Multiscale Pedotransfer Functions for Soil Water Retention

Raghavendra B. Jana, Binayak P. Mohanty,* and Everett P. Springer

Parametric soil water retention and hydraulic conductivity functions are often used for predicting soil hydrologic behavior using hydrologic, hydroclimatic, and contaminant transport models. The prediction accuracy of any such model is critically dependent on the quality of the input parameters. Limited availability of (detailed) soil hydraulic data for large-scale hydroclimatic models (with grids ranging from several kilometers to several hundred kilometers) is a major challenge. To address this need, pedotransfer functions (PTFs) have been used to estimate the required soil hydraulic parameters from other available or easily measurable soil properties. While most previous studies derive and adopt these parameters at matching spatial scales (1:1) of input and output data, we have developed a methodology to derive soil water retention functions at the point or local scale using the PTFs trained with coarser scale input data. This study was a novel application of an artificial neural network (ANN)-based PTF scheme across two spatial support scales within the Rio Grande basin in New Mexico. The ANN was trained using soil texture and bulk density data from the SSURGO database (scale 1:24,000) and then used for predicting soil water contents at different pressure heads with point-scale data (1:1) inputs. The resulting outputs were corrected for bias before constructing the soil water characteristic curve using the van Genuchten equation. A hierarchical approach with training data derived from multiple clustered subwatersheds (with varying spatial extent) was used to study the effect of the increase in spatial extent. The results show good agreement between the soil water retention curves constructed from the ANN-based PTFs and field observations at the local scale near Las Cruces, NM. The robustness of the multiscale PTF methodology was further tested with a separate data set from the Little Washita watershed region in Oklahoma. Overall, ANN coupled with bias correction was found to be a suitable approach for deriving soil hydraulic parameters at a finer scale from soil physical properties at coarser scales and across different spatial extents. The approach could potentially be used for downscaling soil hydraulic properties.

ABBREVIATIONS: ANN, artificial neural network; LW, Little Washita; PTF, pedotransfer function.

868

Soil hydraulic properties are needed in global- and regionalscale circulation models for hydrologic and climate forecasting. They are also necessary in point- and nonpoint-source contaminant transport models. Prediction accuracy of these models is highly dependent on the quality of the model parameters. Collecting the required soil hydraulic parameters by direct measurement at any model grid scale is expensive and time consuming. Also, to capture the variability of these spatially distributed parameters, a large number of samples needs to be collected. All of these factors make it highly impractical to match direct measurements of the soil hydraulic parameters to the invoked model grids. For these reasons, pedotransfer functions (PTFs) have been advocated to estimate the required soil hydraulic parameters from other available or more easily measurable soil data.

A large number of studies have been performed in the recent past to develop such PTFs and test them against available soil prop-

Vadose Zone J. 6:868–878 doi:10.2136/vzj2007.0055

© Soil Science Society of America

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher.

erties databases (e.g., Rawls et al., 1991; van Genuchten and Leij, 1992; Schaap et al., 1998; Pachepsky et al., 1999; Wösten et al., 2001; Sharma et al., 2006). "Point regression" PTFs use empirically derived regression equations to predict the soil water content at fixed soil water potentials (e.g., Rawls et al., 1982; Ahuja et al., 1985; Tomasella et al., 2000). On the other hand, "function parameter" PTFs predict the parameters of the water retention and hydraulic conductivity functions (e.g., Vereecken et al., 1989; Schaap et al., 1998; Wösten et al., 2001), such as those given by Brooks and Corey (1964), Campbell (1974), and van Genuchten (1980). Both approaches have been widely utilized for various soil databases. Arguably, function parameter PTFs were preferred to point regression PTFs since they generate the complete function of the relationship between the water content and the pressure head. This made it easy to construct the entire soil water retention curve useful for modeling studies. It was recently shown, however, that point regression PTFs perform better than function parameter PTFs (Tomasella et al., 2003). Since the relationship between soil physical properties and soil water retention parameters is rather complicated, the variability in the retention parameters is controlled by different subsets of soil physical properties at different ranges of soil water pressure. Tomasella et al. (2003) suggested that the reason for the relatively poor performance of function parameter PTFs is due to their inability to accurately describe these complex physical relationships at all pressure heads from wet to dry conditions.

Soil texture (sand, silt, and clay percentages) has been popularly used for predicting soil hydraulic properties. It has been shown that the use of detailed particle-size distributions can increase the accuracy of soil hydraulic parameter predictions

R.B. Jana and B.P. Mohanty, Dep. of Biological and Agricultural Engineering, Texas A&M Univ., College Station, TX 77843; and E.P. Springer, Earth & Environmental Sciences Division, Los Alamos National Lab., Los Alamos, NM 87845. Received 23 Mar. 2007. *Corresponding author (bmohanty@tamu.edu).

(Schaap et al., 1998) compared with using soil textural class alone as input (Clapp and Hornberger, 1978); however, detailed particle size data are not easily available in all cases. The most commonly used soil physical properties for PTF applications are soil texture, organic C, and bulk density. Other parameters such as topography and vegetation have rarely been used in developing PTFs (Wösten et al., 2001). Recently, Pachepsky et al. (2001), Leij et al. (2004), and Sharma et al. (2006) included certain available topographical and vegetation attributes, in addition to soil physical parameters, for developing PTFs. While the inclusion of more input parameters for the PTFs provided some improvement in the performance of the transfer function models, the basic soil properties had by far the most effect on the hydraulic properties predictions. Increasing the number of model input parameters also means increasing the complexity of the model, including inherent uncertainties associated with the input data.

Most previous PTF studies have derived and adopted soil hydraulic parameters at matching spatial scales of the input and target data. The primary objective of this study was to develop and test a methodology to derive soil water content values (at saturation, θ_s , residual, θ_r , and field capacity, θ_f) and the van

Genuchten soil water retention function at the point or local (1:1) scale using artificial neural network (ANN)-based PTFs trained with coarser (1:24,000) scale SSURGO soil textural data. A secondary objective was to investigate improvements, if any, in the performance of the ANN-based PTF scheme to predict soil water contents at the local or point scale by including training data from larger spatial extents (scales) in a hierarchical fashion within the Rio Grande basin, New Mexico.

