Received: 13 July 2020 | Accepted: 20 August 2020

DOI: 10.1002/vzj2.20065

Vadose Zone Journal

SPECIAL SECTION

Transdisciplinary Contributions and Opportunities in Soil Physical Hydrology

An empirical soil water retention model based on probability laws for pore-size distribution

Wenjuan Zheng[1] | **Chongyang Shen**[2] | **Lian-Ping Wang**[1,3] | **Yan Jin**[4]

[1] Center for Complex Flows and Soft Matter Research and Dep. of Mechanics and Aerospace Engineering, Southern Univ. of Science and Technology, Shenzhen, Guangdong 518055, China

[2] Dep. of Soil and Water Sciences, China Agricultural Univ., Beijing 100193, China

[3] Dep. of Mechanical Engineering, Univ. of Delaware, Newark, DE 19716, USA

[4] Dep. of Plant and Soil Sciences, Univ. of Delaware, Newark, DE 19711, USA

Correspondence
Chongyang Shen, Dep. of Soil and Water Sciences, China Agricultural Univ., Beijing 100193, China.
Email: chongyang.shen@cau.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 41907008, 41671222; USDA Hatch Funds

ABSTRACT

Knowledge of the soil water retention curve (SWRC) is critical to mathematical modeling of soil water dynamics in the vadose zone. Traditional SWRC models were developed based on bundles of cylindrical capillaries (BCCs) using a residual water content, which fail to accurately describe the dry end of the curve. This study improved and expanded on the traditional BCC models. Specifically, the total water retention was treated as a weighed superposition of capillary and adsorptive components. We proposed a mathematical continuous expression for water retention from saturation to oven dryness, which also allowed for a partition of capillary and adsorptive retention. We further evaluated six capillary retention functions using different probability laws for pore-size distribution—namely, the log-logistic, Weibull, lognormal, two-parameter van Genuchten (VG), three-parameter VG (or Dagum), and Fredlund–Xing (FX) distributions. Model testing against 144 experimental data showed better agreement of the proposed model with experimental observations than the traditional approaches that use the residual water content. The Dagum and FX distributions, which have one more degree of freedom, provided better agreement with experimental data than the other four distributions. The log-logistic and lognormal distributions fitted the experimental data better than the Weibull and VG distribution for loam soils. In addition, the fitted weighting factor w using the log-logistic and lognormal distributions better correlated to soil clay content than the other four distributions. Our study suggests that the log-logistic and lognormal distributions are more suitable to model soils' pore-size distribution than other tested distributions.

1 | INTRODUCTION

Soil water retention characteristics play a crucial role in determining water dynamics in porous media. In vadose zone hydrology, quantification of water movement through partially saturated soils, in the framework of Richards' theory, requires the input of soil hydraulic

Abbreviations: BC, Brooks–Corey; BCC, bundle of cylindrical capillaries; CPF, cumulative probability function; PDF, probability density function; SWRC, soil water retention curve; UNSODA, Unsaturated Soil Hydraulic Database; VG, van Genuchten; FX, Fredlund–Xing.

properties—namely, the soil water retention curve (SWRC) and the hydraulic conductivity function. The SWRC specifies the relationship between soil water content θ or saturation S and soil water potential ψ and can be measured experimentally using traditional methods (e.g., hanging water column and pressure plate apparatus; Klute, 1986) or new devices based on evaporation method (e.g., Hyprop and WP4C; Schindler, Durner, von Unold, & Mueller, 2010). However, experimental measurements are usually laborious and time consuming. Therefore, extensive efforts have been devoted to developing SWRC models, because mathematical equations can be more conveniently incorporated into the Richards' equation than discrete data.

Various approaches have been applied to model SWRC (Assouline & Or, 2013), including (a) expressions based on pore-size distribution or the classical bundle of cylindrical capillaries (BCC), (b) expressions based on particle-size distribution, (c) fractal representation of soil pore space and percolation theory, and (d) pedotransfer function based on easy-to-measure soil physical properties (porosity, the composition of sand, silt, and clay fractions, the content of soil organic matter, etc.). Among these models, BCC-based approaches are the prevailing, such as the widely used van Genuchten (VG; van Genuchten, 1980) and Brooks–Corey (BC; Brooks & Corey, 1964) models.

Although the BCC-type models have been widely used to model water retention by capillary force, they fail to capture the dry end of SWRC where water is held by adsorptive forces (Tokunaga, 2009; Tuller & Or, 2001). The traditional BCC-based SWRC models normally assume a residual water content, θ_r, to account for the residual water that is held by adsorptive forces and remains immobile (Luckner, van Genuchten, & Nielsen, 1989). However, the concept of residual water content contradicts with the general observation that the soil water content eventually reaches 0 at oven dryness (Nimmo, 1991), indicating an inappropriate representation of water retention by noncapillary forces. The contribution of adsorptive water film was later taken into account in various modified SWRC models (Khlosi, Cornelis, Gabriels, & Sin, 2006; Lebeau & Konrad, 2010; Morel-Seytoux & Nimmo, 1999; Peters, 2013; Rossi & Nimmo, 1994; Wang, Ma, & Guan, 2016; Weber, Durner, Streck, & Diamantopoulos, 2019; Zhang, 2011) by applying the empirical expressions developed by Campbell and Shiozawa (1992), Fayer and Simmons (1995), or Fredlund and Xing (1994). However, these modified SWRC models still have various limitations. For example, the models developed by Campbell and Shiozawa (1992), Fayer and Simmons (1995), Khlosi et al. (2006), and Lebeau and Konrad (2010) do not guarantee a zero water content at oven dryness for soil samples with a wide pore-size distribution. The functions proposed by Fredlund and Xing (1994) or Wang et al. (2016) do not allow a partition of capillary and adsorptive components. Peters (2013), Ross and Nimmo (1994), Zhang (2011), and Du (2020) used segmental expressions, which explicitly separate the capillary and adsorptive water retention, but with a discontinuity at a specified suction (i.e., the suction below which the film component is one or the air-entry suction). Weber et al. (2019) overcame the abovementioned drawbacks with a continuous modular framework while also allowing a partition of capillary and adsorptive water retention. The model of Weber et al. (2019) calculated the adsorptive (or noncapillary) retention by integrating the effective saturation function over the pF [$= \log(-\psi)$] space, where ψ is expressed in the unit of centimeters. The integration of common effective saturation functions (e.g., the VG model) does not have analytical solutions and requires numerical calculation. Therefore, the first aim of this study was to propose a simplified analytical expression to describe the SWRC from saturation to oven dryness.

Our second aim was to summarize different capillary water retention functions and incorporate them in the newly proposed SWRC model to evaluate their performance. The capillary water retention functions are based on the BCC concept, which simplifies the soil pore space as BCC tubes with a statistical pore-size distribution. Various probability laws have been used for the probability density function (PDF) $f(r)$, where r is the tube radius, to develop capillary water retention functions in previous studies, including the gamma distribution (Brutsaert, 1966; Diamantopoulos & Durner, 2013; Fredlund & Xing, 1994; Tuller & Or, 2001), lognormal distribution (Brutsaert, 1966; Diamantopoulos & Durner, 2015; Kosugi, 1994, 1996), normal distribution (Laliberte, 1969; Fredlund & Xing, 1994), and Weibull distribution (Assouline, Tessier, & Bruand, 1998). There also existed some capillary water retention functions, where the PDFs were not explicitly specified or were not regarded as a common distribution in the general probability theory. However, the corresponding $f(r)$ can be easily determined by calculating the derivative of the saturation $S(\psi)$ with respect to the matric potential (Fredlund & Xing, 1994). Because these capillary water retention functions were proposed by different researchers, few

Core Ideas

- A novel expression for water retention from saturation to oven dryness was proposed.
- Six capillary retention functions using different probability laws were evaluated.
- Log-logistic and lognormal distributions are suggested to model soil pore-size distribution

TABLE 1 The capillary water retention functions and the corresponding probability density functions (PDFs) and cumulative probability functions (CPFs) of various probability laws

