Journal of Hydrology 563 (2018) 251–259

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Measurement and estimation of the soil water retention curve using the evaporation method and the pseudo continuous pedotransfer function

Amir Haghverdi[a,*], Hasan Sabri Öztürk[b], Wolfgang Durner[c]

[a] Environmental Sciences Department, University of California Riverside, Riverside, CA 92521, USA
[b] Department of Soil Science and Plant Nutrition, Faculty of Agriculture, Ankara University, Turkey
[c] Institute of Geoecology, Technische Universität Braunschweig, Germany

ARTICLE INFO

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Zhiming Lu, Associate Editor

Keywords:
Data mining
Evaporation method
Neural network
Pseudo continuous pedotranfer function
Soil water retention curve

ABSTRACT

HYPROP (Hydraulic Property Analyzer) system works based on the simplified evaporation method, which determines soil hydraulic properties in the laboratory. This system simultaneously monitors the changes in weight of soil samples as well as the soil matric potential at two depths throughout a drying process governed by evaporation from the soil surface. In this study, we examined the performance of the pseudo continuous pedotranfer function (PC-PTF) to estimate the soil water retention curve (SWRC) using high resolution data measured by HYPROP system. The dataset consisted of 7963 measured water retention data points obtained from 81 Turkish soil samples from which 60%, 20% and 20% were randomly selected for training, cross-validation and test subsets, respectively. The best PC-PTF developed in this study with a mean absolute error of 0.023 $m^3\ m^{-3}$ (a root mean square error of 0.033 $m^3\ m^{-3}$) and a correlation coefficient of 0.96 showed promising performance considering the typical performance range (RMSE = 0.034–0.085 $m^3\ m^{-3}$) for parametric PTFs estimating SWRC. The best PC-PTF used soil textural information, soil bulk density, the percentage of stable aggregates, soil organic matter content and the initial water content as the input attributes. PTFs developed in this study also ranked high among previously developed PC-PTFs (with RMSE ranging from 0.027 to 0.056 $m^3\ m^{-3}$) using sparse datasets collected via the traditional equilibrium approach (i.e. sandbox apparatus /pressure plates extractor). We, therefore, recommend further application of the PC-PTF approach for development of PTFs using high resolution data obtained by HYPROP system.

1. Introduction

The soil water retention curve (SWRC) is arguably the most important curve in the soil science, which provides critical information regarding the unsaturated behavior of soils with a wide range of applications in the agro-hydrological studies. In addition, it is often used to estimate the soil hydraulic conductivity, by capillary bundle models (Burdine, 1953; Mualem, 1976), which together are used to study the unsaturated water flow as the controlling components when applying Richards equation. As a result, a significant body of research in the soil science literature has been devoted to the measurement, modeling and estimation of SWRC.

There is no single laboratory device available to measure water retention data over the entire range of SWRC. The standard equilibrium method to measure soil water retention of undisturbed samples in the laboratory is by using the sandbox apparatus, the sand /kaolin box and the pressure plate extractor with typical usage ranges of 0–10, 10–50, and 100–1500 kPa, respectively. However, it has been reported that the pressure plates are susceptible to substantial errors. Recent studies have demonstrated a lack of reliability of the pressure plates when measuring SWRC (in particular in the dry range) due to low plate and soil conductance, loss of soil-plate contact due to the soil shrinkage, blocking of the pores in the ceramic porous plate, and soil dispersion (Bittelli and Solone, 2012; Solone et al., 2012; Cresswell et al., 2008). In addition, the measurements via the equilibrium method are time consuming and it takes weeks, if not months, to reach the equilibrium at each soil matric potential. Moreover, measurement campaigns based on the equilibrium methods typically yield only a limited number of water retention data points (usually less than 10) per soil sample.

The evaporation method (Wind 1968; Wendroth et al. 1993; Halbertsma 1996) is an alternative option, which concurrently produces the soil water retention and hydraulic conductivity data through monitoring of the water content (evaporation) and matric potential dynamics at various heights of a soil sample exposed to evaporation. Schindler (1980) showed that the measurement of the soil matric potential at only two depths produces promising results for a wide range

* Corresponding author.
E-mail address: amirh@ucr.edu (A. Haghverdi).

https://doi.org/10.1016/j.jhydrol.2018.06.007
Received 13 March 2018; Received in revised form 1 June 2018; Accepted 4 June 2018
Available online 05 June 2018

of soil textures (Schindler et al., 1985). The range of typical tensiometers (less than 70 kPa) limits the operational range of the evaporation method. The extended evaporation method (Schindler et al., 2010a,2010b), however, extends the range up close to the wilting point using new cavitation tensiometers and by considering the air-entry value as an additional matric potential measurement (Schindler et al., 2016). Schindler et al. (2012) conducted a study to compare SWRCs determined from the traditional equilibrium methods with those obtained with the extended evaporation method. They found the results of the two approaches to be comparable, reported no systematic deviation between the water retention data and only observed less than a 2% deviation in average water content between 0 and −1500 kPa.

HYPROP system (Hydraulic Property Analyzer, METER Group, Inc., Pullman, WA, USA) is a semi-automated commercial laboratory instrument that works based on a simplification of the evaporation method proposed by Schindler (1980) which was shown to adequately determine the hydraulic properties of the most soils (Peters et al., 2015; Peters and Durner, 2008). HYPROP has two advantages over the traditional equilibrium method: it generates high resolution water retention data (over 100 water retention data points in the 0–100 kPa range) and the measurement cycle for each sample is typically completed in a few days. Schelle et al. (2013) evaluated and compared several water retention measurement methods on a wide range of soil texture classes including sand, silt, silt loam, clay loam and clay. They reported comparable water retention data between the hanging water column method and HYPROP in the wet to moderate water content range, yet noticed some tendency for the over estimation of the water content in the retention data measured by the pressure plate extractor. Zhuang et al. (2017) observed an excellent agreement between the soil hydraulic properties measured by HYPROP against the multistep flux and the hanging water column experiments.

Pedotransfer functions (PTFs) are statistical models mostly developed to estimate soil hydraulic characteristics from readily available basic properties of the soil. In many practical applications, PTF-estimated soil hydraulic characteristics are often used when adequate measured data are not available. In general, PTFs derived using artificial neural network (ANN) models have been shown to outperform PTFs developed using regression models (Vereecken et al., 2010). This superior performance is mainly attributed to the ability of ANNs to model PTF input-output relationships without any a priori functional form (Vereecken et al., 2010). PTFs are developed to either estimate the soil water content at a few predefined soil matric potentials, i.e. point PTFs, or to continuously estimate the soil water content at any desired soil matric potential, typically parametric PTFs (Wösten et al., 2001). Parametric PTFs estimate parameters of soil hydraulic equations which are subsequently used to estimate water contents across a wide range of soil matric potentials. The main disadvantage of parametric PTF is once the PTF is developed user cannot change the preselected soil hydraulic equation, which may or may not adequately represent the real shape of SWRCs for all samples. In addition, it is sometimes challenging to correlate the parameters of soil hydraulic equations to basic soil properties (Minasny and McBratney, 2002). In recent years, the pseudo continuous PTF (PC-PTF, Haghverdi et al., 2012, 2014) and the *k*-nearest neighbor (*k*-NN) PTF (Haghverdi et al., 2015a) were proposed as two alternative approaches to estimate SWRCs. These approaches allow withdrawal and replacement of soil hydraulic equations by an appropriate data mining method (PC-PTF) or potential concurrent use of multiple soil hydraulic equations (*k*-NN PTF).

