Division – Soil Processes and Properties | Commission – Soil Chemistry

Pedotransfer functions: the role of soil chemical properties units conversion for soil classification

Fernanda Reis Cordeiro[(1)]* (iD)**, Fernando Vieira Cesário**[(2)] (iD)**, Ademir Fontana**[(3)] (iD)**, Lúcia Helena Cunha dos Anjos**[(1)] (iD)**, Ana Carolina Barbosa do Canto**[(2)] (iD) **and Wenceslau Geraldes Teixeira**[(2)] (iD)

[(1)] Universidade Federal Rural do Rio de Janeiro, Departamento de Solos, Seropédica, Rio de Janeiro, Brasil.
[(2)] Universidade Federal Fluminense, Departamento de Engenharia Agrícola e Meio Ambiente, Niterói, Rio de Janeiro, Brasil.
[(3)] Empresa Brasileira de Pesquisa Agropecuária, Embrapa Solos, Jardim Botânico, Rio de Janeiro, Rio de Janeiro, Brasil.

ABSTRACT: Chemical soil analysis data can be expressed by weight (i.e., gravimetric basis) or volume (i.e., volumetric basis) of the fine earth (sieved ≥2 mm), resulting in different units, $cmol_c\ kg^{-1}$ and $cmol_c\ dm^{-3}$, respectively. The research problem is that the difference between methods to express the same soil properties hinders the comparison of results and database or dataset standardization. This paper aims to develop pedotransfer functions (PTF) to obtain the density of fine earth, which will then be used for conversion data expressed in volumetric to gravimetric basis, or vice versa, that will be applied to compare results and to standardize databases with different units. Soils samples, including profiles of the main soil orders in Brazil such as *Latossolos* (Ferralsols or Oxisols) and *Argissolos* (Acrisols or Ultisols), from the states of Rondônia, Roraima, and Mato Grosso do Sul (132 horizons) were selected and weighed (in triplicate) to obtain the fine earth mass contained in a volume of 10 cm^3. The mass values were used to calculate the fine earth density. Spearman's correlation analysis was used between the density and nine soil properties (coarse sand, fine sand, total sand, silt, clay, clay dispersed in water, clay dispersion, particle density, and organic carbon). The total sand, clay, and organic carbon showed the best correlations, therefore they were selected to construct the pedotransfer functions. Nonlinear regression techniques were used to obtain the models (PTFs) to predict density, which was used for unit conversion. As a result, the residual standard error (RSE) statistics of the models were: 0.0920, 0.1231, and 0.1633 g cm^{-3}, respectively for PTF1 (using total sand as a predictor), PTF2 (using clay), and PTF3 (using organic carbon). Independent data was used to evaluate the accuracy of the models by residue analysis and the RSE. For the validation, the lowest RSE obtained was from the PTF1, so the best performance. Thus, to convert values of the chemical properties from a volumetric to gravimetric basis, the value must be divided by the predicted density. While, the conversion from gravimetric to volumetric basis requires that the value be multiplied by the predicted density. The PTFs using the properties total sand, clay, and organic carbon as predictor variables, allowed conversion of analytical data of soil samples expressed in the volumetric basis to gravimetric and vice versa, which can be used for dataset or database standardization.

Keywords: data standardization, soil analysis, nonlinear regression.

INTRODUCTION

The majority of soil chemical and physical analyses start with the preparation of samples in the laboratory, through drying and maceration, followed by sieving (≥2 mm) to obtain the called fine earth (Teixeira et al., 2017). To carry out the chemical analysis, a certain weight or volume of fine earth is taken to be analyzed. In general, there are two ways to obtain the sample: (i) using the fine earth mass (i.e., 10 g), thus expressing results in a gravimetric basis, that is, $cmol_c$ kg^{-1}, g kg^{-1} or mg kg^{-1} (Teixeira et al., 2017); and (ii) using the fine earth volume (i.e., 10 cm^3) and expressing the results in a volumetric basis, that is, $cmol_c$ dm^{-3} or mg dm^{-3} (Silva et al., 2009).

The most common procedure to evaluate soil fertility in Brazil and countries such as the U.S.A., is to sample volumetrically the fine earth by using a device that consists of a small cylinder with a 10 cm^3 volume (Silva et al., 2009; Soil Survey Staff, 2014a). This device gives convenience and speed in the process of sample preparation for chemical analysis. The adoption of this method began in Brazil and Latin America since 1966 with the support of North Carolina State University (USA) to evaluate more samples in less time and return fertility recommendations to farmers faster (Soil Science Department, 1966). Therefore, the manuals for recommendations of lime and fertilizers adopted in Brazil express the chemical properties on a volumetric basis (e.g., Freire et al., 2013; Prezotti et al., 2013; CQFS-RS/SC, 2016). However, in soil survey analyzes for pedological purposes, the results are commonly expressed on a gravimetric basis. Details about the chemical properties for soil classification and chemical analyses units are presented in the publications IUSS Working Group (2015), Santos et al. (2018), Soil Survey Staff (2014b), Kellogg Soil Survey Laboratory Methods Manual (Soil Survey Staff, 2014a), and the Manual for soil and water analysis (Buurman et al., 1996).

These two soil sampling procedures, which result in data expressed with different units, make difficult data comparisons and standardization of soil dataset or database. Thus, the conversion of the data analyzed on a mass basis to the volumetric basis, or vice versa, is necessary to evaluate data with more accuracy. The variation of results expressed by mass or volume depends of the fine earth fraction density (Mehlich, 1972), which is related to the granulometric composition and organic carbon in the sample, as pointed by Stewart et al. (1970), Qiao et al. (2019), and Patton et al. (2019).

The problems of reporting the analytical soil results in different units was discussed by Mehlich (1972), as well as the differences when the results were expressed as soil mass or volume. The author reports the importance of the volume weight (i.e., the fine earth density) for the conversion of units obtained on a volumetric or gravimetric basis and emphasizes that volumetric base data will only be equal to the gravimetric data if the fine earth density from the soil sample is precisely 1.0 g dm^{-3}. Further, the multiplication of the gravimetric data by the fine earth density converts the data to the volumetric unit. For soils that do not have 1.0 g dm^{-3} density, it is necessary to obtain the fine earth density to convert the units.

An option for conversion of these units is the use of pedotransfer functions (PTF), wich involves the application of statistical models. Soil data properties, which are difficult, expensive, and laborious to be obtained, can be predicted by PTF using other easily accessible and economically affordable soil properties (McBratney et al., 2002).

The PTF has been widely used around the world. Qiao et al. (2019) used PTF to estimate soil density at the Loess Plateau in China, formed from deep deposits of sediments. The authors used organic carbon, texture, and soil depth as explanatory variables for the prediction of soil densities. Ottoni et al. (2018) developed PTF to estimate soil hydraulic conductivity for a database of Brazilian tropical soils and European temperate soils, using as predictor variables soil texture and effective porosity. Dobarco et al. (2019) developed PTF to estimate available soil water from a French soil dataset, with contents of sand,

clay, organic carbon, and bulk density as predictor variables. Patton et al. (2019) used PTF to estimate the bulk density of the thin fraction of soils from the Reynolds Creek Critical Zone Observatory, USDA Agricultural Research Service site. The percentage of organic carbon measured in soils derived from felsic and mafic lithologies and the particle size distribution were used as predictor variables. The authors point to the importance of incorporating these predictor variables to provide a reliable fine fraction density and, consequently, to estimate soil carbon stocks.

Studies using PTFs in Brazilian soils show consistent results for several soil properties. Nascimento et al. (2015) developed PTF using different variables such as sand, silt, clay, soil density, and organic carbon to estimate water content at 33 and 1500 kPa IN Xanthic Oxisols and Ultisols from a coastal tableland landscape database, from different locations in Brazil. Other studies used soil predictors, such as granulometry, organic carbon content, soil density, texture, and moisture, to estimate the water retention curve in different Brazilian soils (van Den Berg et al., 1997; Gaiser et al., 2000). Beutler et al. (2017) used organic carbon and clay content in Histosols and in other soils with high organic matter content, to predict the soil bulk density obtained in non-deformed soil samples. Likewise, Benites et al. (2007) also predicted soil bulk density by using organic carbon, clay, total nitrogen, and the sum of bases for Brazilian soils in general.

