ_{FC}, and inflection point, $\theta_{\inf}$ [29].

In method 1, the effective porosity which was determined based on the Ahuja et al. [13] approach was used as a predictive variable in development of NNs model:

$$\phi_{eff} = \theta_s - \theta_{FC} \tag{2}$$

Dexter [29] explained that at the inflection point the behavior of soil moisture curve changes. For soil drying between saturation and the inflection point, it is mainly structural pores that are emptying. However, for soil drying below the inflection point, it is mainly textural pores that are emptying. In addition to the Ahuja et al. [13] approach, a new definition of effective porosity proposed by Han et al. [20] based on the inflection point of soil water retention curve was used as follows (method 2):

$$\phi_{eff} = \theta_s - \theta_{\inf} \tag{3}$$

B. Neural networks model

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 [22] and commonly consists of three layers, an input layer, a hidden layer, and an output layer.

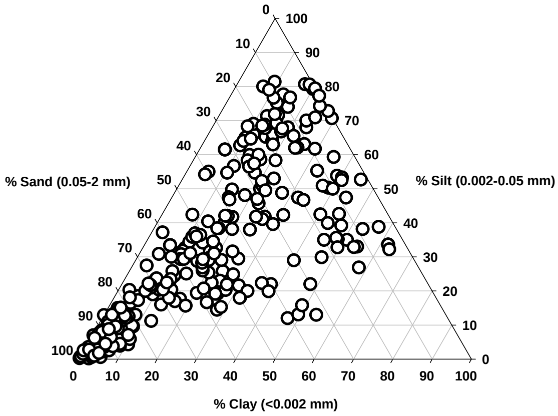


Fig. 1 Soil texture triangle of the UNSODA database samples used in this study

In several studies, feed forward networks in which the weighted connections feed activations only in the forward direction from an input layer to the output layer have been successfully used in combination of back propagation training algorithm which is a gradient descent algorithm [30-32].

In this study, a Multi Layer Perceptron (MLP) neural networks model with one hidden layer was used to estimate saturated hydraulic conductivity. This neural networks model consisted of a LOGSIG activation function in the hidden layer and a TANSIG activation function in the output layer. The number of Epochs was selected by MATLAB software for training the neural networks model. Furthermore, the best training model was selected by changing the number of neurons in the middle layer.

In this study in order to avoid over training, total of 311 was divided into three groups as 187 (60%) to develop, 62 (20%) to validate, and 62 (20%) to test the neural networks model. Input predictive variables such as clay and sand percentages, bulk density, effective porosity and vG retention model parameters such as θ_r, α and n of the development data set were standardized, and then were used to train the MLP neural networks model. As it was mentioned before, two different definitions of effective porosity were used to develop two NNs models and their performances were compared with each other. All the processes were carried out by Neural Network toolbox of MATLAB software [33].

C. Statistical criteria

In order to evaluate the developed neural networks model, the statistical parameters such as mean square error (MSE) and correlation coefficient (R^2) were calculated as follows:

$$MSE = \sum_{i=1}^{n} \left(O_i - P_i\right)^2 \Big/ n \tag{4}$$

$$R^2 = \sum_{i=1}^{n} (O_i - O_{ave})(P_i - P_{ave}) / \sqrt{\sum_{i=1}^{n} (O_i - O_{ave})^2 (P_i - P_{ave})^2} \tag{5}$$

International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:4, No:2, 2010

The less the MSE value, the better train the model.

III. Results and Discussion

Variation of K_s in the logarithmic scale versus different input predictive variables of neural networks model such as sand and clay contents, bulk density, vG model parameters and effective porosity calculated by using two different approaches for all 311 soil samples are shown in Fig. (2). This figure shows that when sand and clay contents increase, saturated hydraulic conductivity increases and decreases, respectively. No clear trend is evident between K_s and bulk density. By removing the zero values of residual water content, Fig. (2) also indicates that when residual water content increases, K_s decreases. However, there is a lot of scatter in the data. The result is in agreement with the result of Schaap et al. [34]. The authors presented average values of residual water content and K_s for 9 soil textural classes and showed that if clay content increases, residual water content increases as well and K_s decreases. The results also show that when vG model parameter n increases, K_s increases as well.

