Development and comparative analysis of pedotransfer functions for predicting soil water characteristic content for Tunisian soil

Jabloun Mohamed and Sahli Ali*

Abstract - An accurate determination of the soil hydraulic characteristics is crucial for using soil water simulation models. However, these measurements are time consuming which makes it costly to characterise a soil. As an alternative, pedotransfer functions (PTFs) often prove to be good predictors for soil water contents. The purpose of this study is (i) to evaluate three well-known and accepted parametric PTFs used to estimate soil water retention curves from available soil data [1]-[2]-[3], and (ii) to derive and validate, for Tunisian soils, a more accurate point PTFs; the proposed PTFs were developed for four levels of availability of basic soil data (particle fractions, dry bulk density and organic matter content) and provide estimation for water content at 0, 100 and 1500 kPa pressure. A total of 147 Tunisian soil samples were divided into two groups; 109 for the development of the new PTFs and 38 for comparing the reliability of the tested PTFs against the derived ones. This data set contains measured soil water retention data, the dry bulk density, sand, silt, clay percentages and the organic matter content. The accuracy and reliability of the predictions were evaluated by the coefficient of determination (R^2) and the Root Mean Square Error (RMSE) between the measured and predicted values. The developed PTFs have a good accuracy with R^2 and RMSE ranged from 0.40 to 0.86 and from 2.32 to 4.85%, respectively. We found that adding bulk density and soil organic matter as predictors in addition to sand, silt and clay percentages increased the prediction capability by 1%. For the performance test, the water retention contents were better predicted by the developed than by the tested PTFs. The RMSE ranged from 2.57 to 6.87 for the derived PTFs compared with 3.34 to 13.61 for the tested PTFs. HYPRES had the lowest RMSE among the tested PTFs. For the derived PTFs, The R^2 values varied from 0.34 to 0.68. The lowest RMSE as well as the highest R^2 values were found for the PTF that used the detailed data.

***Index Terms*— Field capacity, Pedotransfer function, Tunisian soils texture, Wilting point.**

INTRODUCTION

Knowledge of the soil hydraulic properties is indispensable to solve many soil and water management problems related to agriculture, ecology, and environmental issues. These properties are needed to describe and predict water and solute transport, as well as to model heat and mass transport near the soil surface. One of the main soil hydraulic properties is the water retention curve, as it expresses the relationship between the matric potential and the water content of the soil. It can be considered of great importance in present-day agricultural, ecological, and environmental soil research. Unfortunately, direct measurement of this property is labor intensive and impractical for most applications in research and management, generally cumbersome, expensive and time consuming, especially for relatively large-scale problems. Alternatively, the indirect methods are increasingly used to estimate the soil hydraulic characteristics. One of the indirect methods, which are used to estimate the soil hydraulic properties, is Pedotransfer function (PTF). Pedotransfer functions can be defined as predictive functions of certain soil properties from other easily measured properties such as particle-size distribution (sand, silt and clay content), organic matter or organic C content, bulk density, porosity, etc. [4]-[5].

Pedotransfer functions may be categorized into "class" and "continuous" PTFs [6]. Class PTFs predict certain soil properties based on the soil classes to which the soil sample belongs [7]. Continuous PTFs predict certain soil properties as a continuous function of one or more measured variables. This latter type of pedotransfer function can also be classified as single point and parametric regressions [6]-[8]. Single point PTFs predict a soil property at a special point of the water retention curve [9] or available water capacity [10]. The parametric PTFs aim to predict the parameters of a model, and the most widely used soil hydraulic model is the van Genuchten function [11]:

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{\left(1 + (\alpha h)^n\right)^m} \qquad (1)$$

Where, θ_r and θ_s are the residual and saturated water content, α is the scaling parameter, n is the curve shape factor and m is an empirical constant, which can be related to n by:

$$m = 1 - 1/n. \qquad (2)$$

Many attempts have been made to correlate the parameters of this equation to basic soil properties, e.g. [1]-[2]-[3]-[12]. The latter type of PTFs is more suitable for modelling purposes as the equations describe the whole of the water-release curve, thus allowing the computation of water content at arbitrary pressures. But, there are indications that the parametric approach may lead to lower accuracy in water retention predictions compared with the point-based approach [13].

PTFs can be obtained by various mathematical methods. Until recently, most PTFs were derived through multiple regression methods [1]-[6]-[14]. But the artificial neural networks approach [15]-[16] is becoming more and more popular. An advantage of neural networks over traditional PTFs is that they do not require a priori model concept. The optimal and possibly nonlinear relations that link input data (particle-size data and bulk density, etc.) to output data (hydraulic parameters) are obtained and implemented in an iterative calibration procedure. As a result, neural network models typically extract the maximum amount of information from the data [17]; performances equivalent or superior to PTFs derived by regression-type methods have been reported [16]-[18].

Regardless of the methodology used to derive them, PTFs are developed on the basis of databases of a limited number of soil samples. Consequently any PTF is likely to give less accurate or possibly even very poor predictions if used outside the range of soils from whose they were derived [8]. Thus, the predictive ability of PTFs is somewhat related to the similarity between the data set used in developing and testing the PTF [19]. It is important, therefore, to evaluate how well the PTFs will perform when applied outside the range of the data that were used to derive them.

The objective of this study was, therefore, (i) to evaluate the general applicability and the prediction accuracy of three of the most commonly cited parametric PTFs to predict the water retention curve to a tunisian area, where climatic and geological conditions are different from those that prevailed for the establishment of the tested PTFs; (ii) to derive and validate, for Tunisian soils, a more accurate PTF to estimate the characteristic soil water content needed for irrigation planning; (iii) to compare prediction reliability of the derived and tested PTFs.

MATERIALS AND METHODS

THE DATA SET

The soil data set used for this study was that of [20] and [21]. It contains Cambisols, Vertisols, Calcisols and Fluvisols mainly from the Ariana plain with some from Mornag, Tunisia [20]. The data set available contained 147 horizons, including basic soil properties: three texture fractions as defined by USDA system sand (2 - 0.05 mm), silt (0.05 - 0.002 mm), and clay (<0.002 mm); bulk density, organic matter content and water content at saturation, field capacity and at wilting point.

