Effects of various input levels and different soil water retention curve models on water content estimation using different statistical methods

Hossein Bayat and Eisa Ebrahimi

ABSTRACT

This study investigated the impact of different input variables on the predictability of the water content using soil water retention curve (SWRC) models. The particle and aggregate size distribution model parameters were calculated by fitting the Perrier model to the related distributions for 75 soil samples. Nine SWRC models were fitted to the experimental data and their coefficients were obtained. The regression method was used to estimate the coefficients for nine SWRC models at three input levels. Cluster analysis classified the SWRC models into more homogeneous groups according to the accuracy of their predictions. The SWRC estimated using the Gardner model had the highest accuracy, but it was not an appropriate model for the soils because of its low fitting accuracy. Boltzman, Campbell, and Fermi models obtained the highest accuracy after the Gardner model. The Durner model yielded the lowest prediction accuracy due to the lack of correlation between the input variables and coefficients in this model. Thus, the water content predictions obtained using different SWRC models varied because different input variables were employed.

Key words | cluster analysis, multivariate regression, soil water retention curve model

Hossein Bayat (corresponding author)
Eisa Ebrahimi
Department of Soil Science, Faculty of Agriculture,
Bu Ali Sina University,
Hamedan,
Iran
E-mail: *h.bayat@basu.ac.ir*

NOMENCLATURE

Parameter	Description	Unit	Parameter	Description	Unit
BD	Bulk density	gr cm^{-3}	N	Number of samples	
			q	Number of model parameters	
α_{PSD}, L_{PSD}, D_{PSD}	Parameters of the particle size distribution (PSD) model		θ_r	Residual water content	cm^3 cm^{-3}
α_{Agg}, L_{Agg}, D_{Agg}	Parameters of the aggregate size distribution model		Ψ	Soil suction	kPa
			α Gardner	Fitting parameter related to the air entry value	kPa^{-1}
θ_1	Balance water content	gr			
θ_s	Saturated water content	cm^3 cm^{-3}	α Campbell	Air entry value	kPa
			n Campbell	Slope of the soil water retention curve (SWRC)	
θ_{pre}	Predicted water content	cm^3 cm^{-3}	α van Genuchten	Inverse of the air entry value	kPa^{-1}
θ_{mean}	Average measured water content	cm^3 cm^{-3}	n van Genuchten	Related to the pore size distribution of the soil	
θ_{meas}	Measured water content	cm^3 cm^{-3}	m van Genuchten	Related to the asymmetry of the model	

doi: 10.2166/nh.2015.219

Parameter	Description	Unit
α Fermi	The α parameter equals the soil suction when the effective degree of saturation equals 0.5	kPa
n Fredlund-Xing	Curve-fitting parameter related to the slope at the inflection point of the SWRC	
m Fredlund-Xing	Fitting parameter	
α Fredlund-Xing	Fitting parameter related to the air entry value	
Ψ_r Fredlund-Xing	Suction corresponding to the residual water content	kPa
c Dexter	Asymptote of the function equal to the residual water content	gr gr^{-1}
A_1 Dexter	Proportional to the amount of matrix pore spaces	gr gr^{-1}
A_2 Dexter	Proportional to the amount of structural pore spaces	gr gr^{-1}
Ψ_1 Dexter	Specific suction when the matrix pore spaces are empty	kPa
Ψ_2 Dexter	Specific suction when the structural pore spaces are empty	kPa
w Dexter	Gravimetric water content	gr gr^{-1}

INTRODUCTION

The soil water retention curve (SWRC) is defined as a nonlinear relationship between the soil water content and matric potential. The SWRC is one of the most important soil hydraulic properties, which depends on the pore size distribution, particle size distribution (PSD), and soil structure (Tuller & Or 2003). Numerous models have been developed to describe the SWRC and nine of these models are shown in Table 1.

The Gardner (1956) model is a continuous function for describing the SWRC. The Gardner model has low flexibility to fit the experimental SWRC data. Campbell's (1974) model assumes an abrupt air entry value, but although an abrupt air entry value may be applicable to coarse-textured soils, the air entry value for fine-textured soils is gradual. The van Genuchten (1980) model is one of the most popular models for describing the SWRC and this model estimates the sigmoid shape of the SWRC very well. The high flexibility of the van Genuchten model, the physical meaning of its parameters, and its capacity for describing the SWRC of a wide range of soils are its main advantages. Fredlund & Xing (1994) introduced a model that is similar to the van Genuchten model, which employs five parameters and can be fitted to a wide range of soils at different suctions.

The Boltzman model (McKee & Bumb 1984) has a simple form, but both the shape and position of the SWRC are affected by the two parameters utilized in the model, which makes it difficult to fit the SWRC using this model (Sillers *et al.* 2001). In addition, the Fermi model was introduced by McKee & Bumb (1987), which is simpler and more flexible than the Boltzman model. The Fermi model provides a continuous function, whereas the Boltzman model does not assume a continuous function.

The Durner (1994) and Seki (2007) models are bimodal functions based on the van Genuchten model and a log-normal soil pore size distribution, respectively. An advantage of bimodal models is their flexibility, but they are complex and difficult to fit to experimental data. In addition, the Dexter model (Dexter *et al.* 2008) is a double-exponential bimodal function based on the hierarchical pore size distribution, which divides the soil porosity into four parts. The Dexter model is not suitable for well-graded sands, but it describes the SWRC in agricultural soils very well and its parameters have physical meaning.

Each of the SWRC models has unique features, but they differ in the number of coefficients and the forms of the equations employed. In addition, direct measurement of the SWRC in the field or laboratory is time-consuming and expensive, while the results of SWRC measurements may lack the requisite precision due to temporal and spatial variability (Merdun *et al.* 2006).

Recently, soil hydraulic properties have been estimated extensively using pedotransfer functions (PTFs) because of their facile application and low requirements in terms of time and resources (Wosten *et al.* 1995; Minasny & McBratney 2002; Minasny *et al.* 2004). Estimating soil hydraulic properties using other methods such as constant head well permeameter method in the estimation of field-saturated hydraulic conductivity (Archer *et al.* 2014) would be time-consuming and expensive. PTFs are predictive

Table 1 | The SWRC models and their parameters

Model	Function	Parameters
Gardner (1956)	$\theta = \theta_r + (\theta_s - \theta_r)\left(\frac{1}{1+(\alpha\psi)^n}\right)$	θ_s, θ_r, α, n[a]
Campbell (1974)	$\psi = \alpha\left(\frac{\theta_s}{\theta}\right)^n$	θ_s, n, α
van Genuchten (1980)	$\theta = \theta_r + (\theta_s - \theta_r)\left(\frac{1}{(1+(\alpha\psi)^n)^m}\right)$ $m = 1 - \frac{1}{n}$	θ_s, θ_r, α, n, m
Boltzman (McKee & Bumb 1984)	$\theta = \theta_r + (\theta_s - \theta_r)\exp\left(\frac{\alpha - \psi}{n}\right)$	θ_s, θ_r, a, n
Fermi (McKee & Bumb 1987)	$\theta = \theta_r + (\theta_s - \theta_r)\left(\frac{1}{1+\exp\left(\frac{\psi - \alpha}{n}\right)}\right)$	θ_s, θ_r, a, n
Fredlund & Xing (1994)	$\theta = \theta_s\left(1 - \frac{\ln\left(1+\frac{\psi}{\psi_r}\right)}{\ln\left(1+\frac{10^6}{\psi_r}\right)}\right)\left(\frac{1}{\left(\ln\left(e+\left(\frac{\psi}{\alpha}\right)^n\right)\right)^m}\right)$	θ_s, a, n, m, Ψ_r
Durner (1994)	$s_e = \sum_{i=1}^{k} w_i\left[\frac{1}{1+(\alpha_i\psi)^{n_i}}\right]^{m_i}$	θ_s, $\theta_r w$, α_1, n_1, α_2, n_2
Seki (2007)	$s_e = \sum_{i=1}^{k} w_i Q\left[\frac{\ln\left(\frac{\psi}{\psi_{m_i}}\right)}{\sigma_i}\right]$ $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty}\exp\left(\frac{-t^2}{2}\right)dt$	θ_s, θ_r, w, Ψ_{m1}, σ_1, Ψ_{m2}, σ_2
Dexter *et al.* (2008)	$w = c + A_1 e^{\left(\frac{-\psi}{\psi_1}\right)} + A_2 e^{\left(\frac{-\psi}{\psi_2}\right)}$	c, A_1, h_1, A_2, h_2

[a]All of the abbreviations are defined in the Nomenclature.

functions for certain soil properties, which use data obtained from soil surveys or readily measured soil properties (Wösten *et al.* 2001). For example, Stirk (1957) estimated the permanent wilting point using the clay content, while Nielson & Shaw (1958) determined a parabolic relationship between the clay content and permanent wilting point. Gupta & Larson (1979) estimated the water content at matric potentials of 4 to 1,500 kPa using the PSD and organic matter content. Bayat *et al.* (2011, 2013b) used the fractal parameters of the PSD and aggregate size distribution to estimate the SWRC. Many other researchers have used different input variables to estimate the SWRC with different techniques, such as multiple regression and artificial neural networks (Wosten *et al.* 1995; Pachepsky *et al.* 1996; Koekkoek & Boolting 1999; Mayr & Jarvis 1999; Tomasella *et al.* 2000), group methods for data handling (Pachepsky *et al.* 1998), and multi-objective group methods for data handling (Bayat *et al.* 2011).

However, only a few studies have investigated water content predictability using different SWRC models with

parametric PTFs at different input levels. In particular, Khlosi *et al.* (2008) compared the water content predictability with eight SWRC models using the regression technique, while Ebrahimi *et al.* (2013) compared the water content predictability using ten SWRC models with the artificial neural networks technique.

The regression technique is a statistical method for estimating a relationship between a dependent variable and one or more independent variables (or predictors) (Freedman 2005). Regression is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning methods such as artificial neural networks. The technique is also used to understand which among independent variables are related to the dependent variable, and to explore the forms of these relationships (Armstrong 2012).

