

Evaluation of adaptive neural-based fuzzy inference system approach for estimating saturated soil water content

Fereshte Haghighi Fashi¹

Received: 23 October 2016 / Accepted: 2 November 2016 / Published online: 10 November 2016
© Springer International Publishing Switzerland 2016

Abstract The saturated soil water content (θ_s) is an important parameter in hydrological studies. In this paper, adaptive neural-based fuzzy inference system (ANFIS) was used for estimation of soil saturation percentage of some flood spreading areas in Iran. Soil particle size distribution (sand%, silt%, and clay%), bulk density and medium porosity (0.2–30 μm) were used to develop saturated soil water content pedotransfer functions (PTFs). Then, contributions of various member functions (MFs) were assessed on estimation of θ_s . The results showed that the member function type has an important role in performance of ANFIS approach. In the present investigation, Gaussian curve (gaussmf) was found to be superior over the other MFs in estimating θ_s . In all of the θ_s PTFs, correlation between estimations of θ_s and corresponding observations was the low. R^2 values between measured and PTF-estimated θ_s using ANFIS approach did not increase as some input predictors were used in the PTFs (from PTF₁ to PTF₅). Based on the results, it is suggested that ANFIS model can be applied for reasonable estimation of θ_s and there is a need for obtain more information of the proposed approach especially for the selection of best member functions. Therefore, a good performance may be obtained when best member function would be selected in addition to the more effective PTF inputs.

Keywords ANFIS · Member function · PTFs · Saturated soil water content

Introduction

The estimation of saturated soil water content (θ_s) is important for the knowledge of hydrological behaviour of watersheds. Measuring this parameter is time- and finance-consuming. Therefore, indirect estimation of the soil hydraulic properties has been concerned. Pedotransfer functions (PTFs) are functions applied to estimate the soil hydraulic properties using basic soil properties such as sand, silt, and clay percentages, bulk density (BD) and other easy to measure properties (Bouma and Van Lanen 1987). A detailed review of PTFs can be found in ver-eecken et al. (2010). Multiple-linear regression method (Mayr and Jarvis 1999; Tomasella et al. 2000), group Method of Data Handling (Pachepsky and Rawls 1999) and neural network analysis (ANN) (Schaap et al. 1998; Minasny and McBratney 2002; Minasny et al. 2004) have been used to develop θ_s PTFs. Tamari et al. (1996) have presented a review of ANN applicability to estimate soil hydraulic properties. Two problems that should be concerned in developing ANN-based models are included: (1) the learning algorithm may not get optimum weights to minimize prediction errors, (2) a number of weights that are difficult to easy interpretation (Schaap et al. 2001). Therefore, ANN approach is not always applicable to develop the accurate and reliable θ_s PTFs (Twarakavi et al. 2009). Adaptive neuro-fuzzy inference system (ANFIS) can be an appropriate approach using the learning capability of an ANN for parameter optimization. ANFIS have been used for database management, signal processing, hydrologic applications (Jang 1993; Nayak et al. 2004) and developing PTFs for estimation of paddy soils CEC (Seyedmohammadi et al. 2016). Researchers have not assessed the potential of this approach in saturated soil water study. There is no discussion in the literature of

✉ Fereshte Haghighi Fashi
haghighif634@yahoo.com

¹ University of Tehran, Tehran, Iran

applications of ANFIS approach for developing θ_s PTFs and evaluation of member function types. An important advantage of this technique is that it does not require the model structure to be known a priori compared with most of the approaches (Nayak et al. 2004). Nayak et al. (2004) showed that the ANFIS outperforms ANNs and other traditional methods in terms of prediction errors, computational speed and efficiency. The objective of this study was to evaluate the applicability of ANFIS-based PTFs and different member functions in estimating θ_s to overcome the problems with the other approaches in some soils of Iran's flood spreading areas. In addition, the effect of different input variables to develop θ_s PTFs was investigated.

