EVALUATION OF THE CROPLAND MODELING COMPONENT OF THE U.S. NATIONAL SCALE CEAP PROJECT: ESTIMATION OF SOIL HYDRAULIC PROPERTIES

Attila Nemes[1,2], Dennis Timlin[2], Bruno Quebedeaux[1] and Vangimalla Reddy[2]

[1]Dept. Plant Science and Landscape Architecture, Univ. of Maryland, College Park, MD [2]USDA-ARS Crop Systems and Global Change Lab, Beltsville, MD

CONTACT

Attila Nemes
USDA-ARS
Crop Systems and Global Change Lab
10300 Baltimore Ave., Bldg 001, BARC-W
Beltsville, MD 20705
USA
Phone: +1 301 504-5177
Fax: +1 301 504-5823
Email: Attila.Nemes@ars.usda.gov

INTRODUCTION OF PROJECT, GENERAL OBJECTIVES OF EVALUATION

The USDA-NRCS is partnering with other US government agencies and Universities (ARS, NASS, FSA, Texas A&M Univ.) to conduct a national scale assessment of environmental benefits of conservation practices. The resulting Conservation Effects Assessment Project (CEAP) uses the National Resources Inventory as statistical framework and the Agricultural Policy Environmental Extender (APEX, Williams et al., 2000) simulation model to evaluate on-site benefits of conservation practices in cultivated croplands. Use of a simulation model for a range of crops and management systems requires robust parameterization and a large number of assumptions and abstractions, which require prior assessment. An independent evaluation of the cropland component of the national-scale assessment is being performed at the USDA-ARS Crop Systems and Global Change Laboratory and the University of Maryland. The goal of this collaboration is, to provide an in depth review of the modeling approach: the input databases, model output, and the model's processes, assumptions and abstractions.

INTRODUCTION, OBJECTIVES

Soil hydraulic properties comprise a significant part of the input to environmental simulation models. Such models are frequently used in projects that address public concerns regarding agricultural production and the environment, as is the case in the CEAP project.

Determination of soil hydraulic properties by field or laboratory measurements for large-scale studies is not feasible. This is especially true when different scenarios are considered, in which case obtaining measured data is impossible. In such cases, the use of pedotransfer functions (PTFs) offers means to obtain estimates of soil hydraulic properties.

Choice to use a particular PTF is usually driven by geographic validity and availability of input data. In this study we evaluate some aspects of the performance of the currently applied linear regression type pedotransfer function to generate soil hydraulic properties for US soils. We point out a limitation of linear regression based PTFs. We examine an alternative solution – a k-Nearest Neighbor non-parametric estimation technique - to obtain soil water retention information for US soils. Implementation of the latter technique in the CEAP simulation approach could be considered.

MATERIALS AND METHODS

VARIABLES OF INTEREST:

- Field capacity (FC – approximated by -33 kPa lab water retention measurement)
- Wilting point (WP – approximated by -1500 kPa lab water retention measurement)
- Available water holding capacity (AWC) = FC-WP

CURRENTLY APPLIED PTF IN CEAP: *Rawls et al. (1982)*

$$FC = 0.2576 - 0.002 * SAND + 0.0036 * CLAY + 0.0299 * OM$$

$$WP = 0.026 + 0.005 * CLAY + 0.0158 * OM$$

- Development data: 1323 soils (5350 horizons) from 32 states
- PTF considered to be valid for the contiguous US
- Linear regression PTF is widely used

Figure 1. Distribution of soils used by Rawls et al. (1982) to develop PTFs

AN ALTERNATIVE PTF SOLUTION: *Nemes et al. (2006, 2007)*

The *k-Nearest Neighbor software* (Nemes et al. 2008) uses a non-parametric technique (Nemes et al. 2006, 2007) to estimate -33 and -1500 kPa water retention. The technique includes searching a reference database for a small number of soils that are most similar to the target soil, based on the selected input attributes. The estimated value of the output attribute is then calculated as the weighed average of the output attribute of the selected nearest soils. Weighting is based on the degree of similarity (e.g. d1…d4 in Figure 2) of the selected soils to the target.

$$d_i = \sqrt{\sum_{j=1}^{x} \Delta a_{ij}^2}$$

d_i - 'distance' of the ith soil from the target soil;
Δa_{ij} - the difference of the i-th soil from the target soil in the j-th soil attribute.
x - the number of input attributes

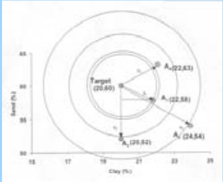


Figure 2. Schematic representation of the k-Nearest Neighbor technique to find the best match for the target soil, using two input attributes..

DATA USED TO TEST PTF PERFORMANCE: Soil Survey Staff, (1997).