Materials and Methods

Study Area

The Rio Grande basin, from its headwaters in southern Colorado to the New Mexico-Texas border, is the focus of field and modeling studies by the National Science Foundation's Science and Technology Center for Sustainability of Semi-Arid Hydrology and Riparian Areas and Los Alamos National Laboratory (Winter et al., 2004). This region, having an area of approximately 90,000 km² in New Mexico, was used for our case study (Fig. 1). We used point-scale (1:1) soil physical and hydraulic properties measured at the Las Cruces trench site (Wierenga et al., 1989), situated within the Rio Grande-Mimbres subwatershed region, for our data at the local or point scale. We developed a multiscale soil physical and hydraulic properties database by compiling (i) the point-scale (1:1) data from the Las Cruces trench site, and (ii) coarse-scale (1:24,000) SSURGO soil survey data for the Rio Grande river basin from the NRCS. The database was developed in Geodatabase format, thus making the spatial data accessible through the ESRI ArcMap software (ESRI, Redlands, CA).

Point- or Local-Scale Soil Properties Data

The Las Cruces trench is located at New Mexico State Ranch, roughly 65 km northeast of the city of Las Cruces (Fig. 1). The trench is 26.4 m long, 4.5 m wide, and 6 m deep, and is situated in undisturbed soil. Using in situ and laboratory methods, Wierenga et al. (1989) developed a comprehensive database of fine-scale (1:1) soil properties using 594 disturbed soil samples and 594 associated soil cores taken from nine distinct soil layers identified along the north wall of the trench. Additional samples were taken from three vertical transects along this wall. The data set included saturated hydraulic conductivity, the soil water retention function, particle size distribution, and the bulk density of each layer. Besides Las Cruces trench site data, no other complete soil properties data set was available for the Rio Grande river basin at this scale (Jana et al., 2005), thus limiting our model test bed to the Las Cruces location. For point- or local-scale testing of the hierarchical PTFs developed using the coarse-scale soil properties across clustered subwatersheds up to the Rio Grande river basin, we used replicated values across the 26-m-long trench.

Coarse-Scale Soil Properties Data

The coarse-resolution (1:24,000) soil properties data were derived from the Soil Survey Geographic (SSURGO) database (Soil Survey Staff, 2007). SSURGO is the most detailed soil map-



Fig. 1. Study area showing the Rio Grande Basin in New Mexico and location of the city of Las Cruces (trench site).

ping database compiled by the NRCS, containing georeferenced spatial and attribute data for soils. Since the database covers a large areal extent, the soil property data in SSURGO are based on soil type rather than the spatial location. The SSURGO database was created by field methods, using observations along soil delineation boundaries and traverses, and by determining map unit composition using field transects. Aerial photographs were interpreted and used as the field map base, while multiple readings were taken for each property within each map unit. The number of readings differ between map units based on such factors as size of the soil polygon, variations in topography, and changes in vegetation, among others. Low, high, and representative values for the observed readings were then entered into the database for that particular soil type or map unit. The procedure is in line with the specifications of the USDA-NRCS National Soil Survey Handbook (Natural Resources Conservation Service, 2001). Maps were made at scales ranging from 1:12,000 to 1:31,680 (www.il.nrcs.usda.gov/technical/soils/soilfact.html [verified 9 Oct. 2007]). In our study, we used representative values for all parameters from the maps with a scale of 1:24,000. The initial SSURGO data were measured at the point scale, but the values reported in the database are averaged values. Although recent studies (Zhu and Mohanty, 2002a,b; Mohanty and Zhu, 2007) have indicated that appropriate (site-specific) upscaling schemes are necessary for deriving effective soil parameters at larger scales, averaging multiple sample values across soil map units to arrive at the spatially representative parameter values for the 1:24,000 scale is a simple form of upscaling adopted by SSURGO. Hence, the available SSURGO database arguably provides a generic form of upscaled (coarse-resolution) values for the parameters.

The hydraulic parameters used from the SSURGO database are the water content at satiation (θ_s), the water content at a pressure of 1.5 MPa (θ_r), and the water content at 33.3 kPa (θ_f). Retention measurements closest to 33.3 kPa at the Rio Grande trench site were taken at -300 cm pressure. Hence, the water content at that value was used for θ_f , while other water contents

are for the same pressure heads as those of SSURGO. Furthermore, the values used from the SSURGO database were for the topsoil layer (0–6 cm). To compare similar data from the two scales (coarse and fine) at matching depths, only values from corresponding layers were used from the Las Cruces trench site.

Validation Study Area

The Little Washita (LW) watershed in Oklahoma was selected for validation of the multiscale PTF methodology. This watershed has been the focus of a number of intensive field studies, including the Southern Great Plains 1997 hydrology experiment and the Soil Moisture Experiment 2002 campaign. The SSURGO data for Oklahoma and the pointscale soil property data from the LW watershed, collected by Mohanty et al. (2002), were used together in our validation study (Fig. 2). The LW region has rolling topography with a variety of vegetative covers. As such, the spatial distribution of point-scale data across the LW watershed was in sharp contrast to the local conditions (small spatial extent and limited soil variety) of the Las Cruces trench site. These site-specific differences, and the availability of spatially distributed localor fine-scale soil hydraulic properties data, made the LW region ideal for validating the multiscale PTF approach. Seventy pointscale measurements at 3- to 9-cm depth were chosen across the LW watershed to form the fine-scale (1:1) data. SSURGO data from Caddo, Comanche, and Grady counties in Oklahoma were used for the coarse-scale (1:24,000) data.

Artificial Neural Network–Linear Regression Based Pedotransfer Function

In this study, the ANN analysis was performed using the Neural Network Toolbox of MATLAB (The MathWorks, Natick, MA). The networks were designed with one input layer, one hidden layer, and an output layer. Although ANN techniques have been well established, for the sake of completeness, a brief description of the ANN approach adopted here is given below.

A neural network typically consists of an input layer, an output layer, and one or more hidden layers linking these two. The hidden layer extracts useful information from inputs and uses them to predict the outputs. The ANN is schematically represented in Fig. 3. The input layer consists of four input parameters (soil physical properties). These are fed to the hidden layer, consisting of four neurons. The inputs are multiplied by the layer weights w and summed with the layer bias b. This summation is then fed to the transfer function f. Outputs from the hidden-layer transfer functions are subjected to the same treatment at the output layer. The output-layer transfer function produces the required output, the soil water content.

Neurons may use any differentiable transfer function to generate their output. The log-sigmoid (Log), tan-sigmoid (Tan), and linear (Lin) transfer functions are the more popularly used functions. These three functions are generally used as they are mathematically convenient and allow the ANN to model both strong and weakly nonlinear relationships. Because of the complexity of soil hydraulic properties across scales, we adopted all three (Log, Tan, and Lin) transfer functions in this



 $\mathsf{F}\mathsf{iG}$ 2. Validation study area in Oklahoma and Little Washita (LW) watershed data locations.