Description of ANFIS

Since the detailed review of adaptive networks of ANFIS have been provided by Jang (1993), only a briefly description is presented in this paper. The adaptive network is a multi-layer feed-forward network in which each node performs a particular function on incoming signals and a set of parameters pertaining to this node (Jang 1993). This structure has five layers including fuzzy layer, product layer, normalized layer, de-fuzzy layer and total output layer. A detail about the functioning of the ANFIS is as follows (Jang 1993):

Layer 1 Every node (i) in this layer is an adaptive node with node function. $O_{1,i}$ is the output of the *i*th node of the layer 1:

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(x) \quad \text{for } i = 1, 2, \text{ or} \\ O_{1,i} &= \mu_{B_{i-2}}(y) \quad \text{for } i = 3, 4, \end{aligned} \quad (1)$$

where *x* (or *y*) is the input to node *i*, and A_i is a linguistic label associated with this node. Therefore, $O_{1,i}$ is the membership function of A_i . The bell shaped membership functions (MF) can be computed as follows:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x-c_i}{a_i} \right|^{2b}}, \quad (2)$$

where a_i , b_i and c_i are the parameter set to be learnt.

Layer 2 Each node in this layer denoted as Π which multiply incoming signals and sending the product out.

$$O_{2,i} = w_i = (\mu_{A_i}(x)) \times (\mu_{B_i}(y)) \quad i = 1, 2, \quad (3)$$

where each node represents the firing strength of the rule. w_1 , w_2 are the weight functions of the next layer.

Layer 3 Each node in this layer is labeled as N . $O_{3,i}$ indicates the Layer 3 output as shown in Eq. 4:

$$O_{3,i} = \bar{w} = \frac{w_i}{w_1 + w_2}. \quad (4)$$

Layer 4 This layer's nodes are adaptive with node functions.

$$O_{4,i} = w_i f_i = w_i(p_i x + q_i y + r_i), \quad (5)$$

where w_i denotes the output of layer 3 and p_i , q_i , r_i are the parameters set.

Layer 5 This layer's single node computes the overall output as the summation of all incoming signals as:

$$O_{5,i} = \sum_i \bar{w} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, \quad (6)$$

where $O_{5,i}$ denotes the layer 5 output.

Materials and methods

The θ_s PTFs were developed using the database established by Mahdian and Kamali (2010). Some flood spreading areas in Iran were chosen and soil properties data including sand%, silt%, clay%, Organic carbon and bulk density (BD) were collected. Then, the content of pores of medium size (0.2–30 μm) was calculated. Soil properties have been measured based on USDA standards. Data were divided into approximately 80% for training and the 20% for testing. Then, the performance of ANFIS-based PTFs and different member functions for θ_s predictions were assessed. In the present study, the Sugeno fuzzy model (Takagi and Sugeno 1985) and back propagation algorithm were applied for training. A set of models was developed based on the ANFIS approach for the water content at θ_s . The selection of best inputs is an important issue in the ANFIS-based models development. ANFIS-based PTFs were developed by varying the above mentioned input variables. In the first set of models (P_1) (SSC), sand%, silt%, and clay% were employed as inputs. The second set of models (P_2) (SSCBD) employed sand%, silt%, clay% and BD as input variables, while the input data of the third set of models (P_3) were included medium porosity (0.2–30 μm) in addition to those inputs used for P_1 . The fourth set of models (P_4) used all input variables mentioned above together. The fifth set of models (P_5) used all input variables mentioned above together in addition to organic carbon (OC). Moreover, considering a hierarchical approach, SSCBD θ_{33} model employed sand, silt and clay percentages, bulk density and water retention point at –33 kPa potential. SSCBD $\theta_{33}\theta_{1500}$ added the water content at –1500 kPa to the input variables of the SSCBD θ_{33} model. In the present investigation, different eight MFs (Gaussian curve (gaussmf), Π -shaped (pimf), Triangular-shaped (trimf), Trapezoidal-shaped (trapmf), Generalized bell-shaped (gbellmf), composed of product of two sigmoidally shaped (psigmf), composed of difference between

Table 1 Summary statistics of input data for training and testing ANFIS-based PTFs

| Variables | Training set | | | | Testing set | | | |
|---|--------------|-------|-------|--------|-------------|-------|-------|-------|
| | Min | Max | Mean | SD | Min | Max | Mean | SD |
| Sand (%) | 17.00 | 79.00 | 59.95 | 13.226 | 62.00 | 79.00 | 71.50 | 5.212 |
| Silt (%) | 3.00 | 49.00 | 26.97 | 9.392 | 9.00 | 26.00 | 17.22 | 4.969 |
| Clay (%) | 2.00 | 35.00 | 13.08 | 6.349 | 6.00 | 15.00 | 11.27 | 1.876 |
| BD (g cm ⁻³) | 1.02 | 1.95 | 1.44 | 0.168 | 1.50 | 1.66 | 1.55 | 0.030 |
| OC (%) | 0.02 | 2.70 | 0.55 | 0.562 | 0.16 | 0.64 | 0.26 | 0.089 |
| FC (cm ⁻³ cm ⁻³) | 0.05 | 0.37 | 0.16 | 0.057 | 0.16 | 0.21 | 0.18 | 0.012 |
| PWP (cm ⁻³ cm ⁻³) | 0.02 | 0.24 | 0.08 | 0.038 | 0.07 | 0.11 | 0.09 | 0.009 |
| P (μm) | 0.01 | 0.16 | 0.08 | 0.029 | 0.08 | 0.11 | 0.09 | 0.009 |
| θ _s (cm ⁻³ cm ⁻³) | 0.17 | 0.49 | 0.27 | 0.069 | 0.20 | 0.43 | 0.25 | 0.041 |

