



# Evaluation of point and parametric pedotransfer functions for predicting soil water retention and availability in soils of southwestern Saudi Arabia

Hesham M. Ibrahim<sup>1,2</sup> · Ali M. Al-Turki<sup>1</sup>

Received: 12 May 2017 / Accepted: 3 July 2018 / Published online: 12 July 2018  
© Saudi Society for Geosciences 2018

## Abstract

Simulation of water flow and solute transport in the soil environment requires accurate estimates of the soil water retention characteristics (SWRC). In this study, multiple linear regression (MLR) was used to develop two site-specific pedotransfer functions (PTFs), a point (MLRP) and a parametric (MLRF), using soil properties of 43 soil samples collected from the Jazan region, southwest of Saudi Arabia. The accuracy of the developed PTFs and four existing PTFs to determine the SWRC, predict soil water content (SWC) at  $-10$ ,  $-33$ , and  $-1500$  kPa, and estimate available water content (AWC) was assessed. The MLRP and the Schaap PTFs produced the best estimate of SWC, with smaller root mean square error (RMSE) ( $0.023$ – $0.053$   $\text{cm}^3 \text{cm}^{-3}$ ), and larger D-index ( $0.8$ – $0.9$ ) and Akaike information criterion (AIC) ( $-303.7$  to  $-240.7$ ), respectively. The largest prediction errors in the estimation of SWC were observed at matric potential close to field capacity (FC). For the AWC, the Schaap PTF provided the best prediction (RMSE =  $0.014$   $\text{cm}^3 \text{cm}^{-3}$ , D-index =  $0.93$ , AIC =  $-359.9$ ), followed by the MLRP PTF (RMSE =  $0.027$   $\text{cm}^3 \text{cm}^{-3}$ , D-index =  $0.83$ , AIC =  $-302.1$ ), whereas the Vereecken, Gupta and Larson, and the MLRF PTFs produced less accurate predictions of the AWC. The MLRP PTF proved to be more accurate compared to other tested PTFs in the prediction of SWC at both FC and permanent wilting point (PWP). In contrast, the MLRF PTF produced a relatively large error in the estimation of SWC at FC.

**Keywords** Hydraulic parameters · Pedotransfer functions · Soil water retention · Available water content

## Introduction

Accurate measurement or estimation of soil hydraulic parameters is needed to determine the availability of soil water to plants, and to model water flow and solute transport in the vadose zone. The most important soil hydraulic properties are hydraulic conductivity and soil water retention characteristics (SWRC). Unfortunately, direct measurement of these properties is time consuming and expensive, especially for research studies conducted at regional scales (Vereecken 1995). Despite the availability of many

empirical relations (e.g., Brooks and Corey 1964; Campbell 1974; van Genuchten 1980; Russo 1988) to model the SWRC, accurate estimate of the parameters required for these models is cumbersome and time consuming. During the last decades, several efforts have been made to develop low-cost, accurate, and quick methods to estimate soil hydraulic parameters from easily measured and available soil data (Nielsen et al. 1986). Mathematical models, called pedotransfer functions (PTFs), are employed to relate soil water retention properties with easily measured soil properties such as sand, silt, and clay contents, organic matter, and bulk density (Bouma 1989). PTFs can be categorized into two main groups: point and parametric PTFs (Tietje and Tapkenhinrichs 1993). Point PTFs predict soil water retention properties at a specific water potential (e.g.,  $-10$ ,  $-33$ , and  $-1500$  kPa). Parametric PTFs predict the parameters of a soil water retention model such as those of van Genuchten (1980). The two most common methods used to develop PTFs are multiple linear regression (MLR) and artificial neural networks (ANN). MLR

✉ Hesham M. Ibrahim  
habdou@ksu.edu.sa

<sup>1</sup> Department of Soil Science, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia

<sup>2</sup> Department of Soils and Water, Faculty of Agriculture, Suez Canal University, Ismailia 41522, Egypt

methods (Gupta and Larson 1979; Rawls et al. 1982; Lin et al. 1999; Mayr and Jarvis 1999; Tomasella et al. 2000) determine the most fundamental input data related to the soil water retention using stepwise regression techniques. On the other hand, ANN methods (Vereecken et al. 1989; Tamari et al. 1996; Schaap and Leij 1998; Schaap et al. 1998, 2001; Haghverdi et al. 2014) use an iterative calibration procedure to obtain the optimal and possibly nonlinear relations between input data (soil properties) and output data (hydraulic parameters), without the need for a priori regression model (Schaap and Leij 1998).

The validity of the developed PTFs is largely dependent on the data set that was used to generate the functions (Wosten et al. 1999), which in turn represents the geographic origin and the basic soil properties in the sampling area. Most of the available PTFs were developed for temperate regions, and consequently are expected to provide poor estimation of the hydraulic soil parameters when used in regions of different edaphoclimatic characteristics (Hodnett and Tomasella 2002). In this regard, Tietje and Tapkenhinrichs (1993) tested the accuracy of 13 PTFs in estimating the water retention parameters of a wide range of soils and found large variability in the performance of the tested PTFs, which ranged from good to poor, and the lowest accuracy was observed with the estimation of the residual water content ( $\theta_r$ ) and the shape parameter ( $\alpha$ ) of the van Genuchten equation. Cornelis et al. (2001) compared nine PTFs to estimate the SWRC using a data set of 298 undistributed soil samples collected from 69 different soils covering a wide range of textures. They found that most PTFs provided better prediction accuracy for the water content near saturation and at the dry end of the SWRC, especially when the PTFs were developed and calibrated based on data sets with average soil properties similar to those of the collected soil samples. They concluded that caution should be considered when PTFs are used in a geographical context different from the one where the data set used to develop the PTFs was collected. Khodaverdiloo et al. (2011) estimated the SWRC of calcareous soils using a number of PTFs that were developed with and without  $\text{CaCO}_3$  as one of the predictors of the PTFs. They reported no significant differences in the accuracy of the estimation of the SWRC when  $\text{CaCO}_3$  was included or eliminated from the developed PTFs. The accuracy of the developed PTFs was compared with the Rosetta PTFs of Schaap et al. (2001) and was found to provide more accurate estimate of the SWRC as indicated by a lower average RMSE (Khodaverdiloo et al. 2011). Abbasi et al. (2011) evaluated the accuracy of 13 point and parametric PTFs in estimation of SWRC and total available water in saline and saline-alkali soils, and found that the PTFs of Saxton et al. (1986) and Campbell (1985) provided the best estimates of SWRC and total available water, respectively.

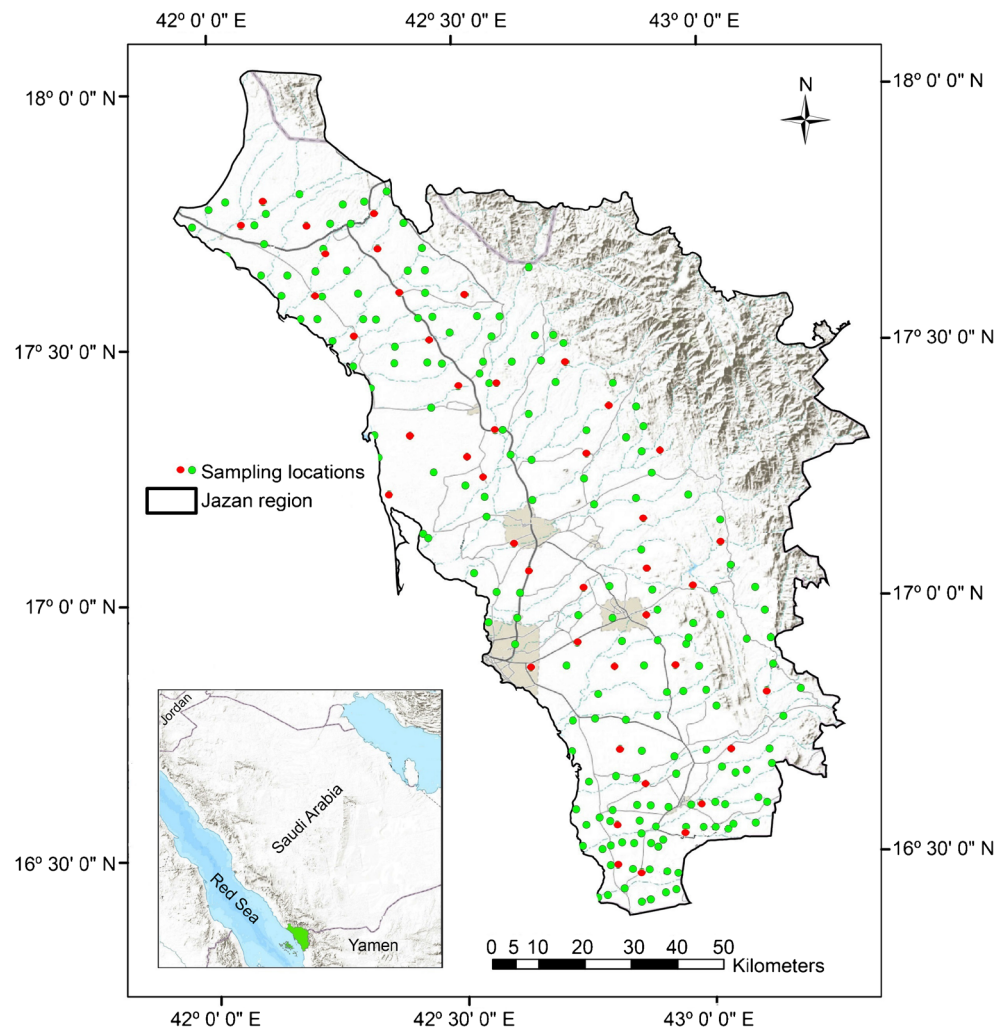
Most of the studies that evaluate PTFs focus on the assessment of the estimation error in comparison with some measured values of the SWRC at specific water potentials and limited soil locations. However, the main sources of the estimation error remain unclear as it can be attributed to the differences between data sets used to derive the PTFs, or differences in the algorithms and the predictors used to develop the PTFs (Nemes et al. 2003). The determination of the SWRC is not a goal itself, rather it is needed as a basic input data for the assessment of plant available water, and simulation models in agricultural, hydrological, and environmental research (Minasny and Hartemink 2011). Therefore, any errors generated during the estimation of the parameters of the SWRC will propagate when implementing the values into simulation models, leading to lower accuracy of the output of the simulation process (Botula et al. 2012; Wang et al. 2015). Regardless of the size and variability of the data set used to develop and calibrate the PTFs, it is important to evaluate the applicability and accuracy of the PTFs when applied outside the range of data used to develop them. For accurate evaluation of the PTF performance, it is preferable to use a small data set of relevant (local) soil properties rather than a large data set of general soil properties (Nemes et al. 2002). Therefore, the objective of this study was to develop two site-specific PTFs for the Jazan region southwest of Saudi Arabia, and to compare the performance of the developed PTFs with that of other existing PTFs. The specific objectives were to (a) compare the performance of two developed site-specific PTFs with that of four existing PTFs to determine the SWRC, and (b) evaluate the applicability of using point and parametric PTFs to predict soil water content at field capacity (FC) ( $-10$  and  $-33$  kPa) and permanent wilting point (PWP) ( $-1500$  kPa), and estimate available water content (AWC) in soils of Jazan region, southwestern Saudi Arabia.

