Deriving and validating pedotransfer functions for some calcareous soils

Habib Khodaverdiloo [a], Mehdi Homaee [b,*], Martinus Th. van Genuchten [c], Shoja Ghorbani Dashtaki [d]

[a] Department of Soil Science, Urmia University, Urmia, Iran
[b] Department of Soil Science, Tarbiat Modares University, Tehran 14115-336, Iran
[c] Department of Mechanical Engineering, COPPE/LTTC, Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil
[d] Department of Soil Science, Shahrekord University, Shahrekord 88186/34141, Iran

ARTICLE INFO

SUMMARY

The unsaturated soil hydraulic properties are needed for many different applications in soil hydrology. Pedotransfer functions (PTFs) have proven to be useful to indirectly estimate these parameters from more easily obtainable soil data. Until now no studies have been conducted to derive or verify PTFs for calcareous soils, which hydraulically may not behave the same as non-calcareous soils. The objectives of this study were to assess the influence of soil $CaCO_3$ on the soil water retention characteristics of some calcareous soils, and to derive PTFs for these soils. Two data sets were used to derive and evaluate the established PTFs. Data set 1 containing 220 samples was employed as a calibration set for multiple linear regression. An independent data set containing 55 soil samples from a different location served to verify the derived PTFs. No significant difference in accuracy was found between the PTFs with and without $CaCO_3$ in terms of estimating specific soil water retention values or the van Genuchten soil hydraulic parameters. Compared with the Rosetta PTFs of Schaap et al. (2001), the derived point and parametric PTFs provided better accuracy with average *RMSE* values of 0.028 and 0.107 $cm^3\,cm^{-3}$, respectively.

1. Introduction

Estimates of the unsaturated soil hydraulic properties are needed for many agronomic, environmental and engineering applications involving water flow and contaminant transport processes in the vadose zone. A broad array of laboratory and field methods are currently available for direct measurement of the hydraulic properties (e.g., Dirksen, 2000; Dane and Topp, 2002). Unfortunately, most or all methods are too time consuming and costly for most applications (Wösten et al., 2001). Furthermore, due to high temporal and spatial variability in the hydraulic characteristics, a large number of samples is generally required to accurately characterize field conditions.

An alternative to direct measurement is the use of indirect methods in which the hydraulic properties are estimated from more easily measured or more readily available soils data (Bouma, 1989; Wösten et al., 2001). Methods that may be used for this purpose include statistical pore-size distribution models, inverse models and especially pedotransfer functions (PTFs). PTFs (Bouma and van Lanen, 1987; Bouma, 1989) are based on the premise that the hydraulic properties can be related to basic soil properties such as the particle size distribution, bulk density and/or organic matter content. From a prediction point of view, PTFs can be distinguished into point and parametric PTFs. Point PTFs predict the water content at pre-defined soil water pressure heads, while parametric PTFs assume that the hydraulic properties can be described adequately with a hydraulic model containing a limited number of parameters.

Iran, as many other countries, does not have sufficient soil hydraulic data for deriving local PTFs for relevant agricultural or environmental modeling applications. However, several large databases such as UNSODA (Leij et al., 1996), HYPRES (Lilly, 1997; Wösten et al., 1999), WISE (Batjes, 1996) and the USDA-NRCS ped-on soil characterization database (Soil Survey Staff, 2010a) have been established to facilitate analyses of the soil hydraulic properties. Consequently, it is attractive to use such databases to develop PTFs for regions with less available data. Although the advantages of existing PTFs are well discussed in the literature (e.g. Tietje and Tapkenhinrichs, 1993; Espino et al., 1995; Wösten et al., 1995; Minasny et al., 1999), their applicability to calcareous soils has not been addressed thus far (Homaee and Farrokhian Firouzi, 2008). Also, practical applications of most PTFs is often hampered by their very specific data requirements.

Hydrologists in countries with insufficient soil hydraulic or taxonomic data are often confronted with situations where one or more PTF input parameters is not available. The Rosetta package (Schaap et al., 2001) implements five hierarchical PTFs to predict the WRC, as well as the saturated and unsaturated hydraulic

conductivity. The hierarchy in PTFs allows prediction of van Genuchten hydraulic parameters and the saturated hydraulic conductivity using limited (soil textural class only) to more extended (texture, bulk density, and one or two water retention points) input data. This makes Rosetta quite useful for countries having data limitations.

Calcareous soils are widely distributed over the world and physically may not behave the same as non-calcareous soils. Calcareous soils contain 5% or more (by volume) inorganic carbon or CaCO$_3$ equivalent (Soil Survey Staff, 1999). Soils of this type are the dominant formation in many areas, particularly in arid and semiarid regions. According to the FAO, about 65% of soils in Iran are calcareous (FAO/UNDP, 1972). Inorganic carbon in soils commonly exists in the form of the carbonate minerals calcite (CaCO$_3$), dolomite [Ca, Mg(CO$_3$)$_2$], and as magnesium calcites (Ca$_{1-x}$Mg$_x$CO$_3$). Other less common forms are aragonite (CaCO$_3$) and siderite (FeCO$_3$). Carbonates in soils may be of primary (inherited from parent materials) or secondary (pedogenic) origin. Secondary carbonates usually consists of aggregates made up of silt- and clay-sized calcite crystals, which are easily identified using grain mount techniques. Large crystals of calcite or dolomite are mostly of primary origin (Doner and Lynn, 1989).