Taking into account that the access to soil samples already analyzed by fine earth mass or fine earth volume may be impracticable, the use of PTF helps to obtain the fine earth density and then to convert the units. To date, there are no studies proposing PTF for the prediction of the fine earth density to convert units expressed in the volumetric to gravimetric basis, or vice versa. Therefore, this paper aims to propose PTFs to obtain the fine earth density (hereinafter called just density) of mineral soils, which will be used for conversion of soil chemical properties expressed in volumetric basis to gravimetric basis, or vice versa.

MATERIALS AND METHODS

Soil samples selection and measurement of fine earth density (density)

Profiles representing soil orders from Brazil, including the two most common (*Latossolos and Argissolos*), were selected from recent Brazilian Soil Classification and Correlation Meeting studies (Cardoso et al., 2012; Batista et al., 2018; Lumbreras et al., 2019), which occurred in Rondônia, Mato Grosso do Sul, and Roraima States (Figure 1). The sample selection considered an amplitude of the parameters that, according to the literature, are related to density, which are granulometry and carbon content. The soils selected (Table 1) were classified as *Latossolos* (Oxisols), *Argissolos* (Ultisols), *Plintossolos* (Plinthic Subgroups), *Neossolos* (Entisols), *Espodossolos* (Spodosols), *Gleissolos* (Entisols with aquic properties), and *Vertissolos* (Vertisols), according to Brazilian Soil Classification System (SiBCS) (Santos et al., 2018) and the Soil Taxonomy (Soil Survey Staff, 2014b), respectively.

According to the Köppen classification system, the climate conditions are tropical Aw for all profiles located at Rondônia and Mato Grosso do Sul; and tropical Af and Am in the profiles from Roraima. Samples of the surface and subsurface horizons were selected to comprehend the majority of diagnostic horizons defined in the SiBCS. Hence, they were included 3 profiles of Rondônia, 16 profiles of Mato Grosso do Sul, and 16 profiles of Roraima, totaling 132 samples. The organic horizons (O and H) were not included due to their particularities regarding the predictive properties (e.g., sand, silt, and clay) and the high levels of organic matter.

To measure the density, the soil sample mass contained in a 10 cm^3 device was weighted (precision scale ± 0.001 g) using triplicates. Then, the density was calculated
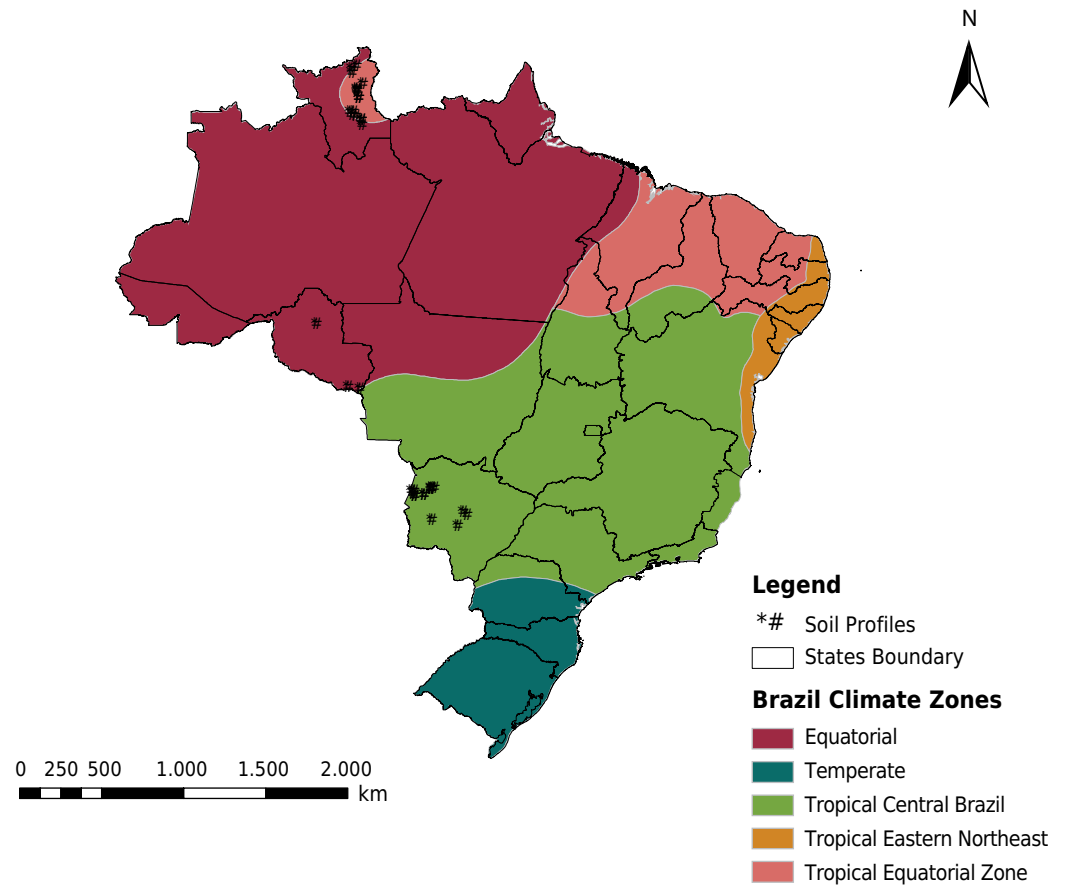
Figure 1. Soil map profiles of the data set and the Brazil climate zones.

from the corresponding mass values of each soil sample and the volume of the volumetric device.

Selection of the prediction variables

The following properties were selected from available analyses of the soil samples: coarse sand, fine sand, total sand, silt, clay, water dispersed clay, clay dispersion, particle density, and organic carbon. The analyzes of these properties were done in the laboratory following the methods described in Donagemma et al. (2011). Briefly, bulk soil samples were ground and passed through a 2 mm mesh sieve, to prepare fine earth samples, to be used for the following physical analyses: granulometry (with sodium hexametaphosphate as a dispersant for total clay) and clay dispersed in water, both using the densimeter method, and sieves for fractioning sand particle classes, with the silt obtained by difference. The particle density was obtained using the volumetric balloon, with alcohol, method. Total organic carbon was determined by the wet combustion method.

The properties chosen are associated with the particle size (granulometry) and the amount of soil organic matter, which are the main components of the solid soil phase. Additionally, the particle density reflects partially the mineralogy of the sample, since it is related to the amount of iron oxides, silicate clays, and quartz in the soil, as well as the amount of soil organic matter. Thus, they correspond to the possible predictive variables of density and are also properties routinely measured in the laboratories for soil analyses. Boxplots (Figure 2) were used to graphically summarize the descriptive statistics of the properties for density prediction.

Spearman correlation (Spearman, 1904) was used to support our prior pedological knowledge to identify which properties were mostly associated with density, and

Table 1. Soil profiles number and parent material of each data reference source

Source	Profile	Parent material	Source	Profile	Parent material
RCC Rondônia (2017)	08	Sandy clayey sediments		01	Sandy clayey sediments
	12	Crystalline rocks of Jamari Group		02	Sandy clayey sediments
	15	Sandy clayey alluvial sediments		04	Sandy sediments
RCC Roraima (2018)	01	Granite		05	Sandy sediments
	02	Ferruginous sandstone		06	Sandy sediments
	03	Ferruginous sandstone		07	Limestone
	04	Ferruginous sandstone		08	Colluvium from granite decomposition
	05	Sandy sediments		09	Sandy clayey sediments
	06	Volcanic acid and intermediate rocks	RCC Mato Grosso do Sul (2012)	10	Limestone
	07	Acid and intermediate rocks		11	Siltstones and limestones
	08	Sandy clayey sediments		12	Sandy sediments
	09	Sandy clayey sediments		13	Sandy clayey sediments
	10	Basaltic and intermediate rocks		15	Limestone
	11	Granite		16	Basalts
	12	Colluvial-alluvial, sandy clayey sediments		17	Sandstones
	13	Alluvial sediment of sandy texture		18	Basalts
	14	Acid and Intermediate rocks			
	15	Sandy sediments			
	16	Clayey alluvial sediments			

then the correlogram was developed (Figure 3). The correlation values are presented at the intersection between rows and columns with the variables, ranging from -1 to 1. The circle and its diameter are proportional to the correlation value. The blank intersections indicate that the correlation was not significant ($p>0.05$), while colored intersections (i.e., red or blue) indicate that there was a significant correlation between the respective variables.

