

**Soil Science**

Issue: Volume 163(10), October 1998, pp 765-779

Copyright: © 1998 Wolters Kluwer Health | Lippincott Williams & Wilkins

Publication Type: [Technical Articles]

ISSN: 0038-075X

Accession: 00010694-199810000-00001

Keywords: Water retention, hydraulic conductivity, prediction, uncertainty estimates, neural networks

[Technical Articles]

## **DATABASE-RELATED ACCURACY AND UNCERTAINTY OF PEDOTRANSFER FUNCTIONS**

Schaap, Marcel G.; Leij, Feike J.

### **Author Information**

USDA-ARS, U.S. Salinity Laboratory, 450 West Big Springs Rd., Riverside, CA 92507-4617. Dr. Schaap is corresponding author. E-mail:

[mschaap@ussl.ars.usda.gov](mailto:mschaap@ussl.ars.usda.gov)

Received Jan. 28, 1998; accepted June 10, 1998.

### **Abstract**

Pedotransfer functions (PTFs) are becoming a more common way to predict soil hydraulic properties from soil texture, bulk density, and organic matter content. Thus far, the calibration and validation of PTFs has been hampered by a lack of suitable databases. In this paper we employed three databases (RAWLS, AHUJA, and UNSODA) to evaluate the accuracy and uncertainty of neural network-based PTFs. Sand, silt, and clay percentages and bulk density were used as input for the PTFs, which subsequently provided retention parameters and saturated hydraulic conductivity,  $K_s$  as output. Calibration and validation of PTFs were carried out on independent samples from the same database through combination with the bootstrap method. This method also yielded the possibility of calculating uncertainty estimates of predicted hydraulic parameters. Calibration and validation results showed that water retention could be predicted with a root mean square residual (RMSR) between 0.06 and 0.10  $\text{cm}^3 \text{ cm}^{-3}$ , the RMSR of  $\log(K_s)$  was between 0.4 and 0.7  $\log(\text{cm day}^{-1})$ . Cross-validation was used to test how well PTFs that were calibrated for one database could predict the hydraulic properties of the other two databases. The results showed that systematically different predictions were made when the RMSR values increased to between 0.08 and 0.13  $\text{cm}^3 \text{ cm}^{-3}$  for water retention and to between 0.6 and 0.9  $\log(\text{cm day}^{-1})$  for  $\log(K_s)$ . The uncertainty in predicted  $K_s$  was one-half to one order of magnitude, whereas predicted water retention points had an uncertainty of about 0.04 to 0.10  $\text{cm}^3 \text{ cm}^{-3}$ . Uncertainties became somewhat smaller if the PTFs were calibrated on all available data. We conclude that the performance of PTFs may depend strongly on the data that were used for calibration and evaluation.

Pedotransfer functions (PTFs) are often used to estimate soil hydraulic properties when direct measurements are too expensive or cumbersome. PTFs transform soil properties such as texture, bulk density, and organic matter content (among others) into water retention and saturated or unsaturated hydraulic conductivity. Many different types of PTFs have been developed (cf [Rawls et al. 1991](#)), but all need to be calibrated on existing data before true predictions can be made. Consequently, the utility of a PTF can only be established objectively on independent data, i.e., data not used for calibration. In past years, several such evaluations have been carried out. [Tietje and Tapkenhinrichs \(1993\)](#) tested 13 PTFs for water retention using a German database of 1079 samples. They found that the PTF of [Vereecken et al. \(1989\)](#) was applicable to many soils and provided relatively low root mean square residuals ( $0.05 \text{ cm}^3 \text{ cm}^{-3}$ ), but it gave a systematic underestimation of water contents. In another study encompassing more than 25,000 samples, [Kern \(1995\)](#) evaluated six water retention PTFs. Results of this study show that water contents at pressure heads of -10, -33, and -1500 kPa were often over- or underestimated. [Tietje and Hennings \(1996\)](#) used a German data set of 1161 samples to test six PTFs to predict  $K_s$ . None of the tested PTFs had a consistently better prediction for all textural classes. Using a dataset of 1209 samples, [Schaap et al. \(1998\)](#) found that neural network-based PTFs usually provided better predictions than PTFs published by [Rawls and Brakensiek \(1985\)](#), [Vereecken et al. \(1989\)](#), [Cosby et al. \(1984\)](#), [Brakensiek et al. \(1984\)](#), and [Saxton et al. \(1986\)](#).

Although some PTFs perform better than others, there seem to be no clearly superior and generally applicable PTFs. There are several reasons for this. First, PTFs require different input data and predict a variety of hydraulic properties (e.g., water retention points or parameters of water retention functions). Second, the distribution of hydraulic and basic soil data in a calibration data set often differs from the distribution of data in an independent validation set because of geographical or other practical constraints. It is possible that if the calibration-validation sequence was reversed (i.e., calibration on the validation data set and testing on the calibration data set), entirely different PTFs would have been found. As a result, a PTF may yield good or bad evaluation results, depending on the data used. A third reason may be that the hydraulic parameters used for calibration and validation may have been established using different methods, which may lead to PTFs that make systematically different predictions. Finally, PTFs may yield different results because of an inherent uncertainty in their predictions. A PTF is commonly calibrated on a data set that is just one realization of a population of soils. A slightly different PTF would result if another realization was available for calibration. As a result, PTF predictions have an associated uncertainty, which, if quantifiable, would provide valuable information about the reliability of the PTF.

The issues mentioned above may cause one to question the applicability and reliability of PTF predictions. The first objective of this paper is, therefore, to test systematically how well predictions of PTFs that are calibrated on three different databases agree with each other. The second objective is to establish the confidence intervals that are associated with the predictions.

To meet these objectives we developed neural network-based PTFs to predict parameters in the retention function of [van Genuchten \(1980\)](#) and saturated hydraulic conductivity,  $K_s$ . As input for these PTFs, we used sand, silt, and clay percentages and bulk density. Neural network models were used because, contrary to more traditional PTFs, they require no *a priori* model concept and provide the best possible (nonlinear) mapping from input data to output data ([Hecht-Nielsen 1991](#); [Schaap and Bouten 1996](#)). Neural networks thus serve as benchmarks that indicate how well hydraulic parameters can be predicted from basic soil properties for a certain data set.

The PTFs were developed in three stages: calibration, validation, and cross-validation. Calibration and validation were carried out on two independent randomized data sets of the same database generated with the bootstrap method ([Efron and Tibshirani 1993](#)). The bootstrap method further yielded confidence intervals of each PTF prediction. In the third cross-validation stage, the calibrated models were tested on data from two different databases. Finally, we developed PTFs calibrated and validated using data from all three databases.

## MATERIALS AND METHODS

### Data Sets

Three databases, RAWLS, AHUJA, and UNSODA, were used in this study. All three databases were heterogeneous: they consisted of data from different sources and different measurement procedures; essential variables were sometimes missing, and water retention data were available at various levels of detail. For each sample, we required the availability of sand, silt, and clay percentages and bulk density. To obtain reliable fits of the [van Genuchten \(1980\)](#) curve, we needed at least six water retention points; samples with fewer than six points or clearly chaotic retention data were discarded.

Figs 1A-C show the textural distribution of the selected samples in each database.

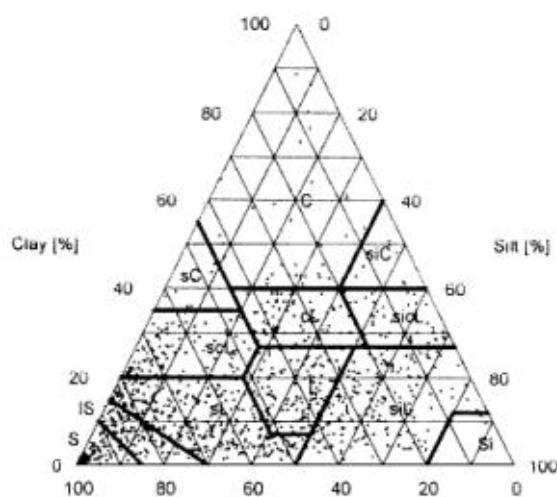
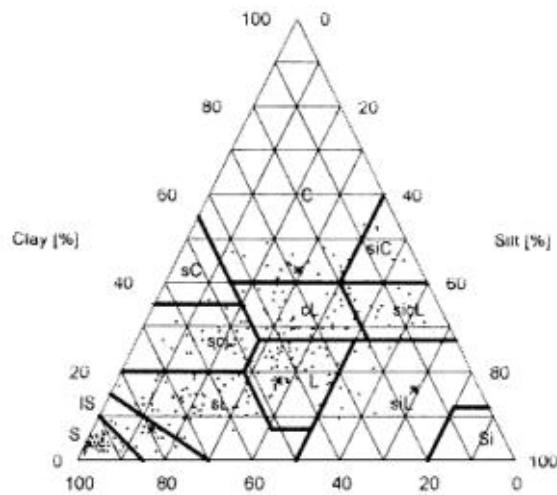
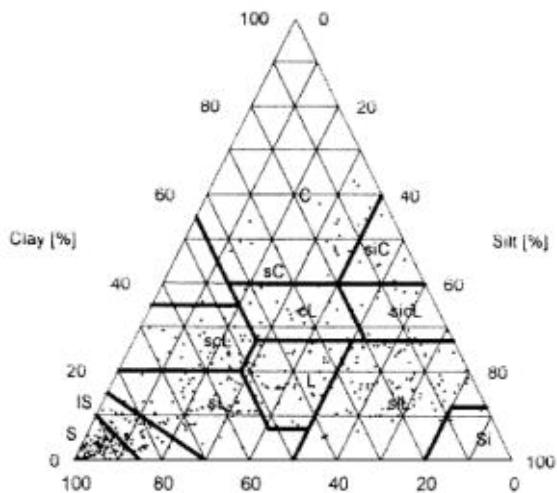
**RAWLS****A****Sand [%]****AHUJA****B****Sand [%]****UNSODA****C****Sand [%]**

Fig. 1. Textural distributions for the RAWLS (A), AHUJA (B), and UNSODA (C) data sets (S: sand, C: clay, Si: silt, L: loam, s: sandy, c: clayey, si: silty, l: loamy). The symbols in Fig. 1B represent the textural distributions of the four samples used in Figs 3, 4, and 5.

