WATER RESOURCES RESEARCH, VOL. 46, W01501, doi:10.1029/2009WR007939, 2010

Can texture-based classification optimally classify soils with respect to soil hydraulics?

Navin K. C. Twarakavi,[1] Jirka Šimůnek,[2] and M. G. Schaap[3]

Received 3 March 2009; revised 18 August 2009; accepted 31 August 2009; published 1 January 2010.

[1] In the past, texture-based classification of soils has been used for grouping soils in variably saturated water flow and solute transport studies. Classification of soils becomes especially important for large-scale studies where the spatial and temporal variability in the hydraulic properties of soils exceeds the field sampling capabilities. Although soil-texture-based classification has been widely used, questions remain about the validity of its use from a hydraulic perspective. In this study, we attempt to answer the following questions: (1) what is the optimal number of (soil hydraulic) classes that can adequately classify the soils from a hydraulic standpoint, and (2) how does such a classification compare to the soil texture classification currently used? To investigate these questions, the commonly used *k*-means clustering algorithm was integrated with the ROSETTA pedotransfer functions to predict the so-called soil hydraulic classes. The optimal soil hydraulic classifications and the associated uncertainty were estimated for numbers of soil hydraulic classes varying from 2 to 30. It was concluded that the optimal number of soil hydraulic classes is 12. The optimal soil hydraulic classes were represented in a ternary diagram called the soil hydraulic triangle. While there exist some surprising similarities in classification between the soil texture triangle and the soil hydraulic triangle for soils with high sand percentages (sand >60%), the opposite is true for soils with low sand contents. From a hydraulic standpoint, the texture-based classification does not classify soils well when there is a considerable impact of capillary forces. The soil texture and hydraulic classes were analyzed for accuracy using two databases. Compared to the soil texture classes, it was found that the soil hydraulic classes marginally improve the accuracy of classification. Even though the improvement is only marginal, it was observed that the optimality of soil texture triangle for hydraulic studies cannot be assured because of the nonuniform distribution of data across various textural possibilities in the two databases. As an extension of this research, we have also estimated the average soil hydraulic parameters for the different optimal soil hydraulic classes.

Citation: Twarakavi, N. K. C., J. Šimůnek, and M. G. Schaap (2010), Can texture-based classification optimally classify soils with respect to soil hydraulics?, *Water Resour. Res.*, *46*, W01501, doi:10.1029/2009WR007939.

1. Introduction

[2] Understanding hydrological processes, especially at larger scales, requires grouping soils on the basis of some relevant criteria. One of the most common approaches for characterizing soils is based on their textural properties. Soil texture describes the relative proportion of different grain size particles in a soil and is commonly represented by relative proportions of sand, silt, and clay contents. Soil texture classification has been addressed frequently, with the earliest research dating back to the beginning of the 20th century [e.g., *Atterberg*, 1905, 1912]. One of the major reasons for the popularity of texture-based classification of soil is that textural characteristics are among the most easily measured soil properties. Skilled practitioners can determine the textural class manually. Texture-based classification categorizes soils into so-called soil texture classes. Although several texture-based classifications are used currently (U.S. Department of Agriculture (USDA) and International Soil Science Society (ISSS)), the soil texture classification system proposed by the USDA is perhaps the most widely recognized. The USDA system classifies soils into 12 soil texture classes. This soil texture classification is conveniently represented using the ternary diagram as proposed by *Davis and Bennett* [1927]. Figure 1 shows a ternary diagram of the soil texture triangle with the USDA-based soil texture classes. A ternary diagram represents the different soil texture classes as a function of sand, silt, and clay percentages, in which the size boundaries are set at 2 μm for clay/silt, at 50 μm for silt/sand, and at 2000 μm for sand/gravel.

[1]Department of Agronomy and Soils, Auburn University, Auburn, Alabama, USA.
[2]Department of Environmental Sciences, University of California, Riverside, California, USA.
[3]Department of Soil, Water and Environmental Sciences, University of Arizona, Tucson, Arizona, USA.

0043-1397/10/2009WR007939$09.00

2. Hydraulic Validity of Soil Texture Classes

[3] Studies related to the hydraulic behavior of soils have typically used soil-texture-based information. Several

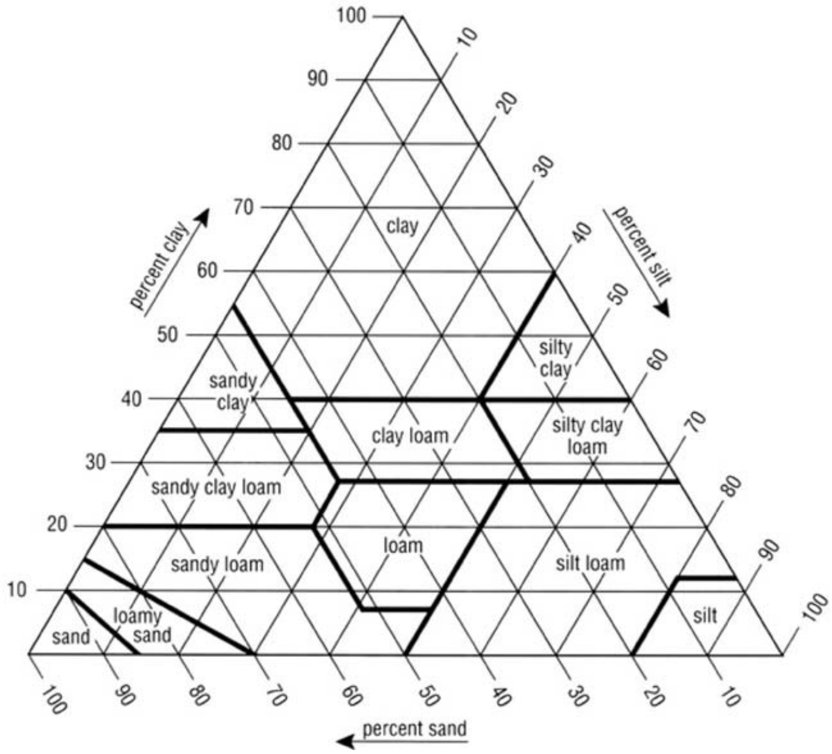

Figure 1. A ternary diagram of the soil texture triangle showing the different USDA-based soil texture classifications.

approaches are possible. For example, *Carsel and Parrish* [1988], *Rawls et al.* [1982], *Wösten et al.* [1995], and *Schaap and Leij* [1998] estimated the average soil hydraulic parameters for different textural classes, whereas other studies (for reviews, see *Wösten et al.* [2001] and *Pachepsky and Rawls* [2004]) developed pedotransfer functions (PTFs) that provided estimated hydraulic parameters as a function of continuous texture and other soil variables. In the latter case, the performance of each PTF was quantified on a textural class basis [e.g., *Schaap and Leij*, 1998]. In most or all of these studies, it was assumed that soil texture is the dominant soil variable in determining hydraulic properties, while other soil variables, such as bulk density or organic matter content, have only a secondary effect. However, a basic (but unproven) premise of these studies is that the traditional soil texture *classification* is optimal from a soil hydraulic standpoint. The hidden assumption behind using the soil-texture-based classes in variably saturated flow studies is that these classes provide the best possible grouping of soils. In other words, it is assumed that the soils within each soil texture class are as similar as possible to each other from the perspective of hydraulic behavior. Also, it is assumed that the number of classes used (that is, the 12 textural classes) is the optimal number of classes. Because of the wide use of soil texture classes in soil physics and hydrology, an investigation of the optimality of soil texture classes becomes important. Even though the *National Resources Conservation Service* (*NRCS*) [1975] discusses soil taxonomy classification, the following quotes [*NRCS*, 1975, p. 7] still hold true for soil texture classification and its relevance for hydrological purposes.