Nevertheless, it is still unclear whether changing the input variables influences the water content predictability when utilizing SWRC models based on the regression technique. Therefore, the present study investigated the interactions between the input variables and performance of different SWRC models when predicting the soil water content using the regression technique, as well as comparing the water content predictability with different SWRC models.

MATERIALS AND METHODS

Sampling and measurement of the PSD and aggregate size distribution

In this study, 75 disturbed and undisturbed soil samples were collected from Guilan province in Iran (Figure 1). The samples were taken from the top soils and subsoils at depths that depended on the thickness of the soil, i.e., from 10 to 35 cm and 20 to 45 cm, respectively.

Hydrometer and sieving methods were used to measure the PSD in ranges of 0–0.05 mm and 0.05–2.0 mm, respectively (Gee & Or 2002). A soil hydrometer (ASTM 152H; West Conshohocken, PA, USA) was used to obtain hydrometer readings at nine time points (0.5, 1, 3, 10, 30, 60, 90, 120, and 1,440 min). A dry sieve series with five diameters (0.05, 0.15, 0.25, 0.5, and 1.0 mm) was used to measure the distributions of sand particles. Bulk density (BD) was measured using the core method (Grossman & Reinsch 2002). The aggregate size distribution was measured in the same manner as the PSD, except dispersing agents were not used during the measurements and wet sieving was employed instead of dry sieving (Millan *et al.* 2007).

Measurement of the SWRC

A sand box apparatus (model 08.01; Eijkelkamp Co., The Netherlands) was used to measure the water content at matric suctions of 0, 1, 2, and 5 kPa. A pressure plate instrument (model 1600, 5 Bar; and model 1500F2, 15 Bar; ELE International Co., UK) was used to measure the water content at matric suctions of 10, 25, 50, 100, 500, 1,000, and 1,500 kPa (Dane & Jan 2002). The undisturbed and disturbed soil samples were used to obtain soil water content measurements at matric suctions of $\leq$100 kPa and $>$100 kPa, respectively.

SWRC models

Nine SWRC models were fitted to the experimental data obtained for each soil sample. The models were selected based on their simplicity, previously reported fitting accuracy, and novelty. These models are shown in Table 1. The models of Kosugi (1994), Groenevelt & Grant (2004), and Brooks & Corey (1964) were not estimated due to their large fitting errors.

Fitting the SWRC models

Various software packages were used to fit the models to the SWRC experimental data obtained for all 75 soil samples. The models of Durner (1994) and Seki (2007) were fitted to the SWRC data using a program written by Seki (2007). The van Genuchten (1980) model was fitted to the SWRC data using the RETC program (van Genuchten *et al.* 1991). The Gardner (1956) and Boltzman (McKee & Bumb 1984) models were fitted to the SWRC data using the DataFit program (Oakdale Engineering 2008). The models of Campbell (1974), Fredlund & Xing (1994), Fermi (McKee & Bumb 1987), and Dexter *et al.* (2008) were fitted to the SWRC data using the Solver function in the Excel program. The soil moisture unit was based on the volume ratio in all of the models, except for the Dexter model. In the latter case,

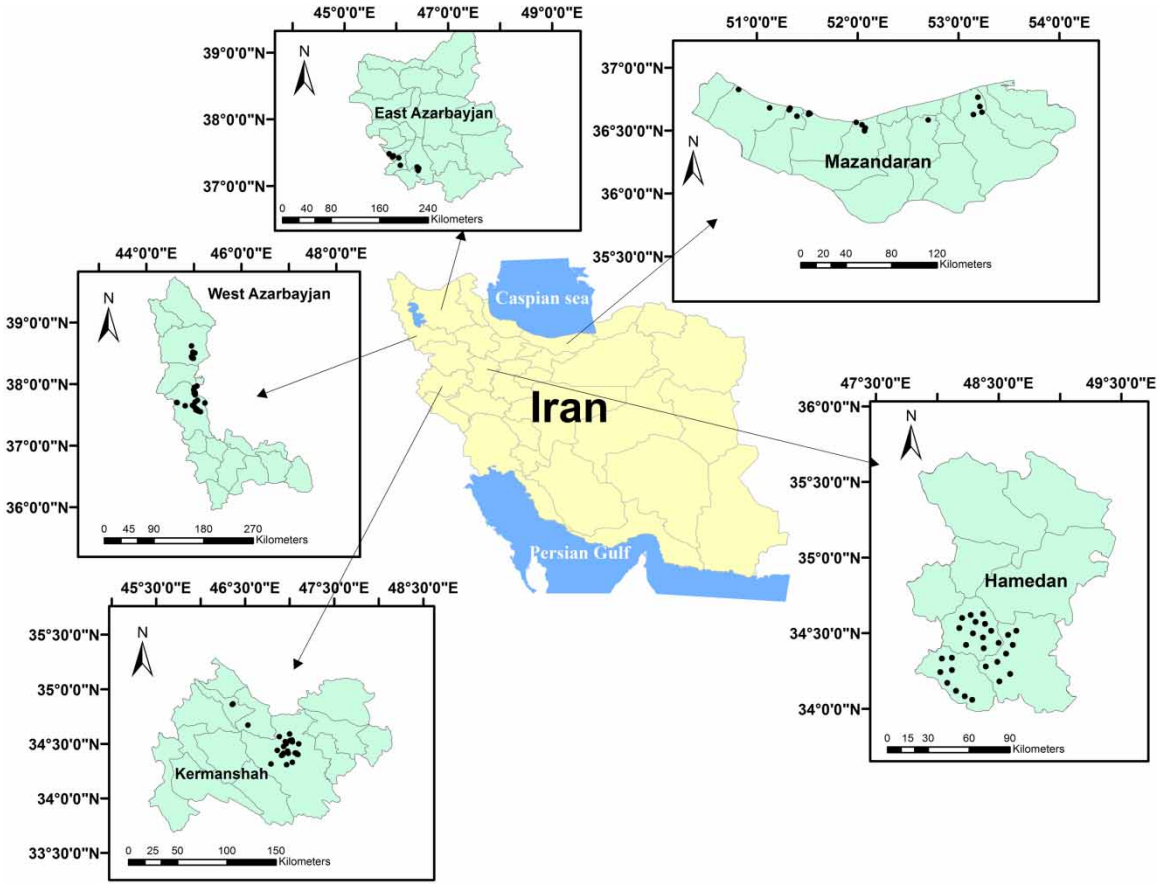

Figure 1 | Study area locations in Guilan province, Iran.

the moisture content was calculated using the Dexter model and converted into a volume ratio in order to compare the accuracy of the predictions obtained among all of the models (Dexter *et al.* 2008).

PSD and aggregate size distribution models

The fractal model of Perrier *et al.* (2002) was fitted to the PSD and aggregate size distribution data. The PSD and aggregate size distribution model parameters (fractal parameters) were calculated as follows:

$$m(x \le x_i) = (\alpha^{D-3})(L^{D-3})x_i^{3-D}, \quad (1)$$

where $m(x \le x_i)$ is the soil mass (or soil mass percentage) formed by units (primary particles, fragments, or micro-aggregates) with a characteristic diameter, x, which is smaller than or equal to a defined diameter, x_i; α is a similarity ratio, and L is the initiator size. After fitting the model above to the PSD and aggregate size distribution, their coefficients were denoted by the subscripts of PSD (D_{PSD}, α_{PSD}, and L_{PSD}) and Agg (D_{Agg}, α_{Agg}, and L_{Agg}), respectively.

Developing PTFs

Normality testing was performed on the variables, where the sand (%) and BD (gr cm^{-3}) were transformed to $sand^{0.5}$ and BD^2 to normalize them. The reason for this is their non-normal distribution and the fact that in developing PTFs using the regression method, all variables must be normally distributed. Whole soil samples were classified into fine and moderate textural groups based on the classification system employed by the United States Department of Agriculture. The textural classes of clay and silty clay were included in the fine textural group, and the textural classes of clay loam, silty clay loam, silt loam, loam, and sandy loam

were included in the moderate textural group. The predictive capacity was investigated using nine SWRC models for all of the soil samples, the fine textural group, and the moderate textural group.

The detailed process employed for estimating the SWRC using the nine SWRC models and their comparison is depicted in Figure 2. In order to develop the PTFs, the coefficients of SWRC models were introduced one by one as dependent variables in DataFit 9 (2008), where clay, sand, BD, and the PSD and aggregate size distribution model parameters were used as the input variables. DataFit is a science and engineering tool that simplifies the tasks of data plotting, regression analysis (curve fitting), and statistical analysis. What sets DataFit apart from similar curve fitting and regression programs is its ease of use (Oakdale Engineering 2008).

The silt content was not used as a predictor in order to avoid multicollinearity, and multicollinearity was examined in all of the regression models. Approximately 50% of the soil samples ($n = 38$) were selected randomly for training and the remaining 50% ($n = 37$) were used for testing. In the training step, regression PTFs (models) were developed using DataFit. Training involves determining the coefficients of regression equations to minimize the differences between observed and predicted values of output variables for the training data set (Konikow & Bredehoeft 1992). Testing is a comparison of model results with observations for a data set that is independent from those used in the training. As such, testing demonstrates the ability of a model to represent cause-and-effect relations for a particular data set. In fact, testing evaluates how well a model represents the real system (Konikow & Bredehoeft 1992).