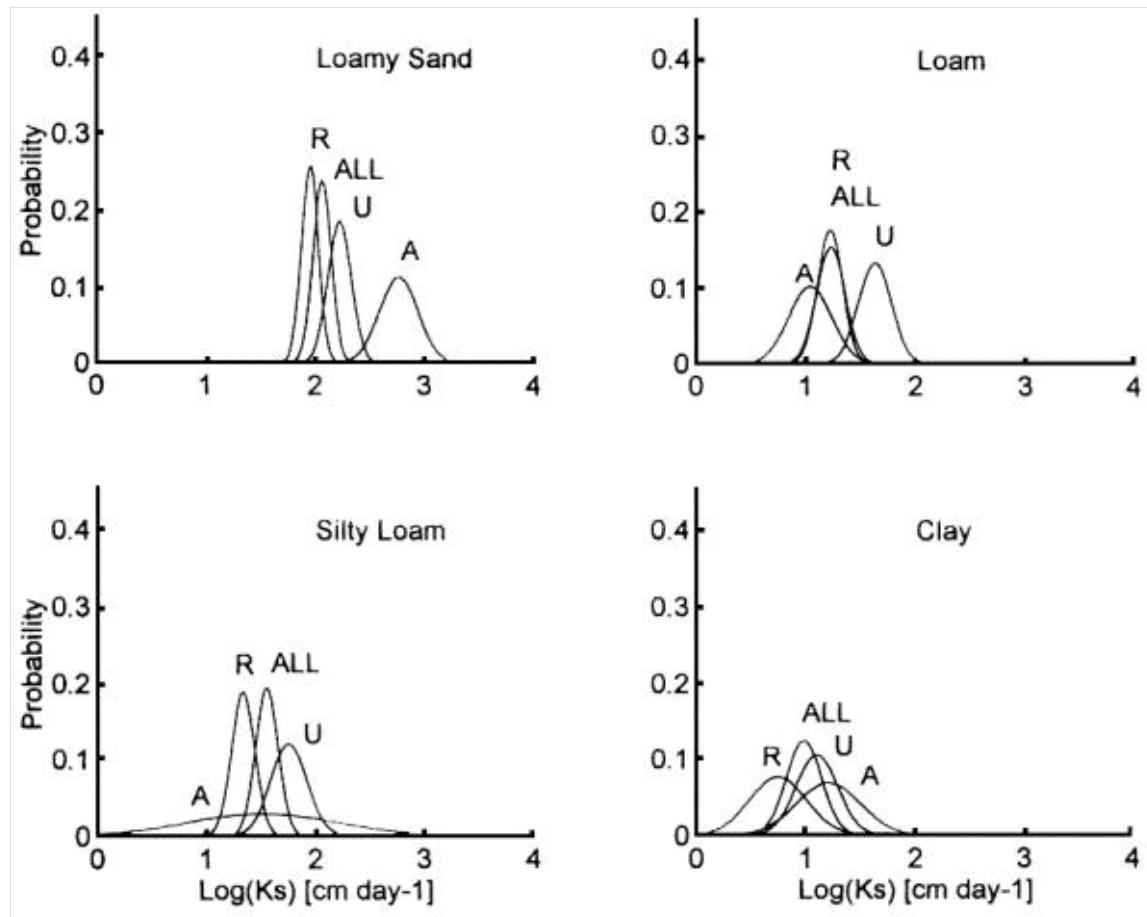


Fig. 3. Probability distributions of  $\log(K_s)$  as predicted by neural network models calibrated on the four data sets (R: RAWLS, A: AHUJA, U: UNSODA, ALL: all available data).

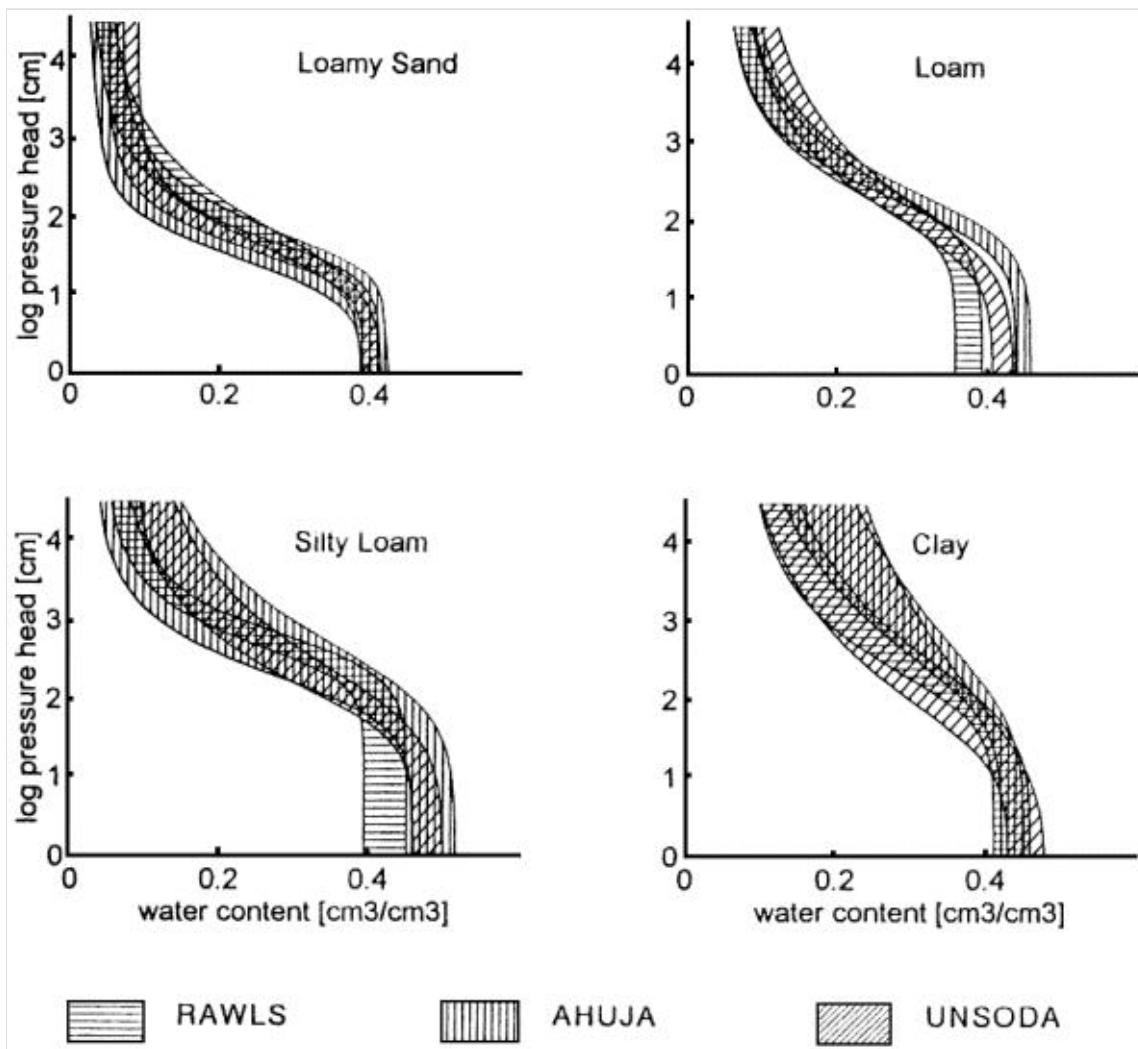


Fig. 4. Confidence intervals for water retention as predicted by neural network models calibrated on the three data sets