Based on the correlogram (Figure 3), the properties with the highest values of significant correlations (negative or positive) with density was selected as predictors, to obtain the pedotransfer functions: total sand (TS) with a correlation of 0.85, clay with -0.76, and organic carbon (OC) with -0.55. Since the original structure and porosity of the soil samples is modified in the preparation process, the most prominent factors influencing the density were the total sand, clay, and organic carbon contents. The relation of these predictors with the fine fraction density were also discussed by Stewart et al. (1970), Qiao et al. (2019), and Patton et al. (2019), and it is corroborated by the results obtained in the correlation.

Pedotransfer function models

To construct the pedotransfer functions, firstly three linear models were fitted using TS, clay, and OC as explanatory variables. The three models were evaluated using a fitted vs. residual plot. The fitted vs. residual plot helps to identify if the relationship between the variables is linear (i.e., linearity) (Ritz and Streibig, 2008). All models exhibit a clear u- shaped pattern (see supplementary materials – S1), and therefore nonlinear exponential models were performed.
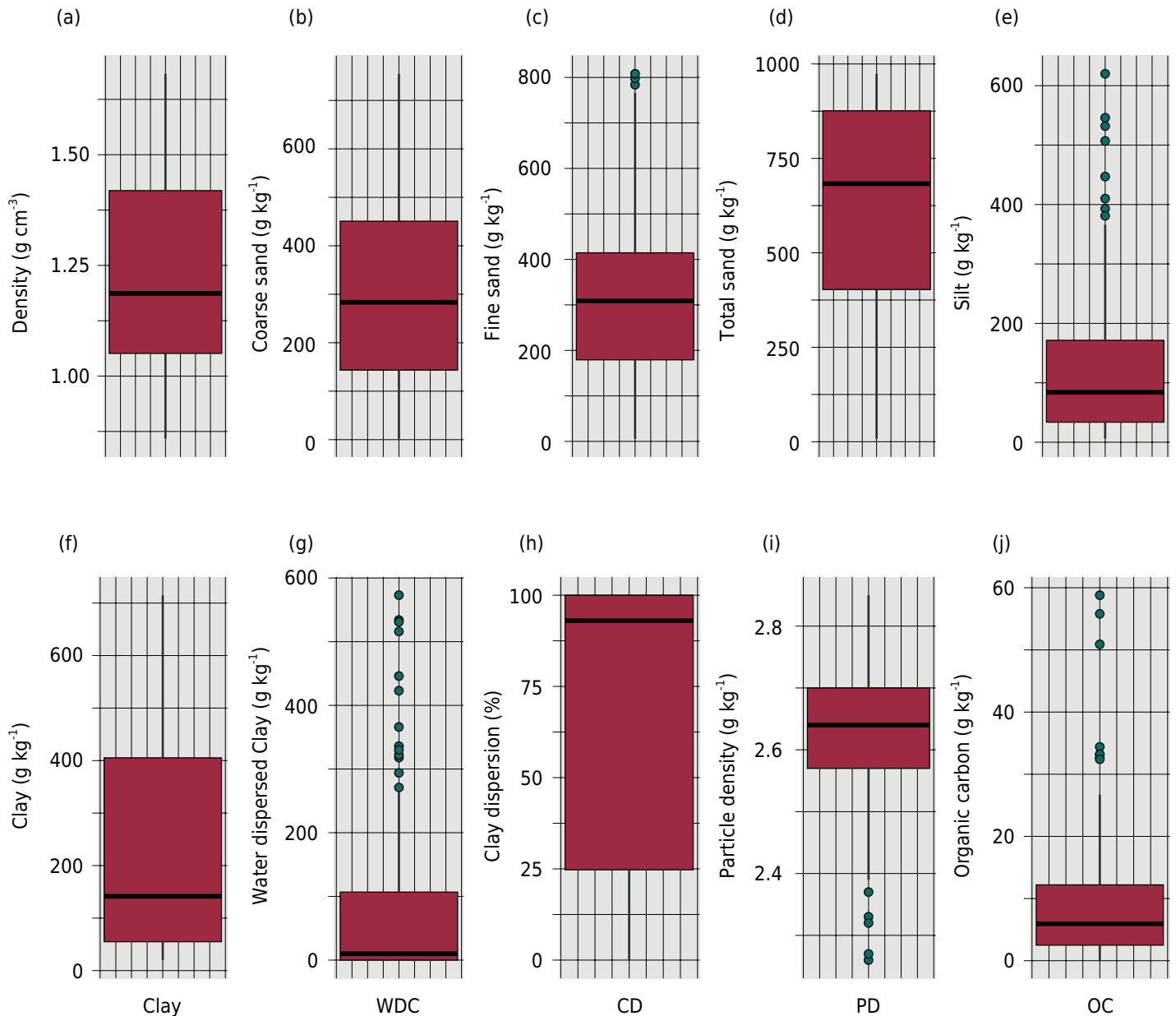
Figure 2. Soil properties boxplots (*n* = 132) of different soil classes of Brazil. PD: particle density; OC: organic carbon.

We look for equations that could be associated with the pedological/environmental field of study. Therefore we found the role of reciprocal yield-density functions related to the fundamental relationship between crop yield and plant population (Farazdaghi and Harris, 1968) and also the plant growth models equations (Paine et al., 2012). All these equations resemble the phenomena modeling by soil variables in relation to density. Among the four equations (e.g., Farazdaghi equations, rational model equations, reciprocal quadratic and exponential models) the exponential models (Equation 1) were selected to represent the increase or decrease in density as a function of total sand, clay, or organic carbon.

