Toward improving estimates of field soil water capacity from laboratory-measured soil properties

Attila Nemes[A], Yakov Pachepsky[B], Dennis Timlin[C]

[A]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD, USA, Email attila.nemes@ars.usda.gov
[B]Environmental Microbial and Food Safety Lab, USDA Agricultural Research Service, Beltsville, MD, USA, Email yakov.pachepsky@ars.usda.gov
[C]Crop Systems and Global Change Lab, USDA Agricultural Research Service, Beltsville, MD, USA, Email dennis.timlin@ars.usda.gov

Abstract

Different recommendations exist world-wide on which – if any - pressure head should be used in laboratory measurements to approximate the ‘field capacity’ (FC) of the soil. Literature often deems any such pressure heads to be inadequate to approximate FC for soils of all textures. We used a data collection from the literature to evaluate if corrections can be made to improve the estimation of FC from -33 kPa water retention (W33). Regression tree modeling coupled with jack-knife cross validation was used to identify the best predictors – sand, silt, clay and the measured W33 value – to estimate the difference between W33 and FC. Such predictions were then successfully used to adjust the W33 value as the estimate of FC. An improvement in estimating FC was seen in general statistical terms, and texture specific bias was also greatly reduced. Such a solution may allow the reliable use of a single pressure head in the laboratory to approximate FC, which may be the only feasible option for large scale studies.

Key Words

Field capacity, drained upper limit, pedotransfer function, regression tree, jack-knife, re-sampling

Introduction

Field capacity (FC) - i.e. the content of water remaining in a soil that has been wetted with water and after free drainage is negligible - is an important soil hydraulic parameter that has multiple uses in hydrological, meteorological, agronomical, and environmental predictions and modelling. Measurement of FC is unfeasible in large scale projects, therefore estimating FC is common practice. The customary way to estimate FC is to equate it to soil water contents measured in the laboratory at a predefined soil water pressure head. Different values of such pressure head have been employed in different countries, e.g. -5 kPa in United Kingdom (White, 2006) and France (Le Bas *et al*., 2007), -6 kPa in Brazil (Ajayi *et al*., 2009), -10 kPa in Australia (White, 2006) and Sweden (Kätterer *et al*., 2006) and even varied among different authors in the same country. In the United States the recommended value of such pressure head is -33 kPa (Kirkham, 2005).

There have been continual reports that the laboratory measured water content at -33 kPa can be a poor predictor of FC (e.g., Haise *et al*., 1955; Rivers and Shipp 1978). The objective of this work was to use a sizeable data collection (a) to evaluate the accuracy of using laboratory measured water content at -33 kPa as the predictor of FC, and (b) to develop corrections to laboratory measured water content at -33 kPa - via examining the prediction residuals - that may allow a more reliable estimation of FC.

Materials and Methods

We used the data collection described and used by Ratliff *et al.* (1983), Cassel *et al.* (1983) and Ritchie *et al.* (1987). Data were assembled in 15 U.S. states for 61 soil profiles representing 6 soil orders. For each soil profile, the in-situ drained upper limit (DUL) and lower limit (LL) were measured at various depths. Following the field capacity concept (Veihmeyer and Hendrickson, 1931), DUL can be interpreted as the equivalent of FC. Cassel *et al.* (1983) report a wide range of soil properties that have been measured in the National Soil Survey Laboratory (Lincoln, NE) for these locations using standard procedures (USDA-NRCS SCS, 1972).

We categorized the field collected information and laboratory measured properties into three groups. *Field observations* comprised data on depth, US taxonomy order, master soil horizon notation, land use type, drainage and permeability classes. *Simple laboratory based data* included 3 particle-size classes (sand, silt, clay), texture classes derived from those classes, organic carbon content, bulk density, coefficient of linear extensibility (COLE) and the ratio of clay content to water retention measured at -1500 kPa pressure

(CLratio). *Detailed laboratory based data* included additional details about the particle-size distribution (PSD) by describing PSD using 8 particle-size classes. Water retention measured at -33 kPa pressure was used as a separate individual variable. Total of 243 samples had all the abovementioned data. The distribution of samples by USDA texture classes are shown in Table 1.