No.	Probability laws	$S(\psi)$	$F(r)$	$f(r)$	$F(\alpha)$
1	log-logistic[a]	$[1+(\psi_0/\psi)^{-n}]^{-1}$	$[1+(r/\alpha)^{-n}]^{-1}$	$\frac{n}{r}\frac{(n/\alpha)^n}{[(r/\alpha)^n+1]^2}$	0.5
2	Weibull[b]	$1-\exp\left[-(\psi_0/\psi)^n\right]$	$1-\exp\left[-(r/\alpha)^n\right]$	$n\alpha^{-n}r^{n-1}\exp\left[-(r/\alpha)^n\right]$	$1-1/e$
3	Lognormal[c]	$\frac{1}{2}\left\{1+\mathrm{erf}\left[\frac{\ln(\psi_0/\psi)}{\sqrt{2}n}\right]\right\}$	$\frac{1}{2}\left\{1+\mathrm{erf}\left[\frac{\ln(r/\alpha)}{\sqrt{2}n}\right]\right\}$	$\frac{1}{rn\sqrt{2\pi}}\exp\left[-\frac{\ln^2(r/\alpha)}{2n^2}\right]$	0.5
4	VG[d]	$\left[1+(\psi_0/\psi)^{-n}\right]^{1/n-1}$	$\left[1+(r/\alpha)^{-n}\right]^{1/n-1}$	$\frac{n-1}{r}\frac{(n/\alpha)^n}{[(r/\alpha)^n+1]^{(2-1/n)}}$	$2^{1/n}-1$
5	Dagum[d]	$\left[1+(\psi_0/\psi)^{-n}\right]^{-m}$	$\left[1+(r/\alpha)^{-n}\right]^{-m}$	$\frac{nm}{r}\frac{(n/\alpha)^{nm}}{[(r/\alpha)^n+1]^{m+1}}$	$1/2^m$
6	FX[e]	$\ln\left[e+(\psi_0/\psi)^{-n}\right]^{-m}$	$\ln[e+(r/\alpha)^{-n}]^{-m}$	$\frac{mn\alpha^n\ln[e+(r/\alpha)^{-n}]^{-m-1}}{\left[e+(r/\alpha)^{-n}\right]r^{n+1}}$	$1/\ln^m(e+1)$
7	Gamma[f]	$1-\exp(\frac{\psi_0}{\psi})\sum_{n=0}^{n-1}\frac{(\psi_0/\psi)^n}{n!}$	$1-\exp(-\frac{r}{\alpha})\sum_{n=0}^{n-1}\frac{(r/\alpha)^n}{n!}$	$\frac{\alpha^n}{\Gamma(n)}r^{n-1}\exp(-r/\alpha)$	
8	Power[g]	$(\psi_0/\psi)^n$	$(r/\alpha)^n$	$n\alpha(r/\alpha)^{n-1}$	1

Note. S, saturation; ψ, matric potential; r, radius of the pore; α, characteristic pore size; ψ_0, onset matric potential to drain the pores with the characteristic pore size α; n and m, shape factors; Γ, gamma function.

[a] Assouline et al. (1998).

[b] Brutsaert (1966).

[c] Kosugi (1996)

[d] van Genuchtenn (1980).

[e] Fredlund and Xing (1994).

[f] Tuller and Or (2001).

[g] Brooks and Corey (1964).

attempts have been made to compare their performance using the same sets of experimentally measured SWRCs. To bridge this research gap, this study compared six capillary water retention functions that are based on different probability laws—namely, the log-logistic, Weibull, lognormal, two-parameter VG model of van Genuchten (1980), three-parameter VG (or Dagum), and Fredlund and Xing (1994) (denoted as FX) distributions by testing them against the same datasets. Table 1 summarized these capillary water retention functions.

The rest of the manuscript is organized as follows. The proposed model to consider adsorptive water retention is described in the Section 2, and the application of each distribution law in modeling the capillary water retention in literature is summarized. Then, the capability of the proposed models based on different probability laws to predict SWRC is evaluated using 144 SWRCs obtained from Rudiyanto, Minasny, Shah, Setiawan, and van Genuchten (2020).

2 | MODEL DESCRIPTION

Water is retained in soils by capillary and adsorptive forces; therefore, comprehensive understanding soil water retention requires two components accounting for both mechanisms. Traditionally, models were developed by neglecting the adsorptive component so capillary water retention could be readily calculated by using the function $S_{cap}(\psi)$ as listed in Table 1. Then, water content is related to $S_{cap}(\psi)$ by

$$\theta = \theta_s S_{cap}(\psi) \tag{1}$$

or

$$\theta = (\theta_s - \theta_r)\, S_{cap}(\psi) + \theta_r \tag{2}$$

where θ_s and θ_r are saturated and residual water contents. While θ_s is available from the Unsaturated Soil Hydraulic Database (UNSODA) database, θ_r is defined as the water content beyond which water is assumed to be immobile. Despite the convenience of the assumption in applications, it is widely accepted that residual water exists in the form of water film adsorbed on the particle surface and is mobile. When the adsorptive water retention is considered, the water content can be written as (Weber et al., 2019)

$$\theta = \theta_s w \Gamma_{cap}(\psi) + \theta_s(1-w)\Gamma_{film}(\psi) \tag{3}$$

where w is a weighting factor and $w = 1$ indicates no film water retention, and $\Gamma_{cap}(\psi)$ and $\Gamma_{film}(\psi)$ are the effective saturation functions due to capillary and adsorptive water retention. $\Gamma_{cap}(\psi)$ is given by

$$\Gamma_{cap}(\psi) = \left[S_{cap}(\psi) - S_0\right] / (1 - S_0) \quad (4)$$

where $S_{cap}(\psi)$ is the capillary water saturation and $\Gamma_0 = \Gamma(\psi_0)$ is the saturation at oven dryness, which corresponds to a matric potential of $\psi_0 = -6.3 \times 10^6$ cm (or pF = 6.8) according to experimental observation (Schneider & Goss, 2012). The rescaling approach is used to ensure a zero water content at oven dryness (Iden & Durner, 2014). Below, we present a simplified expression $\Gamma_{film}(\psi)$ for adsorptive water retention and summarize different capillary water retention functions $S_{cap}(\psi)$ and incorporate them in the newly proposed SWRC model (i.e., Equation 3) to evaluate their performance.

2.1 | Capillary retention function $S_{cap}(\psi)$

Various expressions of $S_{cap}(\psi)$ have been used to describe the capillary water retention based on the BCC concept. These expressions simplify the soil pore space as BCC tubes. At the single pore scale, the capillary tube assumption provides a determinant connection between the size of the pore and the capillary pressure (the capillary pressure equals the negative value of the matric potential ψ if the adsorptive pressure is neglected) according to the Young–Laplace law, which specifies the capillary pressure and the radius r of capillary tube as

$$\psi = -\frac{2\sigma \cos\beta}{r} = -\frac{C}{r} \quad (5)$$

where β is the contact angle at the air–water–solid triple contacting point, σ is the surface tension of soil fluid, and both β and σ are assumed to be independent of ψ; thus, the numerator in Equation 5 is usually treated as a constant for a particular soil sample.

At the sample scale, the saturation (i.e., the volumetric water content θ normalized by the saturated volumetric water content θ_s and/or the residual water content θ_r) can be expressed by further upscaling the relationship (Equation 5) using a continuous statistical pore-size distribution function. Given the volumetric PDF of a bundle of capillary tube $f(r)$, the saturation S_{cap} is simply the cumulative probability function (CPF) $F(r)$:

$$S_{cap} = \frac{\int_{r_{min}}^{r} f(r)\mathrm{d}r}{\int_{r_{min}}^{r_{max}} f(r)\mathrm{d}r} = \frac{F(r) - F(r_{min})}{F(r_{max}) - F(r_{min})} = F(r) \quad (6)$$

The expression of $S_{cap}(\psi)$ can be determined by replacing r by C/ψ in the function of $F(r)$, where $C = 0.114$ N m^{-1} assuming the surface tension of water $\sigma = 0.072$ N m^{-1} and $\theta = 0$.