Fig. 3. Artificial neural network model: *w* represents the layer weights, *b* the bias, *f* is the transfer function, and θ is the output (water content). After Demuth et al. (2005).

study to evaluate their performance in estimating nonlinear soil water retention functions. The algorithms for the three transfer functions used here are (Demuth et al., 2005)

$$Log = \log \operatorname{sig}(n) = \frac{1}{1 + \exp(-n)}$$
[1]

$$Tan = \tan \operatorname{sig}(n) = \frac{2}{1 + \exp(-2n)}$$
[2]

$$\operatorname{Lin} = \operatorname{purelin}(n) = n$$
[3]

A "feed-forward backpropagation" type neural network has been used previously to develop PTFs (Pachepsky et al., 1996; Schaap et al., 1998; Sharma et al., 2006). A feed-forward network is one in which the flow of data through the network is in one direction only. Backpropagation refers to the process of feeding the output of a neuron back to itself so that it may learn. The backpropagation network learns by examples in small steps. A set of inputs and corresponding outputs are given to the network to train it in recognizing the desired results. By adjusting the weights iteratively, the network is trained for each input–output combination until the overall error decreases below a predetermined value.

Using SSURGO data, we trained the ANNs for estimating soil water contents at different pressure heads (θ_s , θ_r , and θ_f). The variable learning rate algorithm "traingdx" was used for backpropagation training of the ANNs. This algorithm is faster and more reliable than traditional "train" and "traingd" algorithms in that it ensures stability of learning. Early stopping, a technique to prevent overfitting of the data, was also enabled. When early stopping is enabled, the ANN monitors the error on the validation data set. If the ANN overfits the data, validation errors rise. When the validation error increases for a specified number of iterations, the training is terminated and the weights and biases at the minimum of the validation error are returned. In our study, we specified 5% (500 iterations) as the limit for rise in validation error.

The SSURGO data pool available for the Rio Grande-Mimbres subwatershed containing the Las Cruces trench site consisted of 6685 sets of values. Each data set consisted of the training inputs (sand, silt, and clay content and ovendry bulk density) and the corresponding target outputs (θ_{e} , θ_r , and θ_f). One thousand random sets of data values were selected for the ANN training from this data pool by means of a bootstrapping process that was terminated once the required number (1000) of values had been reached. While any random selection algorithm could be used to extract the training data from this large data pool, we used bootstrapping since our methodology was designed in part for applications where such a large data pool may not be available (e.g., remotely sensed data). Conducting several replicated model runs, we observed that further increase in the size of the training data set (>1000 and within the available data pool) did not provide any further improvement in the training. Moreover, by keeping a low ratio of selected to available data sets, we ensured randomness of the bootstrapped selections. Five hundred random sets of data values were further extracted for use as a validation data set to enable early stopping in the ANN. Finally, using the trained neu-

ral networks with the SSURGO-based coarse-resolution data sets, predictions of soil water contents (θ_s , θ_r , and θ_f) were made at the point resolution for the Las Cruces trench site with 50 point data sets of sand, silt, and clay contents and oven-dry bulk density at the depth of 0 to 6 cm.

Using a hierarchical approach, we enhanced the data pool for the ANN training from one subwatershed to the entire Rio Grande river basin in 10 clusters. Table 1 shows the number of available SSURGO data sets in the data pool for the clusters. At each clustering level, 1000 sets of data were extracted by bootstrapping and used to train the ANNs. Subsequently, ANN predictions of the soil water contents (θ_s , θ_r , and θ_f) were made at the point resolution for the Las Cruces trench site at each level of clustering of the subwatersheds having different spatial extents.

The ANN predictions were evaluated using different statistical indicators. Correlations between predicted and target values for θ_s , θ_r , and θ_f at the point scale were determined. Following Ines and Hansen (2006), the prediction errors relative to the target (measured) values were split into random and systematic components. This was done to allow corrections in the predictions due to systematic bias. Bias here refers to the systematic component of the error and is not to be mistaken for bias in the ANN architecture. According to Willmott (1982), the MSE,

TABLE 1. Number of values available for bootstrapping from the SSUR-GO data pool at different cluster levels in the Rio Grande basin.

Cluster no.	Area	Values in SSURGO data pool
	km ²	no.
1	4,133	6,396
2	12,430	12,721
3	20,001	15,031
4	31,875	17,625
5	39,912	18,730
6	46,749	19,379
7	55,060	24,890
8	67,159	28,956
9	78,696	32,694
10	90,294	36,022

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$$
[4]

can be decomposed into a random component not correctable by linear transformation,

$$MSE_{R} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_{i} * - y_{i})^{2}$$
[5]

and a systematic component,

$$MSE_{S} = MSE - MSE_{R}$$
[6]

where *n* is the number of samples, y_i and \hat{y}_i are the target and ANN-predicted parameter values respectively, and \hat{y}_i^* is \hat{y}_i after bias correction by linear regression. Such a bias correction provides a proportional shifting and brings the mean of the ANN-predicted values closer to that of the target values at the local scale. Before applying the bias correction, we applied the Kolmogorov–Smirnov algorithm to verify that the residuals between the target and ANN-predicted water content values were normally distributed. Values of the correlation coefficient (*R*) and RMSE were used for evaluating the transfer function model since they are generally accepted measures of prediction accuracy.

Using point-scale data, previous studies (Schaap et al., 1998; Tomasella et al., 2000; Nemes et al., 2003) attributed any bias in the ANN prediction of the soil hydraulic properties to differences in quality, textural composition, and origin of the data between

the training and target sets. In this study, using two independent and characteristically different (similar vs. different textural distributions at the fine and coarse scales) from the Rio Grande basin and Little Washita watershed, we suggest that the systematic component of the error in the simulated values is a function of the scale of support of the observed data. Nevertheless, we realize that the bias due to differences in the support scale may also include some of the other site-specific reasons as suggested in other studies. We trained the ANN using coarse-scale (1:24,000) data and then used this training to estimate the soil hydraulic properties at a finer scale (1:1), resulting in a systematic bias. This is in line with findings in other geoscience applications where a systematic bias was attributed to scaling (e.g., Kanamaru and Kanamitsu, 2007). Thus, we adopted model calibration by correcting for the systematic bias using a simple linear regression approach. The bias-corrected values for each parameter having the best correlation (among the Log, Tan, and Lin transfer functions) were then used to fit the van Genuchten model for the soil water characteristics curve,

$$\theta(\psi) = \theta_{r} + (\theta_{s} - \theta_{r}) \left[\frac{1}{1 + (\alpha \psi)^{n}} \right]^{1 - \frac{1}{n}}$$
[7]

The van Genuchten curve-shape parameters α and *n* were estimated by iteratively solving for

the $\theta_{\rm f}$ value at a pressure head (ψ) of -300 cm using known (ANN-predicted) values for $\theta_{\rm s}$ and $\theta_{\rm r}$. The iterative solver was constrained within the parameter ranges (i.e., between upper and lower limits) for the particular soil type (at the local scale). Here, α was constrained between 0.001 and 0.5 while *n* was constrained between 1 and 4.

