Taylor & Francis
Taylor & Francis Group

Arid Land Research and Management

ISSN: 1532-4982 (Print) 1532-4990 (Online) Journal homepage: http://www.tandfonline.com/loi/uasr20

Simple estimation of available water capacity in soils of semiarid and subhumid environments

Josefina Luisa De Paepe, Alfredo Angel Bono & Roberto Alvarez

To cite this article: Josefina Luisa De Paepe, Alfredo Angel Bono & Roberto Alvarez (2017): Simple estimation of available water capacity in soils of semiarid and subhumid environments, Arid Land Research and Management

To link to this article: https://doi.org/10.1080/15324982.2017.1408153

Published online: 12 Dec 2017.

ARID LAND RESEARCH AND MANAGEMENT
https://doi.org/10.1080/15324982.2017.1408153

Taylor & Francis
Taylor & Francis Group

Simple estimation of available water capacity in soils of semiarid and subhumid environments

Josefina Luisa De Paepe[a], Alfredo Angel Bono[b], and Roberto Alvarez[a]

[a]Facultad de Agronomia, Universidad de Buenos Aires, CONICET, Buenos Aires, Argentina; [b]EEA INTA Anguil, Anguil – La Pampa, Argentina

ABSTRACT

Soil available water capacity (AWC) is an important factor affecting soil productivity in semiarid and subhumid environments and is mainly determined by the soil textural composition. As the soils of these environments usually present fairly uniform textures across depth, we hypothesized that it would be possible to accurately estimate the whole-profile AWC using surface information. Our objective was to test this hypothesis in the Argentine Semiarid Pampas. Information was collected from 152 sites where AWC was measured in 20 cm layers up to a depth of 140 cm or up to the upper limit of the petrocalcic horizon, when present. In each case, whole profile AWC was estimated using a one-step and a two-step approach, comparing multiple regression and artificial neural networks as modeling techniques. Both modeling methods were effective ($R^2 > 0.76$, $P < 0.05$), however the former was chosen as no special software is required to run it, thus favoring simplicity. Models showed a strong interaction between surface AWC and soil depth and a simple nomogram was developed to estimate whole-profile AWC. Sampling and laboratory efforts should be significantly reduced using the model proposed in this paper.

ARTICLE HISTORY
Received 7 December 2016
Accepted 18 November 2017

KEYWORDS
Semiarid soils; water storage capacity; water storage estimation

Introduction

In semiarid and subhumid regions, where rain-fed agriculture is conducted, moisture limits crop dry matter production during part of the growing season (FAO 1991). Soils are coarse textured (Noy-Meir 1973), the B-horizon is almost absent (USDA 2014), and carbonates can be present depending upon root penetration depth and amount of rainfall received (FAO 1975). Under these conditions, crop water supply is strongly related to soil available water capacity (AWC) (Wong and Asseng 2006). Whole-profile AWC information is relevant, as it enables farmers to make knowledgeable management decisions regarding crop selection and irrigation that would help maximize profitability (Hanson, Rojas, and Shaffer 1999). AWC also allows forecasting yield (Keating et al. 2003) and is a required input for simulation models (Williams, Jones, and Dyke 1990).

AWC determination is laborious, time consuming, and requires access to special equipment (Klute 1986). Due to these limitations, many pedotransfer functions were developed to establish relationships between routinely measured soil properties and hydraulic properties at specific matric potentials (Rawls, Brakensiek, and Saxton 1982;

CONTACT Josefina Luisa De Paepe depaepe@agro.uba.ar Universidad de Buenos Aires Facultad de Agronomia, CONICET, Buenos Aires 1417, Argentina.

Bouma 1989). Many of these functions are frequently used without any local validation (Quiring and Legats 2008; Havrylenko et al. 2016). If these functions are not effective locally, it is necessary to measure the AWC or develop new estimation methods.

Soil textural composition is a strong controlling factor of AWC (Schaap, Leij, and van Genuchten 1998). Usually, in sandy soils of semiarid areas, minor textural variations related to depth are found. This is the case in the Argentine Semiarid Pampas, where extensive sampling performed for National Soil Surveys (INTA 1980; INTA 1989; INTA 2003) and a recent study (Romano, Alvarez, and Bono 2016) showed a rather uniform soil textural composition across depth. Therefore, we hypothesized that surface soil layer AWC might be an effective predictor of whole-profile AWC. Although previous attempts have been successful at estimating whole profile soil moisture content using surface information (Arya, Richter, and Paris 1983; Bono and Alvarez 2012), this procedure has not yet been tested for AWC estimates.

The Argentine Pampas constitutes one of the most important grain producing regions of the world (Satorre and Slafer 1999). Approximately 50% of this area (30 Mha) is semiarid–subhumid, with coarse textured soils (Hall et al. 1992). A worldwide analysis showed that crop root distribution is concentrated in the first 150 cm of the soil profile (Jackson et al. 1996). Consequently, most of the crop water consumption usually comes from the upper 140 cm of the soil profile, especially in sandy semiarid environments (Cutforth et al. 2013). This has also been determined for some extensive crops in the Semiarid Pampas (Fagioli 1973, 1983). Although the point of zero water extraction is approximately at a depth of 2 m (Dardanelli et al. 1997), in this region good predictions of yield have been attained when measuring water retention properties in the 0–140 cm soil layer (Bono, De Paepe and Alvarez 2011). Currently, there are no locally suitable methods for whole profile AWC estimates, and pedotransfer functions taken from the literature are currently being used without previous validation (Alvarez 2009). The objective of this research was to develop a simple statistical approach to determine whole-profile AWC measuring only surface AWC.

Materials and methods

Study area

The Argentine Pampas, is a vast 50 Mha plain located between 28°S and 40°S and 57°W and 68°W (Figure 1). Its natural vegetation is grassland. Mean annual temperature ranges from 14°C in the South to 23°C in the North and mean annual rainfall from 500 mm in the West to 1200 mm in the East. Soils were formed on Quarternary wind-transported loessoid deposits (Teruggi 1957). Superficial water may have played some role in the redistribution of this originally eolian sediment but the general phenomenon was a deposition of sandy materials in the West and finer particles in the East (Teruggi 1957). This sedimentary process determined the soil textural range from sandy in the West to clayey in the East. Clay neoformation is not significant in the region (Alvarez and Lavado 1998) and illite is the most common clay mineral (Alvarez and Lavado 1998; Berhongaray et al. 2013). Mollisols are the predominant soils formed on these loess-like materials (Alvarez and Lavado 1998). In many places along the West and the South of the region a petrocalcic horizon appears within the soil profile (Teruggi 1957). In areas with annual rainfall greater than 500 mm, rain-fed crops are mostly cultivated (Hall et al. 1992). Around 60% of the

Pampean area is under cropping with soybean (*Glycine max* (L.) Merr.), wheat (*Triticum aestivum* L.), corn (*Zea mays* L.), and sunflower (*Helianthus annus* L.) as main crops (MinAGRO 2017).