Jain et al. (2004) showed that a three-layer feed forward ANN with one input predictor representing the soil matric potential and one output node for the soil water content provides equally or more accurate estimation of SWRC as compared to some widely-used soil hydraulic functions. Haghverdi et al. (2012) combined the ANN topology suggested by Jain et al. (2004) with the topology of ANN point PTF to develop ANN PC-PTF. They showed that PC-PTF performed slightly better than parametric PTFs using datasets with limited water retention data per sample. The unique topology of PC-PTF allows to combine different datasets with various measured water retention points and to utilize datasets with uneven/missing measured water retention points among soil samples. PC-PTF has shown promising results for soils from Iran and Australia (Haghverdi et al., 2012), Turkey and Belgium (Haghverdi et al., 2014), USA (Haghverdi et al., 2015b), and Vietnam (Nguyen et al., 2017). Haghverdi et al. (2014) and Nguyen et al. (2017) examined several statistical approaches (e.g. ANN, support vector machines, *k*-NN, and multiple linear regression) and concluded ANN to be the best modeling approach to derive PC-PTFs.

In recent years, HYPROP system has been increasingly used by researchers from different parts of the world (e.g. Schwen et al., 2014; Grath et al., 2015; Herbrich and Gerke, 2017; Schindler et al., 2012; Zhuang et al., 2017, Fields et al., 2016). Therefore, in the near future, large datasets will likely be compiled to develop PTFs using HYPROP data. Recently, Schindler and Müller (2017) published a database of 173 soil hydrological data from 71 sites collected from all around the world via the common evaporation method, the extended evaporation method and HYPROP system. There remains the opportunity to develop more performing PTFs when high resolution water retention data produced by HYPROP system are used.

Unlike parametric PTF, PC-PTF does not incorporate certain assumptions about the shape of the SWRC and only utilizes measured water retention data and relies on the power of ANN models to determine the real shape of the SWRC. Therefore, PC-PTF has to cover more nonlinearity and more complex relationships among input-input and input–output attributes, which in part is covered by soil hydraulic equations in parametric PTF (Haghverdi et al., 2012, 2014). Consequently, the number of soil samples, the density of the measured water retention data and their distribution in wet, moist (midrange) and dry ranges of the curve, and the best combination of input predictors should be identified carefully when developing PC-PTFs. Our main hypotheses for this study are that (i) PC-PTF can accurately model SWRC using high resolution data measured by HYPROP, and (ii) High resolution data provided by HYPROP result in equally or more accurate estimations of SWRC compared to sparse datasets collected via the traditional equilibrium approach (i.e. the sandbox apparatus along with the pressure plates extractor).

In addition, we examined the following three key questions:

1. What combination of input attributes provides the best estimation of SWRC?
2. How does abundant measured water retention data in wet and moist regions of the curve versus lack of data in the dry range of the curve affect the estimation of SWRC by PC-PTF?
3. Does adding an arbitrary oven-dry water retention point improve the performance of PC-PTF in the dry range of the curve?

2. Materials and methods

2.1. Data collection and laboratory analyses

A total of 135 undisturbed (100 cm^3 stainless-steel soil sampling cylinder) and disturbed soil samples were collected from the soil surface (0–30 cm) mainly from the areas surrounding Ankara, Turkey in 2010. Some basic soil properties were measured by using the disturbed soil samples in the soil physics laboratory at Ankara University, Turkey. Soil organic matter content (OM) and soil texture were determined by the modified method of Walkley and Black (Jackson, 1958) and the hydrometer method (Gee and Bauder, 1986), respectively. The percentage of stable aggregates (SA) was measured by the wet sieving apparatus (Eijkelkamp Agrisearch Equipment, Giesbeek, The Netherlands). The undisturbed samples were brought up to saturation in the laboratory and subsequently were put in the sandbox apparatus and the pressure plate extractor (Eijkelkamp Agrisearch Equipment, Giesbeek, The Netherlands) to measure the water content at 8 matric potentials

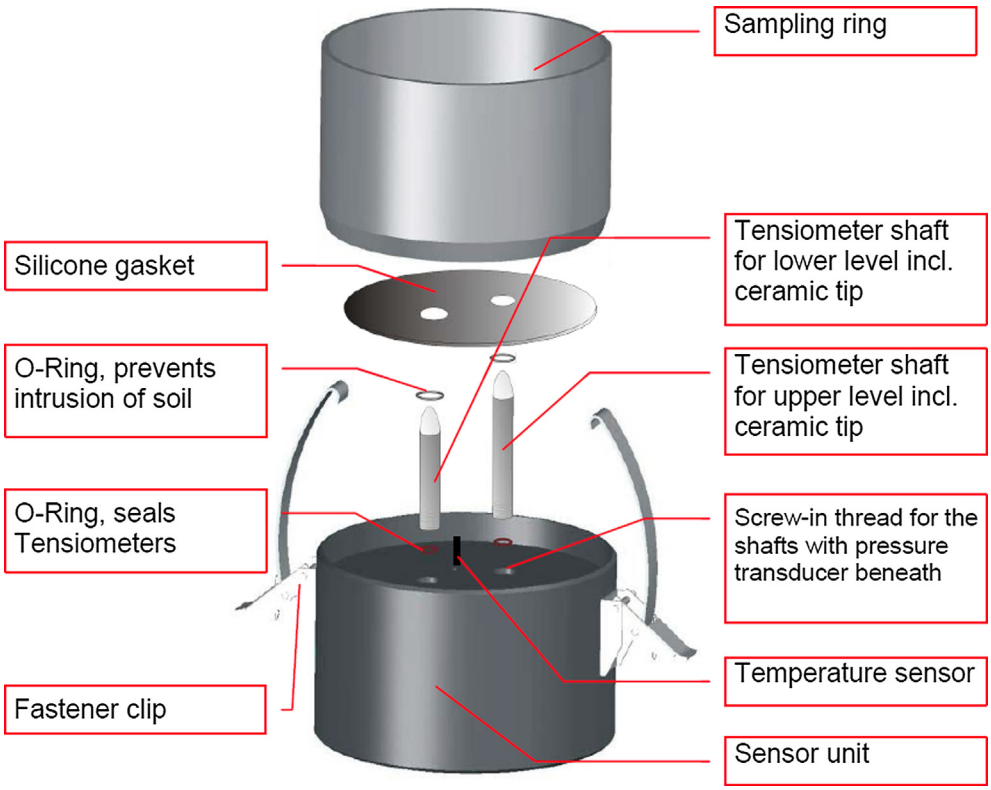

Fig. 1. Components of HYPROP system main body (adapted from HYPROP user manual, Meter Group Inc., 2018).

ranging from −5 to −1500 kPa. The soil cores were later used to determine bulk density (BD) (Blake and Hartge, 1986). More details about this soil dataset is outlined in Haghverdi et al. (2014).

