Scaling Water Retention Curves for Soils with Lognormal Pore-Size Distribution

K. Kosugi and J. W. Hopmans*

ABSTRACT

The scaling theory approach has been widely used as an effective method to describe the variation of soil hydraulic properties. In conventional scaling, reference retention curves and scaling factors are determined from minimization of residuals. Most previous studies have shown that scaling factors are lognormally distributed. In this study, we derived physically based scaling factors, assuming that soils are characterized by a lognormal pore-size distribution function. The theory was tested for three sets of retention data. Two data sets included samples of a sandy loam soil, and one set included samples of a loamy sand soil. Individual soil water retention data were fitted to the retention model proposed by Kosugi (1996). The parameters of the model are the mean and variance of the log-transformed pore-radius distribution. Scaling factors and parameters of the reference curve were computed directly from the parameters of individual soil water retention functions. Assuming that (i) the soil pore radius of a study area is lognormally distributed and (ii) soil samples are obtained from random sampling of effective soil pore volume from the study area, we have proposed a theoretical interpretation of the lognormal scaling factor distribution. Scaling results for all three data sets compared well with those obtained using the conventional scaling method.

Describing water flow in soils requires knowledge of the soil hydraulic properties. The hydraulic properties of unsaturated soil are represented by the water retention characteristic (the relationship between the volumetric soil water content, \( \theta \), and the capillary pressure head, \( h \), and the unsaturated hydraulic conductivity, \( K \), function. Both properties are variable in heterogeneous soils, or in field plots that are apparently homogeneous. The concept of similar media was introduced by Miller and Miller (1956) to develop scaling theory for the analysis of such variations in field soils. Scaling provides a means to relate hydraulic properties of different soils to those of a reference soil using scaling factors. In recent years, the scaling theory approach has been widely used as an effective method to describe the variation of soil hydraulic properties.

Previous studies proposed different methods to determine scaling factors for soil hydraulic properties. While Russo and Bresler (1980) suggested the scaling method to compute reference hydraulic properties directly from observed hydraulic data, most previous studies employed functional models for the reference retention and conductivity curves. Warrick et al. (1977) adopted polynomial functions to express reference \( h(\theta) \) and \( K(\theta) \) curves. Simmons et al. (1979) used a logarithmic function for the scaling of the water retention characteristic. The combined soil water retention–hydraulic conductivity model proposed by Brooks and Corey (1964) was used by Ahuja and Williams (1991) in the scaling of \( \theta(h) \) and \( K(h) \) relationships. Clausnitzer et al. (1992) proposed a method of simultaneous scaling of \( h(\theta) \) and \( K(\theta) \) curves by employing the combined soil water retention–hydraulic conductivity model proposed by van Genuchten (1980). Moreover, Zhang et al. (1993) showed scaling results for four different soil water retention models. Most studies showed that scaling factors are lognormally distributed (e.g., Warrick et al., 1977; Hopmans, 1987; Clausnitzer et al., 1992; Zhang et al., 1993).

Abbreviations: C, conventional; PB, physically based; PDF, probability density function; REV, representative elementary volume; RSS, residual sum of squares.
On the basis of the lognormal distribution of scaling factors, stochastic models have been proposed to analyze the effects of variable soil hydraulic properties on saturated and unsaturated soil water flow. By employing Monte Carlo simulation and using the scaling factor as a lognormally distributed random variable, Clapp et al. (1983) concluded that heterogeneity in soil hydraulic properties may account for approximately 75% of the observed standard deviation of field water content. Ahuja et al. (1984) examined infiltration phenomena by using lognormally distributed scaling factors for the saturated hydraulic conductivity. Assuming that the scaling factor distribution function is lognormal, the statistical properties of soil water regime (Hopmans and Stricker, 1989; Kim et al., 1997) and solute transport (van Ommen et al., 1989; Bresler and Dagan, 1979) for large soil domains (agricultural field, watershed) were analyzed. Most recently, Nielsen et al. (1998) concluded that scaling opportunities to describe field soil water behavior continue to appear both promising and provocative.

Thus, many studies have demonstrated the potential of scaling to describe variability of soil hydraulic properties. However, these studies have not emphasized the statistical significance of the scaling factor distribution. No theoretical interpretation has been proposed for the apparent lognormal distribution of scaling factors. This is most likely so because most previous studies used empirical curve-fitting equations for soil hydraulic properties, which do not address the physical significance of their parameters. The objectives of this study were to present a physically based method of scaling soil water retention curves using the physically based retention model introduced by Kosugi (1996) and to propose a theoretical interpretation for scaling factor distributions. The Kosugi (1996) model assumes the soil pore radii to be lognormally distributed. Consequently, the parameters of the retention model have physical significance and are directly related to the statistical properties of the soil pore-size distribution. The lognormal soil pore-size distribution has been assumed in some previous studies. Based on the fact that many soils show a lognormal particle-size distribution, Brutsaert (1966) proposed the lognormal distribution to describe pore-size distribution. Gardner (1956) introduced the possibility of characterizing soil structure using a lognormal pore-size distribution, assuming a relationship between aggregate size and pore size. Most recently, Nimmo (1997) proposed such a relationship and subsequently derived a model to describe the soil structural influence of soil water retention using a lognormal aggregate-size distribution model. Pachepsky et al. (1995) derived the fractal dimension of soil pores assuming a lognormal pore-size distribution.