Table 1. Number of samples and texture class-wise differences in laboratory measured water retention at -33 kPa (W33) and field measured drained upper limit (DUL). (s - sand; ls – loamy sand; sl – sandy loam; scl – sandy clay loam; sc – sandy clay; l – loam; cl – clay loam; sil – silt loam; si – silt; sicl – silty clay loam; sic – silty clay; c – clay).

	s	ls	sl	scl	sc	l	cl	sil	si	sicl	sic	c
n	1	8	13	11	0	39	40	63	4	39	19	6
mean(W33-DUL) [vol. %]	-3.96	-3.36	-1.31	-2.25	-	0.24	0.21	2.67	3.40	1.80	1.54	3.94
st.dev.(W33-DUL) [vol. %]	N/A	4.15	2.44	3.46	-	4.07	6.85	5.27	6.58	3.58	4.80	3.09

We used regression tree modeling to find ways of improving the estimation of DUL using the selected data. Regression tree modelling is an exploratory technique that uncovers structure in data by first partitioning data into two groups. Each group is then further subdivided into two subgroups, providing groups as homogeneous as possible at each of the levels (Clark and Pregibon, 1992). Regression trees can use both categorical and numerical variables as predictors and have been used in the estimation of soil properties by e.g. McKenzie and Jacquier (1997), Rawls and Pachepsky (2002) and Lilly *et al*. (2008).

The optimal use of a tree model requires a criterion to halt further partitioning of the data to avoid over-fitting. In preliminary runs, we used random re-sampling combined with a trial-and-error approach and root-mean-squared residuals (RMSR) as decision criterion to optimize tree pruning for the current task, i.e. to estimate DUL. The ratio of the development and test data set size had also been optimized simultaneously. As a result, in the subsequent sections of this study, tree development will be stopped when the tree reaches 10 terminal nodes and calculations that involve re-sampling are performed using a training data set (N=220) and an independent test data set (N=23) that are used in a recurrent "jack-knife" cross validation scheme (i.e. randomized subset selection without replacement). In order to facilitate the estimation of uncertainty of our subsequent findings, calculations are performed on one hundred alternative training/test data set pairs.

Results and Discussion

Figure 1a shows the 1:1 comparison of W33 and DUL values in the data set, and a simple linear regression equation that best describes the data. There is considerable scatter around the 1:1 line as well as some bias; W33 tends to underestimate DUL where DUL is small and overestimate it where DUL is large. Typically coarse textured soils retain less water at a given suction than finer textured ones; hence the lower values in Figure 1a are those mostly of sands and other sandy soils. Table 1 also reflects such underestimation of DUL by W33 for the coarse textured samples. This observation agrees with the general recommendation reflected in literature; i.e. that in order to approximate DUL for coarse textured soils, a higher pressure should be used in laboratory measurements (e.g. Cassel and Nielsen, 1986 and therein).

If the root mean squared residual (RMSR) and mean residual (MR) is calculated directly from the scatter data in Figure 1a, an RMSR of 5.18 vol. % and an MR of -1.03 vol. % is found. To test various alternatives to improve the DUL estimates from W33, the first choice was to use the simple linear equation in Figure 1a as the correction factor. When DUL is calculated as 0.7433*W33+6.7746, the obtained RMSR is 4.715 and the overall bias is removed (MR=0), which are both improvements. One other solution to account for variation by texture is to correct the W33-based estimate of DUL according to the mean difference between W33 and DUL in each texture class, as shown in Table 1. When that is done, we obtain an RMSR of 4.58 and an MR of 0.04 (vol. %). To test such correction on independent data, we also generated the RMSR and MR using the re-sampling and cross-validation scheme outlined above. In each of the 100 alternative runs, 220 samples were analyzed for statistical differences between W33 and DUL and the texture class-based correction was then applied to the 23 independent test samples. The mean RMSR and MR for the independent test data set were 4.83 and 0.07 (vol. %) respectively.

