ACCURACY AND UNCERTAINTY IN PTF PREDICTIONS

Marcel G. Schaap

GEBJ Salinity Laboratory (USDA/ARS), 450 Big Springs Road, Riverside, CA 92507, USA
Tel.: +1-(909)-369-4844; E-mail:mschaap@ussl.ars.usda.gov

Several statistical criteria are available to evaluate and quantify the potential usefulness of pedotransfer functions (PTFs) in making accurate and reliable estimates of soil hydraulic properties. In this section we will deal with three closely related terms: optimization, accuracy, and uncertainty. Optimization criteria are the objective functions used to calibrate empirical parameters in PTFs. Accuracy criteria relate to the statistical methods used to test the performance of indirect methods, preferably using data distinct from those used for calibration. Uncertainty estimates provide information about the probability distribution of estimated hydraulic quantities. Thus, a PTF can make accurate estimates (i.e., it produces the correct values *on average*), but it may not be reliable when, for example, real-world variability is larger than estimated with the PTF. Alternatively, a PTF can be deemed inaccurate when it produces estimates that, on average, differ systematically from observations. Optimization criteria are the objective functions used to calibrate PTFs. Although optimization criteria are in a strict sense not measures to independently evaluate PTFs, their discussion in this chapter is warranted since optimization and accuracy and reliability criteria are closely related. Because PTF calibration is normally done before evaluation, we will discuss optimization criteria first.

1. OPTIMIZATION CRITERIA

All indirect methods, whether physically based or empirical, contain parameters that must be optimized using databases containing both independent data (predictors) and dependent data (hydraulic properties or parameters in hydraulic functions). Virtually all indirect methods are optimized by least-squares methods, and define the objective function as

$$SSQ = \sum_{N} (\varsigma - \varsigma(\mathbf{b})')^2 \qquad (1)$$

where N is the number of data points and $\mathbf{b}$ is a parameter vector containing the empirical PTF coefficients. For point-based methods ς and ς' represent data points for measured and

estimated water retention or hydraulic conductivity, respectively. In the case of parametric methods, ς and ς' represent parameters of hydraulic functions, such as those in the Brooks–Corey (1964, BC) or van Genuchten (1980, VG) relationships. It is often necessary to transform some of the hydraulic parameters before the optimization of Equation (1) is carried out. For parametric methods it is common to deal with log-transformed values of α, n and λ in the BC or VG equations (e.g., Rawls and Brakensiek (1985) and Wösten et al. (1999), among many others). The general rationale for these transformations is to convert the distribution of the parameters into a more statistically normal distribution (Carsel and Parrish, 1988). For both point and parametric PTFs, log-transforms can also be used in conjunction with Equation (1) to remove a bias towards large conductivity values. For example, depending upon soil type, the saturated conductivity (K_s) can vary by several orders of magnitude. Optimization according to Equation (1) without a log-transform would yield then a PTF that is generally more accurate for highly permeable soils than for low permeability soils.

The optimization of point-based PTFs according to Equation (1) is relatively straightforward. For parametric methods the situation is more complicated since *two* distinct optimizations are necessary. First the parameters in the BC, VG or other hydraulic functions must be optimized using observed water retention and/or conductivity points by minimizing the objective function given by Equation (1). This optimization is only possible if the number of observations exceeds the number of free parameters in the hydraulic function. For this reason, most parametric methods require rather detailed retention or unsaturated conductivity data sets. The actual PTFs are constructed in the second optimization by choosing appropriate functions and minimizing the sum of squares between estimated and previously fitted hydraulic parameters, again using an objective function of the type given by Equation (1).

Given a calibration data set, subtle differences may arise when point- or parametric PTFs are optimized. Being optimized directly to observed – but possibly transformed – water retention or conductivity points, point-PTFs yield optimal results in terms of those observed quantities. Parametric PTFs on the other hand estimate derived quantities (i.e., hydraulic parameters) and not quantities that are directly observed such as retention points. Because the hydraulic functions often do not perfectly fit the hydraulic characteristics, parametric PTFs typically perform slightly worse than point PTFs (Schaap and Bouten, 1996). This fact, combined with the strong non-linearity of the water retention and hydraulic conductivity equations, means that an optimum SSQ in terms of *parameters values* does not guarantee an optimum SSQ in terms of estimated water contents or conductivities.