## Materials and methods

### Soil sampling and analysis

The study was carried out in Jazan region (latitude  $16^\circ 54' 34.8''$  N and longitude  $42^\circ 33' 5.4''$  E) located in southwestern Saudi Arabia (Fig. 1). The total area of the region is  $8764 \text{ km}^2$ . The region is characterized by hot dry climate with average temperatures of  $40$  and  $21$  °C during summer and winter time, respectively. The average annual rainfall is  $232$  mm, mostly during Nov–Jan and some seasonal rainfall during April–May. The data set utilized in this research was collected from a previous study (Al-Turki et al. 2008), and consisted of 219 surface (0–30 cm)

**Fig. 1** Location map of the Jazan region and the sampling locations. Both red and green dots represent the large dataset, and the red dots only represent the small dataset



soil samples that were collected from the region (Fig. 1). Particle size distribution (PSD) (sand (Sa), 2–0.05 mm; silt (Si), 0.05–0.002 mm; clay (C), < 0.002 mm) and bulk density (BD) were measured on the soil samples according to the principles of Dane and Topp (2002). Total calcium carbonate (CaCO<sub>3</sub>) was determined according to the method of Loeppert and Suarez (1996). Saturation percent (SP) was measured gravimetrically in saturated soil paste, and the electrical conductivity (EC) was determined in the saturated soil paste extract (Rhoades 1996). Organic carbon (OC) was determined according to Nelson and Sommers (1982), and a conversion factor of 1.724 was used to estimate the percentage of organic matter (OM) in the soil samples.

A small data set (43 soil samples) representative of the large data set was used to determine the water retention data at matric potentials of –10, –33, –61, –102, –306, –510, –714, –1020, and –1500 kPa using pressure plate extractor (Soilmoisture Equipment, Santa Barbara, CA). The SWRC

was determined by fitting the retention data to the equation of van Genuchten (1980):

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

Where  $\theta(h)$  is the volumetric water content (cm<sup>3</sup> cm<sup>-3</sup>) at a specified matric potential,  $h$ (cm);  $\theta_r$  and  $\theta_s$  are the residual and saturated water contents (cm<sup>3</sup> cm<sup>-3</sup>), respectively;  $\alpha$  (cm<sup>-1</sup>) and  $n$  (–) are shape parameters of the SWRC. The nonlinear least-squares fitting procedure in the RETC program (van Genuchten et al. 1991) was used to estimate the unknown parameters ( $\theta_r$ ,  $\theta_s$ ,  $\alpha$ , and  $n$ ) from measured soil water retention data. For each soil sample, the water content at FC was calculated as the water content at a matric potential of –10 and –33 kPa for light (sand; loamy sand) and medium (sandy loam; sandy clay loam; loam; silt loam; clay loam) soil textures,

respectively. The PWP was calculated as the water content at a matric potential of  $-1500$  kPa for all soil samples. The AWC was calculated as the difference between FC and PWP.

## Pedotransfer functions

### Site-specific PTFs

The MLR method was used to develop two site-specific PTFs based on measured basic soil properties (Sa; Si; C; BD; SP; EC;  $\text{CaCO}_3$ ; OM). Logarithmic and reciprocal values of these soil properties were also included in the MLR, and the number of variables was reduced using a backward stepwise method to ensure that the final equation includes the most significant variables that explain a large portion of the parameter variability. Selection and elimination criteria were set to be  $P < 0.05$ , and  $P > 0.1$ , respectively. The first site-specific PTF that was developed is a point PTF (MLRP) that was used to estimate the soil water content at  $-10$ ,  $-33$ , and  $-1500$  kPa. The second site-specific parametric PTF (MLRF) was developed to estimate the hydraulic parameters ( $\theta_r$ ,  $\theta_s$ ,  $\alpha$ , and  $n$ ) of the van Genuchten equation using the parameters that were obtained from the fitting of the measured soil water retention data.

### Selected point and parametric PTFs

Four well-documented PTFs were also included in this study (Gupta and Larson 1979; Rawls et al. 1982; Vereecken et al. 1989; and Schaap et al. 2001), referred to hereafter as Gupta and Larson, Rawls, Vereecken, and Schaap PTFs. The first three PTFs use MLR to link water content at specific matric potential (Gupta and Larson; Rawls) to basic soil properties, or to estimate the hydraulic parameters (Vereecken) of the SWRC. The Schaap PTF uses ANN to estimate the hydraulic parameters of the van Genuchten equation using five hierarchical approaches based on available soil properties. The five hierarchical approaches range from texture class; sand, silt, and clay percentages; bulk density; WC at  $-10$  kPa; and WC at  $-1500$  kPa. The Schaap PTF is implemented in a computer program (Rosetta) that combines ANN analyses with the bootstrap method to assess the uncertainty in the predicted estimates of the hydraulic parameters. Out of the five hierarchical approaches available in the Rosetta program, three were used in this study: H1 (texture class); H2 (Sa, Si, C); H3 (Sa, Si, C, BD,  $\theta_{-10\text{kPa}}$ ,  $\theta_{-1500\text{kPa}}$ ). Model inputs of basic soil properties, along with the number and origin of soils used in the development of the selected PTFs, are listed in Table 1.

## Evaluation criteria

Several statistical indices were used to evaluate the performance of the selected and developed PTFs, which include the root mean square error (RMSE), the geometric mean error (GMER), the mean relative error (MRE), the D-index, and the Akaike information criterion (AIC) (Akaike 1974; Loague and Green 1991) defined as:

$$\text{RMSE} = \left\{ \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right\}^{1/2} \quad (2)$$

$$\text{GMER} = \exp \left\{ \frac{1}{n} \sum_{i=1}^n \ln \left( \frac{\hat{y}_i}{y_i} \right) \right\} \quad (3)$$

$$\text{MRE} = \left\{ \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \right\} \times 100 \quad (4)$$

$$\text{D-index} = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n \left\{ |\hat{y}_i - \bar{y}_i| + |y_i - \bar{y}_i| \right\}^2} \quad (5)$$

$$\text{AIC} = n \ln \left\{ \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right\} + 2k \quad (6)$$

where  $y_i$  and  $\hat{y}_i$  are the measured and predicted values, respectively;  $\bar{y}_i$  is the average measured value;  $n$  is the total number of observations; and  $k$  is the number of parameters in the PTF. The accuracy of the estimates of soil water content at the different matric potentials ( $-10$ ,  $-33$ , and  $-1500$  kPa) also was assessed using the Nash-Sutcliffe coefficient of efficiency (NSCE) described as (Nash and Sutcliffe 1970):

$$\text{NSCE} = \sum_{i=1}^n \left( \frac{y_i - \bar{y}_i}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \right)^2 \quad (7)$$

