

## Functional Evaluation of Pedotransfer Functions Derived from Different Scales of Data Collection

A. Nemes,\* M. G. Schaap, and J. H. M. Wösten

### ABSTRACT

Estimation of soil hydraulic properties by pedotransfer functions (PTFs) can be an alternative to troublesome and expensive measurements. New approaches to develop PTFs are continuously being introduced, however, PTF applicability in locations other than those of data collection has been rarely reported. We used three databases were used to develop PTFs using artificial neural networks (NNs). Data from Hungary were used to derive *national scale* soil hydraulic PTFs. The HYPRES database was used to develop *continental scale* PTFs. Finally, a database containing mostly American and European data was used to develop *intercontinental scale* PTFs. For each database, 11 PTFs were developed that differed in detail of input data. Accuracy of the estimations was tested using independent Hungarian data. First, soil water retention at nine values of matric potential were estimated. Root mean squared residuals (RMSRs) using different inputs ranged from 0.02 to 0.06 m<sup>3</sup> m<sup>-3</sup> for national scale PTFs, while international scale PTFs had RMSRs from 0.025 to 0.088 m<sup>3</sup> m<sup>-3</sup>. Estimated water retention curves (WRCs) were then used to simulate soil moisture time series of seven Hungarian soils. Root mean squared residuals during a growing season ranged from 0.065 to 0.07 m<sup>3</sup> m<sup>-3</sup>, using different PTF estimates. Simulations using laboratory-measured WRCs had RMSR of 0.061 m<sup>3</sup> m<sup>-3</sup>. Such small differences in the accuracy of simulations make international PTFs an alternative to national PTFs and measurements. However, testing of the international PTFs with a specific model for specific soil and land use remains desirable because of uncertainty in soil representation in such databases.

**M**ODELING WATER and solute transport has become an important tool in simulating agricultural productivity as well as environmental quality. The use of models, however, is often limited by the lack of accurate information on soil hydraulic properties. Soil water and solute transport models typically require data on soil water retention and hydraulic conductivity. Measurement of these properties is relatively time-consuming and costly, especially when data are needed for large areas of land. For many applications, the estimation of hydraulic characteristics with PTFs can be an alternative as most of the PTFs use input data that are easily and/or routinely collected.

A prerequisite for the development of PTFs is the availability of a source database that contains potential predictors as well as hydraulic properties. Most PTFs available in the literature use soil texture, bulk density

(D<sub>b</sub>), and organic matter (OM) contents as predictors; additional parameters are rarely used (Rawls et al., 1991; Wösten et al., 2001). Estimation of hydraulic characteristics is mostly limited to water retention points or parameters and saturated hydraulic conductivity. A small number of PTFs were proposed for the estimation of unsaturated hydraulic conductivity. Pedotransfer functions are usually published in a tabular form for particular soil classes, as linear or nonlinear regression equations, or more recently, distributed as computer codes resulting from NN analysis (e.g., Schaap et al., 2001). For an overview of the current status of PTFs, we refer to Wösten et al. (2001).

Many PTFs have been developed in recent decades. Independent databases have been used to evaluate various PTFs that were developed elsewhere. For example, Tietje and Tapkenhinrichs (1993) and Kern (1995) evaluated different PTFs for the estimation of water retention. Tietje and Hennings (1996) tested PTFs for the estimation of saturated conductivity. Some recent comparisons include Imam et al. (1999), Cornelis et al. (2001), and Wagner et al. (2001). Imam et al. (1999) compared three PTFs to compute the water holding capacity of inorganic soils. Cornelis et al. (2001) compared nine PTFs to estimate the soil moisture retention curve. Wagner et al. (2001) evaluated the performance of eight PTFs to estimate unsaturated soil hydraulic conductivity. The latter two studies ranked PTFs, noting that the PTF performance in both cases could be influenced by the geographical preference of the source data sets.

A limitation of most studies that evaluate PTFs is that it remains unclear what the main sources of the estimation errors are. In those studies it is not clear whether differences between data sets used to derive PTFs (size, origin, reliability), differences between the algorithm of PTF development (e.g., different regression types vs. NN models) or differences among the predictors cause a particular PTF to perform better than others. Schaap and Leij (1998) cross-validated NN models by developing PTFs using the same algorithm and the same predictors on data of three independent databases. They found that PTFs derived from one database gave systematically different estimations for the other two data sets, but that estimations improved somewhat when PTFs were derived from all available data. They concluded that the performance of a PTF depends on both the derivation and evaluation data sets and that origin,

A. Nemes, USDA-ARS Hydrology and Remote Sensing Lab., 10300 Baltimore Ave. Bldg. 007 Rm. 124, BARC-WEST, Beltsville, MD 20705-2350, USA; M.G. Schaap, USDA-ARS, George E. Brown Jr. Salinity Lab., 450 W. Big Springs Road, Riverside, CA 92507; J.H.M. Wösten, ALTERRA Green World Res., P.O. Box 47, 6700 AA Wageningen, The Netherlands. Received 22 Apr. 2002. \*Corresponding author (anemes@hydrolab.arsusda.gov).

**Abbreviations:** D<sub>b</sub>, bulk density; OM, organic matter; NN, neural network; PTF, pedotransfer function; MR, mean residual; RMSR, root mean squared residual; WRC, water retention curve.

size, and other data characteristics may determine the performance of PTFs.

Water retention and hydraulic conductivity are not the final aim but are intermediate characteristics needed to calculate other soil properties with more practical meaning. Functional evaluation of estimated soil hydraulic data helps to characterize the contribution of such data to the inaccuracy and uncertainty of simulations. A number of studies used soil water simulation models to evaluate the performance of estimated soil hydraulic characteristics through the simulation of different aspects of soil behavior (e.g., Wösten et al., 1995; Espino et al., 1996; Hack-ten Broeke and Hegmans, 1996; van Alphen et al., 2001; Soet and Stricker, 2003).

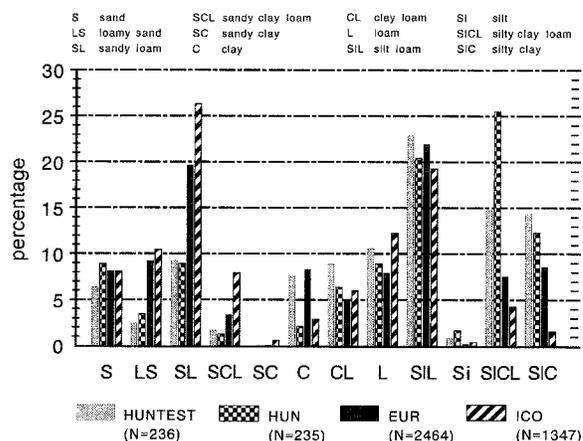
Many countries or regions in the World do not have a sufficient amount of soil hydraulic data for agricultural or environmental modeling purposes or to develop PTFs. Nations or regions with appropriate resources, however, have collected considerable amounts of soil information. Most often, these data are used in individual studies at local or national scale. However, in recent years, development of international databases (e.g., Wösten et al., 1999; Nemes et al., 2001) enabled, among other possibilities, the development of international PTFs. A potential benefit of having extensive international databases is that those may include, but are not limited to, soils that are similar in their properties and were developed under similar soil-forming conditions to the soils of the area of planned application.