A solution to this problem is to optimize parametric PTFs directly in terms of the observable quantities rather than optimizing the parameter values. This, in effect, will replace the two-step optimization approach for parametric PTFs with a one-step optimization method. We illustrate this approach using the work of Scheinost et al. (1997) and Minasny et al. (1999) who developed PTFs to estimate VG water retention parameters using "extended non-linear regression." Three steps are necessary in this approach. Firstly, four simple equations (f_1 through f_4) are selected that express the four VG retention parameters (θ_r, θ_s, α, n) in terms of predictors (texture parameters, porosity, organic matter content). Secondly, these equations are substituted in a modified

van Genuchten (1980) equation, thus yielding (Scheinost et al., 1997)

$$\theta(h) = f_1(\text{clay}, C) + \frac{f_2(\text{clay}, \phi) - f_1(\text{clay}, C)}{[1 - \{f_3(d_g)h\}^{f_4(\sigma_g)}]} \qquad (2)$$

The functions f_1 through f_4 are given by

$$f_1 : \theta_r = r_1\text{clay} + r_2 C \qquad (3)$$

$$f_2 : \theta_s = s_1\phi + s_2\text{clay} \qquad (4)$$

$$f_3 : \alpha = a_1 + a_2 d_g \qquad (5)$$

$$f_4 : n = n_1 + n_2/\sigma_g \qquad (6)$$

where $\{r_1, r_2\}$, $\{s_1, s_2\}$, $\{a_1, a_2\}$, and $\{n_1, n_2\}$ are eight unknown parameters associated with the four VG retention parameters. The predictors d_g and σ_g are texture parameters (according to Shirazi et al., 1988), C is organic carbon content and ϕ is the porosity. In the third step, the eight unknown parameters are optimized according to Equation (1).

The extended non-linear regression approach has several advantages. Firstly, no fitted retention parameters are required because f_1 through f_4 are not optimized directly. Therefore, this optimization method can be applied also to databases containing fewer than four points for each retention characteristic. Secondly, this method still estimates VG parameters (f_1 through f_4) thus keeping the flexibility of parametric methods. Thirdly, the method allows the optimization of a parametric PTF directly on observed water contents. A potential drawback is that the number of free parameters must be kept to a minimum otherwise the eight-parameter system is difficult to optimize and may even yield non-unique solutions. This requires that f_1 through f_4 be kept relatively simple, possibly leaving these inadequate for certain conditions. However, Minasny et al. (1999) showed that this is not a major problem as this method provided similar results as a neural network approach. Minasny and McBratney (2002) demonstrated that a similar approach can be used for artificial neural network-based PTFs.

Although Equation (1) is used for the calibration of almost all PTFs, it is possible to use alternative objective functions. When calibrating PTFs, one is often confronted with databases that contain "outlier" points that deviate from the general trend. Outlier points can be caused by several factors such as undocumented measurement errors, or the presence of a few soils that have different characteristics from the larger population (e.g., some fine-textured soils in a larger dataset of coarse textured soils). Such outlier data points may have a large contribution to the sum of squares, and may lead to poor calibrations. Short of deleting outlier points one may use Robust Estimation (Press et al., 1986), which is relatively insensitive to outliers. Instead of minimizing the sum of squares (Eq. (1)), some variants of Robust Estimation minimize the sum of the absolute value of the difference between estimated and observed quantities. We refer to Press et al. (1986) for more information about robust parameter estimation.

2. CRITERIA FOR EVALUATING THE ACCURACY OF PTFs

The accuracy of PTFs is commonly tested using independent data sets. That is, the real value of a PTF can only be evaluated using data that were not used for the calibration of the PTF in question. When a PTF is developed for a particular data set, this PTF should perform well on that data set. However, success on other data sets is not guaranteed as illustrated by Schaap and Leij (1998) who showed that the performance of a PTF depends on the data set used for calibration *and* the one used for evaluation. In many cases, soils databases are organized along local (e.g., case studies), regional, or national boundaries and may have a bias to agricultural soils. Also, current international databases have a serious bias towards soils from temperate climates (Imam et al., 1999). For example, very little data is available for tropical soils (Epebinu and Nwadialo, 1993; Tomasella and Hodnett, 1997). Because it is difficult to define truly representative soils databases, the development, calibration and evaluation of PTFs may lead to somewhat arbitrary results. It is, therefore up to the user of PTFs to decide whether or not a particular PTF – or group of PTFs – is suitable for a particular application. This means that users should somehow obtain a data set and test the PTF using accuracy criteria.