The RMSE is a measure of the accuracy of the PTFs, and the estimation improves when the RMSE value is close to 0. Both the GMER and the MRE are a measure of the bias in predicted values with underestimation for less than 1 and negative, or overestimation for more than 1 and positive values of the GMER and MRE, respectively. D-index equals to 1 indicates perfect matching between measured and predicted values. The AIC is a useful statistic measure when comparing the performance of different models as it takes into consideration the number of adjustable parameters in each model. The model with the lower AIC value is the best model. The PTFs were ranked based on the value of each statistical index. All indices were given equal weight, and a final ranking was calculated as the mean of the rankings of all indices. For the NSCE coefficient, a



**Table 1** List of the selected pedotransfer functions (PTFs) and their inputs

Reference	PTF	No. of samples	Origin of soils	Model inputs						
				Sa	Si	C	BD	OM	$\theta_{-10kPa}$	$\theta_{-1500kPa}$
Gupta and Larson (1979)	Point	43	USA	+	+	+	+	+		
Rawls et al. (1982)	Point	1323	USA	+		+		+		
Vereecken et al. (1989)	Parametric	182	Belgium	+		+	+	+		
Schaap et al. (2001)	Parametric	2134	USA and Europe	+	+	+	+		+	+

Sa, sand; Si, silt; C, clay; BD, bulk density; OM, organic matter;  $\theta$  soil water content at  $-10$  and  $-1500$  kPa

value of 1 indicates perfect model performance, while 0 or negative values indicate less accuracy (i.e., the mean of the measured data is representing the property better than the model predictions).

## Results and discussion

### Soil properties and measured water retention

Soil property ranges of the large data set are presented in Table 2. Results showed that the collected data set contained ample variations in the measured soil properties. For example, sand content varied from 9.6 to 91.1%, silt from 3.9 to 59.3%, clay from 5.1 to 31.2%, OM from 0.25 to 0.95%, and BD from 1.31 to 1.69 g cm<sup>-3</sup> (Table 2). The existence of large variability in the soil properties of a data set collected

for the development of PTFs is favorable to ensure better accuracy in the estimates of the developed PTFs (Pachepsky and Rawls 1999). Table 2 shows that all texture classes that were found in the large data set were also represented in the small data set. The only exception was the SiCL texture class. All basic soil properties of the small data set, which was used to obtain water retention data, fit well within the ranges of the soil properties of the large data set. Table 3 shows the water retention parameters and soil water content at  $-10$ ,  $-33$ , and  $-1500$  kPa per soil textural class for the small data set of the Jazan region. The van Genuchten equation was used to fit the measured water retention data.  $R^2$  values of the fitting were in the range 0.95–0.99, indicating that the van Genuchten equation adequately described the water retention characteristics of the different soils in the small data set. Mean values of  $\theta_r$  and  $\theta_s$  were larger in medium (SaL, SaCL, L, SiL, and CL) as compared to light (Sa

**Table 2** Basic soil properties of the large data set and the small data set used to generate the soil water retention data

Texture class	No. of samples	Sand (%)	Silt	Clay	SP	CaCO <sub>3</sub>	OM	EC (dS m <sup>-1</sup> )	BD (g cm <sup>-3</sup> )
Large dataset									
Sa	30	91.1 ± 2.4	3.9 ± 1.7	5.1 ± 1.9	29.2 ± 4.9	3.53 ± 3.3	0.33 ± 0.21	3.3 ± 2.8	1.69 ± 0.01
LSa	43	83.3 ± 3.6	9.8 ± 2.4	6.9 ± 2.3	28.9 ± 4.4	1.76 ± 2.7	0.28 ± 0.14	7.5 ± 3.2	1.65 ± 0.02
SaL	52	65.5 ± 8.2	23.1 ± 8.9	11.4 ± 3.7	29.6 ± 3.9	3.78 ± 5.7	0.44 ± 0.21	23.3 ± 5.9	1.57 ± 0.03
SaCL	5	50.3 ± 9.9	19.5 ± 8.1	30.2 ± 11.7	31.9 ± 8.8	2.65 ± 2.9	0.25 ± 0.15	4.5 ± 2.1	1.50 ± 0.04
L	42	41.1 ± 7.1	41.8 ± 5.3	17.1 ± 5.2	34.1 ± 5.3	2.72 ± 3.5	0.50 ± 0.35	13.6 ± 2.8	1.46 ± 0.03
SiL	34	27.2 ± 9.1	58.5 ± 5.9	14.4 ± 5.5	38.4 ± 5.2	7.59 ± 4.8	0.65 ± 0.51	9.5 ± 3.9	1.39 ± 0.04
SiCL	5	9.6 ± 6.1	59.3 ± 6.4	31.1 ± 4.8	39.9 ± 6.2	5.99 ± 4.5	0.95 ± 0.73	7.11 ± 2.3	1.31 ± 0.03
CL	8	32.2 ± 6.6	37.7 ± 5.7	31.2 ± 3.5	38.3 ± 6.6	4.51 ± 10.3	0.48 ± 0.35	11.8 ± 2.8	1.42 ± 0.03
Small dataset									
Sa	7	92.3 ± 2.5	2.7 ± 1.9	5.1 ± 2.7	30.1 ± 4.8	6.12 ± 1.4	0.46 ± 0.18	2.5 ± 2.1	1.69 ± 0.01
LSa	2	77.4 ± 0.8	15.6 ± 0.9	7.1 ± 0.2	31.6 ± 4.1	6.42 ± 2.2	0.25 ± 0.21	1.1 ± 0.2	1.62 ± 0.01
SaL	14	64.8 ± 7.1	24.3 ± 6.8	10.9 ± 2.9	30.2 ± 3.3	6.62 ± 3.9	0.38 ± 0.49	21.2 ± 4.5	1.57 ± 0.03
SaCL	1	34.3	15.1	50.6	34.8	8.8	0.28	8.1	1.43
L	6	40.5 ± 5.4	40.1 ± 5.7	19.4 ± 5.8	36.2 ± 4.6	4.57 ± 3.8	0.39 ± 0.32	4.9 ± 4.2	1.46 ± 0.02
SiL	9	25.8 ± 5.5	59.7 ± 6.1	11.4 ± 3.1	37.4 ± 3.5	4.36 ± 3.4	0.48 ± 0.32	1.7 ± 0.3	1.39 ± 0.02
CL	4	31.3 ± 7.3	38.1 ± 6.2	30.6 ± 2.2	41.1 ± 3.4	8.84 ± 10.1	0.58 ± 0.37	14.1 ± 2.4	1.41 ± 0.03

SP, saturation percent; CaCO<sub>3</sub>, calcium carbonate; OM, organic matter; EC, electrical conductivity; BD, bulk density; Sa, sand; L, loam; Si, silt; C, clay. Values presented are means followed by standard deviation (±1SD)

**Table 3** Water retention parameters of the van Genuchten equation and measured water content at  $-10$ ,  $-33$ , and  $-1500$  kPa

Texture class*	van Genuchten parameters**				Water content ( $\text{cm}^3 \text{cm}^{-3}$ ) at		
	$\theta_r$ ( $\text{cm}^3 \text{cm}^{-3}$ )	$\theta_s$ ( $\text{cm}^3 \text{cm}^{-3}$ )	$\alpha$ ( $\text{cm}^{-1}$ )	$n$ (-)	$-10$ kPa	$-33$ kPa	$-1500$ kPa
Sa	$0.022 \pm 0.01\#$	$0.325 \pm 0.02$	$0.044 \pm 0.012$	$1.518 \pm 0.14$	$0.098 \pm 0.03$	$0.066 \pm 0.03$	$0.036 \pm 0.02$
LSa	$0.020 \pm 0.01$	$0.373 \pm 0.01$	$0.041 \pm 0.004$	$1.375 \pm 0.01$	$0.130 \pm 0.01$	$0.082 \pm 0.01$	$0.033 \pm 0.01$
SaL	$0.027 \pm 0.01$	$0.394 \pm 0.01$	$0.040 \pm 0.011$	$1.467 \pm 0.10$	$0.171 \pm 0.02$	$0.119 \pm 0.03$	$0.048 \pm 0.03$
SaCL	0.028	0.410	0.009	1.481	0.228	0.187	0.101
L	$0.037 \pm 0.02$	$0.387 \pm 0.02$	$0.029 \pm 0.013$	$1.383 \pm 0.03$	$0.221 \pm 0.03$	$0.184 \pm 0.02$	$0.079 \pm 0.03$
SiL	$0.031 \pm 0.01$	$0.369 \pm 0.02$	$0.027 \pm 0.014$	$1.403 \pm 0.04$	$0.232 \pm 0.03$	$0.184 \pm 0.04$	$0.074 \pm 0.02$
CL	$0.053 \pm 0.02$	$0.406 \pm 0.01$	$0.035 \pm 0.006$	$1.341 \pm 0.09$	$0.233 \pm 0.02$	$0.189 \pm 0.02$	$0.102 \pm 0.04$

\*Texture class: *Sa*, sand; *L*, loam; *Si*, silt; *C*, clay

\*\*Estimated by fitting the measured soil water retention data to the van Genuchten equation using the RETC program

#Values are means of all soil samples in the representative texture class followed by standard deviation ( $\pm 1\text{SD}$ )

and LSa) textured soils. The strong retention of the medium-textured soils is related to its larger total porosity, which increases the ability of the soil to hold water. Values of  $\alpha$  increased with the increase in the sand percentage, or in the presence of well-structured soils. Maximum  $\alpha$  values reached  $0.044$ ,  $0.041$ , and  $0.035 \text{ cm}^{-1}$  for the Sa, LSa, and CL soils, respectively. Larger  $\alpha$  values indicate the occurrence of sudden change in the water content as pores are emptying faster in response to small changes in the matric potential (Hodnett and Tomasella 2002). Values of  $n$  were highly dependent on the texture of the soil, and showed a gradual increase as the texture changed from CL to Sa as a result of the increase in the size of the soil particles.