Qualitative and quantitative knowledge of Ca and Mg carbonates is useful in many agronomic and environmental studies, such as for investigations of root growth, water movement, contaminant transport, soil pH (Nelson, 1982) and the nature of the exchange complex (Goh and Mermut, 2006). Studies that considered the effects of calcium carbonate precipitation on reductions in pore space suggest that neither water retention nor the hydraulic conductivity is affected significantly by calcium carbonate precipitation (Frenkel et al., 1978). However, from water retention and adsorption energy points of view, carbonates of the clay size can be treated as silt particles (Soil Survey Staff, 2010b). Thus, the hydraulic properties of calcareous soils may well differ from non-calcareous soils. Also, calcium carbonate is routinely removed from soil when soil texture is determined. The properties of soils with large CaCO$_3$ contents hence may well change during standard hydraulic analyses, which may further complicate the use of PTFs for non-calcareous soils. Hence, the applicability of available PTFs, such as those in the Rosetta package, to calcareous soils remains largely untested.

The objectives of this study were to (1) derive and verify pedotransfer functions of several typical calcareous soils, and to test whether the inclusion of soil CaCO$_3$ content as an input parameter would improve the accuracy of the PTFs, and (2) to investigate if the Rosetta PTFs, which were not derived specifically for calcareous soils, can be applied to calcareous soils.

2. Materials and methods

2.1. Soil data and soil characterization

A large set of disturbed and undisturbed samples were collected from the top 30 cm soil horizon of the Karaj, Karaj-Najm Abad, Mohammad Abad, Mard Abad and Nazar Abad soil series near the city of Karaj, Iran. The soils were classified as Torriorthents, Calcixerepts, Haplocalcids and Haplocambids according to USDA Soil Taxonomy (Soil Survey Staff, 2010b). In total, 275 undisturbed cores (7.5 cm diameter by 5.0 cm height) and 275 disturbed samples were collected. The disturbed samples were first air-dried and then sieved through a 2 mm sieve. The particle size distributions of the samples were obtained using the hydrometer method (Gee and Or, 2002), while CaCO$_3$ contents were determined using pressure calcimetry (Sherrod et al., 2002).

To determine water retention values, the soil cores were slowly saturated by soaking them in water at a level just below the top of

the core. This procedure was carried out over a period of 2 days to obtain as complete saturation as possible by minimizing entrapped and dissolved air in the cores. Soil water contents at saturation were determined by oven-drying the saturated samples at 105 °C for 24 h. After saturation, cores were placed on a saturated ceramic plate inside a pressure plate apparatus and subjected to consecutive pressure heads of −100, −330, and −1000 cm to allow desaturation of the samples and measurement of the water content. Soil water retention values at −3000, −5000, and −15,000 cm pressure heads were similarly determined using a pressure plate apparatus on disturbed soil samples in three replicates.

Particle size distributions of selected soil samples (ranging from those having the lowest to the highest CaCO$_3$ contents) were determined after removing their calcium carbonate. Calcium carbonate was removed using diluted HCl following procedures proposed by Kroetsch and Wang (2006) for removal of carbonates prior to determining particle size distributions. The HCl-treated soil samples were air-dried and grinded, after which the particle size distributions and water retention values at −3000, −5000, and −15,000 cm pressure heads were obtained, again using the pressure plate apparatus. A t-test was performed to verify statistical differences between the amounts of retained water and the soil particle size distributions of paired samples (i.e., samples before and after calcium carbonate removal).

To obtain a uniform description of all WRCs, we used the soil water retention equation of van Genuchten (1980):

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{(1 + |\alpha h|^n)^{1-1/n}} \qquad (1)$$

where θ is the volumetric water content (–), h is the soil water pressure head (L), θ_r and θ_s are the residual and saturated water contents (–), respectively, and α (L^{-1}) and n (–) are shape factors. The nonlinear least-squares optimization program RETC of van Genuchten et al. (1991) was used to fit the hydraulic parameters in Eq. (1) to the measured retention data.

The geometric mean d_g (L) and geometric standard deviation δ_g (–) of the soil particle diameters were determined using methods proposed by Shirazi and Boersma (1984):

$$d_g = \exp a, \quad a = 0.01 \sum_1^j f_i \ln M_i \qquad (2)$$

$$\delta_g = \exp b, \quad b^2 = 0.01 \sum_1^j f_i \ln^2 M_i - a^2 \qquad (3)$$

where j is the number of soil separates (e.g., clay, silt, sand), M_i is the mean diameter (in mm) of particles in soil separate i, and f_i is the percentage of total soil mass having a diameter equal to or less than M_i.

The entire soils database was divided into two subsets. Subset 1 containing 80% of the samples ($n = 220$) from the Karaj, Karaj-Najm Abad, and Mard Abad soil series was used as the calibration set. Subset 2 with the remaining 20% ($n = 55$) involving independent samples from the Mohammad Abad and Nazar Abad soil series was employed as a validation set. Table 3 gives details on the properties of the two data sets.

2.2. PTF development and WRC estimation

Multiple linear regression analyses were performed on the calibration subset to correlate soil water retention values at pressure heads of 0, −100, −330, −1000, −3000, −5000, and −15,000 cm, as well as the van Genuchten parameters, with available soil physicochemical parameters to develop several pairs of PTFs, i.e., PTFs with and without using CaCO$_3$ as one of the predictors of the PTFs. The

paired PTFs hence refer to using $CaCO_3$ as input to the developed PTFs, not to whether $CaCO_3$ was or was not removed from the samples prior to the water content measurements.