Sampling and determinations

Soil samples were collected from 152 sites during the 2000–2006 period in the Pampean subregion called the Semiarid Pampas (Figure 1a). Annual rainfall ranged from 500 to 800 mm along the sampled sites, thus comprising semiarid to subhumid climate conditions. Soil moisture regimes, in this subregion were classified as aridic, ustic or udic depending upon location (SAGyP-INTA 1990; Liu et al. 2012) following the West-East climatic gradient. These sites were broadly distributed over an area of approximately 18 Mha which allowed for regional coverage and were selected because they are subject to commonly used agricultural practices. The sampled soils were classified as Entic Haplustoll (65%), Typic Hapludoll (18%), Entic Hapludoll (11%), and others (6%) according to Soil Taxonomy (USDA 2014). Soil texture showed little variation as depth increased, as described before (INTA 1980, 1989, 2003). Six deep soil Hapludoll and Haplustoll profiles from sampled sites were randomly chosen from the National Soil Surveys (INTA 1980, 1989, 2003) to illustrate their sand percentage across depth (Figure 1b and c). Although both soil types present a broad range of sand content, varying from 25% to almost 80%, a consequence of the loess deposition gradient that follows the wind direction, this percentage remains quite constant at different depths. Hapludolls appear to be deeper than Haplustolls and on average, the latter seem to have more sand in their profile.

At each experimental site a pit was dug for soil profile description and bulk density determination via the cylinder method (Blake and Hartge 1986). Four samples per site were

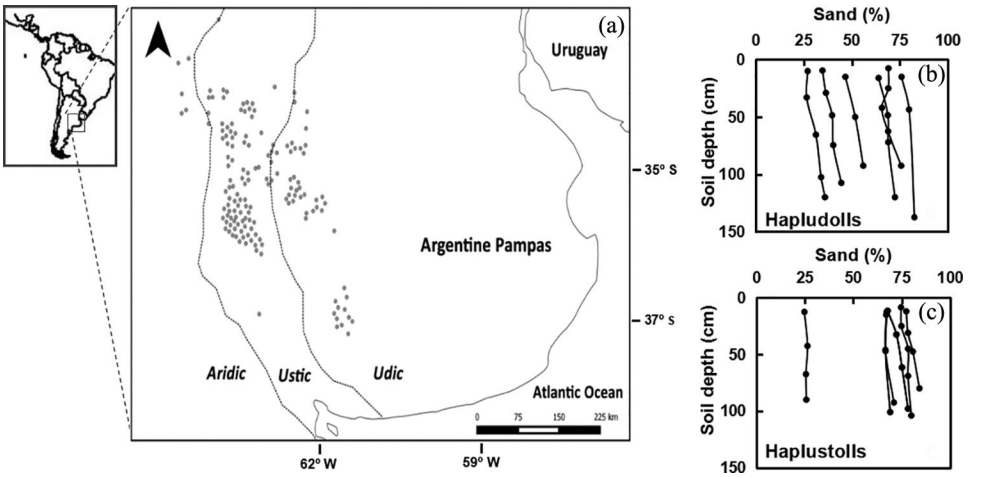

Figure 1. Location of the Pampas in the center of Argentina showing Udic, Ustic and Aridic soil moisture regimes according to National Soil Surveys (SAGyP-INTA, 1990). Gray dots indicate the 152 sampled sites. (b) and (c) show sand percentages with depth of Hapludoll and Haplustoll soil profiles chosen at random from a previous extensive soil sampling performed for National Soil Surveys. This information is published in the National Soil Surveys of the provinces that correspond to the location of the sampled sites (INTA 1980, 1989, 2003).

taken from the pit walls every 20 cm down to a depth of 140 cm. When soils presented a petrocalcic horizon, samples were taken to its upper limit. Samples were manually homogenized and air dried in the laboratory. Texture (Gee and Bauder 1996) and organic matter (Nelson and Sommers 1996) were determined in the soil surface layer (0–20 cm). Hydraulic properties were determined at each site with the suction cell method as described by Klute (1986). For determination of hydraulic properties, soil samples were repacked in 5 cm in diameter × 1 cm high cylinders. Approximately 25 g of soil were packed in the cylinders to attain a bulk density of 1.25 g cm^{-3}. This bulk density value was close to the mean bulk density of 1.29 g cm^{-3} for soils of this area (Alvarez et al. 2014). Cylinders were placed on a porous ceramic plate and wetted by capillarity until a sheet of water was observed on the sample surface. The soil was weighted approximately 48 h after that natural soil saturation was reached, when water dripping stopped. The plate was placed in a pressure pot (Soil Moisture equipment corporation, San Barbara California – USA) that applied a pressure of 33 kPa to determine field capacity and a pressure of 1500 kPa for wilting point determination. Eight artificially packed replicates per site and depth layer were measured using different samples for the determination of field capacity (FC) and wilting point (WP). When the coefficient of variation of these replicates was larger than 10%, the hydraulic properties were measured again. Finally, samples were oven-dried at 105°C during approximately 48 h and weighed. Weight difference between water stabilized and dried samples determined the soil water content at both mentioned matric potentials.

AWC was calculated as the difference between FC and WP, and gravimetric water capacity was transformed into volumetric water capacity (mm) using bulk density. Surface soil layer AWC ($AWC_{0–20\ cm}$) was calculated as the difference between FC and WP of the upper 20 cm soil layer. The sum of AWC of all soil layers to the sampling depth of 140 cm or to the upper limit of the petrocalcic horizon represented the whole-profile AWC ($AWC_{0–140\ cm}$). Whole-profile AWC excluding the surface layer was also calculated ($AWC_{20–140\ cm}$).

One-step and two-step soil $AWC_{0–140\ cm}$ estimation

Two modeling methods were contrasted: multiple regression vs. artificial neural networks. For both statistical methods, a one-step and a two-step approach were tested. In the one-step approach $AWC_{0–140\ cm}$ was estimated using $AWC_{0–20\ cm}$, surface texture and organic carbon contents and sampling depth as predictors. In the two-step approach, $AWC_{20–140\ cm}$ was estimated in the first step using all previously mentioned predictors and in the second step $AWC_{0–140\ cm}$ was calculated as the sum of measured $AWC_{0–20\ cm}$ and estimated $AWC_{20–140\ cm}$.

Polynomial regressions that included linear, quadratic, and interaction effects were used (Shen et al. 2003). Multicollinearity among variables was checked based on the variance inflation factor (VIF) value (Neter, Wasserman and Kutnet 1990). Variables were considered as uncorrelated when the VIF value was 1, moderately correlated when the VIF value ranged from 1 to 5 and highly correlated when it was larger than 5. Stepwise regression adjustments were tested to select the simplest model with the highest R^2 (Schaap and Bouten 1996). The significance of the model was determined at the $P < 0.05$ level through the F test, and terms were included in the regression only when they were statistically significant at the $P < 0.05$ level. A visual inspection of the relationship of

residuals vs. fitted values was conducted to verify homogeneity (Zuur et al. 2009). Neural networks are simpler than process based models and have become a common method for determining complex input–output relationships (Silveira et al. 2013). A common neural network structure that includes input, hidden, and output layers was used. Feed-forward networks were adjusted, using the back-propagation algorithm for weight fitting (Kaul, Hill, and Walthall 2005). Linear transfer functions connected the input layer to the hidden neuron layer and the output neuron layer to the network output while a sigmoidal transfer function connected the hidden layer to the output layer (Lee et al. 2003). The procedure described by Alvarez (2009) was followed for network architecture simplification, scaling methods, learning rates, and epoch size. The weight of the independent variables on the dependent variable was assessed based on the sensitivity ratio. Only inputs with a ratio larger than 1 were selected (Miao, Mulla, and Robert 2006). Neural networks were fitted using Statistica Neural Networks (version 2011 StaSoft).