### The development of site-specific PTFs

The number of basic soil properties that was retained in the final MLR equations varied between 2 and 9 out of the 24 predictors used in the MLR analyses. The accuracy of the generated site-specific PTFs was assessed by the value of  $R^2$ , which ranged between  $0.61$ – $0.71$  and  $0.37$ – $0.67$  for the MLRP and MLRF, respectively (Table 4). For the MLRP, the equation for the water content at  $-10$  kPa ( $\theta_{-10\text{kPa}}$ ) contained only *SP* and *BD* as predictors, and explained 68% of the variance in the water content. The equations for the water content at  $-33$  kPa ( $\theta_{-33\text{kPa}}$ ) and  $-1500$  kPa ( $\theta_{-1500\text{kPa}}$ ) contained 3 (*SP*, *Si*, *C*) and 4 (*Si*, *OM*, *CaCO<sub>3</sub>*, *BD*) soil properties as predictors, and explained 71 and 61% of the variance in the water content, respectively. Similar findings were reported by Obalum and Obi (2013). For the MLRF, more basic soil properties were retained in the regression equations (Table 4). However, the percentage of the variance explained was, in general, smaller than that for the MLRP, especially for the equation to determine the  $n$  parameter, which explained only 37% of the variance. Pearson correlations analysis showed consistent results

and indicated that water content at the three matric potentials was always negatively correlated with *Sa*, and positively correlated with *Si* and *C* contents (Table 5). In addition, *BD* and *SP* showed consistent negative and positive correlations, respectively, with the water content at the three matric potentials. The maximum negative ( $-0.81$ ) and positive ( $0.62$ ) correlations with the *BD* and *SP*, respectively, were obtained with the water content at  $-10$  kPa, indicating that the water content at this high potential is largely influenced by the soil structure. Correlations between basic soil properties and the hydraulic parameters of the van Genuchten equation revealed that the residual water content ( $\theta_r$ ) was positively correlated with *Si* ( $0.41$ ), *C* ( $0.61$ ), and *SP* ( $0.40$ ), and negatively correlated with *Sa* ( $-0.57$ ) and *BD* ( $-0.58$ ), respectively. The shape parameter  $\alpha$  was found to be positively correlated with *Sa* ( $0.52$ ) and *BD* ( $0.50$ ), and negatively correlated with *Si* ( $-0.49$ ) and *SP* ( $-0.34$ ). No significant correlations were found between the saturated water content ( $\theta_s$ ) or the shape parameter  $n$  and the measured soil properties (Table 5).

### Evaluation of the performance of the PTFs

The accuracy of the site-specific and selected PTFs to estimate soil water content at  $-10$ ,  $-33$ , and  $-1500$  kPa was evaluated using several statistical indices (Table 6). At  $-10$  kPa, the lowest RMSE between measured and predicted water contents was observed with the site-specific MLRP PTF ( $0.028 \text{ cm}^3 \text{ cm}^{-3}$ ). The second and third best PTFs were the Schaap (H3) and the site-specific MLRF PTFs, which showed intermediate RMSE ( $0.053$  and  $0.095 \text{ cm}^3 \text{ cm}^{-3}$ , respectively) as compared to the other evaluated PTFs. All PTFs overestimated the soil water content at  $-10$  kPa as indicated by the  $> 1.0$  and the positive GMER and MRE, respectively. The MLRP and the Schaap (H3) PTFs showed the minimum overestimation in soil water content ( $1.02$ ,

**Table 4** Regression parameters and coefficients of the site-specific point multiple linear regression (MLRP) and parametric multiple linear regression (MLRF) PTFs

PTF*	Regression parameters and coefficients**	R <sup>2#</sup>
MLRP	$\theta_{-10kPa} = 0.158 + 0.0637 \times \ln SP - 0.458 \times \ln BD$	0.68
	$\theta_{-33kPa} = -29.159 - 0.129 \times SP + 0.000868 \times Si + 8.429 \times \ln SP + 0.0342 \times \ln C + \frac{130.908}{SP}$	0.71
	$\theta_{-1500kPa} = 0.432 - 0.00177 \times Si - 0.00199 \times OM - 0.00644 \times \ln CaCO_3 + 0.0162 \times \ln OM - 0.734 \times \ln BD$	0.61
MLRF	$\theta_r = 0.293 - 0.000118 \times EC - 0.00152 \times Si - 0.0064 \times OM - 0.00549 \times \ln CaCO_3 + 0.0664 \times \ln OM - 0.624 \times \ln BD + \frac{0.0647}{OM}$	0.67
	$\theta_s = 7.536 + 0.548 \times C + 0.549 \times Si + 0.663 \times Sa + 0.0267 \times \ln C - 74.994 \times \ln BD - \frac{56.135}{BD}$	0.56
	$\alpha = -40.774 - 0.0527 \times SP + 0.383 \times C + 0.386 \times Si + 0.427 \times Sa - 8.7 \times BD + 3.694 \times \ln SP + 0.0778 \times \ln C + \frac{61.994}{SP} + \frac{0.0838}{C}$	0.55
	$n = -793.809 - 2.97 \times C - 2.963 \times Si - 3.448 \times Sa - 0.605 \times \ln SP + 31.077 \times \ln Sa + 0.0766 \times \ln OM + 703.468 \times \ln BD + \frac{375.533}{Sa} + \frac{1055.507}{BD}$	0.37

\*Multiple linear regression: point (MLRP) and parametric (MLRF) PTFs

\*\* $\theta_{-10kPa}$ ,  $\theta_{-33kPa}$ , and  $\theta_{-1500kPa}$  are the soil water contents at - 10, - 33, and - 1500 kPa, respectively;  $\theta_r$  and  $\theta_s$  are the residual and saturated water contents;  $\alpha$  and  $n$  are shape parameters of the soil water retention curve;  $Sa$ , sand;  $Si$ , silt;  $C$ , clay;  $SP$ , saturation percent;  $BD$ , bulk density;  $CaCO_3$ , calcium carbonate;  $EC$ , electrical conductivity;  $OM$ , organic matter

# Coefficient of determination

0.01, and 1.16, 2.42 for the GMER and the MRE, respectively). The values of the statistical indices D-index and AIC also showed that the MLRP and the Schaap (H3) were the best PTFs to estimate the soil water content at - 10 kPa (0.898, - 303.7, and 0.802, - 240.8, respectively). The final ranking listed in Table 6 showed that the MLRP PTF performed the best followed by the Schaap (H3) PTF. The performance of the Gupta and Larson and Rawls PTFs was close to that of the MLRF PTF with an average RMSE that ranged between 0.122 and 0.136 cm<sup>3</sup> cm<sup>-3</sup>, and average AIC that ranged between - 174.6 and - 161.4. The Vereecken PTF showed the worst performance as indicated by the largest RMSE and AIC (0.194 cm<sup>3</sup> cm<sup>-3</sup>, - 137.1, respectively). The Vereecken PTF ranked eighth among all the tested PTFs, this can be mainly attributed to the fact that the properties of the soils used in this study are not in the range of those that were used to develop this PTF. These

results are consistent with Liao et al. (2011) who reported that the Vereecken PTF produced the worst estimate (largest RMSE and absolute mean error) of the soil water retention of 107 soil samples characterized by mean clay content and organic matter of 17.1 and 1.4%, respectively. In addition, Mermoud and Xu (2006) evaluated the Vereecken et al. (1989) and three other PTFs and found that all of the evaluated PTFs failed to accurately estimate water retention in light-textured soils characterized by low clay and organic matter contents. Most of the soils in the Jazan region are characterized by low clay and organic matter contents (Table 2), which could explain the larger error in the estimation of soil water content by the Vereecken PTF. Inclusion of OM as input to PTFs greatly improves the estimates of soil water retention (Rawls et al. 2003). OM greatly affects the pore size distribution through the enhancement of the soil structure. Therefore, soils containing

**Table 5** Pearson correlations between basic soil properties and the hydraulic parameters of the van Genuchten equation and the measured water content at - 10, - 33, and - 1500 kPa

