Modelling of permanent wilting point from routine soil properties on a typical alfisol

Modelo para predecir el punto de marchitamiento permanente a partir de las propiedades de un suelo alfisol típico

Affinnih Kehinde[1], Olaniyan John[1], Ahamefule Henry[1] and Anwanane Nnenne[1]

[1]University of Ilorin, Faculty of Agriculture, Department of Agronomy, PMB 1515 Ilorin, Kwara State Nigeria.

*Corresponding author: affinnih.ko@unilorin.edu.ng; kehkennyy@yahoo.com

Abstract

Soil water holding capacity at permanent wilting point is imperative for plant water stress in specific soil type. This study was undertaken to formulate a regression model or equation for predicting permanent wilting points (PWP) of soils on a typical *Alfisol* of basement complex origin at the Teaching and Research Farm of the University of Ilorin. A total of forty five (45) disturbed and forty five (45) undisturbed soils samples were collected along a toposequence (upper, middle and bottom slope) at 3 depths: 0 cm – 30 cm, 30 cm – 60 cm, and 60 cm – 90 cm. Soil properties of the disturbed and undisturbed samples were determined using basic experimental methods and/ or calculated using reputable techniques. The measured soil properties include the proportions of soil separates, bulk density, total porosity, PWP and organic matter. Three different models were developed for predicting PWP of soil using regression model technique. There was no significant relationship between PWP and soil separates, bulk density and total porosity. However, only the silt content was positively correlated with PWP (r=0.22; p<0.05). Although, model three of PWP with the highest adjusted coefficient of determination (0.2952) emerged as the optimal choice. The model clarifies 30 % of part of variance in the mean square error of PWP with sand, silt and clay contributing statistically to the model. This implies that additional variables and techniques such as spatial and machine learning aside those used in the present study would provide a more reliable pedotransfer function for predicting PWP in the soil.

Keywords: Pedotransfer function, regression models, soil moisture content, toposequence, water coefficient.

Resumen

La capacidad de retención de agua del suelo en el punto de marchitamiento permanente es fundamental para el estrés hídrico de las plantas en un tipo de suelo específico. Este estudio se llevó a cabo para formular un modelo de regresión o ecuación para predecir los puntos de marchitamiento permanente (PWP) de los suelos en un Alfisol típico de origen complejo basal en la Granja de Enseñanza e Investigación de la Universidad de Ilorin. Se recogió un total de cuarenta y cinco (45) muestras de suelos alterados y cuarenta y cinco (45) no alterados a lo largo de una toposequencia (pendiente superior, media e inferior) a 3 profundidades: 0 cm - 30 cm, 30 cm - 60 cm, y 60 cm - 90 cm. Las propiedades del suelo de las muestras perturbadas y no perturbadas se determinaron utilizando métodos experimentales básicos y/o se calcularon utilizando técnicas acreditadas. Las propiedades del suelo medidas incluyen las proporciones de suelo separadas, densidad aparente, porosidad

total, PWP y materia orgánica. Se desarrollaron tres modelos diferentes para predecir la PWP del suelo utilizando la técnica del modelo de regresión. No se observó ninguna relación significativa entre la PWP y los separadores del suelo, la densidad aparente y la porosidad total. Sin embargo, solo el contenido de limo se correlacionó positivamente con PWP (r=0.22; p<0.05). Aunque, el modelo tres de PWP con el coeficiente de determinación ajustado más alto (0.2952) emergió como la elección óptima. El modelo aclara el 30 % de parte de la varianza en el error cuadrático medio de PWP con arena, limo y arcilla contribuyendo estadísticamente al modelo. Esto implica que variables y técnicas adicionales, como el aprendizaje espacial y automático, aparte de las utilizadas en el presente estudio, proporcionarían una función de pedotransferencia más fiable para predecir la PWP en el suelo.

Palabras clave: Función de pedotransferencia, modelos de regresión, contenido de humedad del suelo, toposequencia, coeficiente hídrico.

Introduction

The essence of modeling lies in simplifying complex realities into manageable representations called models (Kinoshita et al., 2012). These models capture key elements that are essential for understanding or predicting a specific outcome (Van Looy et al., 2017). For example, in soil science, models can represent the behavior of soil systems (Obi et al., 2012). One key parameter in such models is the permanent wilting point (PWP) (Kinoshita et al., 2012). This represents the moisture level in soil at which foliage permanently shrivel due to their inability to conduct moisture from soil. Understanding soil water retention, including the PWP, is crucial for optimizing water management practices in agriculture, ultimately leading to sustainable and improved production. Knowledge of soil moisture content is vital to decide plant variety to be grown, available water for plant growth, water stress as well as solute movement, evapotranspiration, cropping systems, tillage management, infiltration, drainage and irrigation scheduling and other assorted hydro-physical processes.

Soil water holding capacity is critical for simulating the hydrological behaviour of landscapes and assessing the suitability of soil for various applications (Kukal et al., 2023). This ability to retain water is primarily determined by capillarity, the physical phenomenon arising from the interaction between water molecules and soil particles. The effectiveness of capillarity, in turn, depends heavily on the structure of soil pores.

Several key factors collectively influence soil water retention, including texture (the size and proportion of particles greatly impact pore size and water holding capacity), structure (the organization of particle and aggregates affects pore connectivity and water movement), bulk density (the density of soil particles influences the volume of pores available for water storage), organic matter (OM) content (organic matter increases pore space and water holding capacity), etc. (Amsili et al., 2022; Vereecken et al., 1989). Since water retention is pretentious by these aforementioned physical properties, empirical relationships accelerated by advances in computer modelling (Minasny et al., 1999) could be developed for their prediction in soils, which can help in making informed decisions about land use and management. In line with this, pedotransfer function (PTF), defined as converting available data (those we have) into useful information (what we need) was devised by Bouma (1989). This allowed for the creation of functions that predict the values of specific soil properties based on other, more readily and economically measurable properties. These PTFs are often established from empirical observations whose applicability would however be restricted to the datasets employed in generating the model (Donatelli et al., 1996; Wosten et al., 1999). The general form of the linear regression equation is:

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 + b_6 x_6 \quad (1)$$

where Y denotes the dependent variable such as water content at selected water potential, b_0 the intercept, b_1 to b_6 are the regression coefficients and x_1 to x_6 represents the independent variables signifying the basic soil properties.