Results and Discussion

The ANN was trained using soil texture (sand, silt, and clay contents) and bulk density data from the SSURGO database (1:24,000 scale) by clustering data from 10 hierarchical spatial extents, and subsequently used for predicting the soil water contents (θ_s , θ_r , and θ_f) for the point- or local-scale (1:1) inputs at the Las Cruces trench site. Results are presented and discussed individually for the saturated water content, the residual water content, and the field capacity water content, reflecting the significance of applying ANN-based PTFs across spatial scales as the network is trained at one scale and rendered to predict soil water content values at another scale.

Saturated Water Content

Table 2 shows the results of statistical analyses between the observed and ANN-simulated values of θ_s . The best correlation between observed and simulated θ_s was most often obtained when we used the sigmoid transfer functions. This indicates that a linear transfer function is unable to capture the relationship between the inputs and the saturated water content. The *R* value for the

TABLE 2. Statistical analysis of artificial neural network predicted fine-resolution saturated volumetric water content values from the Las Cruces trench site.

Cluster	Transfer function†	Mean	SD	R	RMSE	MSE	Random error	Systematic error
		— m ³	m ^{_3} —					%
1	Linear	0.354	0.003	0.193	2.56×10^{-2}	$6.553 imes 10^{-4}$	78.5	21.5
	Log	0.332	0.007	0.537	2.29×10^{-2}	$5.234 imes 10^{-4}$	74.2	25.8
	Tan	0.365	0.041	0.553	4.09×10^{-2}	1.673×10^{-3}	22.3	77.7
2	Linear	0.350	0.002	0.142	$2.43 imes 10^{-2}$	5.890×10^{-4}	88.9	11.1
	Log	0.357	0.003	0.558	2.63×10^{-2}	$6.910 imes 10^{-4}$	54.0	46.0
	Tan	0.379	0.006	-0.256	4.46×10^{-2}	1.989×10^{-3}	25.6	74.4
3	Linear	0.339	0.003	0.089	2.31×10^{-2}	$5.342 imes 10^{-4}$	99.2	0.8
	Log	0.324	0.008	0.548	2.66×10^{-2}	7.069×10^{-4}	53.7	46.3
	Tan	0.341	0.003	0.333	2.22×10^{-2}	4.948×10^{-4}	96.8	3.2
4	Linear	0.340	0.002	-0.012	$2.32 imes 10^{-2}$	5.396×10^{-4}	99.1	0.9
	Log	0.374	0.003	0.246	$3.90 imes 10^{-2}$	1.518×10^{-3}	33.0	67.0
	Tan	0.337	0.038	0.520	3.28×10^{-2}	1.075×10^{-3}	37.2	62.8
5	Linear	0.342	0.003	0.043	2.31×10^{-2}	$5.313 imes10^{-4}$	100.4	-0.4
	Log	0.353	0.007	0.494	$2.30 imes 10^{-2}$	5.309×10^{-4}	77.2	22.8
	Tan	0.369	0.009	0.573	$3.30 imes 10^{-2}$	1.091×10^{-3}	33.3	66.7
6	Linear	0.354	0.002	-0.127	2.61×10^{-2}	6.832×10^{-4}	76.9	23.1
	Log	0.362	0.002	-0.072	3.06×10^{-2}	$9.335 imes 10^{-4}$	57.1	42.9
	Tan	0.350	0.011	0.562	2.06×10^{-2}	4.224×10^{-4}	87.8	12.2
7	Linear	0.356	0.002	-0.013	2.70×10^{-2}	7.284×10^{-4}	73.4	26.6
	Log	0.357	0.008	0.561	$2.48 imes 10^{-2}$	6.166×10^{-4}	60.6	39.4
	Tan	0.336	0.008	0.451	2.15×10^{-2}	4.637×10^{-4}	93.0	7.0
8	Linear	0.370	0.002	0.107	3.59×10^{-2}	1.291×10^{-3}	40.9	59.1
	Log	0.373	0.001	0.521	$3.86 imes 10^{-2}$	$1.493 imes 10^{-3}$	26.4	73.6
	Tan	0.353	0.006	0.505	2.34×10^{-2}	5.469×10^{-4}	72.9	27.1
9	Linear	0.352	0.002	-0.042	$2.53 imes 10^{-2}$	6.381×10^{-4}	83.8	16.2
	Log	0.365	0.007	0.512	3.06×10^{-2}	9.384×10^{-4}	42.4	57.6
	Tan	0.355	0.009	0.500	2.39×10^{-2}	$5.733 imes 10^{-4}$	70.3	29.7
10	Linear	0.355	0.002	0.117	2.64×10^{-2}	6.984×10^{-4}	75.4	24.6
	Log	0.356	0.009	0.513	$2.42 imes 10^{-2}$	5.878×10^{-4}	68.1	31.9
	Tan	0.371	0.008	0.544	$3.54 imes 10^{-2}$	1.256×10^{-3}	30.2	69.8
Target		0.342	0.023					

+ Log = log-sigmoid; Tan = tan-sigmoid.

best transfer function output at each cluster level was found to be consistently >0.5, thus indicating a reasonably good performance of the transfer function. The highest R value observed was 0.573 (Cluster 5 with the Tan transfer function). Also, in most cases, we observed that the systematic errors were less than the random errors. From a physical perspective this is quite intuitive in that the saturated water content (which includes gravitational, capillary, and hygroscopic water coexisting at saturation) depends on all of the pore space created by the various components of soil texture (sand, silt, and clay) and the bulk density, irrespective of the scale of support of the input data.