Table 1 summarizes the functions of $S_{cap}(\psi)$, the corresponding CPF $F(r)$, and PDF $f(r) = \mathrm{d}F(r)/\mathrm{d}r$. The application of each model is reviewed and their common features are discussed below

1. *The three-parameter VG (or the Dagum) function*. The three-parameter VG model was originally proposed by van Genutchten (1980) to numerically describe the SWRC. In the general probability theory, the corresponding distribution was given the name after Camilo Dagum, who was the first to apply this distribution law to study personal income (Dagum, 1977). The Dagum distribution has been extensively used in various fields such as income and wealth data, meteorological data, and reliability and survival analysis. The Dagum model provides good fit to SWRC of many soils, particularly for data near saturation (van Genuchten & Nielsen, 1985), and thus became very popular in soil science applications. It is one of the most widely used models in vadose zone research, as it has the largest flexibility covering a large range of soil textures. It is currently implemented in several vadose zone models (e.g., HYDRUS and SWAP). In addition, the VG model is the most popular model in the development of pedotransfer function since the 1990s (Van Looy et al., 2017).
2. *The two-parameter VG function*. The two-parameter VG model, where the parameter m is substituted by $1 - 1/n$, is more commonly used in practice than the three-parameter VG model. The purpose of fixing $m = 1 - 1/n$ was to obtain a closed-form equation for relative hydraulic conductivity (van Genuchten, 1980). Although reducing one degree of freedom provides convenience for deriving the analytical form of hydraulic conductivity function, the flexibility for fitting the SWRC is reduced. The two-parameter VG model has been widely used for calculating capillary water retention in the modified models that incorporated adsorptive water retention (Peters, 2013; Rudiyanto et al., 2015; Zhang, 2011). To distinguish between the two VG models, we use "Dagun model" instead of the three-parameter VG model, whereas the VG model specifically refers to the two-parameter VG model.
3. *The log-logistic function*. The log-logistic model is a special case of the Dagum model with a fixed m value (i.e., $m = 1$). It was first proposed to describe the SWRC by Brutsaert (1966). The log-logistic distribution has since been successfully applied to describe SWRC (Haverkamp, Vauclin, Touma, Wierenga, & Vachaud, 1977). However, little attention has been paid to this

model, since the VG model was proposed and widely adopted.

4. *The Weibull function.* A Weibull-type distribution of pore sizes was reported by Assouline et al. (1998), which correlated pore-size distribution to particle-size distribution using a uniform random fragmentation process. It was shown that the Weibull distribution exhibited increased flexibility and better reproduced the measured data both in the high and the low water content ranges compared with the VG model.
5. *The lognormal function.* Application of the lognormal distribution to model SWRC was discussed in detail by Kosugi (1994, 1996), which showed that the lognormal distribution performed similarly to the VG model. The lognormal distribution was further modified to incorporate the adsorptive water film by Lebeau and Konrad (2010) and Peters (2013). Diamantopoulos and Durner (2015) applied the lognormal distribution to a capillary bundle model with a triangular geometry. One of the advantages of using the lognormal distribution is that it allows a closed-form expression for hydraulic conductivity function.
6. *The FX function.* Fredlund and Xing (1994) developed an empirical SWRC model, which became widely used recently (Wang, Ma, Guan, & Zhu, 2017; Weber et al., 2019). The FX model allows a slower drop in the low water content range than the VG model, and thus produces a better fitting to experimental data. The FX model contains three parameters: two shape factors (n and m) and one parameter ψ_0 representing a suction value higher than the air-entry value as shown in the graphic analysis in Fredlund and Xing (1994). Although air-entry value corresponds to the critical matric potential where the saturation starts to drop, the value of ψ_0 is more negative than the critical matric potential due to the reason that $S(\psi = \psi_0)$ is always <1 as long as m is positive.
7. *The gamma function.* The possibility of using the gamma distribution to model pore-size distribution was discussed in Brutsaert (1966) and Fredlund and Xing (1994). However, the application of this distribution in developing BCC-based SWRC models did not happen until more recently, when the gamma distribution was incorporated into capillary models considering an angular cross-section (Diamantopoulos & Durner, 2013; Tuller & Or, 2001; Zheng, Yu, & Jin, 2013). It was found that the gamma distribution was only suitable for sandy soils (Diamantopoulos & Durner, 2015), because one of the model parameters n is limited to integers and sometimes treated as a fixed value thus reducing the model's flexibility.
8. *The power function.* The power law was proposed by Brooks and Corey (1964) (referred to as the BC model) and is also one of the most widely used models in vadose zone hydrology. Despite its popularity, the BC model contains a discontinuity in the air-entry pressure, which can cause numerical difficulties in derivatives of SWRC.

The review of the pore-size distribution functions shows that the Dagum and FX distributions have three parameters, whereas other distributions have two. In general, one of the parameters, α, is related to a characteristic length, whereas the other parameters n in all the two-parameter distributions and m in the Dagum and FX distributions are related to the shape of the distribution curve. The value of $F(r)$ at $r = \alpha$ is listed in the last column of Table 1 in order to show the location of $r = \alpha$ in the cumulative distribution curve. For the power law, $F(\alpha) = 1$, and thus α can be defined as the maximum pore size that corresponds to the air entry matric potential. The values of $F(\alpha)$ for the other models are <1, meaning that α represents a characteristic value that is less than the maximum pore size. For the VG, Dagum, and FX distributions, the values of $F(\alpha)$ depend on the shape parameter n or m. The distributions, except the power law, are continuous for the non-negative variable (i.e., the size of the pore, r, and the value of the CPF ranges within [0,1]). We tested six continuous distribution laws and their corresponding capillary water retention functions in this study. We did not consider the gamma distribution, because the parameter n is limited to integers and thus is less flexible than other models (Diamantopoulos & Durner, 2015), nor the power law, as discontinuity in the SWRC derivative introduces difficulties when numerically solving the Richards' equation.

2.2 | Adsorptive water retention function S_{film}

The adsorptive water content decreases nearly linearly toward 0 on the semi-log scale in the dry end of the SWRC (Campbell & Shiozawa, 1992; Schneider & Goss, 2012). Peters (2013) used a piecewise function to describe the adsorptive water retention:

$$\Gamma_{\text{film}}(\psi) = \begin{cases} \left(1 - \frac{\ln(2)}{\ln(1+\psi_0/\psi_r)}\right)^{-1} \left(1 - \frac{\ln(1+\psi_0/\psi_r)}{\ln(1+\psi_0/\psi_r)}\right) & \psi > \psi_r \\ 1 & \psi \le \psi_r \end{cases} \quad (7)$$

where ψ_r and ψ_0 are the matric potential values corresponding to the residual water content and a water content value of 0, respectively. The piecewise function is not continuously differentiable at ψ_r. Weber et al. (2019) proposed

a new function to alleviate the problem of discontinuity,

$$\Gamma_{\text{film}}(\psi) = 1 - \frac{\log_{10}(e)\int_{\psi}^{\psi_r}\frac{S_{\text{cap}}(h)-1}{h}dh}{\log_{10}(e)\int_{\psi_0}^{\psi_r}\frac{S_{\text{cap}}(h)-1}{h}dh} \quad (8)$$

In this study, we adopted the following simple function to describe the adsorptive water retention,

$$\Gamma_{\text{film}}(\psi) = 1 - \frac{\ln(1+c\psi/\psi_r)}{\ln(1+c\psi_0/\psi_r)} \quad (9)$$

where c is a constant that allows S_{film} to be 0 when $\psi = \psi_0$ and approaches 1 when $\psi = \psi_r$. Wang et al. (2016) has shown that model performance did not change sensitively to the value of c; therefore, we adopted the value of 0.01 as presented in Wang et al. (2016). We propose to incorporate Equation 8 into Equation 3 to calculate the total water retention:

$$\theta = \theta_s w\left[\frac{S_{\text{cap}}(\psi) - S_0}{1 - S_0}\right] + \theta_s(1-w)\left[1 - \frac{\ln(1+c\psi/\psi_r)}{\ln(1+c\psi_0/\psi_r)}\right] \quad (10)$$

The new model allows a partition of capillary and film components and is continuous over saturation to oven dryness. It does not introduce any additional parameters compared with Equations 2 or 3, incorporating Equation 6 by Peters (2013) or Equation 7 by Weber et al. (2019).