Several studies can be found in the literature that compared and evaluated PTFs (Tietje and Tapkenhinrichs, 1993; Tietje and Hennings, 1995; Kern, 1995). These studies have used different criteria for evaluating the accuracy of indirect methods. Most commonly, criteria such as correlation coefficients, mean errors and root mean square errors are used.

Correlation coefficients (R) or coefficients of determination (R^2) are relatively simple statistics that provide insight in how well data sets of estimated and measured (fitted) hydraulic points (parameters) are related. Although the statistical significance of the value of R or R^2 depends on the size of the data set being tested they give a quick normalized impression of how much of the variance between measured and estimated hydraulic data is explained by the PTF.

Mean errors (ME) and root mean square errors (RMSE) are often used to measure the match of estimated hydraulic properties to such measurable quantities as water contents and hydraulic conductivities. General expressions for ME and RMSE are

$$ME = \frac{1}{N} \sum_N (\varsigma' - \varsigma) \tag{7}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_N (\varsigma' - \varsigma)^2} \tag{8}$$

respectively, where N is the number of datapoints, and – for example – ς' and ς are the estimated and measured water contents or conductivities, respectively. When parametric PTFs are used the estimated parameters are first converted to conductivities or water contents at the pressure head *points* for which measurements are available. Similar to the optimization criteria in the pervious section, ς' and ς may pertain to transformed variables. We note that in the case of log-transformations, the resulting RMSE and ME values will have dimensionless units (Tietje and Hennings, 1996; Schaap and Leij, 2000). When no transformations are applied the RMSE and ME values will have the same units as the measured data.

The ME and RMSE criteria serve several purposes. For example, mean errors quantify systematic errors as they reflect the average deviation between estimated and measured hydraulic data. RMSE values quantify the root of the average bivariate variance between estimated and measured data. RMSE values can thus be interpreted as a random error (not unlike a standard deviation). For a truly well-behaved PTF, both ME and RMSE should be as low as possible. A negative ME indicates that the PTF underestimates the quantity being evaluated, whereas positive values indicate an overestimation. We note, however, that the ME pertains to an *average* over N data points. It may very well be that while the ME is zero the PTF overestimates hydraulic quantities for fine-textured soils and produces overestimates for coarse-textured soils. Likewise in the case of water retention characteristics, it may be that a PTF overestimates water contents near saturation, but underestimates water contents in the dry range (Figure 1). To partially resolve this problem, absolute mean errors (AME) may be computed:
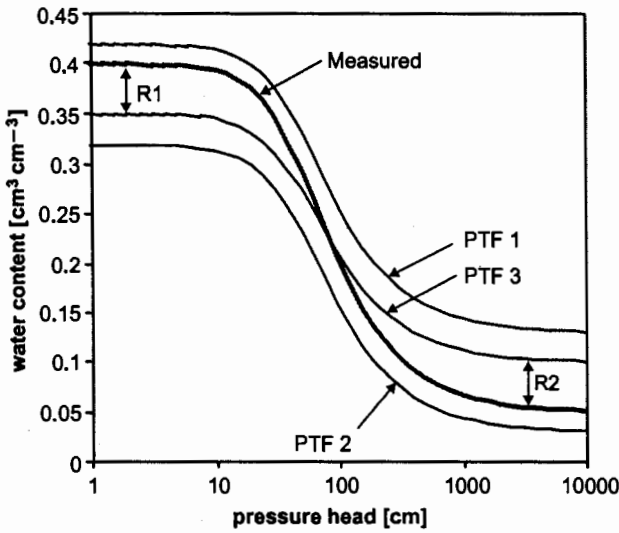

