An Acad Bras Cienc (2021) 93(4): e20200646 DOI 10.1590/0001-3765202120200646
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

SOIL SCIENCE

X-ray fluorescence spectrometry applied to digital mapping of soil fertility attributes in tropical region with elevated spatial variability

LUCAS BENEDET, MATHEUS S. NILSSON, SÉRGIO HENRIQUE G. SILVA, MARCELO H.P. PELEGRINO, MARCELO MANCINI, MICHELE D. DE MENEZES, LUIZ ROBERTO G. GUILHERME & NILTON CURI

Abstract: Portable X-ray fluorescence (pXRF) spectrometry offers valuable information for prediction models of soil fertility attributes spatial variation, although this approach is yet scarce in tropical regions. This study aims to predict and build spatial variability maps of soil pH, remaining phosphorus (P-Rem), soil organic matter (SOM) and sum of bases (SB) using pXRF results through stepwise multiple linear regression (SMLR) and Random Forest (RF) in a highly variable tropical area. Composite samples from soil A horizon were collected at 90 points throughout the campus of the Federal University of Lavras, Minas Gerais, Brazil, for pH, P-Rem, SOM, SB and pXRF analyses. RF predictions showed the highest accuracies, especially for P-Rem and SB (R^2 values of 0.66 and 0.55, respectively). Attributes that showed higher R^2 in punctual predictions also exhibited higher R^2 in spatial predictions. Data obtained from pXRF in tandem with RF can be used to assist prediction models for soil fertility attributes, consequently enabling the digital mapping of such attributes and helping to improve the knowledge about the spatial variability of such attributes in soils of tropical climate. This technique can therefore assist in the identification and orientation of adequate management practices in tropical agricultural practices.

Key words: Proximal sensor, random forest, spatial prediction, tropical soils.

INTRODUCTION

Modern agricultural systems search for optimal efficiency in the utilization of the necessary products and energy to guarantee higher productivity with minimum or null environmental impact (Lopes & Guilherme 2016). Therefore, the application of fertilizers and amendments should be based on soil analyses, correcting deficiencies and avoiding excesses (Hedley 2015). Following on, precision agriculture, when performed correctly, promotes better management of agricultural lands, compartmentalizing productive areas in homogeneous zones according to soil type, fertility, slope, etc. in order to both drive and apply specific management practices suitable to each zone (Mulla 2013). The development and efficacy of precision agriculture were only possible due to the introduction of new technologies, like the Global Positioning System (GPS), proximal sensors, Geographic Information Systems (GIS) and the utilization of advanced software and precise equipment (Hedley 2015). Such technology is vital, as they offer fast and robust information about soil attributes, plant

nutrition and climate conditions. Also, they are the foundation of decision-making and creation of detailed maps (Mulla 2013, Mancini et al. 2019). However, information about soil attributes is based on traditional laboratory analyses.

Although larger number of samples provides better soil characterization of the area of interest, the number of samples to be analyzed may be constrained since conventional laboratory analyses are laborious and time-consuming (Weindorf et al. 2014). Furthermore, such methods require a large quantity of reagents, which often demand special attention regarding their adequate disposal in order to avoid environmental pollution (Weindorf et al. 2016). Under such context, proximal sensors are being increasingly used for the creation of soil attributes prediction models that are faster and can work on a larger scale (Mancini et al. 2019). One of these sensors that has drawn soil scientist's attention is the portable X-ray fluorescence (pXRF) spectrometer (Silva et al. 2017, Qu et al. 2019).

The pXRF identifies and quantifies the elemental contents in a sample using the principle of fluorescence induction in atoms. According to this principle, when an X-ray beam hits an atom, electrons from the inner shells move towards the outer shells, and, when they return to their original state, energy, thus named fluorescence, is released (Weindorf et al. 2014, Ribeiro et al. 2017, Silva et al. 2021). The kind and intensity of the emitted wavelength are directly related to unique characteristics of each element, which permits their identification and quantification in studied samples (from Mg to U in the Periodic Table) (Sharma et al. 2014, Weindorf et al. 2014, 2016). This process is fast and non-destructive, requiring minimal sample preparation (Zhu et al. 2011).

Elemental results delivered by pXRF can be directly utilized to determine heavy metal content in contaminated areas or agricultural land (Hu et al. 2017), or even applied to prediction models of diverse soil attributes, such as pH (Sharma et al. 2014), cation exchange capacity (Sharma et al. 2015), calcium sulfate (Weindorf et al. 2013), texture (Zhu et al. 2011, Silva et al. 2020), and soil classes (Benedet et al. 2020b). Additionally, data obtained from pXRF and resulting predictions can be spatialized and used to generate maps, allowing for the quantification of soil attributes in an area (Qu et al. 2019, Duda et al. 2017), or along soil profiles (Silva et al. 2018b, Hartemink et al. 2020).

Most of these analyses are performed by utilizing statistical techniques and machine learning algorithms. Common examples are the multiple linear regression (MLR), which analyses the effect of two or more independent variables upon a dependent variable (Tabachnick & Fidell 1996); and Random Forest (RF), a non-parametric technique based on the creation of several decision trees to optimize the prediction performance of models (Breiman 2001). Many works have shown that the algorithm used in the modeling process significantly influences the prediction accuracy of soil attributes (Heung et al. 2014, Weindorf et al. 2016, Mancini et al. 2019, Teixeira et al. 2018).

In spite of the increasing use of pXRF in the prediction of soil attributes, few studies extend its applications to spatial predictions (Duda et al. 2017), especially in tropical soils. The growth of precision agriculture in Brazil favors the development and adoption of proximal sensors in search for faster estimation of soil attributes and their temporal and spatial distribution, contributing to a low-cost and correct application of fertilizers and amendments. Therefore, this study aims to evaluate the capacity of pXRF data to predict pH, remaining P (P-Rem), soil organic matter (SOM) and sum of bases (SB) via stepwise MLR (SMLR)

and RF in an area with very high spatial variability of soils and land uses in order to map the intricate distribution of these attributes across tropical landscapes. The thorough knowledge of the distribution of soil fertility parameters permits the precise identification of optimal fertilizer input requirements in each specific area, thus contributing to decision making concerning precision agriculture, at lower costs due to the use of proximal sensors. We hypothesize that pXRF in tandem with machine learning algorithms will allow for reliable spatial predictions, enabling the adequate mapping of soil fertility attributes and further leading to the optimization of tropical crop management.

MATERIALS AND METHODS

Study area

This study was conducted at the Federal University of Lavras' campus, located at the city of Lavras, Minas Gerais state, Brazil, between latitudes 21° 13'07'' and 21° 14'27''S, and longitudes 44° 59'24'' and 44° 57'29''W, comprising an area of 314.5 hectares (Fig. 1). This area presents very high spatial variability regarding soil classes, parent material, relief and land uses (Curi et al. 2017) (Fig. 1). Soil composite sampling encompassed areas with annual plantations, pasture, coffee plantations, candeia plantations, orchards, eucalyptus and pinus plantations, areas with native vegetation, riparian forest and non-vegetated areas (Ferreira et al. 2013).

According to the Köppen classification system, climate in the region is Cwa with mean annual temperature of 20.4° C and mean annual rainfall of 1460 mm (Alvares et al. 2013). The study area has altitude of 920 m, approximately, and is characterized by high-altitude tropical climate with hot and humid summers and dry and mild winters.