3 | MODEL TESTING

Measured SWRCs were used to analyze and test the proposed model (Equation 3) combined with each of the six empirical capillary water retention functions listed in Table 1. A total of 144 soil samples were obtained from UNSODA and Rudiyanto et al. (2020). The soil samples were selected with data available near saturation to, when possible, near dryness ($\psi < -10^4$ cm). The analyzed soil samples cover a wide range of soil textures, including sand, loamy sand, sandy loam, loam, silt loam, silt, sandy clay loam, clay loam, silt clay loam, silty clay, and clay, as shown in the marked soil triangle (Figure 1).

The models were fitted to the experimental data using the nonlinear curve fitting routine *lsqcurvefit* in Matlab. There are two adjustable model parameters with the log-logistic, Weibull, lognormal, and VG distribution functions and three for the Dagum and FX distribution functions assuming $\theta_r = 0$ (i.e., Equation 1). With Equations 2 or 3, the residual water content θ_r or the weighting factor w was treated as an additional adjustable model parameter,

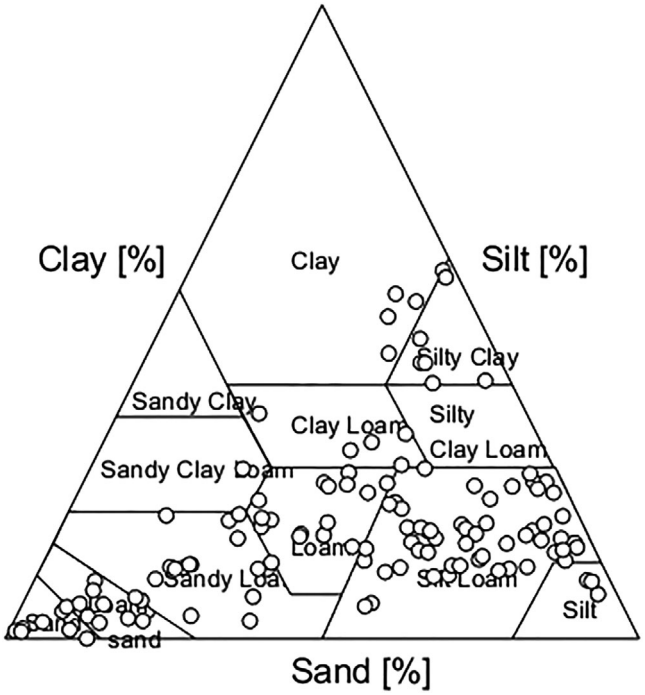

FIGURE 1 The soil triangle marking the tested soil samples by circles

respectively. The performance of different SWRC models is assessed by the RMSE. The RMSE is an indicator for the overall error of the evaluated model, which is defined as

$$\text{RMSE}_\theta = \sqrt{\sum_{i=1}^{N}\left(\theta_{m,i} - \theta_{f,i}\right)^2 \Big/ N} \quad (11)$$

where N is the number of data points for each soil sample, and $\theta_{m,i}$ and $\theta_{f,i}$ are the measured and fitted soil water content of the ith data pair for each soil. A smaller RMSE value indicates a better fit to the experimental data. The measurement error for water retention data (water content) is ~0.01 cm^3cm^{-3}; therefore, an acceptable RMSE value should be close to 0.01 m^3cm^{-3} (Peters, 2013). The uniqueness of the fitting was confirmed by rerunning the program with different initial parameter estimates.

4 | RESULTS AND DISCUSSION

4.1 | Illustrative examples

To evaluate the performance of the models, we chose six representative soil samples, including a sand, sandy loam, loam, silt loam, clay loam, and clay sample. Figure 2 shows the modeled SWRCs using the newly developed model (Equation 10), where curves with different

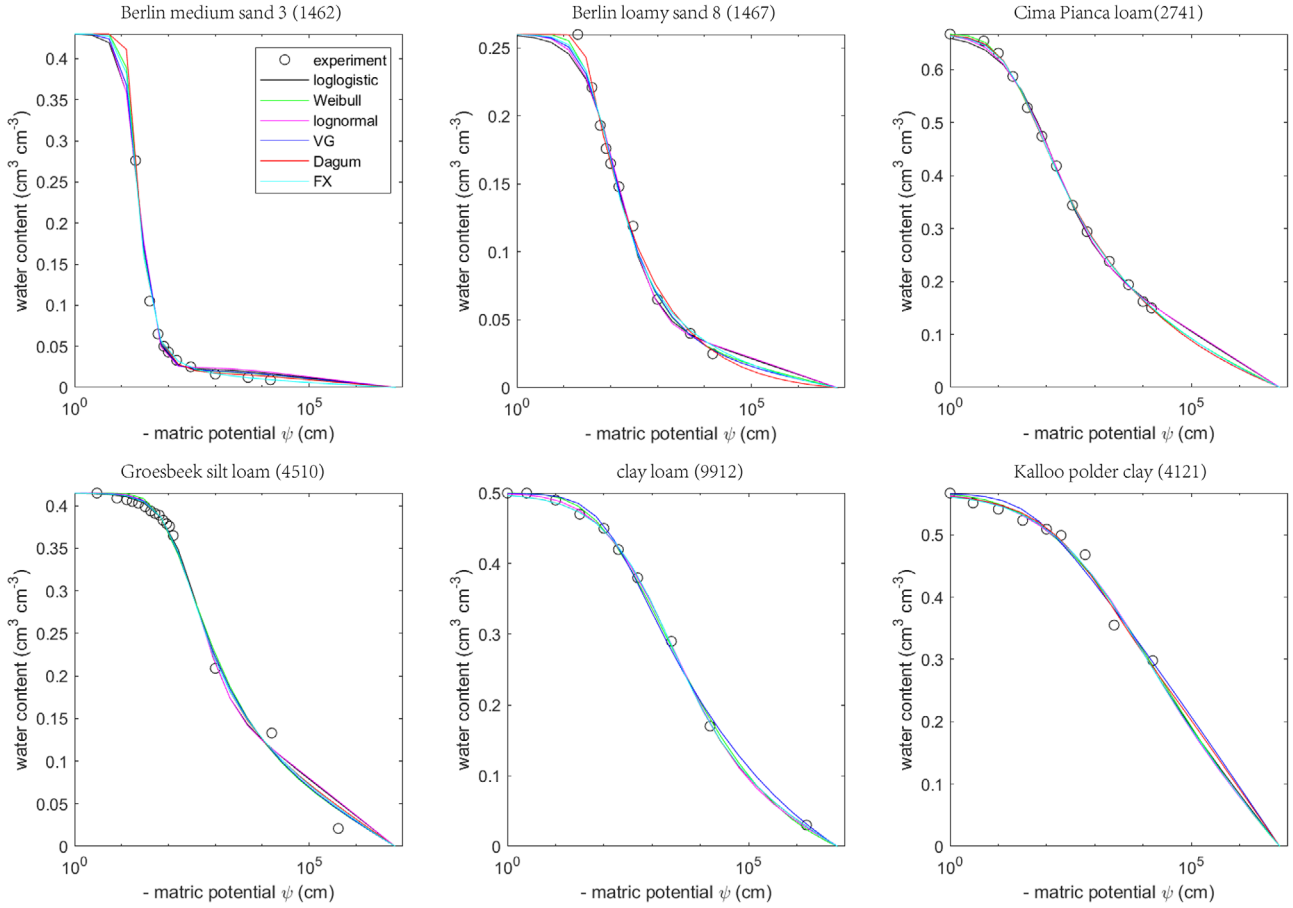

FIGURE 2 The experimental and modeled soil water retention curves for the six representative soil samples using Equation 10. VG, van Genuchten; FX, Fredlund–Xing

colors represent modeling results with different probability laws (i.e., the log-logistic, Weibull, lognormal, VG, Dagum, and FX distributions). The differences in the six probability laws are mainly found in two regions where water content initially started to drop and in the low water content region, while all curves coincided at the intermediate water contents. The two three-parameter (i.e., the Dagum and FX) distributions better captured the wet end than the four two-parameter distributions (i.e., the log-logistic, Weibull, lognormal, and VG). The two-parameter distributions tended to underestimate in the high-water-content region for coarse-textured soil samples (e.g., Berlin medium sand 8 and Cima Pianca loam) while overestimating for the finer-textured samples (e.g., the clay loam and Kalloo polder clay). The lognormal and log-logistic distributions predicted earlier air entry and lower water contents at the dry end than the other distributions. At the low water content beyond the measurement range (e.g., $\psi < -10^5$ cm), there was a substantial difference in the shape of SWRCs among the six distributions, particularly for fine-textured soil samples (e.g., Kalloo polder clay).