**THEORY**

After establishing the functional form of the soil water retention model for soils with a lognormal pore-size distribution, we show how statistical theorems provide physically based parameters for the soil water retention function of a study field. Based on the similar media concept, the distribution of physically based scaling factors is derived that, combined with the reference retention curve, characterizes the spatial variability of soil water retention data.

**Lognormal Distribution Model for Soil Pore-Size Distribution and Water Retention Curve**

The probability density function (PDF) of soil pore radius $r, p(r)$, is defined as (Brutsaert, 1966)

$$p(r) = \frac{dS_e}{dr}$$

where $S_e$ is the effective saturation

$$S_e = (\theta - \theta_t)/(\theta_s - \theta_t)$$

described by $\theta_t$ and $\theta_s$, denoting the saturated and residual volumetric water content ($L^3 L^{-3}$), respectively, and the dimension of $p(r)$ is $L^{-1}$. In Eq. [1], $p(r) dr = dS_e$, represents the volume of pores of radius $r \rightarrow r + dr$ per unit effective pore volume of soil. The effective pore volume is defined as the product of the total soil volume and the effective porosity, $(\theta_s - \theta_t)$ of the soil. Integrating Eq. [1] yields the cumulative pore-radius distribution function:

$$S_e(r) = \int_0^r p(r) dr$$

Consequently, $S_e(r)$ represents the effective saturation of the soil when the pore volume fraction occupied by pores with a radius equal to or smaller than $r$ is filled with water. Here, $r$ is associated with the capillary pressure head, $h$ (L), $(h > 0$ for unsaturated soil) by the capillary pressure function

$$h = 2\gamma \cos \beta / \rho g r$$

where $\gamma$ is the interfacial tension, $\beta$ is the contact angle, $\rho$ is the density of wetting fluid, and $g$ is the acceleration of gravity. The value of $A$ for air–water–soil systems is 0.149 cm$^2$. On the basis of the direct correspondence of $r$ and $h$, the function $S_e(r)$ is transformed into $S_e(h)$, which is the soil water retention curve.

The PDF of the natural logarithm of pore radius $(\ln r)$ is expressed as

$$f(\ln r) = \frac{dS_e}{d\ln r}$$

Consequently, $f(\ln r) d \ln r = dS_e$ represents the volume of full pores of log-transformed radius $\ln r \rightarrow \ln r + d \ln r$ per unit effective pore volume of medium. Similarly, the PDF of the natural logarithm of pore capillary pressure head $(\ln h)$ can be defined as

$$g(\ln h) = \frac{dS_e}{d\ln h}$$

Hence, $g(\ln h) d \ln h = dS_e$ represents the volume of full pores in which water is retained by log-transformed capillary pressure $\ln h \rightarrow \ln h + d \ln h$ per unit effective pore volume of medium. The relationship between $p(r), f(\ln r)$, and $g(\ln h)$ is expressed by

$$p(r) dr = f(\ln r) d \ln r = g(\ln h) d \ln h = dS_e$$

An analytical expression for the effective saturation, $S_e$, is derived by integrating $f(\ln r)$, or $g(\ln h)$:

$$S_e(\ln r) = \int_{\ln r}^{\ln r_{max}} f(\ln r) d \ln r$$

$$S_e(\ln h) = -\int_{\ln h_{min}}^{\ln h} g(\ln h) d \ln h$$

Kosugi (1996) applied a lognormal distribution law to the PDF of the soil pore radius to derive a combined soil water
where $\ln r_m$ and $\sigma^2$ are the mean and variance of $\ln r$, respectively. The parameter $m_r$ is the geometric mean pore radius and is equal to the median pore radius [$S_e(\ln r_m) = 0.5$] based on the assumption of the lognormal distribution. Figure 1a shows a graphical representation of Eq. [10]. The corresponding expression for $g(\ln h)$ is

$$g(\ln h) = \frac{1}{\sqrt{2\pi\sigma_h}} \exp\left[-\frac{(\ln h - \ln h_m)^2}{2\sigma^2_h}\right]$$

Equation [11] indicates that $\ln h$ obeys the normal distribution $N(\ln h_m, \sigma_h^2)$. The mean of $g(\ln h)$, $h_m$, is related to the mean of $f(\ln r)$, $\ln r_m$ by $\ln h_m = \ln A - \ln r_m$, whereas the variance of $g(\ln h)$, $\sigma^2_h$, is identical to the variance of $f(\ln r)$. The distribution described by Eq. [11] is shown in Fig. 1b. Substituting Eq. [10] and [11] into Eq. [8] and [9], respectively, and integrating yields expressions for the cumulative distribution functions $S_e(\ln r)$ and $S_e(\ln h)$:

$$S_e(\ln r) = \frac{1}{2} \left[ 1 + \text{erfc}\left(\frac{\ln r - \ln r_m}{\sqrt{2\sigma^2_r}}\right) \right]$$

$$S_e(\ln h) = \frac{1}{2} \left[ 1 + \text{erfc}\left(\frac{\ln h - \ln h_m}{\sqrt{2\sigma^2_h}}\right) \right]$$

where $\text{erfc}$ denotes the complementary error function.