Fig. (2) also shows the variation of K_s as a function of two definitions of effective porosity. This figure indicates that when effective porosity increases, K_s increases as well. However, the values of K_s versus $\theta_s - \theta_{FC}$ values are in a more restricted range compared with the values of K_s versus $\theta_s - \theta_{\inf}$ values.

In this study, organic matter (OM) was not included as a predictive variable for developing the neural networks model because of the unavailability of this measurement for all samples of data set. Although, some researches have showed that when organic matter increases, saturated hydraulic conductivity decreases. Nemes et al. [35] investigated the influence of organic matter on the estimation of saturated hydraulic conductivity. Their results showed a strong indication that there is a negative relationship between OM and K_s. The authors justified this explanation by the fact that organic matter retains soil water well and does not allow water to flow freely. On the other hand, OM may also affect the pore size distribution of the soil through soil structure development which also affects the soil hydraulic conductivity.

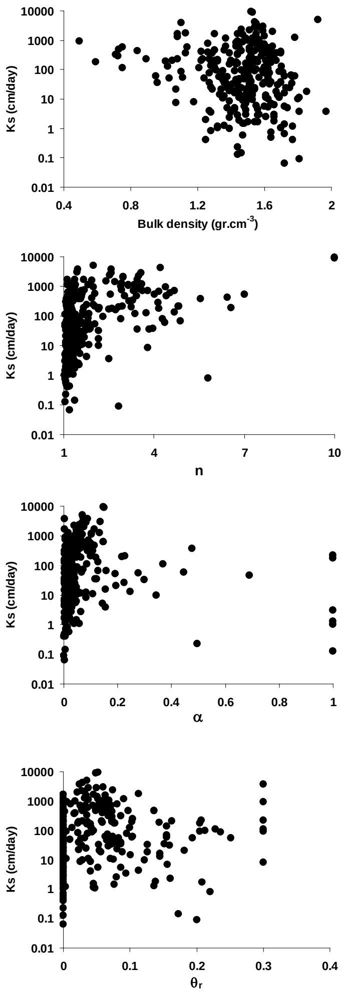
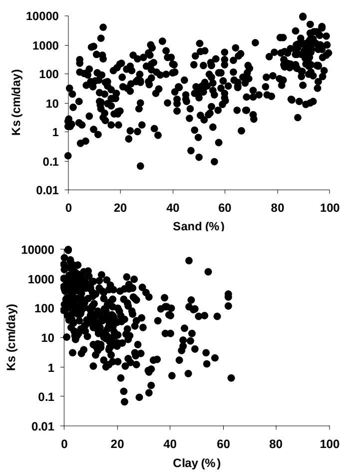


Fig. 2 Variation of saturated hydraulic conductivity versus input parameters of neural networks model for 311 soil samples



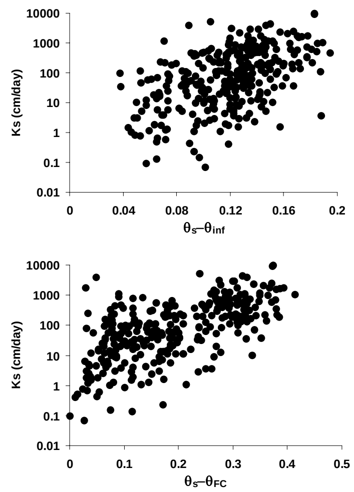


Fig. 2 Continued

The statistical parameters of neural networks model development, validation and test for methods 1 and 2 are presented in Tables 1 and 2, respectively. As it was mentioned before, the number of neurons in the hidden layer was determined by try and error method. In addition, the best model of neural networks was selected based on the two statistical parameters e.g. MSE and R^2 calculated for development, validation and test steps. The less the MSE value and the higher R^2 value, the better the model.

The MSE and R^2 values presented in Table 1 show that the best structure for method 1 is neural networks model with 44 neurons in the hidden layer. The MSE values of development and test steps were calculated 0.001 and 0.002 [(cm/day)^2], respectively, indicating that neural networks model has been developed well and estimated saturated hydraulic conductivity accurately. The R^2 values were also calculated for development and test steps 0.98 and 0.94, respectively, which show that the estimated and measured K_s values are highly correlated.