The remaining disturbed sieved samples were sent to the institute of environmental sciences at Technical University of Braunschweig in Germany for additional analyses using HYPROP system (Fig. 1). The remaining soil materials were enough to successfully prepare 81 samples (out of the original 135 soil samples) for the measurement of soil hydraulic properties using HYPROP system. The samples were packed in stainless-steel cylinders (approximately 250 cm^3 with an inside diameter of 8 cm and a height of 5 cm) to BDs close to the field condition and then were saturated. A small auger was used to make two holes for the two vertically aligned tensiometers, designed for the optimum implementation of the evaporation method. The tensiometers were positioned such that the center of the soil sample (2.5 cm) was in the middle between the tensiometers' tips (i.e. 1.25 cm and 3.75 cm). Then, the bottom side of the sample was closed throughout the experiment, but the upper part was open to the atmosphere for evaporation. The soil matric potentials were recorded at different points of time (approximately once a minute for the first hour and every ten minutes afterwards) at the two depths and the weight of the sample were measured twice a day following the single balance mode described in HYPROP operation manual (Meter Group Inc., 2018). The measurement campaign lasted on average 9 days for each sample (varying between 5 and 13 days for different samples depending on the soil type). The data were post-processed to construct SWRC by relating the mean soil water content (obtained from sample mass) to the mean soil matric potential (obtained from the tensiometers at the two depths). The basic assumptions for the analytical approach were the linear distribution of the soil matric potential and the water content through the soil column, and the linear changes of weighs and water tensions between any two evaluation points (Pertassek et al., 2011).

Fig. 1 illustrates the main components of HYPROP system. Three distinct stages can be recognized in a typical dataset obtained during HYPROP measurement campaigns (Fig. 2): (I) the initial stage in which

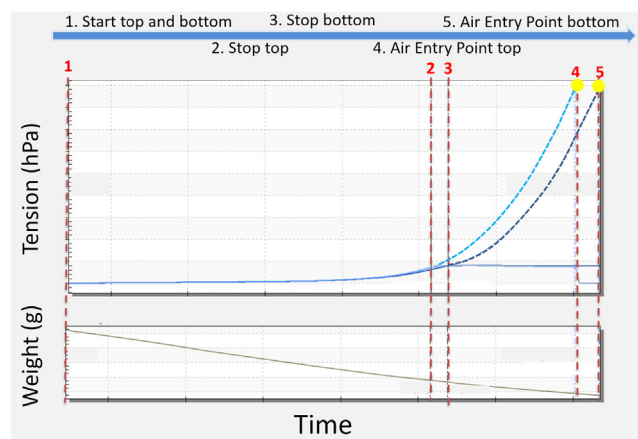

Fig. 2. Dynamics of tension and weight for the top and bottom tensiometers throughout a typical measurement campaign by HYPROP system. Dashed blue lines are interpolated using support points (yellow dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the measured soil matric potential reflects the real tension of the soil surrounding tensiometers, (II) the "vapor pressure stage" in which tensiometers readings are no longer representative of the real soil tension, (III) the "air-entry stage" which takes place when the tension in the soil surrounding tensiometers becomes greater than the air-entry pressure of the ceramic material and therefore the measured soil matric potential reduces to zero. HYPROP uses the initiation of the stage three as an extra soil matric potential measurement point which is the underlying idea behind the extended evaporation method (Pertassek et al., 2015).

Table 1 summarizes the characteristics of the soils used in this study. Fig. 3 illustrates the textural distribution of the samples and the water retention data measured by HYPROP system (on average 98

Table 1
Characteristics of the soils used in this study.

Attribute	Min.	Max.	SD	Average
Sand (%)	5.9	83.6	17.2	35.3
Silt (%)	5.2	57.6	8.6	30.5
Clay (%)	9.4	62.2	14.9	34.2
BD (Mg m^{-3})	0.69	1.33	0.13	0.98
SOM (%)	0.01	3.85	0.65	1.21
SA (%)	1.50	75.40	18.72	30.72
pF	−2.00	3.89	0.76	1.75
SWC (m^3 m^{-3})	0.09	0.74	0.11	0.42
IWC (m^3 m^{-3})	0.29	0.74	0.08	0.52

BD: soil bulk density; SOM: soil organic matter content; SA: percentage of stable aggregates; SD: standard deviation; SWC: soil water content measured by HYPROP system; IWC: initial water content; pF = Log (h) where h is the soil matric potential (cm water column).

water retention data points were measured per soil sample in the range between saturation and close to the permanent wilting point with an average pF of 1.75). We observed some differences between SWRCs obtained via the equilibrium and HYPROP approaches (Öztürk et al., 2013) which is mainly attributed to the differences in soil structures caused by the use of repacked soil samples for the evaporation experiment. However, the methodological approach of PC-PTF (as described next) is not affected by the sample preparation method (i.e. the undisturbed versus repacked cylinders), hence its practical applicability is not hampered by the use of the repacked samples in this study.

2.2. Deriving pseudo continuous pedotransfer functions

Fig. 4 shows the typical topology of ANN PC-PTF. ANN PC-PTF has an extra input node compared to the traditional point and parametric PTFs. The extra node provides the soil matric potential information to ANN, which allows PTF to estimate water retention at any given soil matric potential. Other input attributes are a set of basic soil properties, which allows PTF to distinguish among soils and understand how differences among soil basic information impact the curves.

Neurosolution 7.1.1.1 (NeuroDimension Inc., Gainesville, FL, USA) was used to develop three-layer feed-forward perceptron ANN models. The transfer functions were the hyperbolic tangent and linear for the hidden and the output layers, respectively. The learning algorithm was Levenberg-Marquardt (Bishop, 1995). The maximum epoch was set to 1000 and the number of runs was set to 3 while the weights were updated after the presentation of the entire training set. The best weights were loaded automatically for testing. To eliminate the possibility of

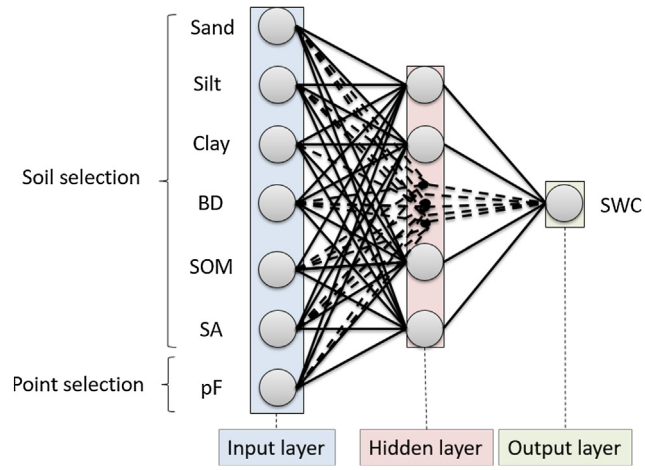

Fig. 4. Topology of the three-layer feed-forward artificial neural network (ANN) pseudo continuous pedotransfer function (PC-PTF). BD: bulk density; SOM: soil organic matter content; SA: percentage of stable aggregates; pF: Log (h) where h is the soil matric potential (cm water column); SWC: soil water content.

the over-training, training was terminated when the mean square error of the cross-validation subset either began to increase or showed no improvement after 100 iterations. The number of neurons of the hidden layer was changed from 1 to 20. Soil samples were randomly partitioned into 5 folds: three for training, one for cross-validation and one for test subsets. The model development process was repeated 5 times in order to use all 5 folds as test subsets. When using PC-PTFs to estimate SWRCs, saturation (h = 0 kPa or pF = 0) and the permanent wilting point (h = −1500 kPa or pF: 4.18) were considered as the upper and lower bounds, respectively.