Residual Water Content

The results of statistical analyses between the observed and simulated values of θ_r are tabulated in Table 3. Based on correlation (*R*) between observed and simulated values, the Tan transfer function yielded a relatively better output for θ_r for most cases of spatial clustering. The Lin function yielded better outputs in two cases. We also found, however, that the *R* values for θ_r were generally low, with an overall highest value of 0.456 (for Cluster 8 with the Tan transfer function). On the other hand, the systematic errors were consistently much higher than the random errors. This indicates that there is a bias between the observed and simulated θ_r values. This bias, attributed to the difference in support scale between the training data and the application data, was easily corrected by linear regression, as mentioned above. Correcting for the bias brought the mean of the ANN-predicted values closer to that of the target. A sample illustration is provided in Fig. 4. Note, however, that no changes are apparent in the R values. In physical terms, this may suggest that the residual water content, dominated by hygroscopic water, depends only on the fine pore spaces created by certain fractions of soil particle sizes (Arya and Paris, 1981). Soil texture (sand, silt, and clay content) does not provide complete details of the particle size distribution or, in turn, the pore size distribution, including the arrangement, tortuosity, and connectivity of pores. This explains the limited success in predicting the residual soil water retention. It is noteworthy, however, that even when the R values are low at the individual point locations, the ANN–linear regression based PTF method provides a matching average prediction at the field scale (i.e., the average across 50 points at the Las Cruces trench site).

Water Content at Field Capacity

Performance analyses of ANN-based transfer functions for θ_f are shown in Table 4. The best correlation for θ_f was obtained with the Lin transfer function model for all but two of the clustering levels. For this parameter at an intermediate pressure range of the nonlinear soil water retention curve, however, we observed very low values of *R* compared with those for water contents under wet (θ_s) or dry (θ_r) conditions. In general, ranges of systematic errors or prediction bias for θ_f fell somewhere between those of θ_s and θ_r . These findings further suggest that θ_f predictions have more uncertainty than those for θ_s or θ_r . As with the residual water content, limited success of the ANN for the water content at field capacity can be attributed to having incomplete

TABLE 3. Statistical analysis of artificial neural network predicted fine-resolution residual volumetric soil water content values from Las Cruces trench site.

Cluster	Transfer function†	Mean	SD	R	RMSE	MSE	Random error	Systematic error
		— m ³ n	n ⁻³ —					% ———
1	Linear	0.052	0.007	0.173	3.08×10^{-2}	9.464×10^{-4}	11.6	88.4
	Log	0.052	0.006	-0.046	3.10×10^{-2}	9.631×10^{-4}	11.6	88.4
	Tan	0.055	0.006	0.054	2.75×10^{-2}	7.582×10^{-4}	14.8	85.2
2	Linear	0.057	0.007	0.158	2.63×10^{-2}	6.902×10^{-4}	16.0	84.0
	Log	0.049	0.006	0.285	3.30×10^{-2}	1.088×10^{-3}	9.5	90.5
	Tan	0.068	0.009	0.415	1.64×10^{-2}	2.676×10^{-4}	34.6	65.4
3	Linear	0.058	0.007	0.176	2.54×10^{-2}	6.432×10^{-4}	17.0	83.0
	Log	0.056	0.006	0.278	2.67×10^{-2}	7.118×10^{-4}	14.6	85.4
	Tan	0.066	0.011	0.445	1.79×10^{-2}	3.194×10^{-4}	28.1	71.9
4	Linear	0.051	0.007	0.178	3.15×10^{-2}	9.912×10^{-4}	11.0	89.0
	Log	0.052	0.005	0.164	2.98×10^{-2}	8.883×10^{-4}	12.4	87.6
	Tan	0.057	0.004	0.144	2.51×10^{-2}	6.322×10^{-4}	17.5	82.5
5	Linear	0.052	0.007	0.185	3.01×10^{-2}	$9.089 imes 10^{-4}$	12.0	88.0
	Log	0.055	0.006	0.233	2.75×10^{-2}	7.539×10^{-4}	14.1	85.9
	Tan	0.057	0.005	0.287	2.53×10^{-2}	6.394×10^{-4}	16.2	83.8
6	Linear	0.058	0.007	0.197	2.51 × 10 ⁻²	6.285×10^{-4}	17.3	82.7
	Log	0.052	0.006	0.020	3.08×10^{-2}	9.510×10^{-4}	11.8	88.2
	Tan	0.051	0.007	0.339	3.07×10^{-2}	$9.429 imes 10^{-4}$	10.5	89.5
7	Linear	0.054	0.007	0.197	2.88×10^{-2}	8.289×10^{-4}	13.1	86.9
	Log	0.058	0.005	0.262	2.47×10^{-2}	6.087×10^{-4}	17.2	82.8
	Tan	0.069	0.007	0.395	1.50×10^{-2}	2.254×10^{-4}	41.9	58.1
8	Linear	0.054	0.007	0.198	2.85×10^{-2}	$8.139 imes 10^{-4}$	13.3	86.7
	Log	0.062	0.007	0.388	$2.05 imes 10^{-2}$	4.194×10^{-4}	22.7	77.3
	Tan	0.071	0.009	0.456	1.42×10^{-2}	2.005×10^{-4}	44.2	55.8
9	Linear	0.056	0.007	0.219	2.64×10^{-2}	6.970×10^{-4}	15.4	84.6
	Log	0.052	0.005	0.016	3.08×10^{-2}	9.457×10^{-4}	12.0	88.0
	Tan	0.076	0.007	0.448	1.06×10^{-2}	1.132×10^{-4}	78.6	21.4
10	Linear	0.057	0.007	0.202	2.62×10^{-2}	6.842×10^{-4}	15.8	84.2
	Log	0.056	0.006	0.092	2.67×10^{-2}	$7.104 imes 10^{-4}$	15.7	84.3
	Tan	0.069	0.005	0.398	1.46×10^{-2}	2.125×10^{-4}	44.2	55.8
Target		0.080	0.011					

+ Log = log-sigmoid; Tan = tan-sigmoid.

www.vadosezonejournal.org · Vol. 6, No. 4, November 2007

soil particle size distribution data.

Overall, using soil texture and bulk density as inputs in the multiscale transfer function models, better predictions (as reflected by higher R values) were obtained for $\theta_{\rm c}$ than for $\theta_{\rm f}$ or $\theta_{\rm r}$. This is because of the fact that θ_s depends on the total pore space rather than specific pore arrangements, as is the case for θ_f and θ_r . The pore volume information is indirectly available to the ANN through the bulk density input. Various pore sizes and their distribution, including structural anomalies, can cause the field capacity and residual water content values to be different, even within the same soil type. The presence of organic matter may also influence their values. The fact that our ANN training does not account for the presence of macropores and mesopores (which affect θ_{f}) or organic matter could further explain some of the reductions in the correlation between observed and simulated values for θ_f and θ_r .