$$y = \beta_0 + \beta_1 \, exp \, (\beta_2 x) \hspace{6cm} \text{Eq. 1}$$

The equation has three parameters for describing the curve: β_0 the intercept, β_1 the relative increase or decrease rate, and β_2 the exponential rate. The exponential model (Equation 1) was fitted using the TS, Clay, and OC as explanatory variables, totalizing three PTFs. The parameters for each PTFs were obtained using previous information
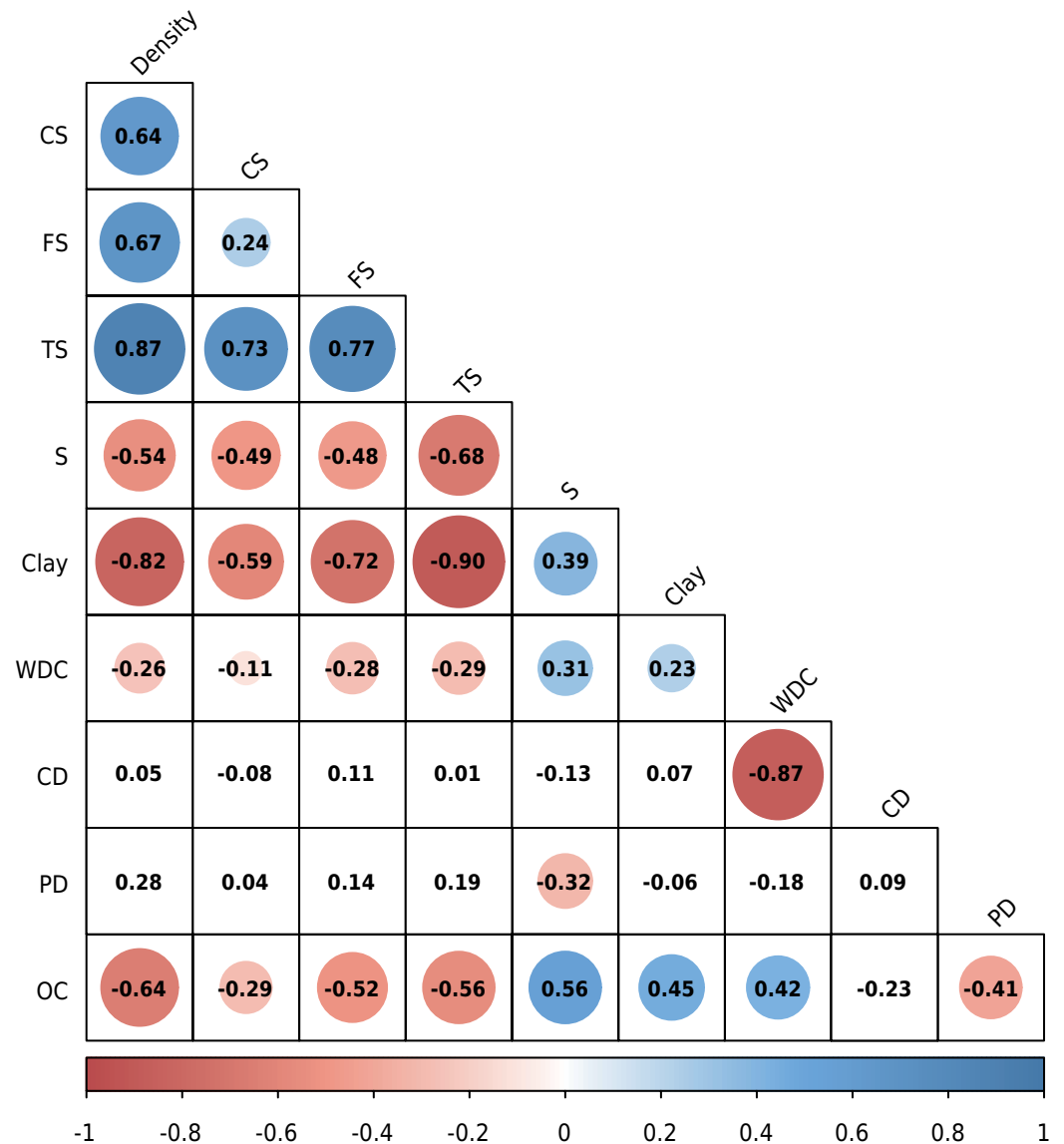
Figure 3. Correlogram between density and the nine possible predictive variables for the development of pedotransfer functions. Density: density of fine earth; CS: coarse sand; FS: fine sand; TS: total sand; S: silt; WDC: water dispersed clay; CD: clay dispersion (%); PD: particle density; and OC: organic carbon.

to allow an adequate guess of starting values that were evaluated graphically to get plausible candidate model parameter values. For this purpose, nlstools provides a graphical function called preview, which can be used to assess the suitability of the chosen starting values, before fitting the model. This same approach was used by Ritz and Streibig (2008). The graphical examination was used jointly with the residual sum of squares result. Additionally, self-start functions were used to support the previous choice, except for β_0. The choice of good starting values is essential here because of the novelty in development PTFs for a predictive model. Following, the three PTFs were fitted (i.e., for TS, Clay, and OC), and the model's residual standard error (RSE) statistics were obtained; after that, the RSE (Equation 2) was used to compare the exponential models and to select the best (Dalgaard, 2002). The RSE was calculated by the square root of the residual sum of squares divided by the degrees of freedom, and the lowest values indicate a better fit.

$$RSE = \sqrt{\frac{1}{d.f.} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}$$

Eq. 2

in which the $\hat{y}_i$ are the predicted values, y_i are the mean values, and *d.f.* is the degrees of freedom.

The standardized residuals in a normal Q-Q plot and the plot of the fitted values versus the standardized residuals were used to evaluate the normality of the residuals of the nonlinear models. The Normal QQ Plot compares the distribution of standardized residuals and the standard normal distribution. It is expected to obtain a straight line by intercepting the y-axis at zero and slope of 1 as the inference of the data normality satisfied (Dalgaard, 2002). All analyses were performed using the software R (R Core Team, 2018).

Models validation

The evaluation of the accuracy of the models, PTF validation was performed using data of exchangeable potassium (K^+), available in both mg dm^{-3} and mg kg^{-1}, of 88 samples taken from 0.00-0.05 and 0.10-0.15 m layers from a *Latossolo Amarelo* (Xanthic Oxisols), from the National Forest of *Tapajós*, Pará State - Brazil (Cesário, 2018). The validation was assessed through predicted vs. measured graphs and the residual standard error (RSE) for each PTF.

RESULTS

The non-linear models between TS and density show an increasing exponential pattern (Figure 4a), indicating that enhancing sand content increases density. On the other hand, the Clay and OC non-linear models show a decreasing exponential pattern, indicating that decreasing values of Clay or OC (Figures 4b and 4c) increase density.

Three different models were developed to offer options for conversion since not all data (TS, Clay or OC) might be available. The models were: i) PTF1, which uses TS as the predictor variable; ii) PTF2, using clay as a predictor; and iii) PTF3, with OC as a predictor (Table 2). The residual standard error (RSE) was calculated for each PTF. The PTF1 had the lowest RSE, 0.092 g cm^{-3}, followed by PTF2 and PTF3 with 0.1231 and 0.1633 g cm^{-3}, respectively. However, the dataset used for the development of PTF2 has some low clay contents, which can lead to high-density predictions, so it is recommended to use it with caution. The sequence for RSE is PTF1 < PTF2 < PTF3. This result shows the importance of TS in the density prediction, considering that RSE of PTF1 was 56 % smaller than that of PTF3.

The residuals of the three PTFs were analyzed graphically, and figure 4 shows the residuals graphs (e.g., fitted values vs. standardized residuals) and the Q-Q Plot for the
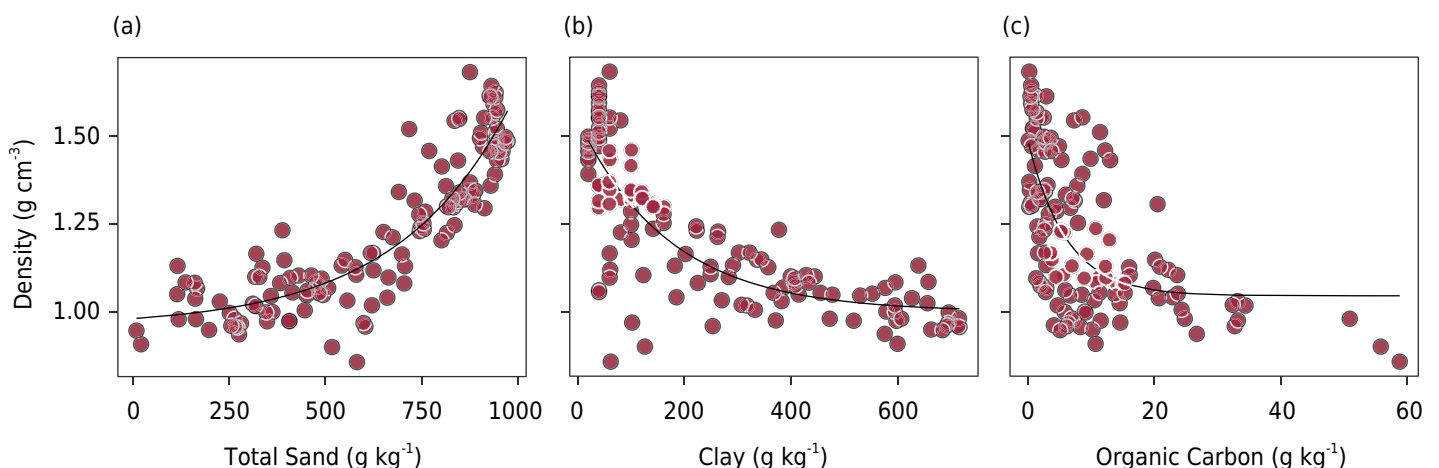

Figure 4. Scatterplot of total sand vs. density (a); clay vs. density (b); and organic carbon vs. density (c).