Two filters were designed and implemented using Microsoft Excel to convert raw outputs of ANN models to final PTF outputs. The first filter replaced negative estimations with zero, while the second filter adjusted the curvature of SWRCs to ensure that no increase in the water content occurs as the soil dried from fully saturated down to the permanent wilting point and the soil matric potential becomes more negative (Fig. 5). For each soil, the second filter was implemented twice from the dry-to-wet end and from the wet-to-dry end such that any incorrect water retention point was replaced with the water retention of the previous measured point (i.e. a lower suction for the wet-to-dry end and a higher suction for the dry-to-wet end). The final output was the average of the two implementations.

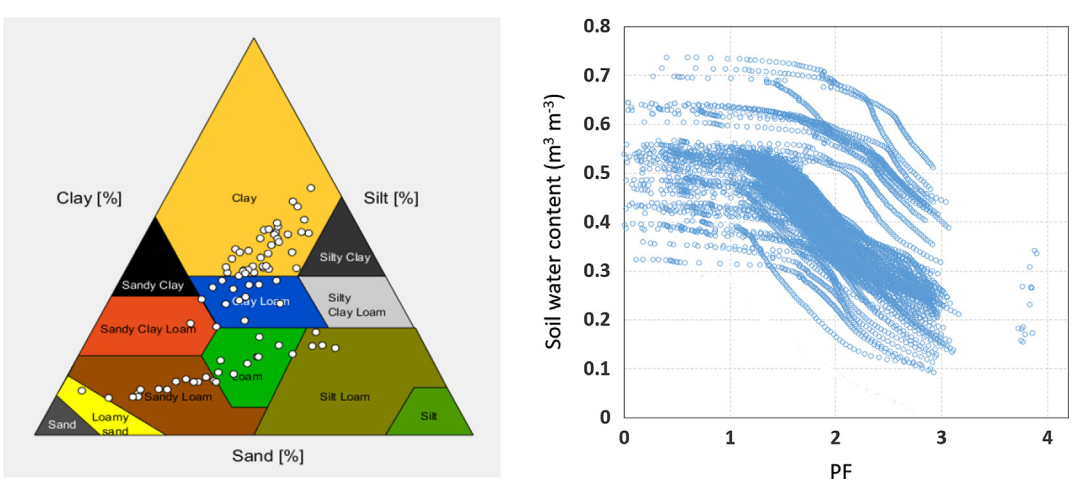

Fig. 3. Left: Soil textual classes of soil samples used in this study. Right: Water retention data measured by HYPROP system throughout this study. pF = Log (h) where h is the soil matric potential (cm water column).

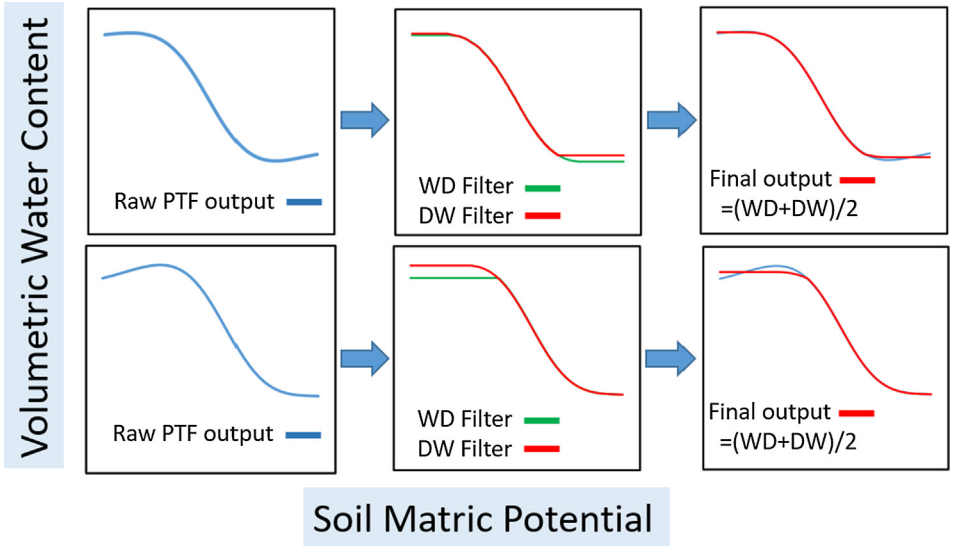

Fig. 5. A step-by-step illustration of how the second filter adjusts raw outputs of ANN PC-PTFs to generate final PTF estimated SWRCs for two hypothetical soil samples. WD: wet-to-dry end; DW: dry-to-wet end.

Table 2
Combinations of input attributes used to develop artificial neural network (ANN) pseudo continuous pedotransfer functions (PC-PTF).

Model	Input attributes
1	SSC
2	SSC, BD
3	SSC, SOM
4	SSC, SA
5	SSC, SA, BD, SOM
6	SSC, SA, BD, SOM, IWC

SSC: silt, sand, clay content; BD: bulk density; SA: percentage of stable aggregates; SOM: soil organic matter content; IWC: initial water content.

Table 2 summarizes 6 models developed in this study using different combinations of input attributes (hereafter referred to as M1, M2, M3, M4, M5, and M6). The first three combinations include the typical input attributes (i.e. soil texture, BD and OM) used in the literature to develop PTFs for SWRC. Two additional inputs including the initial soil water content (IWC) and AS were incorporated in the remaining models. Correlation among input attributes was assumed to have no negative impact on the performance of PC-PTF due to the insensitive nature of ANNs to multicollinearity (De Veaux and Ungar, 1994). IWC represents the water content of the completely saturated samples, which is calculated from the total loss of water by evaporation and oven drying. Note that, BDs of repacked soils in Germany were highly correlated ($r = 0.85$) but not identical to BDs of the original undisturbed soils collected in Turkey. Therefore, we used BDs of repacked soils to develop PTFs.

A sensitivity analysis was performed on M6, which contained all input predictors to better understand the relative importance of each input and to illustrate how the estimated water content value varies in response to variation of an input. Best ANN weights were loaded and sensitivity analysis was performed on the training data (the process was repeated 5 times using all combinations of the training data). Each input predictor was varied in 50 steps above and below its mean (upper limit: mean + SD, lower limit: mean-SD) while all other inputs kept fixed at their respective means. This process was repeated for each input predictor and the standard deviation of the estimations was reported as sensitivity.

Our initial modeling attempts revealed that lack of data in the dry part of the curve caused PC-PTF estimated SWRCs to quickly level off, which resulted in a relatively high residual water content for some soils. To address this issue, we added the oven-dry point (water content = 0 $m^3\ m^{-3}$ and pF = 6.8) as an additional arbitrary data point to all samples and derived a new set of M6 PTFs.