The objective of this study was to test the hypothesis that a PTF developed from an international database cannot be used at smaller, national scale, and country- or region-specific PTF is necessary. We developed PTFs for the estimation of water retention using data stored in two international databases and developed similar country-specific PTFs using data from Hungary. All PTFs were tested using a second, independent Hungarian data set. The performance of international and national PTFs were first tested on sum of square residuals between measured and estimated retention characteristics. Next, we used different PTF estimates to simulate soil moisture time series of seven Hungarian soils. Sum of square residuals were evaluated between simulated water contents and water contents observed in the field.

## MATERIALS AND METHODS

### Data Sets

Three databases were used to provide soil hydraulic data collected at three different scales. The HUNSODA database (Nemes, 2001) comprises soil data collected solely in Hungary. The database holds soil water retention characteristics for 576 soil horizons. The HYPRES database (Wösten et al., 1999) contains data of 12 countries of Europe, and holds measured soil hydraulic characteristics for 4030 soil horizons. A third intercontinental database provided data from several parts of the World. This database was previously used for the development of the ROSETTA pedotransfer program (Schaap et al., 2001) and contains 2134 soil samples derived from the UNSODA database (Nemes et al., 2001) and two other databases that originate from the United States (cf. Schaap and Leij, 1998).



**Fig. 1.** Textural composition of the test data set (HUNTEST) and the Hungarian (HUN), European (EUR), and intercontinental (ICO) data sets of pedotransfer function development according to the USDA soil textural classification system (USDA, 1951). The number of samples ( $N$ ) in each data set are in parentheses.

All three databases were filtered to select soils that are (i) mineral soils; (ii) have data available on soil texture,  $D_b$ , and OM content; and (iii) have at least four measured soil water retention points [ $\theta(h)$ ]. This selection left us with Hungarian ( $N = 471$ ), European (EUR,  $N = 2464$ ), and intercontinental (ICO,  $N = 1347$ ) data sets. The Hungarian data set was further split randomly to generate a data set to develop PTFs (HUN,  $N = 235$ ) and a data set to test PTFs (HUNTEST,  $N = 236$ ). The HUN, EUR, and ICO data sets were used to formulate two additional data sets to develop PTFs. Because Hungarian data were not present in either of the international scale databases, the assumption is that any information developed from these databases would not be valid for Hungary. To investigate the effect of data from a country being represented in an international database we also combined a copy of the HUN data set with a copy of each of the two International data sets to form two new data sets later referred to as EUR + HUN ( $N = 2699$ ) and ICO + HUN ( $N = 1582$ ).

Figure 1 shows the number of samples and the textural composition in the HUN, EUR, and ICO data sets that are used to develop PTFs and in HUNTEST, the data set used for testing PTF performance. Texture classes are not represented equally in all data sets. The two international data sets hold a considerably larger proportion of loamy sand, sandy loam, sandy clay loam samples, whereas soils with silty clay and silty clay loam texture are better represented in the Hungarian data. Sandy clays and silts are, however, poorly represented in all three data sets. Table 1 shows the summary statistics of some variables of the above data sets that were used in the PTF development and evaluation. Differences in texture and OM content are apparent, with decreasing average silt, clay, and OM contents from the Hungarian data through the European to the intercontinental scale data. Data in the different data sets were obtained from different sources and had different numbers and positions of points at the WRCs. To obtain uniform description of all the WRCs, the volumetric soil water content,  $\theta$ , as a function of matric potential,  $h$ , was described with the van Genuchten equation (van Genuchten, 1980):

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

where subscripts  $r$  and  $s$  refer to residual and saturated values, and  $\alpha$ ,  $n$ , and  $m$  are curve shape parameters, where  $m = 1 - 1/n$ . Parameters of the above equation were fitted to

**Table 1.** Mean, SD, and median of some soil properties of the test data set (HUNTEST) and the Hungarian (HUN), European (EUR), and intercontinental (ICO) data sets.

Property	Unit	HUNTEST			HUN			EUR			ICO		
		Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
USDA sand	%	29.00	26.38	21.15	30.79	28.76	20.80	40.12	31.09	31.66	51.78	26.49	51.90
USDA silt	%	43.72	17.99	47.85	44.40	19.93	51.40	37.75	21.93	35.85	31.59	21.82	30.50
USDA clay	%	27.28	16.32	26.11	24.81	14.68	25.80	22.14	17.17	18.00	16.63	12.36	14.20
Organic matter	g kg <sup>-1</sup>	17.72	16.66	16.45	18.43	16.67	17.20	15.93	19.31	9.50	13.82	22.07	3.45
Bulk density	Mg m <sup>-3</sup>	1.43	0.16	1.47	1.44	0.16	1.46	1.45	0.22	1.49	1.40	0.25	1.44
$\theta_{air}^{\dagger}$	m <sup>3</sup> m <sup>-3</sup>	0.47	0.06	0.46	0.46	0.06	0.45	0.43	0.08	0.42	0.41	0.09	0.41
$\theta$ (-10 kPa) <sup>†</sup>	m <sup>3</sup> m <sup>-3</sup>	0.40	0.09	0.40	0.38	0.08	0.39	0.33	0.12	0.34	0.31	0.12	0.32
$\theta$ (-33 kPa) <sup>†</sup>	m <sup>3</sup> m <sup>-3</sup>	0.35	0.10	0.36	0.33	0.10	0.34	0.29	0.12	0.29	0.25	0.11	0.26
$\theta$ (-1500 kPa) <sup>†</sup>	m <sup>3</sup> m <sup>-3</sup>	0.20	0.09	0.20	0.18	0.08	0.19	0.17	0.10	0.16	0.11	0.06	0.10

<sup>†</sup> Derived from fitted van Genuchten parameters.

the individual WRCs using the Simplex method (Nelder and Mead, 1965). Average water contents at four different matric potentials (0 kPa, -10 kPa, -33 kPa, -1500 kPa) show a decreasing trend from the Hungarian to the intercontinental data, as can also be expected from the differences in texture and OM content. Besides being geographically more diverse than the EUR data set, the ICO data set is also the one being more distant from the HUN data set in its properties, based on Table 1.

### Neural Networks

Pedotransfer functions have been developed using a wide variety of techniques (cf. Rawls et al., 1991; Wösten et al., 2001). One recently used technique is the analysis by artificial NNs (e.g., Pachepsky et al., 1996; Tamari et al., 1996; Minasny et al., 1999). As most studies found that the predictive capabilities of NN PTFs were equivalent or superior to different regression-type PTFs, we used NNs in this study.

A NN model consists of many simple computing elements (neurons or nodes) that are organized into subgroups (layers) and are interconnected as a network by weights. A model typically consists of an input layer, an output layer, and one (or more) *hidden* layer(s) that connect(s) the input and output layers. The number of nodes in the input and output layers correspond to the number of input and output variables of the model, the number of hidden nodes can be varied freely. Data flow goes from the input layer through the hidden layer(s) to the output layer. A node in the hidden and output layers receives multiple inputs, typically from all nodes of the previous layer. Within the node, each input is weighted and combined to produce a single value as the output of that node, which is then directed to all the nodes of the next layer, or outputted if it was a node of the output layer. The weight matrices are obtained through a calibration (training) procedure, which can then be used to make estimations on independent data. For a more thorough description on NNs, we refer the reader to Hecht-Nielsen (1990) or Haykin (1999).

Following Schaap and Leij (1998), we used a three-layer back-propagation NN model. The number of nodes in the hidden layer was set to six. Eleven different models were developed to estimate water retention through the parameters of the van Genuchten equation, separately from each of the five data sets outlined above. This is to avoid a possible bias while applying one particular set of input parameters. Models were built up gradually, from simple to more complex, using the most commonly used predictors. These predictors were sand, silt, and clay contents,  $D_b$ , OM contents, as well as water retention points [ $\theta(h)$ ] at three different matric potentials ( $h = -10, -33, -1500$  kPa). Details on the input parameters used in each model can be seen in Table 2.