Figure 1. Examples of systematic and relative errors for a typical water retention characteristic. Shown are "measured" data (thick line) and estimates with three hypothetical PTFs (PTF 1, PTF 2, PTF 3). PTF 1 shows an overall overestimation, PTF 2 exhibits an overall underestimation, while PTF3 shows an underestimation near saturation and an overestimation in the dry range. Although the errors produced by PTF 3 at the points R1 and R2 are of similar size (about $0.05 \ cm^3 \ cm^{-3}$), the relative errors at these points are different: 12.5% at R1 and 100% at R2.

$$AME = \frac{1}{N} \sum_{N} |\varsigma' - \varsigma| \qquad (9)$$

AME are always equal to or greater than zero but do not provide information about the sign of the systematic errors.
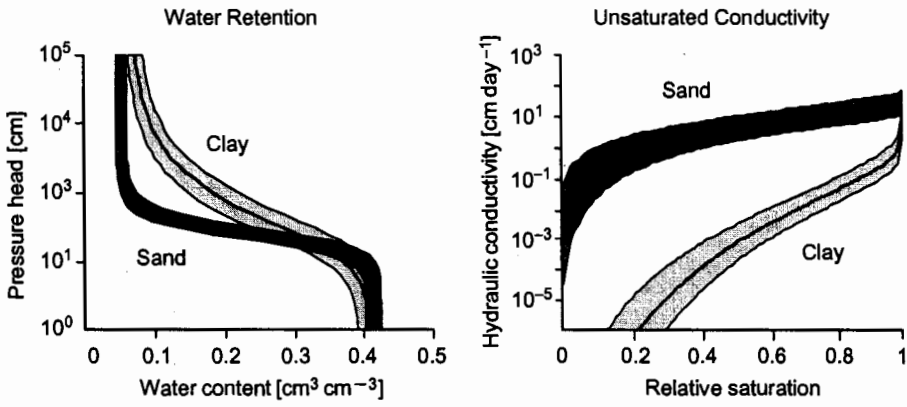
Figure 2. Examples of confidence intervals as generated with the Bootstrap method for water retention (left) and the unsaturated hydraulic conductivity (right). Shown are estimates and 90% confidence intervals for a sand (95% sand, 2.5% clay) and a clay (25% sand, 60% clay); both soils have a bulk density of 1.4 g cm^{-3}.

Occasionally it is useful to provide an impression about the relative size of the systematic errors. For example, a PTF that produces a systematic error of 0.05 cm^3 cm^{-3} for a measured saturated water content of 0.40 cm^3 cm^{-3} leads to a relative error of 12.5% (R1 in Figure 2). However, a PTF that produces the same error for a measured retention point at 0.05 cm^3 cm^{-3} yields a relative error of 100% (R2 in Figure 2). Such errors can be quantified using the relative mean errors (RME) (Williams et al., 1992; Kern, 1995)

$$\text{RME} = \frac{1}{N} \sum_N \frac{\varsigma' - \varsigma}{\varsigma} \tag{10}$$

If mean errors are present then the RMSE given by Equation (8) includes both systematic and random errors. To better separate the systematic from the random errors one may compute unbiased root mean square errors (Tietje and Hennings, 1996)

$$\text{URMSE} = \sqrt{\frac{1}{N} \sum_N (\varsigma' - \varsigma - \text{ME})^2} \tag{11}$$

where the ME is derived from Eq. (7). Similar as ME, URMSE may still include systematic errors for particular parts of the data set.

The error measures discussed thus far are well suited for comparison of parametric PTFs, i.e., PTFs that estimate hydraulic parameters that can be converted to water retention or conductivity points for which measurements are available. A more complicated situation arises when point-PTFs are evaluated and compared since these usually estimate different numbers of water retention points at different pressure heads (Rawls et al., 1982, 1983; Ahuja et al., 1985; Minasny et al., 1999). This may thus lead to the situation where no measured water retention points are available for a particular PTF, or vice versa, no PTF estimate is available for pressure heads for which measured

data are available. The available data is then not used in an optimal way. Also, the error criteria discussed previously above will be evaluated on a different number of points for each PTF (3–12 points are estimated by Rawls et al. (1982, 1983), Ahuja et al. (1985), and Minasny et al. (1999)). Results for different point PTFs may thus be difficult to compare.