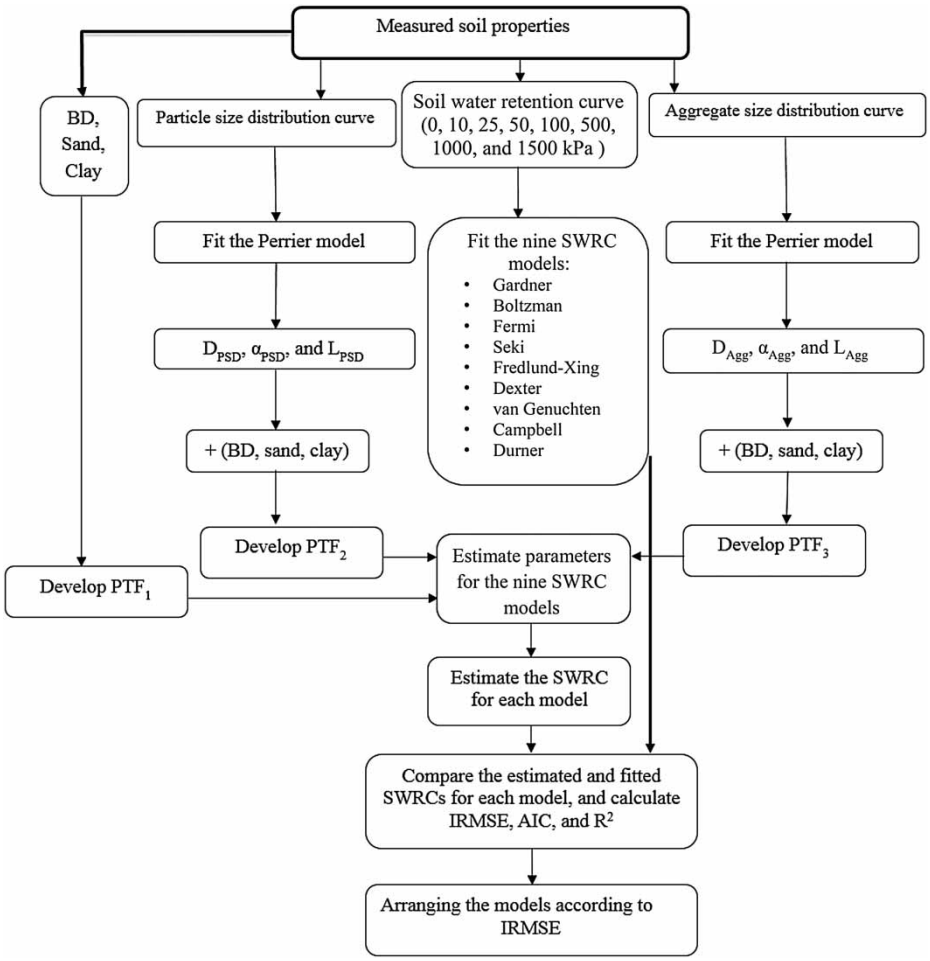

Figure 2 | Detailed process employed to estimate the SWRC using nine SWRC models and their comparison.

The prediction of each coefficient was performed at three levels with different inputs. The input variables for each level were as follows:

- Input variables for level 1: clay, sand, and BD.
- Input variables for level 2: clay, sand, BD, and the three parameters in the PSD model (D_{PSD}, α_{PSD}, and L_{PSD}).
- Input variables for level 3: clay, sand, BD, and the three parameters in the aggregate size distribution model (D_{Agg}, α_{Agg}, and L_{Agg}).

In the next step, the estimated values of the coefficients in each SWRC model for each level were used to calculate the estimated SWRC for the corresponding level. The estimated SWRC for the three input levels were then compared curve-by-curve with each other and the measured SWRC. 'Fitting accuracy' and 'prediction accuracy' imply the correspondence between measured-fitted and fitted-predicted values of water content (SWRC), respectively.

Cluster analysis

In this study, cluster analysis was performed to classify the SWRC models within the groups (clusters) in such a way that the SWRC models in the same group (called a cluster) are more similar in terms of IRMSE, AIC, and R^2 to each other than to those in the other groups (clusters). Therefore, the SWRC models that were classified in the same class are similar to each other and can be used for the same purposes (e.g., estimating water content). The SWRC models were partitioned into more homogeneous groups or clusters according to the accuracy of the water content estimates obtained by the models. The grouping of the SWRC models was based on evaluation criteria for the PTFs, i.e., Akaike's information criterion (AIC) (Akaike 1974), coefficient of determination (R^2), and integrated root mean squared error (IRMSE) (Tietje & Tapkenhinrichs 1993), using cluster analysis in SPSS 16. SPSS provides powerful statistical analysis, data and text mining, predictive modeling, and decision optimization using descriptive menus and simple dialog boxes to help users anticipate change and take action to improve outcomes (SPSS Inc. 2007). Ward's clustering method with squared Euclidian distance was used to group all of the models with the aim of minimizing the within-cluster variances and the squared Euclidian distances between groups. For example, if there are two clusters (groups) of SWRC models, it means that there are two groups of SWRC models with the minimum within-cluster (group) variance.

Evaluation criteria

Three criteria, i.e., AIC (Akaike 1974), coefficient of determination (R^2), and IRMSE (Tietje & Tapkenhinrichs 1993), were used to evaluate the accuracy of the predictions. AIC consider the number of parameters and penalizes models with additional parameters; therefore, AIC facilitates model selection based on parsimony (Ungaro *et al.* 2005). The model output is more accurate when the AIC value is smaller. The validation of the models was based on the average of the evaluation criteria in the testing and training steps:

$$IRMSE = \left[\frac{1}{b-a}\int_a^b (\theta_p - \theta_m)^2 d\psi\right]^{0.5} \tag{2}$$

$$R^2 = \left[\frac{\left[N\left(\sum_{i=1}^{N}\theta_m\theta_P\right) - \sum_{i=1}^{N}\theta_m\sum_{i=1}^{N}\theta_P\right]}{\sqrt{\left[N\left(\sum_{i=1}^{N}{\theta_m}^2\right) - \left(\sum_{i=1}^{N}\theta_m\right)^2\right]\times\left[N\left(\sum_{i=1}^{N}\theta_P^2\right) - \left(\sum_{i=1}^{N}\theta_P\right)^2\right]}}\right]^2 \tag{3}$$

$$AIC = N\ln\left(\frac{\sum(\theta_p - \theta_m)^2}{N}\right) + 2q \tag{4}$$

where θ_m and θ_p are the observed and predicted water contents, respectively (water content in the integration (Equation (2)) intervals from a to b), q is the number of model parameters, and N is the number of observations for each curve.

The root mean square error (RMSE) or IRMSE has been used as a standard statistical metric to measure model performance in many research studies. The RMSE measures total-error including bias. Therefore, it is more appropriate to represent model performance (Chai & Draxler 2014).

Since the RMSE criterion is more reliable than R^2 and low values of R^2 do not necessarily show the increment of the error in the model (Kozak *et al.* 1996), we focused on IRMSE (and to some extent AIC) for the interpretation of the result. However, a combination of metrics are often required to assess model performance (Chai & Draxler 2014).

RESULTS AND DISCUSSION

Table 2 shows the statistical analysis of the input variables and the parameters of the SWRC models. High standard deviations were observed for the clay and sand contents, thereby demonstrating the high variability in soil texture among the samples analyzed. The mean sand content was lower than the mean clay content.

Describing the correlations between the input and output variables

The correlation coefficients between the predictor variables and the parameters in the SWRC models are shown in Table 3. This section focuses on the correlations between the inputs and outputs for the SWRC models that obtained the best SWRC estimates.

The θ_r and n parameters of the Fermi model had significant correlations ($P < 0.05$) with the parameters of the aggregate size distribution model (α_{Agg} and L_{Agg}). Sillers *et al.* (2001) reported that the n parameter in the Fermi model is related to the shape of the SWRC. Similarly, Dexter (2004) showed that the aggregate size distribution has a significant effect on the shape of the SWRC. These correlations demonstrate the close relationship between the soil texture and structure, and water retention. The model of Perrier *et al.* (2002) performed well in quantifying the soil texture and structure. In addition, there was a negative relationship between θ_s in the van Genuchten model and BD. This may be attributable to the fact that increasing the BD eliminates larger soil pores and θ_s is related to larger pores (Sillers *et al.* 2001; Tuller & Or 2003).

The θ_r parameter in the Boltzman model had a significant relationship with all of the predictors, except for the parameters of the aggregate size distribution model. There were significant relationships between θ_s in the Seki model with clay, α_{PSD}, $D_{PSD,}$ L_{PSD}, and D_{Agg}. Correspondingly, there were significant correlations between w_1 in the Seki model with L_{Agg}, and σ_2 in the Seki model with $L_{PSD.}$ These correlations may be associated with the similarity between the particle, aggregate, and pore size distributions, as reported by Alemi (1981).

General results of the cluster analysis

C columns in Tables 4–6 show the cluster analysis result. The values of 1, 2, 3, and 4 in the C column show the groups 1, 2, 3, and 4, respectively. The reason for different group numbers in various input levels in Tables 4–6 is the closeness or the difference between the IRMSE of the SWRC models.

In most of the input levels of all the textural groups (all soil samples: fine and moderate textural groups), the Durner and/or Dexter bimodal models were classified into separate groups (Tables 4–6) with the lowest prediction accuracy (Tables 4–6). The models of Gardner, Fermi, Campbell, and Seki, and to some extent Boltzman, were classified mostly in the same group (Tables 4–6) with the highest prediction accuracy (Tables 4–6), thereby demonstrating their suitability for the development of parametric PTFs. Because, by considering the values in C columns (group number) of Tables 4–6, it is clear that the group number for the models of Gardner, Fermi, Campbell, and Seki, and to some extent Boltzman, were mostly 1, it means that they were classified mostly in the same group (Tables 4–6). Since, the above-mentioned models have mostly the least IRMSE, then they have the highest accuracy.

In most of the clusters (except for the fine textural group), the van Genuchten and Fredlund–Xing models were classified into the groups that differed from those with the lowest or highest accuracy, thereby indicating the superiority of other models such as Fermi and Seki for developing parametric PTFs compared with the van Genuchten and Fredlund–Xing models.