Soil properties#	$\theta_r$ (cm <sup>3</sup> cm <sup>-3</sup> )	$\theta_s$ (cm <sup>3</sup> cm <sup>-3</sup> )	$\alpha$ (cm <sup>-1</sup> )	$n$ (-)	$\theta_{-10kPa}$ (cm <sup>3</sup> cm <sup>-3</sup> )	$\theta_{-33kPa}$ (cm <sup>3</sup> cm <sup>-3</sup> )	$\theta_{-1500kPa}$ (cm <sup>3</sup> cm <sup>-3</sup> )
Sa	-0.57**	0.01	0.52**	0.01	-0.81**	-0.79**	-0.56**
Si	0.41**	-0.03	-0.49**	0.09	0.73**	0.69**	0.39**
C	0.61**	0.06	-0.26	-0.20	0.55**	0.60**	0.63**
BD	-0.58**	0.01	0.50**	0.01	-0.81**	-0.79**	-0.56**
SP	0.40**	-0.01	-0.34*	-0.06	0.62**	0.62**	0.52**
EC	-0.13	0.05	0.16	-0.02	-0.08	-0.12	-0.13
CaCO <sub>3</sub>	-0.18	0.08	-0.07	-0.06	0.01	-0.09	-0.15
OM	0.09	0.05	0.03	0.10	0.04	0.06	0.08

#  $Sa$ , sand;  $Si$ , silt;  $C$ , clay;  $BD$ , bulk density;  $SP$ , saturation percent;  $EC$ , electrical conductivity;  $CaCO_3$ , calcium carbonate;  $OM$ , organic matter

\* Significant at  $P < 0.05$

\*\* Significant at  $P < 0.01$

**Table 6** Statistical evaluation of the measured and predicted water contents ( $\text{cm}^3 \text{cm}^{-3}$ ) at  $-10$ ,  $-33$ , and  $-1500$  kPa

PTFs	SWC	RMSE	GMER	MRE	D-index	AIC	Final rank
Gupta and Larson	- 10 kPa	0.136 (7)	1.51 (6)	11.09 (7)	0.525 (6)	- 161.4 (7)	7
Rawls		0.122 (6)	1.55 (7)	10.79 (6)	0.520 (7)	- 174.6 (6)	6
MLRP		0.028 (1)	1.02 (1)	0.01 (1)	0.898 (1)	- 303.7 (1)	1
Vereecken		0.194 (8)	1.78 (8)	17.11 (8)	0.379 (8)	- 137.1 (8)	8
Schaap (H1)		0.103 (5)	1.24 (3)	6.62 (3)	0.685 (3)	- 193.4 (3)	3
Schaap (H2)		0.102 (4)	1.31 (4)	7.25 (5)	0.673 (4)	- 189.4 (4)	4
Schaap (H3)		0.053 (2)	1.16 (2)	2.42 (2)	0.802 (2)	- 240.8 (2)	2
MRLF		0.095 (3)	1.33 (5)	7.18 (4)	0.575 (5)	- 186.9 (5)	5
Gupta and Larson		- 33 kPa	0.126 (6)	1.85 (6)	11.05 (6)	0.567 (6)	- 167.8 (6)
Rawls	0.089 (5)		1.63 (5)	7.6 (5)	0.660 (5)	- 201.7 (5)	5
MLRP	0.033 (1)		0.96 (1)	-0.84 (2)	0.903 (1)	- 284.6 (1)	1
Vereecken	0.246 (8)		2.53 (8)	22.24 (8)	0.348 (8)	- 116.7 (8)	8
Schaap (H1)	0.073 (3)		1.31 (3)	4.33 (3)	0.780 (3)	- 222.7 (3)	3
Schaap (H2)	0.075 (4)		1.36 (4)	4.89 (4)	0.777 (4)	- 217.2 (4)	4
Schaap (H3)	0.039 (2)		1.15 (2)	0.65 (1)	0.893 (2)	- 268.0 (2)	2
MRLF	0.139 (7)		1.89 (7)	12.09 (7)	0.493 (7)	- 153.5 (7)	7
Gupta and Larson	- 1500 kPa		0.054 (6)	1.32 (6)	2.05 (6)	0.713 (4)	- 241.0 (6)
Rawls		0.058 (7)	1.89 (7)	4.44 (7)	0.617 (7)	- 240.7 (7)	7
MLRP		0.023 (1)	1.07 (2)	0.01 (1)	0.862 (2)	- 313.4 (1)	1
Vereecken		0.354 (8)	5.99 (8)	31.78 (8)	0.151 (8)	- 85.4 (8)	8
Schaap (H1)		0.042 (5)	1.25 (4)	0.92 (4)	0.625 (6)	- 271.1 (3)	4
Schaap (H2)		0.041 (4)	1.28 (5)	1.07 (5)	0.654 (5)	- 269.4 (4)	5
Schaap (H3)		0.026 (2)	0.99 (1)	- 0.91 (3)	0.875 (1)	- 300.5 (2)	2
MRLF		0.038 (3)	1.08 (3)	0.69 (2)	0.753 (3)	- 265.0 (5)	3

SWC, soil water content ( $\text{cm}^3 \text{cm}^{-3}$ ); RMSE, root mean square error ( $\text{cm}^3 \text{cm}^{-3}$ ); GMER, geometric mean error; MRE, mean relative error (%); AIC, Akaike information criterion. The number in parenthesis represents the ranking order of the PTFs

larger contents of OM tend to retain more water than soils with lesser OM at the same matric potential (Nemes et al. 2003). Most of the soils in the Jazan region are characterized by low organic matter. The mean value for the organic matter content ranged between 0.25 and 0.95% (Table 2). Nevertheless, the Gupta and Larson, Rawls, and Vereecken PTFs were developed based on soils that contain a much larger organic matter content that ranged between 0 and 23%, 0.1 and 12.5%, and 0.02 and 11.4%, respectively. This could explain the general trend of these PTFs to overestimate soil water content in the soils of the Jazan region.

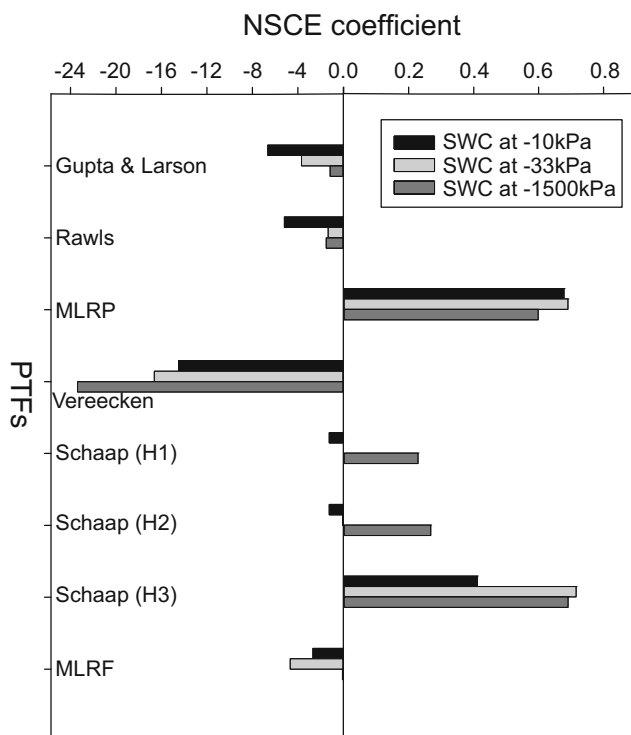
At  $-33$  and  $-1500$  kPa, the same trend was observed and the MLRP PTF provided the lowest RMSE (0.033 and  $0.023 \text{ cm}^3 \text{ cm}^{-3}$ , respectively) values, followed by the Schaap (H3) PTF ( $0.039$  and  $0.026 \text{ cm}^3 \text{ cm}^{-3}$ , respectively). The overall trend to overestimate the soil water content also continued, except with the MLRP PTF at  $-33$  kPa and the Schaap (H3) PTF at  $-1500$  kPa where slight underestimation was observed. Previous studies have reported the tendency of the Gupta and Larson, Rawls, and Vereecken PTFs to overestimate soil water content at both FC and PWP (Kern 1995; Nguyen et al. 2015). The shape of the SWRC in the saturated

range is highly dependent on soil structure, texture, and clay mineralogy, whereas in the dry range soil texture mostly defines the shape of the SWRC (Bruand 2004). The Gupta and Larson, Rawls, and Vereecken PTFs were developed for temperate soils characterized by medium-textured soils, high organic carbon, and well-aggregated structure. Therefore, it is expected that the application of these PTFs to predict water content in soils of dry regions will yield higher estimation errors, especially in the saturated range at matric potentials  $> -33$  kPa. The D-index (0.903, 0.862 and 0.893, 0.875), and the AIC ( $-284.6$ ,  $-313.4$ , and  $-268$ ,  $-300.5$ ) values also confirmed that the MLRP and the Schaap (H3) PTFs provided the best estimate of soil water content among all tested PTFs at  $-33$  and  $-1500$  kPa, respectively (Table 6). The final ranking of the accuracy of the evaluated PTFs to estimate soil water content at  $-33$  and  $-1500$  kPa was in agreement with the performance of the PTFs at  $-10$  kPa, and showed that the MLRP and the Schaap (H3) PTFs gave better predictions. The ranking of the five statistical indices was quite consistent for most of the evaluated PTFs, only slight modifications in the final ranking order of the Rawls and MRLF PTFs were observed (Table 6).