NNs were combined with the data selection procedure of

the bootstrap method (Efron and Tibshirani, 1993) to generate internal calibration-validation data set pairs for an early stopping procedure. The bootstrap method is a nonparametric technique that simulates alternative (replica) data sets out of a single data set. Given a data set of size  $N$ , the bootstrap method generates replica data sets, also of size  $N$ , by random selection with replacement. Some samples are included more than once, while others are not selected into a particular replica data set. The replica data set is used to calibrate the NN model while data not in the replica data set are used for validation to stop the calibration process when a minimum error is reached. Multiple realization of subsets can help to avoid bias toward any particular calibration-validation data set pairs. We generated 50 replica data sets, each of which was used to calibrate the NN models. This procedure provided 50 subestimates that could be slightly different from each other. The final estimate of a PTF for each value was then calculated by averaging the 50 subestimates of the value. All NN modeling was performed with the Neural Network Toolbox in MATLAB (Demuth and Beale, 1992).

### Functional Evaluation using Simulated Water Content Time Series

Data on water regime in seven Hungarian soils were used to evaluate PTFs in their ability to provide parameters for water transport modeling. Soil water contents were measured in a Dystric Haplustept (GDL; Farkas et al., 1999), an Aquic Kandiuustalf (DHR2), and two Aquic Calcicustepts (DHR3 and DHR4; Czinege, 2000), an Udertic Haplustoll (KM1), a Pachic Udertic Haplustoll (KM1K; Tóth and Várallyay, 2001), and a Leptic Natrustoll (NYL249; Tóth and Kuti, 2002). These profiles represent arable land as well as pasture. Five different crops covered the seven fields. The profiles had three to five distinct genetic horizons. Data on soil properties and land use are collected in Table 3.

**Table 2.** Input parameters of the various neural network models. SSC, sand, silt, and clay content, %;  $D_b$ , bulk density, Mg m<sup>-3</sup>; OM, organic matter content, g kg<sup>-1</sup>;  $\theta(x)$  is soil water content (m<sup>3</sup> m<sup>-3</sup>) at matric potential  $x$  (-kPa).

Model	Input variables
M1	SSC
M2	SSC + BD
M3	SSC + BD + OM
M4	SSC + BD + $\theta$ (1500)
M5	SSC + BD + $\theta$ (33)
M6	SSC + BD + $\theta$ (10)
M7	SSC + BD + $\theta$ (10), $\theta$ (33)
M8	SSC + BD + $\theta$ (10), $\theta$ (1500)
M9	SSC + BD + $\theta$ (33), $\theta$ (1500)
M10	SSC + BD + $\theta$ (10), $\theta$ (33), $\theta$ (1500)
M11	SSC + BD + OM + $\theta$ (10), $\theta$ (33), $\theta$ (1500)

**Table 3. Properties of soils used in the simulations.**

Soil	Soil taxonomy	Crop	USDA sand		USDA silt		USDA clay		Organic matter		Bulk density		
			Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	
			%									— Mg m <sup>-3</sup> —	
			g kg <sup>-1</sup>										
GDL	Dystric Haplustepts	Corn	63.2	65.4	19.0	20.5	14.1	16.7	15.0	28.1	1.56	1.66	
DHR2	Aquic Kandistalfs	Alfalfa	21.5	56.9	24.6	44.9	18.4	33.6	6.6	25.3	1.58	1.68	
DHR3	Aquic Calcustepts	Alfalfa	30.1	95.9	1.8	35.5	1.5	34.5	1.0	19.1	1.48	1.74	
DHR4	Aquic Calcustepts	Alfalfa	50.6	96.1	1.8	26.8	2.1	22.7	1.0	13.2	1.49	1.58	
KM1	Udertic Haplustolls	Sunflower	2.9	4.5	53.9	57.2	39.3	41.7	9.4	53.2	1.39	1.53	
KM1K	Pachic Udertic Haplustolls	Winter wheat	3.3	5.1	55.2	60.7	35.9	40.3	0.7	64.8	1.40	1.52	
NYL249	Leptic Natrustolls	Pasture	3.6	7.3	54.4	67.1	27.2	42.1	2.7	15.6	1.46	1.66	

Soil water contents were measured at five to 10 depths per profile (to a maximum depth of 80–100 cm), eight to 10 times a year, using auger. Simulations of one-dimensional flow were performed using the SWAP model (version 2.07d) (van Dam et al., 1997). The lower boundary conditions were set as: free drainage for GDL, and measured ground water levels given as input for the other six profiles. Ground water levels were measured at variable intervals. Linear interpolation was performed to derive daily values of groundwater levels as input. The upper boundary conditions were controlled by daily weather data collected from nearby weather stations. Simulation year was 1997 for GDL, 2000 for the three DHR soils and 1999 for KM1, KM1K, and NYL249. The simple crop growth routine of SWAP was used for each field. Factors to characterize plant growth were derived by adjusting general factors suggested by van Dam et al. (1997), Tiktak et al. (2000), and therein to local conditions. Soil hydraulic characteristics were described according to the Mualem-van Genuchten model. Measured saturated hydraulic conductivity ( $K_s$ ) coupled with an assumed  $L = 0.5$ , as suggested by Mualem (1976), was used to describe unsaturated hydraulic conductivity [ $K(h)$ ] for each soil. Simulations were run to calculate the soil-water profile of each soil, using different WRCs, as estimated by the various PTFs, but keeping two of the parameters (i.e.,  $K_s$  and  $L$ ) constant for each run with the same soil. As a control, we also ran the same simulations using laboratory-measured WRCs, and applying the same  $K_s$  and  $L$  values as above. Simulated moisture content data were then compared with measured data at corresponding depths and dates. Two weights were introduced to account for the different number of layers and days at which moisture contents were measured, to allow each soil to contribute equally to the averaged values.

### Evaluation Criteria

The calibrated NN models were used to make estimates of the van Genuchten parameters of the 236 soils of the HUNTEST data set. In turn, these parameters were converted to water contents at the matric potentials that correspond to those available for the original WRC measurements. Accuracy of the estimations was evaluated using two measures. The mean residual (MR) can quantify systematic errors between measurements and estimations and the RMSRs can give the accuracy of the estimations in terms of standard deviations. These measures are calculated as:

$$MR = (1/N) \sum_{i=1}^N (\theta_i - \hat{\theta}_i) \quad [2]$$

and

$$RMSR = \sqrt{(1/N) \sum_{i=1}^N (\theta_i - \hat{\theta}_i)^2} \quad [3]$$

where  $N$  is the number of estimated and measured values,  $\theta$

and  $\hat{\theta}$  are measured and estimated water contents, respectively.