SD Standard deviation, BD bulk density, FC field capacity, PWP permanent wilting point, OC organic carbon, P medium porosity (0.2–30 μm)

two sigmoidal (dsigmf) and Gaussian combination (gaussmf)) were assessed for developing θ_s ANFIS-based PTFs. The PTFs performance was investigated based on the RMSE, ME, and R². The MFs of the ANFIS takes values between 0 and 1. Therefore, input data were normalized so as to lie in the range of 0–1 (Masters 1993). The models development was done using MATLABTM (The MathWorks Inc. USA).

Result and discussion

PTFs performance of the saturated soil water content

The brief statistics of soil properties of the studied flood spreading areas are presented in Table 1. In this research, studied soils were coarse-textured and the bulk densities ranged from 1.02 to 1.95 g cm⁻³. The goodness of fit statistics on the normalized values (Table 2) for all the PTFs is presented in Table 3. In this research, it was observed that the variation of the performance criteria was significant for ANFIS-based PTFs with different MF types (Table 4). Ignoring the various input data set, the ANFIS-based models were able to estimate soil hydraulic

Table 2 Summary statistics on the normalized values of studied soil hydraulic properties for training and testing pedotransfer functions

| Variables | Training set | | | | Testing set | | | |
|----------------|--------------|------|------|------|-------------|------|------|------|
| | Min | Max | Mean | SD | Min | Max | Mean | SD |
| FC | 0.00 | 1.00 | 0.37 | 0.18 | 0.35 | 0.50 | 0.43 | 0.04 |
| PWP | 0.00 | 1.00 | 0.27 | 0.18 | 0.22 | 0.40 | 0.33 | 0.04 |
| θ _s | 0.00 | 1.00 | 0.32 | 0.21 | 0.09 | 0.81 | 0.25 | 0.12 |

SD Standard deviation, BD bulk density, FC field capacity, PWP permanent wilting point, OC organic carbon, P medium porosity (0.2–30 μm)

properties with a reasonable accuracy, as can be evidenced by relatively the low RMSE values (Table 3). The fairly same RMSE values for estimation of θ_s using different combinations of input variables were observed in Table 3 and Fig. 1. Soil structure and pore-size distribution may extensively influence soil water content in the air-entry and capillary region of the soil water retention curve (Merdun et al. 2006), which were only indirectly accounted for in the θ_s PTFs by BD and medium porosity. Medium pore-size was included as an input variable to develop ANFIS-based PTFs of θ_s, which was not effective to improve the model accuracy.

Effect of input variables

From Table 3, it is apparent that all of the PTFs perform similarly as the RMSE did not vary significantly amongst different combination of inputs except. One other observation is that R² values between measured and PTF-estimated θ_s using ANFIS approach did not increase as some input predictors were used in the PTFs (from P₁ to P₅). In all of the θ_s PTFs, correlation between estimations of θ_s and corresponding observations was the low (Tables 3, 4). Nemes et al. (2003) and Vereecken et al. (2010) have stated that including extra input variables can improve the efficiency of ANN models. However, the results of the present study showed that increasing the number of input variables did not necessarily lead to improve the model performance and better estimations. This is consistent with earlier findings of Nemes et al. (2006) and Haghverdi et al. (2012). As mentioned above, there are many other important parameters such as soil salinity and clay mineralogy that may significantly influence the soil water content (Nemes et al. 2003), which were not accounted for as input data for water retention PTFs in the present research. Vegetation and topography characteristics may significantly affect soil hydraulic properties (Sharma et al. 2006), which were