Regression models were fitted using 75% of randomly selected data (training set), and validated against the independent remaining 25% (validation set). Artificial neural network models were fitted using 50% of the data (training set), 25% of the data were used for early stopping of weight fitting (test set), and finally the remaining independent 25% was used to validate the model (validation set) (Park and Vlek 2002). Cross-validation was used to avoid overlearning (Özesmi, Tan, and Özesmi 2006). The generalization capacity of the fitted polynomial regressions and network models was assessed by comparing the R^2 of training and validation datasets, which were statistically contrasted (Kleinbaum and Kuper 1979). Ordinates and slopes of regressions between observed vs. estimated data were also statistically tested using Integrated Resources for Evaluating Numerical Estimates (IRENE) (Fila et al. 2003). A range of statistics for model evaluation was used (Willmott 1981), as it is better to perform model comparisons based on a combination of statistics (Lin et al. 2002). To quantify the deviation of the regressions and neural networks estimated results from the observed data, the root mean squared error (RMSE) (Kobayashi and Salam 2000) (Eq. 1), the mean difference (MD) (Kobayashi and Salam 2000) (Eq. 2) and the modeling efficiency (ME) (Tedeschi 2006) (Eq. 3) were calculated:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \tag{1}$$

$$\mathrm{MD} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) \tag{2}$$

$$\mathrm{ME} = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x})^2 - \sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(x_i - \hat{x})} \tag{3}$$

where n is the number of measurements and i is the ith observation, x_i and y_i the measured and estimated, by the regression and neural network models, values respectively. $\hat{x}$ the mean of the measured values. When estimated values are closer to measured values, RMSE and MD, which had the same units as the studied variable, decrease. Conversely, better estimates are achieved when, on a relative scale, ME increases from 0 to 1.

A sensitivity analysis was performed to estimate the relative impact of the independent variables on $AWC_{0-140\ cm}$ discarding the 10 and 90 percentiles to avoid extreme values as described by Alvarez and Grigera (2005). Results were represented graphically by a regression surface plot and by a nomogram. Simple regression and correlation analysis were used for inspecting the relationships between variables, assessing their significance by the F test ($P < 0.05$).

Results

Measured soil variables

The sampled sites covered a broad soil variation range along the Semiarid Pampas (Table 1). In the first 20 cm soil layer, about 40% of the sites had more than 2% organic matter content. Most soils presented limited clay content as only 10% of them had more than 20% clay particles. High levels of sand were generally measured, as more than 90% of the soils had more than 40% of sand particles. Extremely high sand contents of 88% were reached in some cases. Almost 75% of the sampled soils were deep and could be sampled up to a 140-cm depth. Only one site was very shallow and could only be sampled up to 40 cm because of the presence of a petrocalcic horizon. Soil bulk density was low and varied on average between 1.22 g cm^{-3} in the superficial soil layer and 1.29 g cm^{-3} in the deepest sampled layer.

The average FC was almost twice as large as the WP in the upper soil layer with both hydraulic properties varying greatly among sites. Average $AWC_{0-20\ cm}$ was variable and represented 18% of $AWC_{0-140\ cm}$. The AWC of the shallowest soil was 26 mm while the highest AWC value of 235 mm corresponded to a deep soil profile. AWC measurements taken every 20 cm from the surface to a depth of 140 cm exhibited limited variability with depth (Figure 2). AWC of all sampled soil layers contributed proportionally to the $AWC_{0-140\ cm}$. For example, soils with 70% sand had an AWC of 10 mm per 20 cm soil layer while soils with 40% sand had an AWC of 20 mm. Variability among soils was high with coefficients of variation ranging from 20 to 50% depending on the variable. The soil textural composition had a 15-fold range of variation (clay content) and $AWC_{0-140\ cm}$ had a ninefold range. Soil organic matter was highly and significantly related to clay content but no significant relation was detected with $AWC_{0-20\ cm}$ (Figure 3). Deeper soils had larger AWC, but this association was not strong. $AWC_{0-20\ cm}$ was highly correlated with $AWC_{0-140\ cm}$ and $AWC_{20-140\ cm}$ (Table 2).

Table 1. Mean values and range of soil variables measured in the sampled sites ($n = 152$). Field capacity and wilting point correspond to gravimetric moisture contents. $AWC_{0-20\ cm}$ = available water capacity in the 0–20 cm layer; $AWC_{0-140\ cm}$ = available water capacity in the 0–140 cm layer. CV = Coefficient of variation.

	Organic matter (%)	Clay (%)	Sand (%)	Sampling (cm)	Bulk density (g cm^{-3})	Bulk density (g cm^{-3})	Field capacity (%)	Wilting point (%)		
	0–20 cm	0–20 cm	0–20 cm	depth	0–20 cm	120–140 cm	0–20 cm	0–20 cm	$AWC_{0-20\ cm}$ (mm)	$AWC_{0-140\ cm}$ (mm)
Minimum	0.43	2.01	20.2	40	0.91	1.04	5.37	2.48	5.68	25.9
Mean	1.97	11.7	57.9	127	1.22	1.29	14.5	7.22	17.6	99.9
Maximum	5.69	30.8	87.5	140	1.46	1.49	23.6	14.7	38.8	234.7
CV	41.1	50.2	26.2	19	10.7	8.17	30.8	38.6	35.1	46

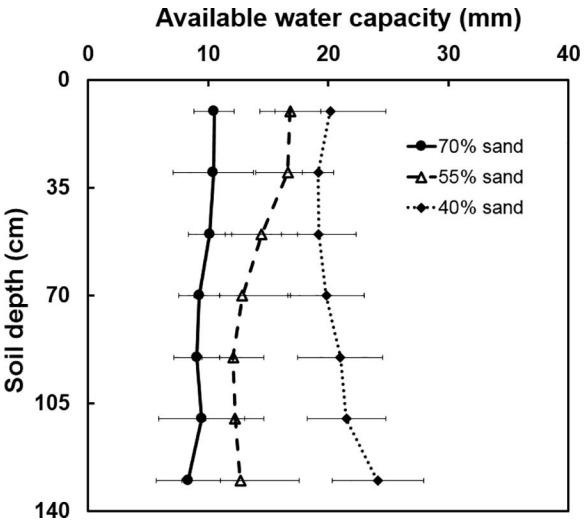

Figure 2. Measured soil available water capacity across the sampled sand gradient of the Semiarid Pampas. Each curve represents the average soil available water capacity per soil layer of five randomly chosen sampled soils. The sand gradient was chosen to represent the 10, 50 and 90 percentiles of the data set and correspond to the content of the upper 0–20 cm soil layer. Horizontal bars represent standard deviations.