Data selection:

- Sand/Silt/Clay content, Bulk Density, Organic Carbon Content, -33 kPa and -15000 kPa water retention.
- Checked for missing and inconsistent data
- Can be considered representative for the contiguous US
- Randomly split into two:
  - 30000 samples as reference data for k-Nearest Neighbor
  - 9632 samples to test both PTF methods

Table 1. Distribution of selected physical and hydraulic properties of 39632 soils of the NRCS NSSC National Characterization Data

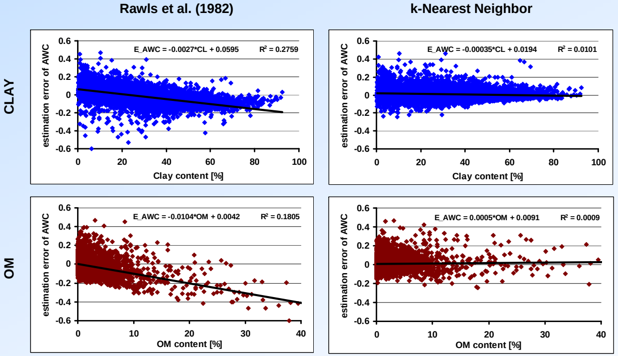
Properties	Unit	MIN	MAX	AVG	SD	MEDIAN
USDA Sand	%	0.000	98.600	34.869	25.071	31.400
USDA Silt	%	0.000	95.500	39.413	19.100	38.600
USDA Clay	%	0.200	94.900	25.718	15.972	23.500
Bulk Density	g/cm^3	0.100	2.370	1.431	0.244	1.440
Org. Mat.	%	0.000	93.096	1.542	2.862	0.707
θ(-33 kPa)	m^3/m^3	0.019	0.865	0.325	0.100	0.331
θ(-1500 kPa)	m^3/m^3	0.002	0.739	0.191	0.103	0.174

RESULTS

Traditional comparison of PTFs will not reveal all information that will become influential while the PTF is actually applied. (Table 2.)

Estimation errors of -33 and -1500 kPa water content - and the AWC derived from those values - were evaluated in terms of their correlation to input variables.

Using the linear regression PTF, -33 and -1500 kPa estimates can be biased to the opposite direction; which will propagate to an enhanced bias when the AWC value is calculated. (Table 3, Figure 3)



Figures 3. Correlations between input properties (CLAY content above; OM content below) and estimation errors while deriving an estimate of available water holding capacity (AWC) using the linear regression based PTF (left panels) and the k-Nearest Neighbor PTF (right panels)

Table 2. Root mean squared residuals and mean residuals of AWC estimations using the two PTFs

	Rawls et al. (1982)	k-Nearest Neighbor
RMSR	0.0847	0.0579
ME	-0.0131	0.0099

Table 3. Correlation coefficients between input properties and estimation errors while estimating water retention at -33 and -1500 kPa pressure and derived available water holding capacity (AWC)

	Rawls et al. (1982)			k-Nearest Neighbor		
	-33 kPa	-1500 kPa	AWC	-33 kPa	-1500 kPa	AWC
Sand	0.0539	0.0447	**0.2057**	>0.0001	0.0034	**0.0023**
Silt†	0.0043	0.0116	**0.0272**	0.0032	0.0030	**0.0004**
Clay	0.0850	0.0432	**0.2759**	0.0050	0.0008	**0.0101**
Org. Matter	0.5080	0.3787	**0.1805**	0.0208	0.0276	**0.0009**

†not a direct input to Rawls et al. (1982)

CONCLUSIONS, OUTLOOK

- A single linear regression based PTF is likely to introduce bias over the range of applicable soil properties in terms of available soil water content – shown on the example of Rawls et al. (1982)
- Such behaviour is very likely not unique for the above PTF but is expectable from other linear regression based PTFs, because of possible non-linearity in the underlying data relationships and because linear regression parametric PTFs are best-fit functions and tend to generalize (i.e. bias) towards the database means
- The non-parametric k-Nearest Neighbor estimation technique may provide a feasible alternative with significantly lower bias. Such reduced bias is obtained largely because the estimate is formulated from existing data of a few similar soils and not from a function that is generally applicable for the entire data domain
- Functional testing of the performance of the k-Nearest Neighbor technique for the purposes of CEAP is underway
- Potential other alternatives that are planned to be explored: a group of existing PTF equations used in parallel as an ensemble (e.g. Guber et al. 2006)

REFERENCES

Guber, A.K., Ya. A. Pachepsky, M.Th. van Genuchten, W.J. Rawls, J. Simunek, D. Jacques, T.J. Nicholson, and R.E. Cady. (2006). Field-Scale Water Flow Simulations Using Ensembles of Pedotransfer Functions for Soil Water Retention. Vadose Zone J. 5: 234-247.

Nemes, A., W.J. Rawls and Ya.A. Pachepsky. (2006). Use of a non-parametric nearest-neighbor technique to estimate soil water retention. Soil Sci. Soc. Am. J. 70(2): 327-336.

Nemes, A., W.J. Rawls, Ya.A. Pachepsky and M.Th. van Genuchten. (2007). Sensitivity Analysis for the Non-Parametric Nearest Neighbor Technique to Estimate Soil Water Retention. Vadose Zone J. 5: 1222-1235.

Nemes, A., R.T. Roberts, W.J. Rawls, Ya.A. Pachepsky and M.Th. van Genuchten. (2008). Software to estimate –33 and –1500 kPa soil water retention using the non-parametric k-Nearest Neighbor technique. Environmental Modelling and Software 23: 254-255.

Rawls, W.J., D.L. Brakenciek and K.E. Saxton. (1982). Estimation of soil water properties. Trans. ASAE 25(5) 1316-1320 and 1328.

Williams, J.R. et al., 2000. The APEX model. BRC Report NO. 00-06. Temp, TEXAS, Blackland Research Center.