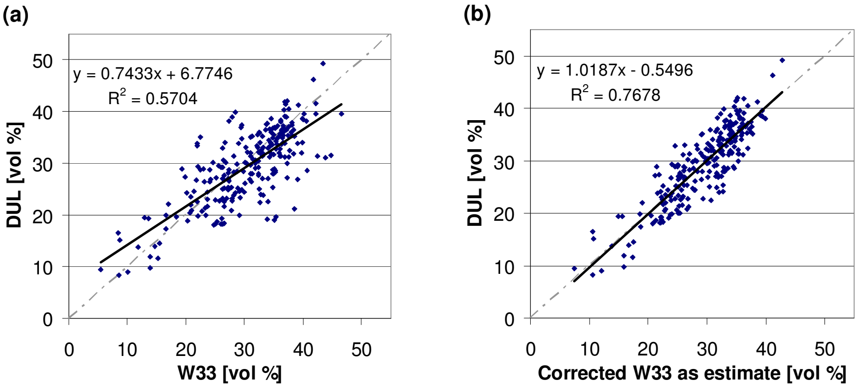

Figure 1. (a) Laboratory measured water retention at -33 kPa (W33) vs. field measured drained upper limit (DUL) for the 243 samples. Solid line and equation represent the best fit linear equation. (b) Estimate formulated after correcting W33 by the best regression tree description of ε, where ε=DUL-W33.

Table 2. Estimation RMSR and ME (and their standard errors) in estimating ε - formulated as ε=DUL-W33 (vol. %) - using various input groups. (W33 - laboratory water retention at -33 kPa; DUL - field measured drained upper limit).

Lab	Lab	Field	W33	TEST DATA				TRAINING DATA			
(simple)	(detailed)			RMSR	ST. ERR	ME	ST. ERR	RMSR	ST. ERR	ME	ST. ERR
x				4.900	0.087	-0.025	0.118	3.989	0.012	0	0
x		x		5.013	0.083	-0.091	0.119	3.984	0.012	0	0
x			x	4.412	0.064	-0.010	0.102	3.652	0.008	0	0
x		x	x	4.465	0.063	-0.018	0.108	3.647	0.009	0	0
x	x			5.092	0.086	-0.213	0.120	3.885	0.013	0	0
x	x	x		5.223	0.086	-0.284	0.125	3.857	0.013	0	0
x	x		x	4.808	0.074	-0.146	0.115	3.765	0.016	0	0
x	x	x	x	4.891	0.076	-0.167	0.121	3.757	0.015	0	0
			x	4.855	0.085	-0.067	0.114	4.404	0.011	0	0
		x		5.167	0.097	-0.039	0.121	4.513	0.014	0	0
		x	x	4.548	0.079	0.022	0.103	3.929	0.009	0	0
			x	4.855	0.085	-0.067	0.114	4.404	0.011	0	0
x(†)			x	4.367	0.063	-0.015	0.102	3.655	0.008	0	0
x(‡)			x	4.377	0.066	-0.028	0.101	3.693	0.009	0	0

†: sand, silt, clay, org. carbon, clay/-1500kPa water retention ratio used only ‡: sand, silt, clay used only

We then calculated ε for each sample as ε=DUL-W33. The value of ε is the correction needed to adjust W33 as the estimate of DUL. We evaluated the use of a hierarchically decreasing amount and variety of input variables to estimate ε (Table 2). Of the initial grouping of input data, using the *simple laboratory* data group and W33 itself gave one of the most accurate (training data) – and the most reliable (test data) – results. Use of more input data clearly had no advantage and signaled 'over-fitting' even for the training data. We then reduced the amount of data to sand, silt, clay content and W33, the error-estimates did not get worse. The three-class particle-size distribution appears to control much of the explainable variability in W33-DUL, while information related to e.g. soil depth (depth, horizons), taxonomic grouping or drainage/permeability classification could not explain any additional variability. When W33 is adjusted by using the texture+W33 model to estimate the W33-DUL difference, using re-sampling and the cross-validation scheme, a mean RMSR of 4.25 and MR of 0.01 (vol. %) is obtained, which is a substantial improvement from the direct estimation of DUL from W33. Moreover, the general texture related bias could be virtually eliminated, as seen in Figure 1b.