2.3. Evaluation statistics

The mean absolute error (MAE) and the correlation coefficient (r) were chosen as the main statistics to evaluate the accuracy of the models, yet the root mean square error (RMSE) was also calculated to allow comparison with those studies that did not report MAE values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |E_i - M_i| \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (E_i - M_i)^2} \tag{2}$$

$$r = \frac{\sum_{i=1}^{n} (E_i - \overline{E})(M_i - \overline{M})}{\sqrt{\sum_{i=1}^{n} (E_i - \overline{E})^2 \sum_{i=1}^{n} (M_i - \overline{M})^2}} \tag{3}$$

where E and M are the estimated and measured water content values ($m^3\ m^{-3}$) using PTF and HYPROP system, respectively, $\overline{E}$ and $\overline{M}$ are the mean estimated and measured water content values, and n is the total number of the measured water retention points for all samples ($n = 7963$). The evaluation statistics were calculated only for the test data for all the models outlined in Table 2.

3. Results

Table 3 summarizes the performance evaluation statistics for all 6 models (calculated during the test phase only). The MAE ranges from 0.023 to 0.105 $m^3\ m^{-3}$ and the correlation coefficient varies from 0.15 to 0.96. M6 followed by M5 show the best performances. The lowest estimation accuracy belongs to M3 and M1.

Fig. 6 illustrates the estimated water content data by the best (M6) and the worst (M3) PC-PTF models. Fig. 7 depicts the results of the sensitivity analysis. PC-PTF is the most sensitive to input attributes IWC and pF. It shows moderate sensitivity to the soil textural components (i.e. percentages of sand, silt and clay) and BD. The lowest relative sensitivity is observed to SOM and SA. The variation in sensitivity

Table 3
Performance of neural network (ANN) pseudo continuous pedotransfer functions (PC-PTF) developed in this study.

	PTF models					
	M1	M2	M3	M4	M5	M6
MAE	0.093	0.055	0.105	0.078	0.043	0.023
r	0.41	0.76	0.15	0.58	0.85	0.96
RMSE	0.129	0.080	0.159	0.107	0.061	0.033

The mean absolute error (MAE), the correlation coefficient (*r*) and the root mean square error (RMSE) were calculated using Eqs. (1)–(3), respectively.

among the 5 versions of M6 (developed using different realizations of the input data) is high for the soil textural components and low for SOM, SA, IWC, and the soil matric potential.

Fig. 8 illustrates the performance of the PTFs across textural classes against the number of measured water retention points for each class. The lowest number of data in our dataset belongs to the loamy sand (1% of data) and the sandy clay loam (1% of data) textural classes, which shows a relatively higher error. Clay is the dominant class (47% of data) followed by sandy loam (17% of data) both showing the lowest error among all classes. The error remarkably decreases as the number of the data points increases but levels off such that the performance of PTFs for the loam (13% of data), the clay loam (16%) and the sandy loam (17%) classes only slightly differs from that for the clay class (47% of data).

Fig. 9 illustrates SWRCs estimated by M6 across the textural classes, using the original dataset and the modified dataset with the oven-dry point included. Two fitted SWRCs are also depicted for each sample using the van Genuchten soil hydraulic model (van Genuchten, 1980): (i) the original curve with the residual water content estimated during the curve fitting process and (ii) the modified curve which ensured the water content matched zero at the oven-dry point (pF: 6.8). The PTF estimated SWRCs for the clay texture, as the dominant class in our dataset, show a high agreement with the fitted curves, while the differences between the curves increase for the textural classes with lower data percentage share. Impact of additional water retention point (i.e. the oven-dry point) differs among the soil samples even for samples from the same textural class. For instance, incorporation of the oven-dry point in the training process causes slight, moderate and high changes in SWRCs for the clay samples 21, 10, and 13, respectively.

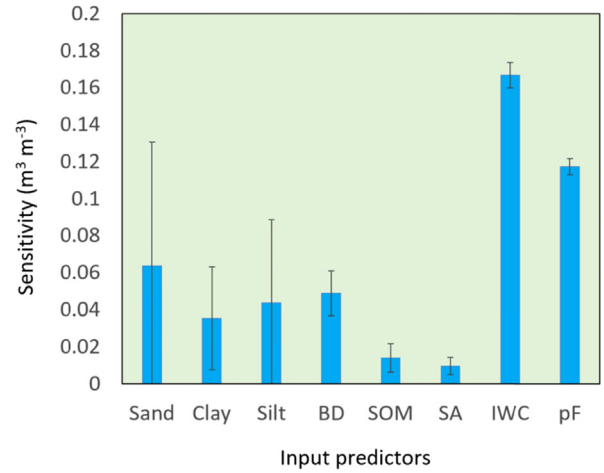

Fig. 7. Sensitivity of the neural network (ANN) pseudo continuous pedotransfer function (PC-PTF) model 6 (M6) to the input predictors. BD: bulk density; SOM: soil organic matter content; SA: percentage of stable aggregates; IWC: initial water content; pF: Log (*h*) where *h* is the soil matric potential (cm water column).

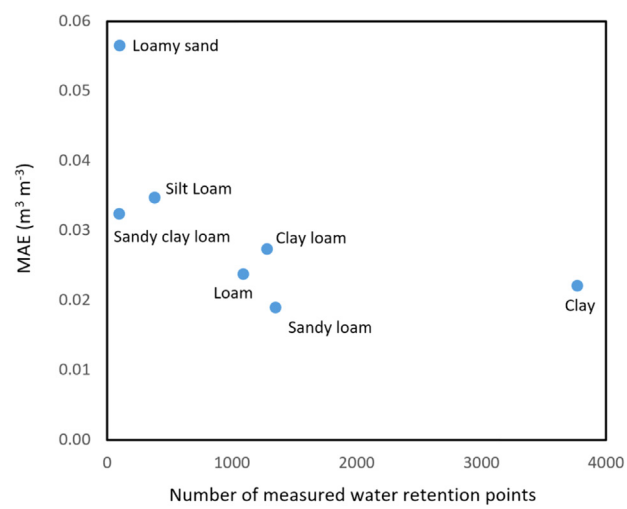

Fig. 8. Relationship between the number of data points for each textural class in the training dataset and the estimation accuracy of artificial neural network (ANN) pseudo continuous pedotransfer functions (PC-PTF).