> For the different purposes of the soil survey, classes are needed that can be grouped or subdivided and regrouped to permit the largest number and the most precise predictions possible about responses to management and manipulation.
>
> Classifications are contrivances made by humans to suit their purposes. They are not themselves truths that can be discovered. A perfect classification would have no drawbacks when used for the purpose intended. Each distinctly different purpose, to be served best, demands a different classification.
>
> As knowledge expands, new facts or closer approximations of truths not only make improvements in classification possible but also make some changes imperative. Thus, classifications are not static but require change as knowledge expands.

[4] To understand the motive and objectives of soil texture classification, we conducted an extensive literature review but were unable to gather precise references. One of the earliest references discussing the soil texture triangle was the USDA *Soil Survey Manual* [*Soil Survey Staff*, 1951], which states that soil texture classes in present use are defined in terms of size distribution as determined by mechanical analysis in the laboratory. The report also mentions that definitions and boundaries for the soil texture classes resulted from long experience and special research for maximum general use for soil definitions and interpretations. It is also acknowledged that the standards are not yet perfect. *Soil Survey Staff* [1951] goes on to stress that textural classes must be used exclusively to express the differences in particle size distribution and that any other use (such as for soil structure or hydrology) will result in the loss of their fundamental significance. It is clear that the hydraulic validity of the classification of soils based on the soil texture triangle needs to be scrutinized. In this study, we strive to provide answers to the following two questions. (1) Is an optimal classification of soils based on hydraulic characteristics (hereafter called soil hydraulic classes) different from soil-texture-based

classification? (2) With respect to the hydraulic characteristics, what is the optimal number of soil hydraulic classes?

[5] In this study, we do not question the optimality of the USDA particle size boundaries (2, 50, and 2000 μm) or the optimal number of such particle sizes. *Skaggs et al.* [2001] seem to suggest that four instead of three particle size fractions are needed to adequately characterize a particle size distribution. This would mean that the relationship between texture and hydraulic parameters could be improved but that the current ternary diagrams can no longer be used.

[6] Here, we summarize the methodology of this research. The objective of this research is achieved using the following steps: (1) obtaining soil data sets with relevant hydraulic properties that adequately represent the different soil texture possibilities (that is, all possible combinations of sand, silt, and clay percentages); (2) using appropriate techniques, grouping the soils in the data set into an optimal number of classes based on soil hydraulic properties; and (3) comparing the resulting classification of soils into soil hydraulic classes with the soil texture classes as described by the USDA system.

3. Materials and Methods

3.1. Soil Hydraulic Properties

[7] We consider the soil hydraulic parameters as described by the van Genuchten–Mualem model [*van Genuchten*, 1980]:

$$\theta(h) = \begin{cases} \theta_r + \dfrac{\theta_s - \theta_r}{[1 + |\alpha h|^n]^m} & h < 0 \\ \theta_s & h \geq 0, \end{cases} \tag{1a}$$

$$S_e(h) = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r}, \tag{1b}$$

$$K(h) = K_s S_e^l \left[1 - (1 - S_e^{1/m})^m\right]^2, \tag{1c}$$

$$m = 1 - 1/n, \qquad n > 1, \tag{1d}$$

where $\theta(h)$ is the volumetric water content ($L^3\ L^{-3}$) as a function of the pressure head h (L); θ_s and θ_r are the saturated and residual volumetric water contents, respectively; $S_e(h)$ is the soil water saturation (dimensionless) expressed at the pressure head h; α (L^{-1}) and n (dimensionless) are van Genuchten shape parameters; $K(S_e)$ is the unsaturated hydraulic conductivity ($L\ T^{-1}$) as a function of the soil water saturation; K_s is the saturated hydraulic conductivity ($L\ T^{-1}$); and l is an empirical parameter commonly assumed to be 0.5.

3.2. The *k*-Means Clustering Algorithm

[8] The *k*-means clustering algorithm [*MacQueen*, 1967] is one of the simplest unsupervised clustering algorithms for solving the well-known clustering problem. The procedure follows a simple method for classifying a given data set of size N into a specific, predefined number of clusters (assumed as k clusters). If one considers a data set that contains N data points, each having p attributes, the objective of the *k*-means clustering algorithm is to group the data set into k clusters such that the objective function, $\xi(k)$, described in equation (2), is minimized:

$$\xi(k) = \sum_{i=1}^{k} \sum_{j=1}^{n_k} \| \mathbf{x}_{ij} - \mu_i \|^2, \tag{2}$$

where k is the number of clusters determined a priori, n_k is the number of points belonging to the cluster i, $\mathbf{x}_{ij}$ is the vector of attributes of the jth data point of the cluster i (a vector of hydraulic parameter values), μ_i is the mean value of the attributes for the data points in the cluster i (also called cluster centers), and $\| \ \|$ describes the Euclidean distance (a squared distance) between two data points.

[9] The algorithm for the *k*-means clustering consists of the following steps: (1) for a predefined number of k clusters, initialize them by placing the k cluster centers randomly in the attribute space; (2) assign each data point to the nearest cluster center on the basis of the Euclidean distance measure; (3) for each cluster, recalculate the new centers as the mean of the data points assigned to it; (4) repeat steps 2 and 3 until the cluster centers no longer move. One may note that the *k*-means algorithm does not necessarily find the most optimal clustering result and is significantly sensitive to the position of the random cluster centers that are selected initially. Several variations of the *k*-means clustering algorithm that emphasize obtaining the global minimum of the objective function have been proposed. In this study, we use *Hartigan and Wong*'s [1979] algorithm because it performs better in obtaining the global minimum. In order to make the clustering results more robust, the *k*-means algorithm was run multiple times (here 1000 times) with different initial cluster centers, and the average clustering was used.