Table 2 also shows the calculated values of MSE and R^2 for method 2. The results indicate that the best structure of neural networks model for method 2 is when 6 neurons were used in the hidden layer. The MSE values of development and test steps were calculated 0.0004 and 0.001 [(cm/day)^2], respectively, which are less than the MSE values for method 1.

The less MSE values of method 2 in comparison with method 1 shows that method 2 in which the new definition of effective porosity presented by Han et al. [20] was used estimated saturated hydraulic conductivity slightly better than method 1. The low and high values of MSE and R^2 values for both methods 1 and 2 also indicate that both methods estimated saturated hydraulic conductivity well.

IV. Conclusion

In this study, available soil physical properties such as soil texture data (sand and clay contents), soil water retention curve e.g. vG retention model parameters, bulk density, and effective porosity were used to develop neural networks model in estimation of saturated hydraulic conductivity. Based on two definitions of effective porosity, two neural networks models were developed and their performances were compared with each other. It was found that method 2 in which neural networks model had been developed based on the new definition of effective porosity estimated saturated hydraulic conductivity slightly better than method 1.

TABLE I
STATISTICAL PARAMETERS CALCULATED FOR THE DEVELOPMENT, VALIDATION, AND TEST STEPS OF METHOD 1

No. of neurons	No. of epochs	Development		Validation		Test	
		MSE	R^2	MSE	R^2	MSE	R^2
2	18	0.003	0.92	0.002	0.91	0.004	0.51
3	15	0.006	0.89	0.002	0.51	0.005	0.77
4	8	0.009	0.96	0.004	0.43	0.005	0.50
6	18	0.001	0.96	0.004	0.86	0.003	0.91
7	10	0.004	0.71	0.016	0.59	0.018	0.54
9	10	0.004	0.98	0.004	0.48	0.003	0.47
10	12	0.001	0.98	0.002	0.72	0.007	0.78
11	9	0.003	0.97	0.005	0.50	0.008	0.87
12	19	0.001	0.92	0.003	0.93	0.013	0.85
13	19	0.001	0.98	0.005	0.41	0.002	0.93
14	11	0.002	0.98	0.002	0.56	0.006	0.37
15	13	0.000	0.99	0.003	0.62	0.005	0.67
16	7	0.008	0.71	0.009	0.56	0.012	0.67
17	25	0.001	0.98	0.005	0.84	0.001	0.81
21	15	0.002	0.96	0.002	0.90	0.008	0.47
22	16	0.000	0.99	0.002	0.71	0.003	0.90
23	20	0.002	0.93	0.002	0.88	0.001	0.98
24	24	0.000	0.98	0.006	0.64	0.006	0.82
25	10	0.002	0.96	0.001	0.96	0.007	0.89
28	6	0.003	0.97	0.001	0.96	0.001	0.97
29	21	0.000	0.94	0.008	0.52	0.019	0.65
30	9	0.003	0.99	0.006	0.43	0.006	0.20
32	9	0.002	0.99	0.002	0.48	0.003	0.88
33	14	0.002	0.98	0.002	0.50	0.006	0.40
34	10	0.013	0.87	0.003	0.07	0.016	0.69
36	10	0.003	0.96	0.006	0.31	0.002	0.93
37	10	0.002	0.99	0.011	0.13	0.010	0.78
38	8	0.002	0.99	0.005	0.70	0.002	0.67
40	13	0.001	0.99	0.004	0.52	0.002	0.95
41	11	0.001	0.98	0.006	0.32	0.002	0.93
42	9	0.005	0.99	0.007	0.49	0.005	0.47
44	**13**	**0.001**	**0.98**	**0.003**	**0.88**	**0.002**	**0.94**
45	8	0.004	0.99	0.009	0.76	0.002	0.62
50	17	0.001	0.99	0.001	0.44	0.005	0.85

TABLE II
STATISTICAL PARAMETERS CALCULATED FOR THE DEVELOPMENT, VALIDATION, AND TEST STEPS OF METHOD 2.