The sand fraction was determined by dry sieving and the silt and clay fractions by the pipette method after pre-treatment with hydrogen peroxide and sodium hexametaphosphate. Bulk density was determined from undisturbed soil cores collected in 100 cm3 cylinders after drying at 105°C for 24 h. Organic matter, another variable used in the PTFs, was estimated by wet digestion in acid dichromate and automatic titration with iron sulphate. Water retention data were obtained from undisturbed soil cores using the same methodology: tension tables between saturation and -10 kPa, and pressure chambers between -10 kPa and -1500 kPa. The distribution of soil texture across the horizons used in the study is shown in Fig. 1.

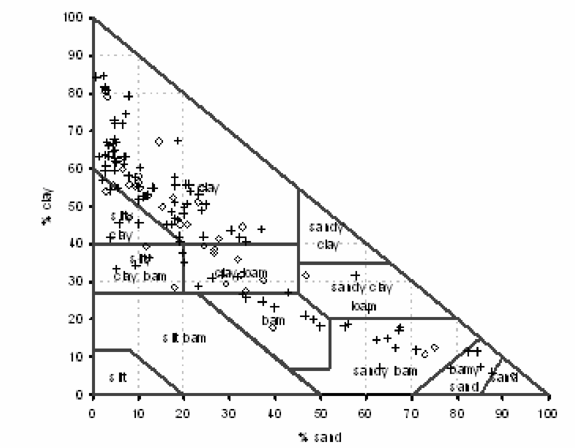
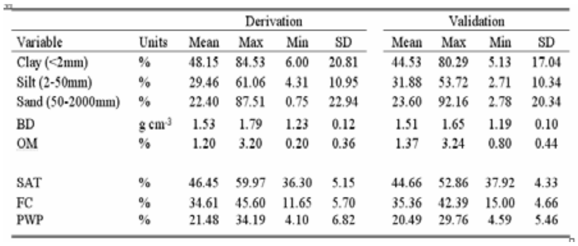


Fig 1. The distribution of soil texture within the derivation (+) and validation (◊) data sets (144 points), shown on the textural triangle

The selected data were subdivided into two parts. Derivation data set (109 samples) was used for evaluating the use of the tested PTFs and for the development of the new point PTFs. Validation data set was set aside as an independent data set that served to objectively evaluate and compare the reliability of the tested PTFs against the derived ones.

Table 1. Ranges of soil texture (according to USDA classification) and other basic information in the data sets used for pedotransfer functions derivation and validation.



		Derivation				Validation			
Variable	Units	Mean	Max	Min	SD	Mean	Max	Min	SD
Clay (<2mm)	%	48.15	84.53	6.00	20.81	44.53	80.29	5.13	17.04
Silt (2-50mm)	%	29.46	61.06	4.31	10.95	31.88	53.72	2.71	10.34
Sand (50-2000mm)	%	22.40	87.51	0.75	22.94	23.60	92.16	2.78	20.34
BD	g cm^{-3}	1.53	1.79	1.23	0.12	1.51	1.65	1.19	0.10
OM	%	1.20	3.20	0.20	0.36	1.37	3.24	0.80	0.44
SAT	%	46.45	59.97	36.30	5.15	44.66	52.86	37.92	4.33
FC	%	34.61	45.60	11.65	5.70	35.36	42.39	15.00	4.66
PWP	%	21.48	34.19	4.10	6.82	20.49	29.76	4.59	5.46

This data set consisted of 38 horizons. Table 1 summarises the characteristics of both data sets in terms of means, maximum, minimum and standard deviations (SD) for the data sets used in the derivation and validation of the PTFs. The data in Table 1 suggest that both data sets are largely similar.

THE TESTED PTFS

In the present study, three frequently used parametric PTFs are evaluated, namely those by [1], [2] and [3]. The three PTFs will be referred to hereinafter as HYPRES, Vereecken and Saxton, respectively. All of them are based on a series of multiple linear regression equations that link the water content at a certain matric-potential to some of the soil physical properties, such as particle-size distribution, organic matter content and bulk density. PTFs equations are not reported here for the sake of brevity, but can be found in the cited publications.

Reference [1] used multiple linear regression to predict the parameters of the Van Genuchten equation using data from 4030 horizons from all over Europe (the HYPRES database). A class as well as a continuous PTF were developed. The class PTF was obtained by subdividing the database into 11 soil textural classes, and it gives the Van Genuchten parameters [11] in tabular format, whereas the continuous PTF does not consider any grouping. Input data needed are sand, silt and clay content, bulk density, organic matter content and a qualitative variable, indicating whether topsoil or subsoil is considered. Only the continuous PTF was evaluated.

Reference [2] presented several equations for estimating parameters of a modified form of the Van Genuchten function [11] by introducing the simplifying assumption m = 1 (Eq. 1). The parameter estimation was performed through multiple regression using two sets of soil properties as predictor variables: a first set composed of the sand and clay fraction, the carbon content and the bulk density and a second set taking into account more detailed information on the particle-size distribution. The former approach was used in the present study; the organic matter content was converted into carbon content by dividing it by a coefficient of 1.724.

Reference [3] used the data of [22] to derive equations that cover the whole range of matric-potential values, instead of only 12 selected values. They used the regression approach of [22] that is least input-data demanding, from which they removed the influence of the BD and in which the OM was fixed to 0.66%, the average value reported by [22]. Thus, the approach in [3] only needs texture (sand, clay) data. This method may be applied for soils of 5<sand<30% with 8<clay<58%, and 30<sand<95% with 5<clay<60%.

These PTFs were used to predict the water content at saturation (SAT), field capacity (FC) and at permanent wilting point (PWP).

THE DERIVED PTFS

The soil water contents at saturation, field capacity and wilting point are used to calculate the water depth that should be applied by irrigation [23], and to determine water availability, which is a crucial factor in assessing the suitability of a land area for producing a given crop [24]. Since the aim of this study was to find a suitable PTF to include in an irrigation scheduling model, it would, therefore, be useful to develop new point PTFs instead of parametric ones to predict the water content at these characteristic matric-potentials.