One-step approach for soil $AWC_{0-140\ cm}$ estimation

The best fitted regression model for $AWC_{0-140\ cm}$ estimates used two soil variables as predictors, $AWC_{0-20\ cm}$ and soil depth:

$$AWC_{0-140\,cm}\ (mm) = -3.29 + 0.0464\ ^{*}\ AWC_{0-20\,cm}\ (mm)\ ^{*}\ \text{soil depth}\ (cm) \quad (4)$$

The structure of this model is logical as it is very similar to the result obtained when dividing $AWC_{0-20\ cm}$ by 20 which implies having an AWC value per soil centimeter, and multiplying the result by the soil depth. When the model of Eq. (4) and the latter calculation were compared, very similar R^2, ME, and RMSE values were attained but the MD was lower using Eq. (4). The regression model accounted for 78% of $AWC_{0-140\ cm}$ variation, a performance that could not be significantly improved by the best artificial neural network fitted (Figure 4 and Table 3). The inputs of the neural network were the same as the predictors used in the regression model and network architecture included two neurons in the hidden layer.

The different modeling techniques showed a strong interaction between $AWC_{0-20\ cm}$ and soil depth. The linear regression of observed vs. estimated values of the regression and the neural network models had an intercept and a slope that were not significantly different from 0 and 1, respectively ($P < 0.05$). Differences between R^2 of training and validation data sets were not significantly different ($P < 0.05$) and RMSE, MD, and ME were also similar (Table 3).

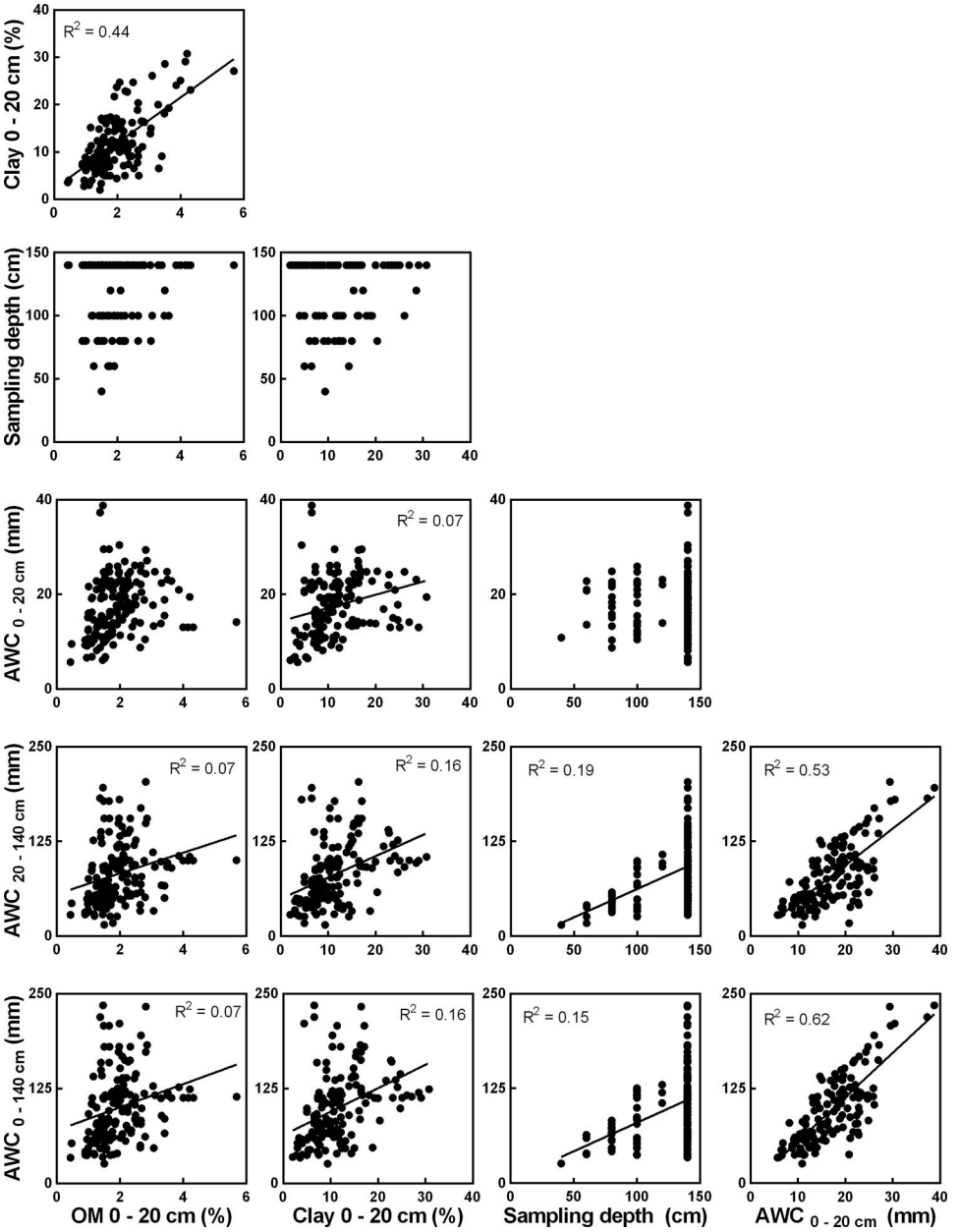

Figure 3. Scatter plots of the relationships among measured soil properties. Only significant linear regressions were drawn ($P < 0.05$). Available water content (AWC) of the 0–140 cm depth is the AWC to the maximum sampling depth that in most cases was 140 cm but in soils with petrocalcic horizon was to the upper limit of that horizon.

Table 2. Linear relationship between available water capacity of the 0–20 cm soil layer ($AWC_{0–20\ cm}$) and the available water capacity of the 0–140 cm ($AWC_{0–140\ cm}$) or 20–140 cm ($AWC_{20–140\ cm}$) layers. Both the intercepts and slopes are significantly different from 0 and 1, respectively ($P < 0.05$).

Dependent variable	Intercept	Slope	Independent variable	R^2
$AWC_{0–140\ cm}$	−2.94	5.83	$AWC_{0–20\ cm}$	0.62
$AWC_{20–140\ cm}$	−3.03	4.84	$AWC_{0–20\ cm}$	0.53