**Statistical Characterization of Soil Sampling**

We consider that soil samples are collected at different locations within a study area. We treat the total effective soil pore volume in the study area as the population, whereas each soil sample is obtained from random sampling of effective soil pore volume from the population. Based on previous studies (Kosugi, 1996 and 1997), the PDF of $\ln r$ for each soil sample is expected to describe the normal distribution:

$$f_i(\ln r) = \frac{1}{\sqrt{2\pi\sigma_r}} \exp\left[-\frac{(\ln r - \ln r_m)^2}{2\sigma^2_r}\right]$$

where $\ln r_m$ and $\sigma_r$ denote the mean and variance of $\ln r$, respectively, for soil sample $i (i = 1, 2, \ldots, I; I$ is total number of soil samples). Soil sample size is assumed to be constant (i.e., each soil sample has the same effective pore volume). It is assumed that the PDF of $\ln r$ for the population (i.e., the study area) is represented by the normal distribution:

$$f(\ln r) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(\ln r - \ln r_m)^2}{2\sigma^2}\right]$$

where $\ln r_m$ and $\sigma$ denotes the mean and variance of $\ln r$, respectively, for the population.

In most soil sampling procedures, the sample size is small relative to the population size (i.e., the effective pore volume of a soil sample is small relative to that of the study field). In such cases, the sample mean is normally distributed with a mean value equal to the population mean (Spiegel, 1975) only if the samples are drawn from a normally distributed...
population. Consequently, under the conditions assumed in this study, the sample mean, $\ln r_{m}$, is normally distributed with its mean equal to the population mean, $\ln r_{m}$, or

$$\ln r_{m} = \text{Mean}(\ln r_{m}) = \frac{1}{l} \sum_{i=1}^{l} \ln r_{m, i} \quad [18]$$

The population variance, $\sigma_{i}^{2}$, can be written as the sum of the mean of the sample variances, $\text{Mean}(\sigma_{i}^{2}) = (1/l) \sum_{i} \sigma_{i}^{2}$, and the variance of the sample means, $\text{Var}(\ln r_{m})$ (see Appendix A for its derivation), or

$$\sigma_{i}^{2} = \text{Mean}(\sigma_{i}^{2}) + \text{Var}(\ln r_{m}) \quad [19]$$

As the volume of a soil sample approaches the representative elementary volume (REV) of the study area, the average of $\ln r$, $\ln r_{m}$, for all samples will coincide, since geometrical aspects are preserved as the centroid of the REV is moved from one place to another (Baveye and Sposito, 1984). In this situation, $\text{Var}(\ln r_{m})$ will be zero and the value of $\sigma_{i}^{2}$ will be equal to $\text{Mean}(\sigma_{i}^{2})$.

**Physically Based Scaling Approach**

Based on the similar media concept, the microscopic structures of soils are assumed to be identical (Warrick et al., 1977; Kutilek and Nielsen, 1994), and the soils differ only by their microscopic length scale, which is characterized by a scaling factor. The scaling factor, $\alpha_{i}$, is defined as the ratio of a microscopic characteristic length, $\lambda_{i}$, of sample $i$ to the characteristic length, $\bar{\lambda}$, of a reference soil:

$$\alpha_{i} = \lambda_{i}/\bar{\lambda} \quad [20]$$

Defining the pore radius $r$ as the microscopic characteristic length, Eq. [20] becomes

$$\ln \alpha_{i} = \ln r - \ln \bar{r} \quad [21]$$

where $r$ and $\bar{r}$ are the largest water-filled pore radii at a specific volumetric water content for sample $i$ and the reference soil, respectively. Based on the similar media concept, each soil sample has identical saturated and residual water content values. Hence, Eq. [21] holds for all $r$ and $\bar{r}$ at equal effective saturation values.