Cluster analysis results for the three input levels of the fine textural group soil samples (Table 5) showed that only the Durner or Dexter models with the lowest prediction accuracy were classified into separate classes,

Table 2 | Statistical analysis of nine SWRC models, including the parameters and soil properties (i.e., input variables) for all samples

Model	Parameter	Mean	Min	Max	SD	Unit
Dexter	c[a]	0.173	0.000	0.503	0.08	$grgr^{-1}$
	A_1	0.354	0.232	8.444	0.97	$grgr^{-1}$
	ψ_1	2,426	479	8,790	4,193	kPa
	A_2	0.069	0.000	0.530	0.95	$grgr^{-1}$
	ψ_2	74.22	1.430	1,833	228	kPa
Fermi	θ_r	0.177	0.062	0.345	0.06	cm^3cm^{-3}
	θ_s	0.566	0.341	0.781	0.12	cm^3cm^{-3}
	α	648	1.005	3,085	1,651	kPa
	n	923	39	4,308	824	
van Genuchten	θ_r	0.056	0.000	0.448	0.09	cm^3cm^{-3}
	θ_s	0.571	0.252	0.710	0.15	cm^3cm^{-3}
	α	0.014	0.000	0.200	0.03	kPa^{-1}
	n	1.643	1.040	6.560	1.05	
Boltzman	θ_r	0.185	0.060	0.400	0.05	cm^3cm^{-3}
	θ_s	0.548	0.240	0.800	0.10	cm^3cm^{-3}
	α	169.0	1.380	1,647	221	kPa
	n	1,451	6.390	4,623	1,007	
Gardner	θ_r	0.164	0.006	2.000	0.12	cm^3cm^{-3}
	θ_s	0.544	0.250	0.790	0.24	cm^3cm^{-3}
	α	3.718	0.003	9.380	2.56	kPa^{-1}
	n	1.026	0.250	3.750	0.79	
Fredlund-Xing	θ_s	0.532	0.252	0.776	0.01	cm^3cm^{-3}
	α	2,780	12.99	63,557	9,133	kPa
	n	0.993	0.273	4.097	5.93	
	m	1.421	0.111	5.778	6.98	
Campbell	θ_s	0.55	0.30	0.73	0.10	cm^3cm^{-3}
	α	15.70	3.72	49.32	7.91	kPa
	n	0.10	0.01	0.37	0.05	
Seki	θ_r	0.175	0.018	0.380	0.07	cm^3cm^{-3}
	θ_s	0.552	0.250	0.730	0.09	cm^3cm^{-3}
	w_1	0.487	0.017	0.999	0.26	
	Ψm_1	864	10.22	12,248	1,736	
	σ_1	9.215	0.058	428	55	
	Ψm_2	7,773	8.120	8,322	1,791	
	σ_2	4.511	0.050	48.77	10.46	
Durner	θ_r	0.170	0.050	0.400	0.07	cm^3cm^{-3}
	θ_s	0.549	0.250	0.730	0.09	cm^3cm^{-3}
	w_1	0.452	0.017	0.890	0.23	
	α_1	1.688	0.010	42.67	7.32	kPa^{-1}
	n_1	6.499	1.056	113	17.80	
	α_2	5.455	0.013	9.680	2.73	kPa^{-1}
	n_2	10.32	1.200	49.37	12.01	
Soil properties (i.e., input variables)	Clay	33	13	58	11.20	%
	Sand	17	1	59	10.03	%
	BD	1.285	0.733	1.704	0.24	$grcm^{-3}$
	A_{psd}	0.793	0.631	1.226	0.09	
	D_{psd}	0.867	0.738	0.936	0.04	
	L_{psd}	0.792	0.630	1.225	0.09	
	A_{agg}	1.177	0.654	2.094	0.20	
	D_{agg}	2.814	2.707	2.954	0.07	
	L_{agg}	1.178	0.652	2.095	0.20	

[a]All of the abbreviations are defined in the Nomenclature.

Table 3 | Matrix of Pearson's correlation coefficients (r) between the predictors and output variables

Model	Parameter	Sand	Clay	BD	α_{PSD}	D_{PSD}	L_{PSD}	α_{Agg}	D_{Agg}	L_{Agg}
Dexter	c[a]	−0.22	0.23	0.03	−0.26**	0.18	−0.18	−0.29	0.16	0.04
	A_1	−0.01	−0.04	0.15	−0.01	−0.02	−0.02	−0.04	0.03	−0.03
	ψ_1	−0.11	0.05	−0.05	−0.15	0.09	−0.08	0.36**	0.06	−0.03
	A_2	0.01	0.02	−0.22	0.02	−0.01	0.01	0.06	−0.03	0.01
	ψ_2	−0.11	0.04	−0.03	−0.15	0.08	−0.08	0.02	0.05	−0.04
Fermi	θ_r	−0.07	0.20	0.04	−0.06	0.14	−0.06	0.33*	0.09	0.32*
	θ_s	−0.06	0.04	0.19	0.02	0.03	0.02	−0.25	0.16	−0.25
	α	−0.16	0.11	0.02	−0.19	0.13	−0.19	−0.26	−0.24	−0.27
	n	−0.12	0.15	0.20	−0.15	0.17	−0.15	−0.30*	0.04	−0.30
van Genuchten	θ_r	−0.15	0.22	−0.10	−0.22	0.27	−0.21	−0.08	0.14	0.21
	θ_s	−0.06	−0.08	−0.35*	0.00	−0.02	0.00	−0.04	−0.03	0.19
	α	0.013	0.18	0.25	0.07	0.12	0.08	−0.04	−0.12	0.12
	n	0.13	−0.22	0.03	0.12	−0.17	0.13	−0.23	−0.16	0.13
Boltzman	θ_r	−0.38**	0.63**	0.30**	–0.41**	0.54**	–0.41**	0.09	0.31	0.05
	θ_s	−0.11	−0.11	0.02	−0.05	−0.09	−0.05	−0.14	−0.28	−0.17
	α	0.01	0.01	0.18	−0.06	0.04	−0.05	−0.06	−0.04	−0.01
	n	−0.01	−0.01	0.04	−0.12	0.03	−0.12	−0.26	−0.09	−0.24
Gardner	θ_r	−0.14	0.11	−0.02	−0.15	0.12	−0.15	0.00	−0.21	0.02
	θ_s	−0.04	−0.07	0.03	−0.09	−0.07	−0.09	−0.16	−0.19	−0.23
	α	−0.01	−0.08	0.19	−0.09	−0.03	−0.09	−0.15	0.02	−0.27
	n	−0.09	0.01	−0.13	−0.08	0.04	−0.08	−0.01	−0.13	0.70
Fredlund–Xing	θ_s	−0.13	0.03	0.11	−0.14	0.03	−0.14	−0.13	−0.09	−0.17
	α	−0.10	0.22	0.02	−0.17	0.17	−0.17	−0.05	−0.09	−0.07
	n	0.01	−0.03	0.05	0.05	0.00	0.05	0.03	−0.11	0.08
	m	0.17	−0.12	−0.12	0.08	−0.15	0.07	−0.10	−0.21	−0.20
Campbell	θ_s	−0.13	0.07	0.05	−0.20	0.06	−0.20	−0.20	−0.14	−0.16
	α	−0.05	0.08	−0.03	−0.06	0.08	−0.05	0.06	0.09	−0.07
	n	0.05	0.00	−0.17	0.05	0.00	0.05	−0.20	−0.13	−0.07
Seki	θ_r	−0.20	0.010	0.00	−0.24	0.11	−0.24	0.17	−0.25	0.00
	θ_s	−0.22	0.42**	0.22	−0.29*	0.37**	–0.29*	0.06	0.40**	0.00
	w_1	0.18	−0.10	0.10	0.21	−0.04	0.23	0.26	0.13	0.38*
	Ψm_1	−0.06	0.17	0.09	−0.08	0.12	−0.08	0.13	0.04	0.15
	σ_1	0.08	−0.13	−0.24	0.07	−0.10	0.07	−0.08	−0.07	−0.04
	Ψm_2	−0.14	0.11	0.15	−0.09	0.08	−0.09	0.14	0.05	0.10
	σ_2	0.20	0.00	0.30	0.26	−0.13	0.27*	−0.02	0.21	−0.01
Durner	θ_r	−0.19	0.13	−0.03	−0.25	0.14	−0.25	0.21	−0.26	0.02
	θ_s	−0.08	0.22	0.08	−0.14	0.17	−0.14	−0.14	0.32*	−0.19
	w_1	−0.16	0.02	−0.9	−0.18	0.14	−0.17	0.11	−0.06	0.28
	α_1	0.00	−0.13	−0.12	−0.07	0.10	−0.07	0.09	0.22	0.09
	n_1	0.12	−0.09	0.05	0.18	−0.17	0.18	−0.08	−0.22	−0.04
	α_2	−0.24	0.02	−0.24	−0.18	0.14	−0.18	0.32*	−0.07	0.43**
	n_2	−0.09	0.09	0.14	−0.06	0.13	−0.06	0.13	−0.02	0.20

* and ** denote significant correlations at $P < 0.05$ and $P < 0.01$, respectively.

[a]All of the abbreviations are defined in the Nomenclature.

whereas all of the other models were classified into the same class. Therefore, in contrast to all of the soil samples and the moderate textural groups, the SWRC models had little effect on SWRC estimation by PTFs for the fine textural soils. Thus, users should be more careful when selecting SWRC models for developing parametric PTFs using all soil samples and moderate textural groups.

Table 4 | Accuracy of the water content estimates obtained by the SWRC models for all soil samples

Level 1					Level 2					Level 3				
Model	IRMSE[$,#] ($cm^3cm^{-3}$)	R^2	AIC[b]	C	Model	IRMSE (cm^3cm^{-3})	R^2	AIC	C	Model	IRMSE (cm^3cm^{-3})	R^2	AIC	C
Gardner	0.096[a]	0.895	−1,095	1	Gardner	0.086[a]	0.939	−1,713	1	Gardner	0.089[a]	0.936	−1,666	1
Boltzman	0.173[ab]	0.924	−1,289	1	Campbell	0.120[ab]	0.996	−1,368	1	Campbell	0.109[ab]	0.999	−1,428	1
Fermi	0.177[ab]	0.789	−925	2	Fermi	0.131[abc]	0.849	−1,172	2	Fermi	0.121[ab]	0.862	−1,142	2
Seki	0.206[bc]	0.779	−768	2	Seki	0.134[abc]	0.785	−1,336	2	Seki	0.174[bc]	0.838	−905	2
Fredlund-Xing	0.237[bc]	0.943	−1,096	3	Fredlund–Xing	0.158[bc]	0.834	−451	3	Boltzman	0.198[c]	0.738	−1,188	3
Dexter	0.247[bc]	0.809	−908	2	van Genuchten	0.166[bc]	0.827	−966	2	Fredlund–Xing	0.217[c]	0.865	−567	2
van Genuchten	0.256[bc]	0.955	−1,069	3	Boltzman	0.178[c]	0.955	−1,269	1	van Genuchten	0.241[c]	0.870	−768	2
Campbell	0.265[c]	0.996	−950	3	Dexter	0.245[d]	0.701	−878	3	Dexter	0.541[d]	0.788	−318	4
Durner	0.508[d]	0.938	−353	4	Durner	0.431[e]	0.534	−373	4	Durner				

Lower case letters compare the prediction accuracy of the SWRC models based on the IRMSE, within the related input level. Where letters are different (within each column), means were significantly different at the 0.05 level.
[$]IRMSE is the integrated root mean squared error, R^2 is the coefficient of determination, and AIC is Akaike's information criterion.
C columns show the cluster analysis result. The values of 1 and 2, in the C column show the groups 1 and 2, respectively.
[#]The values of these criteria (IRMSE, AIC, and R^2) are the averages from the training and testing steps.