The diverse characteristics of soil across a toposequence, a sequence of soils along a slope, arise from a complex interplay of natural and human influences. Geological processes

establish the initial foundation, while soil formation further shapes these properties. Land use and management practices over time leave their mark, and natural forces like erosion and deposition further sculpt the landscape, leading to the unique mosaic of soil properties observed across the toposequence (Phillips, 2007).

Research has shown a strong link between soil particle size, its derivatives, solute transport properties, other characteristics (Amsili et al., 2024; Mbagwu et al., 1983; Ogban & Ekerette, 2001), as well as even mineralogy (Souza et al., 2009). Understanding these relationships could help model soil properties related to the PWP (-1500 kPa), the critical soil moisture threshold for plant survival. This study purposes to create a regression model for predicting the permanent wilting point in *Alfisols* soils at the Teaching and Research Farm of the University of Ilorin. By establishing these relationships, the model could aid in optimizing water management and ensure timely irrigation in agricultural practices.

Materials and methods

Description of Study Area

This research was carried out on a toposequence (upper, middle, and bottom slope positions) at the Teaching and Research Farm of the University of Ilorin, Ilorin, Nigeria. The region lies within the Southern Guinea Savanna zone (Lat. 9° 29' N, long. 4° 35' E, 307 m elevation) and features a tropical climate with bimodal rainfall (1000-1240 mm annually) and temperatures ranging from 20 °C - 35 °C (Kolo et al., 2012). The dominant soil type is gravelly Alfisols, formed over a basement complex (Olaniyan, 2003). Historically used for agriculture, the site was in a fallow state during the sampling period.

Soil Sampling and Analysis

Soil were sampled along the toposequence (upper, middle, and bottom slope positions) at three depths (0 cm -30 cm, 30 cm -60 cm, and 60 cm - 90 cm). Fifteen mini-pits were excavated, with five located at each slope position. Ninety samples were collected (45 disturbed, 45 undisturbed). Undisturbed samples were collected using metallic cylinders (8.3 cm height, 5.5 cm diameter). The soil within the cylinders was held in place with calico and rubber bands, then carefully labelled. A soil auger was used to collect disturbed samples and put in labelled polythene bags. These samples were transported to the laboratory for further physical and chemical properties determination using standard procedures.

Preparation of Soil Samples

The disturbed soil samples were air-dried, then ground and sieved to pass through a 2 mm mesh size sieve prior to analysis.

Soil Properties Analysis

Particle Size Analysis

Particle size analysis was conducted by the hydrometer method reported by Gee & Or (2002) employing sodium hexametaphosphate (calgon) as dispersant.

Bulk Density

Core method was employed for the determination of Bulk density, following the procedure outlined by Blake & Hartge (1986). First, undisturbed soil was dried using hot air oven at 105 °C till constant weight was attauned. Bulk density was computed using the formula:

$$\rho_b = Ms/Vt \tag{2}$$

Where, ρ_b = b*ulk density (kg/m^3)*,

Ms = oven-dried soil mass (kg),

Vt = volume of soil (m^3)

Total Porosity

Total porosity was calculated as:

$$\text{Ø} = [1 - \rho_b/\rho_s] \tag{3}$$

Where, Ø =Total porosity (m^3/m^3)

ρ_b,= *bulk density* (kg/m^3),

ρ_s= p*article density asummed to b*e 2650 kg/m^3

Soil Organic Matter

The determination of soil organic carbon (OC) was done using the Walkley-Black wet oxidation method (Nelson & Sommers, 1982). To convert the measured OC content into soil organic matter (OM), a standard conversion factor of 1.724 was applied. This factor accounts for the approximate 58 % carbon composition of SOM (Brady & Weil, 1999).

Permanent Wilting Point (PWP)

The determination of PWP was done using a procedure adapted from Odu et al. (1986). 300 g of air-dried soil that had been made to pass through a 2 mm sieve from each sampling point was moistened in pots, and three maize seeds were planted in each pot. After thinning to one seedling per pot, an aluminum ring was placed around the base of each plant before the opening of the coleoptile and pressed slightly into the soil. The plants were then allowed to grow until they reached the four-leaf stage. Next, the soil surface of each pot was sealed with a ¼-inch layer of molten paraffin wax. Finally, cotton wool was used to fill any gaps between the seedling stem and the aluminum ring, ensuring that water loss could only occur through the plant. The seedlings were left to grow and monitored for clear signs of wilting. Upon initial wilting, plants were shaded overnight to observe if they recovered. If wilting persisted, PWP was confirmed. The samples were then weighed prior to oven-drying and afterwards, with the difference in weight (as percentage dry soil) determining the PWP.

Statistical Tool

Correlation Analysis

When two or more quantities vary in sympathy such that movement in one tend to be accompanied by analogous movement in the others, at that time such quantities are ascribed to be correlated.

A multiple correlation was employed for the exploration of the relationship between measured pair of values. The statistical tool was given as:

$$r = ((Z_{ij})(Z_{ik}))/n \tag{4}$$

and $Z_{ij} = X_{ij} - X)/S_j \tag{5}$

is the standardized values of the response variable.

$$S_j = \sqrt{((\sum(X_{ij} - X_j))/(n-1))} \tag{6}$$

is the sample variance of the response variable

This correlation coefficient is generally called *Karl Pearson's* coefficient of correlation. Correlation coefficient quantifies the level of association (linear relationship) amid two or more random variables. Correlation coefficient may be positive or negative. A significant positive correlation coefficient indicates that an increase in one variable is accompanied by a similar increase in the other. While significant negative correlation coefficient indicates an inverse relationship between the two random variables in question. On the other hand, a non-significant correlation coefficient implies zero relationship between the variables under study.