No. of neurons	No. of epochs	Development		Validation		Test	
		MSE	R^2	MSE	R^2	MSE	R^2
1	8	0.017	0.89	0.007	0.063	0.004	0.47
3	23	0.002	0.95	0.003	0.882	0.012	0.18
4	25	0.001	0.96	0.005	0.925	0.002	0.69
5	14	0.002	0.98	0.004	0.425	0.006	0.51
6	**19**	**0.000**	**0.95**	**0.004**	**0.898**	**0.001**	**0.98**
7	20	0.001	0.98	0.002	0.934	0.002	0.62
8	8	0.007	0.94	0.006	0.766	0.006	0.88
9	18	0.001	0.98	0.006	0.889	0.004	0.45
11	15	0.001	0.98	0.004	0.447	0.005	0.92
12	9	0.005	0.94	0.002	0.890	0.002	0.33
13	8	0.017	0.96	0.005	0.201	0.005	0.11
16	42	0.001	0.98	0.003	0.909	0.002	0.53
17	11	0.003	0.95	0.002	0.208	0.015	0.82
18	14	0.001	0.95	0.012	0.265	0.027	0.47
19	11	0.005	0.95	0.001	0.889	0.006	0.01
20	17	0.000	0.99	0.001	0.891	0.007	0.34
21	7	0.006	0.95	0.008	0.558	0.007	0.90
22	15	0.001	0.99	0.023	0.762	0.001	0.98
24	17	0.002	0.93	0.003	0.895	0.001	0.96
25	26	0.000	0.99	0.002	0.927	0.009	0.81
27	14	0.001	0.99	0.002	0.566	0.004	0.90
30	12	0.001	0.96	0.008	0.795	0.016	0.82
31	7	0.008	0.99	0.004	0.321	0.007	0.05
32	59	0.000	0.99	0.001	0.942	0.005	0.60
33	15	0.000	0.99	0.002	0.451	0.005	0.84
34	13	0.001	0.99	0.004	0.446	0.001	0.97
36	51	0.001	0.96	0.001	0.954	0.012	0.52
38	25	0.001	0.87	0.001	0.970	0.002	0.94
40	12	0.001	0.99	0.000	0.408	0.003	0.87
41	11	0.000	0.99	0.006	0.524	0.005	0.86
43	7	0.020	0.98	0.005	0.037	0.014	0.28
44	7	0.008	0.99	0.004	0.186	0.008	0.48
45	10	0.001	0.99	0.005	0.740	0.009	0.33
48	15	0.000	1.00	0.003	0.735	0.005	0.33

ACKNOWLEDGMENT

The authors acknowledge Dr. M.Th. van Genuchten for providing the UNSODA database used in this study.

REFERENCES

[1] A. Hazen. Discussion of 'Dams on sand foundations' by A. C. Koenig, Trans. Am. Soc. Civ. Eng., 1911, 73:199-203.

[2] J. Kozeny. Ueber kapillare Leitung des Wassers im Boden. Wien, Akad. Wiss., 1927, 136:271.

[3] P. C. Carman. The determination of the specific surface of powders. J. Soc. Chem. Ind. Trans., 1938,57:225.

[4] P. C. Carman. Flow of gases through porous media, Butterworths. Scientific Publications, London, 1956.

[5] G. S. Campbell. Soil Physics with Basic. Elsevier, New York, 1985.

[6] H. Vereecken, J. Maes, J. Feyen. Estimating unsaturated hydraulic conductivity from easily measured soil properties. Soil Sci., 1990, 149:1–12.

[7] J. H. M. Wösten. Pedotransfer functions to evaluate soil quality. In: Gregorich, E.G., Carter, M.R.ŽEds.., Soil Quality for Crop Production and Ecosystem Health. Developments in Soils Science, vol. 25, Elsevier, Amsterdam, 1997, pp. 221-245.

[8] J. H. M. Wösten, A. Lilly, A. Nemes, C. Le Bas. Development and use of a database of hydraulic properties of European soils. Geoderma, 1999, 90:169-185.

[9] B. J. Cosby, G. M. Hornberger, R. B. Clapp, T. R. Ginn. A statistical exploration of soil moisture characteristics to the physical properties of soils. Water Res. Res., 1984,20:682-690.

[10] D. L. Brakensiek, W. J. Rawls, G. R. Stephenson. Modifying SCS hydrologic soil groups and curve numbers for rangeland soils. ASAE, 1984, Paper No. PNR-84-203, St. Joseph, MI.

[11] M. G. Schaap, F. J. Leij, M. Th. van Genuchten. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, J. Hydrol., 2001, 251:163-176.