The site-specific point PTF (MLRP) provided more accurate estimation of the soil water content at  $-10$ ,  $-33$ , and  $-1500$  kPa than the parametric PTFs (MLRF). The MLRF PTF contained more input variables, which possibly could increase the estimation error due to the larger spatial variability of the data collected from a relatively large area. In addition, the relationship between the parameters of the van Genuchten equation and soil properties is not linear. Therefore, regression analysis could produce larger errors in the estimation of the van Genuchten parameters (Minasny et al. 1999). Both PTFs provided better estimate of the soil water content close to saturation ( $-10$  kPa) and at the dry end of the SWRC ( $-1500$  kPa), whereas the largest error was observed close to FC ( $-33$  kPa). The RMSE of the MLRP PTF reached 0.028, 0.033, and 0.023  $\text{cm}^3 \text{cm}^{-3}$  for soil water contents at  $-10$ ,  $-33$ , and  $-1500$  kPa, respectively (Table 6). The same trend was also observed with the MLRF PTF, which had the largest RMSE (0.139  $\text{cm}^3 \text{cm}^{-3}$ ) at FC, and much less RMSE at  $-10$  kPa (0.095  $\text{cm}^3 \text{cm}^{-3}$ ) and at  $-1500$  kPa (0.038  $\text{cm}^3 \text{cm}^{-3}$ ). Similar results were reported by Cornelis et al. (2001) who evaluated 9 PTFs and found that the largest errors in predicting SWRC was obtained at a matric potential between  $-10$  and  $-31$  kPa (i.e., close to FC).

The NSCE values of the accuracy of the different PTFs in estimating soil water content (Fig. 2) showed that only the MLRP and the Schaap (H3) PTFs gave positive values that ranged between 0.41 and 0.72 at all three matric potentials ( $-$



**Fig. 2** The Nash-Sutcliffe coefficient of efficiency (NSCE) of the soil water content at  $-10$ ,  $-33$ , and  $-1500$  kPa as estimated by the selected different PTFs

$10$ ,  $-33$ , and  $-1500$  kPa). All other PTFs had negative NSCE values, with the exception of the Schaap (H1) and (H2) PTFs at  $-1500$  kPa. Consistent with previous results, the Vereecken PTF showed the lowest accuracy in predicting the soil water content at all matric potentials, with NSCE values in the range of  $-14.4$  to  $-23.4$  (Fig. 2). Assuming the average soil water content, based on measurement at discrete locations, at any matric potential is known, this average will produced an NSCE value of 0 when used to represent the soil water content at any location other than the measurement locations. This indicates that only the MLRP and the Schaap (H3) PTFs can reliably be used to predict the soil water content at FC ( $-10$  and  $-33$  kPa) and at PWP ( $-1500$  kPa) in the soils of the Jazan region. Based on the assessment at all three matric potentials, the final ranking of the accuracy of the evaluated PTFs ordered from larger to smaller was MLRP > Schaap (H3) > Schaap (H1) > Schaap (H2) > MLRF > Rawls > Gupta and Larson > Vereecken PTFs.

### Prediction of AWC

The estimates of the soil water content that were generated by the different PTFs at FC ( $-10$  and  $-33$  kPa) and at PWP ( $-1500$  kPa) were used to predict the AWC in the surface layer of the soils of the Jazan region. Predicted and measured AWC were compared and the accuracy of the prediction was evaluated using the statistical performance criteria defined earlier. Concerning the overall accuracy of the evaluated PTFs in predicting the AWC, Table 7 presents the statistical indices used to compare between measured and predicted AWC. The best agreement between measured and predicted AWC was obtained with the Schaap (H3) PTF as indicated by the lower RMSE (0.014  $\text{cm}^3 \text{cm}^{-3}$ ), and maximum D-index and AIC values of 0.934 and  $-359.9$ , respectively. The MLRP PTF performance was the second best with RMSE of 0.027  $\text{cm}^3 \text{cm}^{-3}$ , and D-index and AIC values of 0.826 and  $-302.1$ , respectively (Table 7). The lowest accuracy in predicting the AWC was observed with the Vereecken, Gupta and Larson, and the MLRF PTFs, with larger RMSE (0.069–0.133  $\text{cm}^3 \text{cm}^{-3}$ ), AIC ( $-220.4$  to  $-163.8$ ), and smaller D-index (0.641–0.341) values. The prediction bias as measured by the GMER and the MRE values was mostly positive indicating an overestimation in the prediction of AWC, and ranged between 1.05–2.39 and 0.43–11.39, respectively. The only exception was with the MLRP and the Vereecken PTFs, which showed a slight (GMER = 0.91, MRE =  $-0.81$ ) and moderate (GMER = 0.61, MRE =  $-1.62$ ) underestimation in the prediction of AWC, respectively. Figure 3 presents the predicted vs. measured AWC as estimated by the different PTFs for all the 43 soil samples collected in the small data set of the Jazan soils. For the Schaap (H3) and the MLRP PTFs, most of the points are aligned along the 1:1 line, which confirms the good

**Table 7** Statistical evaluation of the measured and predicted available water contents ( $\text{cm}^3 \text{cm}^{-3}$ )

PTFs	RMSE	GMER	MRE	D-index	AIC	Final rank
Gupta and Larson	0.093 (7)	2.34 (7)	8.69 (7)	0.457 (7)	-194.5 (7)	7
Rawls	0.042 (3)	1.51 (6)	3.17 (6)	0.731 (3)	-262.6 (3)	4
MLRP	0.027 (2)	0.91 (3)	-0.81 (2)	0.826 (2)	-302.1 (2)	2
Vereecken	0.069 (6)	0.61 (5)	-1.62 (3)	0.641 (6)	-220.4 (6)	6
Schaap (H1)	0.048 (4)	1.11 (4)	2.14 (5)	0.648 (5)	-250.3 (4)	5
Schaap (H2)	0.049 (5)	1.08 (2)	2.11 (4)	0.661 (4)	-249.7 (5)	3
Schaap (H3)	0.014 (1)	1.05 (1)	0.43 (1)	0.934 (1)	-359.9 (1)	1
MRLF	0.133 (8)	2.39 (8)	11.39 (8)	0.341 (8)	-163.8 (8)	8

*RMSE*, root mean square error ( $\text{cm}^3 \text{cm}^{-3}$ ); *GMER*, geometric mean error; *MRE*, mean relative error (%); *AIC*, Akaike information criterion. The number in parenthesis represents the ranking order of the PTFs

agreement between measured and predicted AWC, and subsequently the reliability of the application of these PTFs to predict AWC in the investigated soils. For all other PTFs, the values are more scattered indicating that these PTFs failed to predict the AWC in the study area with acceptable accuracy. In addition, most of the points are located above the 1:1 line, which confirms the overestimation in the prediction of AWC. From Fig. 3, it can be clearly seen that the Gupta and Larson and the MLRF PTFs showed the maximum overestimation in the prediction of the AWC. This is also confirmed by the larger GMER and MRE of (2.34, 2.39) and (8.69, 11.39) that were obtained for the prediction of AWC with the Gupta and Larson and the MLRF PTFs, respectively.