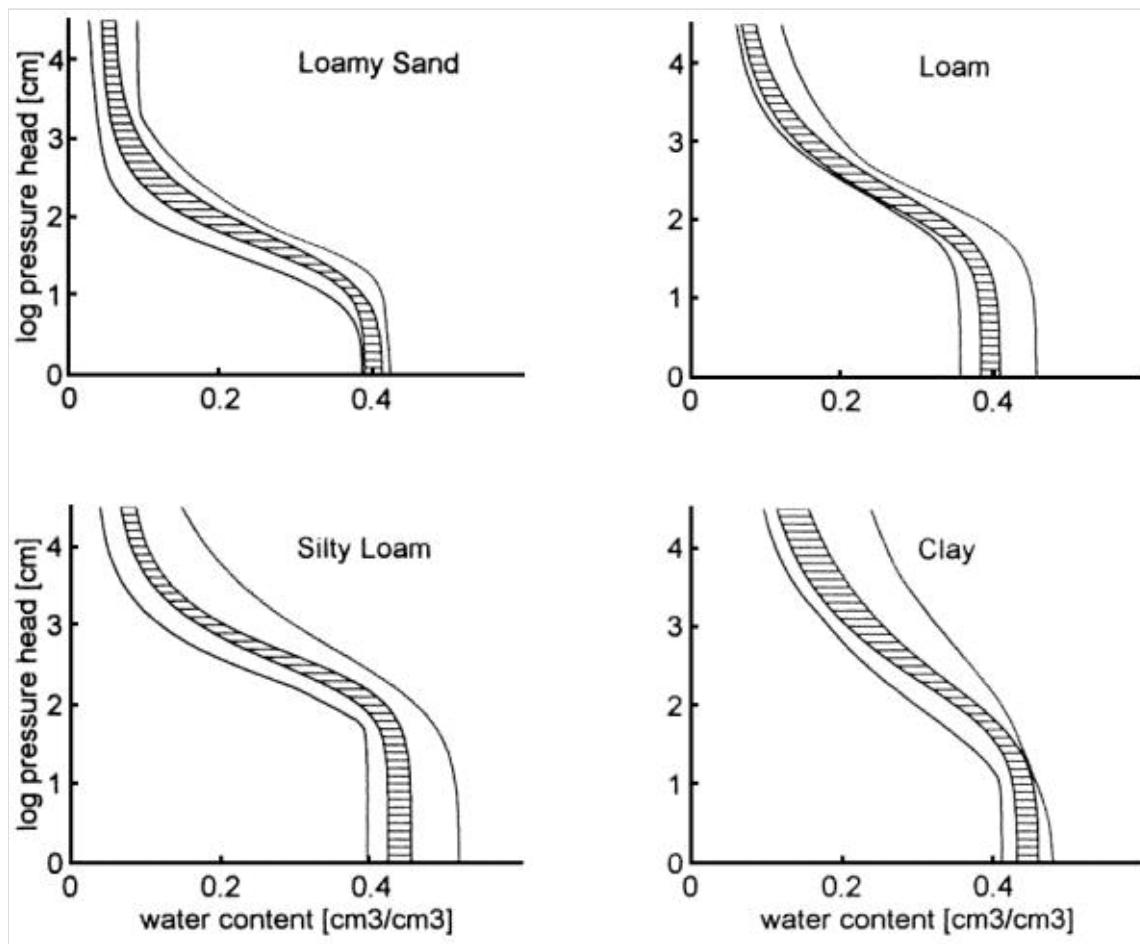


Fig. 5. Confidence intervals for water retention as predicted by neural network models calibrated on all available data (ALL).

The RAWLS database ([Fig. 1A](#)) was compiled from about 30 sources in the United States (W.J. Rawls, USDA-ARS Hydrology lab., Beltsville, MD, 1996, personal communication). The complete database contains the following information on 4515 samples: source of data (publication), state, county, soil series name, slope, land use, soil horizon, depth of sampling, structure (grade, size, type), % sand, % silt, % clay, % organic carbon, % gravel, bulk density, total porosity,  $K_s$ , and four to 12 soil water retention points at various pressure heads. We selected 1209 samples for water retention, of which 620 had measured  $K_s$  values. The same data were used by [Schaap et al. \(1998\)](#).

The AHUJA database (L.R. Ahuja, USDA-ARS Great Plains Systems Research Unit, Fort Collins, CO, 1996, personal communication) contains 393 soil samples with a sandy to clay texture ([Fig. 1B](#)). The data consist of: silt and clay percentages, bulk density, porosity,  $K_s$ , and up to 13 water retention points. We selected 371 samples for water retention and  $K_s$ .

The UNSODA database ([Fig. 1C](#)) was compiled by [Leij et al. \(1996\)](#) from many international sources. UNSODA consists of 791 entries, with water retention and saturated and unsaturated hydraulic conductivity data as well as particle-size distribution and bulk density. We selected 554 samples for water retention data, of which 315 had  $K_s$  data.

In addition to considering the databases separately, all selected samples were pooled into two data sets containing 2134 samples for water retention and 1306 samples for  $K_s$ . These data will be addressed as ALL.

## Hydraulic Parameters

The hydraulic parameters to be predicted consist of  $K_s$  and four parameters of the water retention equation by [van Genuchten \(1980\)](#): [Equation \(1\)](#) where  $h$  is the pressure head (cm),  $[\theta_r]$  and  $[\theta_s]$  are residual and saturated water contents in  $\text{cm}^3 \text{ cm}^{-3}$ ,  $[\alpha]$  ( $\text{cm}^{-1}$ ), and  $n$  are curve-shape parameters. [Equation \(1\)](#) was fitted to measured data with the Simplex or Amoeba algorithm ([Nelder and Mead 1965](#); [Press et al. 1988](#)). The fits were subject to the following constraints:  $0.0 \leq [\theta_r] \leq 0.3 \text{ cm}^3 \text{ cm}^{-3}$ ;  $0.6[\phi] \leq [\theta_s] \leq [\phi] \text{ cm}^3 \text{ cm}^{-3}$  (where  $[\phi]$  is the total porosity);  $0.0001 \leq [\alpha] \leq 1.000 \text{ cm}^{-1}$ ;  $1.001 \leq n \leq 10$ . Logarithms of  $[\alpha]$ ,  $n$ , and  $K_s$  parameters were used in the presented analyses.

$$\theta(h) = \theta_r + \frac{(\theta_s - \theta_r)}{[1 + (\alpha h)^n]^m} \quad (m = 1 - 1/n)$$

Equation 1

## Neural Networks

The term neural network relates to a large collection of numerical techniques that resemble biological neural systems. Many different types of neural network exist, each with a particular range of applications ([Hecht-Nielsen 1991](#); [Haykin 1994](#)). Feed-forward backpropagation networks and radial basis functions have been used previously to predict soil hydraulic properties ([Pachepsky et al 1996](#); [Schaap et al. 1998](#); [Tamari et al. 1996](#)). In this paper we chose feed-forward neural networks because they are more easily combined with the bootstrap method and more efficient to calibrate.

Feed-forward neural networks consist of an input, a hidden, and an output layer, all of which contain nodes ([Fig. 2](#)). The number of nodes in input and output layers corresponds to the number of input and output variables of the model. Following [Schaap and Bouting \(1996\)](#), we used six hidden nodes. All input nodes  $j = 1..J$ , with the input variables  $x_1..x_J$ , are connected to all hidden layer nodes  $k = 1..K$  by weights,  $w_{jk}$ . At the hidden nodes, the input values and weights are multiplied and summed according to [Eq. \(2\)](#). Additionally, a bias value,  $x_0$ , (equal to 1) and weights,  $w_{0k}$ , are used to offset the sum,  $S_k$ . The hidden node output,  $H_k$ , is calculated with [Eq. \(3\)](#). [Equations \(2\) and \(3\)](#)