## RESULTS AND DISCUSSION

### Estimation of Water Retention Curves

Results of estimations with the 11 models are summarized in Fig. 2. Root mean squared residuals in Fig. 2a exhibit a trend of improvement from Model M1 to Model M11. More input information generally leads to better estimates. When one retention point was included in the list of inputs, the water content at  $-1500$  kPa was the worst additional predictor as compared with water contents at  $-33$  and  $-10$  kPa. A possible reason for this is that water content at  $-1500$  kPa is more dependent on soil texture than on soil structure, whereas water contents at  $-33$  and  $-10$  kPa are also largely influenced by soil structure. Soil structure is known to have influence on water conditions and as soil texture is among the basic input parameters for all models in our study, it seems to be more beneficial to include such additional input parameters that have a relation to soil structure as well. Estimations using data from the HUN data set show RMSR values ranging from 0.02 to 0.06 m<sup>3</sup> m<sup>-3</sup>, with only the soil texture (M1) model having RMSR  $>0.045$  m<sup>3</sup> m<sup>-3</sup>. These are relatively accurate estimations when compared with estimations that appear in literature (Wösten et al., 2001). For models M1 to M4, estimations using the EUR data set provide RMSR values that are 0.02 m<sup>3</sup> m<sup>-3</sup> greater than the RMSR using the HUN data set. Difference between the RMSR of the EUR and HUN data sets using any other models was  $\approx 0.01$  m<sup>3</sup> m<sup>-3</sup>. The accuracy of estimations was improved marginally for all models when Hungarian data were added to the European data set (i.e., the EUR + HUN data set was used). With the ICO data, RMSR values range from 0.045 to almost 0.09 m<sup>3</sup> m<sup>-3</sup>; that is, 0.02 to 0.04 m<sup>3</sup> m<sup>-3</sup> greater than RMSR values of PTFs developed using the HUN data set. There are, however, large improvements in some cases when the Hungarian data were added to the ICO data set (ICO + HUN). This is especially visible for M3, where OM content was added to the models. Although the representation of Hungarian data in the ICO + HUN data set is  $\approx 15\%$ , thus greater than in the EUR + HUN data set, it only improved estimations somewhat, but even then it never performed better than any of the smaller scale data sets. This is presumably due to the fact that many soils in the ICO data set come from areas where conditions that

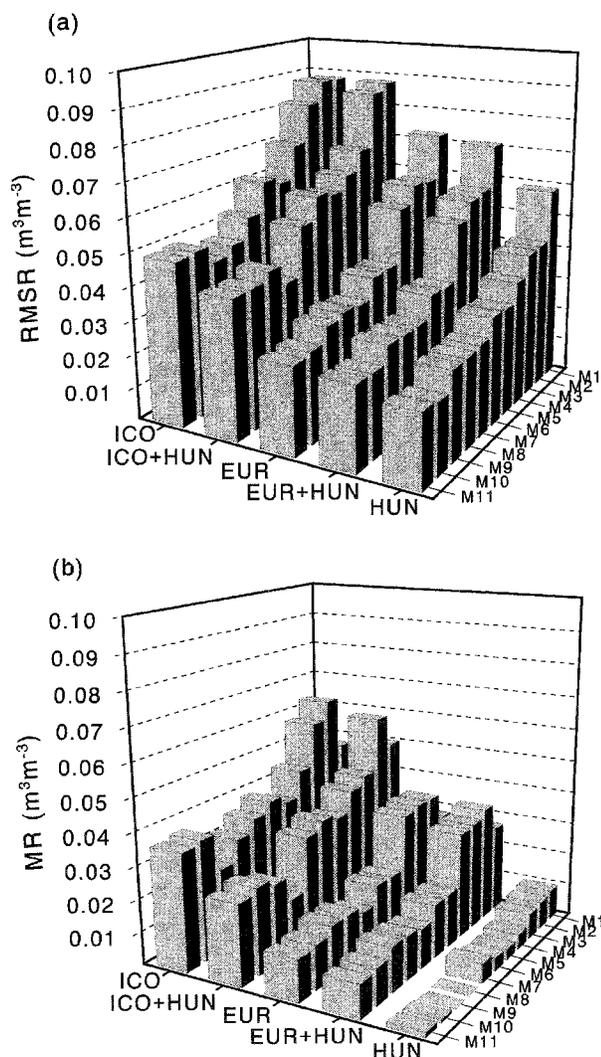


Fig. 2. Estimation errors in terms of (a) root mean squared residuals (RMSRs) and (b) mean residuals (MRs) of each of the eleven models and five data sets used to develop pedotransfer functions. HUN, Hungarian data set; EUR, European data set; ICO, intercontinental data set. Input parameters of models M1, M2 ... M11 are listed in Table 2.

govern soil development may be far from the Hungarian conditions. Because of geographical proximity, soil-forming conditions of other European countries may differ less from the Hungarian conditions. This may result in smaller differences in soil properties, as shown in Table 1, explaining the better estimations.

Mean residual values also show a trend of improvement as the list of input variables increases (Fig. 2b). Most bias is introduced by the PTFs developed using the ICO (and ICO + HUN) data sets (MR is between 0.016 and 0.063  $\text{m}^3 \text{m}^{-3}$ ). PTFs developed using the European scale data sets showed bias from 0.009 to 0.036  $\text{m}^3 \text{m}^{-3}$ , with slightly more accurate estimations when Hungarian data were included. Bias for the HUN data set always remained  $<0.01 \text{ m}^3 \text{m}^{-3}$ . It is interesting to see that bias always remained positive, reflecting an underestimation of water contents (cf. Eq. [3]). This may be caused by the considerable differences among

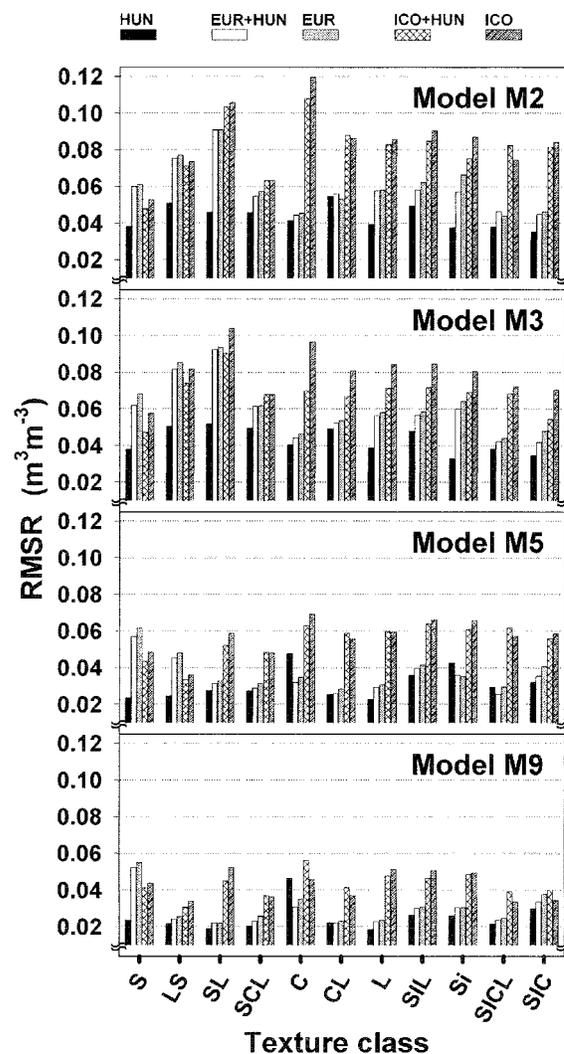


Fig. 3. Root mean squared residuals (RMSRs) for each USDA soil texture class using four selected models (M2, M3, M5, M9) and each of the data sets used to develop pedotransfer functions. HUN, Hungarian data set; EUR, European data set; ICO, intercontinental data set. Input parameters of the four models are listed in Table 2.

data sets in OM contents (see Table 1). In general, soils with greater OM contents tend to retain more water at the same matric potentials than soils with lesser OM contents, which may be a direct effect of greater OM contents or an indirect effect through the improvement in soil structure stability. For this reason, data sets with mostly lesser OM contents may estimate lesser water contents. This seems to be justified when we notice that improvements in MR caused by the inclusion of Hungarian data in the larger scale data sets are largest when OM was one of the input parameters (i.e., Models 3, 11). It is especially clear in the example of the ICO set with considerably lesser average OM contents than the HUN (and HUNTEST) data set (mean: 18.43 vs. 13.82  $\text{g kg}^{-1}$ ; median: 17.20 vs. 3.45  $\text{g kg}^{-1}$ ).