[10] As an unsupervised learning technique, the *k*-means clustering algorithm requires a predefined number of clusters. One may estimate an optimal number of clusters in the *k*-means clustering algorithm using the *elbow criterion*. We would like to point out that there are a number of other approaches for finding optimal clustering and that we chose the elbow criterion because of its popularity and simplicity of use. The elbow criterion suggests that one should choose an optimal number of clusters such that adding one additional cluster does not add sufficient information. One may note that the value of the objective function of the *k*-means clustering algorithm decreases as the number of required clusters increases. However, the rate of decrease in the objective function is less and less dramatic as the number of clusters increases. The elbow criterion suggests an intuitive selection of the optimal number of clusters such that the decrease in the value of the objective function is not significant enough. In this study, we use the elbow criterion to select the optimal number of soil hydraulic classes.

3.3. Methodology

[11] A major hurdle in our research was obtaining a comprehensive soil data set with hydraulic properties that cover the various textural combinations. A number of the soil data sets tend to be clustered in the sand-loam-silt categories (the lower half of the soil texture triangle), with minimal data points in the clay-dominated classes.

[12] In this study, we used the following procedure for classifying soils from a hydraulic standpoint. First, we estimated the five soil hydraulic parameters (θ_r, θ_s, α, n,

and K_s) throughout the entire soil texture triangle using the ROSETTA PTFs such that the various soil texture possibilities (i.e., combinations of sand, silt, and clay percentages) were considered. The data set was created by varying the sand, silt, and clay percentages by 1%, leading to a total of 5151 data points. The sand, silt, and clay percentages were used as inputs to ROSETTA. We used only these percentages as input to ROSETTA so that the results of the clustering process could be shown in a ternary diagram as a function of sand, silt, and clay contents.

[13] The ROSETTA PTFs use artificial neural networks and can predict the necessary soil hydraulic parameters (as well as their uncertainties) as a function of the soil texture (sand, silt, and clay fractions). One may note that a number of PTFs are currently available and might also be used. However, we used ROSETTA because it is one of the more widely accepted PTFs. Although we based our study on generated data, we should emphasize that this does not implicitly include the USDA textural classes as only sand, silt, and clay percentages were entered into ROSETTA and soil texture classification was not.

[14] ROSETTA predicted the mean value and the associated uncertainties of these soil hydraulic parameters for each of the 5151 soil data points. Once the mean and standard deviation (uncertainty) values of the five soil hydraulic parameters at each of the 5151 soil data points were available, they were used to develop 100 data sets of soil hydraulic parameter estimates using the Monte Carlo sampling approach. For the 100 data sets (each having 5151 data points with the estimated soil hydraulic parameters), we estimated the following hydraulic properties: (1) field capacity (θ_{fc}), (2) time needed to reach field capacity from saturated condition (called drainability, t_{fc}), and (3) capillary pressure at which field capacity is attained (h_{fc}). Appendix A discusses the simulation approach in HYDRUS used to estimate the aforementioned hydraulic properties. It took about 800 h of run time on a 3 GHz, 3.25 GB RAM Intel Quad Core Windows XP personal computer to carry out the required 515,100 simulations in HYDRUS. Using the soil hydraulic parameters, we also estimated the permanent wilting point (θ_{pwp}) as the water content in the soil at a capillary pressure of -15 bars. Along with estimated θ_{fc}, t_{fc}, h_{fc}, and θ_{pwp}, we used the saturated water content (θ_s) and the saturated hydraulic conductivity (K_s) as properties to group the soils. While θ_{fc}, h_{fc}, θ_{pwp}, and θ_s represented the retention characteristics of each soil, K_s and t_{fc} represented the dynamics of variably saturated water flow. At this point, we would like to mention that we purposefully avoided classification of soils by directly using van Genuchten–Mualem soil hydraulic parameters. This is because such an approach would result in the usability of the resulting soil hydraulic classification being much more contingent on the interrelationships between the parameters of the van Genuchten–Mualem model.

[15] For each of the 100 data sets, we grouped the soils using the following properties: θ_{fc}, θ_{pwp}, θ_s, $\log_{10}(K_s)$, $\log_{10}(h_{fc})$, and $\log_{10}(t_{fc})$ for various possible numbers of soil hydraulic classes (k, equation (2)). Prior to the application of the k-means clustering algorithm on each of the 100 data sets, the data were standardized such that each of the soil-hydraulic properties imparted an equal influence on the clustering process. A logarithmic transformation of K_s, h_{fc}, and t_{fc} was used for normalization purposes. For a given number of clusters (k), the cluster boundaries from each of the 100 data sets were collected, and the average boundary was chosen for the final classification. The cluster boundary from the 100 data sets was also used to define the uncertainty of the final soil hydraulic classification. All the points with at least one boundary line in the 100 data sets were considered to be a region of uncertainty in the hydraulic classification.

3.4. Validation Data Sets

[16] In order to understand how well the soil hydraulic triangle classifies soils when compared to the soil texture triangle, we analyzed their performance on different real-world databases. The first database was the one used to train and test the ROSETTA PTFs. The database contained 2134 soil samples, with parameters describing the water retention data. Saturated hydraulic conductivity data was available for 1306 soil samples. The second database was used to develop PTFs by *Minasny et al.* [2004]. This data set had 320 soil samples with water retention data, of which 219 samples had estimates of the saturated hydraulic conductivity. Readers are referred to *Schaap et al.* [2001] and *Minasny et al.* [2004] for a detailed description of the databases.

4. Results and Discussion

[17] As described in section 3.3, 100 data sets with the properties θ_{fc}, θ_{pwp}, θ_s, $\log_{10}(K_s)$, $\log_{10}(h_{fc})$, and $\log_{10}(t_{fc})$ as a function of soil texture were developed. Each data set consisted of 5151 data points that uniformly cover the different textural combinations. Figure 2 shows the mean estimates of these properties from the 100 data sets in a ternary diagram. As mentioned earlier, the k-means clustering algorithm was used to classify the soils for a different, predefined number of soil hydraulic classes, k, ranging from 2 to 30. Figure 3 shows the value of the objective function for the k-means clustering (described in equation (2)) as a function of the number of soil hydraulic classes. Note that the plot shows the mean and standard deviations of the objective function from the 100 data sets as a function of the number of soil hydraulic classes. This plot is also referred to as the elbow criterion, which can be used to intuitively judge the optimal number of soil hydraulic classes. Figure 3 also shows the incremental improvement in classification with an increase in the number of hydraulic classes. It was observed that the optimal number of classes for classifying soils with respect to soil hydraulic properties was 12. We chose the optimal number of soil hydraulic classes to be 12 on the basis of (1) the elbow criterion and (2) consistency with the number of soil texture classes used in the USDA classification (which is 12).