Table 3 The goodness of fit statistics during validation for the best selected PTFs

| Hydraulic parameter | Input variables | R^2 | $RMSE$ | ME | MF | MF type | Epoch |
|---------------------|-----------------------------|--------|--------|--------|----------|----------|-------|
| θ_s | Sand, Silt, Clay | 0.5527 | 0.0733 | 0.0241 | Gauss2mf | Constant | 150 |
| θ_s | Sand, Silt, Clay, BD | 0.5265 | 0.0779 | 0.0437 | gaussmf | Constant | 150 |
| θ_s | Sand, Silt, Clay, P | 0.5218 | 0.0745 | 0.0335 | gaussmf | Constant | 150 |
| θ_s | Sand, Silt, Clay, P, BD | 0.4977 | 0.0792 | 0.0431 | trapmf | Constant | 150 |
| θ_s | Sand, Silt, Clay, P, BD, OC | 0.519 | 0.0804 | 0.0480 | gbellmf | Constant | 150 |

Gaussian curve (gaussmf), Trapezoidal-shaped (trapmf), Generalized bell-shaped (gbellmf), Gaussian combination (gauss2mf)

ME correlation coefficient, $RMSE$ root mean square error

Table 4 The correlation coefficient between the estimated and measured saturated soil water content for different member functions (MF1 to MF8)

| Parameter | PTF | MF1 | MF2 | MF3 | MF4 | MF5 | MF6 | MF7 | MF8 |
|------------|------------------|------|------|------|------|------|------|------|------|
| θ_s | P_1 (SSC) | 0.24 | 0.04 | 0.55 | 0.43 | 0.51 | 0.35 | 0.05 | 0.05 |
| | P_2 (SSCBD) | 0.32 | 0.22 | 0.52 | 0.50 | 0.24 | 0.30 | 0.18 | 0.18 |
| | P_3 (SSCP) | 0.35 | 0.37 | 0.48 | 0.52 | 0.47 | 0.21 | 0.34 | 0.34 |
| | P_4 (SSCPBD) | 0.48 | 0.47 | 0.49 | 0.46 | 0.47 | 0.26 | 0.27 | 0.43 |
| | P_5 (SSCPBDOC) | 0.51 | 0.49 | 0.48 | 0.47 | 0.47 | 0.26 | 0.26 | 0.26 |

P_1 sand, silt and clay; P_2 sand, silt, clay and BD; P_3 sand, silt, clay and porosity; P_4 sand, silt, clay, porosity and BD; P_5 sand, silt, clay, porosity, BD and OC; OC organic carbon; P medium porosity (0.2–30 μm); BD bulk density

ignored in the current study. Considering vegetation and topography characteristics as similar among training and test datasets, if not included as estimators, may improve the performance of the models (Haghverdi et al. 2012).

Tomasella et al. (2000) reported the poor fits of retention points near saturation. Considering RMSE, it was seen that increasing the BD and porosity as predictors did not necessarily improve the performance of the PTFs. It seems that while evaluating the performance of any model for its accuracy and applicability in predicting saturated soil water content, it is not only important to evaluate the number and type of input variables but also the differences in member functions. Including more input variables for developing the θ_s PTFs may be confusing and, especially when they have not significant effect, the development will be complicated. Extra input variables may also be employed to improve the performance of the models, such as soil structure (Minasny et al. 1999), soil chemical properties.

In the case of the ANFIS-based PTFs developed in this study, there are different models with various input data set and MF types. Therefore, it seems ANFIS approach provides more flexible θ_s PTFs. To investigate the above mentioned reason and observed improvement, it is needed to do further research evaluating ANFIS in future researches. In the present study, the ANFIS model took 150–300 epochs, implying considerable savings in computational time for ANFIS models. The results suggested that the model building process can be simplified when an ANFIS

model is developed compared to an ANN, and the ANFIS model preserves the full potential of ANN models in its performance (Nayak et al. 2004). Neuro-fuzzy techniques remove some of the shortcomings of ANNs. ANFIS had a good applicability in this study that is reported by Nayak et al. (2004), Aqil et al. (2007) and Dastorani et al. (2010).

A comparative analysis suggests that the proposed modeling approach may outperform ANNs and other traditional models in terms of computational speed, estimation errors and performance. Despite the general belief that ANN-based models are relatively successful, there are number of weights that do not allow easy interpretation (Schaap et al. 2001). Considering the above mentioned problems associated with ANN-based PTFs, ANN approach is not always able to improve the PTFs accuracy and reliability (Twarakavi et al. 2009).