The advantage of this latter approach is that fairly accurate predictions can be made for specific points along the water retention curve. On the other hand, it offers insight into which soil properties are relevant for predicting the water content at SAT, FC and PWP [6]. In addition, [13] reported that the parametric approach may lead to lower accuracy in water retention predictions compared with the point-based approach.

The point PTFs were developed using the entire calibration data set (n= 109) following the multiple linear regression approach used by [1]. The procedure is outlined here.

Step 1: Multiple linear regression techniques were used to relate the water retention at SAT, CC and PWP to texture, bulk density and organic matter content. Linear, reciprocal, and logarithm of these basic soil properties were used in the regression analysis, and possible interactions were also investigated. Logarithmic transformation of the input variables was found useful by [25]. The equation had the following form:

$$\begin{aligned}\theta_i &= a_{i,1} + a_{i,2}Sa + a_{i,3}Si + a_{i,4}Cl + a_{i,5}OM + a_{i,6}BD \\ &+ a_{i,7}\,{}^{1}\!/_{Sa} + \ldots + a_{i,8}\,{}^{1}\!/_{BD} + a_{i,9}SaSi + \ldots + a_{i,j}SiCl \\ &+ a_{i,j+1}Sa^2 + \ldots a_{i,j+2}Cl^2 + a_{i,j+3}Ln(Sa) + \ldots a_{i,j+4}Ln(Cl)\end{aligned} \tag{3}$$

where X_i is the value of the water content (i = 1 to 3 corresponding to SAT, CC and PWP, respectively); Sa, Si and Cl are, respectively, the percentages of sand (2-0.05 mm), silt (0.05- 0.002 mm) and clay (< 0.002 mm); BD is the bulk density (g cm^{-3}); OM is the percentage of organic matter and $a_{i,j}$ (j = 1...n) are coefficients derived by multiple linear regression.

Step 2: For the sake of parsimony, the number of parameters of Eq. (3) was reduced using stepwise techniques leaving in the final equation only variables that explained a significant proportion of the parameter variability. Practical applications of most PTFs are often hampered by their very specific data requirements. Some authors established PTFs that provided the best results for their data set, which sometimes produced models that require many input variables [26] or detailed particle size distributions [27]-[28]. However, users of PTFs are frequently confronted with situations where one or several input variables needed for a PTF are not available. It would therefore be useful if PTFs could accept input data with varying degrees of detail. Since basic data from soil surveys do not always provide information such as bulk density and organic matter, four equations were derived for each water content, depending on the amount of information available: model M1 included all basic information (sand, silt, clay, bulk density, and organic matter content); model M2 excluded bulk density; model M3 excluded organic matter content; and model M4, excluded both bulk density and organic matter content. The different equations of the newly developed models are given in Annex I.

Step 3: Finally, the proposed PTFs were validated using data not including in the fitting procedure (the validation data set). Comparisons of the performance of the newly derived PTFs with those of [1], [2] and [3] were carried out.

EVALUATION CRITERIA AND UNCERTAINTY ANALYSIS

Usually, a common method to evaluate models is to plot the measured values against the predicted values and the correlation between them is used for model evaluation (coefficient of determination R2). Ideally, this relationship should be linear with a slope of unity and intercept of zero. Although this method may be satisfactory for fitting an empirical model to observed data, it is inadequate for evaluating the performance of mechanistic models [29]. Generally, correlation-based statistics in conjunction with two other statistics, root-mean-squared error (RMSE), and mean error (ME), also called bias, are used to evaluate the performance of models. The RMSE is always positive; it equals zero only if all measured water contents equal the predicted water contents. The RMSE is an index of the correspondence between measures and predicted water contents, which represents the expected magnitude of error. Negative and positive values of ME indicate under and over-estimation of PTFs for a given parameter, respectively. Their definitions are given below:

$$R^2 = 1 - \frac{\sum_1^N (y_i - \hat{y}_i)^2}{\sum_1^N (y_i - \bar{y}_i)^2} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_1^N (\hat{y}_i - y_i)^2} \tag{5}$$

$$ME = \frac{1}{N}\sum_1^N (\hat{y}_i - y_i) \tag{6}$$

Where, y_i denotes the measured value, $\hat{y}_i$ is the predicted value, $\bar{y}_i$ is the average of the measured value y, and N is the total number of observations. These three criteria were used in the evaluation of prediction accuracy of both tested and generated PTFs.

The PTF accuracy is assessed from the correspondence between measured and estimated data for the data set from which a PTF has been developed. In contrast to accuracy, the reliability of a PTF needs to be assessed from the correspondence between measured and estimated data for the data set other than the one used to develop a PTF.

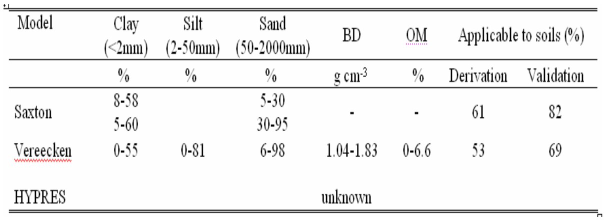
RESULTS AND DISCUSSION

EVALUATION OF THE TESTED PTFS

Initially, we determined whether the samples in the derivation data set fit within the ranges of the tested PTFs (Table 2). Although some soil property values from our data set fell outside these ranges, the evaluation of each PTF was not only conducted on soils within these ranges. It was done on all sampled soils, since one of the objectives of the study was to evaluate PTFs for a whole range of soils in the study area.

When considering the overall applicability of the retained PTFs for evaluation, Table 2 reveals that the PTFs of Saxton and Vereecken cover, respectively, less than 61 and 53% of our derivation soil samples. The Saxton PTF [3] does not include soils with very high clay content (>60%), whereas in case of the Vereecken PTF [2] the soil samples with very high silt and clay content fall out of the range of the derivation data set. Thus, as both the PTF of Saxton and Vereecken were developed on soils with clay content lower than 60 and 55%, respectively, they do not cover the clayey soils of our data set. Since no information about the range of the applied soil properties was available for the HYPRES PTF [1], its applicability could not be tested.