Sampling and soil analyses

For the current study, 90 composite samples from A horizon (0-20 cm) were collected at the study area in a regular grid (Fig. 2). Samples were air-dried and homogenized, and passed through a 2 mm sieve (air-dried fine earth). Next, samples were submitted to laboratory analyses to determine the following soil attributes: soil organic matter (SOM) (Walkley & Black 1934), pH in water (soil solution ratio of 1:2.5), remaining P (P-Rem) (Alvarez et al. 2000), exchangeable Ca^{2+}, Mg^{2+} and Na^{+} contents (McLean et al. 1958), and available K^{+} extracted with Mehlich-1 (Mehlich 1984). The SB was determined by summing the K^{+}, Ca^{2+}, Mg^{2+} and Na^{+} contents.

pXRF analyses

A Bruker S1 Titan LE pXRF was used to determine the elemental contents in the air-dried fine earth (ADFE) samples used for prediction of the results obtained in laboratory. This equipment contains 50kV and 100 µA X-ray Rh tubes. The high energy X-rays emitted by pXRF hit the atoms, displacing electrons from inner to outer orbits. As they return to their original orbit, they emit energy as fluorescence. Since the atoms of each chemical element has unique fluorescence characteristics, a silicon drift detector present in the pXRF unit is able to identify the atoms contained in the analyzed material, according to the type of fluorescence emitted, and the amount of each element, due to the intensity of each fluorescence (Weindorf et al. 2014). Moreover, this analysis require minimal sample preparation, is fast (30 to 90 seconds, in general) and does not produce chemical waste (Weindorf et al. 2014, Ribeiro et al. 2017). Due to the ease of such procedure, pXRF analysis has been increasingly adopted across the world for multiple purposes (Silva et al. 2021). From the analyzed elements, 20 were above the limit of

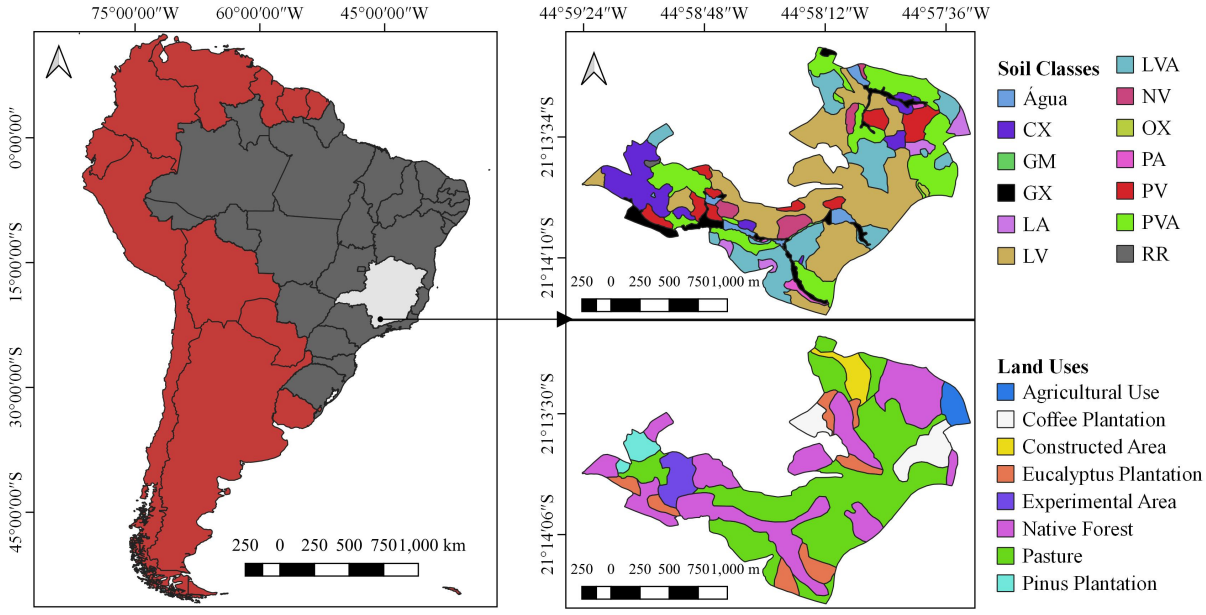

Figure 1. **Geographic location of the study area, at the city of Lavras, state of Minas Gerais, Brazil, soil classes and land uses. CX – Haplic Cambisol, GM – Melanic Gleysol, GX – Haplic Gleysol, LA – Yellow Latosol, LV – Red Latosol, LVA – Red-Yellow Latosol, NV – Red Nitosol, OX – Haplic Organosol, PA –Yellow Argisol, PV – Red Argisol, PVA – Red-Yellow Argisol, RR – Regolithic Neosol.**

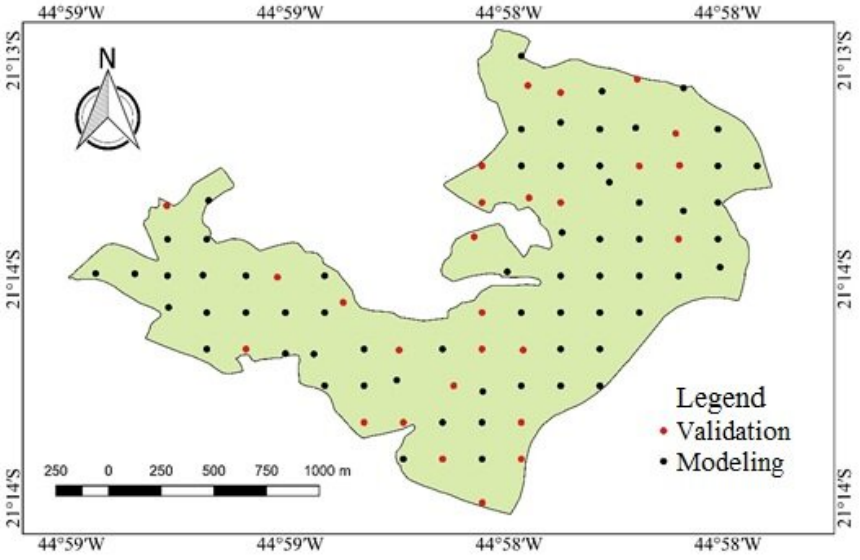

Figure 2. **Sampling points used for modeling and validation collected from the study area located at Lavras, state of Minas Gerais, Brazil.**

detection for all the samples: Al, Si, P Cl, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Rb, Zr, Sr, Nb, Pb, which were utilized to create the prediction models. All analyses were performed in triplicate. The utilized software was the GeoChem, configured in dual soil mode (Trace), for analyses during 60 seconds each (Weindorf et al. 2016).

To measure the equipment accuracy, two materials certified by the National Institute of Standards and Technology (NIST) were used

as reference 2710a and 2711a; additionally, a standard sample (check sample – CS) provided by the pXRF manufacturer was also tested. Through the mentioned reference samples, the pXRF recovery values were calculated for each element according to the formula: recovery value = content obtained by pXRF/certified sample content. Each reference material has its own unique chemical composition and should guarantee that the equipment is providing values comparable to the certified materials. The recovery values (2710a/2711a/CS) for the elements used in this study were (zero values indicates that the equipment did not detect the element or that the element is not present in the certified reference samples): Al (0.80/0.71/0.91), As (0.86/0.68/0), Ca (0.39/0.46/0), Cl (0/0/0), Cr (0/1.11/0), Cu (0.82/0.85/0.89), Fe (0.74/0.80/0.87), K (0.55/0.51/0.88), Mn (0.70/0.68/0.83), Ni (0/1.15/0.96), P (3.85/5.4/0), Pb (1.18/1.05/1.06), Rb (0.92/0.89/0), Si (0.64/0.51/0.87), Sr (1.11/0.83/0), Ti (0.78/0.75/0), V (0.78/1.15/0), Y (0/0/0), Zn (0.95/0.78/0), e Zr (1.12/0/0).