In Fig. 3, RMSR values are stratified by USDA soil texture classes. Results of four selected models (M2, M3, M5, M9) are shown as examples. A general trend of improvement can be seen with increasing input to

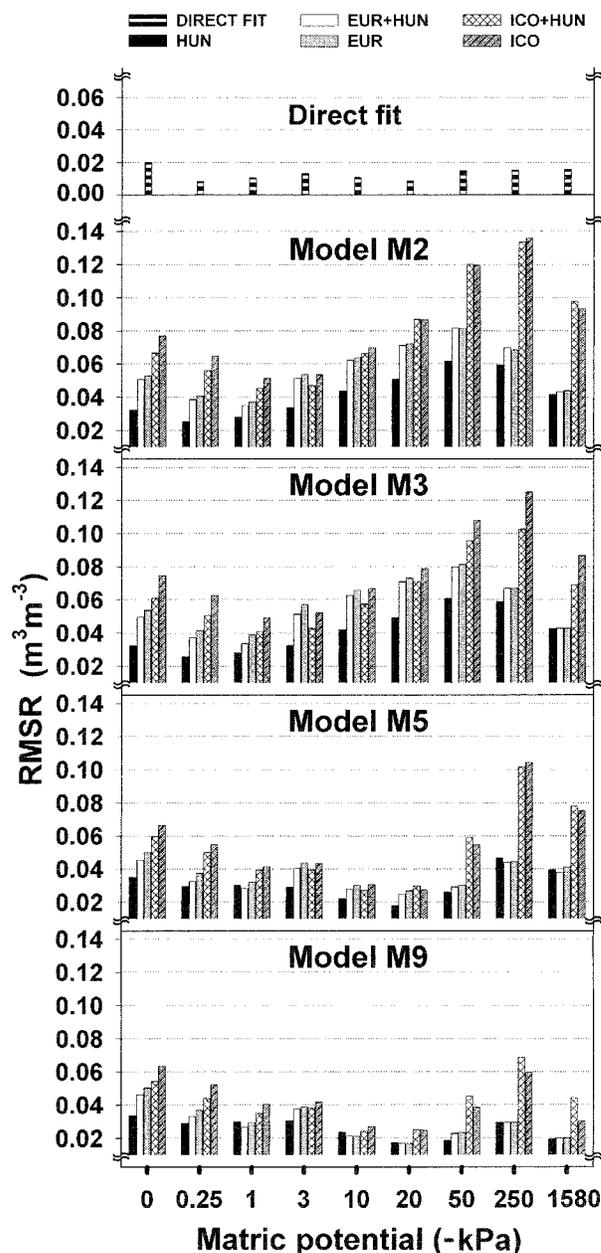
the models. For most texture classes, accuracy of PTF estimations was best using the HUN data set, followed by the EUR + HUN, EUR, ICO + HUN, and ICO data sets. Addition of Hungarian data to the two international sets resulted in marginal improvement only. Pedotransfer functions using the EUR and EUR + HUN data sets provided worst estimations at the coarse end, thus for sand, loamy sand, and sandy loam soils. The large difference between the average OM contents of the HUNTEST and EUR data sets for the sand and loamy sand classes (Table 4) are one possible reason for a systematic overestimation. In fact, inclusion of OM content (as in model M3) made estimations worse for the above texture classes and the sandy clay loam class, as compared with estimations by model M2. Estimations by PTFs developed from the ICO and ICO + HUN data sets with models M2 and M3 that do not use water retention points as input are considerably worse for sandy loam and clay soils than for other texture classes. The texture of Hungarian clay soils is heavier than that of the two international data sets, and their OM content is the greatest of all texture classes (Table 4); thus these soils are largely different from the clay soils of the ICO data set. Besides, the ICO data set is underrepresented in clay soils (Fig. 1). As for the sandy loams, the fact that PTFs from both international data sets make large errors may indicate that Hungarian soils of this texture are, in some way, different from those of the international data sets without the basic soil data explaining it. This is supported by the fact that for these classes estimations improve considerably using models M5 and M9 (thus models that use water retention data also) as compared with models M2 and M3. Measured water contents in the Hungarian data set for the sandy loam texture class are significantly ( $0.08\text{--}0.10\text{ m}^3\text{ m}^{-3}$ ) greater throughout the entire WRC than in the international data sets, without differences in the soil survey data explaining it.

In Fig. 4, RMSR values are shown for the same four models as in Fig. 3, stratified by matric potential values. Clearly, there is improvement in estimations throughout the entire WRC with using more input in the models. In general, a hierarchy of PTFs can be observed: using

**Table 4.** Average organic matter (OM) contents by USDA texture classes of the test data set (HUNTEST) and the Hungarian (HUN), European (EUR), and intercontinental (ICO) data sets.

Texture class	OM content			
	HUNTEST	HUN	EUR	ICO
	g kg <sup>-1</sup>			
Sand	1.81	3.05	14.70	11.11
Loamy sand	3.48	6.68	12.98	15.05
Sandy loam	15.86	14.06	14.99	14.23
Sandy clay loam	9.80	13.43	12.91	5.45
Sandy clay	—	—	9.40	1.94
Clay	25.34	32.26	21.44	7.32
Clay loam	22.60	17.00	22.98	5.20
Loam	19.84	23.00	17.58	17.97
Silt loam	17.97	20.85	11.40	20.75
Silt	21.10	23.98	3.66	3.72
Silty clay loam	14.91	21.34	18.37	5.08
Silty clay	23.05	20.76	22.49	12.61

the HUN data set as input being the best and PTFs using the ICO data set being the worst, with only some exceptions. A trend of special interest is the relatively large RMSR for water contents at saturation and at  $-0.25\text{kPa}$ . While accuracy of estimations at saturation using the HUN data set remains similar to the accuracy at other points in the wet range of the WRC, all PTFs developed on international data sets had larger RMSRs at saturation. Differences in OM content that were detailed above may explain part of the larger errors; however, different structural development and macroporosity of soils in the different data sets used for PTF



**Fig. 4.** Root mean squared residuals (RMSRs) for several matric potential values for the direct fitting of the van Genuchten equation and using four selected models (M2, M3, M5, M9) and each of the data sets used to develop pedotransfer functions. HUN, Hungarian data set; EUR, European data set; ICO, intercontinental data set. Input parameters of the four models are listed in Table 2.

development may also have influenced estimations. In Fig. 4, the RMSRs of the direct fitting of the van Genuchten model to the measured water retention data of the HUNTEST data set are also shown. The possible presence of macroporosity, which is not accounted for in the van Genuchten equation, is reflected by the higher RMSR of the direct fit at saturation. As the comparison of measured and estimated WRCs was made through the use of the van Genuchten model, calculated errors will also include an element of error that result from the nonperfect fit of the van Genuchten model to the measured water retention points.

Large RMSRs for the dryer range of the water retention mean an obvious failure of the ICO (and ICO + HUN) data sets. We found that the largest part of this error originated from soils with textures that were underrepresented in the ICO data set: clay, silt, silty clay, and silty clay loam (c.f. Fig. 1). We note that smaller RMSRs at some matric potentials may have resulted from the fact that some NN models included one, two, or three water retention values at matric potentials close to the estimated values. This is the reason why major improvement is seen from M3 to M5 at matric potentials of  $-10$ ,  $-20$  and  $-50$  kPa and from M5 to M9 at  $-1580$  kPa.