Table 2. Soil property ranges of the data sets used to evaluated pedotransfer functions (PTFs)

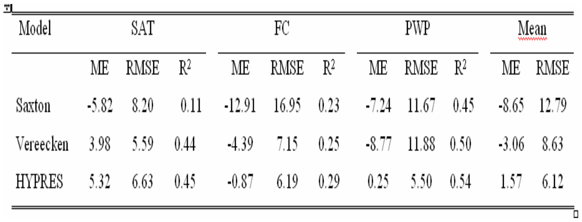


Model	Clay (<2mm)	Silt (2-50mm)	Sand (50-2000mm)	BD	OM	Applicable to soils (%)	
	%	%	%	g cm^{-3}	%	Derivation	Validation
Saxton	8-58 5-60		5-30 30-95	-	-	61	82
Vereecken	0-55	0-81	6-98	1.04-1.83	0-6.6	53	69
HYPRES				unknown			

Although most of the PTFs were not applicable to all the soil samples of the derivation data set, the PTFs were evaluated on the complete derivation data set. The values calculated for the different evaluation criteria are presented in Table 3.

When considering the mean of MEs, it can be observed that the Saxton and Vereecken PTFs tend to underestimate the water content (Table 3). Whereas the HYPRES PTF shows a tendency to overestimate the water content at SAT and PWP. Table 3 reveal that, in the case of the Saxton and Vereecken PTFs, this underestimation occurs at FC and PWP water content. The Vereecken PTF underestimates mainly at moisture contents below saturation which is in contradiction with findings of [30]. He reported a slight tendency to overestimate water content at -1500 kPa matric potential in case of the Vereecken PTF. The overestimation of the HYPRES PTF is pronounced near saturation, that is, at a matric potential of -0.3 kPa which is in concordance with the finding of [31]. However, the Saxton and Vereecken PTFs underestimate considerably at the wilting point, that is, at a matric potential of -1500 kPa. The Saxton PTF also underestimates considerably at the field capacity. The PTF of [1] show slight bias between the data sets, whereas the bias observed for the two other PTFs is considerable.

Table 3. Evaluation criteria of the tested PTFs as computed on the complete derivation data set.



Model	SAT			FC			PWP			Mean	
	ME	RMSE	R^2	ME	RMSE	R^2	ME	RMSE	R^2	ME	RMSE
Saxton	-5.82	8.20	0.11	-12.91	16.95	0.23	-7.24	11.67	0.45	-8.65	12.79
Vereecken	3.98	5.59	0.44	-4.39	7.15	0.25	-8.77	11.88	0.50	-3.06	8.63
HYPRES	5.32	6.63	0.45	-0.87	6.19	0.29	0.25	5.50	0.54	1.57	6.12

As regards the mean of RMSEs, again the HYPRES PTF shows the lowest values, meaning that the predicted water content follows the measured water content relatively well for the three characteristic points. The highest values result from the Saxton and Vereecken PTFs.

The R^2 values reveal a somewhat different pattern in terms of the model's performance. The correspondence between measured and predicted water content is still highest for the HYPRES PTF and lowest for the Saxton PTF especially at SAT point. It can be deduced that the results shown in table 3 clearly show the risks of using PTFs outside the range of texture from which they were derived and for which they are valid.

Most of the models predict best near saturation. The Saxton PTF does not perform very well at SAT, FC and PWP. It has the biggest RMSE and ME (ranging 8.20 from to 16.95 and from -12.91 to 65.82, respectively) and the lowest R^2 (ranging from 0.11 to 0.45). The low performances of the three evaluated PTFs can be attributed to their narrow applicability to the Tunisian soils. Therefore, a more accurate point PTFs were drived to fit the Tunisian soils.

DATA ANALYSIS

Fig. 2 highlights some of the relationships between dependent and independent variables that showed significant ($p < 0.05$) correlations for the derivation data set: in the case of SAT, FC and PWP (%). the only significant correlations were those with clay and sand content.

A negative relationship between the characteristic water content and sand content was determined, whereas a positive relationship between the characteristic water content and clay content was found which is in agreement with the finding of [9]. In addition, they reported a negative effect of the bulk density on the field capacity and permanent wilting point. Whereas we found that the negative effect of the bulk density was only significant on the saturation point. The silt fraction and the organic matter had no significant effect on the variation of the water content at the three characteristic points. The relationship between the permanent wilting point and clay fraction was better than the relationship between the field capacity and clay fraction.

Because the field capacity is the soil moisture content retained in the soil pores against the gravitational force, there might be a close relationship between the field capacity and soil pore size distribution. The clay fraction and organic matter content might affect the field capacity positively by increasing the soil pores retaining the water against the gravitational force, whereas the sand fraction might affect the field capacity negatively by increasing the soil pores allowing the free flow of soil water [9].

ACCURACY AND PERFORMANCE OF THE NEWLY DERIVED PTFS

Results of applying the derived PTFs in order to estimate the soil water content at saturation, field capacity and permanent wilting point from the available basic soil information are shown in Table 4. The RMSE values in Table 4 exhibit a trend of improvement from model M4 to model M1. Including more basic soil parameters in the regression equation generally leads to better estimates.