The accuracies of the paired PTFs were compared using several quantitative measures to provide a basis for selecting the best PTFs for estimating soil water retention values. Water contents at pressure heads of 0, −100, −330, −1000, −3000, −5000, and −15,000 cm were subsequently generated for both the calibration and validation subsets using the selected PTFs, either directly as point PTFs or by inserting the parametric PTFs into Eq. (1). The van Genuchten retention equation was used next to compute water contents at all pressure heads at which water contents had been measured. The water retention parameters were also predicted with the hierarchical PTFs of Rosetta using sand, silt and clay percentages, as well as the bulk density, of the validation subset as input data. We emphasize that the main aim of this study was not necessarily to obtain very accurate predictions as such, but first and foremost to derive PTFs that show possible effects of $CaCO_3$ on water retention.

2.3. Reliability, accuracy and validity of the PTFs

After analyzing for normality in the data, the subset regression procedure was used to select the best group of independent variables. Our general approach was to select the smallest subset that fulfilled certain statistical criteria. The Minitab software (Minitab, 1998) was used to obtain the R-squared (R^2), adjusted R-squared (R^2_{adj}), and C_p statistics for this purpose. The C_p statistic is given by:

$$C_p = \left(\frac{SSE_p}{MSE_m}\right) - (m - 2p) \tag{4}$$

where SSE_p is the sum-of-squared error (SSE) of the best model having p parameters (including its intercept), and MSE_m is the mean square error (MSE) of the model using all m predictors. A small value of C_p indicates that the model is relatively precise (has a small variance) in estimating the true regression coefficients and predicting future responses (Ryan and Joiner, 1994). In this study we choose a subset for which C_p approached the number of predictors, p, in a given regression model (Daniel and Wood, 1980).

The degree of multicollinearity in each derived PTF was tested using the "Variation Inflation Factor", VIF (Hocking, 2003; Ho, 2006). We furthermore used the Durbin–Watson statistic to determine if the residuals showed any significant correlation based on the order in which the data were analyzed.

As predictive equations, PTFs are routinely evaluated in terms of the correspondence between measured and predicted values. When the data used for evaluation are different from those used to develop the equations, the "reliability" of the prediction is evaluated (Pachepsky and Rawls, 1999; Wösten et al., 2001). In this study we evaluated the reliability of the derived PTFs, as well as those obtained with Rosetta, using the independent data set from the Mohammad Abad and Nazar Abad soil series.

A complete assessment of model performance should include one "goodness-of-fit" measure and at least one absolute error measure (Legates and McCabe, 1999). The mean error (ME), root mean square error ($RMSE$), and index of agreement (d) were used to evaluate both the accuracy and the reliability of the derived PTFs. We used the d statistic to circumvent certain problems associated with the commonly used coefficient of determination (R^2). The index of agreement reflects the degree to which the observations are accurately estimated by the predictions (Willmott, 1981). In contrast to R^2, d is a measure of the degree at which model predictions are error free, rather than being a measure of correlation or association. The mathematical expressions of the various statistics are as follows (Willmott, 1981):

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\theta_i^p - \theta_i^o) \tag{5}$$

$$RMSE = \left[\frac{\sum_{i=1}^{n}(\theta_i^p - \theta_i^o)^2}{n}\right]^{1/2} \tag{6}$$

$$d = 1 - \frac{\sum_{i=1}^{n}(\theta_i^p - \theta_i^o)^2}{\sum_{i=1}^{n}\left[\left|\left(\theta_i^p - \overline{\theta^o}\right)\right| + \left|\left(\theta_i^o - \overline{\theta^o}\right)\right|\right]^2} \tag{7}$$

where θ_i^p are the predicted values of the soil water content, θ_i^o are the observed soil water contents, n is the number of samples, while over-lined symbols represent mean values. We furthermore evaluated the relative improvement (RI) of the PTFs with $CaCO_3$ (shown by the subscript $+CaCO_3$) over those without $CaCO_3$ (subscript $-CaCO_3$) as a PTF predictor in terms of the $RMSE$ for all $\theta(h)$ pairs using:

$$RI = \frac{RMSE_{-CaCO_3} - RMSE_{+CaCO_3}}{RMSE_{-CaCO_3}} \times 100\% \tag{8}$$

The relative improvement of region-specific PTFs over Rosetta was similarly evaluated by replacing $RMSE_{-CaCO3}$ and $RMSE_{+CaCO3}$ in Eq. (8) with the $RMSE$s of the region-specific PTFs and the Rosetta predictions, respectively.

3. Results and discussion

Ranges in the particle size distribution, the calcium carbonate content and water retained at pressure heads of −3000, −5000, and −15,000 cm for some selected soil samples before and after calcium carbonate removal are shown in Table 1. A t-test indicated a significant ($p \leqslant 0.001$) increase in retained water at these three pressure heads after removal of $CaCO_3$ from the samples. While any retained water was significantly influenced by the particle adsorption energy at the lower pressure heads, the reason for this may be attributed to the fact that carbonates within the clay size act more as silt from a water holding point of view. A comparison of the particle size distributions in Table 1 of the soil samples before and after $CaCO_3$ removal indicates, surprisingly, that the sand percentage, and hence also the mean soil particle size, increased after removing $CaCO_3$. One may expect the opposite as carbonates often act as a cementing agent to cluster smaller particle into larger units.

One way to explain the particle size increase is that the carbonate removal procedure preferentially eliminated smaller carbonate particles while leaving the larger particles unchanged. Another explanation may be that $CaCO_3$ particles in calcareous soils could act not only as a cementing agent to cluster primary particles into secondary particles, but also as primary particles by themselves. Hence, carbonate removal may have caused the preferential

Table 1
Ranges in soil properties before and after $CaCO_3$ removal (in% of the material prior to $CaCO_3$ removal) for some selected soil samples (values in parentheses are mean values).