The PDF $f(r)$ for each soil was plotted in Figure 3 using the fitted model parameters in Table 2. The range of pore sizes was narrower for the coarser-textured soil samples than for finer-textured soil samples. The pore sizes of Berlin medium sand 3 ranged from 10^{-4} to 10^{-2} cm, whereas the pore sizes of other loam and clayey soils cover more than four orders of magnitudes. The pore-size range generated by the various distributions was close for coarse sand and varied significantly for loam and clayey soils. The pore size corresponding to the peak in the density distribution curve was greater for the sand than that for the finer-textured soil samples. The lognormal and log-logistic distributions led to symmetric shapes of the density distribution curves on the log-scale, whereas the other distributions exhibited shorter tails than the log-logistic and lognormal distributions. The peak was the highest using the Weibull distribution for all the soil samples, except that the Dagum distribution predicted the highest peak for the Berlin medium sand.

Unfortunately, independent and reliable measurements of pore-size distribution may not be possible via experiments. Therefore, the use of models for prediction is quite important for revealing the pore-size distribution. Above results showed that different probability laws have different performances for soils with different soil texture.

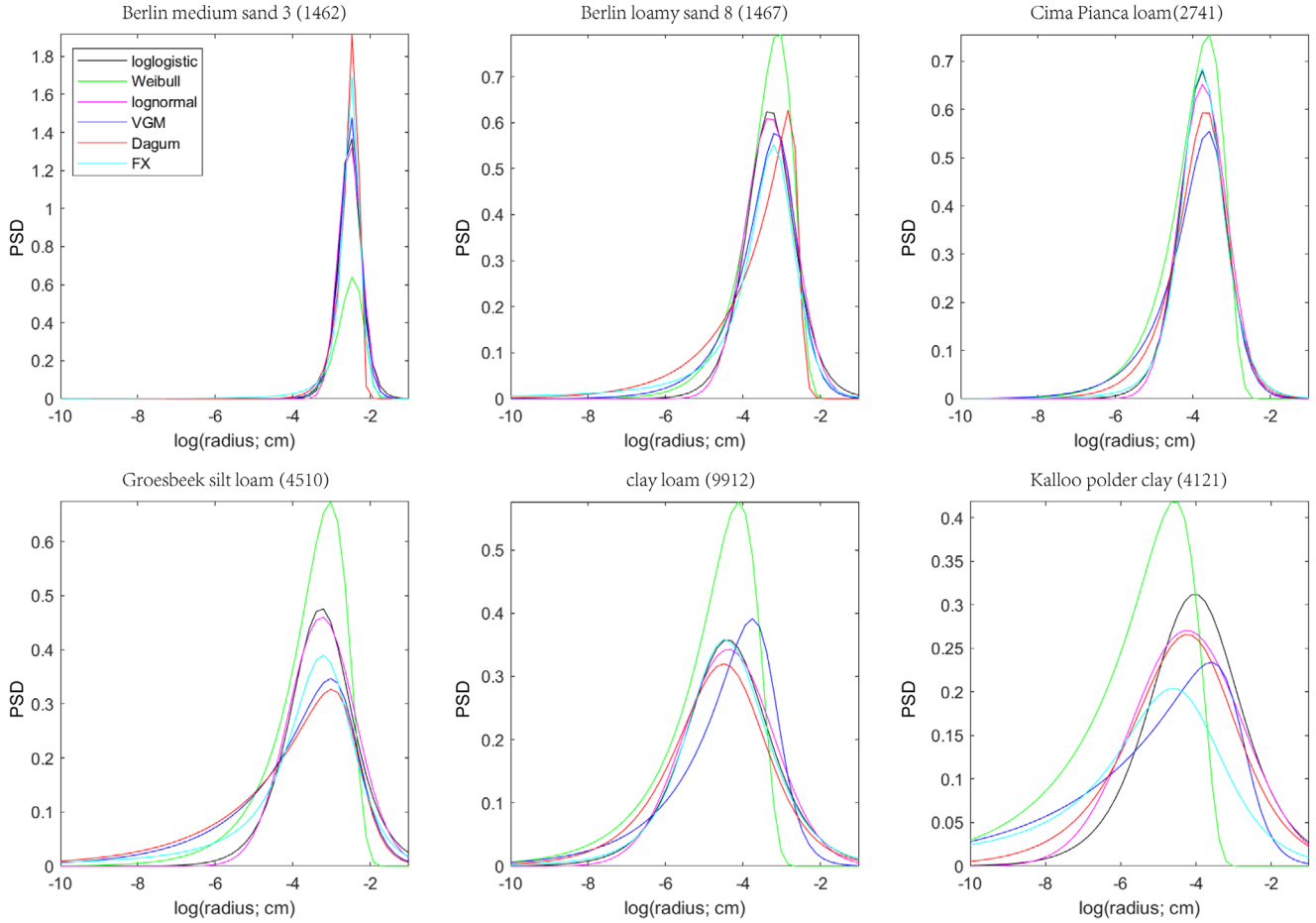

FIGURE 3 The modeled pore-size distribution (PSD) curves for the six representative soil samples using the fitted model parameters. VG, van Genuchten; FX, Fredlund–Xing

Consequently, it is critical to selecting an appropriate function for simulating the pore-size distribution and water dynamics, as well as contaminant transport in soil and sediment environments.

4.2 | Model parameters

The values of model parameters for the six representative soil samples are listed in Table 2 to show their magnitude. In general, as the soil texture becomes finer, more negative ψ_0 values and decreased n and w values are found for all the probability laws, except that n values increases for the lognormal distribution. As stated earlier, w is the weighting factor, and $w = 1$ indicates the dominance of capillary water retention and no adsorptive water retention. The predicted w values indicate a greater contribution of adsorptive water retention for the finer-textured soil samples than for the coarse-textured sample. We further tested the correlation between the weighting factor w and clay content using the data of sand, loamy sand, sandy loam, sandy clay loam, and clay. It appears that w values decreased with the clay contents for the log-logistic, lognormal, and Weibull distributions (Figure 4). The R^2 values for the log-logistic and lognormal distributions were around .62, which was greater than for the Weibull distribution. For the VG, Dagum, and FX distributions, the w values did not correlate to the clay content. The analysis based on the 144 soil samples suggested a general empirical function to predict the weight factor w using the soil clay content when applying the log-logistic or lognormal distribution:

$$w = 0.81 - 1.28 \times [\text{clay}(\%)] \quad (12)$$

Previous results of water sorption isotherm experiments showed a positive correlation between water retention and clay content (Schneider & Goss, 2012), which agrees with results using the log-logistic and lognormal distributions presented in this study. Under dry conditions, the surfaces of soil particles are covered with several molecular layers of water, and thus the water adsorption relies on the specific surface area (Likos & Lu, 2002). The adsorptive water content is correlated to the total percentage of clay, attributing to the fact that the specific surface area is dominated by