Table 2. Pedotransfer functions developed to predict the fine earth density and the respective model indexes. Density: fine earth density (g cm^{-3}); TS: total sand (g kg^{-1}); OC: organic carbon (g kg^{-1}); RSE: residual standard error; SE: standard error of parameters; β_0, β_1, β_2: parameters of the function. All PTF show the same degrees of freedom (129)

	PTF1			PTF2			PTF3		
Equation	Density = 0.9594781 + (0.0229629 * exp (0.0033731*TS))			Density = 1.000383 + (0.545928 * exp (-0.005821 * C))			Density = 1.04583 + (0.43990 * exp (-0.16439 * OC))		
RSE	0.09227			0.1231			0.1633		
Parameters[1]	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
SE	0.0329	0.0123	0.0005	0.0304	0.0340	0.0012	0.0382	0.0462	0.0460

[1] Parameters according to the equation 1.

Table 3. Descriptive statistics of residual values of PTFs

Functions	Mean	SD	Median	Max	Min	Kurtosis	SE
PTF1	0.00	0.09	-0.01	0.30	-0.26	0.89	0.01
PTF2	0.00	0.12	0.01	0.30	-0.52	3.19	0.01
PTF3	0.00	0.16	-0.01	0.40	-0.31	-0.50	0.01

SD: standard deviation; Max: maximum; Min: minimum; and SE: standard error.

three PTFs. For PTF1 and PTF3, no specific patterns were observed in the distribution of points, which confirms the assumption of normality. These results are also corroborated by the Normal Q-Q Plot of the same PTFs. In general, the analysis of PTF1 residues was slightly better than PTF3. Additionally, PTF2 presented a deviation from normality in the Q-Q plot, and the residue graph showed a trend in data distribution with results lower than PTF1 and PTF3. This indicates the lower accuracy of PTF2 in predicting density compared to other PTFs.

The descriptive statistics of residuals of each function (i.e., PTF1, PTF2, and PTF3) are presented in table 3. The maximum, minimum, and mean values (tending to zero), as well as the low value of the standard deviation, show that the accuracies of the models are satisfactory. The PTF1 and PTF3 show low kurtosis values indicating the normal distribution of the residues. The PTF2 had high kurtosis values and non-normal distribution of the residues, which was corroborated by the Q-Q graph for the same function (Figure 4).

Models validation

The three PTFs were applied to independent data, samples from the *Latossolo Amarelo* (Oxisols), used for validation of the model for density prediction. The descriptive analysis of the TS, clay, and OC values are shown in table 4.

Each PTF generated the predicted values of density. Exchangeable potassium (K$^+$) data of National Forest of *Tapajós* samples show a range between 12 to 51 mg dm^{-3} and a median of 27 mg dm^{-3}. The conversion of the unit to mg kg^{-1} was calculated by the ratio of the volumetric data (mg dm^{-3}) and the predicted density of each PTFs. The RSE values for PTF1, PTF2, and PF3, were respectively: 2.74, 2.78, and 2.79 mg kg^{-1}.

From the amplitude of the original data observed and the RSE obtained, PTF1 is the model that presents the best fit, corroborated by the lower values of RSE, followed by PTF2 and PTF3; although these two PTFs also presented low error values and can be used for conversion. Likewise, the graphs of the predicted vs. measured values (Figure 6) show a linear distribution of the values similar for all PTFs and high values of R^2, corroborating the good application of the functions for data conversion.
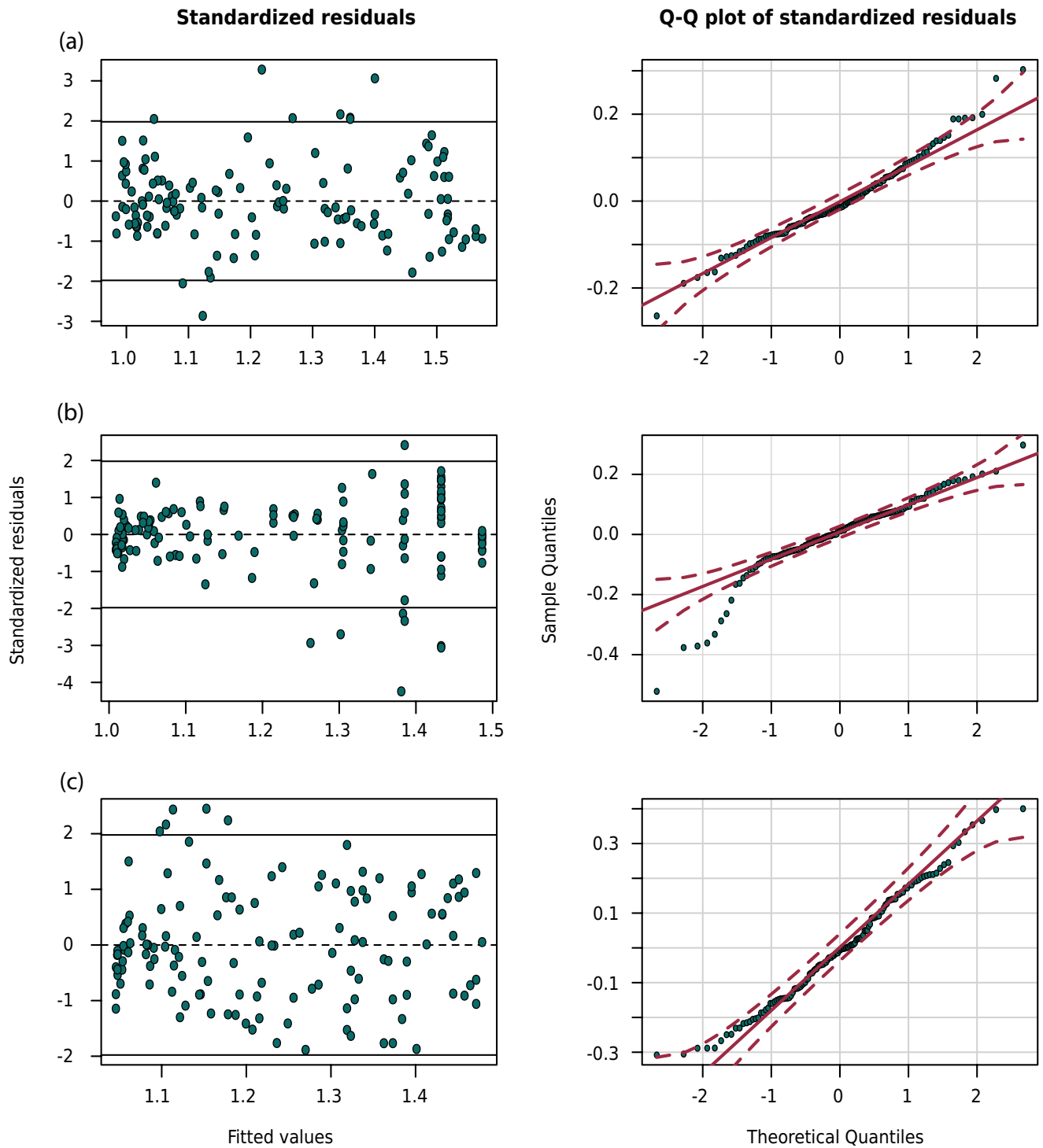
Figure 5. Fitted values vs. Standardized Residuals Plot (left panels) and the standardized residuals in a normal Q-Q Plot (right panels) to (a) PTF1, (b) PTF2, and (c) PTF3.