Two approaches can be followed to facilitate a more objective evaluation and comparison of point PTFs. Saxton et al. (1986) fitted Brooks–Corey (1964) water retention curves to the estimates of the point-PTFs of Gupta and Larson (1979) and Rawls et al. (1982). Such an approach then allows one to evaluate the PTF as a parametric PTF by computing water contents at pressure heads for which measurements are available. The drawback of this approach is that curve fitting point PTF estimates invariably introduces some random errors and perhaps systematic errors as well. Furthermore, this approach is not suitable for PTFs that estimate a limited number of retention points since curve fitting such data is not possible or leads to unreliable results.

Tietje and Tapkenhinrichs (1993) followed a somewhat related approach in their evaluation of point and parametric PTFs. Instead of curve fitting the estimated retention points, they calculated the estimated water contents at the desired pressure heads by linear interpolation between the estimated retention points (using log-transformed pressure heads). Subsequently, they defined the mean errors and root mean square errors by integration over a certain pressure head range as follows

$$\text{IME} = \frac{1}{b-a} \int_a^b (\varsigma' - \varsigma) \mathrm{d}h \tag{12}$$

$$\text{IRMSE} = \sqrt{\frac{1}{b-a} \int_a^b (\varsigma' - \varsigma)^2 \mathrm{d}h} \tag{13}$$

where the integration interval runs form $a = \log(1\ \text{cm})$ and $b = \log(15,000\ \text{cm})$ and h is the pressure head. The integration was carried out numerically. Instead of evaluating ME and RMSE at pressures for which measurements are available, Equations (12) and (13) compute the normalized area between interpolated measured characteristics and interpolated estimated characteristics. Tietje and Tapkenhinrichs (1993) claim three advantages of their approach: (i) the method is applicable to all PTFs, thus making it easier to compare different PTFs objectively, (ii) no specific function is assumed for the water retention characteristic, and (iii) the method is also possible for a limited measured retention points. We note, however, that although log-linear interpolation is a generally very appropriate choice it *does* make an assumption about the underlying shape of the water retention characteristic, especially when only a few measured or estimated points are available. Secondly, though likewise appropriate the integration interval of Equations (12) and (13) is still somewhat arbitrarily chosen. Thirdly, the amount of information available is limited to the number of measured retention points. Interpolation or integration will never extend this information. It seems that interpolation of estimated points to pressure heads for which measured water contents are available yields similar information that is easier to access since no integration is required.

3. EVALUATING THE UNCERTAINTY OF PTF PREDICTIONS

PTFs generally provide estimates of hydraulic properties with only a modest level of accuracy. Overconfidence in PTF estimates should be avoided because the random or systematic errors made by PTFs may for example lead to inaccurate estimates of contaminant transport through soils and sediments. In some cases it would be worthwhile to use a probabilistic approach that associates an uncertainty to a PTF estimate, such as a confidence interval. Three general approaches are possible.

The first approach is the one followed by Wösten and van Genuchten (1988) who estimated water contents and conductivities at 12 pressure heads for coarse, medium and fine-textured soils. Using measured data they computed RMSE values (standard deviations) at each pressure for the three soil groups, and plotted the 90% confidence interval together with measurements and PTF estimates. The measured and estimated characteristics were generally situated within the confidence intervals, which were between 0.05 and 0.20 $cm^3 cm^{-3}$ for water retention, and between one or two orders of magnitude for the unsaturated conductivity. The advantage of this approach is that it can easily be implemented for any PTF. Disadvantages are that measured hydraulic data must be available and – to obtain reliable statistics – that the data must be grouped into classes (e.g., according to texture, or any other relevant subdivision). Because this approach requires measured data, different data sets may lead to different confidence intervals. We note that confidence intervals with this approach are computed from error criteria defined in Section 2.

The second approach is to quantify parameter uncertainty and co-variance in the PTFs during the calibration process. Many statistical optimization methods allow the analytical or numerical quantification of the variance and co-variance matrices of the optimized parameters (Press et al., 1988). Assuming the underlying model is correct, this type of parameter uncertainty relates to random measurement errors in the actual calibration database. Uncertainty in the calibration parameters subsequently leads to uncertainty in the PTF estimates. Unfortunately, this type of parameter uncertainty is often not reported for PTFs.