Table 5 | Accuracy of the water content estimates obtained by the SWRC models for the fine textural group

Level 1					Level 2					Level 3				
Model	IRMSE[$,#] ($cm^3cm^{-3}$)	R^2	AIC[b]	C	Model	IRMSE (cm^3cm^{-3})	R^2	AIC	C	Model	IRMSE (cm^3cm^{-3})	R^2	AIC	C
Gadner	0.113[a]	0.872	−1,520	1	Gadner	0.101[a]	0.924	−1,569	1	Gadner	0.090[a]	0.939	−1,739	1
Boltzman	0.145[a]	0.938	−1,293	1	Campbell	0.118[a]	0.997	−1,410	1	Campbell	0.100[a]	0.999	−1,529	1
Fermi	0.213[ab]	0.740	−833	1	Fermi	0.139[ab]	0.895	−1,138	1	Fermi	0.149[a]	0.906	−996	1
Dexter	0.223[ab]	0.853	−943	1	Seki	0.140[ab]	0.733	−1,349	1	Boltzman	0.169[a]	0.719	−1,232	1
van Genuchten	0.226[ab]	0.957	−1,156	1	van Genuchten	0.148[ab]	0.871	−1,035	1	Seki	0.183[a]	0.832	−882	1
Seki	0.239[ab]	0.731	−702	1	Boltzman	0.154[ab]	0.952	−1,564	1	van Genuchten	0.188[a]	0.891	−934	1
Fredlund–Xing	0.246[ab]	0.967	−1,049	1	Fredlund–Xing	0.162[ab]	0.874	−555	1	Fredlund–Xing	0.230[a]	0.889	−544	1
Campbell	0.373[b]	0.997	−770	1	Dexter	0.247[b]	0.690	−914	1	Dexter	0.582[b]	0.725	−315	1
Durner	0.788[c]	0.926	−75	2	Durner	0.506[c]	0.476	−312	2	Durner				2

Lower case letters compare the prediction accuracy of the SWRC models based on IRMSE, within the related input level. Where letters are different (within each column), means were significantly different at the 0.05 level.
[$]IRMSE is the integrated root mean squared error, R^2 is the coefficient of determination, and AIC is Akaike's information criterion.
C columns show the cluster analysis result. The values of 1, 2 and 3 in the C column show the groups 1, 2 and 3, respectively.
[#]The values of these criteria (IRMSE, AIC, and R^2) are the averages from the training and testing steps.

Table 6 | Accuracy of the water content estimates obtained by the SWRC models for the moderate textural group

Level 1					Level 2					Level 3				
Model	IRMSE[$,#] ($cm^3cm^{-3}$)	R^2	AIC[b]	C	Model	IRMSE (cm^3cm^{-3})	R^2	AIC	C	Model	IRMSE (cm^3cm^{-3})	R^2	AIC	C
Gardner	0.087[a]	0.907	−1,533	1	Gardner	0.080[a]	0.945	−1,870	1	Gardner	0.088[a]	0.935	−1,643	1
Fermi	0.159[b]	0.808	−983	2	Campbell	0.119[ab]	0.995	−1,357	1	Fermi	0.109[a]	0.842	−1,248	1
Boltzman	0.185[bc]	0.921	−1,288	2	Fermi	0.127[abc]	0.828	−1,192	1	Campbell	0.112[a]	0.999	−1,397	1
Seki	0.192[bc]	0.798	−797	2	Seki	0.134[abc]	0.813	−1,323	1	Seki	0.168[ab]	0.844	−918	2
Campbell	0.208[bc]	0.995	−1,161	2	Fredlund–Xing	0.157[bc]	0.819	−426	2	Boltzman	0.207[bc]	0.745	−1,174	2
Fredlund–Xing	0.231[bc]	0.936	−1,123	2	Boltzman	0.184[c]	0.956	−1,271	2	Fredlund–Xing	0.212[bc]	0.855	−574	2
Dexter	0.235[c]	0.787	−956	2	van Genuchten	0.186[c]	0.783	−905	2	van Genuchten	0.270[c]	0.858	−738	2
van Genuchten	0.243[c]	0.954	−1,072	2	Dexter	0.257[d]	0.705	−861	2	Dexter	0.555[d]	0.817	−307	3
Durner	0.369[d]	0.947	−794	3	Durner	0.384[e]	0.548	−432	3	Durner				

Lower case letters compare the prediction accuracy of the SWRC models based on the evaluation criteria (IRMSE, AIC, and R^2), within the related input level. Where letters are different (within each column), means were significantly different at the 0.05 level.

[$]IRMSE is the integrated root mean squared error, R^2 is the coefficient of determination and AIC is Akaike's information criterion.

C columns show the cluster analysis result. The values of 1, 2, 3, and 4 in the C column show the groups 1, 2, 3, and 4, respectively.

[#]The values of these criteria (IRMSE, AIC, and R^2) are the averages from the training and testing steps.

Comparison of the predictive capacity of SWRC models with all of the soil samples

The fitting accuracy for the various SWRC models are shown in Table 7. Different software was used to fit various SWRC models. The different software may affect the fitting accuracy of different models, but their effect is negligible. Some software, such as DataFit and Solver in MS Excel, can be used to fit all regression models, and using them, the fitting accuracy of different models was checked to ensure that the best fitting accuracies were obtained. In addition, Sillers *et al.* (2001) did not report any figure for the Gardner model, which may show lack of fitting in that particular model on the SWRC data.

The results of our unpublished study showed that texture classes slightly affected the fitting accuracy of SWRC models, in a way that the best and poor performing models were almost the same in different texture classes.

Table 4 shows the prediction errors of the models for the overall data set. The Gardner model had the highest accuracy at all the three input levels. The Gardner model is a continuous function with four parameters (Gardner 1956). The SWRC predicted by the Gardner model was very similar to the fitted curve, thereby yielding very low IRMSE values; however, the Gardner model was not an appropriate model for the soils because of its low fitting accuracy (Table 7). This incompatibility occurred because the fitted and estimated moisture contents were lower than the measured values (Figures 3–5). This result agrees with the findings of Bayat *et al.* (2013a) who also reported a low fitting accuracy for the Gardner model.

Table 7 | The fitting accuracy of the nine SWRC models

Model	$R^{2\#}$	RMSE (cm^3cm^{-3})
Dexter	0.968 (0.119)$^{a\$}$	0.021 (0.043)^{a}
Fredlund–Xing	0.974 (0.023)^{a}	0.025 (0.039)^{a}
Campbell	0.841 (0.086)^{b}	0.041 (0.026)^{ab}
Seki	0.989 (0.027)^{a}	0.067 (0.067)^{bcd}
Boltzman	0.961 (0.068)^{a}	0.072 (0.059)^{bcd}
Fermi	0.914 (0.105)^{a}	0.091 (.107)^{bcd}
van Genuchten	0.956 (0.081)^{a}	0.094 (0.079)^{bcd}
Durner	0.831 (0.100)^{b}	0.120 (0.079)^{cd}
Gadner	0.506 (0.217)^{c}	0.308 (0.116)^{e}

$^{\$}$Lower case letters compare the fitting accuracy of the SWRC models based on the evaluation criteria (RMSE and R^2). Where letters are different, means were significantly different at the 0.05 level.

$^{\#}$The values of these criteria (RMSE and R^2) are the averages for all soil samples. The values inside the parentheses show the standard deviations of the evaluation criteria for all soil samples.

The Campbell model had good accuracy (not significantly, $P < 0.05$, different from that of Gardner model) at the second and third input levels, but this model did not obtain accurate predictions at the first input level. The same result was obtained with the fine texture group. The Durner model parameters were not predicted for the third input level due to overestimation. The bimodal models of Seki, Durner, and Dexter did not yield satisfactory accuracy, which may have been due to the pore size distribution in the soil samples. Seki (2007) demonstrated that the fitting accuracy of the bimodal models of Seki and Durner is suitable for soils with a heterogeneous pore size distribution. Furthermore, Seki (2007) reported that the fitting accuracy of his own model was better than that of the Durner model.

After the cluster analysis, the Durner model was classified into a separate class in the first and second levels (Table 4) and its prediction accuracy was significantly ($P < 0.05$) lower than those of other models. The Fermi model was more accurate than the Boltzman model in the second and third input levels (Table 4 and Figure 3), and their difference was significant ($P < 0.05$) in the third input level, whereas the opposite was true in the first input level. The superiority of the Fermi model may be attributed to the simple form of its equation. This finding agrees with the results obtained when these two models were fitted by McKee & Bumb (1984, 1987). The results showed that the accuracy of the water content predictions obtained by the Boltzman model decreased as the number of inputs increased. The cluster analysis confirmed these results, and the Fermi and Boltzman models were classified into various clusters in all three input levels (Table 4).

Similar to the Brooks–Corey model, the Boltzman model does not describe a continuous function for the SWRC. It is possible that the good water content predictions obtained by the Boltzman model in the first input level were due to the high correlations between its parameters and the predictors. Nevertheless, the water content was estimated with sufficient accuracy by the Boltzman model in selected soil samples, as shown in Figures 3–5.