Soil property	Range before $CaCO_3$ removal	Range after $CaCO_3$ removal
C (%)[a]	14–26 (22)	9–29 (20)
Si (%)	40–50 (44)	13–52 (29)
S (%)	28–44 (34)	34–67 (51)
$CaCO_3$ (%)	6.4–28 (11)	–
$\theta_{-3000cm}$ (cm^3 cm^{-3})	0.13–0.17 (0.15)	0.18–0.25 (0.21)
$\theta_{-5000cm}$ (cm^3 cm^{-3})	0.11–0.16 (0.14)	0.15–0.23 (0.19)
$\theta_{-15,000cm}$ (cm^3 cm^{-3})	0.10–0.14 (0.12)	0.13–0.21 (0.16)

[a] C, Si, and S: clay, silt, and sand contents, respectively.

removal of carbonate particles within the clay size. This suggests that determination of soil water retention of acid-treated soil samples may lead to unrealistic results when used for field water management, or for agricultural or environmental modeling applications. Table 1 also shows that in spite of a reduction in the smaller particle sizes after CaCO₃ removal, the retained water increased in the soil samples. This supports the view that calcium carbonate particles within the clay size did not act as clay particles to retain soil water (Soil Survey Staff, 2010b). These results reflect lower adsorption energy of the calcium carbonate than other soil particles of approximately the same size.

A set of regression equations was obtained to estimate the van Genuchten parameters as well as water retained at pressure heads of 0, −100, −330, −1000, −3000, −5000, and −15,000 cm, using the best subset regression method. Since our multiple regression approach assumes normal distributions of the dependent variables, the degree of normality of the data was tested using Ryan and Joiner's (1994) method. The normality tests showed that some of the variables were not normally distributed. Therefore, the data with non-normal distributions were normalized as much as possible using some of the power and other transformations shown in Table 2. This table provides

Table 2
Derived PTFs of calcareous soils that predict water retention at specific pressure heads, as well as van Genuchten parameters, with (+CaCO₃) and without (−CaCO₃) using soil calcium carbonate content as an input parameter. The equations represent the best subset regression results and their adjusted R^2 (R^2_{adj}), root mean square error (*RMSE*), and relative improvement (*RI*) values.

	The derived pedotransfer functions	R^2_{adj} (*RMSE*) **RI** (%)
+CaCO₃	$\theta_s = 0.380 + 0.00620\ C/S + 0.108\ \rho_b - 0.00306\ CaCO_3$	0.57***(0.039) **4.9**
−CaCO₃	$\theta_s = 0.230 + 0.00110\ C/S + 0.177\ \rho_b$[a]	0.42***(0.041)
+CaCO₃	$\theta_{-100cm} = 0.222 + 0.00946\ C/S + 0.106\ \rho_b - 0.00172\ CaCO_3$	0.69***(0.022) **4.3**
−CaCO₃	$\theta_{-100cm} = 0.306 + 0.0123\ C/S + 0.0668\ \rho_b$	0.71***(0.023)
+CaCO₃	$\theta_{-330cm} = 0.720 + 0.177\ \rho_b - 0.936\ d_g^{0.1} - 0.000044\ CaCO_3$	0.72***(0.026) **0.0**
−CaCO₃	$\theta_{-330cm} = 0.724 + 0.176\ \rho_b - 0.938\ d_g^{0.1}$	0.73***(0.026)
+CaCO₃	$\theta_{-1000cm} = 0.612 + 0.142\ \rho_b - 0.796\ d_g^{0.1} - 0.000025\ CaCO_3$	0.76***(0.020) **−2.4**
−CaCO₃	$\theta_{-1000cm} = 0.614 + 0.141\ \rho_b - 0.797\ d_g^{0.1}$	0.77***(0.016)
+CaCO₃	$\theta_{-3000cm} = 0.529 + 0.109\ \rho_b + 0.00247\ \delta_g - 0.723\ d_g^{0.1} - 0.000183\ CaCO_3$	0.77***(0.017) **0.0**
−CaCO₃	$\theta_{-3000cm} = 0.544 + 0.105\ \rho_b + 0.00250\ \delta_g - 0.731\ d_g^{0.1}$	0.78***(0.017)
+CaCO₃	$\theta_{-5000cm} = 0.460 + 0.112\ \rho_b + 0.00230\ \delta_g - 0.648\ d_g^{0.1} - 0.000075\ CaCO_3$	0.73***(0.017) **0.0**
−CaCO₃	$\theta_{-5000cm} = 0.466 + 0.111\ \rho_b + 0.00231\ \delta_g - 0.652\ d_g^{0.1}$	0.74***(0.017)
+CaCO₃	$\theta_{-15,000cm} = 0.422 + 0.131\ \rho_b + 0.00113\ \delta_g - 0.650\ d_g^{0.1} - 0.000242\ CaCO_3$	0.65***(0.021) **0.0**
−CaCO₃	$\theta_{-15,000cm} = 0.402 + 0.136\ \rho_b + 0.00109\ \delta_g - 0.638\ d_g^{0.1}$	0.66***(0.021)
+CaCO₃	$\theta_r = 0.103 - 0.00274\ \delta_g - 0.873\ \theta_{-330cm} + 1.98\ \theta_{-15,000cm} - 0.000736\ CaCO_3$	0.71**(0.026) **0.0**
−CaCO₃	$\theta_r = 0.114 - 0.00328\ \delta_g - 0.837\ \theta_{-330cm} + 1.96\ \theta_{-15000cm}$	0.70**(0.026)
+CaCO₃	$\alpha = 0.0142 + 0.000771\ \delta_g - 0.000612\ CaCO_3$	0.24***(0.008) **18.3**
−CaCO₃	$\alpha = -0.00230 + 0.00129\ \delta_g$	0.06**(0.010)
+CaCO₃	$n = 2.22 - 0.0217\ \delta_g - 3.89\ \theta_{-330cm} + 4.57\ \theta_{-15,000cm} + 0.00292\ CaCO_3$	0.30***(0.153) **0.8**
−CaCO₃	$n = 2.27 - 0.0238\ \delta_g - 3.74\ \theta_{-330cm} + 4.48\ \theta_{-15,000cm}$	0.29***(0.155)