[18] It is of interest to analyze how various soil texture possibilities (sand, silt, and clay percentages) group for different numbers of soil hydraulic classes. Figure 4 shows the location of the soil hydraulic classes generated using the k-means clustering approach for different predefined number of classes. As the number of soil hydraulic classes increases (from 2 (Figure 4a) to 14 (Figure 4f)), one may note a nearly hierarchical and consistent development of the spatial location of the hydraulic classes in the ternary diagram. Even for a lower number of soil hydraulic classes (Figure 4b), a clear distinction is observed between the sand-dominated soils

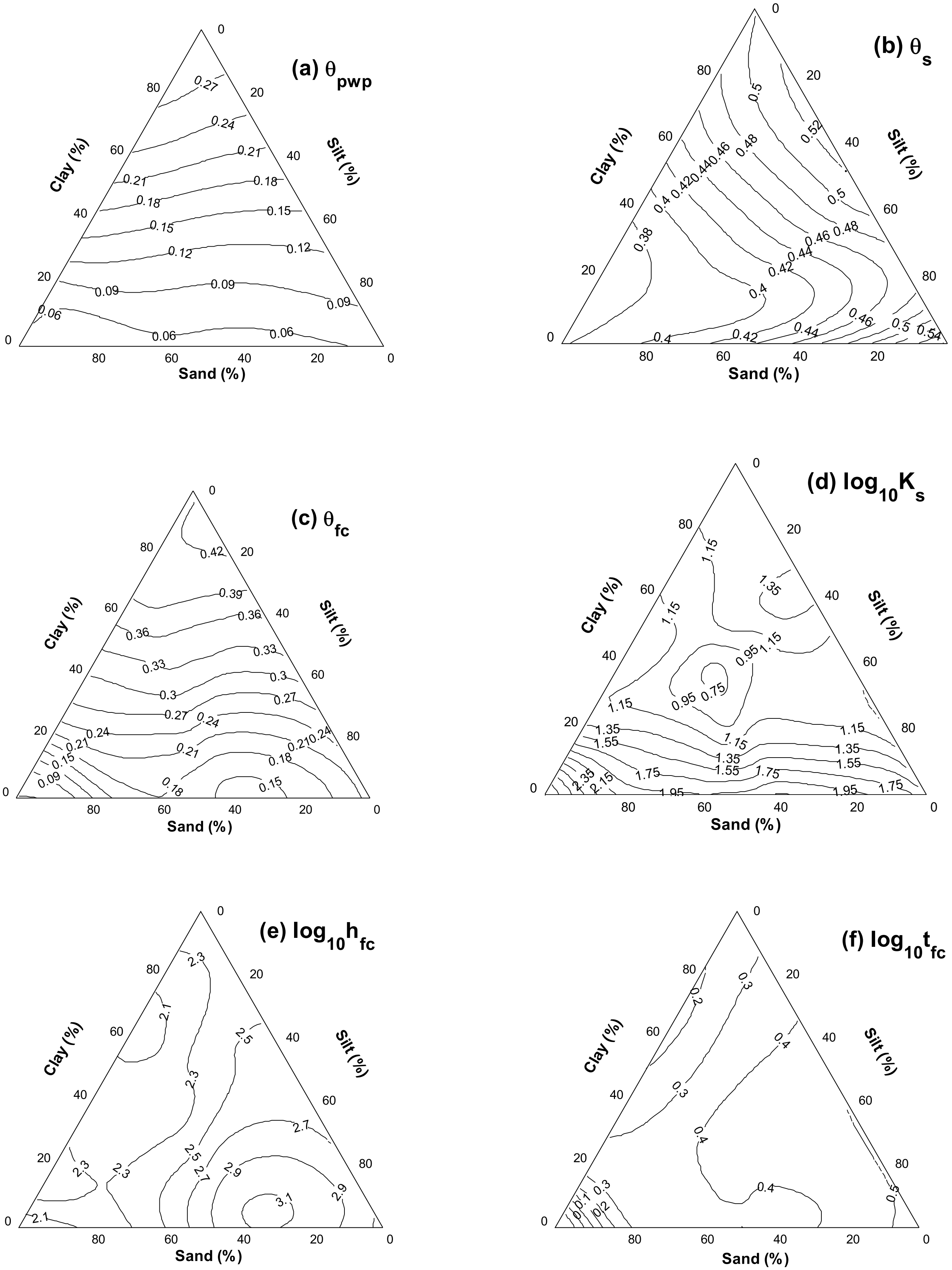

Figure 2. Mean estimates of (a) θ_{pwp}, (b) θ_s, (c) θ_{fc}, (d) $\log_{10}(K_s)$, (e) $\log_{10}(h_{\mathrm{fc}})$, and (f) $\log_{10}(t_{\mathrm{fc}})$ as a function of sand, silt, and clay percentages.

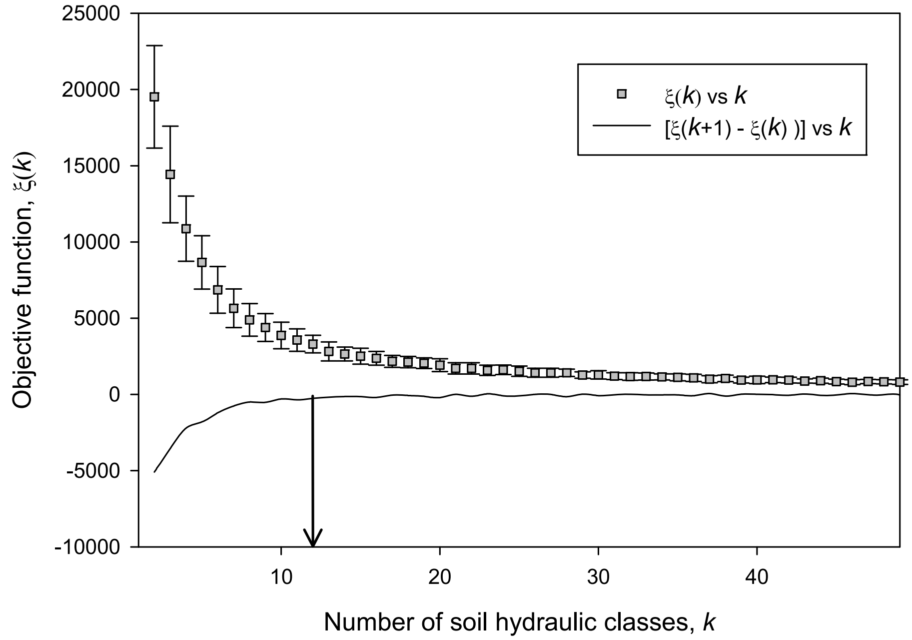

Figure 3. Plot showing the mean and standard deviation of the objective function ($\xi(k)$) as a function of the number of soil hydraulic classes (k). The plot also shows the incremental improvement in the objective function, $[\xi(k+1) - \xi(k)]$, with a unit increase in the number of hydraulic classes.

(sand >60%) and the rest. It is safe to say that soils with higher sand content are clearly differentiated from other soils in terms of hydraulic behavior.