The results demonstrated that the ANFIS showed good performance in terms of various statistical indices and the results are promising. It was also observed that the ANFIS model eases the model developing process. Therefore, it is important to develop the suitable PTFs using appropriate approaches (i.e. ANFIS) after the knowledge of limitations and weaknesses that need to be considered. Finally, a good performance may be obtained when pore size distribution would be measured in addition to the typical PTF inputs. It is noted that there are large uncertainties in estimation of ANFIS-based PTFs due to the lack of data for some textural classes (fine- and medium textured) in the training.

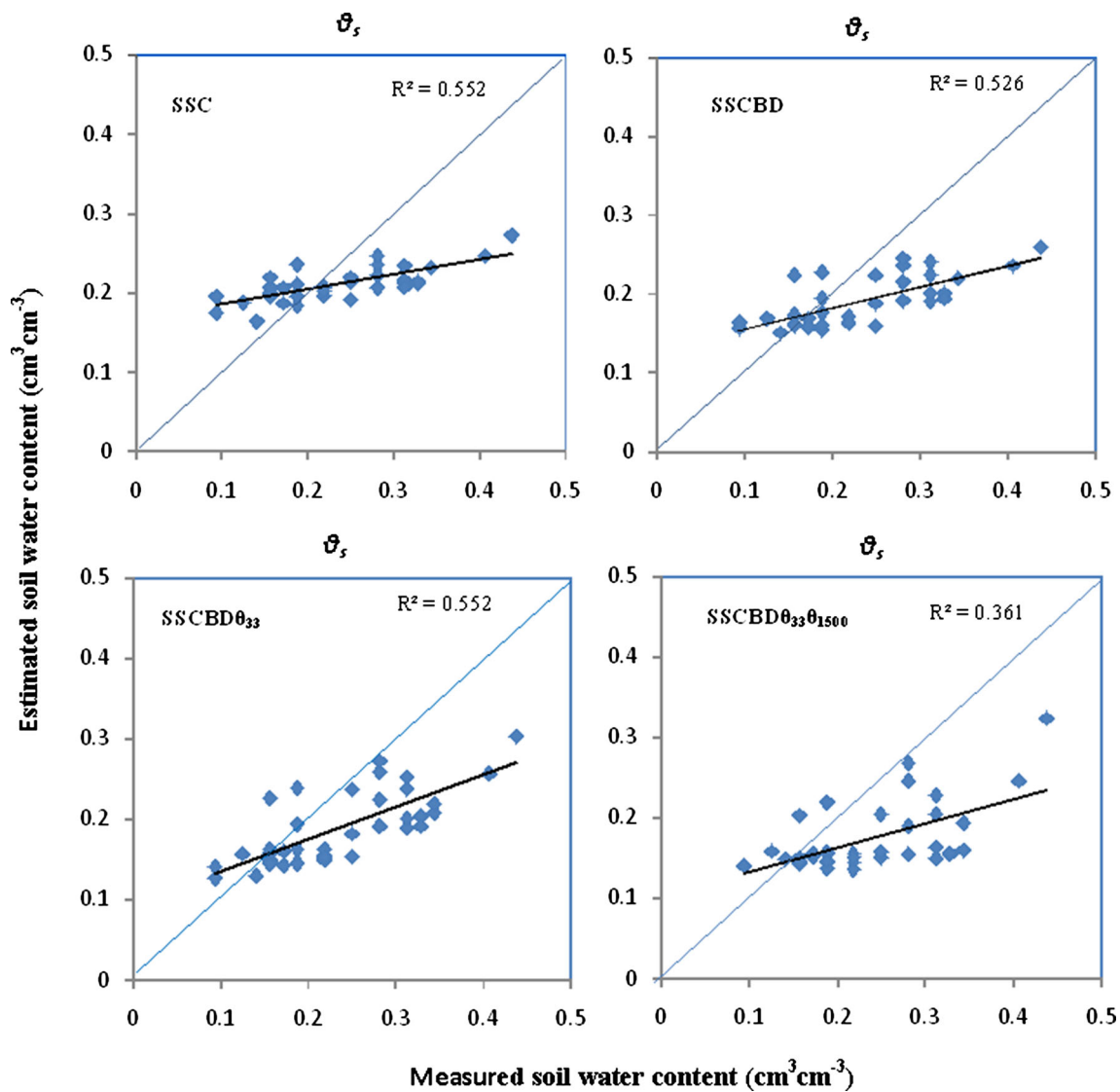


Fig. 1 Measured and estimated soil water content ($\text{cm}^3 \text{cm}^{-3}$) at saturation (θ_s) by ANFIS-based pedotransfer functions. *SSC* sand, silt and clay, *BD* bulk density