The Dexter model is a bimodal model and its parameters have physical meaning. The water contents were predicted

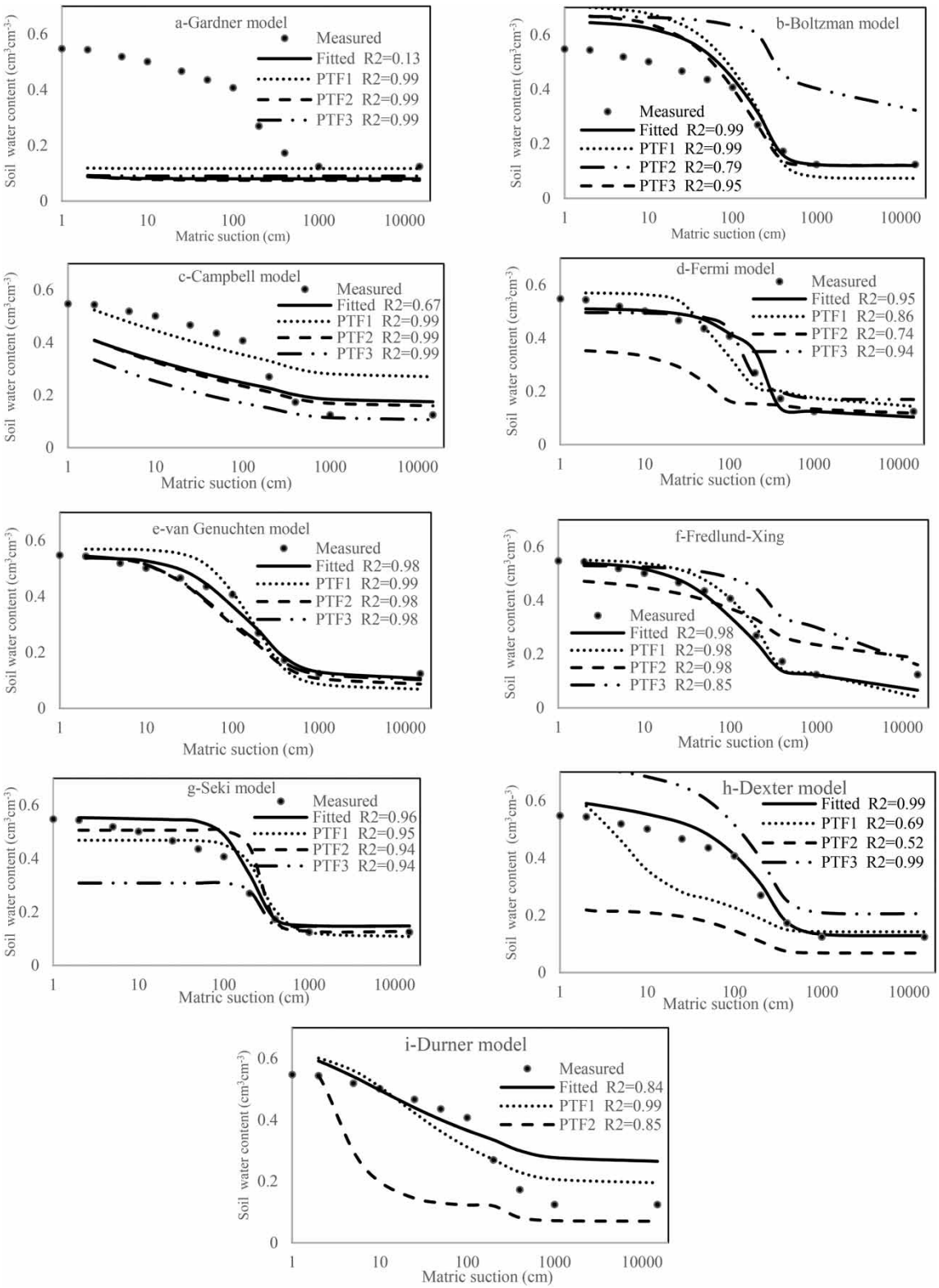

Figure 3 | Comparison of the measured, fitted, and predicted SWRC models for a silt loam soil.

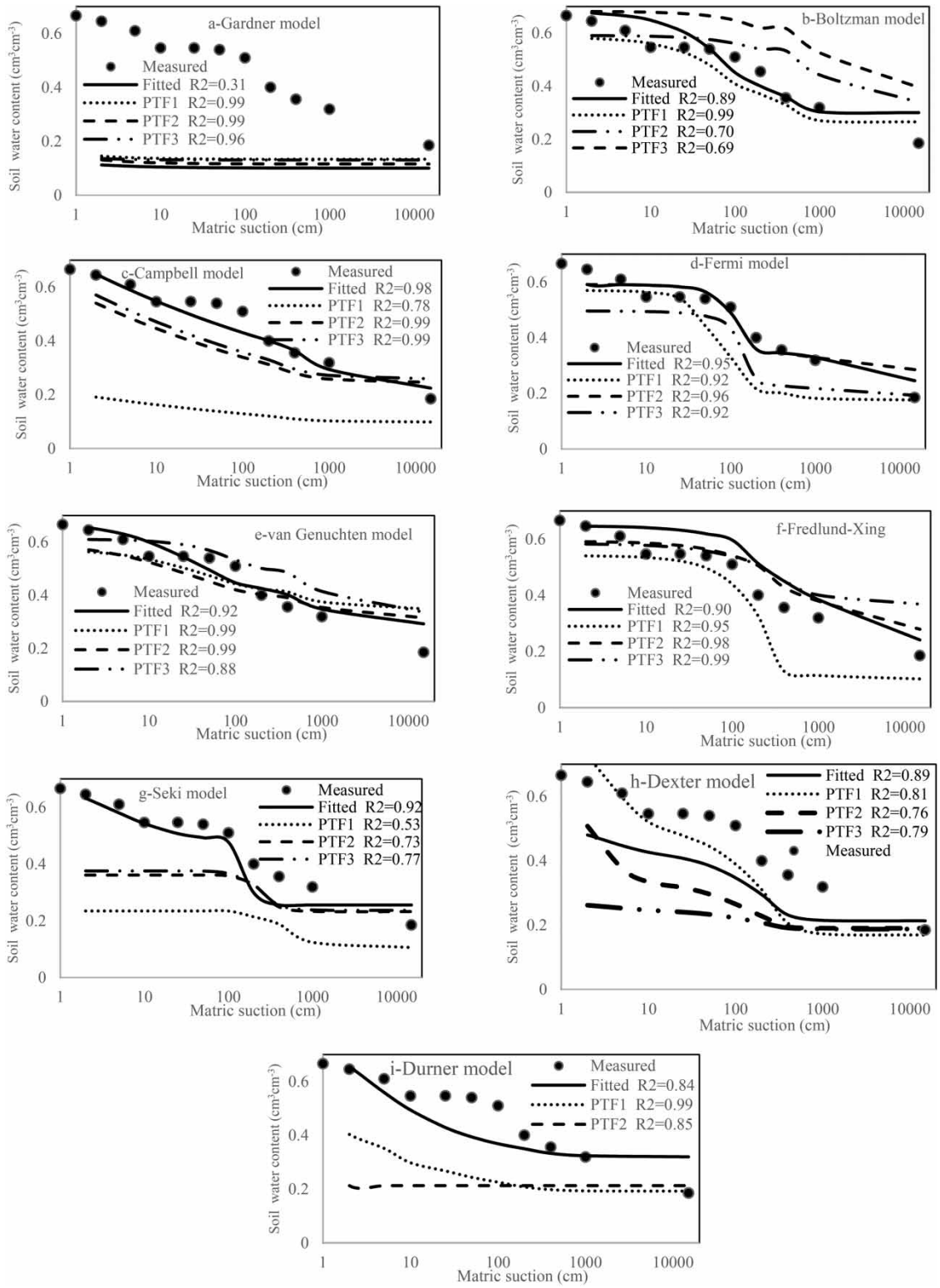

Figure 4 | Comparison of the measured, fitted, and predicted SWRC models for a clay soil.