Statistical analyses

Analyses were performed in four steps: a) descriptive statistics of obtained data; b) creation of prediction models for the studied soil attributes using SMLR and RF; c) evaluation of prediction results and their comparison with laboratory results, in order to select the best model to create soil attribute maps; and d) spatial prediction and validation with laboratory results. Results from pXRF and laboratory analyses were submitted to descriptive statistics to verify the position (minimum, maximum and mean) and dispersion (standard deviation and coefficient of variation) of the data.

Sample analyses were divided randomly in a dataset for modeling purposes and a validation dataset (Fig. 2). A total of 63 samples (70%) were used to build prediction models and 27 (30%) were used for validation, i.e. the validation samples have their observed value (obtained in the laboratory via conventional analysis) compared with the value predicted by the model (estimated value based on both the prediction models and the pXRF elemental results). The separation of both datasets was completely random.

For predictions by the SMLR, the backward method was utilized through the software Sigma Plot, with removal of the least important variables from the model, with 95% probability. This method adds all variables to the modeling process at once, and removes the least important variables, one by one, reaching a final model yielding the best result. Also, this method delivers a final equation that can be used in the future for such predictions. Conversely, RF models are an improvement of decision trees. In this method, several trees are generated (a forest), by randomizing and resampling (out-of-bag) the variables present in each tree (Breiman 2001). These models were created using the R package "randomForest" (Liaw & Wiener 2018) with the rf method. The utilized parameters were: number of trees (ntrees) = 1000; number of variables in each node (nodesize) = 5; number of variables used per tree (mtry) = one third of the number of samples, as suggested by Liaw & Wiener (2018). For RF models, the mean of squared error (MSE) (Eq. 1) and the percentage of variance explained (%Varex) (Eq. 2) were calculated using the out-of-bag (OOB) methodology to determine the importance of each variable in models. In this method, for each interaction, only some predictor variables are used to generate a tree, based on the measure of the importance of predictor variables. Thus, the algorithm can identify the most important variables (Liaw & Wiener 2018); that is, if a variable is removed

from the model, resulting in more error and less precision, such variable is considered more relevant.

$$MSE_{OOB} = \frac{1}{n}\sum_{1}^{n}[y_i - \hat{y}_i^{OOB}]^2 \quad (1)$$

$$\%Var_{ex} = 1 - \frac{MSE_{OOB}}{\hat{\sigma}_y^2} \quad (2)$$

where n is the number of trees, y_i is the ith observed value; $\hat{y}_i^{OOB}$ is the average of the OOB predictions for the ith observation; and $\hat{\sigma}_y^2$ is the variance computed with n as divisor, instead of $(n-1)$.

Finally, the parameters used to validate the predictions delivered by SMLR and RF models were the root mean square error (RMSE) (Eq. 3), mean error (ME) (Eq. 4), along with the coefficient of determination (R^2) and the adjusted R^2 (R^2_{adj}).

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}(e_i - m_i)^2} \quad (3)$$

$$ME = \frac{1}{n}\sum_{1}^{n}(e_i - m_i)^2 \quad (4)$$

where, n is the number of observations; e_i is the ith predicted value; m_i is the ith actual value.

For these calculations, independent data that were not included in the modeling process were used, comprising 30% of the number of samples. The models with the best performance for each predicted soil attribute were chosen based on: highest R^2 and R^2_{adj} scores and lower RMSE and ME values and used for creating the maps of the soil attributes.

Spatial prediction of soil attributes

After selecting the best prediction models for each soil attribute, these were used to predict these attributes throughout the entire study area. Elemental data from the pXRF analyses were spatialized over the study area utilizing the Inverse Distance Weighted (IDW) interpolation method, using the SAGA GIS software. This was necessary since sample data was only available at the points where they were collected and, in order to create maps, data need to be available over all study area. Then, the best prediction models were applied to the interpolated maps, generating maps of the predicted soil properties. These maps were also validated by comparing pixel values with those observed at the same place, using the same validation samples mentioned above. The accuracy of these maps was measured via R^2, RMSE and ME.

RESULTS AND DISCUSSION

Descriptive statistics

The descriptive statistics for SB, SOM, P-Rem and pH obtained from laboratory analyses, including modeling and validation datasets, are shown in Table I. The values of SB, SOM and P-Rem show higher coefficient of variation (CV), but not surpassing 58.4%. The lowest CV was obtained for pH, representing lower variability in studied samples.

The study area presents high diversity of soil types (Fig. 1), reflecting specific pedological processes. Although climate and biologic diversity are rather homogeneous, parent materials, relief and soil formation time vary considerably across the area, resulting in physical and chemical attributes of differential soils (Curi et al. 2017). Additionally, the different land uses and management practices adopted in the area (Fig. 1) influenced these attributes, through tillage, vegetal cover (quality and quantity of C input) and application of products (liming and chemical and organic fertilizers) (Triberti et al. 2016, Han et al. 2019).

According to Araujo et al. (2017), in areas with similar native vegetation cover, the accumulation of SOM is affected especially by altitude (>900

Table I. Descriptive statistics of the soil attributes: pH, sum of bases (SB), soil organic matter (SOM) and remaining phosphorus (P-Rem), obtained from laboratory analysis from soil samples collected at Lavras, Minas Gerais, Brazil.

	Modeling					Validation			
	SB $cmol_c\ dm^{-3}$	SOM %	P-Rem $mg\ dm^{-3}$	pH		SB $cmol_c\ dm^{-3}$	SOM %	P-Rem $mg\ dm^{-3}$	pH
Minimum	1.4	2.1	3.9	4.6		1.8	3.1	3.2	4.0
Maximum	26.6	14.1	35.1	7.2		13.3	11.2	36.1	7.0
Mean	6.2	6.2	22.8	5.8		6.9	6.3	22.7	5.8
SD[1]	3.6	2.4	7.0	0.6		3.2	2.2	7.9	0.7
CV[2] (%)	58.4	38.3	30.9	10.7		46.4	35.6	34.7	11.7

[1]Standard deviation; [2]Coefficient of variation.

m), surpassing even the influence of clay content and parent material in soil C content. Higher altitude zones present more SOM accumulation due to the reduced microbiological activity and vegetal material decomposition rate. Therefore, it is logical to conclude that the variation in SOM contents observed in the study area is related to the difference in vegetation and soil management practices, which promotes alterations in soil C balance (Cates et al. 2016). The conventional preparation of soil and potential rupture of macro- and micro-aggregates leave SOM unprotected from biological activity, favoring its decomposition (Gupta & Germida 2015). Additionally, vegetal residues with low C:N ratio degrade faster, even with high biomass input (Cates et al. 2016).

Obtained from the availability of bases in soil (Ca^{2+}, Mg^{2+}, K^{+}, and Na^{+}), SB is a reliable indicative of soil fertility and is related to the capacity of organic and mineral components to retain and furnish nutrients to the soil solution (Khorshidi & Lu 2017). In a study by Vågen et al. (2016) in Africa, SB values presented high variability, ranging between 1 and 138 $cmol_c\ kg^{-1}$ (mean of 15 $cmol_c\ kg^{-1}$). These authors have also observed higher values of SB in soils with higher pH values and in regions where climate is drier. Additionally, variation in SB values can also be explained by liming and the use of fertilizers in certain areas, due to the increase in Ca^{2+} contents (Han et al. 2019).

Contents of P-Rem vary between 3.2 and 36.1 mg dm^{-3}, with mean of 22.7 mg dm^{-3}. These differences are caused by soil class variability, reflecting the different capacities to adsorb phosphates. This was observed by Bornø et al. (2018), who evaluated P dynamics in multiple soil types. Authors observed low maximum adsorption capacity (211 mg P kg^{-1}) and higher content of available P (35.85 mg P kg^{-1}) in Spodosols, and high maximum adsorption capacity (2242 mg P kg^{-1}) and low available P content (9.98 mg P kg^{-1}) in Alfisols. Such P sorption capacity is strongly influenced by SOM content, texture and mineralogy of the clay fraction (Fang et al. 2017, Bornø et al. 2018).