Table 4. Amplitude and descriptive analysis of data from 88 soil samples, 0.00-0.05 and 0.10-0.15 m layers, from a *Latossolo Amarelo* (Xanthic Oxisols), National Forest of Tapajós, Amazon

Descriptive Analysis	TS	Clay	OC
	g kg⁻¹		
Minimum	14	100	7.0
Median	30	780	20.3
Mean	118	727	25.3
Maximum	898	840	56.3
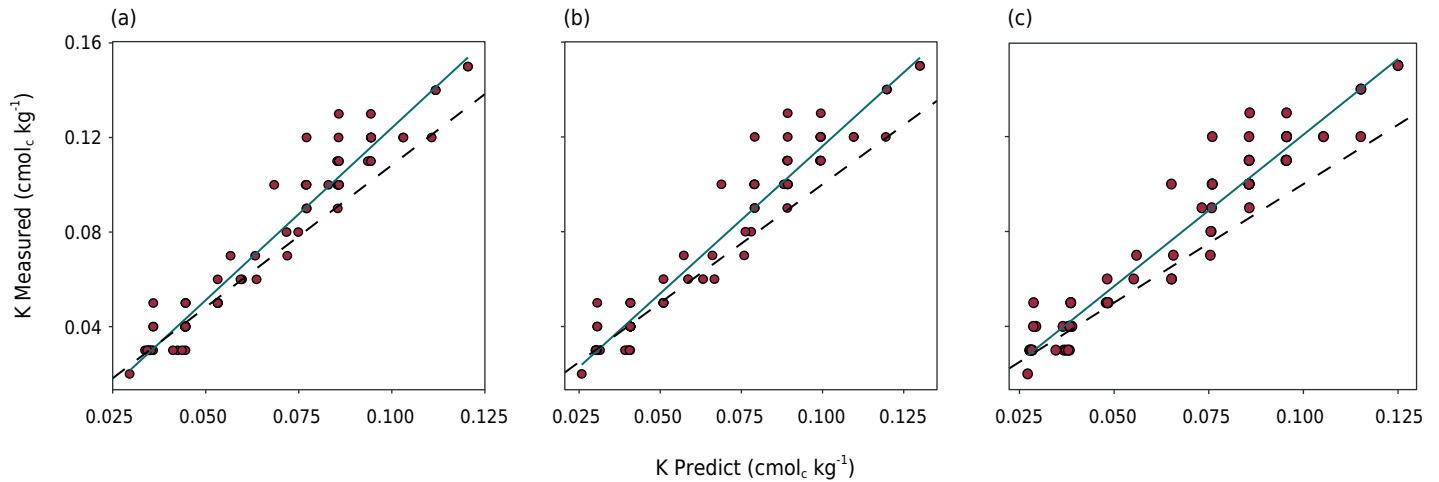
Figure 6. Scatterplots of K[+] of the measured vs predicted data, using PTF1 (a), PTF2 (b), and PTF3 (c). Validation samples from *Latossolo Amarelo* (Xanthic Oxisols), National Forest of Tapajós (AM).

DISCUSSION

As discussed by Mehlich (1972), the volumetric base data will only be equal to the gravimetric data if the fine earth density from the soil sample is precisely 1.0 g dm[-3], for other soils, it is necessary to convert units. Using the PTF1, the density from our data ranged from 0.86 to 1.68 g cm[-3], showing a soil mass variation within 10 cm[-3] of soil.

The US Soil Survey Laboratory Methods Manual (Soil Survey Staff, 2014a) uses a fixed factor (1.45 g cm[-3]) for the conversion of a volume estimate to a weight estimate of the bulk density for the fine earth, this means that no account is taken about the diversity of soil mass contained in a certain volume, which varies according to granulometry and the amount of organic carbon.

The use of PTF contributes to the standardization of chemical properties units in a dataset or database, facilitating the conversion of units in the volumetric base to the gravimetric base or vice versa. In this sense, we used a dataset of soils with anthropic horizons, including data from literature with different units of the chemical properties (unpublished data). This database was created to propose a taxonomic criteria for the anthropic horizon in the Brazilian Soil Classification System – SiBCS (Santos et al., 2018) and the pretic horizon in the World Reference Base for Soil Resource – WRB (IUSS Working Group WRB, 2015). Some of data sources compiled (Corrêa, 2007; Martins et al, 2007; Souza, 2011; Macedo, 2012; Silva et al., 2013; Miranda, 2018; Macedo et al., 2019) expressed chemical properties of horizons on a volumetric basis (as it is recommended for fertility purposes, that takes composite soil samples), while others used gravimetric basis for the expression of the results (unit recommended for soil classification purposes). The conversion was essential to standardize the various chemical properties data from the literature compiled in this database that had units on a volumetric basis.

Thus, for the dataset standardization, the best accuracy pedotransfer function obtained (i.e., PTF1) was applied to convert the calcium plus magnesium (Ca^{2+} + Mg^{2+}) data from the volumetric to the gravimetric base. Table 5 shows the results of the conversion.

As shown in table 5, the largest differences between converted values (i.e., gravimetric basis) and the original values (i.e., volumetric basis) were found for samples with high total sand contents. It corroborates with the data that had a density prediction greater than 1.0 g dm[-3] (Table 5). On the other hand, data with low total sand contents, such as IDs 3, 4, 5, 17, and 18 (Table 5), had a predicted density of approximately 1.0 g dm[-3], thus when applied the conversion of Ca^{2+} + Mg^{2+} data to gravimetric basis, there was

Table 5. Data from literature of calcium plus magnesium ($Ca^{2+}+Mg^{2+}$) in the volumetric base converted to the gravimetric basis through PTF1

Data dentification	Author	Horizon	Layer	Total sand	Predicted density by PTF1	$Ca^{2+}+Mg^{2+}$	Converted $Ca^{2+}+Mg^{2+}$ to $cmol_c\ kg^{-1}$
			m	g kg^{-1}		$cmol_c\ dm^{-3}$	$cmol_c\ kg^{-1}$
1	Souza (2011)	Au	0.00–0.30	430	1.06	28.4	26.8
2	Souza (2011)	Au	0.00–0.15	550	1.11	2.2	1.9
3	Souza (2011)	Au	0.00–0.08	140	0.99	38.8	38.9
4	Souza (2011)	A2u	0.08–0.20	160	0.99	17.6	17.6
5	Souza (2011)	Abiu	0.20–0.60	120	0.99	21.2	21.3
6	Silva (2013)	A1	0.00–0.18	503	1.08	18.0	16.5
7	Silva (2013)	A2	0.18–0.28	490	1.08	18.2	16.8
8	Martins et al. (2007)	A1	0.00–0.12	810	1.31	4.9	3.7
9	Martins et al. (2007)	A2	0.12–0.35	830	1.34	2.3	1.7
10	Miranda (2018)	Aup	0.00–0.05	562	1.11	18.2	16.3
11	Miranda (2018)	A2u	0.05–0.20	519	1.09	14.8	13.5
12	Miranda (2018)	A2u	0.12–0.60	367	1.04	2.0	1.8
13	Miranda (2018)	A2u	0.05–0.20	680	1.19	2.2	1.8
14	Miranda (2018)	A1u	0.00–0.12	517	1.09	18.3	16.8
15	Miranda (2018)	A2u	0.12–0.54	549	1.11	11.7	10.6
16	Miranda (2018)	A3u	0.54–0.64	458	1.07	20.6	19.3
17	Miranda (2018)	Ap	0.00–0.10	131	0.99	24.5	24.6
18	Miranda (2018)	2A3ub	0.60–1.00	130	0.99	26.2	26.2

practically no difference. Which was expected, since mathematically the closer is the density of the fine earth to 1.0 g dm^{-3}, the closer will be values of the chemical properties in volumetric and gravimetric base. However, the predicted thin earth density will not always be equal or close to 1.0 g dm^{-3}, thus chemical properties data on a volumetric basis differs from data on a gravimetric basis, requiring the functions to convert this data or a new measure. Therefore, the use of PTFs assists in the standardization of the database, which can later be used for various purposes, such as soil classification.