A third related approach quantifies parameter uncertainty or the parameter confidence intervals associated with sampling effects in the calibration database. In general, the available calibration database is assumed to be a representative sample of the natural population of soils. In reality, however, the calibration data set is just one sample of this somewhat hard to define population. Calibration of a PTF on another calibration database taken from the same population, will likely lead to a slight variation in the calibration parameters and hence translate into uncertainty in PTF estimates. In most cases, however, only one calibration database making it impossible to directly compute parameter confidence intervals. However, by using Monte Carlo analysis (Press et al., 1986) or the Bootstrap Method (Efron and Tibshirani, 1993) one may gain insight into the parameter confidence intervals without having to collect large quantities of additional data. By using these methods it is also possible to provide uncertainty estimates for each individual PTF estimate (Schaap et al., 2001).

Monte Carlo approaches were used by Carsel and Parrish (1988). In general Monte Carlo approaches involve the generation of many synthetic data sets from a calibrated model. The model is re-calibrated on each synthetic data set and the distributions of the resulting parameters are converted to co-variance matrices. Carsel and Parrish (1988)

developed joint probability distributions of saturated conductivity and van Genuchten (1980) water retention parameters for the 12 USDA soil textural classes. Besides providing uncertainty estimates for the VG parameters for each textural class, Carsel and Parrish (1988) also accounted for correlations among the hydraulic parameters. Results of this type can subsequently be used to generate sets of correlated hydraulic parameters for use in for example Monte-Carlo studies for developing uncertainty estimates of solute transport (Carsel and Parrish, 1988) or Bayesian analyses (Meyer et al., 1999). Minasny et al. (1999) also used a Monte Carlo approach to study the effects of uncertainty in textural data on estimated hydraulic parameters. Depending on how texture was determined (either from imprecise field estimates or more detailed lab measurements) different effects of input data uncertainty were found.

The Bootstrap method is a non-parametric approach that assumes that multiple alternative realizations of a population can be *simulated* from a single (calibration) data set that is available. Thus, contrary to using synthetic data such as the Monte Carlo method, real data are used that is re-ordered in new datasets. The Bootstrap method accomplishes this by repeated random resampling with replacement of the original data set of size N to obtain B alternative data sets, also of size N. The PTF is calibrated B times and, because of slight variations in the alternative data sets, the bootstrap method, therefore leads to B alternative PTFs, resulting in a distribution of estimated hydraulic parameters. These distributions can be further manipulated for confidence intervals and co-variances. The non-parametric nature of the Bootstrap method allows it to be used for a wider range of PTFs than Monte Carlo analysis, such as neural networks which do not have easily identifiable calibration parameters.

Figure 2 shows confidence intervals as generated with the Bootstrap method. Shown are estimated retention and unsaturated hydraulic conductivity characteristics for a coarse textured soil (a sand with 95% sand, 2.5% clay, bulk density $1.4 \, g \, cm^{-3}$) and a fine textured soil (a clay soil with 25% sand, 60% clay, bulk density $1.4 \, g \, cm^{-3}$). The estimated characteristics are surrounded by an area that delineates the 90% confidence intervals. For water retention the 90% confidence area for the sand is smaller than that for the clay, indicating that the PTF estimate for the sand is more reliable ($0.03-0.05 \, cm^3 \, cm^{-3}$ for the sand and $0.05-0.10 \, cm^3 \, cm^{-3}$ for the clay). For unsaturated hydraulic conductivity there is little difference between the confidence intervals for the clay and the sand; in both cases the confidence interval is approximately one to two orders of magnitude. We note that these confidence intervals can be generated without external data and are specific for each individual estimate (i.e., the confidence intervals change with the estimated parameters).

REFERENCES

Ahuja, L.R., Naney, J.W., Williams, R.D., 1985. Estimating soil water characteristics from simpler properties and limited data. Soil Sci. Soc. Am. J. 49, 1100-1105.

Brooks, R.H., Corey, A.T., 1964. Hydraulic Properties of Porous Media. Hydrol. Paper 3. Colorado State University, Fort Collins, CO.

Carsel, R.F., Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. Water Resour. Res. 24, 755-769.

Efron, B., Tibshirani, R.J., 1993. An introduction to the bootstrap. Monographs on Statistics and Applied Probability. Chapman and Hall, New York, NY.