### Functional Evaluation on Water Content Time Series

One possible application of estimated WRCs is their use in numerical models for simulations of water content and solute transport dynamics. Three of the 11 PTFs (M2, M5, and M9) were further used in combination with all five data sets to simulate the soil moisture profiles of seven Hungarian soils. M2 was selected as a widely applicable very simple model that requires only texture and  $D_b$  as input. M5 and M9 were selected as the best models that use one (M5) or two (M9) addi-

tional water retention points as input. Water retention points required by the selected models are often measured as those are in common use for the calculation of water holding capacity. Simulated and measured water contents were then compared at each depth for each available date and for each profile. Root mean squared residuals and MR were calculated as defined earlier, with the replacement of measured and estimated WRC data with field-measured and simulated soil water contents respectively. As a control, allowing further comparisons, simulation was also run using laboratory-measured WRCs, and the same measures were calculated as above.

Table 5 summarizes the results. Averaged RMSR values ranged from 0.046 to 0.093  $\text{m}^3 \text{m}^{-3}$  considering the different data sets of PTF development, and from 0.048 to 0.090  $\text{m}^3 \text{m}^{-3}$  using the measured WRCs. While the focus should be on the averaged errors obtained using the particular data set or NN model, we would like to point out the case of two soils that show the worst results using any input data: GDL and NYL249. One possible cause of that for the GDL soil is the large heterogeneity of the soil (Cs. Farkas, 2002, personal communication), which makes sampling an intricate issue. The RMSR for NYL249 soil was probably greater because it is a soil type (Natrustoll) in which the salt content alters its physical properties. That factor was not accounted for in our PTFs and model runs, and is one of the potential pitfalls of most current PTFs. The average of all profiles is between 0.065 and 0.07  $\text{m}^3 \text{m}^{-3}$  for the five data sets, with the HUN set the best, and the ICO + HUN set the worst. Using measured data, the average RMSR is 0.061  $\text{m}^3 \text{m}^{-3}$ . Errors by the international PTFs were only a fraction larger (0–0.005  $\text{m}^3 \text{m}^{-3}$  on annual average) than errors by the national scale PTFs, and remarkably, the different scale PTFs were only 0.004 to 0.009  $\text{m}^3 \text{m}^{-3}$  worse on an annual basis than laboratory-

**Table 5.** Summary of the simulation results in terms of root mean squared residuals (RMSR) and mean residuals (MR).

Soil	Measured WRC†	Values averaged per data set				Values averaged per model			
		HUN‡	EUR§	EUR + HUN	ICO¶	ICO + HUN	M2	M5	M9
$\text{m}^3 \text{m}^{-3}$									
<b>RMSR</b>									
AVERAGE	0.061	0.065	0.065	0.066	0.067	0.070	0.067	0.066	0.066
GDL	0.067	0.090	0.093	0.088	0.084	0.093	0.084	0.090	0.094
DHR2	0.052	0.064	0.047	0.047	0.064	0.054	0.065	0.055	0.045
DHR3	0.053	0.057	0.084	0.083	0.063	0.065	0.064	0.072	0.076
DHR4	0.049	0.046	0.057	0.057	0.050	0.052	0.047	0.052	0.058
KM1	0.065	0.057	0.049	0.051	0.057	0.055	0.054	0.056	0.051
KMIK	0.053	0.066	0.057	0.061	0.058	0.077	0.067	0.063	0.061
NYL249	0.090	0.074	0.068	0.076	0.093	0.093	0.091	0.076	0.075
<b>MR</b>									
AVERAGE	-0.009	0.004	0.020	0.018	0.005	0.020	0.021	0.007	0.012
GDL	0.049	0.075	0.081	0.076	0.071	0.080	0.071	0.077	0.081
DHR2	0.000	-0.032	-0.003	-0.003	-0.038	-0.027	-0.035	-0.011	-0.016
DHR3	-0.013	0.005	0.032	0.030	0.016	0.022	0.022	0.019	0.022
DHR4	-0.016	0.003	0.023	0.021	0.005	0.014	0.012	0.009	0.019
KM1	-0.034	-0.029	-0.015	-0.017	-0.035	-0.016	-0.019	-0.029	-0.019
KMIK	0.010	0.021	0.025	0.027	0.022	0.041	0.033	0.024	0.025
NYL249	-0.061	-0.016	0.000	-0.006	-0.008	0.022	0.063	-0.040	-0.029

† WRC, water retention curve.

‡ HUN, Hungarian data set.

§ EUR, European data set.

¶ ICO, intercontinental data set.

measured WRCs, which may be a strong argument for PTFs. Inaccuracy of simulations even using original measured data may reflect inaccurate settings of some other input parameters (e.g., vegetation data), nonoptimal numerical solution(s) in the simulation model, or it may as well suggest that our conventional laboratory techniques produce errant WRC data compared with true field conditions (Pachepsky et al., 2001). If errors are grouped by models, M2 to M9, there is improvement, but there is only a difference of  $0.002 \text{ m}^3 \text{ m}^{-3}$  in water content between the best (M9:  $0.065 \text{ m}^3 \text{ m}^{-3}$ ) and the worst (M2:  $0.067 \text{ m}^3 \text{ m}^{-3}$ ).

The trend is similar for the MRs, with GDL and NYL249 showing the largest bias in the simulations with estimated as well as with measured WRCs. On average, PTFs using the HUN and the ICO data sets result in biases that are even smaller than the bias of the measured input data ( $0.009 \text{ m}^3 \text{ m}^{-3}$ ), but for all data sets, bias remains at or under  $0.02 \text{ m}^3 \text{ m}^{-3}$ . Regarding the grouping by M2, M5, and M9 models, M2 shows a somewhat larger bias ( $0.021 \text{ m}^3 \text{ m}^{-3}$ ) whereas the other two remain at or under  $0.012 \text{ m}^3 \text{ m}^{-3}$ . It is interesting to note that there are large differences among the biases of M2, M5, and M9 for the NYL249 Natrustoll. Model M2, which does not use measured water retention data as input, had a bias that was  $\approx 0.1 \text{ m}^3 \text{ m}^{-3}$  greater than M5 and M9. Differences between RMSR values were much smaller, which suggests that M2 was able to find the correct shape of the WRC but not the correct position of the air entry value.

The presentation of averaged values, as in Table 5 for example, may hide large differences between summer and winter periods, wet and dry periods, top- and subsoils, or among certain soil types. Such possible differences were examined. We found no significant difference between simulation errors for wet and dry or for winter and summer periods. There were differences among soils, but no correlation could be shown with texture or any other examined soil properties besides the influence of salinity in one soil. The only significant and systematic difference was the increasing underestimation (i.e., lesser MR) with depth. The change was irregular; however, for most soils MR was more negative by 0.04 to  $0.06 \text{ m}^3 \text{ m}^{-3}$  below 50 cm, than in the top 50 cm. This could be observed with all PTFs, but also when laboratory-measured WRC was used in the simulation model, which indicates that usage of PTFs were not the particular reason for such deviation.