The main chemical properties used as criteria for identification of pretic horizons in the WRB (IUSS Working Group WRB, 2015) are exchangeable calcium and magnesium, available phosphorus, and organic carbon content. These properties are expressed on a gravimetric basis – "*exchangeable calcium plus magnesium content must be greater than 2.0 cmol$_c$ kg^{-1} on thin earth, organic carbon content must be greater than or equal to 10 g kg^{-1} and available phosphorus levels must be greater than or equal to 30 mg kg^{-1} soil*". When applying to the soil samples with IDs 2, 9, 12, and 13, $Ca^{2+}+Mg^{2+}$ values are greater than 2.0 cmol$_c$ dm^{-3}. When the data is converted to the gravimetric base (the unit used in the WRB) the resulting value decreases. Thus, according to WRB, after the conversion, these samples would not fit in the concept of a pretic horizon, since their $Ca^{2+}+Mg^{2+}$ values are less than 2.0 cmol$_c$ kg^{-1}.

The same situation can be observed for the criteria for phosphorus in both WRB and SiBCS. Macedo et al. (2019), for example, evaluated chemical properties of anthropic soils from naturally fertile floodplain areas of the Solimões River in the Brazilian Central Amazon. Phosphorus data were obtained on a volumetric basis (i.e., mg dm^{-3}) and the authors classified these soils according to WRB. In this work, specifically, the phosphorus data are very high and considerably exceed the WRB classification criterion (P above 30 mg kg^{-1}). However, if the phosphorus levels were near 30 mg dm^{-3}, the conversion of these data to the gravimetric basis would be paramount for the proper classification

of the horizons. That is, for soil profile analysis and classification, it is recommended to obtain the data on a gravimetric basis, while the volumetric basis properties are used for soil fertility purposes.

The use of PTF for standardization of the anthropic soil database was just one application example. It can be extended to other data standardization applications when access to samples is restricted or when the option to redo analyses is impracticable to change the units of chemical properties from volumetric base to gravimetric base or vice versa.

Simplified steps to the conversion of data

A. Selection of available data: total sand, clay, or organic carbon.

B. Apply a chosen PTF:

       (i) Pedotransfer Function 1 (PTF1) using Total Sand, or;

       (ii) Pedotransfer Function 2 (PTF2) using Clay, or;

       (iii) Pedotransfer function 3 (PTF3) using Organic Carbon.

C. Divide the value of the properties expressed in a volumetric basis, that is desired to convert, by the density predicted in the previous item, obtaining the value converted to gravimetric basis;

D. Alternatively, multiply the value of the properties expressed on a gravimetric basis and to be converted to predicted density, obtaining the value converted to the volumetric basis.

CONCLUSIONS

The pedotransfer functions obtained by nonlinear regressions, using the properties: total sand, clay, or organic carbon as predictor variables, allowed the conversion of soil chemical properties obtained in the gravimetric base to the volumetric base and vice versa. The pedotransfer function that presented the best precision to predict the fine earth density, which is essential for data conversion, is the one with total sand as a predictor (PTF1), presenting the smallest RSE of 0.092 g cm^{-3}.

The proposed pedotransfer functions can be used to standardize soil datasets or databases and convert chemical soil data expressed in different units.

This paper also highlights the importance of using the proper method of taking fine earth samples (either by weighing or by using the volumetric device) for laboratory analysis.

The selection of samples for elaboration of the functions did not cover all of Brazilian territory. Thus, a proposal for future work is to expand the selection to validated the PTFs in other regions and for all mineral soil orders.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://www.rbcsjournal.org/wp-content/uploads/articles_xml/1806-9657-rbcs-44-e0190086/1806-9657-rbcs-44-e0190086-suppl01.pdf

AUTHOR CONTRIBUTIONS

Conceptualization: ⓘD Wenceslau Geraldes Teixeira (lead) and ⓘD Fernanda Reis Cordeiro (supporting) and ⓘD Ademir Fontana (supporting).

Methodology: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Fernando Vieira Cesário (supporting), and ⓘ Wenceslau Geraldes Teixeira (supporting) and ⓘ Ademir Fontana (supporting).

Software: ⓘ Fernanda Reis Cordeiro (equal) and ⓘ Fernando Vieira Cesário (equal).

Validation: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Fernando Vieira Cesário (lead), ⓘ Ademir Fontana (supporting), ⓘ Lúcia Helena Cunha dos Anjos (supporting), ⓘ Ana Carolina Barbosa do Canto (supporting), and ⓘ Wenceslau Geraldes Teixeira (supporting).

Formal analysis: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Fernando Vieira Cesário (lead), and ⓘ Wenceslau Geraldes Teixeira (supporting).

Investigation: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Ana Carolina Barbosa do Canto (lead), and ⓘ Ademir Fontana (supporting).

Resources: ⓘ Ademir Fontana (lead), ⓘ Wenceslau Geraldes Teixeira (lead), and ⓘ Lúcia Helena Cunha dos Anjos (lead).

Data curation: ⓘ Fernanda Reis Cordeiro (lead) and ⓘ Ana Carolina Barbosa do Canto (supporting).

Writing – original draft: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Lúcia Helena Cunha dos Anjos (lead), ⓘ Ademir Fontana (lead), ⓘ Fernando Vieira Cesário (supporting), and ⓘ Ana Carolina Barbosa do Canto (supporting).

Writing – review and editing: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Lúcia Helena Cunha dos Anjos (lead), ⓘ Fernando Vieira Cesário (lead), and ⓘ Ademir Fontana (supporting)

Visualization: ⓘ Fernanda Reis Cordeiro (lead) and ⓘ Lúcia Helena Cunha dos Anjos (lead).

Supervision: ⓘ Fernanda Reis Cordeiro (lead), ⓘ Lúcia Helena Cunha dos Anjos (lead), ⓘ Ademir Fontana (lead), and ⓘ Wenceslau Geraldes Teixeira (supporting).

Project administration: ⓘ Fernanda Reis Cordeiro (lead) and ⓘ Ademir Fontana (lead).

ACKNOWLEDGMENTS

REFERENCES

Batista KD, Lumbreras JF, Coelho MR, Oliveira VA, Vale Júnior JF. Guia de campo da XI Reunião brasileira de classificação e correlação de solos: RCC de Roraima. Brasília, DF: Embrapa; 2018.

Benites VM, Machado PLOA, Fidalgo ECC, Coelho MR, Madari BE. Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil. Geoderma. 2007;139:90-7. https://doi.org/10.1016/j.geoderma.2007.01.005

Beutler SJ, Pereira MG, Tassinari WS, Menezes MD, Valladares GS, Anjos LHC. Bulk density prediction for Histosols and soil horizons with high organic matter content. Rev Bras Cienc Solo. 2017;41:e0160158. https://doi.org/10.1590/18069657rbcs20160158

Buurman P, van Lagen B, Velthorst EJ. Manual for soil and water analysis. Netherlands: Backhuys Publishers Leiden; 1996.

Cardoso LE, Crispim SMA, Filho AC, Lumbreras JF, Motta PEF, Araújo AR, Fernandes AHBN, Jacomini PKT, Fernandes FA, Coelho MR, Calderano SB, Shinzato E, Soares MTS. Guia de Campo da X Reunião Brasileira de Classificação de Solos: Pantanal e Cerrado. Rio de Janeiro: Embrapa Solos; 2012.

Cesário FV. Funcionamento e propriedades ecossistêmicas da vegetação e dos solos na Floresta Tropical Amazônica [tese]. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2018.

Comissão de Química e Fertilidade do Solo - CQFS-RS/SC. Manual de calagem e adubação para os Estados do Rio Grande do Sul e de Santa Catarina. 11. ed. Porto Alegre: Sociedade Brasileira de Ciência do Solo - Núcleo Regional Sul; 2016.

Corrêa GR. Caracterização pedológica de arqueo-antropossolos no Brasil: sambaquis da região dos Lagos (RJ) e terras pretas do índio na região do baixo rio Negro/Solimões (AM) [dissertação]. Viçosa, MG: Universidade Federal de Viçosa; 2007.

Dalgaard P. Introductory statistics with R. 2nd ed. New York: Springer-Verlag; 2002.