TABLE 2 Fitted parameter values in the six capillary water retention functions for the six representative soil samples

	Soil code					
Parameter[a]	**1462**	**1467**	**2741**	**4510**	**9912**	**4121**
Soil texture	sand	loamy sand	loam	loam	clay loam	clay
Saturated water content θ_s	0.315	0.270	0.397	0.386	0.449	0.557
Composition						
Sand, %	96.4	87.8	46.8	12.7	–	2.0
Silt, %	2.0	7.2	40.4	43.8	–	41.0
Clay, %	1.6	5.0	12.8	43.5	–	57.0
log-logistic model						
$-\psi_0$	24.15	145.33	134.02	419.26	1,739.20	771.86
n	2.43	1.09	0.83	1.18	0.63	0.55
w	0.95	0.84	0.70	0.63	0.72	0.40
Weibull model						
$-\psi_0$	19.64	89.75	78.04	301.36	1,901.40	9,972.50
n	1.93	0.70	0.47	0.58	0.29	0.19
w	0.95	0.89	0.78	0.72	1.00	0.82
log-normal model						
$-\psi_0$	24.10	144.16	129.05	401.26	1,744.11	1,092.69
n	0.68	1.50	2.00	1.41	2.68	3.31
w	0.94	0.84	0.69	0.63	0.72	0.43
VG model						
$-\psi_0$	19.54	50.21	21.58	141.23	116.28	37.94
n	2.98	1.55	1.28	1.48	1.14	0.99
w	0.95	0.91	0.86	0.73	1.00	0.98
Dagum						
$-\psi_0$	14.85	26.12	20.73	191.21	1,604.40	5,0352.00
n	8.24	5.92	1.31	1.36	0.63	0.53
m	0.19	0.06	0.21	0.47	0.95	9.92
w	0.96	1.00	0.86	0.68	0.72	0.31
FX						
$-\psi_0$	18.53	45.43	19.42	160.08	1,097.80	259.06
n	3.59	1.65	1.15	1.34	0.62	0.53
m	1.36	0.54	0.10	0.72	1.80	0.85
w	0.98	1.00	0.98	0.78	0.77	0.53

[a]VG, van Genuchten; FX, Fredlund–Xing; ψ_0, onset matric potential to drain the pores with the characteristic pore size α; n and m, shape factors; w, weighting factor.

the clay content in most soils. However, natural soils are typically composed of a mixture of 1:2 and 1:1 clay types and the water adsorption behavior of the 2:1 clay minerals differs considerably from that of the 1:1 clays; therefore, the data points fail to strictly follow the fitting line as shown in Figure 4. In addition, Schneider and Goss (2012) reported that the clay surface is not the dominant sorption site for soils with a clay content <7%, which could be the reason causing the derivation of data points at low clay contents.

The results comparing the correlation between w and clay content using different probability laws imply that SWRC models based on the log-logistic and lognormal distributions provide more physically meaningful values of w than other distributions.

The parameter ψ_0 can be physically interpreted as the onset matric potential to drain the pores with the characteristic pore size α. Less negative ψ_0 values imply smaller characteristic pore sizes in fine-textured soil samples than those in coarse-textured samples. The definition of α varies among different distribution functions (e.g., α in the log-logistic and lognormal distributions represents the median pore size, whereas it corresponds to the 63rd percentile in the cumulative distribution curve for the Weibull distribution). Therefore, the predicted ψ_0 values using the

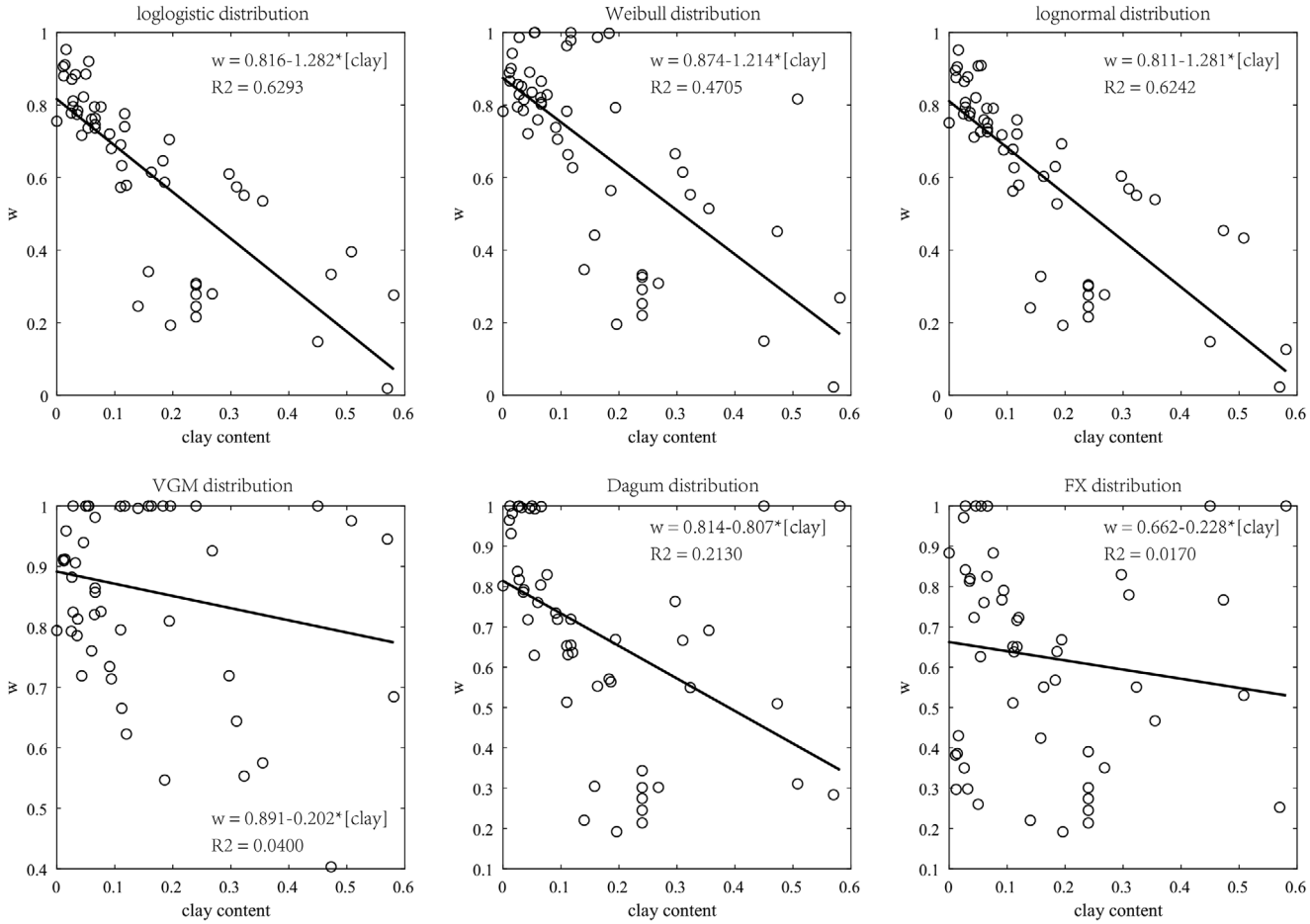

FIGURE 4 The weighting factor w vs. clay content [clay] calculated based on various probability laws. VGM, van Genuchten; FX, Fredlund–Xing

Weibull distribution are smaller than those using the log-logistic and lognormal distributions. For the VG, Dagum, and FX distributions, α corresponds to a percentile of the distribution that depends on the value of the shape factor n or m, and the value of $F(\alpha)$ decreases as n or m increases.