Figure 5 shows an example of the simulations of water content time series for all seven depths of one soil (KM1). Five curves in each subplot show daily model outputs of soil water contents using estimated WRCs from each of the five data sets using the M5 model. An additional curve shows the daily outputs when laboratory-measured WRCs were used in the simulation model. Simulated water contents are marked with symbols at those days for which field-measured water contents were available. Day-to-day changes in the simulated moisture contents in the top layers are apparent. These changes gradually weaken with depth. Differences between simulations using WRCs obtained from

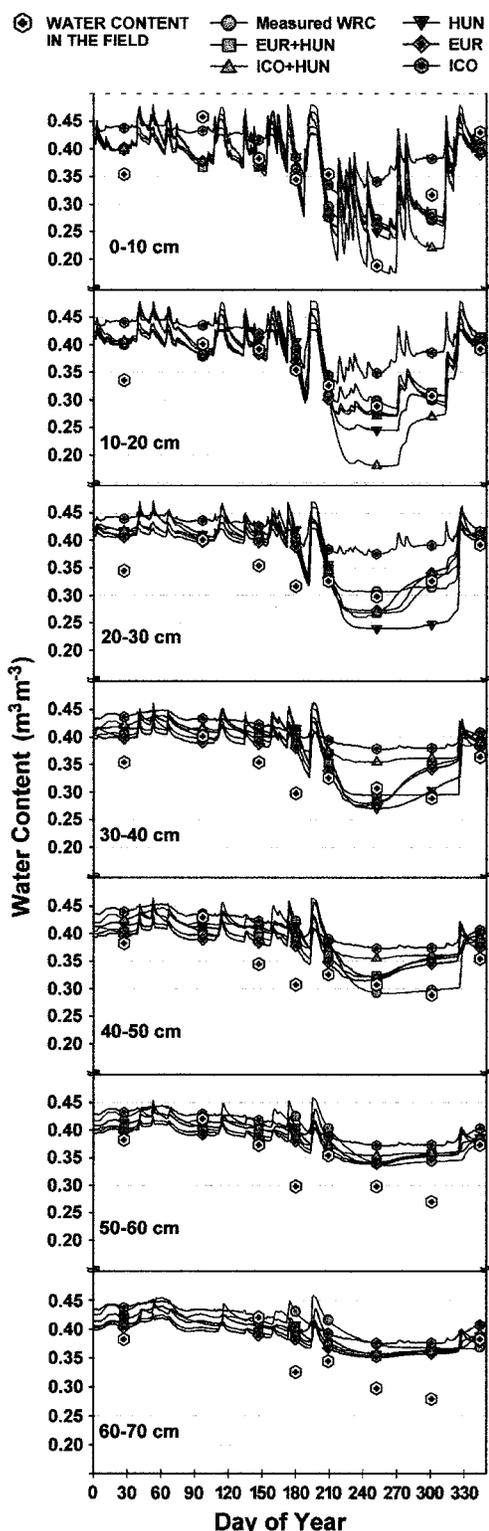
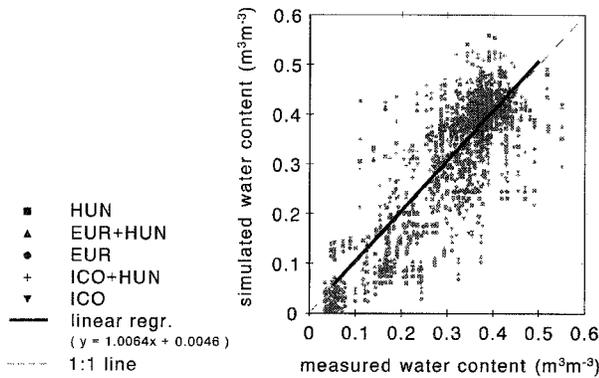


Fig. 5. Example of the simulations of water content time series for different depths of the KM1 soil using measured water retention curves (WRC) and as estimated with model M5 from the different data sets used to develop pedotransfer functions. HUN, Hungarian data set; EUR, European data set; ICO, intercontinental data set.

the different PTFs (and the laboratory measurement) are much larger in the top layers, specially in the dry period of the year. For this soil, at most depths, curves



**Fig. 6.** Measured vs. simulated water contents. All depths of all seven soils are shown using water retention curves estimated with model M5 from each data set of pedotransfer function development. HUN, Hungarian data set; EUR, European data set; ICO, intercontinental data set.

representing PTFs developed from the ICO and ICO + HUN data sets show largest deviations from the field-measured water contents. For the subsoil layers, simulations continuously overestimated water contents during drier periods, no matter which WRC was used. At this site, temperatures  $<0^{\circ}\text{C}$  often occur in the first 30 to 60 d of the year. Large overestimation of field water content at Day 27 in the top horizons probably occurred due to the nonoptimal solution applied in the simulation model for such conditions (J.G. Kroes, 2002, personal communication). In reality, infiltration and soil water flow is limited by temperatures below the freezing point. As also suggested in the previous paragraph, any of the presented characteristics of simulations were not necessarily the same for the other six soils.

Figure 6 summarizes all such measured/simulated data pairs using estimations by the M5 model, as an example, for all depths of all test soils and by all five data sets. Deviation of the regression line from the 1:1 line is very small. The slope and intercept parameters do not differ from 1 and 0, respectively, at a 95% confidence level. Table 6 lists parameters and coefficients of determination ( $R^2$ ) of linear regressions obtained as in Fig. 6, for each group between measured and simulated water contents. For any of the data sets used to develop PTFs, the regression equation introduces very small bias (offset). Bias found for the different development data sets was smaller than the bias using the measured WRCs. For all data sets (and the measured WRCs), the intercept parameter did not differ from 0 at a 95% confidence level. The slope parameter of the line remains around (but always below) one, meaning a slight underestimation in wet periods. This parameter differed from 1 significantly (95% confidence) for the EUR, ICO, and ICO + HUN data sets. When data are grouped by models, superiority of M5 and M9 is apparent over M2, which despite of its small bias shows the largest deviation from the unit slope (0.87). Both parameters of M2 and the intercept parameter of M9 differed from the optimal 0 (intercept) and 1 (slope) at the 95% significance level. Regarding the coefficients of determination, the simulations using measured WRC data show largest  $R^2$  (0.709), but simulations with estimated WRC data

**Table 6.** Parameters of linear regression and coefficients of determination describing the relationship between measured and simulated soil water contents, grouped by data sets and by models.

Data group	Regression coefficients				$R^2$
	Slope		Intercept		
	Mean	SE	Mean	SE	
Measured WRC	0.9807	0.0313	0.0186	0.0102	0.709
HUN	0.9937	0.0191	-0.0014	0.0063	0.691
EUR	0.9634	0.0184	-0.0063	0.0060	0.694
EUR + HUN	0.9671	0.0190	-0.0053	0.0062	0.683
ICO	0.9484	0.0196	0.0123	0.0064	0.659
ICO + HUN	0.9377	0.0208	-0.0024	0.0068	0.627
M2	0.8715	0.0150	0.0135	0.0049	0.626
M5	1.0064	0.0150	-0.0046	0.0049	0.690
M9	1.0082	0.0147	-0.0108	0.0048	0.701

provide in most cases only somewhat lesser values. Worst of the data sets in this respect is ICO + HUN with an  $R^2$  of 0.627.

In summary, differences in simulation accuracy, as a result of PTFs derived from different scale data sets, were much smaller than one would expect from differences obtained while evaluating WRCs only.

## CONCLUSIONS

Soil hydraulic PTFs to estimate water retention characteristics were developed using Hungarian and international data sets and were evaluated using a separate Hungarian data set. PTFs developed using a European data set provided somewhat larger errors than PTFs developed using only Hungarian data. Pedotransfer functions developed using an intercontinental data set (containing data mainly from the USA and Europe) provided much larger errors. Inclusion of data from Hungary in the international data sets resulted in only small improvement of the PTFs. It suggests that having a small set of relevant data, when available, is better than using a large but more general data set.