As an example, Fig. 2 shows the PDF $f(\ln r)$ for a soil sample $i$, $f(\ln r_{m})$, and for a reference soil, $f(\ln r)$. Since the similar media theory assumes that each sample is characterized by a single scale factor only, $\sigma_{i}^{2}$ values for each sample $i$ and the variance of $\ln r$ for the reference soil, $\sigma^{2}$, are assumed identical. Moreover, $\ln r_{m}$ and $\ln r_{m}$ represent the mean of the PDF $f(\ln r)$ and $f(\ln r_{m})$, respectively. Since $r_{m}$ is the pore radius for which the effective saturation $S_{e}$ is equal to 0.5, $r_{m}$ is a suitable representative pore radius to characterize soil water retention curves. Hence, we define $r_{m}$ as the microscopic characteristic length scale rather than an arbitrary pore radius. Consequently, scaling factors are defined by

$$\alpha_{i} = r_{m}/r_{m} \quad \text{or} \quad \ln \alpha_{i} = \ln r_{m} - \ln r_{m} \quad [22]$$

A graphical representation of Eq. [22] is given in Fig. 2, where it is clearly shown how $\ln \alpha_{i}$ can be inferred from the mean values of the PDF for the sample $i$ and the reference soil. Since the sample mean, $\ln r_{m}$, is normally distributed under the conditions assumed in this study, Eq. [22] also shows that $\ln \alpha_{i}$ is normally distributed. That is, the scaling factor obeys a lognormal distribution. From Eq. [22] the variance of $\ln \alpha_{i}$, $\text{Var}(\ln \alpha_{i})$, is equal to the variance of $\ln r_{m}$, $\text{Var}(\ln r_{m})$, so that Eq. [19] becomes

$$\sigma_{i}^{2} = \text{Mean}(\sigma^{2}) + \text{Var}(\ln \alpha_{i}) \quad [23]$$

Thus, the variance of log-transformed scaling factors is defined as the difference of the total variance of $\ln r$ in the study area, $\sigma_{r}^{2}$, and the expected variance of $\ln r$ within soil samples, $\text{Mean}(\sigma^{2})$.

Since by definition the scaling factor for the reference soil is equal to one (Eq. [20]), earlier scaling studies assumed the arithmetic average of scaling factors equal to unity (Peck et al., 1977):

$$\text{Mean}(\alpha_{i}) \equiv \frac{1}{l} \sum_{i=1}^{l} \alpha_{i} = 1 \quad [24]$$

This constraint appears suitable for data with normally distributed scaling factors. However, since scaling factors have been found to be lognormally distributed, we use the constraint that the geometric mean of scaling factors is unity:

$$\left(\prod_{i=1}^{l} \alpha_{i}\right)^{1/l} = 1 \quad \text{so that} \quad \text{Mean}(\ln \alpha_{i}) = \frac{1}{l} \sum_{i=1}^{l} \ln \alpha_{i} = 0 \quad [25]$$

After substitution of Eq. [22] into Eq. [25], it can be shown that $\ln r_{m}$ is computed from

$$\ln \bar{r}_{m} = \ln r_{m} = \text{Mean}(\ln r_{m}) = \frac{1}{l} \sum_{i=1}^{l} \ln r_{m, i} \quad [26]$$

Using the scaling theory approach as outlined above, the PDF of $\ln r$ for the reference soil, $f(\ln r)$, is defined by

$$f(\ln r) = \frac{1}{\sqrt{2\pi \sigma_{r}^{2}}} \exp \left[ -\frac{(\ln r - \ln r_{m})^{2}}{2\sigma_{r}^{2}} \right] \quad [27]$$

where $\ln r_{m}$ is equal to $\text{Mean}(\ln r_{m})$, and $\sigma_{r}^{2}$ is computed as the mean of the sample variances, $\text{Mean}(\sigma^{2})$ (Fig. 2).

**Scaling of Soil Water Retention Curves**

Based on the pore-radius distribution expressed by Eq. [16] and [17], the respective functional expressions for the soil water retention curves, representative for the individual soil samples and the study area, are obtained:

$$S_{e, i} = F_{h} \left( \frac{\ln h_{m, i} - \ln h}{\sigma_{i}} \right) \quad [28]$$

$$S_{e, m} = F_{h} \left( \frac{\ln h_{m, m} - \ln h}{\sigma_{m}} \right) \quad [29]$$

where $\ln h_{m, i} = \ln A - \ln r_{m, i}$ and $\ln h_{m, m} = \ln A - \ln r_{m}$. The parameters $\ln h_{m, i}$ and $\ln h_{m, m}$ represent the mean $\ln h$ values for sample $i$ and the study area, respectively. Based on Eq. [27], the water retention function for the reference soil is defined by

---

![Fig. 2: Schematic representation of the probability density function (PDF) of the lognormal distribution of the pore radius ($\ln r$) for soil sample $i$, $f(\ln r_{m})$, and for reference soil, $f(\ln r)$.](image-url)
where \( \ln \hat{h}_m = (\ln A - \ln r_m) \) represents the mean \( \ln h \) value for the reference soil. Since both \( \ln r_m \) and \( \ln r_m' \) are equal to the mean of \( \ln r_{nj} \) (Eq. [26]), \( \ln \hat{h}_m \) is computed from

\[
\ln \hat{h}_m = \ln h_{mr} = \text{Mean}(\ln h_{nj}) = \frac{1}{I} \sum_{i=1}^{I} \ln h_{nj,i} \tag{31}
\]