4.3 | Comparing model performance

The values of RMSE for each model and for each soil textural group (sand, loam, and clay) are shown in Figure 5. In order to assess model performance at the low water content end, we compared results based on Equations 1, 2, and 10, which represent the physical conditions corresponding to purely capillary water retention, capillary water retention with a residual water content, and water retention by capillary and adsorptive forces, respectively. Figure 5 shows that the RMSE values are consistently greater with the assumption of $\theta_r = 0$ (Equation 1) than those calculated including θ_r as an adjustable parameter (Equation 2) or considering adsorptive water film (Equation 10) at the dry end. This result indicates that considering θ_r or adsorptive water retention in the BCC-based models is essential (Assouline & Or, 2013), particularly for coarse-textured soils (Peters, 2013). Note that Equations 2 and 10 provide better fit to experimental data because they contain one more adjustable parameter than Equation 1.

The RMSE values obtained based on Equation 10 are slightly smaller or comparable with those based on Equation 2, where the numbers of adjustable parameters are the same. Although including θ_r as a model parameter satisfies the fitting propose for SWRC, the physical explanation of θ_r remains problematic (Assouline & Or, 2013). On average, 15% of the tested soils having a fitted θ_r value > 0.2 cm^3 cm^{-3}. An assumption that flow ceases below such a large θ contradicts experimental observations (Nimmo, 1991).

We then compared the performance of the six distribution laws considering adsorptive water retention (Equation 10), which provided the best fit to experimental data compared with Equations 1 and 2. The FX and Dagum distributions consistently fitted the experimental SWRCs better than the other distributions: the log-logstic, Weibull, lognormal, and VG model, which have one less adjustable parameter than the FX and Dagum model. The difference

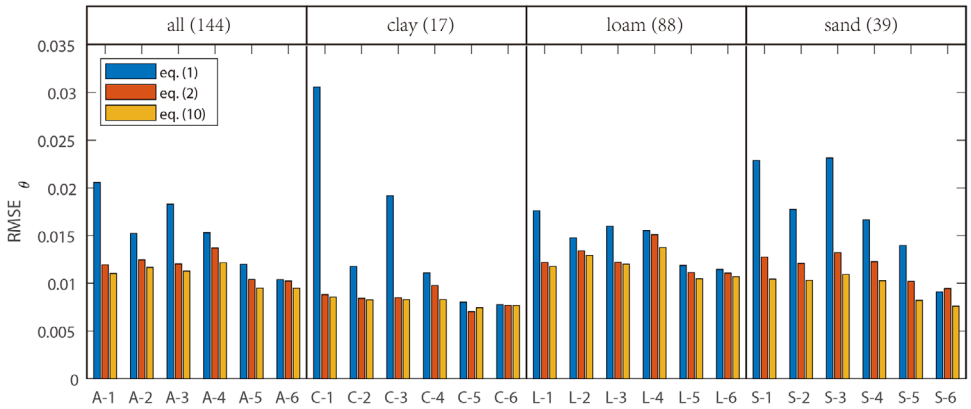

FIGURE 5 The RMSE values calculated from 144 soil water retention curves (SWRCs) for all soil samples (A) and each soil textural group (clay [C], loam [L], sand [S]) as listed in the four columns. The numbers 1–6 in the axis represent the SWRC models based on the log-logistic, Weibull, lognormal, van Genuchten (VG), Dagum, and Fredlund–Xing (FX) distribution, respectively

between the FX and Dagum models, when the adsorptive water retention is considered, is statistically insignificant. The difference among the log-logistic, Weibull, lognormal, and VG distributions varied for different soil textural groups. The log-logistic and lognormal distributions perform similarly to each other, but better than the Weibull and VG distributions for loam soils. However, the difference among the four distributions was not statistically significant for sand and clayey soils.

The RMSE values calculated based on the FX model were around 0.01 cm^3 cm^{-3}, which was close to the assumed measurement error. The difference in the calculated RMSE values from the FX model combing Equations 1, 2, or 10 was not statistically significant for all soil samples. The Dagum model, which has the same number of adjustable parameters as the FX model, cannot predict the SWRCs with $\theta_r = 0$ well. This observation implies that the original FX model is more flexible than the Dagum to model the SWRC of soil samples with all textures. The superior performance of the original FX model that fits to the measured water retention curve even neglecting film adsorption was attributed to its ability to allow a slower drop at the dry end (Fredlund & Xing, 1994). However, the correlation was weak between the weighting factor of film water retention and soil clay content when considering film adsorption in the FX model, as shown in Figure 4.

In summary, the flexibility of the SWRC model in matching the experiential data largely depends on the degree of freedom, whereas the distribution functions with the same number of adjustable parameters cause little difference in the RMSE values. However, the SWRC model based on the log-logistic and lognormal distributions provide the fitted weighting factor values that are better correlated with the clay content than the other four probability laws tested.

5 | CONCLUSION

We proposed an empirical modeling framework for simulating the entire SWRC from saturation to oven dryness. In this framework, total water retention is treated as a weighed superposition of capillary and adsorptive components. Although the capillary component is described using the BCC models, the adsorptive water retention decreases linearly on a semi-log scale. The model formulation is simple and mathematically continuous. Compared with the traditional models that use the residual water content θ_r, this model provides better fits over the whole range of water contents and for all soil types while keeping the same number of parameters as the traditional models.

We reviewed six capillary retention functions that are based on different probability laws for pore-size distribution—namely, the log-logistic, Weibull, lognormal, VG, Dagum, and FX distribution. We found that at least two parameters are essential to effectively describe the pore-size distribution (i.e., a characteristic length and a shape factor), and adding an additional shape factor increases the flexibility of the model (e.g., the Dagum and FX model). The capability of each model to fit experimental data was verified using 144 observed datasets obtained from the UNSODA. The FX and Dagum distributions have one more degree of freedom when fitting to the experimental data; thus, the RMSE values are smaller than those of the other four models. The log-logistic and lognormal distributions can fit slightly better to the experimental data than the Weibull and VG model for loam soils, but the differences among the four models are insignificant for sand and clayey soils. Comparison of the six probability laws also shows that the fitted weighting factor w using the log-logistic and lognormal distributions correlates well to soil

clay content, which agrees with the experimental results from water isotherms.

Based on the capillary bundle concept, the model oversimplifies the pore geometry and thus does not take into account the processes associated with pore connectivity and pore continuity. However, they have been widely used in the soil physics community and have proven useful in practical hydrological applications. We believe that our new model is an improvement to the traditional models by considering adsorptive water retention using a simple and continuous function. Our study also suggests that the log-logistic and lognormal distributions are more suitable to model soils' pore-size distribution than other tested models. Future effort is devoted to predict the unsaturated hydraulic conductivity based on the current framework. The SWRC expressions can be incorporated into the conductivity models of Burdine (1953) or Mualem (1976). The convenience of deriving the unsaturated hydraulic conductivity function from the SWRC expressions should be considered when comparing different pore-size distribution functions, in addition to current evaluation criteria for their preformation on describing the SWRC.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We acknowledge funding by the National Natural Science Foundation of China (No. 41907008 and 41671222) and by the USDA Hatch Funds.

ORCID

Wenjuan Zheng https://orcid.org/0000-0001-9204-1296

REFERENCES

Assouline, S., Tessier, D., & Bruand, A. (1998). A conceptual model of the soil water retention curve. *Water Resources Research*, *34*, 223–231. https://doi.org/10.1029/97WR03039

Assouline, S., & Or, D. (2013). Conceptual and parametric representation of soil hydraulic properties: A review. *Vadose Zone Journal*, *12*(4). https://doi.org/10.2136/vzj2013.07.0121

Brooks, R. H., & Corey, A. T. (1964). *Hydraulic properties of porous media*. Fort Collins: Colorado State University.

Brutsaert, W. (1966). Probability laws for pore-size distributions. *Soil Science*, *101*, 85–92.

Burdine, N. T. (1953). Relative permeability calculations from pore size distribution data. *Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers*, *198*, 71–78.

Campbell, G. S., & Shiozawa, S. (1992). Predicion of hydraulic properies of soils using paricle-size distribuion and bulk density data. In M. Th. van Genuchten & F. J. Leij (Eds.), *Indirect methods for estimating the hydraulic properties of unsaturated soils: International Workshop on Indirect Methods for Esimaing the Hydraulic Properies of Unsaturated Soils* (pp. 317–328). Berkeley: University of California Press.