A test statistic suggested by Morrison (1976) shall be used to test for significance of this correlation coefficient. This test is given as:

$$t = r \sqrt{((n-2)/(1-r^2))} \tag{7}$$

where, r is the so called Karl Pearson's correlation coefficient; n is sample size. It can be shown that when H_o is true, $t \sim t_{\alpha/2}(n-2) \tag{8}$

Model Development

Permanent Wilting Point

An underlying assumption in multiple regression is that the predictions (independent variables) are known without any uncertainty in their given values. Consequently, for the model to be applicable for both the estimation and predictions of the datasets, all assumptions were investigated and necessary violations corrected based on the fit of the models. Generally, the diagnostic plots from the regression models was used to check assumptions including normality of residuals and presence of potential outliers. This was done by using "Residuals versus Fitted" charts to show if there was a trend to the residuals and Shapiro-Wilk's test of residuals normality in the R software

for statistical computing and graphics (R core team, 2022) to confirm residuals normality. To ensure the validity of the developed model, it's crucial to assess whether the model's residuals (errors) follow a normal distribution. Two methods were employed for this purpose. First, a residual plot was visually inspected. Ideally, this plot should exhibit a random scattering of points around zero, indicating no significant relationship between the residuals and any individual predictor variable used in the model (Olorede et al., 2013; Olorede & Mudasiru, 2013). Second, a quantile-quantile (Q-Q) plot was generated. In this plot, if the points, especially those in the central region, fall close to a diagonal line, it suggests good agreement between the observed data and a normal distribution. Standardized residual criterion was employed to check effects of potential outliers in the dataset on the models by removing observations with standardized residuals outside the interval (Barnett & Lewis, 1994; John & Prescott, 1975; Stefansky, 1972). Also, the residuals of the models were standardized, this was achieved by the division of each by the root mean square error of its respective model. The expectation was for the lowest standardized residual to lie within ±1, and the highest to be within ±2; deviations from these ranges indicated potential outliers. All statistical analyses, model fitting and diagnostics were done using the language R version 4.2.2 (R Core Team, 2022).

In multiple linear regression scenarios like this, it is commonly acknowledged that various hypothesis tests concerning the model parameters are valuable for assessing the model's effectiveness. Therefore, the need to describe and test hypotheses about parameters of the proposed regression model collectively as well as individually. This would ascertain whether there is a notable correlation between the response variables (PWP) and a specific subset of the predictors. The built models were thus screened based on number of significant parameters, maximum amount of proportion of variability about the response by these parameters and satisfaction of model assumptions using diagnostic plots.

Rejection of null hypothesis about the full model (model with all parameters) implies that at least one of the predictors donates meaningfully to the model. Rejection of null hypothesis about individual regression parameter (ANOVA table) indicates that the variable cannot be deleted from the model. Coefficient of multiple determination (R^2) measures magnitude of decrease in the variability of permanent wilting point gotten by using the 6 predictors in the model. Merely having a high R-squared value doesn't automatically indicate the regression model's quality (Myers & Montgomery, 1995). Adding an extra predictor continually boosts R-squared, regardless of the statistical significance of the added variable. Consequently, models with high R-squared values might yield inaccurate predictions for new observations or mean response estimates. Hence, certain regression model developers opt for *adjusted* R-squared (Myers & Montgomery, 1995).

Generally, will not constantly escalate as variables increase in the model. If the additional variables are superfluous, the will repeatedly drop. A marked difference between R^2 and is a worthy coincidental that the model includes terms that are not statistically significant. It is imperative to mention that presence of multicollinearity in the data set will make estimations of coefficients from the least squares fit imprecise and statistically insignificant (Martens & Naes, 1989). Hence, when the aim is basically to forecast Y using a set of X variables, multicollinearity is not a significant concern. The predictions are still precise, and the overall R^2 (or) indicates the accuracy of Y values prediction. However, if the objective is to comprehend in what ways various X variables affect Y, then multicollinearity is a big issue. The first challenge that individual p-values may be deceptive; a high p-value may suggest insignificance even if the variable is crucial. Second, confidence intervals on regression coefficients could be wide, potentially encompassing zero. This ambiguity makes it difficult to determine whether an increase in X corresponds to a rise or fall in Y. Furthermore, wide confidence intervals mean that excluding or adding a subject can drastically alter coefficients, possibly even changing their signs.

In some instances, multiple regression results may appear inconsistent. Despite a low overall p-value, individual p-values for all X variables are high. This situation arises when two X variables are highly correlated, essentially conveying the same information such that one becomes redundant once the other is included. However, together, they significantly contribute to the model. Removing both variables would degrade the model fit considerably. Thus, while the overall model fits the data well, neither X variable makes a substantial contribution when added individually. This scenario indicates collinearity among the X variables and manifests as multicollinearity in the results.

Test of Hypothesis about Full Regression Model

Hypothesis 1: [Test of significance encompassing all regression parameters]

Null hypothesis (H$_0$): Model accurately represents data/ Model is suitable

Alternative hypothesis (H$_1$): Not H$_0$

Or mathematically,

Null hypothesis (H$_0$): $\beta_1=\beta_2=...=\beta_{13}=0$ [No single predictor showed a statistically significant contribution to the model]

Alternative hypothesis (H$_1$): $\beta_j \neq 0$ *for at least one j* [At least one predictor showed a statistically significant contribution to the model]

Tests statistic: $F_{ratio}= MS_{Regression}/MS_{Error}$ [Global F-test]

Decision rule: Reject the null hypothesis if and accept the alternative hypothesis at 0.05 significance level, or else do not reject the null hypothesis.

Hypothesis 2: [Test of significance encompassing individual regression coefficient]

Null hypothesis (H$_0$): $\beta_j=0$ [Predictor x$_j$ is not significant given that others are included in model]

Alternative hypothesis (H$_1$): $\beta_j \neq 0$ [Predictor is significant given that others are included in model]

Tests statistic: $t_{value}=\beta j/\sqrt{(C_{jj}\sigma^2)}$ **[Individual t-test]**

C_{jj}=diagonal element of the covariance matrix corresponding to β_j

Decision rule: Reject the null hypothesis if p-value<0.05 and accept the alternative hypothesis at 0.05 significance level, or else do not reject the null hypothesis.

Results and discussion

Results

The results of the descriptive statistics obtained for soil properties is presented in Table 1. Results for Pair-wise correlation of variables measured during PWP experiment is obtainable in Table 2. Results obtained for parameter estimates for the models generated for PWP is accessible in Table 3. Data obtained for the Analysis of Variance (ANOVA) for the multiple regression models established for permanent wilting point are presented in Table 4.