[19] Figure 5 shows the soil hydraulic classification for the optimal number of soil hydraulic classes as determined by the elbow criterion. We refer to this ternary diagram as the *soil hydraulic triangle* because it shows the optimal classification of soils from a hydraulic perspective. We have named the hydraulic classes using an alphanumeric nomenclature. The sand-dominated soils are represented by the letter A, with a number that increases as the sand fraction decreases (A1–A4). Similarly, silt-dominated soils are represented by the letter B, with a number ranging from 1 to 4, and clay-dominated soils are represented by the letter C, with a number ranging from 1 to 4. As mentioned earlier, the optimal soil hydraulic classification is very similar to the soil texture classification (Figure 1) for soils where the sand percentage dominates the silt and clay contents. This is especially true for soils in which the proportion of sand exceeds 60% (A1–A4). Another interesting observation relates to the relatively similar positions of C4 in the soil hydraulic triangle and "clay loam" in the soil texture triangle. However, the similarities between the soil texture triangle and the soil hydraulic triangle are less pronounced for soils where capillary forces have a considerable impact on the water flow. To further analyze the similarities between the soil texture triangle and the soil hydraulic triangle, we compared the relationship of the spatial location of the textural classes and of the optimal hydraulic classes. Table 1 shows the distribution of the ternary space among the soil texture classes and the optimal soil hydraulic classes. In soils where capillary forces are significant, there exists a weaker correlation between the location of the corresponding soil hydraulic and textural classes.

[20] Figure 6 shows the uncertainty in the boundaries of the soil hydraulic triangle shown in Figure 5. The uncertainty in this classification is a testament to the uncertainties of the ROSETTA pedotransfer functions used to develop the data sets. An interesting observation in the uncertainty diagram is that the soil textures that are classified as loam in the soil texture triangle constitute a region of uncertainty for the hydraulic classification.

[21] *Schaap et al.* [2001] estimated the average soil hydraulic parameters for the different soil texture classes using the ROSETTA training database. The average soil hydraulic parameters for different soil texture classes were estimated by averaging of the hydraulic parameters for soils falling under that particular textural class. A similar procedure was applied to the data set used by *Schaap et al.* [2001] to estimate the average soil hydraulic parameters for the different soil hydraulic classes. Table 2 lists the resulting average soil hydraulic parameters for the optimal soil hydraulic classes.

[22] In order to understand how well the soil hydraulic triangle classifies soils compared to the soil texture triangle, we analyzed their performance on the databases of *Schaap et al.* [2001] and *Minasny et al.* [2004]. The soil samples in the two databases were assigned textural classes and hydraulic classes on the basis of the sand, silt, and clay percentages, using the USDA soil texture and the newly developed soil hydraulic triangle. For each soil hydraulic parameter in the two databases, we estimated the cumulative variance of the classifications (CUV) as

$$\mathrm{CUV}(p) = \sum_{i=1}^{k} \left(\gamma_{ip} - \overline{\gamma}_{ip} \right)^2, \qquad (3)$$

where k is the number of hydraulic or textural classes (here, 12), γ_{ip} denotes values of the pth soil hydraulic parameter of the soil samples that fall in the ith hydraulic or textural class, and $\overline{\gamma}_{ip}$ is the mean value of the pth soil hydraulic parameter of the soil samples in the ith hydraulic or textural class. The

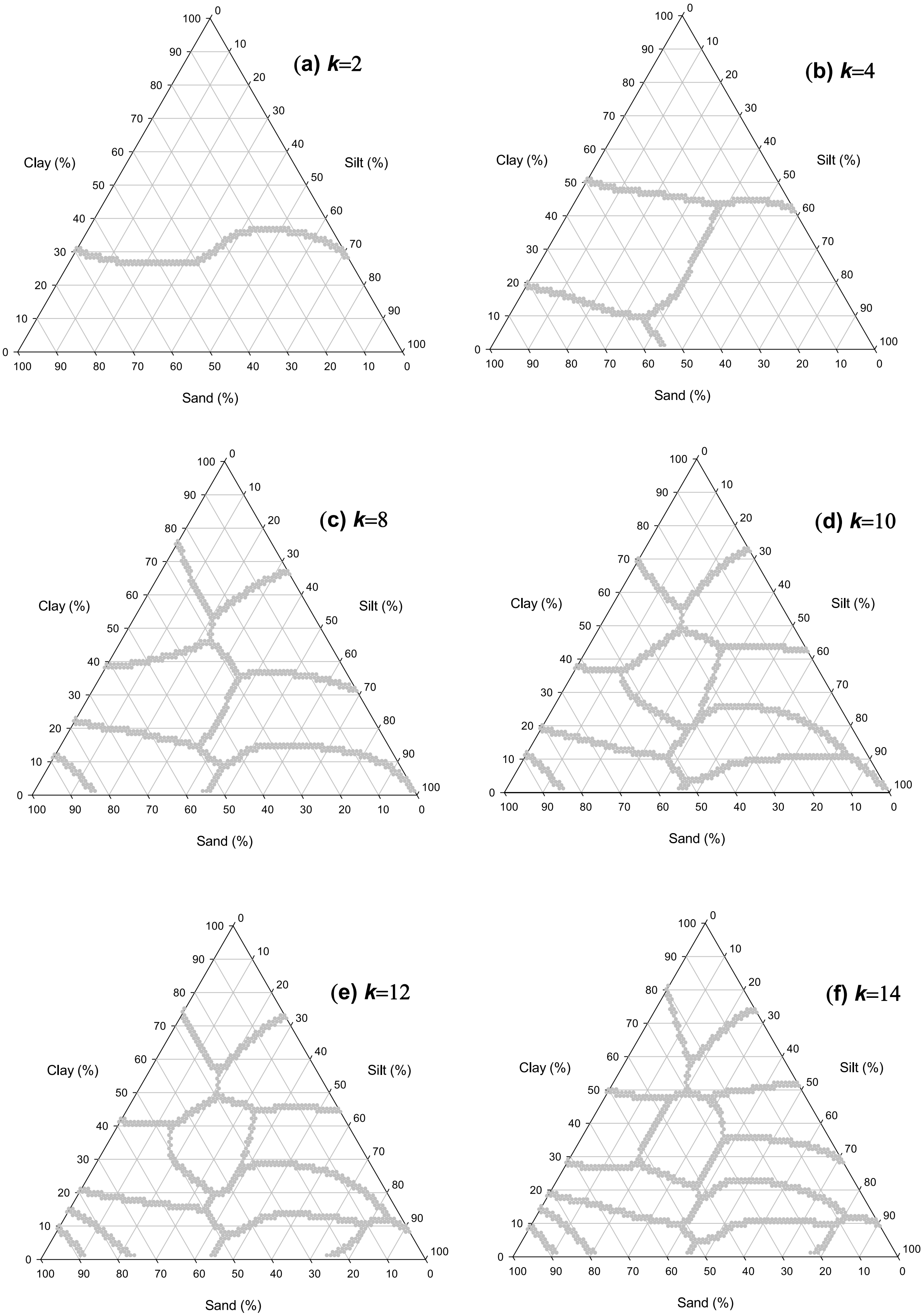

Figure 4. A ternary diagram showing the classification of soils into (a) 2, (b) 4, (c) 8, (d) 10, (e) 12, and (f) 14 classes using the *k*-means clustering algorithm. The boundary was selected as the average boundary from the *k*-means clustering on the 100 data sets.