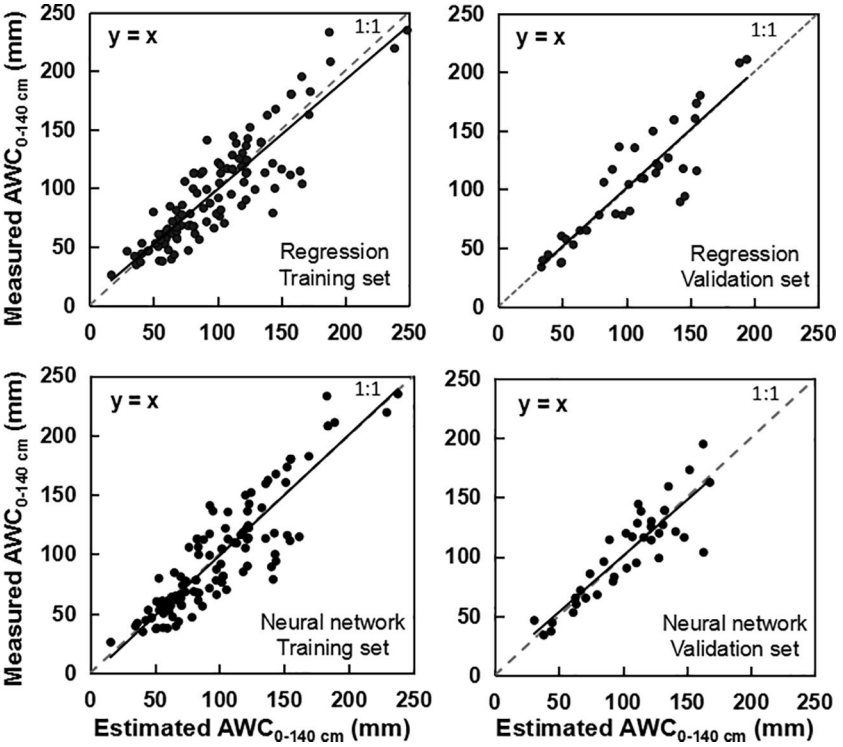

Figure 4. Performance of regression and artificial neural network (ANN) models corresponding to the one step approach developed to estimate $AWC_{0–140\ cm}$ (soil available water capacity of the 0–140 cm layer or to the upper limit of a petrocalcic horizon). The 1:1 relationships (dashed lines) and the fitted functions (solid lines) are shown. Intercept and slope values of the regressions were not different from 0 and 1, respectively and are not shown in the figure.

Table 3. Whole profile soil available water content ($AWC_{0–140\ cm}$) was estimated using regression and an artificial neural network. Coefficient of determination (R^2), root mean squared error in mm (RMSE), mean difference (MD), and modeling efficiency (ME) of the observed vs. estimated values.

$AWC_{0–140\ cm}$ estimation—One-step approach					
		R^2	RMSE	MD	ME
Regression model	Training set	0.78	21.3	1.06	0.77
	Validation set	0.80	21.1	−1.03	0.80
Neural model	Training set	0.79	22.1	1.24	0.79
	Validation set	0.78	18.2	0.16	0.74
$AWC_{0–140\ cm}$ estimation—Two-step approach					
First step: $AWC_{20–140\ cm}$					
		R^2	RMSE	MD	ME
Regression model	Training set	0.72	21.8	1.76	0.72
	Validation set	0.71	22.7	2.07	0.70
Neural model	Training set	0.71	21.6	−1.82	0.71
	Validation set	0.77	21.1	−2.39	0.76
Second step: measured $AWC_{0–20\ cm}$ + estimated $AWC_{20–140\ cm}$					
		R^2	RMSE	MD	ME
Regression model	Training set	0.78	21.8	1.76	0.77
	Validation set	0.76	22.7	2.07	0.75
Neural model	Training set	0.77	21.6	−1.82	0.77
	Validation set	0.82	21.1	−2.39	0.81

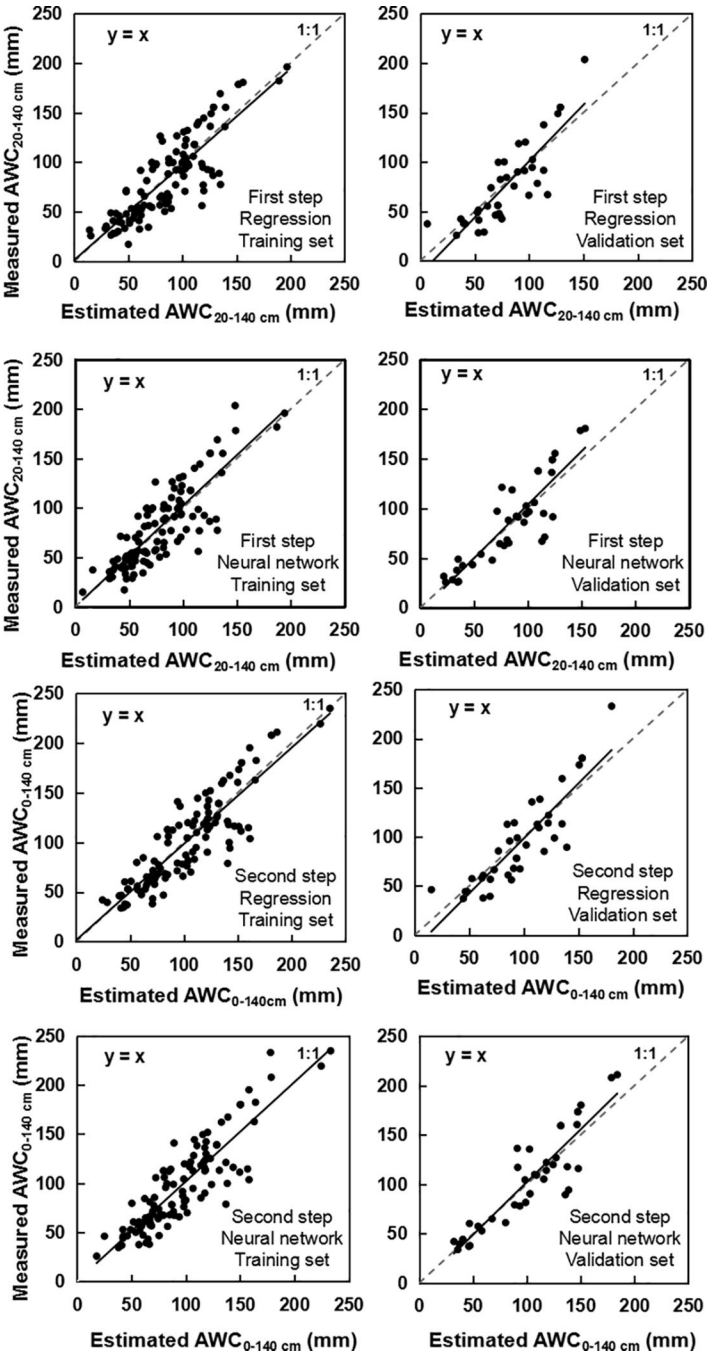

Figure 5. Performance of regression and artificial neural network (ANN) models corresponding to the two step approach developed to estimate $AWC_{20–140\,cm}$ (soil available water capacity of the 20–140 cm layer) and $AWC_{0–140\,cm}$ (soil available water capacity of the 0–140 cm layer). The 1:1 relationships (dashed lines) and the fitted functions (solid lines) are shown. Intercepts and slopes of the regressions were not different from 0 and 1 respectively, and are not shown in the figure.

Two-step approach for $AWC_{20-140\ cm}$ estimation

The best regression model fitted for $AWC_{20-140\ cm}$ prediction was:

$$AWC_{20-140\,cm}\ (mm) = -95.3 + 4.85\ ^{*}\ AWC_{0-20\,cm}\ (mm) + 0.741\ ^{*}\ \text{soil depth}\ (cm) \quad (5)$$

This regression model accounted for 72% of $AWC_{20-140\ cm}$ variability (Figure 5 and Table 3). In the second step, when summing measured $AWC_{0-20\ cm}$ to this latter estimate, it explained 78% of $AWC_{0-140\ cm}$ variability (Figure 5 and Table 3). In this case, the interaction term was not significant. The performance of the neural network model was similar. The selected inputs were the same and the network structure included four neurons in the hidden layer. In the second step, 79% of $AWC_{0-140\ cm}$ variability could be explained (Figure 5 and Table 3). No significant differences ($P < 0.05$) in R^2 between training and validation sets were detected for both modeling techniques and both prediction strategies. RMSE, MD, and ME were also similar (Table 3). The regression of observed vs. estimated values had a slope that did not differ significantly from 1 and an intercept that was not different from 0 ($P < 0.05$).