Conclusions

The use of the water content at -33 kPa as a practical approximate to DUL (and FC) is driven by early recommendations in literature and also by national preference and data availability. Inadequacy of that value to represent DUL for some soil texture groups has been noted in the past – and the use of other water retention point(s) have been recommended. In this study we examined the general suitability of W33 as an estimate to DUL and explored some possibilities to improve such estimate. It appears that a regression tree based grouping of the initial W33-DUL difference by soil particle-size fractions (sand, silt, clay) and W33 can result in a correction that generally improves the initial estimate, while also removes most of the texture based bias noted already in early literature. The presented methodology is also usable to test water contents at other pressure heads or to be tested on data from other parts of the World. More research is planned about the comparison of this technique with an improved direct estimation of DUL, testing of additional potentially influential environmental variables, as well as the extension of the approach to LL and available water content. An international effort seems to be desirable to improve and standardize the estimation of the field capacity value that is widely used in evaluations of the magnitude and consequences of global change.

References

Ajayi AE, de Souza Dias M Jr, Curi N, Araujo CF Jr, Aladenola O, Teixeira Souza T, Vasconcellos Inda A Jr. (2009) Aplicação de diferentes métodos para estimar a resistência de cinco solos. (Comparison of estimation methods of soil strength in five soils). *Revista Brasileira de Ciência do Solo* **33**, 487-495.

Cassel DK, Ratliff LF, Ritchie JT (1983) Models for estimating in-situ potential extractable water using soil physical and chemical properties. *Soil Science Society of America Journal* **47**, 764-769.

Cassel DK, Nielsen DR (1986) Field capacity and available water capacity. In 'Methods of Soil Analysis, Part I.: Physical and Mineralogical Methods', Agronomy Monographs No. 9. (Ed. Klute A), pp. 901-926. (Agronomy Society of America and Soil Science Society of America, Madison, Wisconsin).

Clark LA, Pregibon D (1992) Tree-based models. In 'Statistical Models in S' (Ed. Hastie TJ), pp. 377-419. (Wadsworth, Pacific Grove, CA).

Haise HR, Haas HJ, Jensen LR (1955) Soil moisture of some Great Plains soils. II. Field capacity as related to 1/3-atmosphere precentage and "minimum point" as related to 15- and 26-atmosphere percentages. *Soil Science Society of America Proceedings* **19**, 20-25.

Kätterer T, Andrén O, Jansson P-E (2006) Pedotransfer functions for estimating plant available water and bulk density in Swedish agricultural soils. *Acta Agriculturae Scandinavica* B **56**, 263-276.

Kirkham MB (2005) Principles of Soil and Plant Water Relations. Elsevier.

Le Bas C, King D, Daroussin J (1997) A Tool for Estimating Soil Water Available for Plants Using the 1:1,000,000 Scale Soil Geographical Data Base of Europe. *ITC Journal*, 3–4.

Lilly A, Nemes A, Rawls WJ, Pachepsky YaA (2007). Probabilistic approach to the identification of input variables to estimate hydraulic conductivity. *Soil Science Society of America Journal* **72**: 16-24.

McKenzie NJ, Jacquier DW (1997). Improving the field estimation of saturated hydraulic conductivity in soil survey. *Australian Journal of Soil Research* **35**: 803–825.

Ratliff LF, Ritchie JT, Cassel DK (1983) Field-measured limits of soil water availability and related laboratory-measured properties. *Soil Science Society of America Journal* **47**, 770-775.

Ritchie JT, Ratliff LF, Cassel DK (1987) Soil laboratory data, field descriptions and field measured soil water limits for some soils of the United States. ARS Technical Bulletin.

Rivers ED, Shipp RF (1978) Soil water retention as related to particle size in selected sands and loamy sands. *Soil Science* **126**, 94-100.

Rawls WJ, Pachepsky YaA (2002) Soil Consistence and Structure as Predictors of Water Retention. *Soil Science Society of America Journal* **66**, 1115-1126.

USDA-NRCS Soil Conservation Service (1972) Soil Survey Laboratory Methods and Procedures for Collecting Soil Samples. Report No. 1. U.S. Government Printing Office, Washington D.C.

Veihmeyer FJ, Hendrickson AH (1931) The moisture equivalent as a measure of the field capacity of soils. *Soil Science* **32**, 181-194.

White RE (2006) Principles and practice of soil science: the soil as a natural resource. 4th Ed. Wiley-Blackwell.