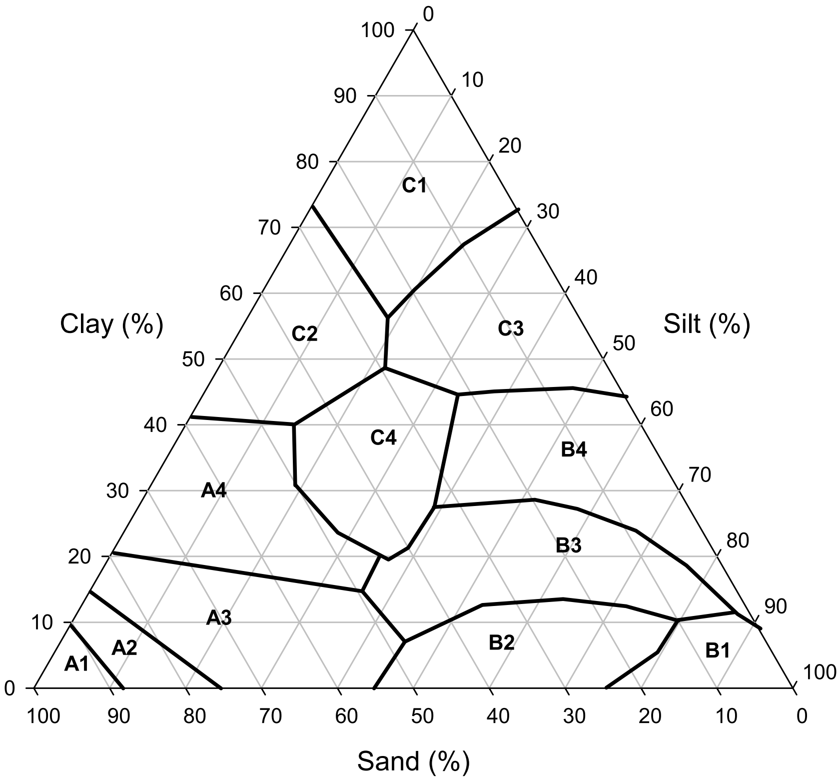

Figure 5. A ternary diagram showing the optimal soil hydraulic triangle with its 12 soil hydraulic classes.

CUV index is very similar to the objective function of the *k*-means clustering algorithm. A lower value for the objective function indicates better classification. Table 3 lists the value of the CUV index for the different soil hydraulic parameters in the two data sets. It was observed that the soil hydraulic triangle performs marginally better than the soil texture triangle in grouping the soil hydraulic parameters for both data sets. It is important to mention that the two databases have data that do not adequately represent the various textural possibilities. The data tend to represent the sand-loam-silt textures more than the clay textures. Given that the soil texture triangle and soil hydraulic triangle are very similar in the sand- and loam-dominated soils, it is not surprising to observe insignificant differences. However, we are not aware of any soil hydraulic databases that represent the various textural possibilities of soils. On the basis of our research, we recognize the lack of a comprehensive soil hydraulic database as a major limitation for the progress of soil sciences, especially in the development of soil classification systems.

[23] For the *Schaap et al.* [2001] data set, the optimal hydraulic and textural classes were also used to group the (1) soil water contents ($\theta(h)$) and (2) soil unsaturated hydraulic conductivity ($\log_{10}K(h)$) for five pressure head classes (in cm) of (0 to -10), (-10 to -100), (-100 to -1000), (-1000 to $-10{,}000$), and ($-10{,}000$ to ∞). For each capillary pressure class, we grouped the $\theta(h)$ and $\log_{10}K(h)$) values by hydraulic or textural class, and the cumulative variance (CUV($\varsigma(H)$)) was estimated as described in equation (4). As

Table 1. Distribution of Textural Possibilities in the Ternary Diagram Among the Soil Hydraulic and Textural Classes[a]

	Sand	Loamy Sand	Sandy Loam	Sandy Clay Loam	Loam	Silt	Silt Loam	Silty Clay Loam	Clay	Sandy Clay	Silty Clay	Clay Loam	Hydraulic Class
A1	1.32												1.32
A2	0.45	2.31											2.76
A3		0.58	9.24		0.54								10.36
A4			1.59	6.6	0.82					1.53			10.54
B1						3.26	0.52						3.78
B2			0.68		0.14		7.16						7.98
B3					4.33		7.05	0.02				0.19	11.59
B4					0.04	0.12	2.27	5.28	0.43		1.63	1.98	11.75
C1									11.82				11.82
C2									7.07	2			9.07
C3									8.08		2.45		10.53
C4				0.85	1.34				2.02	0.16		4.14	8.51
Textural class	1.77	2.89	11.51	7.45	7.21	3.38	17	5.3	29.42	3.69	4.08	6.31	

[a]Values are in percent, representing the percentage of the surface area of the ternary diagram belonging to a particular soil class.

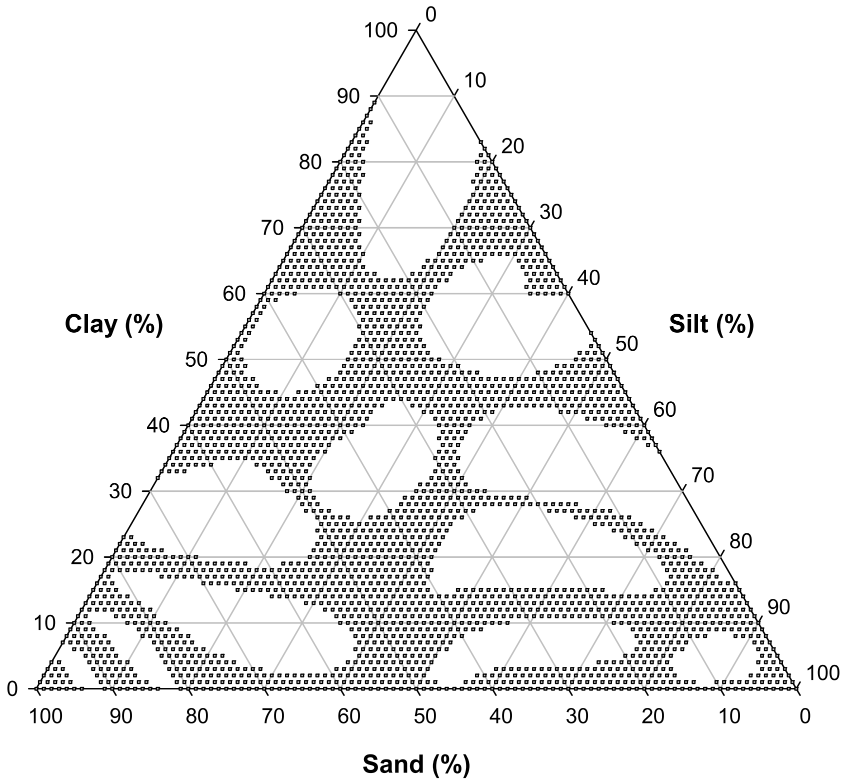

Figure 6. A ternary diagram showing the uncertainty in the optimal soil hydraulic classification shown in Figure 5.