The variance for the reference soil, \( \sigma^2 \), is computed as the mean of the sample variances, \( \text{Mean}(\sigma_i^2) \). Moreover, scaling factors defined by Eq. [22] can be computed directly from

\[
a_i = \frac{\hat{h}_m}{h_{nj,i}} \quad \text{or} \quad \ln a_i = \ln \hat{h}_m - \ln h_{nj,i} \tag{32}
\]

The mean in \( h \) value, \( \ln h_{nj,i} \) for sample \( i \) is related to the mean \( \ln r \) value, \( \ln r_{nj} \), by \( \ln h_{nj,i} = A - \ln r_{nj} \). Since \( \ln r_{nj} \) is normally distributed under the assumed conditions, \( \ln h_{nj,i} \) is also normally distributed. Analogous to Eq. [19], the total variance of \( \ln h \) for the study area, \( \sigma^2 \), can be written in terms of the expected variance of \( \ln h \) within soil samples, \( \text{Mean}(\sigma_i^2) \), and the variance of mean \( \ln h \) values between soil samples, \( \text{Var}(\ln h_{nj}) \): 

\[
\sigma^2 = \text{Mean}(\sigma_i^2) + \text{Var}(\ln h_{nj}) \tag{33}
\]

**MATERIALS AND METHODS**

The proposed method adopts the assumption that each soil sample is obtained from random sampling of effective soil pore volume from a study area comprising the population. This assumption is likely to be valid if soil samples are collected from a study area with no apparent abrupt changes in soil type. Three sets of laboratory-measured soil water retention data that appeared to satisfy this requirement were subjected to the proposed scaling approach. For each data set, parameters \( h_{nj,i} \) and \( \sigma_i \) of the soil water retention model in Eq. [28] were optimized to fit the soil water retention data using the nonlinear least squares optimization procedure that is provided in the spreadsheet program Excel (Wraith and Or, 1998). Values of \( h_{nj,i} \) and \( \sigma_i \) were determined uniquely for every data set, even if various initial parameter estimates were used.

The first two data sets (A and B) included soil water retention curves reported by Carvallo et al. (1976) for 35 soil samples. Soil samples of a Maddock sandy loam (sandy, mixed Udorthentitic Hapludoll) of 345 cm³ were collected at seven depths from five plots in a study area of approximately 100 m² in size. Particle-size distribution data clearly showed differences in soil texture between the surface and subsurface soil. Therefore, the soil samples were categorized into two data sets based on their clay content. Data set A (>8% clay content) contained soil water retention data for 17 soil samples of the surface soil, and data set B (<8% clay content) contained retention data for 18 samples of the subsurface soil.

In the fitting of soil water retention data, the saturated water content, \( \theta_s \), was fixed to the observed value and the residual water content, \( \theta_r \), was considered a fitting parameter. Optimized values of \( \theta_s \) were constrained such that their values were smaller than the minimal observed \( \theta \)-value, but always \( \geq 0 \). Mean values of \( \theta_s \) and \( \theta_r \) for data set A were 0.390 and 0.173 m³ m⁻³, respectively, with corresponding standard deviations of 0.021 and 0.036 m³ m⁻³. Mean \( \theta_s \) and \( \theta_r \) values for data set B were 0.396 and 0.088 m³ m⁻³, respectively, with corresponding standard deviations of 0.024 and 0.022 m³ m⁻³. Since values of both \( \theta_s \) and \( \theta_r \) varied between soil samples, soil water retention data were scaled using effective saturation, \( S_0 \), rather than volumetric water content (Clausnitzer et al., 1992).

The third data set (data set H) included samples of a loamy sand soil of 100 cm³ collected from six locations at three depths (30, 60, and 90 cm) in a 2-m² area in the Hupselse Beek watershed (Hopmans and Stricker, 1987). A total of 43 soil samples were analyzed. The parameters \( h_{nj,i} \) and \( \sigma_i \) were fitted to the Kosugi \( S_r(h) \) model (Eq. [28]) using the observed retention data. Following Hopmans and Stricker (1987), \( \theta_s \) was considered a fitting parameter and \( \theta_r \) was assumed to be zero. Hence, \( S_r \) was numerically equal to the degree of saturation. Mean and standard deviation of \( \theta_s \) were 0.338 and 0.060 m³ m⁻³, respectively.

The parameters for the reference soil water retention curve in Eq. [30] for each of the three data sets were directly computed from the mean values of \( \ln h_{nj,i} \) (Eq. [31]) and \( \sigma^2 = \text{Mean}(\sigma_i^2) \). The scaling factors \( \alpha_i \) were directly computed from Eq. [32].