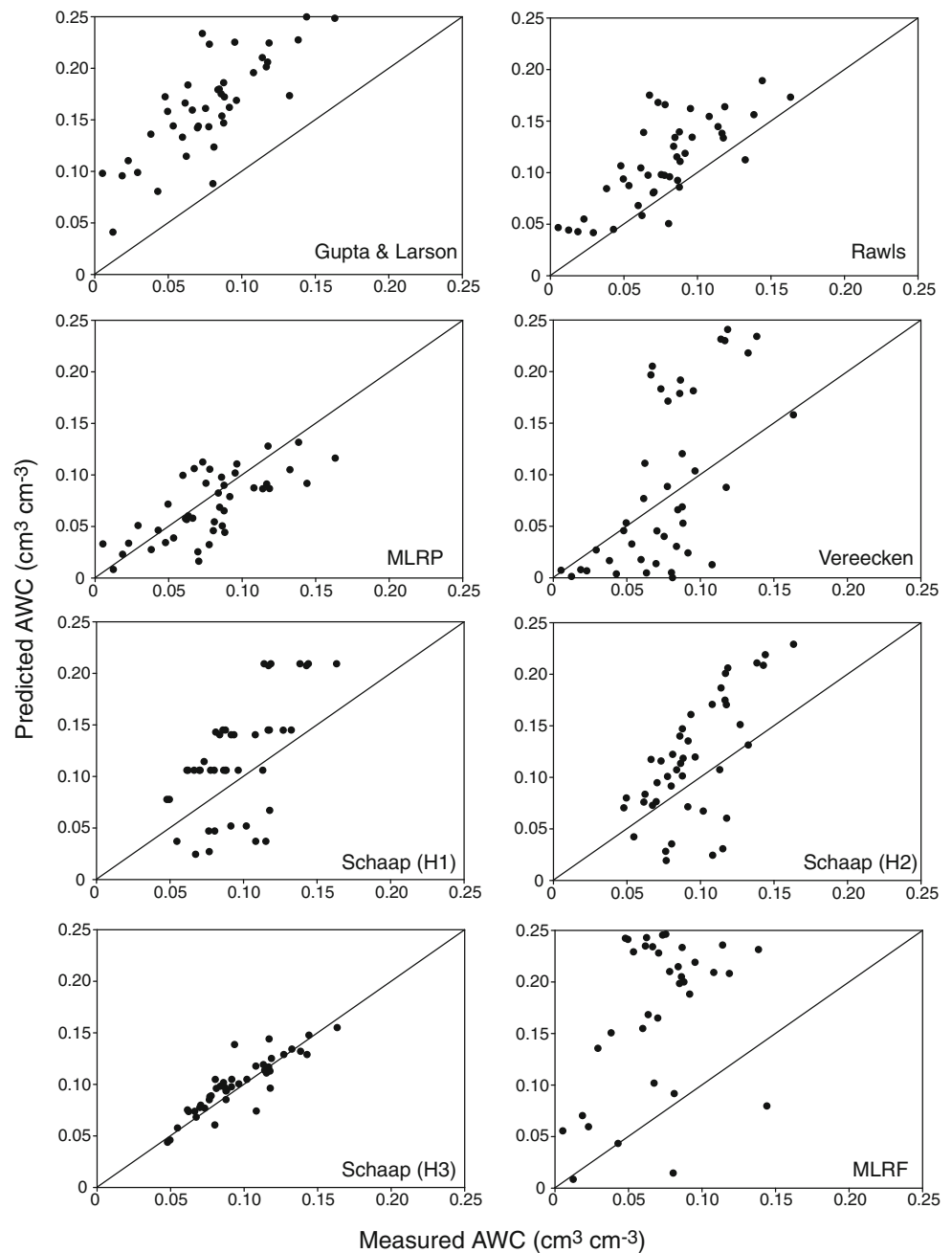
The Schaap (H3) PTF was used to predict the AWC for all 219 soils in the large data set of the Jazan region. Based on the coordinates of the sampling locations, the predicted AWC values were georeferenced to a base map of the Jazan region in ArcGIS 10.3. The geostatistical analyst function within ArcGIS was used to interpolate the AWC values, and to produce a general map showing the AWC distribution in the Jazan region (Fig. 4a). The absolute error in the prediction of AWC was also calculated and presented as a distribution map for the Jazan region (Fig. 4b). AWC in the soils of the Jazan region ranged between 0.043 and 0.178  $\text{cm}^3 \text{cm}^{-3}$ . Larger AWC values (0.125–0.178  $\text{cm}^3 \text{cm}^{-3}$ ) were observed in the mountain soils located in the eastern part of the Jazan region. These soils are characterized by medium-texture and larger amounts of silt, and therefore are expected to retain more water especially at FC. The lowest AWC values (0.043–0.08  $\text{cm}^3 \text{cm}^{-3}$ ) in the Jazan region were mostly observed in the western costal area, where soils are predominately sand in texture with low levels of organic carbon (Fig. 4a). Soils in approximately 90% of the areal extent of the Jazan region had an average absolute error in the prediction of AWC that ranged between 4.8 and 14.4%. The maximum calculated absolute error reached 33.6–38.3%, and was observed in soils characterized by larger amount of clay and high salinity levels (Fig. 4b).

## Summary and conclusions

In this study, MLR was used to develop two site-specific PTFs, a point (MLRP) and a parametric (MLRF), using basic soil properties of 219 soil samples that were collected from the Jazan region in the southwest of Saudi Arabia. We evaluated the accuracy of the two developed PTFs and four existing PTFs, two point (Gupta and Larson; Rawls) and two parametric (Vereecken; Schaap), to determine the SWRC, predict soil water content at -10, -33, and -1500 kPa, and to estimate AWC in soils of Jazan region. The accuracy of the generated site-specific PTFs was determined by the  $R^2$  value, and ranged between 0.61 and 0.68 and 0.37 and 0.67 for the MLRP and MLRF, respectively. Clay and silt contents had significant positive correlation with the soil water content at FC (-33 kPa) and PWP (-1500 kPa). On the other hand, sand content and *BD* had significant negative correlation with the soil water content at FC and PWP. No significant correlations were observed between *EC*, *CaCO<sub>3</sub>*, and *OM* and the soil water content at FC and PWP.

Correlations between basic soil properties and the hydraulic parameters of the van Genuchten equation revealed that the residual water content ( $\theta_r$ ) was positively correlated with *Si* (0.41), *C* (0.61), and *SP* (0.40), and negatively correlated with *Sa* (-0.57) and *BD* (-0.58), respectively. The shape parameter  $\alpha$  was found to be positively correlated with *Sa* (0.52) and *BD* (0.50), and negatively correlated with *Si* (-0.49) and *SP* (-0.34). No significant correlations were found between the saturated water content ( $\theta_s$ ) or the shape parameter *n* and the measured soil properties. The Schaap (H3) and the site-specific MLRF PTFs were able to accurately fit the measurement data to the van Genuchten equation, and provided adequate estimate of the SWRC over the entire range from high (saturation) to low (dry) water potentials. The Schaap (H1) and (H2) PTFs failed to estimate water retention at potentials higher than -60 kPa, whereas the Vereecken PTF failed to accurately estimate the SWRC and provided an overestimation of water content at all water potentials.

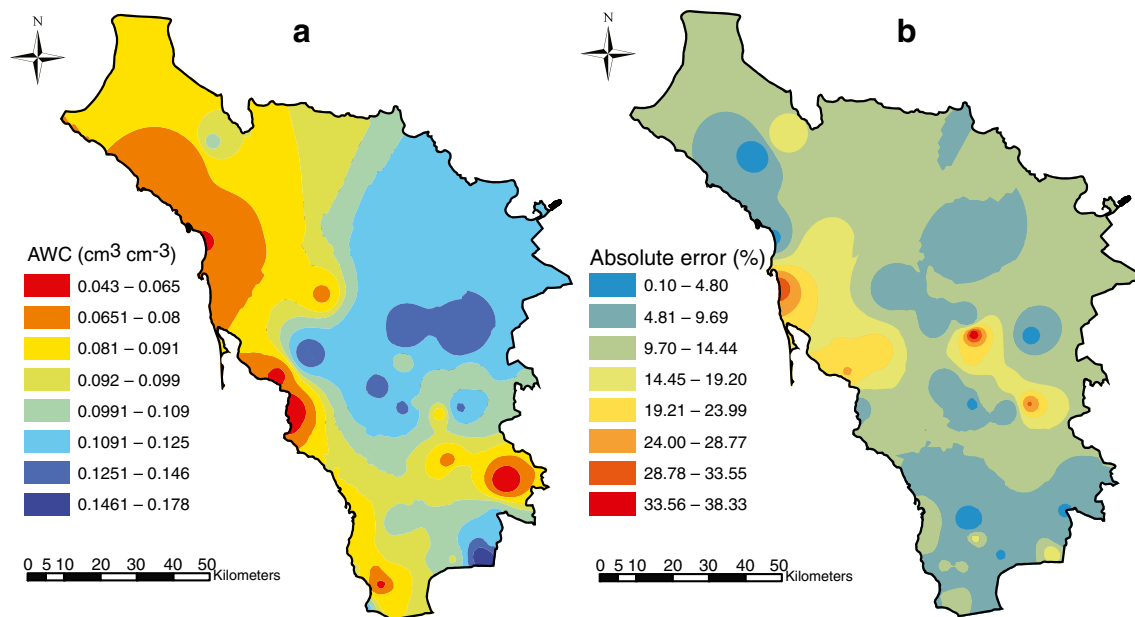
**Fig. 3** Measured vs. predicted available water content ( $\text{cm}^3 \text{cm}^{-3}$ ) as estimated by the different PTFs



The MLRP and the Schaap (H3) PTFs produced the best estimate of soil water content in the Jazan region, with smaller RMSE and larger D-index and AIC values. The performance of Schaap (H1), (H2), and MLRF PTFs was intermediate, whereas the Rawls, Gupta and Larson, and Vereecken PTFs had the largest RMSE and always showed an overestimation in the estimates of soil water content at all matric potentials. In general, the largest prediction errors in the estimation of soil water content were observed at matric potential close to FC ( $-33$  kPa). The reason for the better performance of the Schaap (H3) PTF might be related to the fact that it was developed and calibrated based on a large database of soils from North America and Europe covering a wide range of soil variability.

Therefore, it is highly possible that most of the variability found in the soils of the Jazan region were represented to a certain degree in this database. In addition, the Schaap (H3) PTF is the only PTF, among all the evaluated PTFs, that takes into account water content at two water potentials ( $-33$  and  $-1500$  kPa) as inputs in the PTF, which significantly enhances the accuracy of the prediction. The previous assessment of the accuracy of the evaluated PTFs was confirmed by the predictions of the AWC, which revealed that the Schaap (H3) and the MLRP provided the best estimate of the AWC in the soils of the Jazan region.