Surprisingly, the differences among national and international PTFs largely disappeared when the estimated WRCs were used for simulations of soil water contents. Our findings support the success and effectiveness of PTFs. It has to be decided whether the accuracy of simulations that produce RMSR of 0.065 to 0.07  $\text{m}^3 \text{m}^{-3}$  using PTF-based soil hydraulic properties is satisfactory for a particular application. However, using measured WRCs lead to RMSR  $< 0.01 \text{ m}^3 \text{m}^{-3}$  better than PTFs, which is an argument for PTFs. These indicate that for the presented case study, differences between estimations using different scale data sets (or measured WRCs) were not the main source of error in the simulation model, PTFs errors were overwhelmed by errors resulting from other factors. One should, however, still apply PTFs with care, keeping in mind limitations that were discussed: the fact that not all soils are equally represented in any of the databases and other potential pitfalls set by different structure, clay mineralogy, salinity, and other factors that largely influence soil water status and that are not accounted for in most PTFs.

### ACKNOWLEDGMENTS

The authors wish to thank NATO (CLG 975761) for financial support for travel. One of us (MGS) was supported by the SAHRA science and technology center under a grant from NSF (EAR-9876800). A. Nemes wishes to thank the Ministry of Agriculture, Nature Management and Fisheries for financial support which was provided through the International Agricultural Centre (IAC) of Wageningen, The Netherlands. During this work, A. Nemes was affiliated with the Research Institute for Soil Science and Agricultural Chemistry (RIS-SAC) of the Hungarian Academy of Sciences, Budapest, Hungary. We thank three anonymous reviewers for their constructive comments.

### REFERENCES

- Cornelis, W.M., J. Ronsyn, M. van Meirvenne, and R. Hartmann. 2001. Evaluation of pedotransfer functions for predicting the soil moisture retention curve. *Soil Sci. Soc. Am. J.* 65(3):638–648.
- Czinege, E. 2000. Extension of soil information and knowledge of site-specific nutrient management consulting. (in Hungarian.) PhD Thesis. Univ. of Gödöllő, Gödöllő, Hungary.
- Demuth, H., and M. Beale. 1992. *Neural Network Toolbox Manual*. MathWorks, Natick, MA.
- Efron, B., and R.J. Tibshirani. 1993. *An introduction to the bootstrap*. Monographs on statistics and applied probability. Chapman and Hall, New York.
- Espino, A., D. Mallants, M. Vanclooster, and J. Feyen. 1996. Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties. *Agric. Water Manage.* 29:235–253.
- Farkas, Cs., Cs. Gyuricza, and P. László. 1999. Studies on certain physical soil properties in long-term soil cultivation experiments on a brown forest soil in Gödöllő. (in Hungarian.) *Növénytermelés* 48(3):323–336.
- Hack-ten Broeke, M.J.D., and J.H.B.M. Hegmans. 1996. Use of soil physical characteristics in laboratory measurements or standard series for modelling unsaturated water flow. *Agric. Water Manage.* 29:201–213.
- Haykin, S. 1999. *Neural Networks: A comprehensive foundation*. 2nd ed. Prentice Hall, London.
- Hecht-Nielsen, R. 1990. *Neurocomputing*. Addison-Wesley, Reading, MA.
- Imam, B., S. Sorooshian, T. Mayr, M.G. Schaap, J.H.M. Wösten, and R.J. Scholes. 1999. Comparison of pedotransfer functions to compute water holding capacity using the van Genuchten model in inorganic soils—Report to IGBP-DIS Soil Data Tasks. IGBP-DIS Working Paper no. 22. IGBP-DIS, Toulouse, Cédex, France.
- Kern, J.S. 1995. Evaluation of soil water retention models based on basic soil physical properties. *Soil Sci. Soc. Am. J.* 59:1134–1141.
- Minasny, B., A.B. McBratney, and K.L. Bristow. 1999. Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma* 93:225–253.
- Mualem, Y. 1976. A new model predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12:513–522.
- Nelder, J.A., and R. Mead. 1965. A simplex method for function minimization. *Comput. J.* 7:308–313.
- Nemes, A. 2001. *Unsaturated Soil Hydraulic Database of Hungary: HUNSODA*. *Agrokem. Talajtan* 51(1–2):17–26.
- Nemes, A., M.G. Schaap, F.J. Leij, and J.H.M. Wösten. 2001. Description of the unsaturated soil hydraulic database UNSODA version 2.0. *J. Hydrol.* 251(3–4):151–162.
- Pachepsky, Ya.A., W.J. Rawls, and D. Gimenez. 2001. Comparison of soil water retention at field and laboratory scales. *Soil Sci. Soc. Am. J.* 65:460–462.
- 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., 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., and F.J. Leij. 1998. Database-related accuracy and uncertainty of pedotransfer functions. *Soil Sci.* 163(10):765–779.
- Schaap, M.G., F.J. Leij, and M.Th. van Genuchten. 2001. ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* 251(3–4):163–176.
- Soet, M., and J.N.M. Stricker. 2003. Functional behaviour of pedotransfer functions in soil water flow simulation. *Hydrol. Process.* (In press).
- 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.* 60:771–774.
- Tietje, O., and V. Hennings. 1996. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. *Geoderma* 69:71–84.
- Tietje, O., and M. Tapkenhinrichs. 1993. Evaluation of pedotransfer functions. *Soil Sci. Soc. Am. J.* 57:1088–1095.
- Tiktak, A., F. van den Berg, J.J.T.I. Boesten, M. Leistra, A.M.A. van der Linden, and D. van Kraalingen. 2000. Pesticide emission assessment at regional and local scales: User manual of FOCUS Pearl version 1.1.1. Rijksinstituut voor Volksgezondheid on Milieu Rep. 711401008. Alterra Rep. 28. RIVM, Bilthoven, The Netherlands.
- Tóth, T., and L. Kuti. 2002. Numerical simulation versus repeated field instrumental measurements: A case study of monitoring salinity status in a native sodic grassland with shallow groundwater. *Agrokem. Talajtan* 51:243–252.
- Tóth T., and Gy. Várallyay. 2001. Variability in the soil of a sample area according to salt accumulation factors. (in Hungarian.) *Agrokem. Talajtan* 50:19–34.
- USDA. 1951. *Soil survey manual*. USDA Handbook No. 18. Washington, DC.
- van Alphen, B.J., H.W.G. Bootink, and J. Bouma. 2001. Combining pedotransfer functions with physical measurements to improve the estimation of soil hydraulic properties. *Geoderma* 103:133–147.
- van Dam, J.C., J. Huygen, J.G. Wesseling, R.A. Feddes, P. Kabat, P.E.V. van Walsum, P. Groenendijk, and C.A. van Diepen. 1997. SWAP version 2.0. Theory, simulation of water flow, solute transport and plant growth in the soil-water-atmosphere-plant environment. Tech. Doc. 45. DLO Winand Staring Centre, Rep. 71, Dep. Water Resources, Agric. Univ., Wageningen, The Netherlands.
- 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.
- Wagner, B., V.R. Tarnawski, V. Hennings, U. Müller, G. Wessolek, and R. Plagge. 2001. Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma* 102:275–297.
- Wösten, J.H.M., P.A. Finke, and M.J.W. Jansen. 1995. Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics. *Geoderma* 66:227–237.
- Wösten, J.H.M., A. Lilly, A. Nemes, and C. Le Bas. 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90:169–185.
- 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(3–4):123–150.