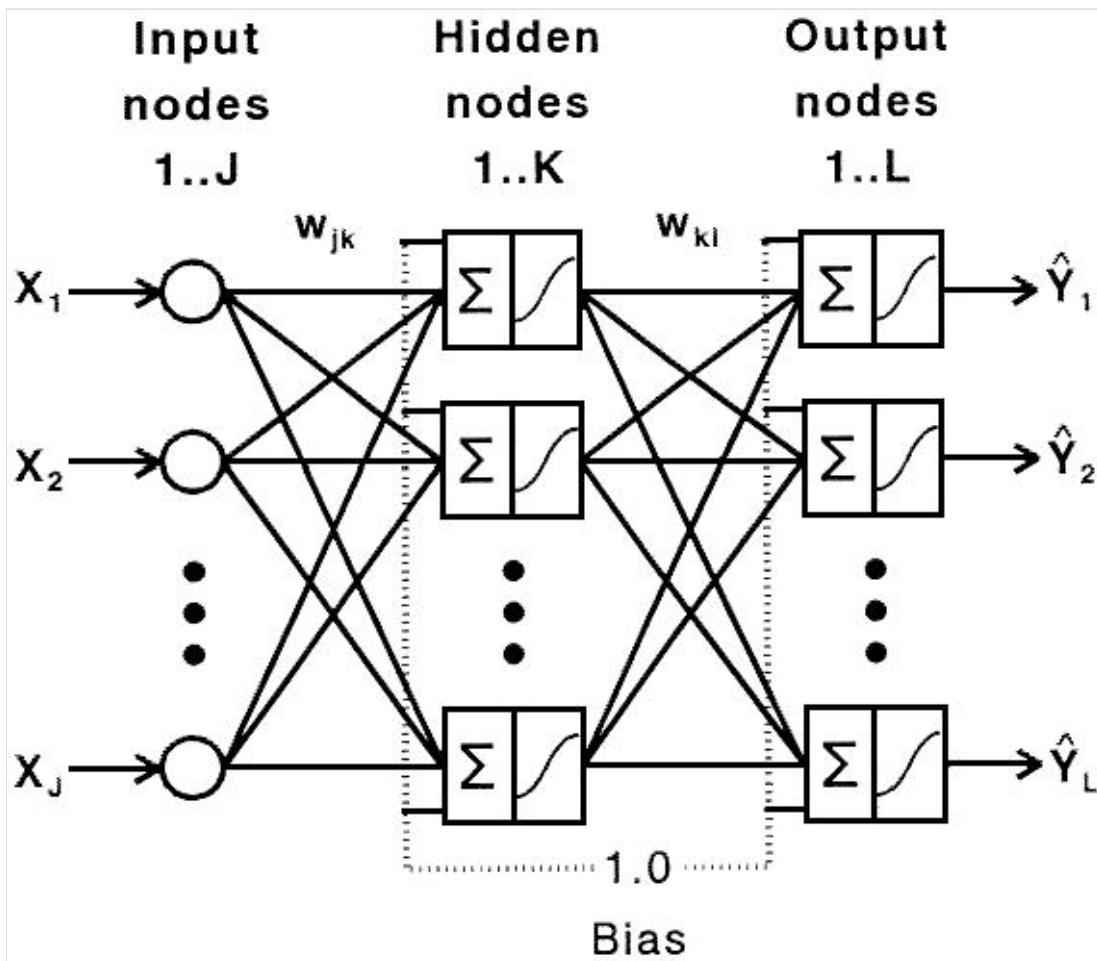


Fig. 2. Schematic overview of the neural network used in this paper.

$$S_k = \sum_{j=0}^J (w_{jk} * \chi_j)$$

Equation 2

$$H_k = \frac{1}{1 + e^{-S_k}}$$

Equation 3

The output nodes  $I = 1..L$  operate in the same way as the hidden nodes. The hidden node outputs,  $H_k$ , are multiplied by the weights  $w_{kl}$ , and the model outputs ( $\hat{Y}_1$ ) are produced in the same way as in  $H_k$  in Eq. (3). The weight matrices  $w_{jk}$  and  $w_{kl}$  are obtained with a calibration procedure based on the Levenberg-Marquardt algorithm (Marquardt 1963; Demuth and Beale 1992). The following objective function is minimized for a data set of  $N_c$  samples: Equation (4) where  $t$  and  $t'$  are the measured and the predicted vectors of hydraulic parameters;  $N_p$  is the number of parameters, which is equal to 1 for  $K_s$  and 4 for the water retention parameters.

$$Ob = \sum_{i=1}^{N_c} \sum_{j=1}^{N_p} (t_{i,j} - t'_{i,j})^2$$

Equation 4

### Calibration, Validation, and Cross-Validation

Two types of neural network were calibrated for the four available databases: one type to predict water retention parameters and the other to predict  $K_s$ . For all neural networks, we used sand, silt, and clay percentages and bulk density as input data.

The bootstrap method ([Efron and Tibshirani, 1993](#)) was essential to obtain independent calibration and validation sets and to calculate confidence intervals of the neural network predictions. The theory behind the bootstrap assumes that multiple alternative realizations of the population can be simulated from the single data set that is available. This is done by repeated random resampling, with replacement of the original data set of size  $N$  to obtain  $B$  alternative data sets, also of size  $N$ . Because the resampling is done with replacement, each sample has a chance of  $1-[(N-1)/N]^N$  to be selected once or multiple times for a particular alternative dataset. These  $B$  alternative data sets, or subsets, therefore, each contain about 63% (for  $N > 50$ ) of the original data. On each subset, a neural network was calibrated leading to a submodel that was subsequently validated on the approximately 37% of the data that were not selected for the calibration subset. [Efron and Tibshirani \(1993\)](#) suggested  $B$  between 50 and 200. We used  $B = 60$  for water retention parameters and  $B = 100$  for  $K_s$ . Confidence intervals were calculating using the standard deviation of the predictions of the  $B$  submodels.

Calibration and validation were performed for each database separately as well as for the total of the data in the ALL data set. Cross-validation was carried out by testing all  $B$  submodels obtained for the RAWLS databases on the AHUJA and UNSODA databases and vice versa. Because there were three databases with corresponding neural network models, three calibration, three validations, and six cross-validations were possible. Cross-validation was not possible for the ALL dataset because it contained all available data.

The combined neural network-bootstrap analysis was carried out with a slightly adapted TRAINLM routine of the neural network toolbox (version 2.0) of the MATLAB®[1](#) package (version 4.0, Math Works Inc., Natick, MA) with additional code to avoid local minima in the error surface of [Eq. \(4\)](#).

### Evaluation Criteria

We used three error criteria to evaluate calibration, validation, and cross-validation results against measured data. The coefficient of determination ( $R^2$ ) indicates how well predicted hydraulic parameters ( $[\theta]_r$ ,  $[\theta]_s$ ,  $\alpha$ ,  $n$ ,  $K_s$ ) matched trends in measured parameters. Bias in predicted hydraulic parameters (i.e., a systematic difference between measured and predicted parameters) was quantified by the mean error (ME): [Equation \(5\)](#) where  $N$  is the number of samples;  $[\zeta]$  is a measured or fitted hydraulic parameter while  $[\zeta']$  is a predicted hydraulic parameter. In the case of  $\alpha$ ,  $n$  and  $K_s$ , logarithmic values were used for  $[\zeta]$  and  $[\zeta']$ . Negative values of  $ME_{[\zeta]}$  indicate that predicted parameters are, on average, larger than measured parameters.

$$ME_{\zeta} = \frac{1}{N} \sum_{j=1}^N (\zeta_j - \zeta'_j)$$

Equation 5

The ability of the predicted hydraulic parameters to match measured retention or  $K_s$  data was evaluated by root mean square residuals ( $RMSR_{[\theta]}$  and  $RMSR_{K_s}$ , for water retention and  $K_s$  respectively). For water retention, the predicted parameters were substituted into [Eq. \(1\)](#), which was evaluated at pressure heads for which water content measurements were available. Subsequently,  $RMSR_{[\theta]}$  values were calculated according to: [Equation \(6\)](#) where  $N$  is the number of samples,  $L$  is the number of measured retention points for each sample and  $[\theta]_{i,j}$  and  $[\theta]_{i,j}'$  are measured and predicted water retention points. Note that there are four hydraulic parameters. The  $RMSR_{K_s}$  values were calculated according to: [Equation \(7\)](#) where  $M$  is the number of samples while  $K_s$  and  $K_s'$  denote measured and predicted values. Smaller RMSR values suggest better predictions. All three error measures are presented as arithmetic averages of  $B$  bootstrap sub models.

$$RMSR_{\theta} = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{L-4} \sum_{j=1}^L (\theta_{ij} - \bar{\theta}_{ij})^2}$$

Equation 6

$$RMSR_{K_s} = \sqrt{\frac{1}{M} \sum_{j=1}^M (\log(K_s) - \log(\bar{K}_s))^2}$$

Equation 7

## RESULTS AND DISCUSSION

### Class Averages

Tables 1 through 3 show the means of fitted retention parameters, and measured  $K_s$  and bulk density values for each textural class in the three databases; standard deviations per textural class are also shown. Although the clay and silt classes are underrepresented, the RAWLS database has the most uniform distribution across the textural triangle. Table 4 provides means and standard deviations for the ALL database (all three databases combined).

Class	N <sup>a</sup>	Water retention					N <sup>a</sup>	$K_s$	
		Bd <sup>b</sup> g cm <sup>-3</sup>	$\theta_e$ cm <sup>3</sup> cm <sup>-3</sup>	$\theta_r$ cm <sup>3</sup> cm <sup>-3</sup>	log( $\alpha_e$ ) log(cm <sup>-1</sup> )	log( $\theta_r$ ) log(cm <sup>-1</sup> )		Bd <sup>b</sup> g cm <sup>-3</sup>	log( $K_s$ ) log(cm day <sup>-1</sup> )
Sand	97	1.46(0.14)	0.044(0.019)	0.415(0.058)	-1.57(0.21)	0.46(0.20)	97	1.44(0.14)	2.71(0.51)
Loamy sand	135	1.51(0.20)	0.040(0.037)	0.395(0.074)	-1.49(0.32)	0.19(0.11)	117	1.52(0.21)	1.92(0.61)
Loam	137	1.39(0.26)	0.052(0.066)	0.354(0.082)	-2.12(0.82)	0.19(0.14)	32	1.50(0.19)	0.99(0.63)
Sandy loam	337	1.42(0.28)	0.031(0.048)	0.389(0.094)	-1.57(0.58)	0.15(0.09)	199	1.53(0.20)	1.53(0.65)
Silt loam	217	1.20(0.28)	0.065(0.062)	0.446(0.109)	-2.51(0.49)	0.26(0.13)	61	1.41(0.15)	1.04(0.54)
Sandy cl. loam	104	1.58(0.19)	0.076(0.075)	0.379(0.068)	-1.80(0.67)	0.13(0.10)	80	1.60(0.20)	1.29(0.71)
Silty cl. loam	47	1.35(0.11)	0.110(0.064)	0.460(0.057)	-2.36(0.39)	0.24(0.11)	10	1.42(0.08)	0.87(0.35)
Clay loam	77	1.41(0.15)	0.092(0.068)	0.441(0.078)	-1.95(0.60)	0.19(0.13)	6	1.47(0.08)	0.67(0.58)
Silt	3	1.27(0.08)	0.077(0.018)	0.501(0.035)	-2.15(0.24)	0.29(0.11)	1	1.38(-)	1.43(-)
Clay	34	1.39(0.13)	0.075(0.078)	0.451(0.070)	-1.93(0.47)	0.11(0.06)	6	1.37(0.07)	0.94(0.31)
Sandy clay	9	1.61(0.08)	0.123(0.095)	0.378(0.041)	-1.40(0.60)	0.09(0.06)	8	1.60(0.08)	1.33(0.33)
Silty clay	12	1.36(0.10)	0.071(0.101)	0.467(0.051)	-2.25(0.34)	0.11(0.05)	3	1.30(0.07)	0.82(0.55)

<sup>a</sup>Number of samples per textural class.<sup>b</sup>Bulk density.<sup>c</sup>Not available.TABLE 1 Average water retention parameters,  $K_s$ , and bulk density values for the textural classes in the RAWLS database.