Reference soil water retention curves and scaling factors were also computed from a “brute force” scaling approach, minimizing the residual sum of squares (RSS) computed from the effective saturation values of the scaled vs. reference retention curves for \( I \) soil samples and \( f(h) \) retention data (conventional method):

\[
\text{RSS} = \sum_{i=1}^{I} \left[ \frac{S_{nj,i}(\alpha h_{nj,i})}{S_r} - \hat{S}_i(\alpha h_{nj,i}) \right]^2 \tag{34}
\]

where \( S_j \) is the \( j \)th effective saturation value of soil sample \( i \) with a capillary pressure head of \( h \), and \( S_r(\alpha h) \) is the effective saturation obtained from substituting \( h = \alpha h \) into the reference soil water retention function that is expressed by Eq. [30]. This method of scaling is similar to the method used by Hills et al. (1992), except that the Kosugi retention model was used instead of the van Genuchten (1980) retention function. Moreover, we imposed the same constraint expressed by Eq. [25], requiring the geometric mean of scaling factors to be unity, while Hills et al. (1992) imposed no constraint. Thus, in the conventional scaling method, the parameters of the reference retention curve, \( \ln h_{nj,i} \) and \( \sigma^2 \), and scaling factors are optimized while minimizing RSS, whereas the physically based approach determines these parameter values directly from \( h_{nj,i} \) and \( \sigma_i \).

**RESULTS AND DISCUSSION**

**Soil Water Retention Parameters**

Values for \( \ln h_{nj,i} \) and \( \sigma^2 \), which together describe the soil water retention function of Eq. [28] for each sample \( i \), are presented in Fig. 3 for all three data sets. The top horizontal scale in Fig. 3 represents the median pore radius \( r_{nj,i} \), which is related to \( h_{nj,i} \) by Eq. [4]. The corresponding mean parameter values, \( \text{Mean}(\ln h_{nj,i}) \) and \( \text{Mean}(\sigma_i^2) \), are presented in Table 1. The value of \( \text{Mean}(\ln h_{nj,i}) \) is equal to both \( \ln h_{mr} \) and \( \ln h_{in} \), representing the mean of \( h \) for the study area and the reference soil, respectively. Moreover, \( \text{Mean}(\sigma_i^2) \) describes the variance of \( h \) for the reference soil. From Fig. 3 and Table 1 it can be seen that the soil water retention curves of data set B are described by the smallest \( \text{Mean}(\ln h_{nj,i}) \) and \( \text{Mean}(\sigma_i^2) \) values among the three data sets. Consequently, data set B has the largest median pore radius and the smallest pore-radius distribution width among the three data sets. The data and Fig. 1a in Carvallo et al. (1976) support the distinct difference in pore-radius distribution between data sets A and B. Indeed, it is
shown in Fig. 3 that the soil samples of data set A (>8% clay content) have large In $h_{m_j}$ and $\sigma_j$ values relative to those of data set B (<8% clay content). In contrast, the retention curves of data set H have the smallest median pore radius, and that set’s pore-radius distribution width is the largest (Table 1).

According to the probability plots of In $h_{m_j}$ shown in Fig. 4, In $h_{m_j}$ values are normally distributed for all three data sets, supporting the assumptions with regard to the mean properties of the sampling distribution. Using the Kolmogorov–Smirnov test, the null (normal) hypothesis was not rejected at the 0.20 level of significance for any of the three data sets. The variance, $\text{Var}(\ln h_{m_j})$, presented in Table 1 were computed from the slopes of the probability plots in Fig. 4. The variance of the sample mean values was largest for data set H; i.e., the variability of soil water retention was largest for this data set.

### Scaling Results

Figures 5 through 7 present the scaling results for all three data sets, with each of the three figures showing the unscaled and scaled retention data, and each using the proposed physically based (PB in Table 2) and conventional (C in Table 2) scaling methods. Values of RSS and the obtained parameters of the reference soil water retention curves are summarized in Table 2. Despite the fact that the RSS value for the PB method is slightly larger than that for the C method, the PB method produced excellent scaling results for every data set. The reductions in RSS were 57.8 to 89.2% for the PB and C methods, respectively. In the physically based approach, all retention data are scaled using $h_{m_j}$ data only. Consequently, the scaled soil water retention data obtained by the PB method match the reference retention curve at $S_c = 0.5$ for all data sets (Fig. 5b, 6b, and 7b). In the conventional approach, all retention data receive equal weight. For every data set, the mean value for the reference soil, In $h_{m_j}$, optimized by the C method is close to its calculated value using the PB method, whereas the variance for the reference soil, $\sigma_j$, optimized by the C method is slightly smaller than the variance calculated by the PB method (Table 2).

Figure 8 presents the probability plots of In $\alpha$, using both the PB and C scaling methods. It can be seen that a lognormal distribution describes the distribution of scaling factors for all three data sets and both scaling methods. The Kolmogorov–Smirnov test did not reject the null hypothesis, i.e., In $\alpha$ is normally distributed, at the 0.20 level of significance for all cases. The statistical properties (mean and variance) of the lognormal scaling factor distribution, as computed from the slope and interception of the linear relationship, are summarized in Table 2. As expected, the variance of In $\alpha$, obtained by the PB method is equal to the variance of In $h_{m_j}$ (Table 1) for all three data sets. From Fig. 8 and Table 2 it can be seen that the scaling factor distribution obtained by the PB method is quite similar to that obtained by the C method for every data set. The scaling requires that the mean of In $\alpha$ be equal to zero in all cases (Eq. [25]), and the scaling factor distributions are characterized by the variance of In $\alpha$ only.