Table 1: Descriptive Statistics of Soil Properties

Soil Properties	Mean	Min	Max	Std dev	C.V
Sand (g.kg^{-1})	636.96	400	885	140.15	22.00
Silt (g.kg^{-1})	198.73	31	270	70.37	35.41
Clay (g.kg^{-1})	166.53	82	350	88.38	53.08
Bulk density (kg.m^{-3})	1402.36	1065	1698	140.19	10.00
Total porosity (m^3.m^{-3})	0.471	0.359	0.598	0.053	11.25
Organic matter (g.kg^{-1})	13.12	1.38	27.80	7.40	56.40
Permanent wilting point (%)	15.15	10.00	33.33	6.71	43.26

Statistical Analysis

Descriptive Statistics Analysis

The results obtained for the proportions of sand, silt and clay, bulk density, porosity, OM and permanent wilting point ranges from 400 g.kg^{-1} to 885 g.kg^{-1}, 31 g.kg^{-1} to 270 g.kg^{-1}, 82 g.kg^{-1} to 350 g.kg^{-1}, 1065 kg.m^{-3} to 1698 kg.m^{-3}, 0.359 m^3.m^{-3} to 0.598 m^3.m^{-3}, 1.38 g.kg^{-1} to 27.80 g.kg^{-1} and 10.00 % to 33.33 %, respectively (Table 1). The data implied a high sand content, indicating a coarse-textured soil, with a loose, crumbly structure while silt ranged between moderate to high; and clay content varied between high and very high. Thus, the soils employed in the present study are characterized as having low water availability, moderate drainage and high water retention following the soil separates. Bulk density, an index of soil compaction was high, thus, the soils has low available water capacity and permanent wilting point since water fills small pores limiting plant availability, thereby plants struggle to extract water from the dense soil. Total porosity rated moderate to high while organic matter content varied between low to moderate, and PWP data classifies the soil as low (10 % - 20 %), moderate (20 % - 30 %), and high (> 30 %).

Pair-wise correlation

A negative relationship between sand and silt (r = -0.870*) and clay (r = -0.9098*) was obtained, though a positive relationship existed between sand and OM (r = 0.668*) (Table 2). The implication of this is that as sand increases, silt and clay will decrease and as sand increases OM will increase. Silt was positively correlated with clay (r =0.603*) and negatively correlated with OM (r = -0.512*). As silt increases, clay will increase and organic matter will decrease. Clay was inversely correlated with OM (r = -0.689*). As clay increases organic matter will decrease. Bulk density was negatively correlated with total porosity (r = -0.999*), when bulk density increases, total porosity decreases at the same rate. Permanent wilting point didn't show any relation with the basic soil properties.

Permanent Wilting Point Model Development

Three models were created for permanent wilting point, each with p-values below 0.05, with the corresponding R^2 adjusted values of 0.2351, 0.2949 and 0.2952 for models 1, 2 and 3 as shown in Table 3. Model 3 of permanent wilting point was selected as the best predictive model since it recorded the highest R^2 adjusted value (0.2952). This means that the model accurately represents the data related to the permanent wilting point with the three predictors included after removal of observations with potential outliers and non-significant predictors bulk density, total porosity and organic matter to assess the model's predictive capability. The model accounts for 30% of the variation in mean squared errors of the permanent wilting point with sand, silt and clay making statistically significant contributions to the model.

Based on the individual statistical significance of each predictor variable, the findings shown in Table 4 for model 3 of permanent wilting point indicate that the null hypothesis for both silt and clay with corresponding p-values of 0.003126

Table 2: Pair-wise correlation results of variables measured during permanent wilting point experiment

	Sand	Silt	Clay	Bulk density	Total porosity	Permanent wilting point	Organic matter
Sand	1						
Silt	-0.87029266*	1					
Clay	-0.908686821*	0.602503779*	1				
Bulk density	0.08398562	0.01706638	-0.162240647	1			
Total porosity	-0.083583321	-0.017237012	0.161766436	-0.999996659*	1		
Permanent wilting point	-0.077695632	0.220709283	-0.016165673	-0.115205715	0.11555406	1	
Organic matter	0.667549445*	-0.512282805*	-0.689291767*	-0.003883358	0.003886232	-0.163623194	1

* Significant at 5% level when r ≥ 0.293955

and 0.010521 respectively can be rejected. The simple interpretation of the findings is that every one of the attributes are statistically significant in the model when others are included and cannot be disregarded. The expression used for model 3 as given in Table 3 is written in equation (9) as:

$$\hat{Y} = -114.52266 + 0.12427 \times Sand + 0.15744 \times Silt + 0.10606 \times Clay \tag{9}$$

For model 3 of permanent wilting point, the residuals are now normal and there are no potential outliers as shown in Figure 1. In fact, the Cook's distance plot now confirms this by having all the Cook's distances for all the predictors less than 1 as shown in Figure 1. The residuals plot and normal quantile-quantile plots presented in Figure 1 also support this. Normality of the

Table 3: Table of Parameter Estimates for Permanent Wilting Point Models

Model No.	P-value of Model	Model Multiple R-square value	Model Adjusted R-square value	Coefficients	Estimates	Std. Error	t value	P-value
1.	0.01892	0.3528	0.2351	(Intercept)	-9.745e+01	4.670e+03	-0.021	0.98348
				Clay	8.702e-02	4.323e-02	2.013	0.05234
				Sand	1.167e-01	4.294e-02	2.718	0.01037 *
				Silt	1.462e-01	4.480e-02	3.262	0.00257 *
				Bulk Density	-3.527e-03	1.767e+00	-0.002	0.99842
				Total Porosity	1.187e+00	4.663e+03	0.000	0.99980
				OM	-1.704e-01	1.236e-01	-1.379	0.17717
2.	0.007154	0.4062	0.2949	(Intercept)	-1.163e+03	4.461e+03	-0.261	0.79601
				Clay	8.598e-02	4.104e-02	2.095	0.04417 *
				Sand	1.146e-01	4.077e-02	2.810	0.00838 *
				Silt	1.478e-01	4.254e-02	3.475	0.00149 *
				Bulk Density	3.988e-01	1.688e+00	0.236	0.81472
				Total Porosity	1.067e+03	4.454e+03	0.240	0.81214
				OM	-1.583e-01	1.174e-01	-1.348	0.18712
3.	0.001555	0.3509	0.2952	(Intercept)	-114.52266	40.08649	-2.857	0.007153 *
				Sand	0.12427	0.04020	3.091	0.003901*
				Silt	0.15744	0.04185	3.762	0.000618*
				Clay	0.10606	0.03923	2.703	0.010521 *