Construction of the Soil Water Retention Curve

Sample graphs for the fitted water retention curves for several point locations at the Las Cruces trench site are shown in Fig. 5. The "Observed" curve



Fig. 4. Illustration of the effect of bias correction. Application of bias correction by linear regression brings the mean of predicted values closer to the mean of target values.

corresponds to the values measured at the point location (from Wierenga et al., 1989). The "Target (vG)" curve corresponds to the van Genuchten equation fitted to the observed water content values using nonlinear regression. The "ANN-Pred (vG)" curve corresponds to the van Genuchten equation fit based on the ANN-predicted θ_{s} , θ_{p} and θ_{r} values, averaged across the 10 spatial clustering levels. The error bars of the ANN-predicted

retention curve show the limits of the variation in the predicted water content at the particular pressure head across the 10 different clusters of the training data. The ANN-predicted values closely matched the target van Genuchten curves and the error bars are quite small. Figure 6 shows the field-average soil water retention curve across the 50 sampling points (at 0–6-cm depth) at the Las Cruces trench site. It can be inferred that the predictions based on an ANN trained with coarse-scale SSURGO data matched the observations at the point to field scale.

As mentioned above, the coarsescale (1:24,000) inputs to the ANN at the subwatershed scale are representative values of a number of observations in a particular soil map unit. These values support a large spatial area. The point- or local-scale (1:1) inputs measured at a spacing of 0.5 m at the Las Cruces trench site have a much smaller support area. The factors that influence soil water retention at the smaller scale of support are different from those at the larger scale of support. Since the ANN model is not a physically based model, however, the use of ANNs to apply PTFs across different spatial scales of support can be viewed as a reasonable and feasible choice. This argument is supported by the "tightness" of the error bars on the ANN-predicted retention curve at the local scale. Having narrow error bars (Fig. 5 and 6) indicates that there is little variation in the predicted values of the water content with an increase in the area of support from which the training inputs were selected.

Figures 7 to 11 show the evolution of average values of the estimated van Genuchten soil water retention parameters α , n, θ_s , θ_r , and θ_f at the local (field or trench) scale using training data pooled from 10 clusters or spatial extents. Little variation was found in the average values of these parameters across these spatial extents. In other words, the effect of the size of the training data pool was minimal.

Validation of Methodology

The Rio Grande basin was the primary focal area for our study; however, the Las Cruces trench site data have the disadvantage of being very localized within the basin. The trench covers an area <120 m² (26.4 m long by 4.5 m

wide). Consequently, there is not much variation in the soil types encountered at the local (field or trench) scale. Hence, we decided to validate the ANN methodology with independent coarse- and fine-resolution data from a different region with different soil genesis and textural compositions.

The Little Washita (LW) watershed region in Oklahoma was selected for this validation study. Bootstrapping was again used to randomly select 1000 training data and 500 valida-

TABLE 4. Statistical analysis of artificial neural network predicted fine-resolution field capacity volumetric water content values from the Las Cruces trench site.

Cluster	Transfer function†	Mean	SD	R	RMSE	MSE	Random error	Systematic error
		$-m^3$	m ⁻³ —					% ——
1	Linear	0.134	0.006	0.178	1.49×10^{-2}	2.213×10^{-4}	71.6	28.4
	Log	0.134	0.010	0.167	1.65 × 10 ⁻²	2.723×10^{-4}	59.5	40.5
	Tan	0.148	0.011	0.127	2.68×10^{-2}	7.204×10^{-4}	22.5	77.5
2	Linear	0.126	0.007	0.198	1.33×10^{-2}	1.757×10^{-4}	89.7	10.3
	Log	0.131	0.005	0.013	1.42×10^{-2}	2.027×10^{-4}	80.2	19.8
	Tan	0.146	0.010	0.156	2.44×10^{-2}	$5.932 imes 10^{-4}$	26.7	73.3
3	Linear	0.125	0.006	0.190	1.32 × 10 ⁻²	1.750×10^{-4}	90.1	9.9
	Log	0.126	0.005	0.141	1.31 × 10 ⁻²	1.725×10^{-4}	92.4	7.6
	Tan	0.132	0.006	0.122	1.46 × 10 ⁻²	2.132×10^{-4}	75.1	24.9
4	Linear	0.126	0.006	0.196	1.31 × 10 ⁻²	1.720×10^{-4}	91.6	8.4
	Log	0.128	0.005	0.195	1.28×10^{-2}	1.633×10^{-4}	97.5	2.5
	Tan	0.127	0.003	0.141	1.27 × 10 ⁻²	1.610×10^{-4}	100.6	-0.6
5	Linear	0.129	0.007	0.204	1.34 × 10 ⁻²	$1.793 imes 10^{-4}$	87.7	12.3
	Log	0.137	0.007	0.223	1.69×10^{-2}	$2.853 imes 10^{-4}$	54.7	45.3
	Tan	0.139	0.005	0.146	1.77 × 10 ⁻²	3.141×10^{-4}	50.6	49.4
6	Linear	0.124	0.007	0.200	1.34×10^{-2}	1.786×10^{-4}	88.0	12.0
	Log	0.132	0.006	0.206	1.41 × 10 ⁻²	$1.978 imes 10^{-4}$	80.0	20.0
	Tan	0.134	0.010	-0.075	1.82 × 10 ⁻²	3.301×10^{-4}	49.5	50.5
7	Linear	0.137	0.007	0.200	1.66 × 10 ⁻²	2.746×10^{-4}	57.3	42.7
	Log	0.135	0.005	0.136	1.52 × 10 ⁻²	2.312×10^{-4}	69.0	31.0
	Tan	0.145	0.010	0.028	2.42×10^{-2}	$5.848 imes 10^{-4}$	27.8	72.2
8	Linear	0.130	0.007	0.203	1.37 × 10 ^{–2}	$1.879 imes 10^{-4}$	83.6	16.4
	Log	0.141	0.008	0.100	2.03×10^{-2}	4.127×10^{-4}	39.0	61.0
	Tan	0.144	0.009	0.070	2.28 × 10 ⁻²	5.184×10^{-4}	31.3	68.7
9	Linear	0.076	0.007	0.148	5.29×10^{-2}	2.801×10^{-3}	5.7	94.3
	Log	0.135	0.003	0.145	1.51 × 10 ^{–2}	2.288×10^{-4}	69.5	30.5
	Tan	0.149	0.009	0.134	2.70×10^{-2}	$7.279 imes 10^{-4}$	21.9	78.1
10	Linear	0.134	0.007	0.203	1.47 × 10 ⁻²	2.161×10^{-4}	72.7	27.3
	Log	0.135	0.005	0.200	1.50 × 10 ⁻²	2.262×10^{-4}	69.1	30.9
	Tan	0.148	0.007	0.077	2.51×10^{-2}	$6.306 imes 10^{-4}$	25.6	74.4
Target		0.127	0.013					

† Log = log-sigmoid; Tan = tan-sigmoid



Fig. 5. Sample soil water retention curves: ψ is the pressure head and θ the water content; NN-Pred are the neural network predicted values. Values at each point are averaged across 10 clustering levels. Error bars show deviation in $\theta(\psi)$ values across the clustering levels.

tion data from a pool of >36,000 SSURGO values from the Little Washita region. The ANNs were trained using the soil physical properties as the inputs and the soil water contents as the targets for the coarse-scale SSURGO data. Using the developed (trained and validated) ANNs with coarse-resolution data, the LW soil property data at 70 locations (Mohanty et al., 2002) were used as the fine-scale inputs for the ANN prediction. The outputs obtained were subjected to linear bias correction after checking for normality of errors, while the soil water retention curve was constructed (Fig. 12) using the same procedure as described for the Rio Grande basin. It is evident that the bias-corrected curve once again closely matches the target values. This finding further attests to the validity of the proposed method for different test sites involving different soil genesis, topography, vegetation, and hydroclimatic conditions, and varying extent of data coverage.