Dobarco MR, Cousin I, Le Bas C, Martin MP. Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty. Geoderma. 2019;336:81-95. https://doi.org/10.1016/j.geoderma.2018.08.022

Donagemma GK, Campos DVB, Calderano SB, Teixeira WG, Viana JHM. Manual de métodos de análise do solo. 2. ed. rev. Rio de Janeiro: Embrapa Solos; 2011.

Farazdaghi H, Harris PM. Plant competition and crop yield. Nature. 1968;217:289-90. https://doi.org/10.1038/217289a0

Freire LR, Balieiro FC, Zonta E, Anjos LHC, Pereira MG, Lima E, Guerra JGM, Ferreira MBC, Leal MAA, Campos DVB, Polidoro JC. Manual de calagem e adubação do Estado do Rio de Janeiro. Rio de Janeiro: Embrapa Solos; 2013.

Gaiser T, Graef F, Cordeiro JC. Water retention characteristics of soils with contrasting clay mineral composition in semi-arid tropical regions. Soil Res. 2000;38:523-36. https://doi.org/10.1071/SR99001

IUSS Working Group WRB. World reference base for soil resources 2014, update 2015: International soil classification system for naming soils and creating legends for soil maps. Rome: Food and Agriculture Organization of the United Nations; 2015. (World Soil Resources Reports, 106).

Lumbreras JF, Silva LM, Anjos LHC, Oliveira VA, Wadt PGS, Pereira MG, Delarmelinda-Honoré EA, Burity KTL. Guia de campo da XII Reunião brasileira de classificação e correlação de solos: RCC de Rondônia. Brasília, DF: Embrapa; 2019.

Macedo AGS. Caracterização e variação temporal da solução do solo em Argissolo Amarelo com horizonte a moderado e a antrópico (Terra Preta de Índio) no município de Iranduba- AM [dissertação]. Manaus: Universidade Federal do Amazonas; 2012.

Macedo RS, Teixeira WG, Lima HN, Souza ACG, Silva FWR, Encinas OC, Neves EG. Amazonian dark earths in the fertile floodplains of the Amazon River, Brazil: an example of non-intentional formation of anthropic soils in the Central Amazon region. Bol Mus Para Emílio Goeldi. Cienc Hum. 2019;14:207-27. https://doi.org/10.1590/1981-81222019000100013

Martins GC, Teixeira WG, Macedo RS, Marques JD. Ocorrência de horizontes antrópicos (Terra Preta de Índio) em Neossolos Quartzarênicos no município de Parintins - AM - Brasil. Brasília, DF: Ainfo; 2007 [cited 15 Jul 2019]. Available from: https://ainfo.cnptia.embrapa.br/digital/bitstream/item/112709/1/Gilvan-Martins-Oral-Terra-Preta.pdf.

McBratney AB, Minasny B, Captle SR, Vervoort RW. From pedotransfer functions to soil inference systems. Geoderma. 2002;109:41-73. https://doi.org/10.1016/S0016-7061(02)00139-8

Mehlich A. Uniformity of expressing soil test results a case for calculating results on a volume basis. Commun Soil Sci Plan. 1972;3:417-24. https://doi.org/10.1080/00103627209366396

Miranda JF. Terras pretas arqueológicas no Médio Amazonas e Alto Solimões: química, mineralogia, micromorfologia e idade [tese]. Viçosa, MG: Universidade Federal de Viçosa; 2018.

Nascimento GB, Anjos LHC, Pereira MG, Fontana A, Santos HG. Funções de pedotransferência do conteúdo de água em Latossolos Amarelos e Argissolos Amarelos. Rev Bras Cienc Agrar. 2015;5:560-9. https://doi.org/10.5039/agraria.v5i4a551

Ottoni MV, Ottoni Filho TB, Schaap MG, Lopes-Assad MLRC, Rotunno Filho OC. Hydro - physical database for Brazilian soils (HYBRAS) and pedotransfer functions for water retention. Vadose Zone J. 2018;17:170095. https://doi.org/10.2136/vzj2017.05.0095

Paine CET, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, Turnbull LA. How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. Methods Ecol Evol. 2012;3:245-56. https://doi.org/10.1111/j.2041-210X.2011.00155.x

Patton NR, Lohse KA, Seyfried M, Will R, Benner SG. Lithology and coarse fraction adjusted bulk density estimates for determining total organic carbon stocks in dryland soils. Geoderma. 2019;337:844-52. https://doi.org/10.1016/j.geoderma.2018.10.036

Prezotti LC, Gomes JA, Dadalto GG, Oliveira JA. Manual de recomendação de calagem e adubação para o Estado do Espírito Santo - 5ª aproximação. Vitória: Incaper; 2013.

Qiao J, Zhu Y, Jia X, Huang L, Shao M. Development of pedotransfer functions for predicting the bulk density in the critical zone on the Loess Plateau, China. J Soils Sediments. 2019;19:366-72. https://doi.org/10.1007/s11368-018-2040-1

R CORE TEAM. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2019. Available from: http://www.R-project.org.

Ritz C, Streibig JC. Nonlinear regression with R. New York: Springer-Verlag; 2008.

Santos HG, Jacomine PKT, Anjos LHC, Oliveira VA, Lumbreras JF, Coelho MR, Almeida JA, Araújo Filho JC, Oliveira JB, Cunha TJF. Sistema brasileiro de classificação de solos. 5. ed. rev. ampl. Brasília, DF: Embrapa; 2018.

Silva CS, Silva-Filho FC, Santos AD, Coscione AR, Vitti A, Boaretto A, Coelho JAPM, van Raij B, Silva CA, Abreu CH, Carmo C, Silva CR, Abreu CA, Gianello C, Andrade-Coelho CA, Pérez D, Casarini D, Silva FC, Prata F, Carvalho FC, Santos L, Cantarella H, Fernandes H, Andrade JC, Quaggio JA, Chitolina J, Cunha L, Pavan MA, Rosias M, Tedesco MJ, Miyazawa M, Abreu MF, Eira PA, Higa R, Massrubá SMFS, Gomes T, Muraoka W, Vieira W, Melo WJ, Barreto W, Boareto APB. Manual de análises químicas de solos, plantas e fertilizantes. 2. ed. Brasília, DF: Embrapa; 2009.

Silva RNP. Interferência de sistemas de manejo na floresta nacional de Caxiuanã, PA nos atributos físicos e químicos na composição isotópica de solo antropizado [tese]. Botucatu: Universidade Estadual Paulista; 2013.

Soil Science Department. Fertilizer requirements of countries in Latin America: Annual Report. North Carolina State: North Carolina State University, 1966.

Soil Survey Staff. Kellogg soil survey laboratory methods manual. Washington, DC: United States Department of Agriculture, Natural Resources Conservation Service; 2014a. (Soil Survey Investigations Report No. 42).

Soil Survey Staff. Keys to soil taxonomy. 12th ed. Washington, DC: United States Department of Agriculture, Natural Resources Conservation Service; 2014b.

Souza KW. Gênese, mineralogia, micromorfologia e formas de fósforo em arqueo-antropossolos da várzea do Rio Amazonas [tese]. Viçosa, MG: Universidade Federal de Viçosa; 2011.

Spearman C. The proof and measurement of association between two things. Am J Psychol. 1904;15:72-101. https://doi.org/10.2307/1412159

Stewart V, Adams W, Abdulla H. Quantitative pedological studies on soils derived from Silurian mudstones: II. The relationship between stone content and the apparent density of the fine earth. J Soil Sci. 1970;21:248-55. https://doi.org/10.1111/j.1365-2389.1970.tb01174.x

Teixeira PC, Donagemma GK, Fontana A, Teixeira WG. Manual de métodos de análise de solo. 3. ed. rev e ampl. Brasília, DF: Embrapa; 2017.

van Den Berg M, Klamt E, van Reeuwijk LP, Sombroek WG. Pedotransfer functions for the estimation of moisture retention characteristics of Ferralsols and related soils. Geoderma. 1997;78:161-80. https://doi.org/10.1016/S0016-7061(97)00045-1