with the objective function of the *k*-means clustering algorithm, a lower value indicates improved clustering:

$$\mathrm{CUV}(\varsigma(H)) = \sum_{i=1}^{k}\left(\varsigma_i(H) - \overline{\varsigma_i(H)}\right)^2. \quad (4)$$

[24] In (4), H (in cm) is one of the pressure head classes, $\varsigma_i(H)$ denotes the $\theta(h)$ (or $\log_{10} K(h)$) values of the ith soil hydraulic or textural class for the pressure head class H, $\overline{\varsigma_i(H)}$ is the mean of $\varsigma_i(H)$, and k is the number of soil hydraulic or textural classes. Table 4 shows the values of the objective function for the different water content and unsaturated hydraulic conductivity groups. Again, the soil hydraulic classes showed a marginal improvement over the textural classes in the classification of soils. Since most of the soils in the database belong to the sand- and loam-dominated textures and texture and hydraulic triangles are very similar in these regions, the observed similarity in the performance of the texture and hydraulic triangles is only applicable to these textures.

5. Summary and Conclusions

[25] Soil classification for hydraulic purposes is an important, but often ignored, topic of interest. In the past, the soil texture triangle has been considered to be the definitive way

Table 2. Mean and Standard Deviation of Soil Hydraulic Parameters for Different Soil Hydraulic Classes Estimated from the *Schaap et al.* [2001] Database[a]

Soil Hydraulic Class	θ_r	θ_S	$\log_{10}(\alpha)$ in $\log_{10}$ (cm^{-1})	$\log_{10}(n)$	$\log_{10}(K_s)$ in $\log_{10}$ (cm/d)
A1	0.055 (0.002)	0.374 (0.008)	−1.479 (0.036)	0.511 (0.06)	2.853 (0.544)
A2	0.053 (0.002)	0.386 (0.007)	−1.474 (0.076)	0.276 (0.055)	2.093 (0.696)
A3	0.051 (0.002)	0.382 (0.01)	−1.54 (0.176)	0.171 (0.015)	1.641 (0.659)
A4	0.055 (0.003)	0.387 (0.009)	−1.672 (0.183)	0.14 (0.019)	1.242 (0.764)
B1	0.057 (0.011)	0.487 (0.031)	−2.034 (0.045)	0.208 (0.009)	1.641 (0.273)
B2	0.053 (0.003)	0.425 (0.03)	−2.278 (0.124)	0.206 (0.024)	1.714 (0.594)
B3	0.056 (0.007)	0.413 (0.026)	−2.234 (0.193)	0.186 (0.021)	1.197 (0.757)
B4	0.073 (0.015)	0.47 (0.026)	−2.016 (0.141)	0.16 (0.024)	1.115 (0.805)
C1	0.072 (0.012)	0.475 (0.013)	−1.94 (0.07)	0.135 (0.01)	1.206 (0.11)
C2	0.091 (0.013)	0.436 (0.025)	−1.444 (0.174)	0.106 (0.008)	1.263 (0.649)
C3	0.069 (0.016)	0.5 (0.037)	−1.888 (0.057)	0.124 (0.01)	1.324 (0.972)
C4	0.064 (0.005)	0.421 (0.012)	−1.83 (0.068)	0.137 (0.017)	0.642 (1.09)

[a]Standard deviations are given in parentheses.

Table 3. Cumulative Variance Estimate for Five Soil Hydraulic Parameters From the *Schaap et al.* [2001] Database and *Minasny et al.* [2004] Database Classified Using the Soil Hydraulic Classification and the USDA Soil Texture Classification[a]

	Schaap et al. [2001]		*Minasny et al.* [2004]	
Parameter	Soil Hydraulic Classification	Soil Textural Classification	Soil Hydraulic Classification	Soil Textural Classification
θ_r	2.052	2.083	1.463	1.5
θ_s	2.401	2.445	0.532	0.536
$\log_{10}(\alpha)$	18.789	18.79	4.509	4.698
$\log_{10}(n)$	4.656	4.714	2.034	2.099
$\log_{10}(K_s)$	26.272	26.988	11.897	11.927

[a]A lower cumulative variance (CUV) estimate indicates better classification.

of classifying soils, even from a hydraulic standpoint. In this research, we have developed a new soil classification that accounts for the hydraulic characteristics of soils. The resulting soil hydraulic triangle has been shown to improve the classification of soils with respect to soil hydraulic characteristics. It was observed that the soil texture triangle is qualitatively very similar to the soil hydraulic triangle, especially for soils where gravity-driven flow dominates capillary forces (such as sands). However, the similarities do not exist for soils where capillary forces dominate the flow through soils. When compared to the soil texture classification, a performance analysis of the classification by soil hydraulic classes on two real-world databases indicated a marginal improvement. Also, the average soil hydraulic parameters for each soil hydraulic class were estimated. We do not consider soil structural effects on soil hydraulic behavior. The results of this research pertain to uniform soils and do not consider other soil structures. The results of this research are therefore conditional on the aforementioned important assumption. From a philosophical perspective, the research further stresses the need to revisit and reevaluate the results from the past in order to successfully move ahead into the future of soil physics. Also, we recognize the lack of a comprehensive soil hydraulic database as a major limitation for the progress of soil sciences, especially in the development of proper soil classification systems.

Appendix A

[26] According to its definition, field capacity (θ_{fc}) is the amount of soil moisture or water content held in soil after excess water has drained away and the rate of downward movement has materially decreased. Several authors [e.g., *Richards and Weaver*, 1944] have suggested suction heads at which soil attains θ_{fc}. However, ambiguity exists as to the accuracy of these suggestions. In this study, we attempt to estimate θ_{fc} using a model that is based on the traditional definition. For the purpose of our analysis, we defined θ_{fc} as the water content at which the drainage from a profile ceases under natural conditions. Since drainage from a soil profile in a simulation never becomes zero, we assume that drainage ceases when the bottom flux from the soil reaches a value that is equivalent to the minimum amount of precipitation that could be recorded. *Dirksen and Matula* [1994] observed that the smallest amount of rainfall measured in meteorological stations was 0.01 cm/d. We used 0.01 cm/d as the value of the bottom flux at which one may assume that θ_{fc} is attained.