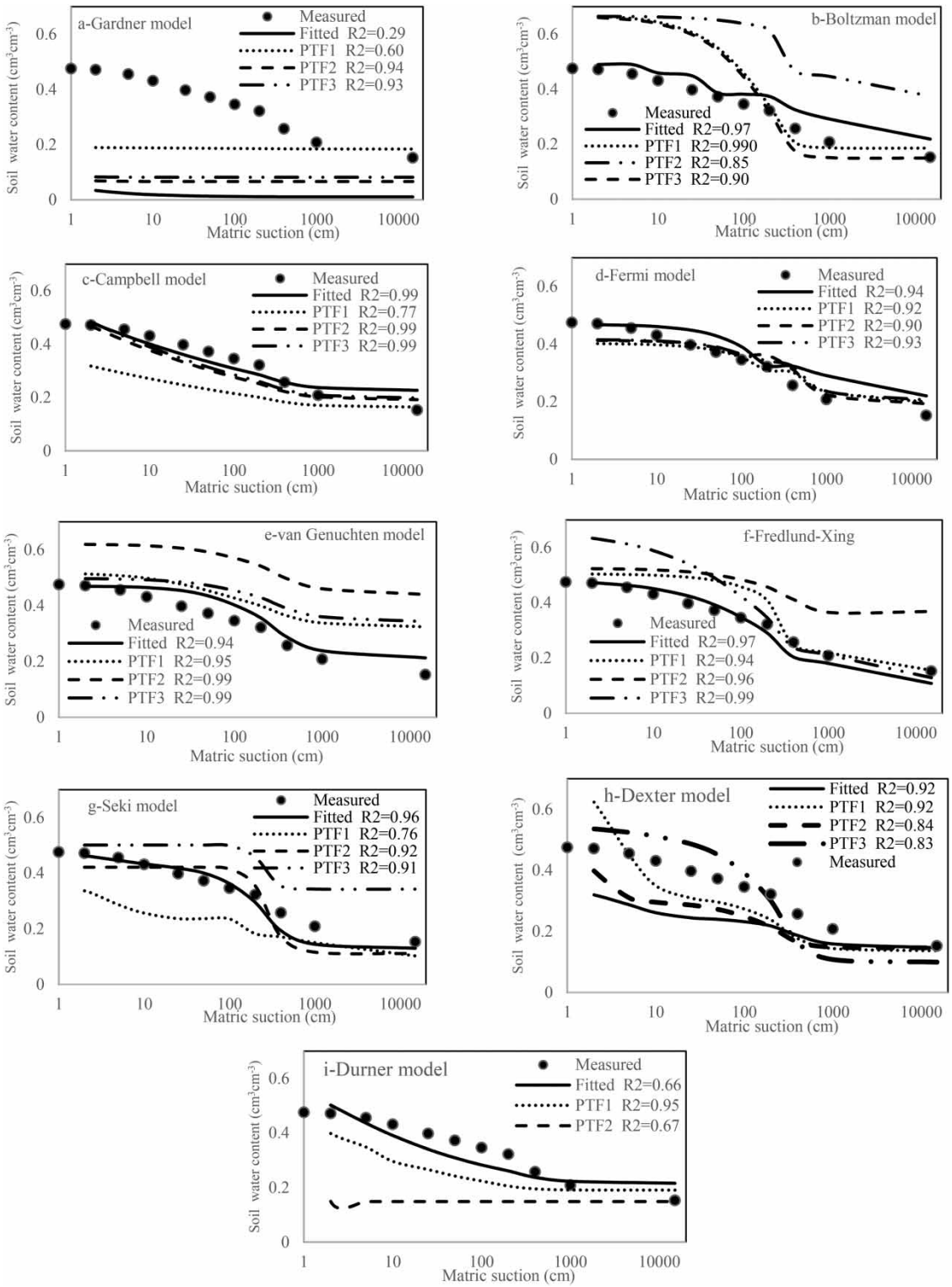

Figure 5 | Comparison of the measured, fitted, and predicted SWRC models for a clay loam soil.

more accurately by the Dexter model in the first input level compared with the other two input levels (Table 4 and Figure 3), which may be attributable to the use of a unimodal model for the PSD and aggregate size distribution, thereby leading to differences between the particle, aggregate, and pore size distributions. The cluster analysis classified the Dexter model into a separate class at the second and third input levels, whereas the Dexter model was classified into the same class as the Seki and Fermi models at the first input level (Table 4). This result demonstrates that the accuracy of predictions is affected by changing the input variables.

Overall, the van Genuchten model did not obtain better results compared with the other models. However, the best predictions by the van Genuchten model among the three input levels were obtained when using the parameters of the PSD model as predictors. Clearly, one explanation for the weak results obtained by the van Genuchten model is the low correlation between its parameters and the predictors (Table 4). Schaap & Leij (1998) predicted the water contents using the van Genuchten model with artificial neural networks, where they obtained an RMSE value of 0.108 (cm^3cm^{-3}) (the mean value for calibration and independent data) when using the clay, silt, and sand contents as inputs. The Seki and Fermi models yielded the same results as the van Genuchten model, and thus they were classified into the same class at the second and third input levels (Table 4).

Tables 5–7 compare the mean values of the IRMSE statistic, which show that there were no significant ($P < 0.05$) differences between the Gardner, Fermi, and Boltzman models at the first input levels with the total data set and the fine textural group. The water contents were predicted with the lowest accuracy by the Campbell and Durner models, and these two models were placed in different classes (significant, $P < 0.05$) at the first input level for the total data set and the fine textural group.

Comparison of the predictive capacity of the SWRC models for the fine textural group

Similar to all of the soil samples (Table 5 and Figure 4), the Gardner model obtained the highest accuracy at all three input levels for the fine textural group, whereas the Durner model had the lowest accuracy (significant, $P < 0.05$) (Table 5 and Figure 4). According to the cluster analysis, all of the models except for the Durner or Dexter models were classified into the same class at the three input levels (Table 5), thereby demonstrating the lack of interaction between the SWRC models and the input levels when estimating the SWRC for fine textural soils. Thus, most of the SWRC models can be used to estimate the SWRC with parametric PTFs at each input level for these soils. Sillers *et al.* (2001) reported that the van Genuchten and Fermi models employ two fitting parameters (α and n), and these models were classified into the same class by the cluster analysis in the present study (Table 5) due to the similarity of their parameters. Patil *et al.* (2012) reported that the Campbell and van Genuchten models achieved the highest accuracy for clayey samples. The findings reported by Patil *et al.* (2012) agreed with the results obtained in the present study at the second and third input levels for the fine textural group (Table 5). However, the results of the present study demonstrated that the Gardner model obtained higher accuracy compared with the models of van Genuchten and Campbell, which agrees with the findings of Omuto (2009). The Campbell model employs three parameters and an abrupt air entry value is assumed, but the air entry value is gradual in the fine textured soils (Tuller & Or 2003). Thus, the assumption of an abrupt air entry value in the Campbell model may explain the higher accuracy of the Gardner model compared with the Campbell model.

Comparison of the predictive capacity of SWRC models for the moderate textural group

Table 6 compares the water content predictions obtained by different SWRC models for moderate texture soils. The Gardner and Durner models yielded the highest and lowest accuracy, respectively, at all three input levels. The models of Fredlund–Xing, Fermi, Boltzman, Seki, and van Genuchten had different degrees of accuracy, and thus their relative order of performance differed at various input levels. These results demonstrate that the input levels interacted with the SWRC models when predicting the soil water content. The results of the cluster analysis (Table 6) showed that the Fredlund–Xing and van Genuchten models were placed into the same class at all three input levels, thereby indicating the similar performance of the two models when estimating the water content (Fredlund & Xing 1994). Moreover, the Seki, Fredlund–Xing, and van Genuchten models were

placed into the same class at the third input level (Table 6), which agreed with the results in Table 6.

McKee & Bumb (1987) reported that the Fermi model is flexible, which may explain why the water content predictions obtained by the Fermi model with few parameters (four parameters) were comparable to the predictions of the Seki model with more parameters (seven parameters). Therefore, increasing the number of SWRC model parameters did not increase their predictability, so the water content predictions of SWRC models may differ in accuracy compared with that of their fitting. Thus, the water contents predicted by different SWRC models with PTFs may be influenced by the properties of soils and the SWRC models. For example, correlations among the SWRC model parameters and the input variables, as well as the form of the curve described by a model may have major effects on the predictions obtained.

Furthermore, the water content predictions produced by the Dexter and Durner models with a high number of parameters tended to be less accurate (significant, $P < 0.05$, at the second and third input levels) than those obtained by the other models considered in the present study. The water content predictions obtained by the Campbell model at the first input level with all of the samples used in this study were characterized by an R^2 value of 0.996 (Table 3), whereas Khlosi *et al.* (2008) obtained an R^2 value of 0.986 using the same model. Thus, the results obtained in the present study were better or comparable to the results reported previously. With most of the PTFs, the AIC results were the same as the IRMSE results. In most cases, the Seki model was superior to the bimodal Durner and Dexter models, and its predictions were considerably more accurate than those of the other bimodal models. This may be explained by the direct correlations between the parameters of the Seki model with the log-normal pore size distribution and the PSD (Seki 2007). However, the accuracy of the predictions obtained by the Seki model was only moderate compared with that of the other models.

Effects of the input levels on the water content predictions obtained by the SWRC models

Table 4 shows the effects of the input levels on the water content predictions produced by each model. At the second input level, the Gardner, van Genuchten, Fredlund–Xing, and Seki models obtained their best results, but they showed their lowest accuracy at the first input level. For example, the IRMSE values of the water content estimation by the Gardner model are 0.096, 0.086, and 0.089 cm^3cm^{-3} in the first, second, and third input levels, respectively (Table 4). Then, at the second input level, the Gardner model obtained the best result (IRMSE = 0.086 cm^3cm^{-3}), but it has the lowest accuracy at the first input level (IRMSE = 0.096 cm^3cm^{-3}). The Boltzman model yielded better results at the first input level compared with the other two input levels and its predictions were reasonably accurate at the second and third input levels. The predictions produced by the Dexter model were similarly accurate at the first and second input levels, but the accuracy was lower at the third input level. The predictions made by the Campbell and Fermi models had the highest and lowest accuracy at the third and first input levels, respectively. The Durner model yielded a higher prediction error at the first input level, but the Durner model parameters were not predicted at the third input level. The predictions obtained by the van Genuchten model were more accurate when using the parameters of the PSD model as predictors at the second input level. Ghanbarian-Alavijeh *et al.* (2010) found that the van Genuchten model yielded better results when using the fractal dimension as a predictor. In addition, Schaap & Leij (1998) estimated the van Genuchten and Gardner models using artificial neural networks and reported that the predictions were improved by increasing the number of input variables.

CONCLUSION

The PTFs were developed to estimate water content through the nine SWRC models using easily measured soil characteristics in three input levels, and the interaction of different input levels and SWRC models was evaluated in the estimation of water content. This study showed that the Gardner model obtained the highest accuracy at all three input levels for all groups. However, it is not a suitable model because large errors occur during its fitting stage. An important note that should be considered in developing and using parametric PTFs is that to estimate water content, models with high fitting accuracy must be selected; otherwise, the estimated water content would be quite different from the measured one. Following the Gardner model, the Campbell model had the highest accuracy at the second input level for all groups, and at the third

input level for all soil samples and the fine textural group. These results demonstrate the high accuracy of the water content predictions produced by the Campbell model using detailed information regarding the soil texture and/or structure as predictors for all groups. However, using detailed information regarding the soil structure as predictors did not yield highly accurate predictions for the moderate textural group. The Boltzman model obtained the highest accuracy at the first input level for all soil samples and fine textural group, thereby demonstrating the suitability of the Boltzman model for predicting the water content with parametric PTFs for all soil samples and the fine textural group when using the basic soil properties as predictors. The Fermi model had the highest accuracy at the first and third input levels for the moderate textural group. In general, the Fermi model could correlate basic or structural soil properties with the water content in the moderate textural group. The water content predictions produced by the SWRC models were changed when various input variables were employed, thereby indicating that interactions occur among the input variables and the SWRC models when estimating the water content using parametric PTFs. Therefore, the most suitable SWRC model should be selected for the parametric estimation of the water content for each specific set of input variables.

ACKNOWLEDGEMENTS

The authors are deeply grateful to the anonymous reviewers and the editors for their helpful comments regarding the manuscript.