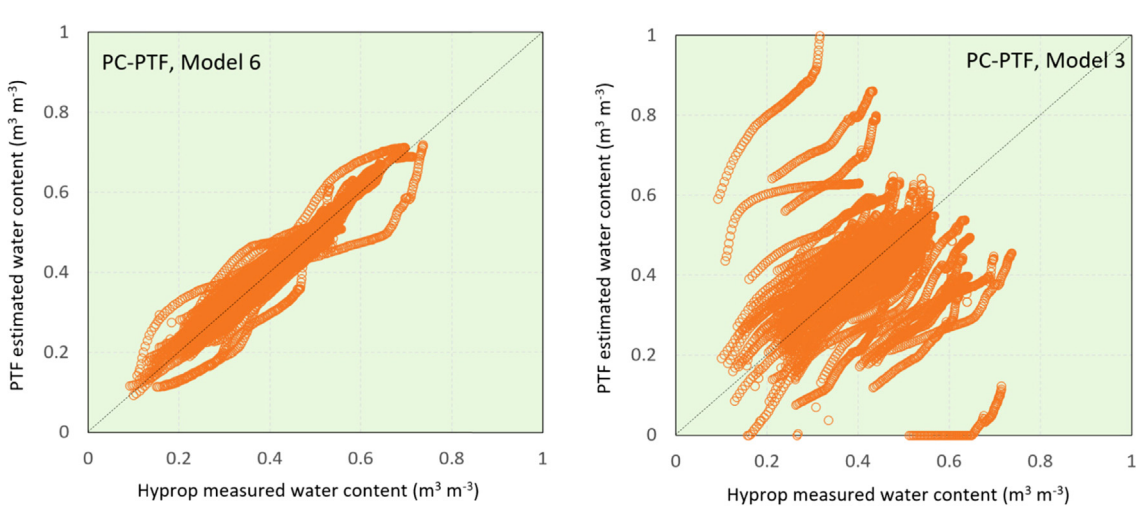

Fig. 6. Scatter of the measured versus estimated water content around 1:1 line for artificial neural network (ANN) pseudo continuous pedotransfer functions (PC-PTF): the best and the worst performance belonged to M6 (left) and M3 (right), respectively.

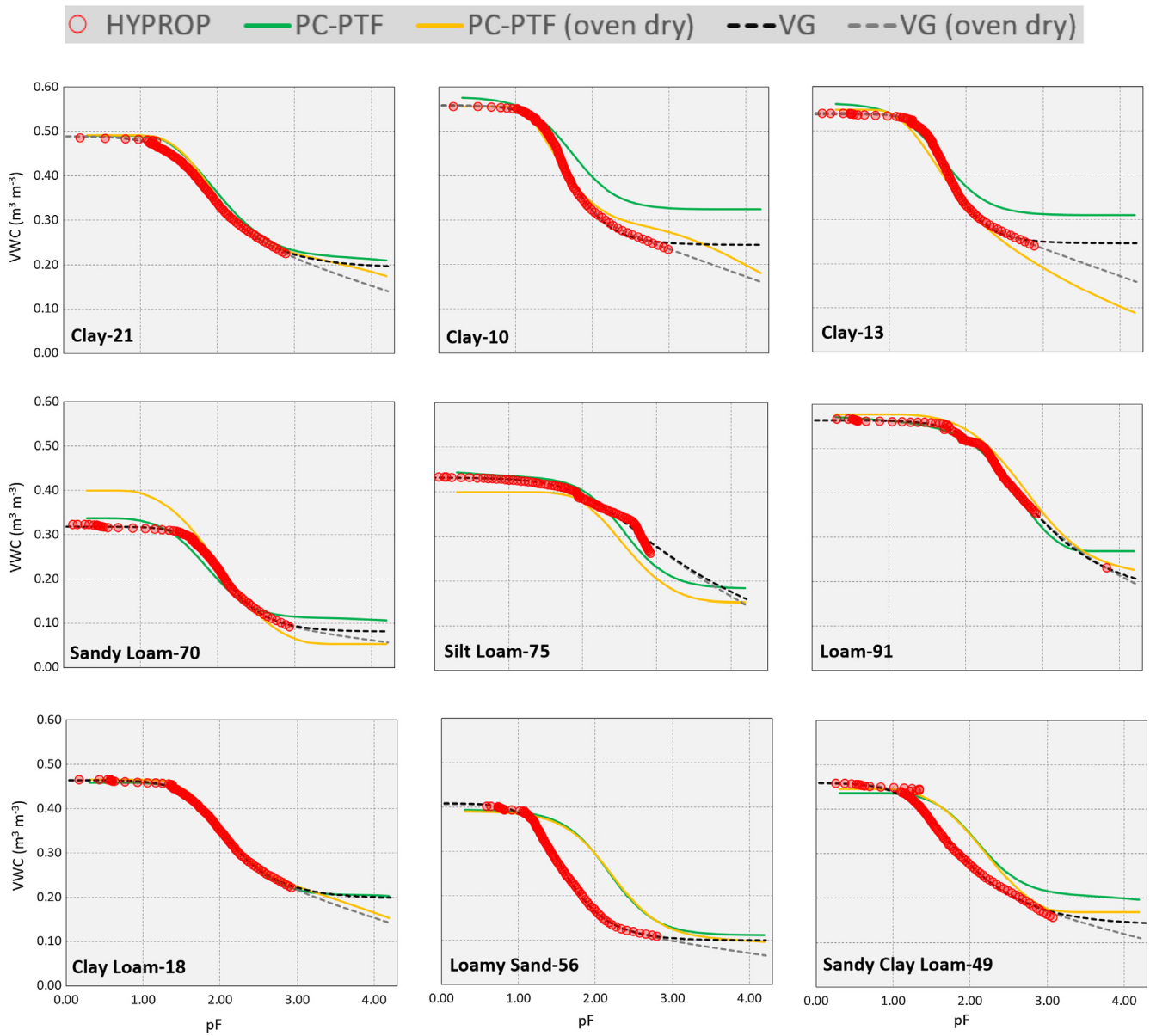

Fig. 9. Estimated SWRCs by neural network (ANN) pseudo continuous pedotransfer functions (PC-PTF) for some soil samples across the textural classes. Red circles: measured data by HYPROP system; Green line: PC-PTF estimated SWRC using the original dataset; Yellow line: PC-PTF estimated SWRC with oven-dry as an additional water retention point; Dashed black line: Van Genuchten (VG; Van Genuchten, 1980) fitted SWRC; Gray dashed line: VG fitted curve with oven-dry as an additional water retention point which ensured the fitted curve matches zero at pF = 6.8. VG curves were fitted using HYPROP-FIT software (Pertassek et al., 2011, 2015). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

The typical RMSE values for the parametric PTFs predicting SWRC vary between 0.034 and 0.085 $m^3\ m^{-3}$ (Vereecken et al., 2010) which puts the best PC-PTF developed in this study (i.e. M6 with RMSE equal to 0.033 $m^3\ m^{-3}$) at a very high rank. This verifies our first hypothesis on the ability of PC-PTF to adequately model SWRC using high resolution measured data via HYPROP system. M6 PC-PTF also ranks high among previously developed PC-PTFs considering the reported RMSEs for PC-PTFs in the literature ranging from 0.027 to 0.056 $m^3\ m^{-3}$ (Haghverdi et al., 2012, 2014, 2015; Nguyen et al., 2017). This finding verifies our second hypothesis on the equal or superior performance of PC-PTF developed using HYPROP data compared to the performance of PC-PTF developed using sandbox/pressure plate data. Haghverdi et al. (2014) emphasized on the high sensitivity of PC-PTF, as a data-driven model, to the number and the distribution of measured water retention data points. HYPROP system provides roughly 10 times more measured water retention data per sample compared to the traditional equilibrium approach and our results show that ANN PC-PTF gains a substantial benefit from this ample amount of data.