For simplicity, the regression model following the one-step approach was chosen to explore the effects of soil factors that significantly affected $AWC_{0-140\ cm}$ in these soils. Using the results of the sensitivity analysis, a regression surface plot and a nomogram were developed. The regression response surface showed that the effect of both soil variables on $AWC_{0-140\ cm}$ was positive and almost linear (Figure 6a). Shallow soils and soils with low $AWC_{0-20\ cm}$ never had high $AWC_{0-140\ cm}$ and, on the contrary, the deepest soils with the highest $AWC_{0-20\ cm}$ had the highest $AWC_{0-140\ cm}$ (Figure 6b). The model showed a positive curvilinear trend when both independent variables presented low values that tended to become linear as values became higher. The regression surface plot is suitable for showing the interrelationships between soil

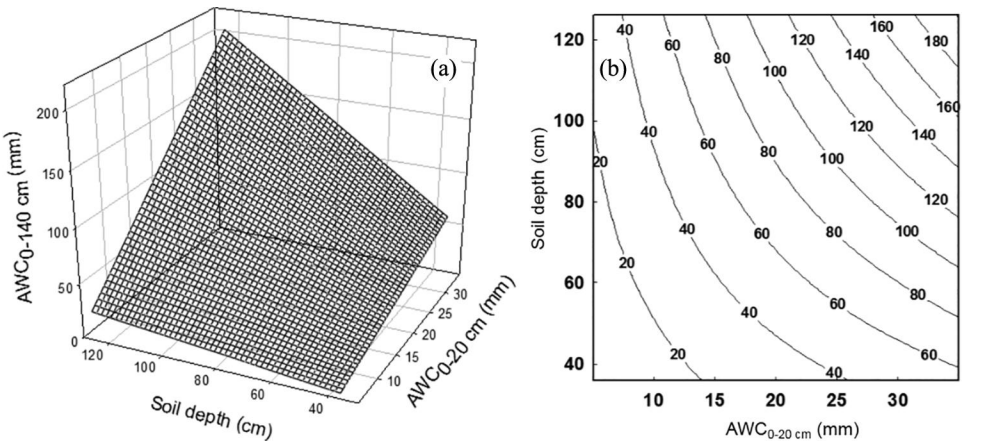

Figure 6. (a) Regression surface plot of $AWC_{0-140\ cm}$ (soil available water capacity of the 0–140 cm layer) as a function of $AWC_{0-20\ cm}$ (soil available water capacity of the 0–20 cm layer) and depth. (b) Nomogram representing the same effects. Numbers on the lines are $AWC_{0-140\ cm}$ (mm). Soil depth corresponds to the maximum sampling depth that in most cases was 140 cm but in soils with petrocalcic horizon was to the upper limit of that horizon.

variables that impacted $AWC_{0-140\ cm}$ while the nomogram is useful to visually estimate $AWC_{0-140\ cm}$.

Discussion

The broad spatial scale covered by our sampling reflected a large variation in soil properties which ensures an ample domain for the $AWC_{0-140\ cm}$ fitted estimation model. Organic matter content was unrelated to $AWC_{0-20\ cm}$. Similarly, a small impact of organic matter on hydraulic soil properties was also found in other studies, effects were statistically significant (Schaap and Bouten 1996; Schaap, Leij, and van Genuchten 1998; Rab et al. 2011). The impact of texture on $AWC_{0-20\ cm}$ was significant in the Semiarid Pampas, something which is consistent with results from other regions (Gupta and Larson 1979; Saxton et al. 1986; Ghorbani-Dashtaki and Homaee 2004). Although surface variables greatly differed between sites, hydraulic properties remained quite constant along the profiles. Our hypothesis was confirmed as we were able to develop an $AWC_{0-140\ cm}$ estimation model with a good performance using surface information. Even though surface $AWC_{0-20\ cm}$ and $AWC_{0-140\ cm}$ correlated well ($R^2 = 0.62$, $P < 0.0.5$) the addition of soil depth as a predictor in the regression model not only significantly improved the estimate by approximately 17% ($R^2 = 0.78$, $P < 0.0.5$) but also revealed a positive interaction between both variables.

The data set used ($n = 152$) was large enough to test neural networks as shown in some previous studies in the Pampas (Alvarez, Steinbach, and Bono 2011; Alvarez and Steinbach 2011). The only requisite was to split the data into training and validation sets as we have done (Bishop 2006) and to perform an early stop of weight fitting to avoid overlearning (Özesmi et al. 2006). Although artificial neural networks are useful when relationships between variables are difficult to model due to nonlinear and complex interactions (Moosavizadeh-Mojarrad and Sepaskhah 2012), our network models did not improve the fit of regression models. Similar results were reported previously for soil water curve retention modeling (Schaap and Bouten 1996). Multiple linear regression methods may perform better than neural networks when applied to estimates of soil water retention curves if relationships between soil properties tend to linearity (Minasny, McBratney, and Bristow 1999) as in our case. Because we found no performance differences between both modeling methods, we chose regressions to avoid the need of special software for simplicity.

Using surface information to estimate whole-profile soil data, as we propose in the one-step approach, is a typical case of a part-whole correlation which should not be considered as a spurious regression and is statistically sound (Sokal and Rohlf 1995). Graphical information provided by the scatterplots of observed vs. estimated values of the fitted models showed that these made reasonable predictions. In order to make model comparison more robust, not only the determination coefficients were used but also other statistics (Willmott 1981; Comrie 1997) that showed similar results. Most of $AWC_{0-140\ cm}$ variance could be explained by the regression and by the neural network models following the one-step approach and the two-step approach did not improve the fit. The RMSE of 21 mm for the regression model represented 20% of the mean $AWC_{0-140\ cm}$. Mean differences were low implying that the estimated values were very close to the observed values and the ME was very good (Mayer and Butler 1993).

The $AWC_{0-140\,cm}$ regression estimate model requires only $AWC_{0-20\,cm}$ determination, limiting laboratory work. Soil depth information can be available from soil surveys or can be generated easily. The coefficients of our model must be considered specific for the Semiarid Pampas and it would be risky to extrapolate them to other soils. However, the methodology appears suitable for other sandy soils for which adequate coefficients should be fitted. Regression models can simply be translated into nomograms avoiding the need of calculation for some users.

Conclusion

In sandy areas it is possible to estimate soil AWC along the whole soil profile by determining only soil water properties at the surface layer. In cases where soil depth may be restricted by horizons that limit root growth, the depth to those horizons must also be assessed to obtain a sound estimate.

Funding

This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas [Grant Number PIP 11220130100084CO] and UBACyT [Grant Number 20020130100484,20020150200015BA].

ORCID

Josefina Luisa De Paepe http://orcid.org/0000-0002-1518-7133
Roberto Alvarez http://orcid.org/0000-0002-4149-4508

References

Alvarez, R. 2009. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *European Journal of Agronomy* 30:70–77. doi:10.1016/j.eja.2008.07.005.

Alvarez, R., and R. S. Lavado. 1998. Climate, organic matter and clay content relationships in the Pampa and Chaco soils, Argentina. *Geoderma* 83:127–41. doi:10.1016/s0016-7061(97)00141-9.