The pH variation tends to be smaller than other soil attributes, as observed by Vågen et al. (2016) in comparison to soil organic carbon values (SOC), clay and SB in soils from Africa. One of the reasons for this low variation, especially in forest areas, is related to the capacity of certain species to neutralize pH,

reducing pH values of alkaline soils (>7.5) and raising pH values of acidic soils (<5.0) (Hong et al. 2018). Authors argue that soils' original pH can influence the liberation of root exudates (HCO_3^-, OH^-, and H^+) and the decomposition of vegetal residues, promoting specific alterations in pH. Concurrently, soil amendments present in some areas of the current study also contributed to higher homogeneity in pH values (Han et al. 2019).

Regarding pXRF analyses, 20 elements were determined in studied samples (Table II). Elements found in highest concentrations were Al, Si, P, K, Ca and Fe. The high contents of these elements, especially Al, Si, K and Fe are related to the chemical composition of the minerals present in such soils (Curi et al. 2017). While Si can be found mostly in quartz (SiO_2) and muscovite ($KAl_2(Si_3Al)O_{10}(OH,F)_2$ dominantly in the sand fraction and kaolinite ($Al_2(OH_4)Si_2O_5$) in the clay fraction, Al can be found in gibbsite ($AlOH_3$) as well as in muscovite and kaolinite (Silva et al. 2018a, Benedet et al. 2020a, Silva et al. 2020). Fe is generally associated with the presence of Fe oxides that are very common in these soils, such as hematite (Fe_2O_3), goethite (FeOOH), magnetite (Fe_3O_4) and maghemite (γ–Fe2O3¬) (Curi et al. 2017). In the case of K and Ca, these elements are mobile in soil, although they can be found in some minerals, such as micas and feldspars (Batista et al. 2018). With the exception of muscovite, other micas as well as feldspars were not commonly found in the soils of the study area. Additionally, although P is naturally provided by soil parent material, its contents tend to be low in the studied soils under natural conditions, due to their intense degree of weathering-leaching, as reported by Silva et al. (2018b) and Ribeiro et al. (2010). This fact explains the low values of this element (Table II). Therefore, P high contents obtained by pXRF are associated with the areas under some anthropic influence, such as the application of fertilizers for agriculture (Weihrauch & Opp 2018).

The variability of elemental contents obtained by pXRF is in accordance with previously described results obtained for SOM, pH, P-Rem and SB, reflecting the multiple soil types, mineralogy, texture, parent material and land uses that comprise the study area (Curi et al. 2017).

Calibration of prediction models

Stepwise Multiple Linear Regression

The equations generated by SMLR (Table III) were promising since the best results were achieved for P-Rem and SB, presenting R^2 values of 0.85 and 0.95, respectively. SOM and pH predictions achieved R^2 scores of 0.75 and 0.67, respectively. Although performance was lower for pH, the R^2 represents sufficient prediction capacity of pXRF data for this attribute. Sharma et al. (2014), for instance, reached R^2 values of 0.57 and 0.77 from two different datasets using soil samples from USA. Additionally, results obtained in this study were superior to those obtained by Silva et al. (2017) in areas of the Brazillian Cerrado for the same attributes using SMLR.

The equations created to predict pH and SOM used the highest number of variables: 8 and 6, respectively (Table III). Equations that used a small number of variables were those created for SB and P-Rem, using 4 and 3 variables, respectively. In these equations, the most highlighted variables were Ca, V, Ni and Si, with 3 occurrences each. The Ca variable was not present only in pH prediction equation, but was present in all others. These results differ from those obtained by Sharma et al. (2014), where Ca data were present in simple and multiple linear regression equations for pH predictions. This demonstrates that variable

Table II. Descriptive analysis of data obtained via pXRF for soils sampled at Lavras, Minas Gerais, Brazil.

	Minimum (ppm)	Maximum (ppm)	Mean (ppm)	SD[1] (ppm)	CV[2](%)
Al	103111	245797	178631	32240	18.0
Si	119532	414752	267958	74750	27.9
P	0	2671	708	510	72.0
Cl	95	935	523	150	28.7
K	801	34101	3217	4421	137.4
Ca	219	66013	5481	7538	137.5
Ti	2428	18308	9583	3654	38.1
V	0	462	114	112	98.2
Cr	0	5910	583	881	151.1
Mn	129	2153	617	397	64.3
Fe	15902	187076	80872	44844	55.5
Ni	0	750	715	118	16.5
Cu	10	140	36	24	66.7
Zn	15	123	41	21	51.2
As	0	9	1	2	200.0
Rb	0	138	11	17	154.5
Sr	0	158	15	19	126.7
Zr	98	332	216	49	22.7
Nb	0	25	7	7	100.0
Pb	0	53	16	15	93.8

[1] Standard deviation; [2]Coefficient of variation.

Table III. Prediction equations for sum of bases (SB), remaining P (P-Rem), soil organic matter (SOM) and pH obtained by SMLR from soil samples collected at Lavras, Minas Gerais, Brazil.

Equation	R^2
SB = 2.58000 + 0.000731*Ca + 0.00397*V + 0.0156*Ni - 0.0324*Cu	0.95
P-rem = 3.338 + 0.0000601*Si + 0.001*Ca - 0.0208*Ni	0.85
SOM = 20.931 - 0.0000212*Al - 0.0000307*Si + 0.000228*Ca + 0.00854*V - 0.0000644*Fe + 0.00819*Ni	0.75
pH = 3.063 + 0.00000717*Si - 0.000206*K+ 0.00314*V -0.00029*Cr - 0.0011*Mn + 0.0000121*Fe + 0.0705*Rb + 0.0283*Sr	0.67

SB = $cmol_c\ dm^{-3}$; SOM = dag kg^1; P-rem = mg dm^{-3}.

importance might change according to the dataset and the studied region. The different Ca importance in both studies can be explained: soils with low pH values are commonly deficient in Ca^{2+}, and liming ($CaCO_3$) usually results in a positive relation between pH increase and the increase in Ca^{2+} content (Behera & Shukla 2015, Vågen et al. 2016, Teixeira et al. 2018). However, Ca^{2+} has little effect upon pH alteration in soils, and is more related to an increase in SB, CEC and base saturation, and reduces toxicity by Al^{3+}, favoring the cellular stability of plant roots (Indrasumunar et al. 2012, Anikwe et al. 2016). Although it is important to note that Ca elemental content values obtained via pXRF represent total contents, which does not represent the element's availability in soil (Weindorf et al. 2016). However, Brazilian soils tend to have low amount of Ca in the crystalline structure of minerals (Resende et al. 2014), thus, most of the Ca determined by pXRF is related to the exchangeable Ca^{2+} (Silva et al. 2019).

According to Imtiaz et al. (2015), the presence of V in soils is associated especially with the presence of titanium-magnetite, since in its crystalline structure Fe can be replaced by V through isomorphic substitution. Besides being present in minerals, V can be adsorbed or be a part of Fe oxides and other clay minerals, having its geochemical behavior highly influenced by pH. Additionally, V has high attraction by SOM and C deposits have highest V contents than other deposits. This helps to explain the importance of this variable when predicting SB, SOM and pH. Regarding Si, its positive participation in P-Rem predictions is mainly related with the lower sorption capacity of P in soils with high sand content (Resende et al. 2014). Soils with higher quartz content present less specific surface area and adsorption sites, restricting their capacity to retain P. Additionally, the positive participation between Si values in the pH equation suggests a positive relation between these variables. Therefore, higher pH in soils with higher Si contents may indicate lower P sorption capacity due to the increase in negative charges (Bornø et al. 2018, Weihrauch & Opp 2018).