Standard deviations are given in parentheses

Class	N <sup>a</sup>	Water retention					N <sup>a</sup>	$K_s$	
		Bd <sup>b</sup> g cm <sup>-3</sup>	$\theta_e$ cm <sup>3</sup> cm <sup>-3</sup>	$\theta_r$ cm <sup>3</sup> cm <sup>-3</sup>	log( $\alpha_e$ ) log(cm <sup>-1</sup> )	log( $\theta_r$ ) log(cm <sup>-1</sup> )		Bd <sup>b</sup> g cm <sup>-3</sup>	log( $K_s$ ) log(cm day <sup>-1</sup> )
Sand	82	1.57(0.07)	0.058(0.018)	0.337(0.035)	-1.32(0.13)	0.54(0.10)	82	1.57(0.07)	3.01(0.45)
Loamy sand	19	1.63(0.08)	0.054(0.041)	0.339(0.043)	-1.33(0.38)	0.25(0.15)	19	1.63(0.08)	2.09(0.69)
Loam	50	1.41(0.11)	0.069(0.049)	0.435(0.036)	-2.03(0.27)	0.17(0.09)	50	1.41(0.11)	0.83(0.95)
Sandy loam	65	1.58(0.13)	0.049(0.038)	0.387(0.060)	-1.58(0.53)	0.18(0.11)	65	1.58(0.13)	1.73(0.64)
Silt loam	12	1.39(0.12)	0.055(0.036)	0.471(0.044)	-2.00(0.33)	0.17(0.09)	12	1.39(0.12)	1.24(0.47)
Sandy cl. loam	36	1.60(0.08)	0.059(0.063)	0.378(0.037)	-1.79(0.64)	0.08(0.06)	36	1.60(0.08)	0.81(0.80)
Silty cl. loam	21	1.31(0.12)	0.048(0.075)	0.472(0.042)	-2.05(0.39)	0.10(0.05)	21	1.31(0.12)	1.09(0.78)
Clay loam	48	1.48(0.15)	0.060(0.075)	0.428(0.045)	-1.87(0.67)	0.11(0.09)	48	1.48(0.15)	0.79(1.08)
Silt	0	-	-	-	-	-	0	-	-
Clay	31	1.51(0.16)	0.121(0.111)	0.419(0.039)	-1.99(0.55)	0.10(0.07)	31	1.51(0.16)	1.03(0.83)
Sandy clay	2	1.58(0.07)	0.060(0.000)	0.374(0.001)	-1.74(0.24)	0.03(0.02)	2	1.58(0.07)	-0.03(1.28)
Silty clay	5	1.25(0.16)	0.125(0.131)	0.525(0.058)	-1.23(0.69)	0.12(0.08)	5	1.25(0.16)	1.15(0.16)

<sup>a</sup>Number of samples per textural class.<sup>b</sup>Bulk density.<sup>c</sup>Not available.TABLE 2 Average water retention parameters,  $K_s$ , and bulk density values for the textural classes in the AHUJA database.

Standard deviations are given in parentheses

Class	N <sup>a</sup>	Water retention					N <sup>a</sup>	$K_s$	
		Bd <sup>b</sup> g cm <sup>-3</sup>	$\theta_i$ cm <sup>3</sup> cm <sup>-3</sup>	$\theta_r$ cm <sup>3</sup> cm <sup>-3</sup>	log( $\alpha$ ) log(cm <sup>-1</sup> )	log(n)		Bd <sup>b</sup> g cm <sup>-3</sup>	log( $K_s$ ) log(cm day <sup>-1</sup> )
Sand	129	1.57(0.11)	0.057(0.038)	0.369(0.042)	-1.44(0.30)	0.51(0.20)	74	1.58(0.13)	2.70(0.74)
Loamy sand	51	1.51(0.15)	0.070(0.044)	0.397(0.038)	-1.43(0.29)	0.37(0.19)	31	1.52(0.15)	2.36(0.59)
Loam	62	1.29(0.31)	0.074(0.098)	0.469(0.111)	-1.53(0.56)	0.11(0.10)	31	1.35(0.32)	1.58(0.92)
Sandy loam	79	1.550(0.16)	0.062(0.077)	0.378(0.056)	-1.58(0.51)	0.19(0.14)	50	1.56(0.15)	1.62(0.70)
Silt loam	103	1.44(0.16)	0.066(0.095)	0.432(0.069)	-1.87(0.51)	0.14(0.12)	62	1.43(0.13)	1.48(0.86)
Sandy cl. loam	41	1.52(0.20)	0.053(0.088)	0.400(0.058)	-1.27(0.73)	0.14(0.17)	19	1.55(0.20)	0.99(1.21)
Silty cl. loam	21	1.28(0.30)	0.088(0.105)	0.541(0.133)	-1.47(0.65)	0.14(0.17)	9	1.40(0.12)	1.14(0.85)
Clay loam	25	1.36(0.30)	0.077(0.093)	0.470(0.115)	-1.22(0.70)	0.16(0.11)	8	1.20(0.42)	1.84(0.89)
Silt	3	1.39(0.02)	0.023(0.033)	0.476(0.093)	-2.21(0.30)	0.16(0.09)	2	1.40(0.03)	1.75(0.20)
Clay	27	1.27(0.24)	0.102(0.125)	0.515(0.090)	-1.51(0.89)	0.08(0.10)	25	1.25(0.25)	1.41(1.05)
Sandy clay	1	1.40(=)	0.300(=)	0.473(=)	-1.59(=)	0.10(=)	0	=	=
Silty clay	12	1.40(0.17)	0.145(0.114)	0.476(0.099)	-1.56(0.50)	0.13(0.13)	6	1.40(0.14)	0.92(0.71)

<sup>a</sup>Number of samples per textural class.<sup>b</sup>Bulk density.<sup>c</sup>Not available.TABLE 3 Average water retention parameters,  $K_s$ , and bulk density values for the textural classes in the UNSODA database.

Standard deviations are given in parentheses

Class	N <sup>a</sup>	Water retention					N <sup>a</sup>	$K_s$	
		Bd <sup>b</sup> g cm <sup>-3</sup>	$\theta_i$ cm <sup>3</sup> cm <sup>-3</sup>	$\theta_r$ cm <sup>3</sup> cm <sup>-3</sup>	log( $\alpha$ ) log(cm <sup>-1</sup> )	log(n)		Bd <sup>b</sup> g cm <sup>-3</sup>	log( $K_s$ ) log(cm day <sup>-1</sup> )
Sand	308	1.53(0.12)	0.053(0.029)	0.375(0.055)	-1.45(0.25)	0.50(0.18)	253	1.53(0.13)	2.81(0.59)
Loamy sand	205	1.52(0.19)	0.049(0.042)	0.390(0.070)	-1.46(0.47)	0.24(0.16)	167	1.53(0.19)	2.02(0.64)
Loam	249	1.37(0.35)	0.061(0.073)	0.399(0.098)	-1.95(0.73)	0.17(0.13)	113	1.42(0.22)	1.06(0.92)
Sandy loam	481	1.46(0.26)	0.039(0.054)	0.387(0.085)	-1.57(0.56)	0.16(0.11)	314	1.55(0.18)	1.58(0.66)
Silt loam	332	1.28(0.27)	0.065(0.073)	0.439(0.093)	-2.30(0.57)	0.22(0.14)	135	1.42(0.14)	1.26(0.74)
Sandy cl. loam	181	1.57(0.18)	0.063(0.078)	0.384(0.061)	-1.68(0.71)	0.12(0.12)	135	1.59(0.18)	1.12(0.85)
Silky cl. loam	89	1.32(0.18)	0.090(0.082)	0.482(0.086)	-2.08(0.59)	0.18(0.13)	40	1.36(0.12)	1.05(0.76)
Clay loam	150	1.42(0.19)	0.079(0.076)	0.442(0.079)	-1.80(0.69)	0.15(0.12)	62	1.44(0.23)	0.91(1.09)
Silt	6	1.33(0.09)	0.050(0.041)	0.489(0.078)	-2.18(0.30)	0.22(0.13)	3	1.39(0.03)	1.64(0.27)
Clay	92	1.39(0.20)	0.098(0.107)	0.459(0.079)	-1.82(0.68)	0.10(0.07)	60	1.40(0.23)	1.17(0.92)
Sandy clay	12	1.59(0.10)	0.117(0.114)	0.385(0.046)	-1.48(0.57)	0.08(0.06)	10	1.60(0.08)	1.06(0.89)
Silty clay	29	1.36(0.15)	0.111(0.119)	0.481(0.060)	-1.79(0.64)	0.12(0.10)	14	1.33(0.16)	0.98(0.57)

<sup>a</sup>Number of samples per textural class.<sup>b</sup>Bulk density.TABLE 4 Average water retention parameters,  $K_s$ , and bulk density values for the textural classes of all the data. Standard deviations are given in parentheses

Tables 1 through 3 and Fig. 1A through C show that there are considerable differences in the distribution of the data among the three data-bases. A Student *t* test demonstrated that the means of hydraulic parameters were often significantly different (at  $P < 0.05$ ) for the same textural classes in different databases (results not shown). No particular patterns of similarity or dissimilarity were found in terms of the number of samples per class or type of hydraulic parameter. It is likely that when PTFs are developed on the individual databases, their predictions will be different.