Dagum, C. (1977). A new model for personal income distribution: specification and estimation. *Economic Applique*, *30*, 413–437.

Diamantopoulos, E., & Durner, W. (2013). Physically-based model of soil hydraulic properties accounting for variable contact angle and its effect on hysteresis. *Advances in Water Resources*, *59*, 169–180. https://doi.org/10.1016/j.advwatres.2013.06.005

Diamantopoulos, E., & Durner, W. (2015). Closed-form model for hydraulic properties based on angular pores with lognormal size distribution. *Vadose Zone Journal*, *14*(2). https://doi.org/10.2136/vzj2014.07.0096

Du, C. Y. (2020). A novel segmental model to describe the complete soil water retention curve from saturation to oven dryness. *Journal of Hydrology*, *584*. https://doi.org/10.1016/j.jhydrol.2020.124649

Fayer, M. J., & Simmons, C. S. (1995). Modified soil-water retention functions for all matric suctions. *Water Resources Research*, *31*, 1233–1238. https://doi.org/10.1029/95WR00173

Fredlund, D. G., & Xing, A. Q. (1994). Equations for the soil-water characteristic curve. *Canadian Geotechnical Journal*, *31*, 521–532.

Haverkamp, R., Vauclin, M., Touma, J., Wierenga, P. J., & Vachaud, G. (1977). A comparison of numerical simulation models for one-dimensional infiltration. *Soil Science Society of America Journal*, *41*, 285–294. https://doi.org/10.2136/sssaj1977.03615995004100020024x

Iden, S. C., & Durner, W. (2014). Comment on "Simple consistent models for water retention and hydraulic conductivity in the complete moisture range" by A. Peters, *Water Resources Research*, *50*, 7530–7534. https://doi.org/10.1002/2014WR015937

Khlosi, M., Cornelis, W. M., Gabriels, D., & Sin, G. (2006). Simple modification to describe the soil water retention curve between saturation and oven dryness. *Water Resources Research*, *42*(11). https://doi.org/10.1029/2005WR004699

Klute, A. (1986). Water retention: Laboratory methods. In A. Klute (Ed.), *Methods of soil analysis. Part 1. Physical and mineralogical methods* (2nd ed., pp. 635–660). Madison, WI: ASA and SSSA. https://doi.org/10.2136/sssabookser5.1.2ed.c26

Kosugi, K. (1994). Three-parameter lognormal distribution model for soil-water retention. *Water Resources Research*, *30*, 891–901. https://doi.org/10.1029/93WR02931

Kosugi, K. (1996). Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resources Research*, *32*, 2697–2703. https://doi.org/10.1029/96WR01776

Laliberte, G. E. (1969). A mathematical function for describing capillary pressure-desaturation data. *Hydrological Sciences Journal*, *14*, 131–149. https://doi.org/10.1080/02626666909493724

Lebeau, M., & Konrad, J. M. (2010). A new capillary and thin film flow model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *46*(12). https://doi.org/10.1029/2010WR009092

Likos, W., & Lu, N. (2002). Water vapor sorption behavior of smectite–kaolinite mixtures. *Clays and Clay Minerals*, *50*, 553–561. https://doi.org/10.1346/000986002320679297

Luckner, L., van Genuchten, M. Th., & Nielsen, D. R. (1989). A consistent set of parametric models for the two-phase flow of immiscible fluids in the subsurface. *Water Resources Research*, *25*, 2187–2193. https://doi.org/10.1029/WR025i010p02187

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*, 513–521. https://doi.org/10.1029/WR012i003p00513

Morel-Seytoux, H. J., & Nimmo, J. R. (1999). Soil water retention and maximum capillary drive from saturation to oven dryness. *Water Resources Research*, *35*, 2031–2041. https://doi.org/10.1029/1999WR900121

Nimmo, J. R. (1991). Comment on the treatment of residual water content in "A consistent set of parametric models for the two-phase low of immiscible luids in the subsurface" by L. Luckner et al. *Water Resources Research*, *27*, 661–662. https://doi.org/10.1029/91WR00165

Peters, A. (2013). Simple consistent models for water retention and hydraulic conductivity in the complete moisture range. *Water Resources Research*, *49*, 6765–6780. https://doi.org/10.1002/wrcr.20548

Rossi, C., & Nimmo, J. R. (1994). Modeling of soil-water retention from saturation to oven dryness. *Water Resources Research*, *30*, 701–708. https://doi.org/10.1029/93WR03238

Rudiyanto, Minasny, B., Shah, R. M., Setiawan, B. I., & van Genuchten, M. Th. (2020). Simple functions for describing soil water retention and the unsaturated hydraulic conductivity from saturation to complete dryness. *Journal of Hydrology*, *588*. https://doi.org/10.1016/j.jhydrol.2020.125041

Rudiyanto, Sakai, M., van Genuchten, M. Th., Alazba, A. A., Setiawan, B. I., & Minasny, B. (2015). A complete soil hydraulic model accounting for capillary and adsorptive water retention, capillary and film conductivity, and hysteresis. *Water Resources Research*, *51*, 8757–8772. https://doi.org/10.1002/2015WR017703

Schindler, U., Durner, W., von Unold, G., & Mueller, L. (2010). Evaporation method for measuring unsaturated hydraulic properties of soils: Extending the measurement range. *Soil Science Society of America Journal*. *74*, 1071–1083. https://doi.org/10.2136/sssaj2008.0358

Schneider, M., & Goss, K.-U. (2012). Prediction of water retention curves for dry soils from an established pedotransfer function: Evaluation of the Webb model, *Water Resources Research*, *48*(6). https://doi.org/10.1029/2011WR011049

Tokunaga, T. K. (2009). Hydraulic properties of adsorbed water films in unsaturated porous media. *Water Resources Research*, *45*(6).

Tuller, M., & Or, D. (2001). Hydraulic conductivity of variably saturated porous media: Film and corner flow in angular pore space. *Water Resources Research*, *37*, 1257–1276. https://doi.org/10.1029/2000WR900328

van Genuchten, M. Th. (1980). A closed-form equation for prediction the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*, 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

van Genuchten, M. Th., & Nielsen, D. R. (1985). On describing and predicting the hydraulic properties of unsaturated soils. *Annales Geophysicae*, *3*, 615–628,

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., … Vereecken, H. (2017). Pedotransfer functions in Earth system science: Challenges and perspectives. *Reviews of Geophysics*, *55*, 1199–1256. https://doi.org/10.1002/2017RG000581

Wang, Y., Ma, J., & Guan, H. (2016). A mathematically continuous model for describing the hydraulic properties of unsaturated porous media over the entire range of matric suctions. *Journal of Hydrology*, *541*, 873–888. https://doi.org/10.1016/j.jhydrol.2016.07.046

Wang, Y., Ma, J., Guan, H., & Zhu, G. (2017). Determination of the saturated film conductivity to improve the EMFX model in describing the soil hydraulic properties over the entire moisture range. *Journal of Hydrology*, *549*, 38–49. https://doi.org/10.1016/j.jhydrol.2017.03.063

Weber, T. K. D., Durner, W., Streck, T., & Diamantopoulos, E. (2019). A modular framework for modeling unsaturated soil hydraulic properties over the full moisture range. *Water Resources Research*, *55*, 4994–5011. https://doi.org/10.1029/2018WR024584

Zhang, Z. F. (2011). Soil water retention and relative permeability for conditions from oven-dry to full saturation. *Vadose Zone Journal*, *10*, 1299–1308. https://doi.org/10.2136/vzj2011.0019

Zheng, W. J., Yu, X., & Jin, Y. (2015). Considering surface roughness effects in a triangular pore space model for unsaturated hydraulic conductivity. *Vadose Zone Journal*, *14*(7). https://doi.org/10.2136/vzj2014.09.0121

How to cite this article: Zheng W, Shen C, Wang L-P, Jin Y. An empirical soil water retention model based on probability laws for pore-size distribution. *Vadose Zone J*. 2020;19:e20065. https://doi.org/10.1002/vzj2.20065