The capacity to adsorb P can also be related to the importance of Ni in the equations. Ni availability in soils is strongly related to texture, type of clay, CEC, SOM, pH and ionic force. Therefore, soils with more adsorption sites are able to retain more Ni, especially those rich in organic groups, due to their high affinity with Ni (Elbana et al. 2018). Notably, such soils tend to present higher SB and SOM. This relation between organic groups and minerals with Ni was also observed by Sun et al. (2018). These authors observed an increase in Ni prediction accuracy when utilizing spectral bands associated with clay and SOM.

Random Forest

The highest MSE_{OOB} scores for RF models were obtained in the order: P-Rem>SB>MOS>pH (Fig. 3), and this same pattern was observed for Var_{exp} values. MSE_{OOB} scores ranged between 0.3 (pH) to 20.7 (P-rem), and Var_{exp} values from 19.5% (pH) to 57.5% (P-rem). Results were similar to those achieved by Wiesmeier et al. (2011), who obtained 72.64 MSE_{OOB} and 61.9% Var_{exp} values when predicting SOM via RF. Interestingly, despite the low MSE_{OOB} scores in SOM and pH predictions, values of $\%Var_{exp}$ were also low, which denotes a reduced predictive capacity of models. Such results differed from those attained by Silva et al. (2017), who achieved models with high $\%Var_{exp}$ and low MSE_{OOB}. This favored higher R^2 scores for the prediction of Ca^{2+}, pH, Al^{3+}, P-rem, effective CEC and base saturation via RF in their work. Thus, low $\%Var_{exp}$ values obtained for SOM and pH in this study indicates the difficulty of models to predict these attributes.

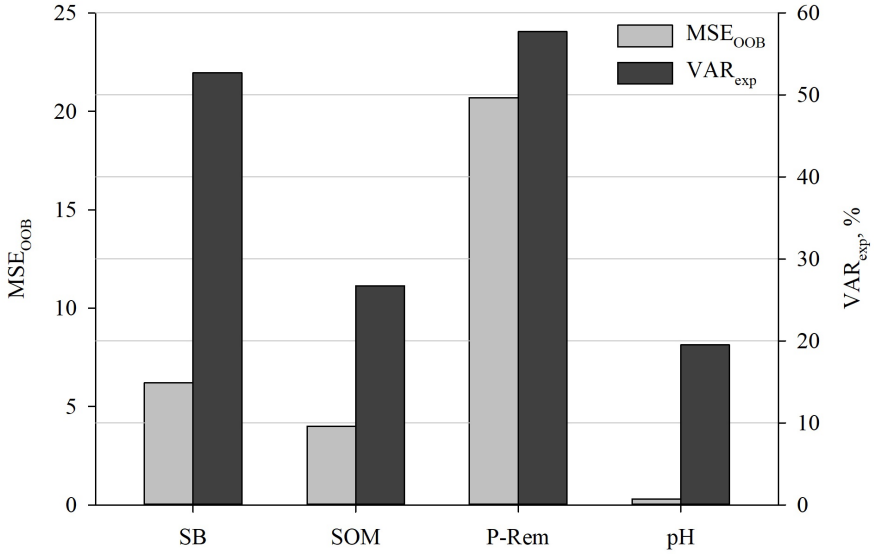

Figure 3. Mean square error (MSE_{OOB}) and variance explained (Var_{exp}) values from Random Forest models trained to predict SB, SOM, P-Rem and pH from soil samples collected at Lavras, Minas Gerais, Brazil.

Analyzing the variables importance for each prediction attribute (Fig. 4) it can be observed that Ca was the most important variable for the SB model. Considering that this variable had also contributed significantly in the SMLR model, this strengthens its importance in SB prediction, independently of the utilized algorithm. Additionally, Cu was also a relevant variable for both SMLR and RF. This might be related to the sorption capacity of soils, as Cu presents high affinity with organic groups and soil minerals (Elbana et al. 2018). Variable contribution was more homogeneous in P-Rem predictions, but Si and Ca were again relevant, along with Ni. This strengthens the relation of P-Rem with soil reactivity. In such context, sandy soils should be highlighted for having low capacity to adsorb phosphates, resulting in higher P-Rem contents (Bornø et al. 2018, Weihrauch & Opp 2018).

Regarding SOM predictions, Al was the most important variable. Such result might be related to the contribution of minerals that contain Al in protecting SOM from degradation through chemical bonding and/or the occlusion of the organic material (Sarkar et al. 2019). Variables Sr and Cr were also considered important in SOM predictions. This may be related to the dynamics of these elements in soil, strongly influenced by soil properties. Availability and mobility of these elements are determined by the content and type of clay mineral and SOM, along with pH and ionic force (Lawson et al. 2016, Choppala et al. 2018).

Strontium is considered mobile in soil, behaving similarly to Ca^{2+}, present in its exchangeable form, especially when there is an increase in its concentration in soils (Smičiklas et al. 2015, Lawson et al. 2016). Thus, the concentration of superficial negative charges at the reactive fractions of soils determines their adsorption capacity. However, decrease in pH, the increase in competition for adsorption sites, SOM oxidation and the flow of water can promote Sr desorption and loss in soils (Smičiklas et al. 2015). Considering that all these processes also influence SOM (Cates et al. 2016, Triberti et al. 2016, Araujo et al. 2017), the differences in Sr contents in soil might be related with SOM contents. Other motive for Sr importance may be related to Ca contents (Lawson et al. 2016). This was observed by

Smičiklas et al. (2015), who stated that Sr also presents affinity to carbonates. Therefore, considering that Ca contents were important to predict these attributes, the relation between Ca and Sr may have contributed for this importance of Sr especially when predicting SB, P-Rem and SOM.

In pH predictions, the Ca variable was not present in the SMLR equation, but was the most important in the RF model. Recall that Ca contents do not have direct relation with pH, but with SB (Indrasumunar et al. 2012, Anikwe et al. 2016). However, subareas where amendments are applied present higher Ca and pH values (Behera & Shukla 2015, Vågen et al. 2016), explaining the importance of this variable.

It is important to recall that variables considered important by models tend to show high correlation with predicted attributes. Therefore, variability in values of such variables should be considered. However, soil use and management alter chemical, physical and biological attributes, and may influence the quality of predictions (Resende et al. 2014, Teixeira et al. 2018, Lopes & Guilherme 2016).

Validation of SMLR and RF models

Analyses of the comparison of R^2, R^2_{adj}, RMSE and ME between observed and predicted data by statistical methods are shown in Figure 5. Differences can be observed between SMLR and RF results, with predictions by RF models showing better performance. This was expected, since decision trees perform non-additive modeling, and work with non-linear relations (Greve et al. 2010). Similar results were obtained by Andrade et al. (2020) when predicting Ca^{2+}, Mg^{2+} and K^{+}. The authors obtained more robust models by utilizing RF in comparison to stepwise generalized linear model (SGLM), with R^2 values increases of 9.23% (Ca^{2+}), 13.73% (K^{+}) and 62.16% (Mg^{2+}).

In this current study, prediction models that presented best results with RF were those for P-Rem and SB, with R^2 scores of 0.66 and 0.55, respectively. Predictions for SOM and pH presented lower R^2 scores (0.47 and 0.26, respectively). This variation in predictive capacity between attributes was also observed by Silva et al. (2017) when utilizing RF with pXRF data obtained from tropical soil samples in Minas Gerais state, Brazil. These authors have also obtained low R^2 scores when predicting de K^{+} (0.39) and potential CEC (0.38), but observed high accuracy for pH (0.89), effective CEC (0.94), Ca^{2+} (0.96) and Al^{3+} (0.86). This also demonstrates the variation of predictive capacity according to the study region and sample variability, which encourages tests in different circumstances.