*implies significant at 1 or 5% level of probability

Table 4: ANOVA Table for Models developed for Permanent Wilting Point

Model No.	Source of Variation	Degree of Freedom	Sum of Squares	Mean Squares	F-values	P-values
	Clay	1	0.07	0.068	0.0041	0.949126
	Sand	1	54.39	54.393	3.3130	0.077813
	Silt	1	203.66	203.665	12.4051	0.001276 *
1.	Bulk Density	1	5.20	5.199	0.3167	0.577419
	Total Porosity	1	0.74	0.735	0.0448	0.833730
	OM	1	31.22	31.221	1.9016	0.177171
	Residuals	33	541.79	16.418		
	Clay	1	1.53	1.525	0.1031	0.7502629
	Sand	1	68.93	68.930	4.6587	0.0385039 *
	Silt	1	209.31	209.309	14.1464	0.0006817 *
2.	Bulk Density	1	14.17	14.167	0.9575	0.3351522
	Total Porosity	1	3.06	3.064	0.2071	0.6521263
	OM	1	26.89	26.886	1.8171	0.1871241
	Residuals	33	473.47	14.796		
	Sand	1	22.68	22.682	1.5338	0.223791
	Silt	1	149.01	149.008	10.0761	0.003126 *
3.	Clay	1	108.08	108.075	7.3082	0.010521 *
	Residuals	35	517.59	14.788		

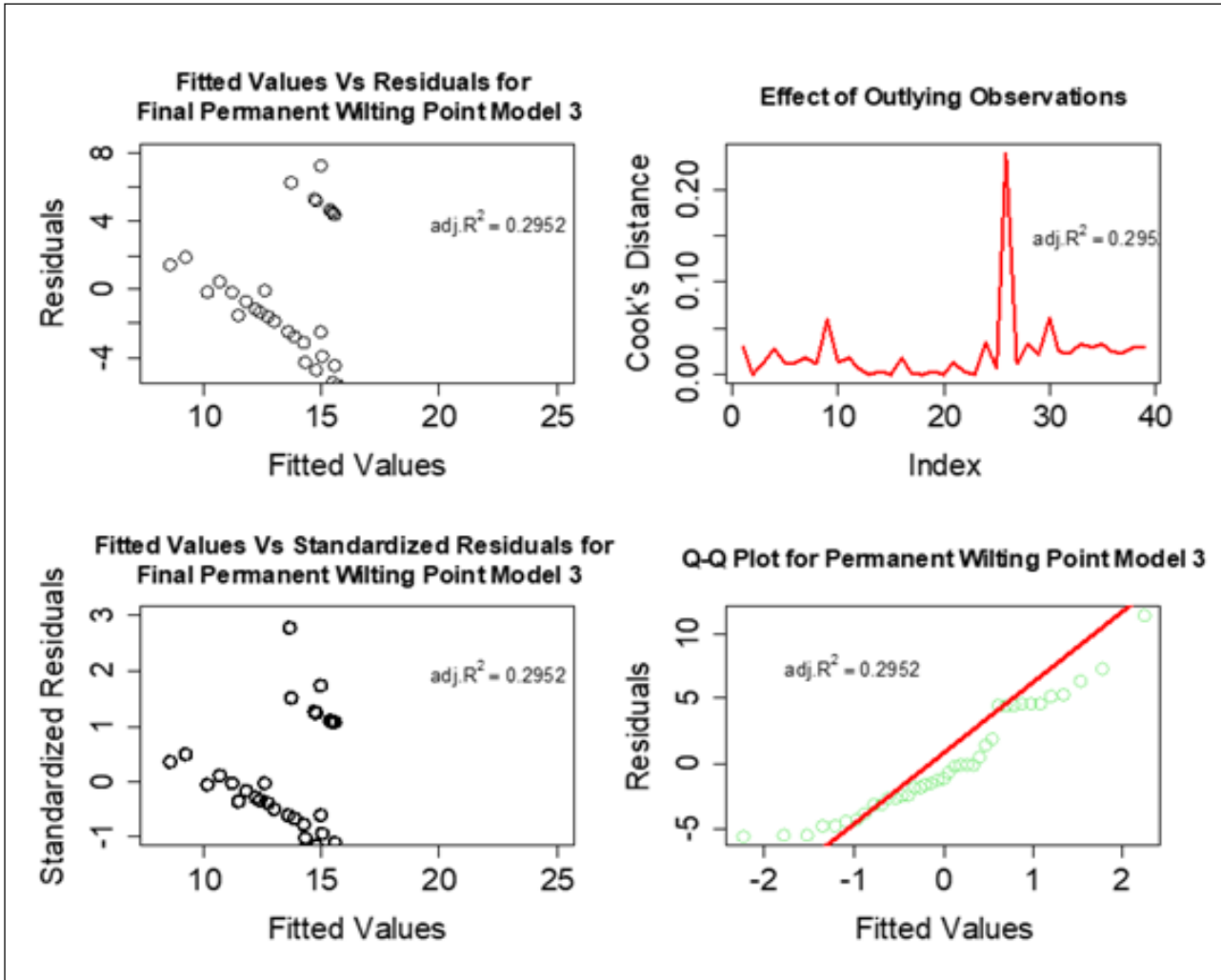*implies significant at 1 or 5% level of probability

Figure 1: Diagnostic Plots for Model 3 of Permanent Wilting Point

residuals is further confirmed by the histogram presented in Figure 2.

The regression equation created for model 3 of permanent wilting point was used in predicting the permanent wilting point values: observed, predicted and residuals of the prediction as revealed in Table 5. Data obtained for observed and predicted field capacity as presented in Table 5 were analysed for extent of relationship and presented a correlation coefficient of 0.5727. Based on the guideline outlined at http://www.westgard.com/lesson42.htm for assessing correlation coefficients, it is observed that when r falls within the range of 0.90 to 1.00, 0.70 and 0.89, 0.50 and 0.69, 0.30 to 0.49, and 0.00 to 0.29, they are said to show very high, high, moderate, low, and little if any correlation, respectively. It indicate that permanent wilting point predictions of model 3 of permanent wilting point have
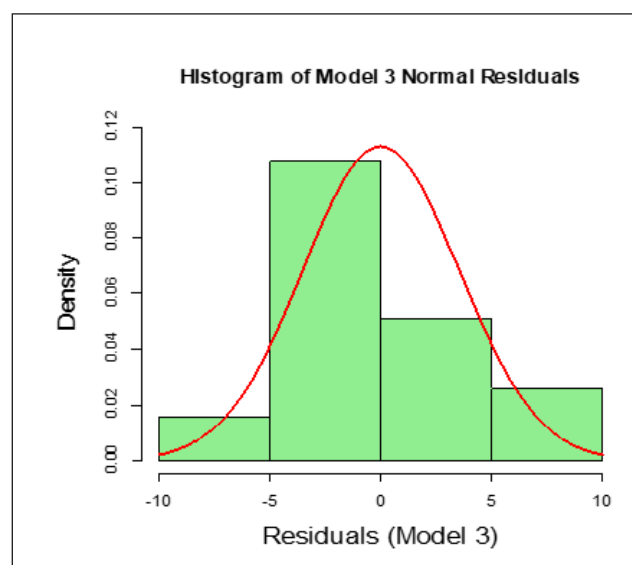