Alvarez, R., and S. Grigera. 2005. Analysis of soil fertility and management effects on yields of wheat and corn in the Rolling Pampa of Argentina. *Journal of Agronomy and Crop Science* 191:321–9. doi:10.1111/j.1439-037x.2005.00143.x.

Alvarez, R., and H. Steinbach. 2011. Modeling apparent nitrogen mineralization under field conditions using regressions an artificial neural networks. *Agronomy Journal* 103:1159–68. doi:10.2134/agronj2010.0254.

Alvarez, R., G. Berhongaray, A. Bono, H. Steinbach, and J. De Paepe. 2014. Rangos de densidad aparente medidos en suelos pampeanos. *Annals of the XXIV Soil Science Congress of Argentina*, 6 pp. Bahía Blanca, Argentina.

Alvarez, R., H. Steinbach, and A. Bono. 2011. An artificial neural network approach for predicting soil carbon budget in agroecosystems. *Soil Science Society of America Journal* 75:965–75. doi:10.2136/sssaj2009.0427.

Arya, L. M., J. C. Richter, and J. F. Paris. 1983. Estimating profile water storage from surface zone soil moisture measurements under bare field conditions. *Water Resources Research* 19:403–12. doi:10.1029/wr019i002p00403.

Berhongaray, G., R. Alvarez, J. L. De Paepe, C. Caride, and R. J. C. Cantet. 2013. Land use effects on soil carbon in the Argentine Pampas. *Geoderma* 192:97–110. doi:10.1016/j.geoderma.2012.07.016.

Bishop, C. M. 2006. *Pattern recognition and machine learning*. Singapore: Springer, 738.

Blake, G. R., and K. H. Hartge. 1986. Bulk density. In *Methods of soil analysis. Part 1. Physical and mineralogical methods*, ed. A. Klute. Madison, WI: SSSA.

Bono, A., and R. Alvarez. 2012. Use of surface soil moisture to estimate profile water storage by polynomial regression and artificial neural networks. *Agronomy Journal* 104:934–8. doi:10.2134/agronj2012.0011.

Bono, A. A., J. L. De Paepe, and R. Alvarez. 2011. In-season wheat yield prediction in the Semiarid Pampa of Argentina using artificial neural networks. In *Progress in Food Science and Technology*, ed. A. J. Greco, Vol. 1, 133–49. New York, USA: Nova Science Publishers, Inc.

Bouma, J. 1989. Using soil survey data for quantitative land evaluation. In *Advances in Soil Science*, ed. B. A. Stewart, 177–213. USA: Springer.

Comrie, A. C. 1997. Comparing neural networks and regression models for ozone forecasting. *Journal of the Air & Waste Management Association* 47 (6):653–63. doi:10.1080/10473289.1997.10463925.

Cutforth, H. W., S. V. Angardi, B. G. McConkey, P. R. Miller, D. Ultrich, R. Gulden, K. M. Vokmar, M. H. Entz, and S. A. Brandt. 2013. Comparing rooting characteristics and soil water withdrawal patterns of wheat with alternative oilseed and pulse crops grown in the semiarid Canadian prairie. *Canadian Journal of Soil Science* 93:147–60. doi:10.4141/cjss2012-081.

Dardanelli, J. J., O. A. Bachmeier, R. Sereno, and R. Gil. 1997. Rooting depth and soil water extraction patterns of different crops in a silty loam Haplustoll. *Field Crops Research* 54:29–38. doi:10.1016/s0378-4290(97)00017-8.

Fagioli, M. 1973. Desarrollo de los aparatos radicales en los cultivos de trigo y maíz en la Región de Pergamino. *RIA* 10:111–35.

Fagioli, M. 1983. Actividad absorbente de los aparatos radicales de cultivos de alfalfa y trigo medida con ^{32}P. Publicación Técnica No. 25 ISSN 03252132. EEA Anguil INTA.

FAO. 1975. *Sandy soils*. Food and Agriculture Organization of the United Nations. Rome, Italy. Bulletin #25. ISBN 92–5-100613-X.

FAO. 1991. *World soil resources: An explanatory note on the FAO world soil resources map at Scale 1: 25,000,000*. FAO World Soil Resources Report 66, FAO, Rome, 58 pp.

Fila, G., G. Bellocchi, M. Acutis, and M. Donatelli. 2003. IRENE: A software to evaluate model performance. *European Journal of Agronomy* 18:369–72. doi:10.1016/s1161-0301(02)00129-6.

Gee, G. W., and J. W. Bauder. 1996. *Particle-size analysis, methods of soil analysis. Part 3: Physical and mineralogical methods*. Madison, Wisconsin, USA: Soil Science Society of America, 383–412.

Ghorbani-Dashtaki, Sh., and M. Homaee. 2004. Using geometric mean particle diameter to derive point and continuous pedotransfer functions. In *EuroSoil*, ed. N. Whrle and M. Scheurer. September 4–12, 2004. Freiburg, Germany. 10 (30):1–10.

Gupta, S., and Larson, W. 1979. Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density. *Water Resources Research*. 15:1633–5. doi:10.1029/wr015i006p01633.

Hall, A. J., C. Rebella, C. Guersa, and J. Culot. 1992. *Field-crop system of the Pampas*. In *Field crop ecosystems of the World*, ed. C. J. Pearson, Vol. 18. Amsterdam: Elsevier, 413–50.

Hanson, J. D., K. W. Rojas, and M. J. Shaffer. 1999. Calibrating the root zone water quality model. *Agronomy Journal* 91:171–7. doi:10.2134/agronj1999.00021962009100020002x.

Havrylenko, S. B., J. M. Bodoque, R. Srinivasan, G. V. Zucarelli, and P. Mercuri. 2016. Assessment of the soil water content in the Pampas region using SWAT. *Catena* 137:298–309. doi:10.1016/j.catena.2015.10.001.

INTA. 1980. *Inventario integrado de los recursos naturales de la Provincia de La Pampa*. Buenos Aires, Argentina: Instituto Nacional de Tecnología Agropecuaria, 493.

INTA. 1989. *Mapa de suelos de la Provincia de Buenos Aires*. Buenos Aires, Argentina: Instituto Nacional de Tecnología Agropecuaria, 525.

INTA. 2003. *Recursos naturales de la Provincia de Córdoba*. Agencia Córdoba Ambiente, 512, Córdoba, Argentina.

Jackson, R. B., J. Canadell, J. R. Ehleringer, H. A. Mooney, O. E. Sala, and E. D. Schulze. 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia* 108 (3):389–411. doi:10.1007/bf00333714.

Kaul, M., R. L. Hill, and C. Walthall. 2005. Artificial neural networks for corn and soybean yield prediction. *Agricultural Systems* 85:1–18. doi:10.1016/j.agsy.2004.07.009.

Keating, B. A., P. S. Carberry, G. L. Hammer, M. E. Probert, M. J. Robertson, D. Holzworth, N. I. Huth, J. N. Hargreaves, H. Meinke, and Z. Hochman. 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* 18:267–88. doi:10.1016/s1161-0301(02)00108-9.

Kleinbaum, D. G., and L. L. Kuper. 1979. *Applied regression analysis and other multivariable methods*. North Scituate, MA: Duxbury Press.