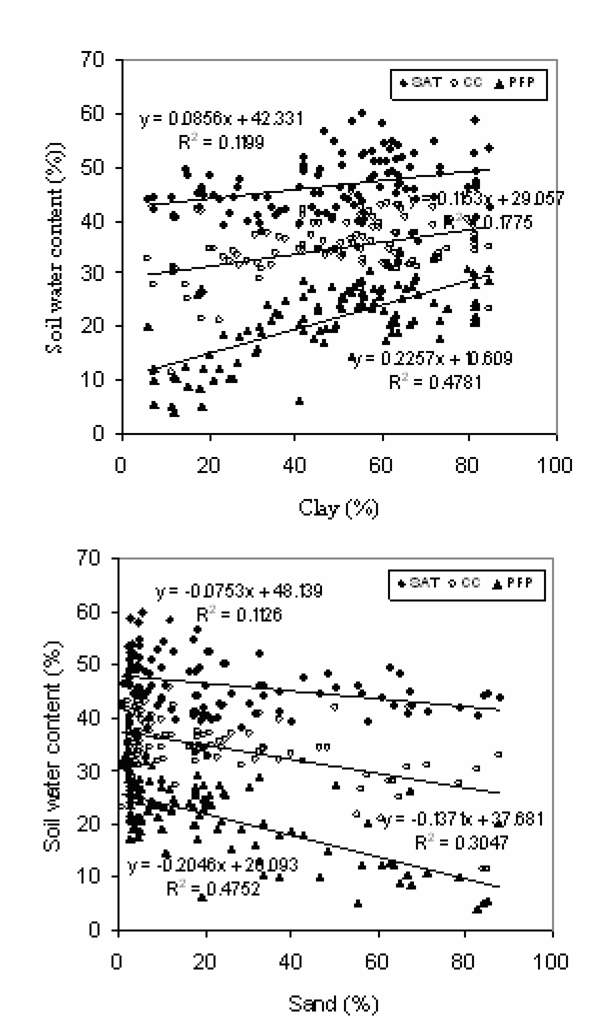


Fig. 2. The relationships between the soil water contents at saturation (SAT), field capacity (FC) and permanent wilting point (PWP) and the clay and sand content.

Estimations using the new PTFs show RMSE values ranging from 2.32 to 4.85. These are relatively accurate estimations when compared with estimations that appear in literature [6]. The RMSE values reported in literature range from 2 to 7 both for FC and PWP [32]. The R² values also show a trend of improvement as the type of soil parameters changes throw model M4 to model M1 and this is true for the three characteristic points. Most bias is introduced by the PTFs M2 and M3 at the permanent wilting point; it's about -0.24 and -0.16, respectively.

Table 4. Evaluation criteria of the derived PTFs as computed on the complete derivation data set.

Model	SAT			FC			PWP			Mean	
	ME	RMSE	R2	ME	RMSE	R2	ME	RMSE	R2	ME	RMSE
M1	0.08	2.73	0.72	0.03	3.21	0.76	-0.02	2.32	0.86	0.03	2.77
M2	0.04	3.89	0.42	-0.02	3.83	0.54	-0.24	4.75	0.62	-0.07	4.18
M3	0.04	2.76	0.71	-0.02	4.25	0.44	-0.16	4.75	0.61	-0.05	4.01
M4	0.002	3.94	0.40	0.09	4.28	0.43	0.03	4.85	0.59	0.04	4.38

In the case of the Model M1, which included all basic information (sand, silt, clay, bulk density, and organic matter);

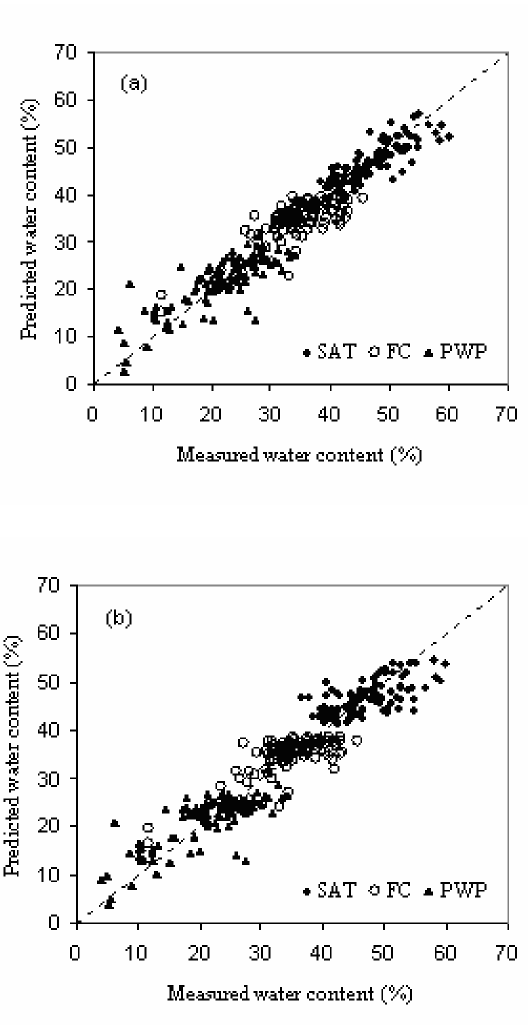


Fig. 3. Predicted water content as a function of measured water content (a) in the case of model M1 (b) in the case of model M2

The level of accuracy is fairly high with the highest R² and lowest RMSE. The observed improvement of including the hole basic data on the accuracy to estimate the water content is similar to the one reported by several study [18]. Of the three characteristic water contents, the permanent wilting point is the one whose estimation is characterized by a higher degree of accuracy ($R^2 = 0.86$ and RMSE = 2.32).

Fig. 3a shows the predicted water content as a function of measured water content at the three characteristic points as predicted by models M1.

Model M2 has a root mean square error of 3.89 at the SAT point whereas model M3 has a value of about 2.76 which is closer to the value of RMSE obtained by the model M1. Hence, omitting BD in model M2 led to an increase in RMSE especially at very low potential, showing the crucial role of BD in the prediction of the water content near saturation. [22] and [33] found similar results when calibrating multiple linear functions and artificial neural network respectively, to predict water content in the same pressure head range. It is important to note also that the organic matter content was a better complimentary predictor of water retention at FC as compared with bulk density (Table 4), and, therefore, if available, may be a preferable predictor of water retention which is in contradiction with findings of [34].