4.1. Effectiveness of input predictors

The sensitivity analysis revealed a relatively low effect of SOM and SA on SWRC, suggesting that these attributes could be excluded from the modeling. The low effect of SOM could be attributed to its low concentration and its narrow range for a majority of our samples, which were composed of the soils of dry Central Anatolia Region of Turkey. On average, the soil textural components (i.e. percentages of sand, silt and clay) had a medium effect on SWRC but the sensitivity of PC-PTF to these components varied among different versions of M6 (developed using different realizations of the input data). This indicates that the

impact of soil textural components varies based on the types of soils used to train PC-PTFs. IWC and pF constantly showed a very high impact on PC-PTF regardless of the training data characteristics. The high effect of pF was expected because it is the main input attribute used by PC-PTF to model the relationship between the soil matric potential and the water content for each soil sample. PC-PTF showed the highest performance when IWC was included as input attribute (compare M5 with M6 in Table 3). This is in line with the observed trend by Vereecken et al. (2010) in their review paper in which they showed the largest improvements in SWRC estimations by parametric PTFs occurred when the water content information was incorporated as an additional input. According to Schaap and Leij, (1998), including one (θ_{33} kPa) and two (θ_{10} and θ_{33} kPa) water retention points as ANN PTFs input predictors, in addition to the soil textural components and BD, improved the predictability of SWRC by 18% and 26%, respectively. In another study, when a larger dataset was used to develop PTFs, Schaap et al., (1998) observed even higher improvement in the performance of ANN PTFs due to inclusion of one (θ_{33} kPa, 31% more accurate) and two (θ_{10} and θ_{33} kPa, 34% more accurate) water retention points. Twarakavi et al. (2009) reported 31% and 35% improvement in the performance of ANN PTF (i.e. Rosetta software, Schaap et al., 2001) when FC (water content at $h = 330$ cm water column) and FC plus PWP (water content at $h = 15000$ cm water column) were included as additional input predictors, respectively.

Although M5 also had a performance in the typical range for PTFs predicting SWRC, excluding IWC roughly doubled the error. In the previous studies (Haghverdi et al., 2012, 2014) only typical input predictors were considered (i.e. soil textural information, BD and SOM) yet high performance for PC-PTFs was observed. We attribute this to the differences between datasets. In all previous studies, data collected via the sandbox apparatus and/or the pressure plate extractor were utilized. This means PC-PTFs were provided with a less complex dataset consisting of a limited number of water retention points per soil in comparison with a much higher number of water retention points per sample in this study. Including IWC as an extra input attribute provided ANNs a measure for the upper limit of the water content for each soil, which substantially improved the overall predictability of SWRCs. More research is needed to see if this finding holds true for bigger datasets with a larger number of soil samples.

4.2. Performance across textural classes

Haghverdi et al. (2014) showed the performance of the PTFs across textural classes was inversely proportional to the number of available samples for each class in the training dataset. This trend was also observed and reported by Vereecken et al. (2010) who reviewed the accuracy of parametric PTFs from a handful of studies and showed that the percentage share of the textural classes in each dataset can affect the accuracy of PTFs. Our finding (Fig. 8) also confirms this trend. We observed that an acceptable performance over textural classes is expected to occur when a minimum adequate number of samples for each class is incorporated in the PTF develop process. Our results suggest including approximately 13% of data should be enough for each textural class, yet further studies with larger datasets would be needed to confirm this threshold.

4.3. Impact of oven-dry information on PC-PTF estimated SWRCs

When actual measured water retention data by HYPROP were used as a criterion, the overall performance of M6 remained unchanged (MAE: 0.023 $m^3\ m^{-3}$) with and without addition of the oven-dry water retention point. This means the overall performance of PC-PTF was not improved in wet and moist (midrange) regions of the curve by supplementing the arbitrary oven-dry information to the dataset. However, when compared against fitted VG model over a wider range of water retention (i.e. from saturation to PWP) the addition of the over-dry point slightly (7% lower MAE) improved the performance of the PC-PTFs. For most of the soils adding the oven-dry point caused PC-PTF to predict lower water content values moving towards the dry part of the curve (examples illustrated in Fig. 9). Therefore, we recommend including the oven-dry point as an extra data point. Note that ideally more measured data points from the dry range of the curve if available should be incorporated into the dataset to achieve the peak performance of PC-PTF. This is particularly important when the complete estimation of the curve form saturation to oven dryness in desired. The extended evaporation technique cannot generate data on a very dry range of the curve, yet a combination of this technique and dew point method (WP4C instrument, METER Group, Inc., Pullman, WA, USA) can be used to efficiently cover the entire range from saturation to oven dryness.

5. Conclusion

A majority of available pedotransfer functions (PTFs) aiming to predict the soil water retention curve (SWRC) have been developed using water retention data measured by the traditional equilibrium method (i.e. sandbox apparatus and pressure plate extractor). HYPROP system generates water retention data in a dynamic process controlled by evaporation, which is inherently different from the data collected using the traditional method when users set the target pressure and measure the water content as equilibrium is reached. Consequently, HYPROP system typically produces about ten times more water retention data in a much shorter measurement cycle relative to the traditional equilibrium approach. The result of our study showed that artificial neural network (ANN) pseudo continuous PTF (PC-PTF) approach can successfully estimate SWRC using high resolution data collected by HYPROP system. We realized adding an arbitrary oven-dry point results in better estimations in the dry range of the curve for some samples, yet it did not improve the overall performance of PC-PTFs in the wet and the moist regions of the curve. The initial water content (saturated water content) was required as an extra input attribute in addition to the basic soil properties to achieve the highest performance of PC-PTFs. We focused on the continuous estimation of the water content between saturation and the permanent wilting point in this study and the data generated by HYPROP system appeared to be adequate for that range. However, if available, we recommend including more measured data points at the dry range of the curve in the PC-PTF development process particularly when estimation of the full curve from saturation to oven dryness is desired.

References

Bishop, Christopher M., 1995. Neural Networks for Pattern Recognition. Oxford University Press.
Bittelli, M., Solone, R., 2012. Measurement errors of water retention curve using pressure plates: consequences on parameterization EGU General Assembly, 22–27 April 2012, Vienna, Austria., EGUGA. 14.6739B, p.6739.
Blake, G.R., Hartge, K.H., 1986. Bulk density, In: Klute, A. (Ed.), Methods of Soil Analysis. Part 1, second ed. Agron. Monogr., 9. ASA and SSSA, Madison, WI, pp. 363–375.
Burdine, N.T., 1953. Relative permeability calculations from pore size distribution data. Trans. Am. Inst. Mineral, Metall. Pet. Eng. 198, 71–78.
Cresswell, H.P., Green, T.W., McKenzie, N.J., 2008. The adequacy of pressure plate apparatus for determining soil water retention. Soil Sci. Soc. Am. J. 72, 41–49.
De Veaux, R.D., Ungar, L.H., 1994. Multicollinearity: A tale of two nonparametric regressions. In: Selecting Models from Data. Springer, New York, NY, pp. 393–402.
Fields, J.S., Owen, J.S., Zhang, L., Fonteno, W.C., 2016. Use of the evaporative method for determination of soilless substrate moisture characteristic curves. Scientia Horticulturae 211, 102–109.
Gee, G.W., Bauder, J.W., 1986. Particle size analysis, In: Klute, A. (Ed.), Methods of Soil Analysis. Part 1, second ed. Agron. Monogr., 9. ASA and SSSA, Madison, WI, pp. 383–411.
Grath, S.M., Ratej, J., Jovičić, V., Curk, B.Č., 2015. Hydraulic characteristics of alluvial gravels for different particle sizes and pressure heads. Vadose Zone J. 14 (3).
Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442, 46–54.
Haghverdi, A., Leib, B.G., Cornelis, W.M., 2015a. A simple nearest-neighbor technique to