Figure 2: Histogram of Model 3 Normal Residuals

Table 5: Summary of Model 3 Prediction of Permanent Wilting Point

Observations	Observed Y (%)	Predicted Y (%)	Residuals (Y- Ŷ)
1.	20	14.85562	5.14438
2.	20	11.90282	8.09718
3	11.11	10.67099	0.43901
4	11.11	14.91809	-3.80809
5	20	14.78278	5.21722
6	11.11	12.90462	-1.79462
7	11.11	11.66634	-0.55634
9	11.11	13.24024	-2.13024
10	11.11	12.90462	-1.79462
11	22.22	14.29642	7.92358
12	11.11	11.66634	-0.55634
14	11.11	13.24024	-2.13024
15	12.5	14.29642	-1.79642
17	11.11	11.24334	-0.13334
18	10	11.56787	-1.56787
19	10	10.24107	-0.24107
20	11.11	9.24597	1.86403
21	11.11	12.56622	-1.45622
22	11.11	11.24334	-0.13334
23	11.11	12.84329	-1.73329
24	11.11	11.81569	-0.70569
25	10	8.49869	1.50131
26	11.11	11.24612	-0.13612
27	12.5	11.61424	0.88576
28	20	12.96962	7.03038
29	11.11	12.82394	-1.71394
30	25	25.00384	-0.00384
31	10	14.12722	-4.12722
33	20	15.30834	4.69166
34	10	15.88322	-5.88322
35	10	14.85562	-4.85562
36	20	15.45402	4.54598
37	10	15.30834	-5.30834
38	20	15.55544	4.44456
40	10	15.30834	-5.30834
42	10	14.85562	-4.85562
43	20	15.45402	4.54598
44	20	15.55544	4.44456
45	20	15.46439	4.53561

Note: The exclusion of observations 8, 13, 16, 32, 39 and 41 was due to the presence of potential

moderate correlation relationship with the observed permanent wilting point values. The graph of observed and predicted permanent wilting point values is presented in Figure 3. This shows the trend of observed and predicted values of PWP.

Variable importance based on the parameter estimates for PWP (Table 3) model showed that sand and silt were the most influential predictor variable for model 1. While, clay was the most influential predictor for model 2, followed by sand, and then clay. Although, model 3 also had the three soil separates as it influential predictor, silt ranked first, followed by sand and then clay. Sand content is the principal determinant of total porosity, hence, there was a negative relationship between sand and PWP. Although, both silt and SOM were of secondary importance as predictor variables of PWP. Silt was relatively more important than SOM for enhancing the prediction of PWP as indicated by a larger standard error (SE) and negative t-value (Table 3). This corroborates with the submission of Amsili et al. (2024) who opined that PWP were mostly defined by texture and SOM.

Silt was the most important variable for the prediction of Θ_{PWP}, followed by sand, and then clay (Model 3: , Table 3). Silt content importance in predicting PWP is logical because is the primary determinant of the proportion of total porosity with a pore diameter equal to less than 0.02 μm, which is the theoretical pore size that can hold water at -1500 kPa in the coarse textured soil studied with relatively low mean clay (166.53 g.kg⁻¹) content (Table 1). This finding was further ascertained by the weak positive (0.22) relationship between silt content and Θ_{PWP} (Table 2). SOM also had a positive influence on Θ_{PWP}, attributable to increase total porosity with a diameter less than 0.20 μm (Libohora et al., 2018) while sand had a negative effect on Θ_{PWP}, because increases result in a smaller proportion of total porosity with a diameter less than 0.20 μm. There was a strong negative relationship between clay and SOM (r=0.69*, p<0.05), which implied that clay content of the soil decreased with SOM and vice-versa. This findings is in agreement with previous research that found that SOM had a small impact on Θ_{PWP} in soil but counter to this submission as the soil in the present study had relatively low clay content (Minasny & Mcbratney, 2018; Saxton & Rawls, 2006). Hence, the statement that texture components- sand, silt and clay were sufficient for predicting PWP is not countenance in the present study evident from the relatively low coefficient of determination of 29.54 %. The low predictive power of this model could therefore be attributed
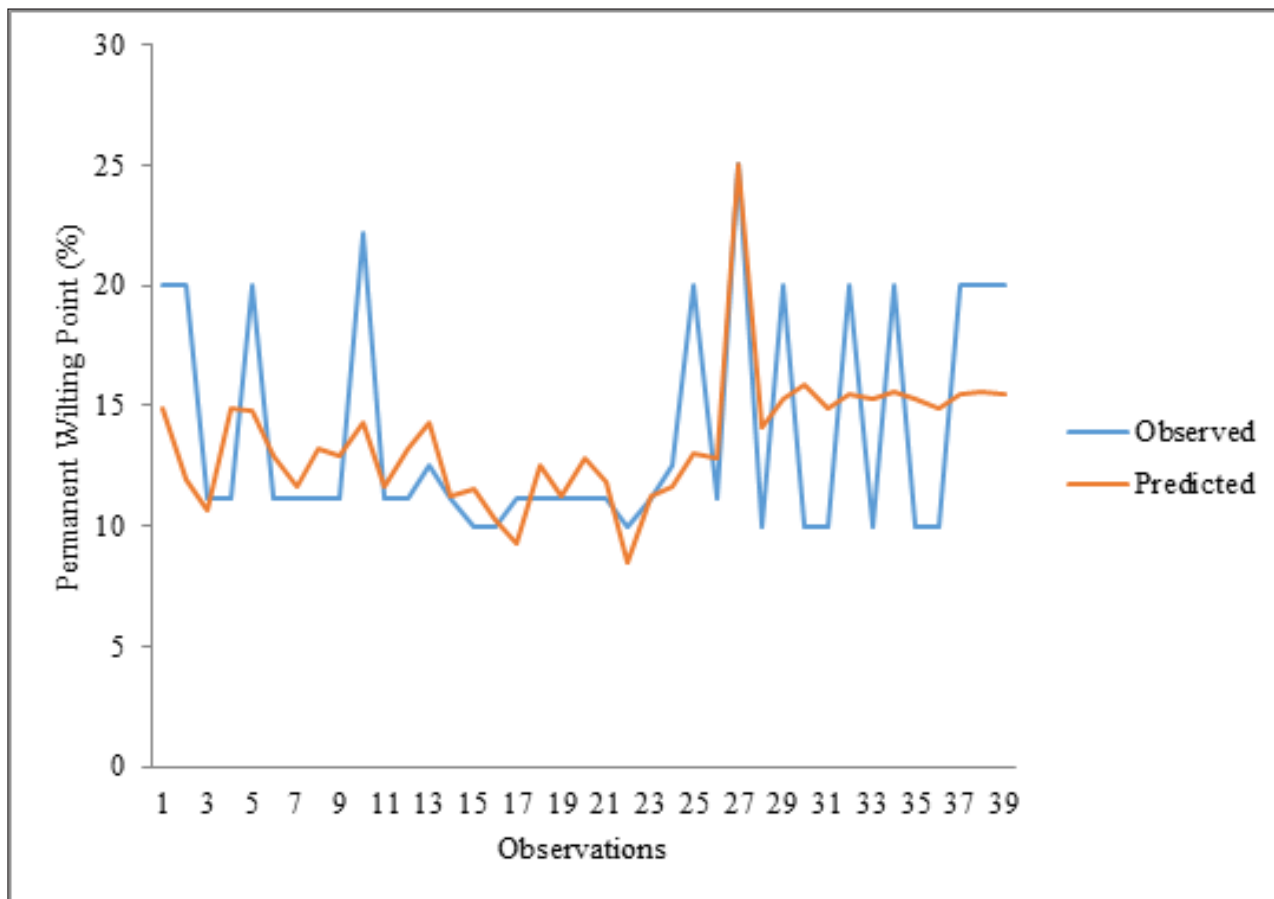
Figure 3: Graph of Observed versus Predicted Permanent Wilting Point