Most studies (e.g., Warrick et al., 1977; Vachaud et al., 1985; Hopmans and Overmars, 1987; Clausnitzer et al., 1992) have shown that scaling factors are approximately lognormally distributed. The same was true for all three soil water retention data sets analyzed in this study. Moreover, it has been demonstrated many times...
Fig. 5. (a) Unsealed soil water retention data, and scaled soil water retention data using (b) physically based (PB) scaling and (c) conventional (C) scaling for data set A.

Fig. 6. (a) Unsealed soil water retention data, and scaled soil water retention data using (b) physically based (PB) scaling and (c) conventional (C) scaling for data set B.

Fig. 7. (a) Unsealed soil water retention data, and scaled soil water retention data using (b) physically based (PB) scaling and (c) conventional (C) scaling for data set H.
Table 2. Residual sum of squares (RSS), parameters for reference retention curves, and statistical properties of scaling factor sets obtained by the proposed physically based (PB) and conventional (C) scaling methods.

<table>
<thead>
<tr>
<th>Data set</th>
<th>RSS</th>
<th>$\ln h_m$</th>
<th>$\phi^2$</th>
<th>Mean($\ln \alpha_o$)</th>
<th>Var($\ln \alpha_o$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.119</td>
<td>0.109</td>
<td>3.914</td>
<td>3.397</td>
<td>1.317</td>
</tr>
<tr>
<td>B</td>
<td>0.106</td>
<td>0.107</td>
<td>3.393</td>
<td>3.397</td>
<td>0.266</td>
</tr>
<tr>
<td>H</td>
<td>1.496</td>
<td>1.011</td>
<td>4.012</td>
<td>4.001</td>
<td>1.064</td>
</tr>
</tbody>
</table>

that the distribution of the saturated hydraulic conductivity ($K_s$) obeys a lognormal distribution (e.g., Kutilek and Nielsen, 1994; Bierkens, 1996). This result would also suggest a lognormal distribution of scaling factors, since the hydraulic conductivity is expected to be proportional to the square of the scaling factor (Jury et al., 1987). This study proposes an interpretation of these empirically derived results. That is, a lognormal scaling factor distribution was derived by assuming that the soil pore radius of a study area is lognormally distributed and that individual soil samples are obtained from random sampling of effective soil pore volume from the study area.

**Upscaling of Soil Water Retention Curves**

According to Eq. [23], the variance of $\ln h$ for the study area, $\sigma_i^2$, is equal to the sum of the expected variance of $\ln h$ within soil samples, Mean($\sigma_i$), and the variance of scaling factors, Var($\ln \alpha_o$), which is equal to the variance of mean $\ln h$ values between soil samples, Var($\ln h_m$) (see Eq. [33]). In Fig. 9, the contributions of Mean($\sigma_i$) and Var($\ln \alpha_o$) to $\sigma_i^2$ value for all three data sets are shown. The $\sigma_i$ values computed as the sum of Mean($\sigma_i$) and Var($\ln \alpha_o$) are summarized in Table 1. Data sets H and B have the largest and smallest $\sigma_i$ values, respectively. In all cases, however, the expected variance within samples is much larger (81-96% of total variance) than the variance between samples.

The retention curve for the whole study area expressed by Eq. [29] can be regarded as the upscaled soil water retention curve, of which the parameters $h_m$ and $\alpha_o$ are computed from Eq. [31] and [33], respectively. Others have attempted to derive upscaled parameters from mostly empirical parameters of the soil water retention curves of individual soil samples (e.g., Yeh and Harvey, 1990; Green et al., 1996) using different averaging techniques (arithmetic, harmonic, or geometric). In the proposed physically based approach of Eq. [31], the parameter $h_m^*$ is computed from the geometric mean of $h_m$, whereas the variance parameter of the upscaled retention function, $\sigma_i^2$, is derived from the arithmetic mean of $\alpha_o^2$ and the variance of $\ln h_m$ (see Eq. [33]). In contrast to describing the effect of soil heterogeneity on the water regime in a study area using the Monte Carlo approach from scaled soil hydraulic properties, the upscaled soil water retention function expressed by
Eq. [29] can potentially characterize field effective soil variability using a single soil water retention curve. Of course, this would require additional investigations, including the upscaling of unsaturated hydraulic conductivity data and the analysis of spatial correlation between samples.

SUMMARY AND CONCLUSIONS

A physically based scaling method is proposed that uses a lognormal pore-size distribution. By using this method, scaling factors and parameters of the reference curve are computed directly from the parameters of individual soil water retention functions, which are the mean and variance of the log-transformed pore-radius distribution. The proposed method was compared with a conventional scaling method for three sets of soil water retention data and was shown to provide excellent scaling results. The scaling factor distribution obtained by the proposed method was quite similar to that obtained by the conventional method.