predict the soil water retention curve. Trans. ASABE 58 (3), 697–705.
Haghverdi, A., Leib, B.G., Washington-Allen, R.A., Ayers, P.D., Buschermohle, M.J., 2015b. High-resolution prediction of soil available water content within the crop root zone. J. Hydrol. 530, 167–179.
Haghverdi, A., Öztürk, H.S., Cornelis, W.M., 2014. Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. Geoderma 226, 31–38.
Halbertsma, J., 1996. Wind's evaporation method, determination of the water retention characteristics and unsaturated hydraulic conductivity of soil samples. Possibilities, advantages and disadvantages. In: Durner, W., Halbertsma, J., Cislerova, M. (Eds.), European workshop on advanced methods to determine hydraulic properties of soils, Thurnau, Germany, June 10–12, 1996. Department of Hydrology, University of Bayreuth, pp. 107.
Herbrich, M., Gerke, H.H., 2017. Scales of water retention dynamics observed in eroded Luvisols from an arable postglacial soil landscape. Vadose Zone J. 16 (10).
Jackson, M.L., 1958. Soil Chemical Analysis. Prentice Hall Inc., Eaglewood Cliffs, N.J.
Jain, S.K., Singh, V.P., van Genuchten, M.T., 2004. Analysis of soil water retention data using artificial neural networks. J. Hydrol. Eng. 9 (5), 415–420.
Meter Group Inc., 2018. Operation manual HYPROP. Available online at: http://library.metergroup.com/Manuals/UMS/Hyprop_Manual.pdf.
Minasny, B., McBratney, A., 2002. The neuro-m method for fitting neural network parametric pedotransfer functions. Soil Sci. Soc. Am. J. 66 (2), 352–361.
Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12, 513–522.
Nguyen, P.M., Haghverdi, A., De Pue, J., Botula, Y.D., Le, K.V., Waegeman, W., Cornelis, W.M., 2017. Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils. Biosyst. Eng. 153, 12–27.
Öztürk, H.S., Durner, W., Haghverdi, A., Walter, B., 2013. Comparison of soil moisture retention characteristics obtained by the extended evaporation method and the pressure plate/sand box apparatus. In: EGU General Assembly Conference Abstracts, pp. 3911.
Pertassek, T., Peters, A., Durner, W., 2011. HYPROP Data Evaluation Software, User's Manual. Umwelt Monitoring Systeme GmbH, München, Germany.
Pertassek, T., Peters, A., Durner, W., 2015. HYPROP-FIT Software User's Manual, V.3.0, UMS GmbH, Gmunder Str. 37, 81379 München, Germany, 66 pp.
Peters, A., Durner, W., 2008. Simplified evaporation method for determining soil hydraulic properties. J. Hydrol. 356 (1), 147–162.
Peters, A., Iden, S.C., Durner, W., 2015. Revisiting the simplified evaporation method: Identification of hydraulic functions considering vapor, film and corner flow. J. Hydrol. 527, 531–542.
Schaap, M.G., Leij, F.J., 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. Soil Tillage Res. 47 (1), 37–42.
Schaap, M.G., Leij, F.J., Van Genuchten, M.T., 1998. Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J. 62 (4), 847–855.
Schaap, M.G., Leij, F.J., Van Genuchten, M.T., 2001. ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251 (3), 163–176.
Schelle, H., Heise, L., Jänicke, K., Durner, W., 2013. Water retention characteristics of soils over the whole moisture range: A comparison of laboratory methods. Eur. J. Soil Sci. 64 (6), 814–821.
Schindler, U., 1980. Ein Schnellverfahren zur Messung der Wasserleitfähigkeit imteilgesättigten Boden an Stechzylinderproben. Arch. Acker- Pflanzenbau Bodenkd. 24, 1–7.
Schindler, U., Bohne, K., Sauerbrey, R., 1985. Comparison of different measuring and calculating methods to quantify the hydraulic conductivity of unsaturated soil. J. Plant Nutrition Soil Sci. 148 (6), 607–617.
Schindler, U.G., Müller, L., 2017. Soil hydraulic functions of international soils measured with the Extended Evaporation Method (EEM) and the HYPROP device. Open Data J. Agric. Res. 3 (1).
Schindler, U., Durner, W., von Unold, G., Müller, L., 2010a. Evaporation method for measuring unsaturated hydraulic properties of soils: extending the measurement range. Soil Sci. Soc. Am. J. 74 (4), 1071–1083.
Schindler, U., Durner, W., Von Unold, G., Mueller, L., Wieland, R., 2010b. The evaporation method: Extending the measurement range of soil hydraulic properties using the air-entry pressure of the ceramic cup. J. Plant Nutrition Soil Sci. 173 (4), 563–572.
Schindler, U., Mueller, L., da Veiga, M., Zhang, Y., Schlindwein, S., Hu, C., 2012. Comparison of water-retention functions obtained from the extended evaporation method and the standard methods sand/kaolin boxes and pressure plate extractor. J. Plant Nutrition Soil Sci. 175 (4), 527–534.
Schindler, U., Mueller, L., von Unold, G., Durner, W., Fank, J., 2016. Emerging measurement methods for soil hydrological studies. In: Novel Methods for Monitoring and Managing Land and Water Resources in Siberia. Springer International Publishing, pp. 345–363.
Schwen, A., Zimmermann, M., Bodner, G., 2014. Vertical variations of soil hydraulic properties within two soil profiles and its relevance for soil water simulations. J. Hydrol. 516, 169–181.
Solone, R., Bittelli, M., Tomei, F., Morari, F., 2012. Errors in water retention curves determined with pressure plates: effects on the soil water balance. J. Hydrol. 470, 65–74.
Twarakavi, N.K., Simunek, J., Schaap, M.G., 2009. Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. Soil Sci. Soc. Am. J. 73 (5), 1443.
Van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated. Soil Sci. Soc. Am. J. 43, 892–898.
Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M.G., van Genuchten, M.Th., 2010. Using pedotransfer functions to estimate the van Genuchten-Mualem soil hydraulic properties: a review. Vadose Zone Journal 9, 1–26.
Wendroth, O., Ehlers, W., Kage, H., Hopmans, J.W., Halbertsma, J., Wösten, J.H.M., 1993. Reevaluation of the evaporation method for determining hydraulic functions in unsaturated soils. Soil Sci. Soc. Am. J. 57 (6), 1436–1443.
Wind, G.P., 1968. Capillary conductivity data estimated by a simple method. In: Proceedings of UNESCO/IASH symposium. Water in the unsaturated zone. Wageningen, The Netherlands, pp. 181–191.
Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 251, 123–150.
Zhuang, L., Bezerra Coelho, C.R., Hassanizadeh, S.M., van Genuchten, M.T., 2017. Analysis of the hysteretic hydraulic properties of unsaturated soil. Vadose Zone J. 16 (5).