Models obtained in the current study can be considered of intermediary predictive capacity. However, several works reveal difficulty in predicting SOM and pH, as noted by Zhang & Hartemink (2020). Sharma et al. (2014) obtained intermediary results for pH prediction studying USA soils, with models presenting R^2 scores ranging between 0.43 to 0.80, with latter result being obtained after the addition of several auxiliary variables to the models (sand, silt and clay contents and SOM). Therefore, the incorporation of variables that represent relations with predicted attributes can increment the models accuracy (Sharma et al. 2015). Additionally, other works have shown intermediary results, i.e., R^2 score of 0.47 for CEC prediction via RF (Chagas et al. 2018), and R^2 score of 0.33 for SOC using diffuse reflectance spectroscopy data via RF (Bhering et al. 2016). Both works concluded that the application of RF in the estimation of soil attributes is promising, agreeing with results obtained in this current study.

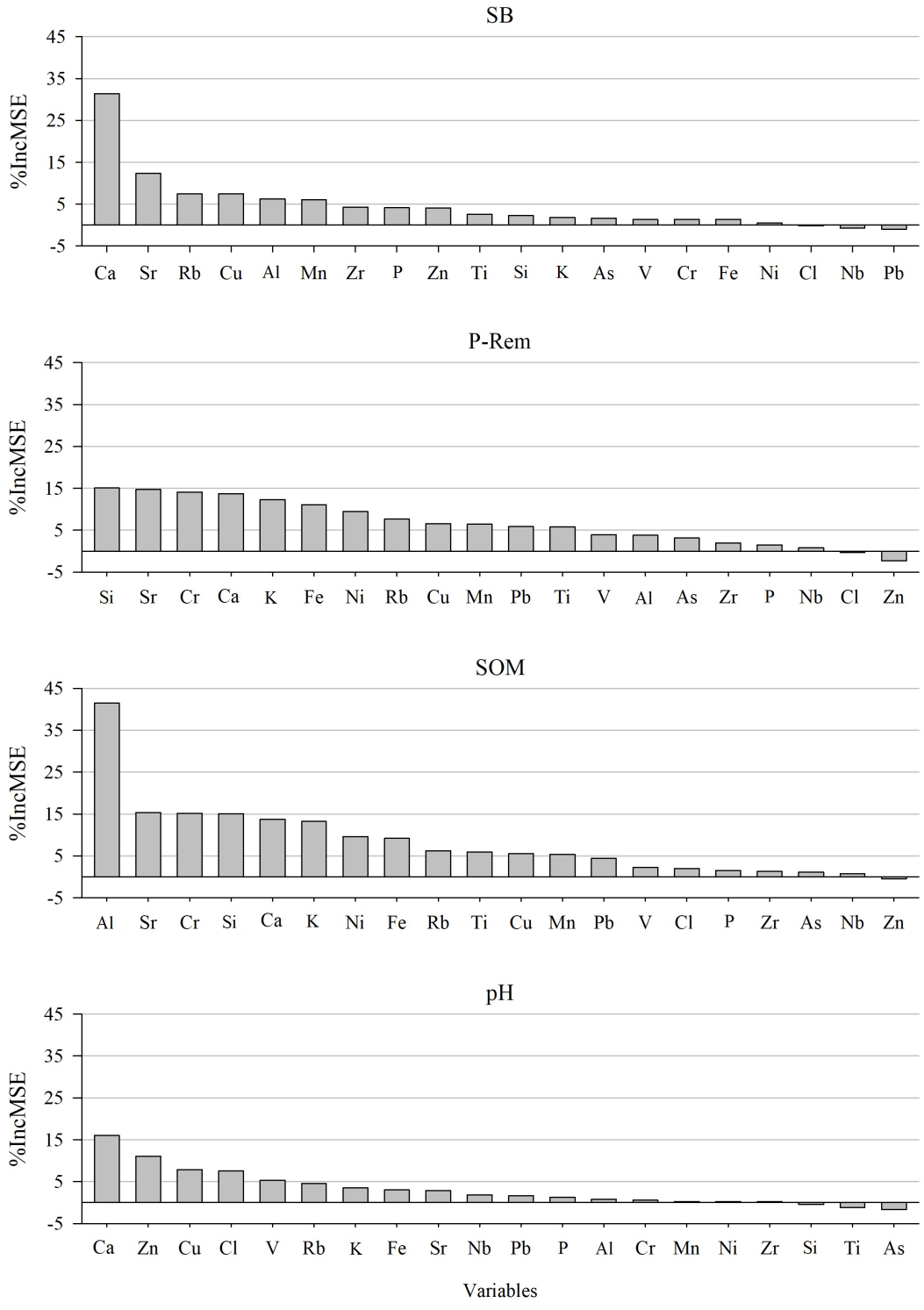

Figure 4. **Most important variables for sum of bases (SB), remaining P (P-Rem), soil organic matter (SOM) and pH predictions via Random Forest from data of samples collected at Lavras, Minas Gerais, Brazil.**

Digital mapping of soil attributes and validation

Models with highest accuracies were utilized for spatial prediction of studied soil attributes and are shown in Figure 6. Analyzing the SOM spatial distribution, it can be observed that lower values are located in pasture subareas with uncovered soil. This happens because such subareas have lower vegetal biomass accumulation and nutrient cycling in comparison to other subareas, which are occupied by eucalyptus plantations or native forests. These different land management systems interfere in the organic matter accumulation in soils (Conceição et al. 2005). However, it is important to notice that some studies demonstrate a positive effect from pasture, with adequate management, in accumulation and conservation of SOM compared to annual plantations (Cates et al. 2016), but this was not observed in the studied area.

Lowest P-Rem values were observed in pasture and coffee plantations, whilst relatively higher values occur under native forests and pinus plantations. SB presented lowest values in pasture, pinus plantations and native forests, gradually increasing annual plantations and reaching its peak contents in coffee plantations. Such variation according to land use occurs due to periodic application of fertilizers and amendments at the cultivated subareas, which increases SB contents (Triberti et al. 2016, Han et al. 2019). In addition, the high values of SOM and pH also contributed to the increase in SB. The presence of SOM contributes to the increase of adsorption sites in the soil, while higher pH values indicate a greater presence of negative charges in these sites, allowing greater adsorption of cations (Sharma et al. 2015, Smičiklas et al. 2015, Bornø et al. 2018). Values for pH showed distribution similar to SB, having its lower values under native forests and forest and annual plantations, being likewise influenced by antropic activities.

The R^2, R^2_{adj}, RMSE and ME score results from spatial predictions of soil attributes were similar to those observed in models validations (Table IV). Best performances were achieved by P-Rem and SB predictions, with R^2 scores of 0.66 and 0.54, respectively, although the values of ME and RMSE were higher. Lowest R^2 values were observed in SOM predictions (0.47), followed by pH (0.28).

Results shown by maps demonstrate the potential and benefits of the utilization of proximal sensors such as pXRF in the assessment of the spatial distribution of soil fertility attributes. These maps offer a simple overview of the variation of studied attributes, allowing for the identification and orientation of practices in each subarea. Thus, management and the application of products can become more efficient, resulting in economic, environmental and social gain.

Table IV. Validation results of the maps generated for soil fertility attributes from samples collected at Lavras, Minas Gerais, Brazil.