[12] W. D. Carrier. Goodbye, Hazen; Hello, Kozeny-Carman. Journal of Geotech. and Geoenviron. Eng., 2003, 129:1054-1056.

[13] L. R. Ahuja, J. W. Naney, R. E. Green, D. R. Nielsen. Macroporosity to characterize spatial variability of hydraulic conductivity and effects of land management. Soil Sci. Soc. Am. J., 1984, 48:699-702.

[14] L. R. Ahuja, D. K. Cassel, R. R. Bruce, B. B. Barnes. Evaluation of spatial distribution of hydraulic conductivity using effective porosity data. Soil Sci., 1989, 148:404-411.

[15] D. P. Franzmeier. Estimation of hydraulic conductivity from effective porosity data for some Indiana soils. Soil Sci. Soc. Am. J., 1991, 55:1801-1803.

[16] I. Messing. Estimation of saturated hydraulic conductivity in clay soils from soil moisture retention data. Soil Sci. Soc. Am. J., 1989, 53:665-668.

[17] D. J. Timlin, L. R. Ahuja, R. D. Williams. Methods to estimate soil hydraulic parameters for regional scale applications of mechanistic methods. In Application of GIS to the Modeling of Non-point Source Pollutants in the Vadose Zone. SSSA Agronomy, Madison, WI, 1996, pp. 185-203.

[18] Y. A. Pachepsky, D. J. Timlin, L. R. Ahuja. Estimating saturated soil hydraulic conductivity using water retention data and neural networks. Soil Sci., 1999, 164:552-560.

[19] R. H. Brooks, A. T. Corey. Hydraulic properties of porous media. Colorado State University, Hydrology, 1964, Paper No. 3, 27 p.

[20] H. Han, D. Gimenez, L. Lilly. Textural averages of saturated soil hydraulic conductivity predicted from water retention data. Geoderma, 2008, 146:121-128.

[21] M. Menhaj. Fundamental of artificial neural network. Amirkabir Press, 2000 (in Persian).

[22] A. Jain, A. Kumar. An evaluation of artificial neural network technique for the determination of infiltration model parameters. Applied Soft Computing, 2006, 6:272-282.

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

[24] M. Doaee, M. Shabanpour sharestani, F. Bagheri. Modeling saturated hydraulic conductivity in clay soils in guilan province (Iran) using artificial neural networks. J. Agric. Sci., 2005, 1:41-48 (in Persian).

[25] K. Parasuraman, A. Elshorbagy, B.C. Si. Estimating saturated hydraulic conductivity in spatially variable fields using neural network ensembles. Soil Sci. Soc. Am. J., 2006, 70:1851-1859.

[26] F. J. Leij, W. J. Alves, M. Th. van Genuchten, J. R. Williams. Unsaturated soil hydraulic database, UNSODA 1.0 user's manual. Rep. EPA/600/R96/095. USEPA, Ada, OK, 1996.

[27] M. Th. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Am. J., 1980, 44:892-898.

[28] M.Th. van Genuchten, F.J. Leij, S.R. Yates. The RETC code for quantifying the hydraulic functions of unsaturated soils. USDA, US Salinity Laboratory, Riverside, CA. United States Environmental Protection Agency, document EPA/600/2-91/065, 1991.

[29] A. R. Dexter. Soil physical quality Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth. Geoderma, 2004, 120:201-214.

[30] E. R. Levine, D. S. Kimes, V. G. Sigillito. Classifying soil structure using neural networks. Ecol. Model., 1996, 92:101-108.

[31] J. Morshed, J. J. Kaluarachchi. Application of artificial neural network and genetic algorithm in flow and transport simulations. Adv. Water Res., 1998, 22:145-158.

[32] H. Yoon, Y. Hyun, K-K. Lee. Forecasting solute breakthrough curves through the unsaturated zone using artificial neural networks. J. Hydrol., 2007, 335:68-77.

[33] The MathWorks, Inc. MATLAB: The Language of Technical Computing. version 7.5, 2008.

[34] M. Schaap, F.J. Leij, M.Th. van Genuchten. Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J., 1998, 62:847-855.

[35] A. Nemes, W. J. Rawls, Y. A. Pachepsky. Influence of organic matter on the estimation of saturated hydraulic conductivity. Soil Sci. Soc. Am. J., 2005, 69:1330-1337.