Epebinu, O., Nwadialo, B., 1993. Predicting soil water availability from texture and organic matter content for Nigerian soils. Commun. Soil Sci. Plant Anal. 24, 633-640.

Imam, B., Sorooshian, S., Mayr, T., Schaap, M.G., Wősten, H., Scholes, B., 1999. Comparison of pedotransfer functions to compute water holding capacity using the van Genuchten model in inorganic soils. Report to the IGBP-DIS soils data tasks, IGBP-DIS working paper #22. IGBP-DIS office, CNRM, 42 avenue G. Coriolis, 31057 Toulouse Cedex, France.

Kern, J.S., 1995. Evaluation of soil water retention models based on basic soil physical properties. Soil Sci. Soc. Am. J. 59, 1134-1141.

Meyer, P.D., Gee, G.W., Rockhold, M.L., Schaap, M.G., 1999. Characterization of soil hydraulic parameter uncertainty. *In*: van Genuchten, M.Th., Leij, F.J., Wu, L. (Eds.), Proceedings of International Workshop, Characterization and Measurements of the Hydraulic Properties of Unsaturated Porous Media. University of California, Riverside, CA, pp. 1439-1451.

Minasny, B., McBratney, A.B., 2002. The neuro-m method for fitting neural network parametric pedotransfer functions. Soil Sci. Soc. Am. J. 66, 352-362.

Minasny, B., McBratney, A.B., Bristow, K.L., 1999. Comparison of different approaches to the development of pedotransfer functions of water retention curves. Geoderma 93, 225-253.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1986. Numerical recipes. The Art of Scientific Computing. Cambridge University Press, Cambridge.

Rawls, W.J., Brakensiek, D.L., 1985. Prediction of soil water properties for hydrologic modeling. *In*: Jones, E.B., Ward, T.J. (Eds.), Watershed management in the eighties. Proc. Irrig. Drain. Div., ASCE, Denver, CO. April 30–May 1, 1985, pp. 293-299.

Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of soil water properties. Trans. ASAE 25, 1316-1320.

Rawls, W.J., Brakensiek, D.L., Soni, B., 1983. Agricultural management effects on soil water processes, part I: soil water retention and Green and Ampt infiltration parameters. Trans. ASAE 26, 1747-1752.

Saxton, K.E., Rawls, W.J., Romberger, J.S., Papendick, R.I., 1986. Estimating generalized soil-water characteristics from texture. Soil Sci. Soc. Am. J. 50, 1031-1036.

Schaap, M.G., Bouten, W., 1996. Modeling water retention curves of sandy soils using neural networks. Water Resour. Res. 32, 3033-3040.

Schaap, M.G., Leij, F.J., 1998. Database related accuracy and uncertainty of pedotransfer functions. Soil Sci. 163, 765-779.

Schaap, M.G., Leij, F.J., 2000. Improved prediction of unsaturated hydraulic conductivity with the Mualem–van Genuchten model. Soil Sci. Soc. Am. J. 64, 843-851.

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251, 163-176.

Scheinost, A.C., Sinowski, W., Auerswald, K., 1997. Regionalization of soil water retention curves in a highly variable landscape, I. Developing a new pedotransfer function. Geoderma 78, 129-143.

Shirazi, M.A., Boersma, L., Hart, J.W., 1988. A unifying quantitative analysis of soil texture: improvement of precision and extension of scale. Soil Sci. Soc. Am. J. 52, 181-190.

Tietje, O., Hennings, V., 1996. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. Geoderma 69, 71-84.

Tietje, O., Tapkenhinrichs, M., 1993. Evaluation of pedotransfer functions. Soil Sci. Soc. Am. J. 57, 1088-1095.

Tomasella, J., Hodnett, M.G., 1997. Estimating unsaturated hydraulic conductivity of Brazilian soils using soil-water retention data. Soil Sci. 162, 703-712.

van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Am. J. 44, 892-898.

Williams, R.D., Ahuja, L.R., Naney, J.W., 1992. Comparison of methods to estimate soil water characteristics from limited texture, bulk density, and limited data. Soil Sci. 153, 172-184.

Wösten, J.H.M., van Genuchten, M.Th., 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. Soil Sci. Soc. Am. J. 52, 1762-1770.

Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90, 169-185.