	SB	SOM	P-Rem	pH
R^2	0.54	0.47	0.66	0.28
R^2_{adj}	0.52	0.45	0.65	0.24
ME	2.74	2.10	19.85	0.31
RMSE	1.66	1.45	4.45	0.55

SB = $cmol_c\ dm^{-3}$; SOM = $dag\ kg^{-1}$; and P-Rem = $mg\ dm^{-3}$.

The very high spatial variability of soil classes, land uses (Fig. 1) and management practices helps to explain the low scores in validation results, requiring additional care in the application of fertilizers and amendments in soils. One of the recent advancements regarding the use of proximal sensors for predicting

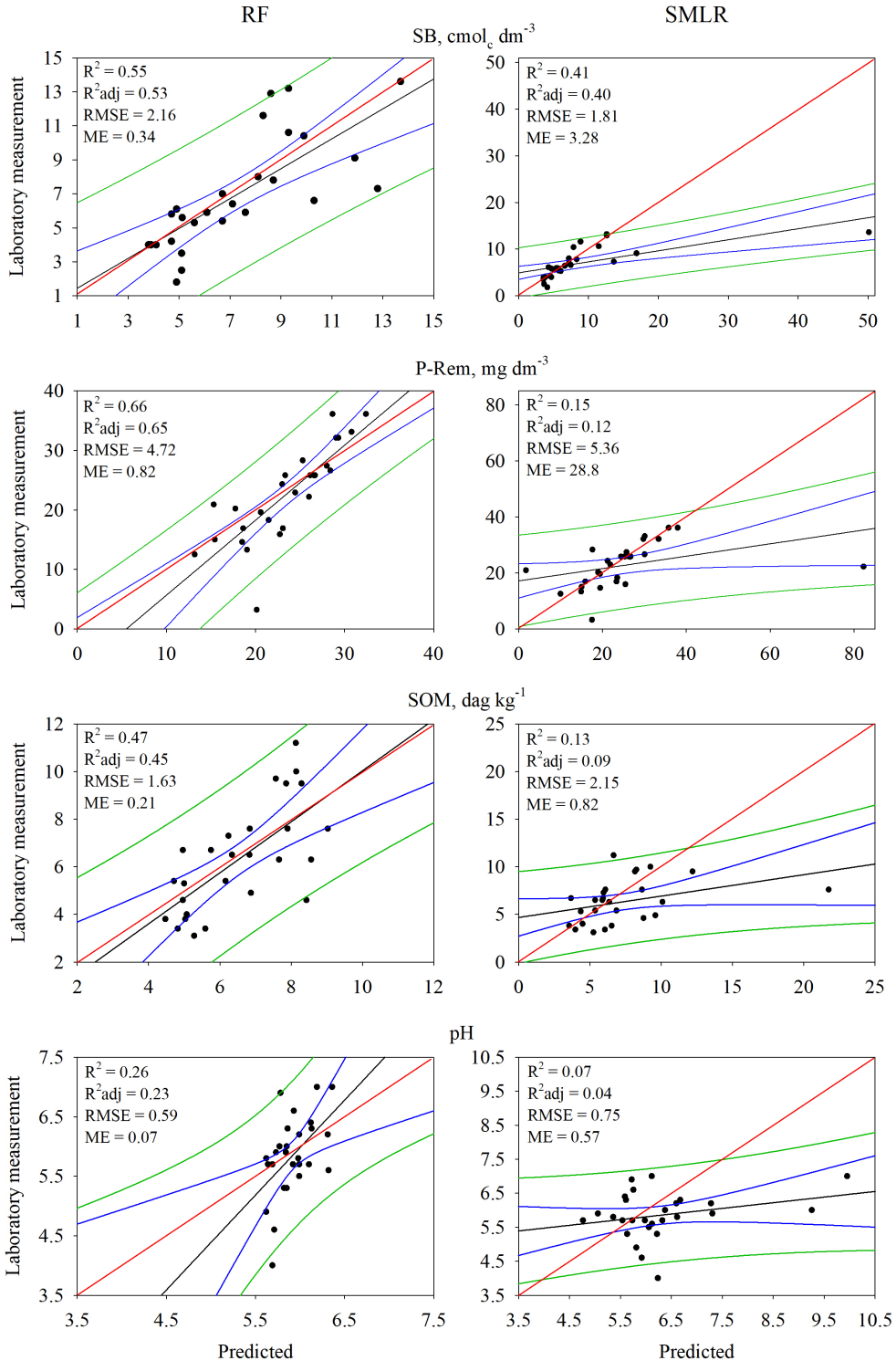

Figure 5. Validation of soil attributes predicted by Random Forest (left) and stepwise multiple linear regression (right) models, from soil samples collected in Lavras, Minas Gerais, Brazil.

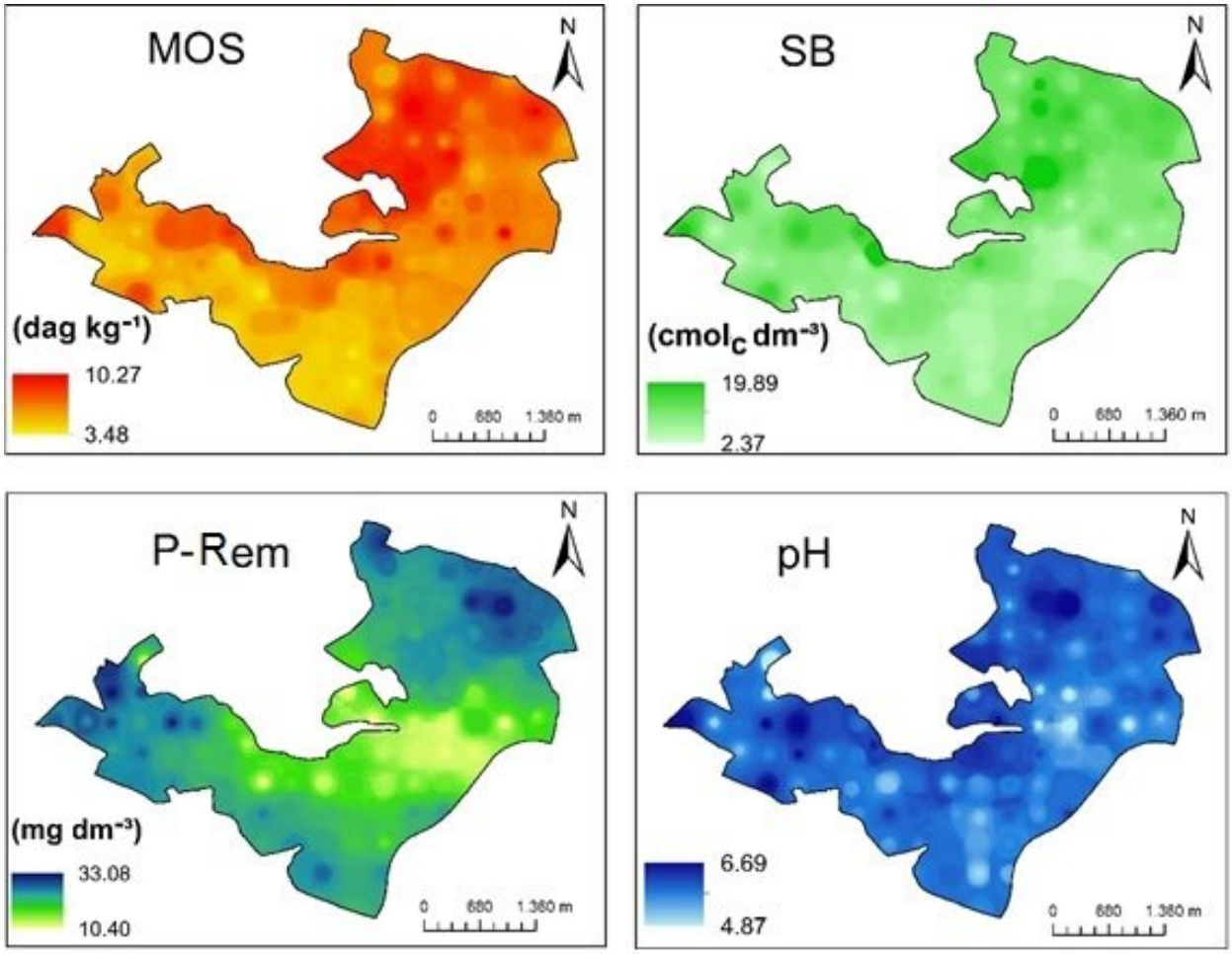

Figure 6. Prediction maps for soil attributes (0-20 cm) via Random Forest from pXRF data obtained from samples collected at Lavras, Minas Gerais, Brazil. SOM: soil organic matter; SB: sum of bases; P-Rem: remaining P.