In contrast to conventional scaling, the physically based scaling provides a theoretical interpretation of the lognormal scaling factor distribution with the assumption that the soil pore radius of a study area is lognormally distributed and that individual soil samples are obtained from random sampling of effective soil pore volume from this study area. The statistical description of a set of soil water retention data is characterized by the variance of log-transformed scaling factors, which is defined by the difference of the total variance of the log-transformed pore radius in the study area and the expected variance of the log-transformed pore radius within soil samples. Moreover, physically based scaling allows upscaling of the physically based parameters of the soil water retention function, thereby providing a hydraulic parameterization representative of the whole study area from individual soil samples. As a result, the physically based scaling can be effectively used for theoretical studies on spatial variability and upscaling of soil water retention characteristics.

ACKNOWLEDGMENTS

The first author sincerely thanks Dr. T. Mizuyama, Dr. S. Kobaishi, and Dr. M. Fujita (all of Kyoto University) for giving him an opportunity to study in the USA. This research was partly supported by International Scientific Research Program of the Ministry of Education, Science, Sports, and Culture of the Japanese Government.

APPENDIX A

The variance of the PDF of ln r for soil sample i is defined as

\[ \sigma_i^2 = \int_{-\infty}^{\infty} (\ln r - \ln r_m)^2 f_i(\ln r) d\ln r \]  \hspace{1cm} [A1]

which is transformed to

\[ \sigma_i^2 = \int_{-\infty}^{\infty} (\ln r - \ln r_m + \ln r_{m*} - r_m)^2 f_i(\ln r) d\ln r \]  \hspace{1cm} [A2]

APPENDIX B

A constant in the capillary pressure function given by Eq. [4], cm²

\( F_a \) the normal distribution function

\( f(\ln r) \) probability density function of log-transformed soil pore radius

\( g(\ln h) \) probability density function of log-transformed soil capillary pressure head

\( h \) soil capillary pressure head, cm

\( h_m \) geometric mean and median in the lognormal distribution of soil capillary pressure head, cm

\( l \) total number of soil samples

\( \ln h_{mi} \) mean of \( \ln h \) for soil sample \( i \)

\( \ln h_{m*} \) mean of \( \ln h \) for study area (= \( \text{Mean}(\ln h_{mi}) \))

\( \ln h_{r*} \) mean of \( \ln h \) for reference soil

\( \ln r_{mi} \) mean of \( \ln r \) for soil sample \( i \)

\( \ln r_{m*} \) mean of \( \ln r \) for study area (= \( \text{Mean}(\ln r_{mi}) \))

\( \ln r_{m} \) mean of \( \ln r \) for reference soil

\( \text{Mean} \) arithmetic mean operator

\( \text{Var} \) variance operator

\( \alpha_i \) scaling factor of soil sample \( i \)

\( \theta_a \) volumetric soil water content, m³ m⁻³

\( \theta_r \) residual soil water content, m³ m⁻³

\( \theta_s \) saturated soil water content, m³ m⁻³

\( \sigma \) standard deviation in the normal distribution of both log-transformed soil pore radius and capillary pressure head

\( \sigma_i^2 \) variance of both \( \ln r \) and \( \ln h \) for soil sample \( i \)
\(\sigma^2\) variance of both \(\ln r\) and \(\ln h\) for study area \[= \text{Mean}(\sigma^2) + \text{Var}(\ln h_m)\)

\(\delta^2\) variance of both \(\ln r\) and \(\ln h\) for reference soil

**REFERENCES**


ERRATUM

Scaling Water Retention Curves for Soils with Lognormal Pore-Size Distribution

K. Kosugi and J. W. Hopmans

The following is a list of errors found in the paper above.

1. On the right-hand side of Eq. [10], [11], [15], [16], [17], and [27], the square root notation must exclude the standard deviation parameter, $\sigma$.

2. The last sentence of the Physically Based Scaling Approach section on p. 1499 should read:

   Using the scaling theory approach as outlined above, the PDF of $\ln r$ for the reference soil, $f(\ln r)$, is defined by

   $$
   \hat{f}(\ln r) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(\ln r - \ln \hat{r}_m)^2}{2\sigma^2} \right]
   $$

[27]

3. In the sentence immediately following Eq. [34] on p. 1500, $S_r(\alpha_i h^{i'})$ must be replaced by $\hat{S}_r(\alpha_i h^{i'})$. Moreover, in the fourth sentence of the Results and Discussion section on p. 1500, $\ln h_m$ must be replaced by $\ln \hat{h}_m$.

4. In the second paragraph of the Scaling Results section on p. 1501, all occurrences of $\ln \sigma_i$ must be replaced by $\ln \alpha_i$.

5. In the first integrand of Eq. [A2] in Appendix A, $r_{m,i}$ must be replaced by $\ln r_{m,i}$. 