to the heterogeneity of soil and complexity of its properties such as soil texture, structure, organic matter and mineral composition coupled with the limited sample size along varying topographic positions which was probably not representative of the entire population, thus, leading to reduced model performance. Also, the relatively high correlation between predictor variable – silt - accounted for a reduced model's explanatory power.

However, the model multiple linear regression (MLR) R^2 value had 5.57 % higher than model adjusted R^2, which indicates inclusion of additional variables beyond sand, silt and clay would not provide meaningful improvements to Θ_{PWP}. Also, diagnostic plots for residuals, standardized residuals and Θ_{PWP}, gave the same adjusted R^2 value (0.30); indicating simple model approach, MLR, employed in the present study for estimating Θ_{PWP} was appropriate. Consequently, MLR would be adequate when 3 to 4 predictor variables are being used. This is in tandem with earlier submission by Amsili et al.

(2024) who also observed that the silt fraction was the most important variable for predicting Θ_{PWP}, owing to its particle size diameter range 2 μm -53 μm, corresponding most closely to the theoretical pore size range that prevent permanent wilting in plant.

Conclusion

The research conducted at the Teaching and Research Farm of the University of Ilorin for developing a model for predicting permanent wilting point (PWP) of the soil formulated three different models for estimating PWP of the soil. Model three () was selected as the optimal choice with the highest R^2 adjusted value of 0.2952. Prediction of permanent wilting point of soil of a representative *Alfisol* developed on basement complex at the Teaching and Research Farm of the University of Ilorin, Ilorin Kwara State will commendably be contingent on dependability of determination of the proportion of particle sizes

of the soil. Consequently, lacking the availability of soil moisture measuring instrument such as Richard's pressure chamber, the point assessment of soil moisture could be produced from easily quantifiable soil properties through employing generated PTFs.

The present study was carried out across three topographic positions whose changes along the sequence would inform variability in soil properties and characteristics. Thus, this variability resulted in systematic underestimation of the soil available water capacity which accounted for the relatively low coefficient of determination that explains just 29.52 % of the soil variance for PWP. Hence, the model is not viable, efficient or effective and should be re-evaluated with the present model serving as an insight for future studies in its current form. Accordingly, it is however advocated that further research be conducted to identify additional predictors, thereby improve the model explanatory power through use of more robust and representative data, consideration of multiple soil properties and characteristics across toposequence, explore non-linear relationships of predictor variables, check for measurement errors to improve measurement accuracy while also using techniques such as spatial and machine learning to enhance model accuracy and generalizability across different soil types and topographic positions, as well as undertake the comparison of the derived model with actual field permanent wilting point experiment to ascertain authentication of the predicted values with actual field values/data.

Conflict of Interest

The authors declare no conflict of interest on the outcome of this research

Funding declaration

Authors' contribution

KA: conceptualization, research design, supervision, literature review and final editing; JO and HA: review; NB: laboratory analysis, data collection and modeling, and draft preparation

ID ORCID and e-mails

Kehinde Olayemi AFFINNIH   affinnih.ko@unilorin.edu.ng

https://orcid.org/0000-0002-3535-5835

References

Amsili, J. P., Van Es, H. M., & Schindelbeck, R. R. (2022). An available water capacity pedotransfer using random forest -2020 Cornell soil health model: Harvard data verse https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/U5DAEP.

Amsili, J. P., Van Es, H. M. & Schindelbeck, R. R. (2024). Pedotransfer functions for field capacity, permanent wilting point and available water capacity based on random forest models for routine soil health analysis. *Communication in soil science and plant analysis, 55*(13)*,* 1–18. https://doi.org/10.1080/00103624.2024.2336573.

Barnett, V., & Lewis T. (1994). *Outliers in Statistical Data*. 3rd edition. Wiley & Sons, New York, XVII, 582p.

Blake, G. R., & Hartge, K. H. (1986). "Bulk density", In: A. Klute (Ed.), *Methods of Soil Analysis. Physical and Mineralogical Methods, part 1*. (pp. 363–375) 2nd edition.

Bouma, J. (1989). Using Soil Survey Data for Quantitative Land Evaluation. In: Stewart, B. A. (eds.), *Advances in Soil Science, 9*. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-3532-3_4.