The site-specific point PTF (MLRP) proved to be more efficient than all other tested PTFs in the prediction of soil



**Fig. 4** Predicted available water content ( $\text{cm}^3 \text{cm}^{-3}$ ) (a), and absolute error of the prediction (%) (b) in the Jazan region as predicted by the Schaap (H3) parametric PTF

water content in the Jazan region at both FC and PWP. In contrast, the parametric site-specific PTF (MLRF) showed acceptable performance at saturation and at the dry end of the SWRC, but failed to estimate water content at FC. Despite the relatively large error in the estimation of SWRC based on parametric PTFs, the predicted hydraulic parameters are usually accurate enough for the simulation of water flow and transport processes in the environment, especially if direct measurement of the hydraulic parameters is not available. The pedoclimatic context of the data sets used to derive the PTFs greatly influence the performance of the PTFs. Therefore, caution should be considered when PTFs developed in this study are being used in soils with basic soil properties outside the range of soil properties that were used to develop and validate the PTFs.

**Funding information** This research was funded by the National Plan for Science, Technology and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, Kingdom of Saudi Arabia, Award Number 12-AGR2575-02.

## References

- Abbasi Y, Alavijeh G, Liaghat AM, Shorafa M (2011) Evaluation of pedotransfer functions for estimating soil water retention curve of saline and saline-alkali soils of Iran. *Pedosphere* 21:203–237
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723
- Al-Turki AM, Al-Farraj AS, Sallam AS, Khalil GA, Al-Malik AS (2008) Survey of land resources and evaluation of water quality in southern Tohama Plains, Kingdom of Saudi Arabia. Final Report, King Abdul Aziz City for Science and Technology
- Botula YD, Cornelis WM, Baert G, Van Ranst E (2012) Evaluation of pedotransfer functions for predicting water retention of soils in the Lower Congo (D.R. Congo). *Agric Water Manag* 111:1–10
- Bouma J (1989) Using soil survey data for quantitative land evaluation. *Adv Soil Sci* 9:177–213
- Brooks RH, Corey AT (1964) Hydraulic properties of porous media. Hydrology paper no. 3. Colorado State University, Fort Collins
- Bruand A (2004) Utilizing mineralogical and chemical information in PTFs. In: Pachepsky Y, Rawls WJ (eds) *Developments in soil science*. Elsevier, pp 153–158
- Campbell GS (1974) A simple method for determining unsaturated hydraulic conductivity from moisture retention data. *Soil Sci* 117:311–314
- Campbell GS (1985) *Soil physics with basics*. Elsevier, New York
- Cornelis WM, 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 JH, Topp GC (ed) (2002) *Methods of soil analysis, part 4, physical methods*. *Soil Sci Soc of Am*. <https://doi.org/10.2136/sssabookser5.4.frontmatter>
- Gupta SC, Larson WE (1979) Estimating soil water retention characteristics from particle size distribution, organic matter content and bulk density. *Water Resour Res* 15:1633–1635
- Haghverdi A, Ozturk HS, Cornelis WM (2014) Revisiting the pseudo continuous pedotransfer function concept: impact of data quality and data mining method. *Geoderma* 226–227:31–38
- Hodnett MG, Tomasella J (2002) Marked differences between van Genuchten soil water retention parameters for temperate and tropical soils: a new water retention pedotransfer functions developed for tropical soils. *Geoderma* 108:155–180
- Kern JS (1995) Evaluation of soil water retention models based on basic soil physical properties. *Soil Sci Soc Am J* 59:1134–1141
- Khodaverdiloo H, Homae M, van Genuchten M, Dashtaki SG (2011) Deriving and validating pedotransfer functions of some calcareous soils. *J Hydrol* 399:93–99
- Liao KH, Xu SH, Wu JC, Ji SH, Lin Q (2011) Assessing soil water retention characteristics and their spatial variability using pedotransfer functions. *Pedosphere* 21:413–422

- Lin HS, McInnes KJ, Wilding LP, Hallmark CT (1999) Effects of soil morphology on hydraulic properties. II. Hydraulic pedotransfer functions. *Soil Sci Soc Am J* 63:955–961
- Loague K, Green RE (1991) Statistical and graphical methods for evaluating solute models: overview and application. *J Contam Hydrol* 7:51–73
- Loeppert RH, Suarez DL (1996) Carbonate and gypsum. In: Sparks DL (ed) *Methods of soil analysis. Part 3 chemical methods*. SSSA, Madison, pp 437–474
- Mayr T, Jarvis NJ (1999) Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Corey type model. *Geoderma* 91:1–9
- Mermoud A, Xu D (2006) Comparative analysis of three methods to generate soil hydraulic functions. *Soil Tillage Res* 87:89–100
- Minasny B, Hartemink AE (2011) Predicting soil properties in the tropics. *Earth-Sci Rev* 106:52–62
- Minasny B, McBratney AB, Bristow KL (1999) Comparisons of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma* 93:225–253
- Nash JE, Sutcliffe JV (1970) River forecasting through conceptual models, part I, a discussion of principle. *J Hydrol* 10:282–290
- Nelson DW, Sommers LE (1982) Total carbon, organic carbon and organic matter. In: Page A et al (eds) *Methods of soil analysis. Part 2, 2nd ed.* Agronomy 9, Amer. Soc. of Agron., Madison, pp 539–579
- Nemes A, Schaap MG, Wosten JH (2002) Validation of international scale soil hydraulic pedotransfer functions for national scale applications. In: Symposium no. 4 on the 17<sup>th</sup> WCSS, paper no. 934, 14–21 August, Thailand
- Nemes A, Schaap MG, Wosten JH (2003) Functional evaluation of pedotransfer functions derived from different scales of data collection. *Soil Sci Soc Am J* 67:1093–1102
- Nguyen PM, Le KV, Botula Y, Cornelis WM (2015) Evaluation of soil water retention pedotransfer for Vietnamese Mekong Delta soils. *Agric Water Manag* 158:126–138
- Nielsen WE, van Genuchten MT, Biggar JW (1986) Water flow and solute transport processes in the unsaturated zone. *Water Resour Res* 22:89–108
- Obalum SE, Obi ME (2013) Moisture characteristics and their point pedotransfer functions for coarse- textured tropical soils differing in structural degradation status. *Hydrol Process* 27:2721–2735
- Pachepsky YA, Rawls WJ (1999) Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Sci Soc Am J* 63:1748–1757
- Rawls WJ, Brakensiek DL, Saxton KE (1982) Estimation of soil water properties. *Trans Am Soc Agric Eng* 25:1316–1320 1328
- Rawls WJ, Pachepsky YA, Ritchie JC, Sobecki TM, Bloodworth H (2003) Effect of soil organic carbon on soil water retention. *Geoderma* 116:61–76
- Rhoades JD (1996) Salinity: electrical conductivity and total dissolved solids. In: Sparks DL (ed) *Methods of soil analysis. Part 3 chemical methods*. SSSA, Madison, pp 417–436
- Russo D (1988) Determining soil hydraulic properties by parameter estimation: on the selection of a model for the hydraulic properties. *Water Resour Res* 24:453–459
- Saxton KE, Rawls WJ, Romberger JS, Papendick RI (1986) Estimating generalized soil water characteristics from texture. *Soil Sci Soc Am J* 50:1031–1036
- Schaap MG, Leij FJ (1998) Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil Tillage Res* 47:37–42
- Schaap MG, Leij FJ, Van Genuchten MT (1998) Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Sci Soc Am J* 62:847–855
- Schaap MG, Leij FJ, Van Genuchten MT (2001) Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J Hydrol* 251:163–176
- Tamari S, Wosten JH, Ruiz-Suarez JC (1996) Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci Soc Am J* 60:1732–1741
- Tietje O, Tapkenhinrichs M (1993) Evaluation of pedo-transfer functions. *Soil Sci Soc Am J* 57:1088–1095
- Tomasella J, Hodnett MG, Rossato L (2000) Pedotransfer functions for the estimation of soil water retention in Brazilian soils. *Soil Sci Soc Am J* 64:327–338
- Van Genuchten MT (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci Soc Am J* 44: 892–898
- van Genuchten MTh, Leij FJ, Yates SR (1991) The RETC code for quantifying the hydraulic functions of unsaturated soils. EPA/600/2–91/065 – U. S. Salinity Laboratory. USDA, ARS, Riverside, CA
- Vereecken H (1995) Estimating the unsaturated hydraulic conductivity from theoretical models using simple soil properties. *Geoderma* 65:81–92
- Vereecken H, Maes J, Feyen J (1989) Estimating the soil water moisture retention characteristics from texture, bulk density, and carbon content. *Soil Sci* 148:389–403
- Wang T, Franz TE, Zlotnik VA (2015) Controls of soil hydraulic characteristics on modeling groundwater recharge under different climatic conditions. *J Hydrol* 521:470–481
- Wosten JH, Lilly A, Nemes A, Le Bas C (1999) Development and use of a database of hydraulic properties of European soils. *Geoderma* 90: 169–185