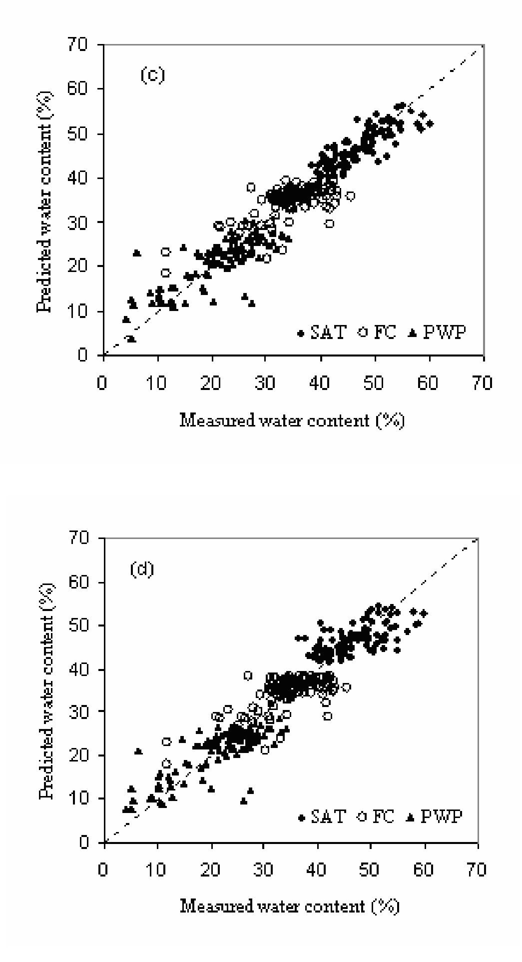


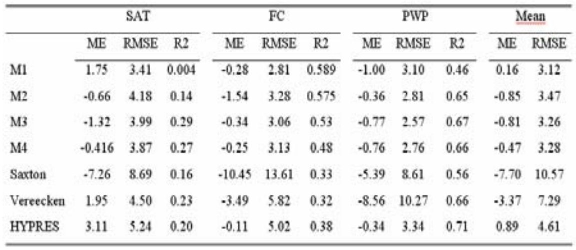
Fig. 3. Predicted water content as a function of measured water content (c) in the case of model M3 (d) in the case of model M4

In addition, he reported that bulk density values could represent the effect of organic carbon on water retention if the primary effect of organic matter was changing bulk density. Reference [22] demonstrated a correlation between bulk density and organic matter content in his data sets and indicated that bulk density effectively substituted organic matter. In our case the correlation between organic matter content and bulk density is very low. That's why using both those properties along with texture as the water retention predictors (model M1) leads to the best overall accuracy (RMSE < 3.2).

Reference [34] concluded that organic carbon and bulk density improve estimates of soil water retention derived from soil texture. [30] and [35], all found that inclusion of organic carbon content as an input to PTFs was useful in improving estimates of soil water at FC and PWP. [36] also saw the need to use organic carbon content in estimating water content at wilting point.

The performance of the newly developed PTF models was evaluated (Table 5) by comparing the RMSE values with predictions of those of Saxton, Vereecken and HYPRES.

Table 5. Evaluation criteria of the derived and tested PTFs as computed on the complete validation data set



	SAT			FC			PWP			Mean	
	ME	RMSE	R2	ME	RMSE	R2	ME	RMSE	R2	ME	RMSE
M1	1.75	3.41	0.004	-0.28	2.81	0.589	-1.00	3.10	0.46	0.16	3.12
M2	-0.66	4.18	0.14	-1.54	3.28	0.575	-0.36	2.81	0.65	-0.85	3.47
M3	-1.32	3.99	0.29	-0.34	3.06	0.53	-0.77	2.57	0.67	-0.81	3.26
M4	-0.416	3.87	0.27	-0.25	3.13	0.48	-0.76	2.76	0.66	-0.47	3.28
Saxton	-7.26	8.69	0.16	-10.45	13.61	0.33	-5.39	8.61	0.56	-7.70	10.57
Vereecken	1.95	4.50	0.23	-3.49	5.82	0.32	-8.56	10.27	0.66	-3.37	7.29
HYPRES	3.11	5.24	0.20	-0.11	5.02	0.38	-0.34	3.34	0.71	0.89	4.61

From these testing results it follows that of the evaluated PTFs Saxton and Vereecken had the same high mean RMSE, whereas the predictions by the HYPRES model gave smaller errors because of a good performance at PWP.

The RMSE values of the M1 model show less good prediction at PWP but show good predictions at SAT and FC as compared to the predictions by the three other derived models (M2, M3 and M4). The results obtained from the validation data set show that the overall predictions capability of the four derived models (M1, M2, M3 and M4) is better the tested PTFs Saxton and Vereecken whereas the HYPRES model has similar errors.

CONCLUSION

In the present study we used multiple linear regression analysis to derive the best relationship between the water retention at saturation, field capacity and wilting point and the basic soil properties. We examined the effect of excluding bulk density, and soil organic matter from predictor parameters. The improvement in predictions using more detailed input soil data (sand, silt, clay, bulk density, and organic matter content) was generally better. The lowest RMSE as well as the highest R2 were found for the model that used the most detailed data.

Reliability of the developed PTF models (M1 to M4) were tested against other well-known PTFs. The validation procedure, using an independent data set, demonstrated the ability of the proposed PTF to accurately predict the water retention curve of Tunisian soils, and also revealed the shortcomings of PTFs derived using soils data from temperate regions in predicting the water retention curves of Tunisian soils. The limitations of [2] and [3] PTFs when applied to Tunisian soils were evident, even within the range of validity (textures) for which they had been derived.

The use of the developed PTF is attractive due to their generally low error level and their flexibility to input parameters. The newly derived pedotransfer functions would allow the examination of for example the effect of reductions in soil organic material content or changes of bulk density (tillage or soil compaction) on the soil hydraulic properties. Such information may then be used in studies like vulnerability of soils to physical or chemical degradation. To improve predictions further additional soil parameters may be included such as parameters describing soil structure. Another alternative could be the development of soil class specific PTF models.

The derived equations were incorporated into a graphical computer program to readily estimate water retention at saturation, field capacity and wilting point. Texture is selected from the texture triangle and text boxes allow introducing the content of the organic matter and the bulk density. The results are dynamically displayed in datasheets or graphs as the inputs are varied. This provides a rapid and visual display of the estimated water holding and transmission characteristics over a broad range of variables.