Any predictive model can only be as good as the input data supplied to it. The error bars in the two figures (Fig. 12 and 13) also reflect this data dependency of the ANN model. The error bars for the ANN-predicted and bias-corrected values at LW (Fig. 12) are comparable to those of the target data. This implies that the ANN is able to capture much of the variation in the soil water content values. This ability is provided to the ANN by the wide range of textures (soil types) present in the simulation inputs from the LW region. On the other hand, the soils at the Las Cruces trench site do not vary much from one measurement location to the next (Table 5). Hence, there is hardly any variation in the simulation inputs. This invariance is then passed on to the predictions. The error bars in Fig. 13 are consequently much smaller for the ANN-predicted and bias-corrected values compared with the target. The distributions of the soil physical properties for the LW region are given



FIG. 6. Average soil water retention curve for the Rio Grande region: ψ is the pressure head and θ the water content; NN-Pred are the neural network predicted values. Values are averaged across 10 clustering levels and across the 50 ground points. Error bars show deviation in $\theta(\psi)$ values across the clustering levels.

in Table 6. It is evident that the fine-scale training data for the LW watershed in the region of the Southern Great Plains 1997 hydrology experiment have much greater variability than the fine-scale data of the Las Cruses trench site in the Rio Grande basin. Furthermore, for the LW region, the statistics of the fine-scale soil hydraulic properties are comparable to the statistics of the coarse-scale data, leading to better predictions with the multiscale PTFs.



Fig. 7. Variation across clustering levels in estimated van Genuchten α values averaged across 50 ground points.



Fig. 8. Variation across clustering levels in estimated van Genuchten n values averaged across 50 ground points.



Fig. 9. Variation across clustering levels in estimated saturated volumetric water content (θ_s) values averaged across 50 ground points.

Conclusions

Using coarse-scale soil property data from the SSURGO database and local-scale soil property data, it has been shown that ANNs can be applied across spatial scales for estimating soil hydraulic properties at the local or field scale while being trained on coarse-scale input data. The simulated soil hydraulic parameters can be further corrected for bias by decomposing



Fig. 10. Variation across clustering levels in estimated residual volumetric water content (θ_r) values averaged across 50 ground points.



Fig. 11. Variation across clustering levels in estimated volumetric water content at field capacity (θ_f) values averaged across 50 ground points.



Fig. 12. Average soil water retention curve for the Little Washita (LW) watershed region: ψ is the pressure head; θ is the soil water content; ANN-Pred are the artificial neural network predicted values before bias correction; BC are the bias-corrected values. Values are averaged across the 70 ground points. Error bars show deviation in $\theta(\psi)$ values across the 70 ground points.

the estimation errors into random and systematic components. The systematic error, attributed to scale differences between the training data set and the simulation application data set, can be eliminated by linear regression. A limitation of the methodology at present is that one would need to know a few "expected" values to use in the bias correction process. The proposed methodology was validated using information from two test sites with different hydrologic characteristics, the Las Cruces trench site in New Mexico and the Little Washita watershed in Oklahoma. Further improvement in the ANN-based predictions across multiple spatial scales may be possible by using Bayesian statistical techniques or genetic algorithms in the ANN training process. Extension of our methodology to data with larger support scales (e.g., using remote sensing techniques) could also help in bet-



Fig. 13. Average soil water retention curve for the Rio Grande region: ψ is the pressure head; θ is the soil water content; ANN-Pred are the artificial neural network predicted values before bias correction; BC are the bias-corrected values. Values are averaged across the 50 ground points. Error bars show deviation in $\theta(\psi)$ values across the 50 ground points.

 $\mathsf{T}\mathsf{ABLE}$ 5. Soil physical properties at the coarse and fine scales for the Rio Grande Basin.

SSURGO	Sa	ind	S	ilt	CI	ay	Bulk d	ensity
data	Mean	SD	Mean	SD	Mean	SD	Mean	SD
				%				_
Cluster 1	49.43	21.49	29.20	19.39	21.37	10.87	1.60	0.12
Cluster 2	48.38	21.48	30.79	13.94	20.83	10.38	1.66	0.19
Cluster 3	47.71	21.75	31.79	14.33	20.41	10.05	1.56	0.14
Cluster 4	50.54	21.50	30.05	14.70	19.41	9.55	1.55	0.13
Cluster 5	57.61	21.50	26.17	16.41	16.22	7.02	1.51	0.13
Cluster 6	49.30	21.09	30.91	14.49	19.79	9.32	1.54	0.13
Cluster 7	49.65	21.88	29.92	14.58	20.43	10.85	1.56	0.13
Cluster 8	50.05	21.53	30.12	14.59	19.82	10.00	1.54	0.13
Cluster 9	48.77	20.54	31.31	13.98	19.92	9.54	1.52	0.14
Cluster 10	50.05	19.52	30.33	13.34	19.62	8.67	1.52	0.14
Average	50.15	21.23	30.06	14.37	19.78	9.63	1.54	0.14
Las Cruces								
trench site point data	81.46	1.67	9.76	1.17	8.78	1.20	1.66	0.05

TABLE 6. Soil physical properties at the coarse and fine scales for the Little Washita watershed.

Dete equiree	Sand		Silt		Clay		Bulk density	
Data source	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	0/_							
				,	0			
SSURGO data	43.31	21.82	37.74	17.31	19.17	8.80	1.78	0.42
Point data	51.91	21.11	33.61	16.41	14.48	6.04	1.40	0.10

ter understanding the effects of the scale difference and, subsequently, in creating a more generalized multiscale ANN pedotransfer model.