d_g: geometric mean of the soil particle diameter (mm).
δ_g: geometric standard deviation of the soil particle diameter (−).
*** and ** indicate statistical significance at the 0.001 and 0.01 confidence level, respectively.
[a] C, S, ρ_b, and CaCO₃: clay content (%), sand content (%), soil bulk density (g cm⁻³), and soil calcium carbonate (%), respectively.

the derived point and parametric PTFs of the calcareous soils with and without $CaCO_3$ as an input parameter, and their RMSE, RI, and $\left(R^2_{adj}\right)$ values. The derived PTFs showed VIF values that were always less than the critical value of 10 often used to indicate multicollinearity among the regressed variables (e.g., Ho, 2006), except for the θ_s and θ_{100cm} PTFs that used both the clay and sand contents as independent predictors. To avoid multi-collinearity in this case, the clay to sand ratio (C/S) was used rather than the original variables. Also, the Durbin–Watson value for all of the derived PTFs was found to be always larger than 1.4, thus indicating that no significant autocorrelation existed among the residuals (Ho, 2006).

Compared with the point PTFs, the accuracy of the derived PTFs for the α and n parameters was not as good. Our results indicate that the derived PTFs could not accurately predict the hydraulic parameters. This may be caused by interdependency amongst the parameters, poor fit of the model to the data, or over-parameterization of the model (van den Berg et al., 1997).

As shown in Table 2, using $CaCO_3$ as an additional input parameter for the PTFs improved the predictions between −2.4% and 18.3%, especially for the α hydraulic parameter. The relatively large improvement for α may be related to changes in the pore-size distribution by secondary $CaCO_3$, which would mainly affect α as discussed earlier. Still, in view of the relatively small differences between the RMSE values (a maximum of 0.031 cm^{-1} for α), this was not a significant improvement, even for α itself which had the highest RI (18.3%). Overall there was no significant difference between the paired PTFs (i.e., the PTFs with and without $CaCO_3$) to estimate soil water retention at either specific pressure heads or via the van Genuchten parameters. This may be due to the fact that the other input parameters (such as bulk density and soil texture) already may have accounted for part of the possible effects of $CaCO_3$ on soil water retention, thus reducing the direct effects of $CaCO_3$ on the PTF results. To further test this hypothesis, we used the PTFs with $CaCO_3$ for α, and those without $CaCO_3$ for the other output variables to evaluate the accuracy of region-specific PTFs, and to compare our results with those obtained using Rosetta.

Table 2 shows that $CaCO_3$ had a negative effect on all water retention values, both at relatively high and low water contents. One possible reason for this is that $CaCO_3$ clusters smaller particles to larger ones, thus possibly increasing secondary porosity and hence the water content, except perhaps at or very near saturation.

At the same time, the low adsorption energy of $CaCO_3$ at the dry end of the curve may lead to its negative correlation with the water content (Table 1).

We note that the PTF's in Table 2 are limited to the Iranian soil samples used in our analysis. Table 3 shows ranges in observed physical and chemical data of the soil samples in the calibration (set 1) and validation (set 2) data sets before $CaCO_3$ removal, including of the soil particle size distribution, the soil bulk density ρ_b, and the $CaCO_3$ content. The values in Table 3 can be used to avoid extrapolation of the PTF predictions outside of the measurement range.

We next applied the derived PTFs, as well as the Rosetta PTFs using sand, silt and clay percentages and bulk density as input data, to the validation subset (set 2). Table 4 provides a quantitative comparison of the reliability of the water contents predicted with the derived point and parametric PTFs and Rosetta. Compared with Rosetta, the region-specific point-PTFs did lead to better accuracy with RMSE < 0.054 cm^3 cm^{-3}. The average RMSEs were 0.028, 0.107, and 0.176 cm^3 cm^{-3} for the point, parametric, and Rosetta PTFs, respectively. The region-specific parametric PTFs had a tendency to overestimate the water contents, while the Rosetta PTFs underestimated the water contents. The average ME values were 0.000, 0.085, and −0.148 cm^3 cm^{-3} for the point and parametric PTFs and Rosetta, respectively. The differences seemed to be larger at the higher water contents (Table 1 and Fig. 1a), which may be interpreted as an inability of the bulk density to reflect the effects of soil structure on the WRC. The average values of the index of agreement, d, were 0.91 and 0.54 for the point and parametric PTFs, respectively, which reflects reasonable performance of the PTFs. The average value of d was 0.32 for Rosetta. In all, the results indicate relatively good performance of the selected point and parametric PTFs in terms of the three statistics (see also Fig. 1a). The average relative improvements of our region-specific point and parametric PTFs over Rosetta were 83% and 36%, respectively. Comparing the average RMSE value of Rosetta in this study (RMSE > 0.102 cm^3 cm^{-3}) with that of Rosetta for its own data (0.068 cm^3 cm^{-3}; Schaap et al., 2001) shows that Rosetta did not perform well for our calcareous soils. We note here that some studies suggested that the accuracy of simulations using PTF-based soil hydraulic properties may be acceptable for selected applications when the RMSE values are on the order of about 0.065–0.070 cm^3 cm^{-3} (e.g., Nemes et al., 2003).