APPENDIX

The equations of the developed PTFs models are given below:

Model M1

$$\theta_{Sat} = -0.8667(Cl) - 1.426(Sa) - 84.2817(BD) - 0.0151(Si)^2 + 0.0012(Sa)^2 - 7.9188\mathrm{Ln}(Si) + 112.0333\mathrm{Ln}(BD) - 0.0064(Cl)(Si) - 0.2835(Cl)(BD) - 0.0068(Si)(Sa) + 0.177(Si)(OM) + 266.768$$

$$\theta_{FC} = 0.0023(Si)^2 - 8.491(BD)^2 - 3.2498(OM)^2 + 153.59021/(Cl) - 101.21431/(Si) -9.02181/(Sa) + 8.52011/(OM) + 20.9002\mathrm{Ln}(OM) + 0.355(Cl)(BD) - 0.2388(Cl)(OM) + 10.1357(BD)(OM) + 16.4788$$

$$\theta_{PWP} = -0.7409(Si) + 0.0126(Si)^2 - 7.4396(BD)^2 - 2.8807(OM)^2 + 136.151/(Cl) - 98.33231/(Si) - 23.99671/(Sa) + 9.03681/(OM) + 20.7999\mathrm{Ln}(OM) + 0.4598(Cl)(BD) - 0.2579(Cl)(OM) + 9.1905(BD)(OM) + 9.8444$$

Model M2

$$\theta_{Sat} = 0.7264(Si) + 0.2026(Sa) - 0.0083(Si)^2 - 13.75491/(Sa) - 7.7387\mathrm{Ln}(Sa) + 2.2103\mathrm{Ln}(OM) - 0.0043(Cl)(Si) + 0.0051(Cl)(Sa) - 0.0047(Si)(Sa) + 53.4646$$

$$\theta_{FC} = 0.2239(Cl) - 57.95441/(Si) - 11.69741/(Sa) + 6.90031/(OM) - 3.5324\mathrm{Ln}(Sa) + 24.0966\mathrm{Ln}(OM) + 0.0031(Cl)(Sa) - 0.1886(Cl)(OM) + 36.7918$$

$$\theta_{PWP} = -181.7238(Cl) - 183.5092(Si) - 182.4525(Sa) - 0.0048(Cl)^2 + 0.0114(Si)^2 - 0.0031(Sa)^2 + 128.78961/(Cl) - 83.0451/(Si) + 6.52931/(Sa) + 9.18951/(OM) + 27.4919\mathrm{Ln}(OM) + 0.0043(Cl)(Si) - 0.2411$$

Model M3

$$\theta_{Sat} = 0.4602(Cl) + 1.1343(Si) - 86.8963(BD) - 0.011(Si)^2 - 9.4193\mathrm{Ln}(Si) + 110.5222\mathrm{Ln}(Da) -0.256(Cl)Da - 0.002(Si)(Sa) + 0.0405(Sa)Da + 135.5837$$

$$\theta_{FC} = 148.39031/(Cl) - 43.85161/(Si) - 5.17411/(Sa) + 16.6718\mathrm{Ln}(Cl) + 0.0011(Cl)(Si) -0.0999(Cl)Da + 0.0025(Si)(Sa) - 24.1522$$

$$\theta_{PWP} = -1.2152(Si) - 0.4877(Sa) - 0.0057(Cl)^2 + 0.0087(Si)^2 + 85.84361/(Cl) - 88.0331/(Si) + 0.0012(Cl)(Si) + 0.2129(Cl)Da + 59.6137$$

Model M4

$$\theta_{Sat} = 0.6658(Si) + 0.1567(Sa) - 0.0079(Si)^2 - 12.31121/(Sa) - 6.4756\mathrm{Ln}(Sa) - 0.0038(Cl)(Si) + 0.0038(Cl)(Sa) - 0.0042(Si)(Sa) + 52.7526$$

$$\theta_{FC} = 118.932(Cl) + 119.0866(Si) + 119.1104(Sa) + 162.31731/(Cl) - 46.21921/(Si) -5.12991/(Sa) + 18.1733\mathrm{Ln}(Cl) + 0.0013(Cl)(Si) + 0.0022(Si)(Sa) - 11939.3493$$

$$\theta_{PWP} = -1.5722(Si) - 0.5423(Sa) - 0.0072(Cl)^2 + 0.0072(Si)^2 - 0.0059(Sa)^2 + 160.14591/(Cl) + 6.60011/(Sa) + 0.0022(Cl)(Si) - 0.0039(Cl)(Sa) + 92.3851$$

REFERENCES

[1] J. H. M. Wösten, A. Lilly, A. Nemes and C. Le Bas, "Development and use of a database of hydraulic properties of European soils", *Geoderma*, vol. 90, pp. 196-185, 1999.

[2] H. Vereecken, J. Maes, J. Feyen and P. Darius, "Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content", *Soil Science*, Vol. 148, pp. 389-403, 1989.

[3] K. E. Saxton, W. J. Rawls, J. S. Romberger and R. I. Papendick, "Estimating generalised soil-water characteristics from texture", *Soil Sci. Soc. Am. J.* vol. 50, pp. 1031-1036, 1986.

[4] J. Bouma, "Using soil survey data for quantitative land evaluation", *Adv. Soil Sci.*, vol. 9, pp. 177– 213, 1989.

[5] A. B. McBratney, B. Minasny, S. R. Cattle and R. Willem Vervoort, "From pedotransfer functions to soil inference systems", *Geoderma*, vol. 109, pp. 41-73, 2002.

[6] J. H. M. Wösten, Ya.A. Pachepsky and W. J. Rawls, "Pedotransfer functions: Bridging the gap between

available basic soil data and missing soil hydraulic characteristics", *J. Hydrol.*, vol. 251, pp. 123-150, 2001.

[7] J. H. M. Wösten, P. A. Finke and M. J. W. Janes, "Comparison of class and Continuous pedotransfer functions to generate soil hydraulic characteristics", *Geoderma*, vol. 66, pp. 227-237, 1995.

[8] M. G. Hodnett and J. Tomasella, "Marked differences between van Genuchten soil water-retention parameters for temperate and tropical soils: a new water-retention pedo-transfer functions developed for tropical soils", *Geoderma*, vol. 108, pp. 155-180, 2002.