ACKNOWLEDGMENTS

We would like to acknowledge support of the Los Alamos National Lab.–Sustainability of Semi-Arid Hydrology and Riparian Areas, NASA (Grant no. 35410), NASA Earth System Science fellowship (NNX06AF95H), the National Science Foundation (CMG/ DMS Grant 0621113), and National Science Foundation-MRI grant (0216275) for this work.

References

- Ahuja, L.R., J.W. Naney, and R.D. Williams. 1985. Estimating soil water characteristics from simpler properties or limited data. Soil Sci. Soc. Am. J. 49:1100–1105.
- Arya, L.M., and J.F. Paris. 1981. A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. Soil Sci. Soc. Am. J. 45:1023–1030.
- Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. Hydrol. Pap. 3. Colorado State Univ., Fort Collins.
- Campbell, G.S. 1974. A simple method for determining unsaturated hydraulic conductivity from moisture retention data. Soil Sci. 177:311–314.
- Clapp, R.B., and G.M. Hornberger. 1978. Empirical equations for some hydraulic properties. Water Resour. Res. 14:601–604.
- Demuth, H., M. Beale, and M. Hagan. 2005. Neural network toolbox users guide. The Mathworks, Natick, MA.
- Ines, A.V.M., and J.W. Hansen. 2006. Bias correction of daily GCM rainfall for crop simulation studies. Agric. For. Meteorol. 138:44–53.
- Jana, R.B., B.P. Mohanty, and E.P. Springer. 2005. Soil hydrologic properties for simulation of semi-arid river basin water balance: A report. Texas A&M Univ., College Station.
- Kanamaru, H., and M. Kanamitsu. 2007. Scale-selective bias correction in a downscaling of global analysis using a regional model. Mon. Weather Rev. 135:334–350.
- Leij, F.J., N. Romano, M. Palladino, and M.G. Schaap. 2004. Topographical attributes to predict soil hydraulic properties along a hillslope transect. Water Resour. Res. 40:1–15.
- Mohanty, B.P., P.J. Shouse, D.A. Miller, and M.Th. van Genuchten. 2002. Soil property database: Southern Great Plains 1997 hydrology experiment. Water Resour. Res. 38(5):1047, doi:10.1029/2000WR000076.
- Mohanty, B.P., and J. Zhu. 2007. Effective averaging schemes for hydraulic parameters in horizontally and vertically heterogeneous soils. J. Hydrometeorol. 8:715–729.
- Natural Resources Conservation Service. 2001. National soil survey handbook, title 430-VI (2001). Nat. Resour. Conserv. Serv., Washington, DC.
- Nemes, A., M.G. Schaap, and J.H.M. Wösten. 2003. Functional evaluation of pedotransfer functions derived from different scales of data collection. Soil Sci. Soc. Am. J. 67:1093–1102.
- Pachepsky, Ya.A., W.J. Rawls, and D.J. Timlin. 1999. The current status of pedotransfer functions, their accuracy, reliability, and utility in fieldand regional-scale modeling. p. 223–234. *In* D.L. Corwin et al. (ed.) Assessment of non-point source pollution in the vadose zone. Geophys. Monogr. 108. Am. Geophys. Union, Washington, DC.
- Pachepsky, Ya.A., D.J. Timlin, and W.J. Rawls. 2001. Soil water retention as related to topographic variables. Soil Sci. Soc. Am. J. 65:1787–1795.
- Pachepsky, Ya.A., D. Timlin, and G. Várallyay. 1996. Artificial neural networks to estimate soil water retention from easily measurable data. Soil Sci. Soc. Am. J. 60:727–773.
- Rawls, W.J., D.L. Brakensiek, and K.E. Saxton. 1982. Estimation of soil water properties. Trans. ASAE 25:1316–1320.
- Rawls, W.J., T.J. Gish, and D.L. Brakensiek. 1991. Estimating soil water retention from soil physical properties and characteristics. Adv. Soil Sci. 16:213–234.
- Schaap, M.G., F.J. Leij, and M.Th. van Genuchten. 1998. Neural network analysis for hierarchical prediction of soil hydraulic properties. Soil Sci. Soc. Am. J. 62:847–855.
- Sharma, S.K., B.P. Mohanty, and J. Zhu. 2006. Including topography and vegetation attributes for developing pedotransfer functions. Soil Sci. Soc. Am. J. 70:1430–1440.

- Soil Survey Staff. 2007. Soil survey geographic (SSURGO) database. Available at www.ncgc.nrcs.usda.gov/products/datasets/ssurgo/ (verified 17 Oct. 2007). NRCS, Washington, DC.
- Tomasella, J., M.G. Hodnett, and L. Rossato. 2000. Pedotransfer functions for the estimation of soil water retention in Brazilian soils. Soil Sci. Soc. Am. J. 64:327–338.
- Tomasella, J., Ya.A. Pachepsky, S. Crestana, and W.J. Rawls. 2003. Comparison of two techniques to develop pedotransfer functions for water retention. Soil Sci. Soc. Am. J. 44:1085–1092.
- van Genuchten, M.Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44:892–898.
- van Genuchten, M.Th., and F.J. Leij. 1992. On estimating the hydraulic properties of unsaturated soils. p. 1–14. *In* M.Th. van Genuchten et al. (ed.) Indirect methods for estimating the hydraulic properties of unsaturated soils. Proc. Int. Worksh. on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA. 11–13 Oct. 1989. U.S. Salinity Lab., Riverside, CA.
- Vereecken, H., J. Maes, J. Feyen, and P. Darius. 1989. Estimating the soil moisture retention characteristics from texture, bulk density, and carbon content. Soil Sci. Soc. Am. J. 1484:389–403.
- Wierenga, P.J., D. Hudson, J. Vinson, M. Nash, A. Toorman, and R.G. Hills. 1989. Soil physical properties at the Las Cruces trench site. NUREG/CR-5441. U.S. Nuclear Regulatory Commission, Washington, DC.
- Willmott, C.J. 1982. Some comments on the evaluation of model performance. Bull. Am. Meteorol. Soc. 63:1309–1313.
- Winter, C., E. Springer, K. Costigan, P. Fasel, S. Mniszewski, and G. Zyvoloski. 2004. Virtual watersheds: Simulating the water balance of the Rio Grande Basin. Comput. Sci. Eng. 6(3):18–26.
- Wösten, J.H.M., Ya.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.
- Zhu, J., and B.P. Mohanty. 2002a. Spatial averaging of van Genuchten hydraulic parameters for steady state flow in heterogeneous soils. Vadose Zone J. 1:261–271.
- Zhu, J., and B.P. Mohanty. 2002b. Upscaling of soil hydraulic properties under steady state evaporation and infiltration. Water Resour. Res. 38:1178, doi:10.1029/2001WR000704.