Table 3
Ranges in particle size distribution, geometric mean particle diameter d_g, geometric mean standard deviation of particle diameter δ_g, $CaCO_3$ content, and soil bulk density ρ_b used as input data for calibration (set 1) and validation (set 2) of the pedotransfer functions.

| Parameter	C (%)	S (%)	Si (%)	d_g (mm)	δ_g (−)	$CaCO_3$ (%)	ρ_b (g cm^{-3})
Range for set 1	11–48	17–54	6–71	0.0067–0.2468	7.35–20.34	4.5–47	1.20–1.83
Range for set 2	12–48	17–53	6–56	0.0068–0.1157	7.58–20.22	6.4–28	1.20–1.83

Table 4
Calculated mean error ME, root mean square error, RMSE, and index of agreement, d, statistics for quantitative validation of the reliability of the derived PTF's and Rosetta at different volumetric water contents (θ) for the validation data set. The values in bold are the relative improvement, RI, of region-specific PTFs over those of Rosetta.

		θ_s	θ_{-100cm}	θ_{-330cm}	$\theta_{-1000cm}$	$\theta_{-3000cm}$	$\theta_{-5000cm}$	$\theta_{-15000cm}$
ME (cm^3 cm^{-3})	Point PTF	0.001	0.000	0.000	−0.001	0.001	0.001	0.000
	Parametric PTF	0.000	0.076	0.150	0.142	0.098	0.076	0.050
	Rosetta	−0.121	−0.260	−0.193	−0.152	−0.118	−0.109	−0.086
RMSE (cm^3 cm^{-3})	Point PTF	0.057	0.025	0.029	0.022	0.019	0.019	0.023
RI (%)		**67**	**91**	**87**	**87**	**86**	**85**	**77**
	Parametric PTF	0.054	0.095	0.174	0.163	0.114	0.090	0.059
		69	**68**	**22**	**6**	**17**	**29**	**42**
	Rosetta	0.172	0.297	0.222	0.174	0.137	0.126	0.102
		−	−	−	−	−	−	−
d (−)	Point PTF	0.892	0.871	0.919	0.933	0.937	0.924	0.893
	Parametric PTF	0.845	0.588	0.377	0.367	0.449	0.506	0.674
	Rosetta	0.363	0.218	0.294	0.315	0.350	0.352	0.401
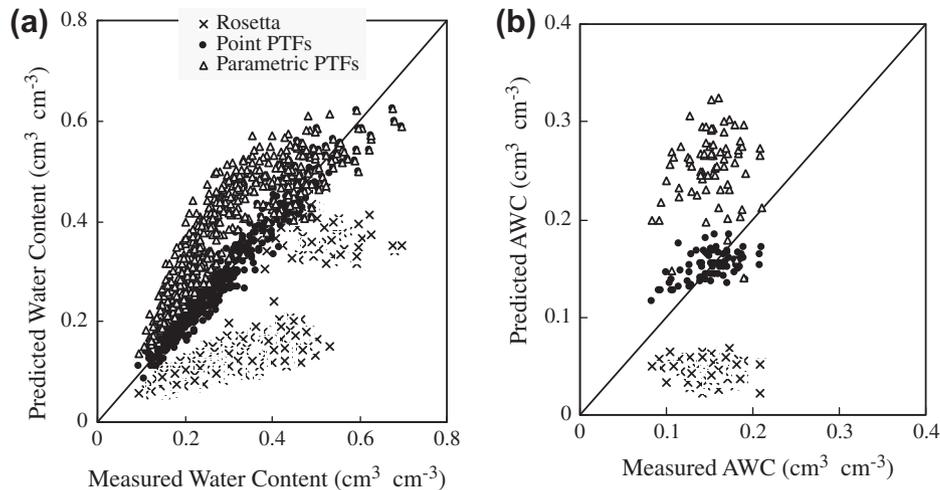
Fig. 1. Measured and predicted soil water retentions (SWR) at pressure heads of 0, −100, −330, −1000, −3000, −5000, and −15,000 cm (all together) (a) and available water content (AWC) (b) using the different PTFs for the validation data set. The solid lines indicate the 1:1 diagonals.
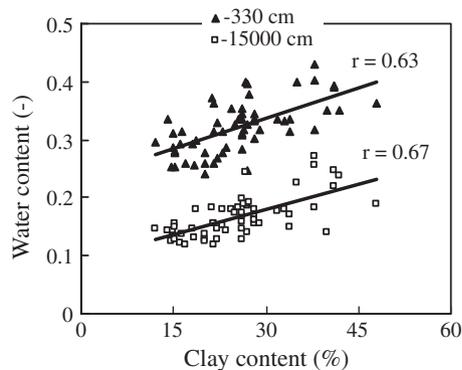
Fig. 2. Relationships between the volumetric water contents at pressure heads of −330 cm (field capacity, FC) and −15,000 cm (permanent wilting point, PWP) with the soil clay content.