## Calibration and Validation

Table 5 contains the results for the calibrations and validations for the three data sets using neural networks. These are denoted in the first column as  $R_{cal}$  and  $R_{val}$  for the RAWLS database,  $A_{cal}$  and  $A_{val}$  for AHUJA,  $U_{cal}$  and  $U_{val}$  for UNSODA. Generally,  $R^2$  values were somewhat lower for the validations than for the calibrations whereas validation RMSR values were somewhat higher than calibration values. The similarities between calibration and validation results mean that the neural network models were robust and not over-parameterized (cf Schaap and Bouting 1996).

Data sets	$\theta_i$		$\theta_r$		log( $\alpha$ )		log(n)		RMSR <sup>c</sup> cm <sup>3</sup> cm <sup>-3</sup>	$R^2$	ME cm day <sup>-1</sup>	RMSR <sup>c</sup> cm day <sup>-1</sup>
	$R_{cal}$	ME cm <sup>3</sup> cm <sup>-3</sup>	$R_{val}$	ME cm <sup>3</sup> cm <sup>-3</sup>	$R_{cal}$	ME cm <sup>3</sup> cm <sup>-3</sup>	$R_{val}$	ME				
RAWLS												
$R_{cal}$	0.21	-0.000	0.56	-0.000	0.40	-0.001	0.41	0.000	0.086	0.61	0.000	0.517
$R_{val}$	0.18	-0.001	0.54	0.000	0.38	0.001	0.38	-0.000	0.087	0.57	0.000	0.537
$A_{cal}$	0.00	-0.004	0.32	-0.024	0.18	0.158	0.16	-0.031	0.127	0.46	-0.296	0.752
$U_{cal}$	0.02	-0.015	0.42	-0.034	0.03	0.400	0.16	-0.010	0.119	0.53	-0.295	0.634
AHUJA												
$A_{cal}$	0.11	0.001	0.80	-0.001	0.52	-0.002	0.80	0.001	0.061	0.67	0.000	0.659
$A_{val}$	0.06	0.002	0.80	-0.000	0.27	-0.010	0.78	0.001	0.062	0.64	-0.001	0.699
$R_{cal}A_{cal}$	0.00	-0.000	0.44	0.008	0.15	-0.126	0.61	0.012	0.092	0.55	0.169	0.819
$U_{cal}A_{cal}$	0.03	-0.006	0.71	-0.009	0.04	0.245	0.74	0.010	0.079	0.55	-0.135	0.822
UNSDA												
$U_{cal}$	0.06	0.003	0.74	-0.000	0.23	-0.003	0.57	0.001	0.096	0.44	0.000	0.761
$U_{val}$	0.02	-0.001	0.70	-0.001	0.16	-0.002	0.55	0.001	0.098	0.41	-0.002	0.775
$R_{cal}U_{cal}$	0.02	0.009	0.64	0.020	0.04	-0.350	0.33	-0.000	0.105	0.40	0.285	0.822
$A_{cal}U_{cal}$	0.01	0.009	0.60	0.008	0.04	-0.153	0.52	-0.018	0.106	0.28	0.067	0.910
ALL data												
$ALL_{cal}$	0.09	-0.000	0.59	-0.000	0.26	0.002	0.49	0.000	0.093	0.53	0.000	0.669
$ALL_{val}$	0.07	-0.000	0.57	-0.000	0.23	0.004	0.48	0.001	0.095	0.51	-0.000	0.685

<sup>a</sup>Mean error.<sup>b</sup>Root mean square residuals.

TABLE 5 Results for calibration, validation and cross-validation, for neural network PTFs, based on input data from the RAWLS (R), AHUJA (A) and UNSODA (U) databases

The  $R^2$  values of measured (i.e. fitted) and predicted retention parameters varied among data sets. The results showed that  $[\theta]_r$  was almost unpredictable from bulk density and sand, silt, and clay percentages. The maximum  $R^2$  for validation is 0.18 for the RAWLS database ( $R_{val}$ ). The  $R^2$  values for validation of  $[\theta]_s$  were 0.54 for the RAWLS database ( $R_{val}$ ) and 0.80 and 0.70 for the AHUJA and UNSODA databases ( $A_{val}$  and  $U_{val}$ ), respectively.  $R^2$  values for  $[\alpha]$  and  $n$  were 0.38 for  $R_{val}$ ; lower  $R^2$  values for  $[\alpha]$  were found for both  $A_{val}$  and  $U_{val}$  (0.27 and 0.16). For the  $n$  parameter we found higher  $R^2$  values (0.78 and 0.55). Validation  $R^2$  values for  $K_s$  were 0.57 ( $R_{val}$ ), 0.64 ( $A_{val}$ ), and 0.41 ( $U_{val}$ ). The differences in  $R^2$  values may reflect the distribution of the databases and the heterogeneity of measurement methods as well as the quality of the data in the three individual databases. UNSODA, in particular, contains data from many different sources and measurement methods that would generally lead to lower  $R^2$  values.

Mean errors (ME) were essentially zero for all calibration and validation results ([Table 5](#)). This was expected because the bootstrap procedure created random calibration and validation subsets from the same database.

The RMSR<sub>θ</sub> values for the validation stage were the lowest for the AHUJA database ( $0.062 \text{ cm}^3 \text{ cm}^{-3}$ ) and higher for the RAWLS and UNSODA database. RMSR<sub>Ks</sub> values were the lowest for the RAWLS database,  $0.537 \log(\text{cm day}^{-1})$  and higher for the other two databases.

[Table 5](#) also shows calibration and validation results for the neural networks that were calibrated using all of the selected data ( $ALL_{cal}$ ,  $ALL_{val}$ ). Since the PTFs were based on all selected data, the  $R^2$ , and RMSR values were within the ranges found for the individual databases. Therefore, the  $ALL_{cal}$  and  $ALL_{val}$  results provide average predictions for all selected data. Similar to the individual databases, the mean errors are (almost) zero, indicating small systematic errors.

### Cross-Validations

The calibration and, especially, validation results provided benchmarks of the best possible neural network model performances for each individual database. Although the validation results were based on data that were not used for calibration, the models are probably biased towards the database from which the calibration and validation data sets were taken. The cross-validations described in this section served as a test to estimate how well calibrated models performed for other databases. Cross-validations are denoted in the first column of [Table 5](#) as a combination of two letters. The first and second capital letters indicate the data sets that were used for calibration (e.g.,  $A_{cal}$ ) and cross-validation (e.g.,  $R_{cval}$ ), respectively.

Cross-validations always resulted in higher RMSR<sub>θ</sub> and lower  $R^2$  values than calibration and validation results for the same data set. The lowest RMSR<sub>θ</sub> values were obtained for cross-validations with the AHUJA database ( $0.092$  and  $0.079 \text{ cm}^3 \text{ cm}^{-3}$ , for  $R_{cal}A_{cval}$  and  $U_{cal}A_{cval}$ , respectively). For the other cross-validations ( $A_{cal}R_{cval}$ ,  $U_{cal}R_{cval}$ ,  $R_{cal}U_{cval}$ , and  $A_{cal}U_{cval}$ ), the RMSR<sub>θ</sub> values were greater than  $0.1 \text{ cm}^3 \text{ cm}^{-3}$ . The difference between the RMSR<sub>θ</sub> for  $R_{cal}U_{cval}$  or  $A_{cal}U_{cval}$  cross-validations and the  $U_{val}$  benchmark is relatively small. Apparently, the models that were calibrated on the RAWLS and AHUJA data sets can perform almost as well as the models that were calibrated on the UNSODA data set itself. Conversely, the other cross-validations ( $U_{cal}R_{cval}$ ,  $A_{cal}R_{cval}$ ,  $R_{cal}A_{cval}$ , and  $U_{cal}A_{cval}$ ) generally produced much larger RMSR<sub>θ</sub> values than the corresponding validations (i.e.,  $R_{val}$  and  $A_{val}$ ).