[27] The HYDRUS-1D software [*Šimůnek et al.*, 2005] was used to estimate the θ_{fc} and the drainability (time needed to reach θ_{fc} from saturated conditions) under conditions of no evaporation. The HYDRUS-1D employs the classical Richards equation [*Richards*, 1931]. For a one-dimensional scenario, the Richards equation is described mathematically as

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(h)\frac{\partial h}{\partial z} - K(h)\right] - S, \tag{A1}$$

where θ is the volumetric water content (dimensionless), h is the soil water pressure head (L), t is time (T), z is the distance from reference datum (L), and $K(h)$ is the unsaturated hydraulic conductivity ($L\ T^{-1}$) as a function of h or θ.

[28] Given estimates of soil hydraulic parameters (α, n, K_s, θ_s, and θ_r), θ_{fc}, h_{fc}, and t_{fc} were estimated using HYDRUS-1D. In a model setup, a 1 cm long one-dimensional profile was considered. The profile was initially considered to be saturated. A no-flux boundary condition was used at the top of the profile. A free-drainage boundary condition at the bottom of the profile was used throughout the simulation. Under these conditions, HYDRUS-1D was used to simulate the changes in the water content in the profile until the flux at the bottom of the profile reached a value of 0.01 cm/d. The

Table 4. CUV Estimate for Water Contents and Unsaturated Hydraulic Conductivities as a Function of Capillary Pressure From the *Schaap et al.* [2001] Database Classified Using Soil Hydraulic Classification and USDA Soil Texture Classification[a]

	Unsaturated Hydraulic Conductivity $\log_{10}$ (cm/d)			Water Content (cm^3/cm^3)		
Capillary Pressure Head $\log_{10}(h)$ $\log_{10}$ (cm)	Number	Soil Hydraulic Classification	Soil Textural Classification	Number	Soil Hydraulic Classification	Soil Textural Classification
<1	335	1.66	1.75	1326	1271.63	1271.90
1–2	1286	10.35	10.57	6782	6834.90	7037.68
2–3	1969	12.12	12.40	7329	4558.64	4571.83
3–4	434	2.07	2.16	3330	938.39	991.34
≥4	38	0.08	0.08	1860	368.28	380.00

[a]A lower CUV estimate indicates better classification.

water content when the bottom flux reached 0.01 cm/d was assumed to be θ_{fc}. The associated time needed to attain the field capacity from saturation was considered the drainability (t_{fc}), and the pressure head at which the field capacity was estimated was h_{fc}. *Twarakavi et al.* [2009] give a detailed analysis of this approach, and readers are referred to their study for further details.

References

Atterberg, A. (1905), Die rationale Klassifikation der Sande und Kiese, *Chem. Ztg.*, *29*, 195–198.

Atterberg, A. (1912), Die Konsistenz und die Bindigheit der Boden, *Int. Mitt. Bodenkunde.*, *2*, 148–189.

Carsel, R. F., and R. S. Parrish (1988), Developing joint probability distributions of soil water retention characteristics, *Water Resour. Res.*, *24*, 755–769, doi:10.1029/WR024i005p00755.

Davis, R. O. E., and H. H. Bennett (1927), Grouping of soils on the basis of mechanical analysis, *Dep. Circ. 419*, U.S. Dep. of Agric., Washington, D. C.

Dirksen, C., and S. Matula (1994), Automated atomized water spray system for soil hydraulic conductivity measurements, *Soil Sci. Soc. Am. J.*, *58*, 319–325.

Hartigan, J. A., and M. A. Wong (1979), A *k*-means clustering algorithm, *Appl. Stat.*, *28*(1), 100–108, doi:10.2307/2346830.

MacQueen, J. B. (1967), Some methods for classification and analysis of multivariate observations, in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, Univ. of Calif. Press, Berkeley.

Minasny, B., J. W. Hopmans, T. H. Harter, A. M. Tuli, S. O. Eching, and D. A. Denton (2004), Neural network prediction of soil hydraulic functions for alluvial soils using multi-step outflow data, *Soil Sci. Soc. Am. J.*, *68*, 417–429.

National Resources Conservation Service (NRCS) (1975), *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*, *U.S. Dep. Agric. Handb.*, vol. 436, 754 pp., U.S. Gov. Print. Off., Washington, D. C.

Pachepsky, Y. A., and W. J. Rawls (2004), *Development of Pedotransfer Functions in Soil Hydrology*, *Dev. Soil Sci.*, vol. 30, Elsevier, Amsterdam.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxton (1982), Estimation of soil water properties, *Trans. ASAE*, *25*(5), 1316–1320.

Richards, L. A. (1931), Capillary conduction of liquids through porous media, *Physics*, *1*, 318–333, doi:10.1063/1.1745010.

Richards, L. A., and L. R. Weaver (1944), Moisture retention by some irrigated soils as related to soil moisture tension, *J. Agric. Res.*, *69*, 215–235.

Schaap, M. G., and F. J. Leij (1998), Database-related accuracy and uncertainty of pedotransfer functions, *Soil Sci.*, *163*, 765–779, doi:10.1097/00010694-199810000-00001.

Schaap, M. G., F. J. Leij, and M. T. van Genuchten (2001), ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, *J. Hydrol.*, *251*(3–4), 163–176, doi:10.1016/S0022-1694(01)00466-8.

Šimůnek, J., M. T. van Genuchten, and M. Šejna (2005), The HYDRUS-1D software package for simulating the movement of water, heat, and multiple solutes in variably saturated media, version 3.0, *HYDRUS Software Ser. 1*, Dep. of Environ. Sci., Univ. of Calif., Riverside.

Skaggs, T. H., L. M. Arya, P. J. Shouse, and B. P. Mohanty (2001), Estimating particle-size distribution from limited soil texture data, *Soil Sci. Soc. Am. J.*, *65*, 1038–1044.

Soil Survey Staff (1951), *Soil Survey Manual*, *U.S. Dep. Agric. Handb.*, vol. 18, 503 pp., U.S. Gov. Print. Off., Washington, D. C.

Twarakavi, N. K. C., M. Sakai, and J. Šimůnek (2009), An objective study of the dynamic nature of field capacity, *Water Resour. Res.*, *45*, W10410, doi:10.1029/2009WR007944.

van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*(5), 892–898.

Wösten, J. H. M., P. A. Finke, and M. J. W. Jansen (1995), Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics, *Geoderma*, *66*, 227–237, doi:10.1016/0016-7061(94)00079-P.

Wösten, J. H. M., Y. A. Pachepsky, and W. J. Rawls (2001), Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics, *J. Hydrol.*, *251*, 123–150, doi:10.1016/S0022-1694(01)00464-4.

M. G. Schaap, Department of Soil, Water and Environmental Sciences, University of Arizona, Tucson, AZ 85721, USA.

J. Šimůnek, Department of Environmental Sciences, University of California, 2208 Geology Building, Riverside, CA 92521, USA.

N. K. C. Twarakavi, Department of Agronomy and Soils, Auburn University, 258 Funchess Hall, Auburn, AL 36832, USA. (twarakavi@auburn.edu)