Conclusion

This study presented the application of an adaptive neuro fuzzy inference system (ANFIS) to saturated soil water content modeling. In this study, the ANFIS approach was successfully employed to develop saturated soil water content PTFs that used various input set. It was observed that an increase in some of the predictors (i.e. medium porosity) did not result in improved saturated soil water content prediction by PTFs. It was concluded that by increasing the BD as input variable to develop the ANFIS-based PTFs of saturated soil water content, the model accuracy was not improved and the observed difference was not statistically significant. The results were encouraging and suggest that an adaptive neuro-fuzzy approach is viable for modeling θ_s . Although estimation errors of

ANFIS for θ_s estimation were relatively similar, the results were different for various member function types to estimate soil water content at θ_s . There is a need for more application of the ANFIS approach especially for fine- and medium textured classes.

References

Aqil M, Kita I, Yano A, Nishiyama S (2007) A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *J Hydrol* 337:22–34

Bouma J, Van Lanen JAJ (1987) Transfer functions and threshold values: from soil characteristics to land qualities. In: Beek KJ, Burrough PA, McCormack DA (eds) *Quantified Land Evaluation Procedures*, vol 6. ITC, The Netherlands, pp 106–110

- Dastorani M, Moghadamnia A, Piri J, Rico-Ramirez M (2010) Application of ANN and ANFIS models for reconstructing missing flow data. *Environ Monit Assess* 166(1–4):421–434
- Haghverdi A, Cornelis WM, Ghahraman B (2012) A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *J Hydrol* 442–443:46–54
- Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference systems. *IEEE Trans Syst Man Cybern* 23:665–685
- Mahdian MH, Kamali K (2010) The investigation of floodwater spreading effects on the changes physical and chemical properties of soil in floodwater spreading areas. *Soil Conservation and Watershed Management Research Institute, Iran*
- Masters T (1993) *Practical neural networks recipes C++*. Academic Press, San Diego
- Mayr T, Jarvis NJ (1999) Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Corey type model. *Geoderma* 91:1–9
- Merdun H, Cinar 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
- Minasny B, McBratney AB (2002) The Neuro-m method 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
- Minasny B, Hopmans JW, Harter T, Eching SO, Tuli A, Denton MA (2004) Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Sci Soc Am J* 68:417–429
- Nayak PC, Sudheer KP, Ranganc DM, Ramasastry KS (2004) A Nero-fuzzy computing technique for modeling hydrological time series. *J Hydrol* 291:52–66
- Nemes A, Schaap MG, Wösten JHM (2003) Functional evaluation of pedotransfer functions derived from different scales of data collection. *Soil Sci Soc Am J* 67:1093–1102
- Nemes A, Rawls WJ, Pachepsky YA (2006) Use of the nonparametric Nearest Neighbor approach to estimate soil hydraulic properties. *Soil Sci Soc Am J* 70:327–336
- Pachepsky YA, Rawls WJ (1999) Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Sci Soc Am J* 63:1748–1757
- Schaap MG, Leij FJ, van Genuchten MT (1998) Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Sci Soc Am J* 62:847–855
- Schaap MG, Leij FJ, Van Genuchten MT (2001) Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J Hydrol* 25:163–176
- Seyedmohammadi J, Esmaeelnejad L, Ramezanzpour H (2016) Determination of a suitable model for prediction of soil cation exchange capacity. *Model Earth Syst Environ* 2:156
- Sharma SK, Mohanty BP, Zhu J (2006) Including topography and vegetation attributes for developing pedotransfer functions. *Soil Sci Soc Am J* 70:1430–1440
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans Syst Man Cybern* 15:116–132
- Tamari S, Wösten JHM, Ruiz-Suarez JC (1996) Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci Soc Am J* 60:1732–1741
- Tomasella J, Hodnett MG, Rossato L (2000) Pedotransfer functions for the estimation of soil water retention in Brazilian soils. *Soil Sci Soc Am J* 64:327–338
- Twarakavi Navin KC, Šimůnek J, Schaap MG (2009) Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Sci Soc Am J* 73(5):1443–1452
- Vereecken H, Weynants M, Javaux M, Pachepsky Y, Schaap MG, van Genuchten MT (2010) Using pedotransfer functions to estimate the van Genuchten-Mualem soil hydraulic properties: a review. *Vadose Zone J* 9:1–26