A more detailed investigation of the correlations of individual parameters showed that cross-validation  $R^2$  values for  $[\theta]_r$  were all essentially zero. Similar results were found by [Tietje and Tapkenhinrichs \(1993\)](#). For  $[\theta]_s$ , a marked reduction in  $R^2$  was found for the  $R_{cal}A_{cval}$  cross-validation when compared with  $A_{val}$ . For  $\log([\alpha])$ , the correlations were low for all cross-validations. Apparently, the relations between basic soil properties and the  $[\alpha]$  parameter were specific for each database. Note, however, that  $R^2$  for validation was also low ([Table 5](#)). Tietje and Tapkenhinrichs also found low correlations for  $[\alpha]$  when testing the PTF of [Vereecken et al. \(1989\)](#) on independent data. We found somewhat better results for  $\log(n)$ ; the  $R_{cal}A_{cval}$ ,  $U_{cal}A_{cval}$ , and  $A_{cal}U_{cval}$  cross-validations had  $R^2$  values close to the corresponding  $A_{val}$  or  $U_{val}$  values. However, the  $A_{cal}R_{cval}$  and the  $U_{cal}R_{cval}$  cross-validations show that the models that were calibrated on the AHUJA or UNSODA databases predict  $\log(n)$  values of the RAWLS database poorly ( $R^2 = 0.16$  in both cases).

Similarly, we found that cross-validation results for  $K_s$  were also inferior to validation results. In terms of RMSR, the best results were found for the  $A_{cal}R_{cval}$  and  $U_{cal}R_{cval}$  cross-validations. Values for  $R_{cal}A_{cval}$ ,  $U_{cal}A_{cval}$ ,  $R_{cal}U_{cval}$ , and  $A_{cal}U_{cval}$  were higher. In terms of  $R^2$ , the worst results were reached when models were cross-validated with the UNSODA data set ( $R_{cal}U_{cval}$  and  $A_{cal}U_{cal}$ ). It should be noted, however, that the benchmark validation  $R^2$  value of 0.44 ( $U_{val}$ ) for UNSODA was already low.

In contrast to the validation results ( $R_{val}$ ,  $A_{val}$ ,  $U_{val}$ ), mean errors for retention parameters and  $K_s$  were usually not equal to zero ([Table 5](#)). The mean error was usually small for  $[\theta]_r$  and somewhat larger for  $[\theta]_s$  but probably still acceptable for practical use. Mean errors of  $\log([\alpha])$  ranged between -0.350 and 0.400  $\log(\text{cm}^{-1})$ , whereas mean errors for  $\log(n)$  were between -0.031 and 0.018. The mean errors for  $\log(K_s)$  were within the -0.3 to 0.3  $\log(\text{cm day}^{-1})$  range. Together with lower cross-validation  $R^2$  values, the non-zero ME values indicate that PTFs calibrated on different data sets make systematically different predictions of hydraulic parameters. The cross-validation results show that it is hard to decide which PTF is the best.

### Confidence Intervals for Selected Samples

Uncertainty estimates of the retention parameters and  $K_s$ , as generated with the bootstrap method, were investigated for a hypothetical loamy sand, loam, silty loam, and clay using calibration results from all four data bases. The location of the samples in the textural triangle is indicated with an asterisk in [Fig. 1B](#). It is easy to use standard deviations of the  $B$  bootstrap estimates to produce normal probability density functions of the predicted parameters such as is done in [Fig. 3](#) for  $\log(K_s)$ . For water retention, however, it is more interesting to evaluate the resulting uncertainty in the water retention curve instead of parameters. Furthermore, water retention parameters are usually correlated, and treating uncertainty estimates of the parameters as independent tends to overestimate the uncertainty in the predicted water retention curve. We, therefore, converted each of the  $B$  predicted sets of water retention parameters to water retention points at 45 logarithmically spaced pressure heads ( $1 < h < 3 \times 10^4 \text{ cm}$ ). The resulting 45 probability distributions of predicted water contents were parameterized with a mean and a standard deviation, which was subsequently used to calculate the 5 and 95% confidence intervals.

[Figure 3](#) shows that the uncertainty for  $\log(K_s)$  is generally one-half to one order of magnitude (one-half to one unit on the horizontal axis). For the silty loam, the uncertainty in the prediction of  $K_s$  was especially large for the neural network calibrated on the AHUJA data-base. For the clay, a large uncertainty was found for the PTFs based on the AHUJA and RAWLS databases. [Tables 1 and 2](#) show that there were only limited  $K_s$  samples available for these classes (12 samples for the silt loam in the AHUJA data-base and six for the clay in RAWLS). The empirical relations that were established in the neural networks may be skewed towards a few, possibly unreliable, samples. As a result, predictions for these classes may differ significantly among the bootstrap model versions and hence lead to a large uncertainty.

Similar effects were found for the confidence intervals for water retention ([Fig. 4](#)). The intervals were relatively small for the loamy sand and the loam samples (about  $0.04 \text{ cm}^3 \text{ cm}^{-3}$ ) and much larger for the silty loam and the clay samples, sometimes more than  $0.10 \text{ cm}^3 \text{ cm}^{-3}$ . Smaller intervals coincide with classes with larger numbers of samples (sandy loam and loam), whereas larger intervals occurred when only a small number of samples was available for calibration (silty loam, clay). The number of samples is not the only factor that determines the probability distributions; the variability of hydraulic measurements within a textural class may also be important.

The distributions of the predictions by the different PTFs did not always overlap. For  $K_s$  ([Fig. 3](#)), the PTFs calibrated on the RAWLS database generally predicted lower  $\log(K_s)$  values than models calibrated on data sets from the AHUJA and UNSODA databases. This is consistent with the mean errors (ME) for  $\log(K_s)$  presented in [Table 5](#). For water retention ([Fig. 4](#)), systematic differences between predicted curves were masked somewhat by the spread in the confidence intervals. For the loamy sand there was much overlap. This is also true for the dry end of the loam, but at the wet end much disagreement in the value for  $[\theta]_s$  was found among the three databases. Other discrepancies were found for the clay and, especially, the silty loam. The clay had a relatively good match at the wet end but a rather broad distribution at the dry end. The lack of coincidence can be explained by systematic differences in the predictions of  $[\theta]_r$ ,  $[\theta]_s$ ,  $\log([\alpha])$  and  $\log(n)$  for the three databases sets. Because the cross-validation results indicated that there is no good way to decide which PTF yields better predictions, the disagreement among the PTF predictions adds to the uncertainty about the predicted hydraulic parameters.

[Figures 3 and 5](#) show predictions and associated uncertainties for the PTFs calibrated on the ALL database. To make a comparison with [Fig. 4](#) possible, we plotted the envelope of the confidence intervals of [Fig. 4](#) in [Fig. 5](#). In the case of  $K_s$  ([Fig. 3](#)), the predicted uncertainties for the loamy sand and the loam were somewhat higher (peak lower) than the uncertainties for the RAWLS database. The uncertainties for the silty loam and the clay were lower (peak higher) than the three separate databases. This improvement occurred because the calibration could be based on more samples. For water retention ([Fig. 5](#)) the predicted uncertainties were all substantially smaller than the envelope of uncertainties from [Fig. 4](#). Further, the size of the uncertainty intervals for the four soils is approximately the same. This indicated that predictions by PTFs based on the ALL database were more reliable than predictions by PTFs based on individual databases. The increased generality of the PTFs based on the ALL database may outweigh their somewhat higher calibration and validation RMSR values, as reported in [Table 5](#).

It should be stressed that [Figs 3, 4, and 5](#) give examples for four hypothetical soils. The position and sizes of the confidence intervals will change according to the input data.

## CONCLUSIONS

In this paper, we investigated the accuracy and reliability of predictions of hydraulic properties by PTFs. The objectives were to discover how dependent the PTF performances were on the distributions of data in four databases (RAWLS, AHUJA, UNSODA, and all three combined in ALL) and to investigate with what certainty predictions were made. To this end, we used a three-step approach: calibration, validation, and cross-validation of neural network-based PTFs. The calibration and validation steps were combined with the bootstrap method ([Efron and Tibshirani 1993](#)) to obtain random calibration and independent validation subsets from the same database and to quantify the uncertainty of PTF predictions. In the third stage, cross-validation, we tested how well calibrated PTFs could predict hydraulic properties from different distributions databases.

The calibration and validations showed that the correlation between predicted and measured  $[\theta]_r$  values was always zero. The two smaller data-bases (AHUJA and UNSODA) also showed weak correlations between predicted and fitted  $[\alpha]$ . The other hydraulic parameters ( $[\theta]_s$ ,  $n$ , and  $K_s$ ) were predicted with better correlations.

Cross-validation error measures were inferior to those found in the validation procedure. Further, we found that the PTFs often predicted systematically different hydraulic parameters. This indicated that the PTFs were optimal for the databases on which they were calibrated but had mediocre performances on different databases. Because most PTFs available in the literature have been developed on different databases, it is likely that similar, systematically different, predictions are made when these PTFs are used to predict soil hydraulic properties. The cross-validations also showed that the perceived quality of a PTF not only depends on the data used for calibration but depends also on the distribution of the data that are used to evaluate the PTF. The PTFs that were developed on the combination of the data in the three data sets (the ALL dataset) did not have systematic errors in their predictions. However, errors are still likely when the PTFs are tested with data with different distributions than the ALL data set.