Brady C. N., & Weil, R. R. (1999). "The nature and properties of soil", 12th edition, Prentice Hall Publishers, London.

Donatelli, M., Acutis M., & Laruccia N. (1996). Pedotransfer functions: Evaluation of methods to estimate soil water content at field capacity and wilting point.

Gee, G.W, & Or, D. (2002). Particle size analysis.

In: Dane, J. H., Toop, G. C. (Eds.), *Methods of Soil Analysis Part 4*, (pp 255–293). SSSA. Book Series No.5. Madison, Wi, USA. Soil Sci. Society of Am.

John, J. A., & P. Prescott (1975). Critical values of a test to detect outliers in factorial experiments. *Journal of the Royal Statistical Society Series C: Applied Statistics, 24*(1), 56–59.

Kinoshita, R., Moebius-Clane, B. N., Van Es, H. M., Hively, W. D., & Bilgillis, A. V. (2012). Strategies for soil quality assessment using visible and near-infrared reflectance spectroscopy in a western Kenya chronosequence. *Soil Science Society of America Journal, 76*(5), 1776–1788. https://doi.org/10.2136/sssaj2011.0307.

Kolo, E., Takim, F. O., & Fadayomi, O. (2012). Influence of planting date and weed Management practice on weed emergence, growth and yield of maize (*Zea mays* L.) in Southern Guinea Savanna of Nigeria. *African crop science journal, 9(*4), 615–627.

Kukal, M. S., Irmak, S., Dobos, R., & Gupta, S. (2023). Atmospheric dryness impacts on crop yields are buffered in soils in the higher available water capacity. *Geoderma 429*, 116270. https://doi.org/10.1016/j.geoderma.2022.116270.

Libohora, Z., Seybold, C., WWysocki, D., Wills, S., Schoeneberger, P., Williams, C., Lindbo, D. Stott, D. & Owens, P. R. (2018). Reevaluating the effects of soil organic matter and other properties on available water-holding capacity using the National Cooperative Soil Survey Characterization Database. *Journal of Soil and Water Conservation, 73*(4), 411–421 https://doi.org/10.2489/jswc.73.4.411.

Martens, H., & Naes, T. (1984). Multivariate Calibration. In: Kowalski, B.R. (eds.), *Chemometrics*. NATO ASI Series. Springer. Dordrecht. https://doi.org/10.1007/978-94-017-1026-8_5.

Mbagwu, J., Lal, R., & Scott, T. W. (1983). Physical properties of three soils in southern Nigeria. *Soil Science*, *136*(1), 48–55.

Minasny, B., Mcbratney, A. B., & Bristow, K. L. (1999). Comparison of different approaches to the development of pedotransfer functions for water-retention curves. *Geoderma, 93*(3–4), 225–253. https://doi.org/10.1016/S0016-7061(99)00061-0.

Minasny, B., & Mcbratney, A. B. (2018). Limited effect of organic matter on soil available water capacity. *European journal of soil science, 69*(1), 39–47. https://doi.org/10.1111/ejss.12475.

Morrison, D. F. (1976) *Multivariate Statistical Methods*. New York, McGraw Hill.

Myers, R. H., & D. C. Montgomery (1995). *Response Surface Methodology: Process and Product Optimization using Designed Experiments*. John Willey and Sons, New York INC.

Nelson, D. W., & Sommers, L. E. (1982). Total carbon, organic carbon and organic matter. In: Page, A. L., Miller, R. H., and Keeney, D. R. (eds) *Methods of Soil Analysis: Chemical and Microbiological Properties. Part 2,* 2nd edition. Madison, WI: ASA-SSSA.

Obi, C. I., Obi, J. C., & Onweremadu, E. U. (2012). *Modeling of Permanent Wilting from Particle Size Fractions of Coastal Plain Sands Soils in Southeastern Nigeria.* International Scholarly Research Notices, 2012, 1-5.

Odu, C. T. I., Babalola, O., Udo, E. J., Ogunkunle, A. O., Bakare, T. A., & Adeoye, G. O. (1986). *Laboratory Manual for Agronomic Studies in Soil, Plant and Microbiology*. University of Ibadan.

Ogban, P. I., & Ekerette, I. O. (2001). Physical and chemical properties of the coastal plain sands soils of southeastern Nigeria. *Nigerian Journal of Soil and Environmental Research*, *2*, 6–14.

Olaniyan, J. O. (2003). *An Evaluation of the Soil Map of Nigeria for Land use Planning in Kwara State* [Doctoral dissertation, University of Arizona]. University of Ibadan, Nigeria.

Olorede, K. O., Mohammed, I. A., & Adeleke, B. L. (2013). Economic Selection of Efficient Level of NPK 16:16:16 Fertilizer for Improved Yield Performance of a Maize Variety in the South Guinea Savannah Zone of Nigeria. *Mathematical Theory and Modeling, 3*(1), 27–39.

Olorede, K. O., & Mudasiru, A. A. (2013). Economic Analysis and Modelling of Effects of NPK

Fertilizer Levels on Yield of Yam. *Mathematical Theory and Modeling, 3*(1), 108–118.

Phillips, J. D. (2007). Development of texture contrast soils by a combination of bioturbation and translocation. *Catena, 70*(1), 92–104. https://doi.org/10.1016/j.catena.2006.08.002

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Saxton, K. E., & Rawls, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil science society of America journal, 70*(5), 1569–78. https://doi.org/10/2136/ssaj2005.0117

Souza, Z. M., Junior, J. M., & Pereira, G. T. (2009). Spatial variability of the physical and mineralogical properties of the soil from the areas with variation in landscape shapes. *Brazilian Archives of Biology and Technology, 52*(2). 305–316.

Stefansky, W. (1972). Rejecting Outliers in Factorial Designs. *Technometrics, 14*(2), 469–479.

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minanasy, B., Mishra, U., Nemes, A.,

Pachepsky, Y. A., Padarian, J., & Vereecken, H. (2017). Pedotransfer functions in earth system science: challenges and perspectives. *Reviews of geophysics, 55*(4), 1199–256. https://doi.org/1.1002/207RG000581

Vereecken, H., Maes, J., Feyen, J., & Darius, P. (1989). Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content. *Soil Sci., 148*(6), 389–403. https://doi.org/10.1097/00010694-198912000-00001

Wosten, J. H. M., Lilly, A., Nemes, A., & Le Bas, C. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, *90,* 169–185. https://doi.org/10.1016/S0016-7061(98)00132-3