The results so far indicate that differences between estimates using PTFs specifically developed for calcareous soils and the Rosetta package were not fully due to the inclusion of soil CaCO$_3$ content into the PTFs, but likely resulted also from other factors such as database-related uncertainty (Schaap and Leij, 1998) as well as from differences between the invoked algorithms (Nemes et al., 2003). However, in the case of the α parameter, part of these differences may be due to the fact that soil CaCO$_3$ content was not a parameter in the hierarchical PTFs of the Rosetta package. We next compared measured retention values with water contents estimated with the point and parametric PTFs and the Rosetta package (Fig. 1a) for the entire retention curve, as well as measured and predicted values of the available water content, AWC, (Fig. 1b), which is defined as the difference between the water contents at −330 cm (field capacity, FC) and −15,000 cm (the permanent wilting point, PWP). We found a significant correlation ($p = 0.001$) between the measured and estimated water contents for the entire retention curve. However, the correlation for available water content was relatively weak. This may be related to the fact that the water content at both pressure heads (−330 and −15,000 cm) increased with increasing clay content. Fig. 2 shows the regression between water content and clay content for two different pressure heads. The best fitted lines were found to diverge slightly. One other reason for the weakness of the correlation may be due to the fact that the water contents

at FC and PWP were both estimated from the regression equations. Thus, any errors in those predictions were already included in the model and hence propagated into the estimates for AWC.

4. Conclusions

In this study we assessed the effect of soil CaCO$_3$ content on water retention, and derived PTFs to predict the water retention curve and the van Genuchten parameters of calcareous soils. Relative improvements of the derived PTFs as compared to the Rosetta PTFs were also evaluated. Our results suggest lower adsorption energy (and hence lower water contents) of soil calcium carbonate than of other soil particles having approximately the same size. We found no significant difference between the paired PTFs (i.e., with and without CaCO$_3$ as an input parameter) in estimating water retention, neither for the point nor parametric PTFs. This may be attributed to the fact that the other input parameters (notably bulk density and soil texture) already may have accounted for part of the possible effects of CaCO$_3$ on soil water retention, thus reducing the direct effects of CaCO$_3$ on the PTF results. Compared with Rosetta, the derived PTFs provided better accuracy, with the RMSE not exceeding 0.174 cm^3 cm^{-3}. Our results indicate that Rosetta may not be very effective in predicting water contents of calcareous soils. A similar conclusion was reached by Homaee and Farrokhian Firouzi (2008) regarding the use of Rosetta to predict water retention data of gypsiferous soils of Iran. This indicates that the performance of a PTF may be influenced by other factors such as the geographical source of the data sets used for its development, differences in measurement techniques, or other dependencies that were not considered in the analysis (e.g. Cornelis et al., 2001; Wagner et al., 2001; Ghorbani Dashtaki et al., 2010). Soil bulk density, for example, is known to play a significant role at or near saturation (Lin et al., 1999). Therefore, further studies are needed to incorporate soil structure as an input parameter to derive PTFs.

References

Batjes, N.H., 1996. Development of a world data set of soil water retention properties using pedotransfer rules. Geoderma 71, 31–52.
Bouma, J., 1989. Using soil survey data for quantitative land evaluation. Adv. Soil Sci. 9, 177–213.
Bouma, J., van Lanen, J.A.J., 1987. Transfer functions and threshold values: from soil characteristics to land qualities. In: Beek, K., Barrough, P.A., McCormack, D.D.

(Eds.), Quantified Land Evaluation, Proc. Workshop, ISSS and SSSA, Washington, DC, 27 April–2 May 1986. Int. Inst. Aerospace Surv. Earth Sci. Publ. No. 6. ITC Publ., Enschede, The Netherlands.

Cornelis, W.M., Ronsyn, J., van Meirvenne, M., Hartmann, R., 2001. Evaluation of pedotransfer functions for predicting the soil moisture retention curve. Soil Sci. Soc. Am. J. 65, 638–648.

Dane, J.J., Topp, G.C. (Eds.), 2002. Methods of Soil Analysis Part 4. Physical Methods. Soil Sci. Soc. Am., Inc., Madison, WI.

Daniel, C., Wood, F., 1980. Fitting Equations to Data: Computer Analysis of Multifactor Data, second ed. Wiley & Sons, Inc., NY.

Dirksen, C., 2000. Unsaturated hydraulic conductivity. In: Smith, K.A., Mullins, C.E. (Eds.), Soil Analysis: Physical Methods, second ed. Marcel Dekker Inc., New York, pp. 183–237.

Doner, H.E., Lynn, W.C., 1989. Carbonate, halide, sulfate and sulfide minerals. In: Dixon, J.B., Weed, S.B. (Eds.), Minerals in Soil Environments, second ed. Soil Sci. Soc. Am., Madison, WI, pp. 279–330.

Espino, A., Mallants, D., Vanclooster, M., Feyen, J., 1995. Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties. Agric. Water Manage. 29, 235–253.

FAO/UNDP. 1972. Calcareous Soils. Report of the Regional Seminar on Reclamation and Management of Calcareous Soils. 27 November–2 December, Cairo, Egypt.

Frenkel, H., Hadas, A., Jury, W.A., 1978. The effect of salt precipitation and high sodium concentration on soil hydraulic conductivity and water retention. Water Resour. Res. 14, 217–222.

Gee, G.W., Or, D., 2002. Particle-size analysis. In: Dane, J.H., Topp, G.C. (Eds.), Methods of Soil Analysis, Part 4 SSSA Book Series No. 5. Soil Sci. Soc. Am., Madison, WI, pp. 255–293.