Previously, estimation of the variability of soil hydraulic parameters was only possible by textural class (e.g., [Carsel and Parrish 1988; Wösten and van Genuchten 1988](#)). Combination of PTF calibration with the bootstrap allows estimates of uncertainty for each individual prediction. We found that uncertainty estimates of the PTFs were sometimes large, especially for samples with properties that were not abundantly present in the calibration databases. For  $K_s$ , the uncertainties were about one-half to one order of magnitude. For water retention, the uncertainties were 0.04 to 0.10  $\text{cm}^3 \text{cm}^{-3}$ . When all available data were used the uncertainties became smaller.

It should be stressed that PTFs are estimation methods that depend strongly on the quality of the calibration data. Whether the accuracy of predicted soil hydraulic properties is sufficient depends on the application for which the predictions are used. Uncertainty estimates will allow insight into the reliability of the prediction and would provide, for example, valuable information for computing confidence intervals of simulated vadose zone water fluxes.

## ACKNOWLEDGMENTS

This research was supported, in part, by NASA, project (95) 805/60110. The authors thank W.J. Rawls at the USDA-ARS Hydrology Lab., Beltsville, MD, and L.R. Ahuja at the USDA-ARS Great Plains Systems Research Unit, Fort Collins, CO, for providing their soil moisture databases.

## REFERENCES

- Brakensiek, D. L., W. J. Rawls, and G. R. Stephenson. 1984. Modifying SCS hydrologic soil groups and curve numbers for rangeland soils. ASAE Paper No. PNR-84-203, St. Joseph, MI. [\[Context Link\]](#)
- Carsel, R. F., and R. S. Parrish. 1988. Developing joint probability distributions of soil water retention characteristics. Water Resour. Res. 24:755-769. [\[Context Link\]](#)
- Cosby B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn. 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resour. Res. 20:683-690. [\[Context Link\]](#)
- Demuth, H., and M. Beale. 1992. Neural Network Toolbox Manual. Math Works Inc., Natick, MA. [\[Context Link\]](#)
- Efron, B., and R. J. Tibshirani. 1993. An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability. Chapman and Hall, New York. [\[Context Link\]](#)
- Haykin, S. 1994. Neural Networks, a Comprehensive Foundation. Macmillan College Publishing Company, New York. [\[Context Link\]](#)
- Hecht-Nielsen, R. 1991. Neurocomputing. Addison-Wesley Publishing Company Inc, New York. [\[Context Link\]](#)
- Kern, J. S. 1995. Evaluation of soil water retention models based on basic soil physical properties. Soil Sci. Soc. Am. J. 59:1134-1141. [\[Context Link\]](#)
- Leij, F. J., W. J. Alves, M. Th. van Genuchten, and J. R. Williams. 1996. The UNSODA unsaturated soil hydraulic database, version 1.0. Report IAGDW12933934, national risk management research laboratory, office of research and

development, US EPA, Cincinnati, OH. (Available at: <http://www.epa.gov/ada/models.html>). [Context Link]

Marquardt, D. W. 1963. An algorithm for least-squares estimation of non-linear parameters. *J. Soc. Ind. Appl. Math.* 11:431-441. [Context Link]

Nelder, J. A., and Mead, R. 1965. A simplex method for function minimization. *Comput. J.* 7:308-313. [Context Link]

Pachepsky, Ya. A., D. Timlin, and G. Varallyay. 1996. Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Sci. Soc. Am. J.* 60:727-733. [Context Link]

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1988. Numerical recipes in C. Cambridge University Press, Cambridge, UK. [Context Link]

Rawls, W. J., and D. L. Brakensiek. 1985. Prediction of soil water properties for hydrologic modeling. In *Watershed Management in the Eighties*. E. B. Jones, and T.J. Ward (eds.). Proc. Irrig. Drain. div., ASCE, Denver, CO, April 30-May 1, 1985, pp. 293-299. [Context Link]

Rawls, W.J., T.J. Gish, and D.L. Brakensiek. 1991. Estimating soil water retention from soil physical properties and characteristics. In *Advances in Soil Science*, Stewart, B.A. (ed.). Springer-Verlag, New York. [Context Link]

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

Schaap, M. G., and W. Bouten. 1996. Modeling water retention curves of sandy soils using neural networks. *Water Resour. Res.* 32:3033-3040. [Context Link]

Schaap, M. G., F.J. Leij, and M.Th. van Genuchten. 1998. Neural network analysis for hierarchical prediction of soil water retention and saturated hydraulic conductivity. *Soil Sci. Soc. Am. J.* 62:847-855. [Context Link]

Tamari, S., J. H. M. Wösten, and J. C. Ruiz-Suárez. 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci. Soc. Am. J.* 60:1732-1741. [Context Link]

Tietje, O., and M. Tapkenhinrichs. 1993. Evaluation of pedotransfer functions. *Soil Sci. Soc. Am. J.* 57:1088-1095. [Context Link]

Tietje, O., and V. Hennings. 1996. Accuracy of the saturated hydraulic conductivity prediction by pedotransfer functions compared to the variability within FAO textural classes. *Geoderma* 69:71-84. [Context Link]

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

Vereecken, H., J. Maes, J. Feyen, and P. Darius. 1989. Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content. *Soil Sci.* 148:389-403. [Buy Now](#) | [Context Link]

Wösten J. M. H. and M. Th. van Genuchten. 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. *Soil Sci. Am. J.* 52:1762-1770. [Context Link]

<sup>1</sup>Trade names are provided for the benefit of the reader and do not imply endorsement by the USDA. [Context Link]

Key words: Water retention; hydraulic conductivity; prediction; uncertainty estimates; neural networks

## IMAGE GALLERY

Select All

Export Selected to PowerPoint

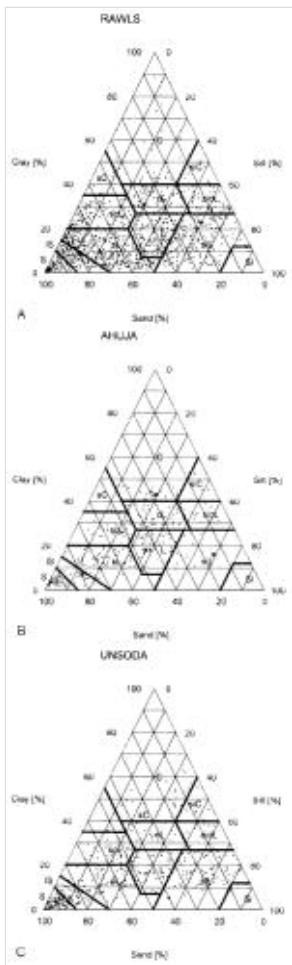


Fig. 1

$$\theta(h) = \theta_s + \frac{(\theta_r - \theta_s)}{[1 + (\alpha h)^m]^m} \quad (m = 1 - 1/n)$$

Equation 1

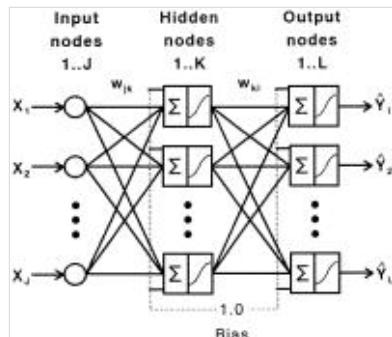


Fig. 2

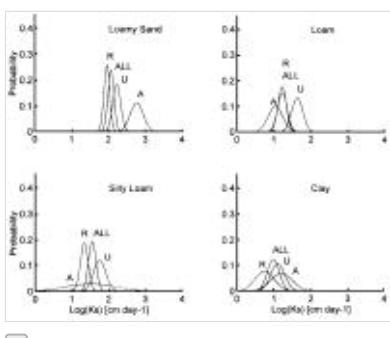


Fig. 3

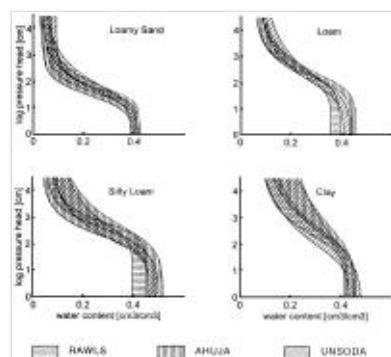


Fig. 5

$$S_k = \sum_{j=0}^J (w_{jk} * \chi_j)$$

$$H_k = \frac{1}{1 + e^{-S_k}}$$

Equation 3

Equation 2

$$Ob = \sum_{i=1}^{N_c} \sum_{j=1}^{N_p} (t_{i,j} - t_{i,j}')^2$$

Equation 4

$$ME_{\zeta} = \frac{1}{N} \sum_{j=1}^N (\zeta_j - \zeta_j')$$

$$RMSR_6 = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{L-4} \sum_{j=1}^L (\theta_{i,j} - \bar{\theta}_{i,j})^2}$$

Equation 6

$$RMSR_{K_0} = \sqrt{\frac{1}{M} \sum_{j=1}^M (\log(K_j) - \log(K'_j))^2}$$

Equation 7

Equation 5

Table 1

Table 2

Table 3

Table 4

Row	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
6	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
8	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
9	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
10	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
11	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
12	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
13	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
14	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
15	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
16	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
18	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
19	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
21	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
22	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
23	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
24	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24

Table 5

Back to Top

© 2017 Ovid Technologies, Inc. All rights reserved.

[About Us](#)   [Contact Us](#)   [Terms of Use](#)

OvidSP\_UI03.27.02.119, SourceID 109790