[9] B. Cemek, R. Meral, M. Apan and H. Merdun, "Pedotransfer Functions for the Estimation of the Field Capacity and Permanent Wilting Point", *Pakistan Journal of Biological Sciences*, vol. 7, n°4, pp. 535-541, 2004.

[10] M. Van den Berg, E. Klämt, L. P. van Reeuwijk and W. G. Sombroek, "Pedotransfer functions for the estimation of moisture retention characteristics of Ferrasols and related soils", *Geoderma*, vol. 78, pp. 161-180, 1997.

[11] M. Th. Van Genuchten, "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils", *Soil Sci. Soc. Am. J.*, vol. 44, pp. 892–898, 1980.

[12] A. C. Scheinost, W. Sinowski and K. Auerswald, "Regionalization of soil water retention curves in a highly variable soilscape: I. Developing a new pedotransfer function", *Geoderma*, vol. 78, pp. 129-143, 1997.

[13] J. Tomasella, Ya. Pachepsky, S. Crestana and W. J. Rawls, "Comparison of Two Techniques to Develop Pedotransfer Functions for Water Retention", *Soil Sci. Soc. Am. J.*, vol. 67, pp. 1085-1092, 2003.

[14] W. J. Rawls and D. L. Brakensiek, "Prediction of soil water properties for hydrologic modelling", in *Proc. Symp. Watershed Management in the Eighties*, New York, 1985, pp. 293-299.

[15] Ya.A. Pachepsky, D. Timlin and G.Varallyay, "Artificial neural networks to estimate soil water retention from easily measurable data", *Soil Sci. Soc. Am. J.*, vol. 60, pp. 727-733, 1996.

[16] B. Minasny, A. B. McBratney and K. I. Bristow, "Comparison of different approaches to the development of pedotransfer functions for water retention curves", *Geoderma*, vol. 93, pp. 225-253, 1999.

[17] M. G. Schaap, F.J . Leij and M. Th. Van Genuchten, "Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions", *J. Hydrol.*, 251, pp. 163-176, 2001.

[18] A. Nemes,, M. G. Schapp and J. H. M. Wösten, "Functional evaluation of pedotransfer functions derived from different scales of data collection", *Soil Sci. Soc. Am. J.*, vol. 67, pp. 1093-1102, 2003.

[19] J. Tomasella, M. G. Hodnett and L. Rossato, "Pedotransfer functions for the estimation of soil water retention in Brazilian soils", *Soil Sci. Soc. Am. J.*, vol. 64, pp. 327-338, 2000.

[20] A Boden, "Fysische karakterisatie van de geïrrigeerde bodems in de treek rond Tunis", Ir. Thesis, Dept. Irrigation and Water Management, Katholieke Universiteit Leuven, Leuven, Belgium, 2000.

[21] E. Van Laare, "Caractérisation des sols de la région de Tunis". Mémoire Ingénieur, Dept. Génie Rural, Université Catholique de Louvain, Louvain La Neuve, Belgique, 2001.

[22] W. J. Rawls, "Estimating soil bulk density from particle size analysis and organic matter content", *Soil Sci.*, vol. 135, pp. 123-125, 1983.

[23] V. E. Hansen, O. W. Israelsen and G. E. Stringham, *Irrigation Principles and Practices*, 4th ed., New York: Wiley, 1980.

[24] C. Sys , E. Van Ranst and J. Debaveye, *Land Evaluation. Part I*, General Administration for Development Cooperation Eds., Brussels, Belgium, 1991.

[25] R. D. Williams, L. R. Ahuja and J. W. Naney, "Comparison of methods to estimate soil water characteristics from soil texture and limited data, *Soil Sci.*, vol. 153, pp. 172-184, 1992.

[26] W. J. Rawls, T. J. Gish and D. L. Brakensiek, "Estimating soil water retention from soil physical properties and characteristics", *Advances in Soil Science*, 16, pp. 213-234, 1991.

[27] L. M. Arya and J. F. Paris, "A physio-empirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data", *Soil Sci. Soc. Am. J.*, vol. 45, pp. 1218-1227, 1981.

[28] R. Haverkamp and J. Y. Parlange, "Predicting the water-retention curve from particle-size distribution: 1. Sandy soils without organic matter", *Soil Sci.*, vol. 142, pp. 325-339, 1986.

[29] K. Kobayashi and M. U. Salam, "Comparing simulated and measured values using mean squared deviation and its components", *Agron. J.*, vol. 92, pp. 345-352, 2000.

[30] J. S. Kern, "Evaluation of soil water retention models based on basic soil physical properties", *Soil Sci. Soc. Am. J.*, vol. 59, pp. 1134–1141, 1995.

[31] W. M. Cornelis, J. Ronsyn, M. Van Meirvenne, and R. Hartmann, "Evaluation of Pedotransfer Functions for Predicting the Soil Moisture Retention Curve", *Soil Sci. Soc. Am. J.*, vol. 65, pp. 638-648, 2001.

[32] Ya.A. Pachepsky and W. J. Rawls, "Accuracy and Reliability of Pedotransfer Functions as Affected by Grouping Soils", *Soil Sci. Soc. Am. J.*, vol. 63, pp. 1748-1757, 1999.

[33] C. D. Børgesen and M. G. Schaap, "Point and parameter pedotransfer functions for water retention predictions for Danish soils", *Geoderma*, vol. 127, pp. 154-167, 2005.

[34] W. J. Rawls, Y. A. Pachepskyb, J. C. Ritchie, T. M. Sobecki and H. Bloodworth, "Effect of soil organic carbon on soil water retention", *Geoderma*, vol. 116, pp. 61-76, 2003.

[35] B. Ambroise, D. Reutenauer and D. Viville, "Estimating soil water retention properties from mineral and organic fractions of coarse-textured soils in the Vosges mountains of France", in *Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, M. Th. Van Genuchten, F. J. Leij and L. J. Lund, ed., University of California, Riverside, California, 1992, pp. 453–462.

[36] A. M. Bell and H. Van Keulen, "Soil pedotransfer functions for four Mexican soils", *Soil Sci. Soc. Am. J.*, vol. 59, pp. 865–871, 1995.