Klute, A. 1986. Water retention: laboratory methods. In *Methods of soil analysis: Part 1. Physical and mineralogical methods*, ed. A. Klute, 635–662, *Agron. Ser*., vol. 9, 2nd ed. Madison, Wis., USA: Am. Soc. Agron.

Kobayashi, K., and M. U. Salam. 2000. Comparing simulated and measured values using mean squared deviation and its components. *Agronomy Journal* 92:345–52. doi:10.1007/s100870050043.

Lee, J. H. W., Y. Huang, M. Dickman, and A. W. Jayawardena. 2003. Neural network modelling of coastal algal blooms. *Ecological Modelling* 159:179–201. doi:10.1016/s0304-3800(02)00281-8.

Lin, L., A. S. Hedayat, B. Sinha, and M. Yang. 2002. Statistical methods in assessing agreement: Models, issues, and tools. *Journal of the American Statistical Association* 97:257–70. doi:10.1198/016214502753479392.

Liu, X., C. L. Burras, Y. S. Kravchenko, A. Duran, T. Huffman, H. Morras, G. Studdert, X. Zhang, R. M. Cruse, and X. Yuan. 2012. Overview of Mollisols in the world: Distribution, land use and management. *Canadian Journal of Soil Science* 92 (3):383–402. doi:10.4141/cjss2010-058.

Mayer, D. G., and D. G. Butler. 1993. Statistical validation. *Ecological Modelling* 68 (1–2):21–32. doi:10.1016/0304-3800(93)90105-2.

Miao, Y., D. Mulla, and P. Robert. 2006. Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. *Precision Agriculture* 7:117–35. doi:10.1007/s11119-006-9004-y.

MinAGRO. 2017. Ministerio de Agroindustria. Secretaría de Ganadería y Pesca. Accessed May 14, 2016. https://datos.magyp.gob.ar/reportes.php?reporte=Estimaciones.

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* 93:225–53. doi:10.1016/s0016-7061(99)00061-0.

Moosavizadeh-Mojarrad, R., and A. R. Sepaskhah. 2012. Predicting soil water retention curve by artificial neural networks. *Archives of Agronomy and Soil Science* 57:3–13. doi:10.1080/03650340903222302.

Nelson, L., and L. Sommers. 1996. Total carbon, organic carbon, and organic matter. In *Methods of soil analysis. Part 3—Chemical methods*, ed. D. L. Sparks. Madison, Wisconsin, USA: Soil Science of America, 961–1010.

Neter, J., W. Wasserman, and M. Kutnet. 1990. *Applied linear statistical models*. Illinois, USA: Irwin, 1172.

Noy-Meir I. 1973. Desert ecosystems: Environment and producers. *The Annual Review of Ecology, Evolution, and Systematics* 4:25–51. doi:10.1146/annurev.es.04.110173.000325.

Özesmi, S. L., C. O. Tan, and U. Özesmi. 2006. Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecological Modelling* 195:83–93. doi:10.1016/j.ecolmodel.2005.11.012.

Park, S. J., and P. L. G. Vlek. 2002. Environmental correlation of three-dimensional soil spatial variability: A comparison of three adaptive techniques. *Geoderma* 109:117–40. doi:10.1016/s0016-7061(02)00146-5.

Quiring, S. M., and D. R. Legats. 2008. Application of CERES-Maize for within season prediction of rainfed corn yields in Delaware, USA. *Agricultural and Forest Meteorology* 148:964–75. doi:10.1016/j.agrformet.2008.01.009.

Rab, M., S. Chandra, P. Fisher, N. Robinson, M. Kitching, C. Aumann, and M. Imhof. 2011. Modelling and prediction of soil water contents at field capacity and permanent wilting point of dryland crop soils. *Soil Research* 49:389–407. doi:10.1071/sr10160.

Rawls, W., D. Brakensiek, and K. Saxton. 1982. Estimation of soil water properties. *Transaction of ASAE* 25:1316–32. doi:10.13031/2013.33720.

Romano, N., R. Alvarez, and A. Bono. 2016. Modeling nitrogen mineralization at surface and deep layers of sandy soils. *Archives of Agronomy and Soil Science* 1:1–13. doi:10.1080/03650340.2016.1241391.

SAGyP-INTA. 1990. Atlas de suelos de la República Argentina. *Escala* 1:500.000 y 1:1.000.000. Vol. I, 731 pp.; Vol. II, 677 pp.

Satorre, E., and G. Slafer. 1999. Wheat production systems of the Pampas. In *Wheat. Ecology and physiology of yield determination*, ed. E. Satorre and G. A. Slafer. NY, USA: Food Products Press, 333–48.

Saxton, K., W. J. Rawls, J. Romberger, and R. Papendick. 1986. Estimating generalized soil-water characteristics from texture. *Soil Science Society of America Journal* 50:1031–6. doi:10.2136/sssaj1986.03615995005000040054x.

Schaap, M. G., and W. Bouten. 1996. Modeling water retention curves of sandy soils using neural networks. *Water Resources Research* 32:3033–40. doi:10.1029/96wr02278.

Schaap, M. G., Leij, F. J., and M. T. van Genuchten. 1998. Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Science Society of America Journal* 62:847–55. doi:10.2136/sssaj1998.03615995006200040001x.

Shen, J., F. Zhang, R. Li, Z. Rengel, and C. Tang. 2003. Orthogonal polynomial models to describe yield response of rice to nitrogen and phosphorus at different levels of soil fertility. *Nutrient Cycling in Agroecosystems* 65:243–51.

Silveira, C. T., C. Oka-Fiori, L. J. C. Santos, A. E. Sirtoli, C. R. Silva, and M. F. Botelho. 2013. Soil prediction using artificial neural networks and topographic attributes. *Geoderma* 195:165–72. doi:10.1016/j.geoderma.2012.11.016.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: The principles and practice of statistics in biological research. W.H. Freeman and Company. Pp. 890. New York, USA.

Tedeschi, L. O. 2006. Assessment of the adequacy of mathematical models. *Agricultural Systems* 89:225–47. doi:10.1016/j.agsy.2005.11.004.

Teruggi, M. E. 1957. The nature and properties of argentine loess. *Journal of Sediment Petrology* 27:322–32.

USDA. 2014. *Keys to soil taxonomy*. 12th ed. Washington, DC, USA: USDA-Natural Resources Conservation Service.

Williams, J. R., C. A. Jones and P. T. Dyke. 1990. The EPIC model. In *EPIC– Productivity impact calculator. 1. Model documentation*, ed. J. R. Williams, 3–86. U.S. Department of Agriculture–Agricultural Research Service; Springfield, Va.: National Technical Information Service Technical Bulletin Number 1768. Washington D.C. USA.

Willmott, C. J. 1981. On the validation of models. *Physical Geography* 2:184–94.

Wong, M., and S. Asseng. 2006. Determining the causes of spatial and temporal variability of wheat yields at sub-field scale using a new method of up scaling a crop model. *Plant and Soil* 283:203–15. doi:10.1007/s11104-006-0012-5.

Zuur, A., E. N. Ieno, N. Walker, A. A. Saveliev, and G. M. Smith. 2009. *Mixed effects models and extensions in ecology with R*. Springer Science and Business Media, 579. New York, USA.