Ghorbani Dashtaki, Sh., Homaee, M., Khodaberdiloo, H., 2010. Derivation and validation of pedotransfer functions for estimating soil water retention curve using a variety of soil data. Soil Use Manage. 26, 68–74.

Goh, T.B., Mermut, A.R., 2006. Carbonates. In: Carter, M.R., Gregorich, E.G. (Eds.), Soil Sampling and Methods of Analysis. CRC Press Taylor and Francis, Boca Raton, FL, pp. 215–223.

Ho, R., 2006. Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS. Chapman & Hall/CRC, p. 403.

Hocking, R., 2003. Methods and Applications of Linear Models. John Wiley & Sons, Hoboken, NJ.

Homaee, M., Farrokhian Firouzi, A., 2008. Deriving point and parametric pedotransfer functions of some gypsiferous soils. Aust. J. Soil Res. 46, 219–227.

Kroetsch, D., Wang, C., 2006. Particle size distribution. In: Carter, M.R., Gregorich, E.G. (Eds.), Soil Sampling and Methods of Analysis. CRC Press Taylor and Francis, Boca Raton, FL, pp. 713–725.

Legates, D.R., McCabe, G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35, 233–241.

Leij, F., Alves, W.J., van Genuchten, M.Th., Williams, J.R., 1996. The UNSODA Unsaturated Soil Hydraulic Database. User's manual, Version 1.0. EPA/600/R-96/095. National Risk Management Laboratory, Office of Research Development, Cincinnati, OH.

Lilly, A., 1997. A description of the HYPRES database (HYdraulic Properties of European Soils). In: Bruand, A., Duval, O., Wösten, J.H.M., Lilly, A. (Eds.), The Use of Pedotransfer Functions in Soil Hydrology Research, Proc. Workshop, 10–12 October 1996, Orléans, France.

Lin, H.S., McInnes, K.J., Wilding, L.P., Hallmark, C.T., 1999. Effects of soil morphology on hydraulic properties: II Hydraulic pedotransfer functions. Soil Sci. Soc. Am. J. 63, 955–961.

Minasny, B., McBratney, A.B., Bristow, K.L., 1999. Comparison of different approaches to the development of pedotransfer functions for water-retention curves. Geoderma 93, 225–253.

Minitab, 1998. Minitab for Windows, Release 12, Minitab Inc., State College, PA.

Nelson, R.E., 1982. Carbonate and gypsum. In: Page, A.L., Miller, R.H., Keeney, D.R. (Eds.), Methods of Soil Analysis. Part 2: Chemical and Microbiological Properties, second ed. Am. Soc. Agron, Madison, WI, pp. 181–198.

Nemes, A., Schaap, M.G., Wösten, J.H.M., 2003. Functional evaluation of pedotransfer functions derived from different scales of data collection. Soil Sci. Soc. Am. J. 67, 1093–1102.

Pachepsky, Y.A., Rawls, W.J., 1999. Accuracy and reliability of pedotransfer functions as affected by grouping soils. Soil Sci. Soc. Am. J. 63, 1748–1757.

Ryan, B.F., Joiner, B.L., 1994. Minitab Handbook. Durbuy Press, p. 483.

Schaap, M.G., Leij, F.J., 1998. Database-related accuracy and uncertainty of pedotransfer functions. Soil Sci. 10, 765–779.

Schaap, M.G., Leij, F.J., van Genuchten, M.Th., 2001. Rosetta: a computer program for the estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251, 151–162.

Sherrod, L.A., Dunn, G., Peterson, G.A., Kolberg, R.L., 2002. Inorganic carbon analysis by modified pressure-calcimeter methods. Soil Sci. Soc. Am. J. 66, 299–305.

Shirazi, M.A., Boersma, L., 1984. A unifying quantitative analysis of soil texture. Soil Sci. Soc. Am. J. 48, 142–147.

Soil Survey Staff, 1999. Soil Taxonomy, A Basic System of Soil Classification for Making and Interpreting Soil Surveys, second ed., USDA Soil Conservation Service, US Gov. Printing Office, Washington, DC. <http://soils.usda.gov/technical/classification/taxonomy>.

Soil Survey Staff, 2010a. National Soil Survey Characterization Data, Soil Survey Laboratory, National Soil Survey Center, USDA, Natural Resources Conservation Service, Lincoln, NE. <http://ssldata.nrcs.usda.gov/>.

Soil Survey Staff, 2010b. Keys to Soil Taxonomy. 11th ed., USDA National Resources Conservation Service, Washington, DC. <http://soils.usda.gov/technical/classification/tax_keys/>.

Tietje, O., Tapkenhinrichs, M., 1993. Evaluation of pedo-transfer functions. Soil Sci. Soc. Am. J. 57, 1088–1095.

van den Berg, M., Klant, E., van Reeuwijk, L.P., Sombroek, G., 1997. Pedotransfer functions for the estimation of moisture retention characteristics of Ferralsols and related soils. Geoderma 78, 161–180.

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

van Genuchten, M.Th., Leij, F.J., Yates, S.R., 1991. The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils. EPA/600/2-91/065, R.S. Kerr Environmental Research Laboratory, US Environmental Protection Agency, Ada, OK. p. 85.

Wagner, B., Tarnawski, V.R., Hennings, V., Müller, U., Wessolek, G., Plagge, R., 2001. Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set. Geoderma 102, 275–297.

Willmott, C.J., 1981. On the validation of models. Phys. Geogr. 2, 184–194.

Wösten, 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., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of database of hydraulic properties of European soils. Geoderma 90, 169–185.

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.