REFERENCES

Akaike, H. 1974 A new look at the statistical model identification. *IEEE Trans Automa. Contr.* **19**, 716–723.

Alemi, M. H. 1981 *Water and Soil*. Tehran University Press, Tehran, Iran, 346 pp. (in Persian).

Archer, N. A. L., Bonell, M., MacDonald, A. M. & Coles, N. 2014 A constant head well permeameter formula comparison: its significance in the estimation of field-saturated hydraulic conductivity in heterogeneous shallow soils. *Hydrol. Res.* **45** (6), 788–805.

Armstrong, J. S. 2012 Illusions in regression analysis. *Int. J. Forecasting* **28** (3), 689–694.

Bayat, H., Neyshabouri, M. R., Mohammadi, K. & Nariman-Zadeh, N. 2011 Estimating water retention with pedotransfer functions using multi-objective group method of data handling and ANNs. *Pedosphere* **21**, 107–114.

Bayat, H., Ebrahimi, E., Rastgo, M., Zare Abiane, H. & Davatgar, N. 2013a Investigating the fitting accuracy of different soil water characteristic models on various soil textural classes. *Water Soil Knowl.* **23** (3), 151–167 (in Persian).

Bayat, H., Neyshabouri, M. R., Mohammadi, K., Nariman-Zadeh, N., Irannejad, M. & Gregory, A. S. 2013b Combination of artificial neural networks and fractal theory to predict soil water retention curve. *Comput. Electron. Agric.* **92**, 92–103.

Brooks, R. & Corey, A. 1964 Hydraulic properties of porous media, Hydrology Paper No. 3. Colorado State University, Fort Collins, CO.

Campbell, G. S. 1974 A simple method for determining unsaturated conductivity from moisture retention data. *Soil Sci.* **117**, 311–314.

Chai, T. & Draxler, R. R. 2014 Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geosci. Model. Dev.* **7**, 1247–1250.

Dane, J. H. & Jan, W. H. 2002 Water retention and storage. In: *Methods of Soil Analysis. Part 4. Physical Methods* (J. H. Dane & G. G. Topp, eds). Soil Science Society of America Book Series No. 5. Soil Science Society of America, Madison, WI, USA, pp. 671–717.

Dexter, A. 2004 Soil physical quality: Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth. *Geoderma* **120** (3), 201–214.

Dexter, A. R., Czyz, E. A., Richard, G. & Reszkowska, A. 2008 A user-friendly water retention function that takes account of the textural and structural pore spaces in soil. *Geoderma* **143**, 143–253.

Durner, W. 1994 Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resour. Res.* **30** (2), 211–223.

Ebrahimi, E., Bayat, H., Neyshaburi, M. R. & Zare Abyaneh, H. 2013 Prediction capability of different soil water retention curve models using artificial neural networks. *Arch. Agron. Soil Sci.* **60** (6), 859–879.

Fredlund, D. G. & Xing, A. 1994 Equations for the soil water characteristic curve. *Can. Geotech. J.* **31**, 521–532.

Freedman, D. A. 2005 *Statistical Models: Theory and Practice*. Cambridge University Press, New York, USA.

Gardner, W. 1956 Mathematics of isothermal water conduction in unsaturated soils. In: *Highway Research Board Special Report 40 International Symposium on Physico Chemical Phenomenon in Soils*. Washington, DC, USA, pp. 78–87.

Gee, G. W. & Or, D. 2002 Particle-size and analysis. In: *Methods of Soil Analysis. Part 4. Physical Methods* (A. D. Warren, ed.). Soil Science Society of America, Madison, WI, USA, pp. 255–295.

Ghanbarian-Alavijeh, B., Liaghat, A., Guan-Huaand, H. V. & van Genuchten, M. T. 2010 Estimation of the van Genuchten soil water retention properties from soil textural data. *Pedosphere* **20** (4), 456–465.

Groenevelt, P. H. & Grant, C. D. 2004 A new model for the soil water retention curve that solves the problem of residual water contents. *European Journal of Soil Science* **55**, 479–485.

Grossman, R. B. & Reinsch, T. G. 2002 Bulk density and linear extensibility. In: *Methods of Soil Analysis. Part 4. Physical Methods* (J. H. Dane & G. Clarke Topp, eds). Soil Science Society of America, Madison, WI, USA, pp. 201–229.

Gupta, S. C. & Larson, W. E. 1979 Estimating soil water retention characteristic from particle size distribution, organic matter content, and bulk density. *Water Resour. Res.* **15**, 1633–1635.

Khlosi, M., Cornelis, W. M., Douaik, A. M., Mualem-van Genuchten, M. T. & Gabriels, D. 2008 Performance evaluation of models that describe the soil water retention curve between saturation and oven dryness. *Vadose Zone Journal* **7**, 87–96.

Koekkoek, E. J. W. & Boolting, H. 1999 Neural network models to predict soil water retention. *Eur. J. Soil Sci.* **50**, 489–495.

Konikow, L. F. & Bredehoeft, J. D. 1992 Ground-water models cannot be validated. *Adv. Water Resour.* **15**, 75–83.

Kosugi, K. 1994 Three parameter lognormal distribution model for soil water retention. *Water Resour. Res.* **30** (4), 891–901.

Kozak, E., Pachepsky, Y. A., Sokolowski, S., Sokolowska, Z. & Stepniewski, W. 1996 A modified number-based method for estimating fragmentation fractal dimensions of soils. *Soil Sci. Am. J.* **60**, 1291–1297.

Mayr, T. & Jarvis, M. N. J. 1999 Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Coorey type model. *Geoderma* **91**, 1–9.

McKee, C. & Bumb, A. 1984 The importance of unsaturated low parameters in designing a hazardous waste site. In: *Hazardous Wastes and Environmental Emergencies Hazardous Materials Control Research Institute National Conference*, Houston, TX, pp. 50–58.

McKee, C. & Bumb, A. 1987 Flow-testing coalbed methane production wells in the presence of water and gas. *SPE Formation Evaluation* **2** (4), 599–608.

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 Tillage Res.* **90**, 108–116.

Millan, H., Gonzalez, P. M., Morilla, A. A. & Perez, E. 2007 Self similar organization of Vertislo microstructure a pore solid fractal interpretation. *Geoderma* **138**, 185–190.

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.

Minasny, B., Hopmans, J. W., Harter, T., Eching, S. O., Tuli, A. & Denton, M. A. 2004 Neural network prediction soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Sci. Soc. Am. J.* **68**, 417–429.

Nielson, D. R. & Shaw, R. H. 1958 Estimation of the 15 atmosphere moisture percentage from hydrometer data. *Soil Sci.* **86**, 103–105.

Oakdale Engineering 2008 DataFit 9.059. PA, USA.

Omuto, C. 2009 Biexponential model for water retention characteristics. *Geoderma* **149** (3), 235–242.

Pachepsky, Y. A., Timlin, D. & Varallyay, G. 1996 Artificial neural network to estimate soil water retention from easily measurable data. *Soil Sci. Soc. Am. J.* **60**, 727–733.

Pachepsky, Y., Rawls, W., Gimenez, D. & Watt, J. P. C. 1998 Use of soil penetration resistance and group method of data handling to improve soil water retention estimates. *Soil Tillage Res.* **49**, 117–126.

Patil, N. G., Pal, D. K., Mandal, C. & Mandal, D. K. 2012 Soil water retention characteristics of vertisols and pedotransfer functions based on nearest neighbor and neural network approach to estimate AWC. *J. Irrig. Drain. Eng.* **138**, 1–10.

Perrier, E. M. A. & Bird, N. R. A. 2002 Modelling soil fragmentation: The pore solid fractal approach. *Soil Tillage Res.* **64** (1–2), 91–99.

Schaap, M. G. & Leij, F. J. 1998 Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil Tillage Res.* **47**, 37–42.

Seki, K. 2007 SWRC Fit a nonlinear fitting program with a water retention curve for soils having unimodal and bimodal pore structure. *Hydrol. Earth Syst. Sci. Discuss.* **4**, 407–437.

Sillers, S. W., Fredlund, D. G. & Noshin, Z. 2001 Mathematical attributes of some soil water characteristic curve models. *Geotech. Geolog. Eng.* **19**, 243–283.

SPSS 16 [computer program] 2007 SPSS, Inc., Chicago, IL, USA.

Stirk, G. B. 1957 *Physical properties of soils of the lower Burdekin valley, North Queensland*. CSIRO Division of Soils Divisional Report 1/57. CSIRO, Canberra, Australia.

Tietje, O. & Tapkenhinrichs, M. M. 1993 Evaluation of pedotransfer functions. *Soil Sci. Soc. Am. J.* **57**, 1088–1095.

Tomasella, J., Hodnett, M. G. & Rossato, L. 2000 Pedotransfer functions for the estimation of soil water retention in Brazilian soils. *Soil Sci. Soc. Am. J.* **64**, 327–338.

Tuller, M. & Or, D. 2003 Retention of water in soil and the soil water characteristic curve. Soil water characteristic. A manual. http://www.engr.uconn.edu/environ/envphys/pdf/vadose_pdf/SWC_revised01.pdf, pp. 1–20.

Ungaro, F., Calzolari, C. & Busano, E. 2005 Development of pedotransfer functions using a group method of data handling for the soil of the Pianura Padano-Veneta region of north Italy: water retention properties. *Geoderma* **124**, 293–317.

van Genuchten, M. T. 1980 A closed form equation predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* **44**, 892–898.

van Genuchten, M. T., Leij, F. J. & Yates, S. R. 1991 *The RETC code for quantifying the hydraulic functions of unsaturated soils*. Robert S. Kerr Environmental Research Laboratory, Office of Research and Development, US Environmental Protection Agency, Ada, OK, USA.

Wosten, J. H. M., Finke, P. A. & Jansen, M. J. W. 1995 Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics. *Geoderma* **66**, 227–237.

Wösten, J. H. M., Pachepsky, Y. & Rawls, W. J. 2001 Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J. Hydrol.* **251**, 123–150.

First received 1 June 2015; accepted in revised form 5 October 2015. Available online 4 November 2015