soil properties is the combination of sensors, e.g., pXRF with visible near infrared diffuse reflectance spectrometer (Vis-NIR DRS) and Nix Pro color sensor, showing considerable increases in prediction accuracy, although these works are still rare in Brazil and in other tropical countries (Benedet et al. 2020b, Wan et al. 2020, Mukhopadhyay et al. 2020). This approach may be very useful in conditions similar to this study to improve the results.

CONCLUSION

Predictions of soil fertility attributes from pXRF data showed higher accuracy when modeled via the RF algorithm compared to SMLR models. In the modeling process, variables Ca, Si and Al were the most relevant for SB, P-Rem and SOM predictions, respectively. Prediction results for pH were less satisfactory.

Analyses performed by pXRF combined with RF modeling were efficient in assessing the spatial variability of soil fertility attributes in the studied tropical area. Such novel method can help improve crop management at reduced cost, which is crucial especially for tropical countries like Brazil, which face a lack of financial resources for such activities. The method presented by this work is

environmentally-friendly, preserves collected samples, and results indicated high efficacy in mapping relevant fertility parameters with greater details and low cost. Although such mapping process is expected to be more difficult in areas with very complex and variable distribution of soil classes, this work demonstrated that this method is sufficiently efficient even in areas with high soil variability.

Thus, this methodology can be applied in the construction of prediction models for soil fertility attributes from spatialized data, allowing for the identification and orientation of adequate management practices specifically for each subarea, even in areas presenting very high spatial variability conditions.

Acknowledgments

The authors would like to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the financial support to develop this research.

REFERENCES

ALVARES CA, STAPE JL, SENTELHAS PC, DE MORAES GONÇALVES JL & SPAROVEK G. 2013. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift 22(6): 711-728. doi:10.1127/0941-2948/2013/0507. URL http://dx.doi.org/10.1127/0941-2948/2013/0507.

ALVAREZ VVH, NOVAIS RF, DIAS LE & OLIVEIRA JA. 2000. Determinação e uso do fósforo remanescente. Boletim Informativo da Sociedade Brasileira de Ciência do Solo 25(1): 27-32.

ANDRADE R, FARIA WM, SILVA SHG, CHAKRABORTY S, WEINDORF DC, MESQUITA LF, GUILHERME LRG & CURI N. 2020. Prediction of soil fertility via portable X-ray fluorescence (pXRF) spectrometry and soil texture in the Brazilian Coastal Plains. Geoderma 357: 113960. doi:https://doi.org/10.1016/j.geoderma.2019.113960. URL https://www.sciencedirect.com/science/article/pii/S0016706119315198.

ANIKWE M, EZE J & IBUDIALO A. 2016. Influence of lime and gypsum application on soil properties and yield of cassava (Manihot esculenta Crantz.) in a degraded Ultisol in Agbani, Enugu Southeastern Nigeria. Soil and Tillage Res 158: 32-38. URL https://www.sciencedirect.com/science/article/pii/S0167198715300489.

ARAUJO MA, ZINN YL & LAL R. 2017. Soil parent material, texture and oxide contents have little effect on soil organic carbon retention in tropical highlands. Geoderma 300: 1-10. doi:https://doi.org/10.1016/j.geoderma.2017.04.006. URL https://www.sciencedirect.com/science/article/pii/S0016706117301076.

BATISTA AH, MELO VF, GILKES R & ROBERTS M. 2018. Identification of Heavy Metals in Crystals of Sand and Silt Fractions of Soils by Scanning Electron Microscopy (SEM EDS/WD-EPMA). Rev Bras Cienc Solo 42: e0170174. URL https://www.scielo.br/j/rbcs/a/zJmSgPDXhRy8vhf5YRQHqjQ/?lang=en.

BEHERA SK & SHUKLA AK. 2015. Spatial Distribution of Surface Soil Acidity, Electrical Conductivity, Soil Organic Carbon Content and Exchangeable Potassium, Calcium and Magnesium in Some Cropped Acid Soils of India. Land Degrad Dev 26(1): 71-79. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ldr.2306.

BENEDET L, FARIA WM, SILVA SHG, MANCINI M, DEMATTÊ JAM, GUILHERME LRG & CURI N. 2020a. Soil texture prediction using portable X-ray fluorescence spectrometry and visible near-infrared diffuse reflectance spectroscopy. Geoderma 376: 114553. URL https://www.sciencedirect.com/science/article/pii/S0016706119330447.

BENEDET L, FARIA WM, SILVA SHG, MANCINI M, GUILHERME LRG, DEMATTÊ JAM & CURI N. 2020b. Soil subgroup prediction via portable X-ray fluorescence and visible near-infrared spectroscopy. Geoderma 365: 114212. URL https://www.sciencedirect.com/science/article/pii/S0016706119324826.

BHERING SB, DA SILVA CHAGAS C, DE CARVALHO JUNIOR W, PEREIRA NR, FILHO BC & PINHEIRO HSK. 2016. Mapeamento digital de areia, argila e carbono orgânico por modelos Random Forest sob diferentes resoluções espaciais. Pesquisa Agropecuária Brasileira 51(9): 1359-1370.

BORNØ ML, MÜLLER-STÖVER DS & LIU F. 2018. Contrasting effects of biochar on phosphorus dynamics and bioavailability in different soil types. Sci Total Environ 627: 963-974. URL https://www.sciencedirect.com/science/article/pii/S0048969718303255.

BREIMAN L. 2001. Random Forests. March Learn 45(1): 5-32. URL https://link.springer.com/article/10.1023/A:1010933404324.

CATES AM, RUARK MD, HEDTCKE JL & POSNER JL. 2016. Long-term tillage, rotation and perennialization effects on particulate and aggregate soil organic matter. Soil and Tillage Res 155: 371-380. URL https://www.sciencedirect.com/science/article/pii/S0167198715300258.

CHAGAS CDS, JÚNIOR WDC, PINHEIRO HSK, XAVIER PAM, BHERING SB, PEREIRA NR & FILHO BC. 2018. Mapping Soil Cation Exchange Capacity in a Semiarid Region through Predictive Models and Covariates from Remote Sensing Data. Rev Bras Cienc Solo 42: e0170183.

CHOPPALA G, KUNHIKRISHNAN A, SESHADRI B, PARK JH, BUSH R & BOLAN N. 2018. Comparative sorption of chromium species as influenced by pH, surface charge and organic matter content in contaminated soils. J Geochem Explor 184: 255-260. URL https://www.sciencedirect.com/science/article/pii/S0375674216301558. Remediation of Polluted Soils - Part 3.

CONCEIÇÃO PC, AMADO TJC, MIELNICZUK J & SPAGNOLLO E. 2005. Qualidade do solo em sistemas de manejo avaliada pela dinâmica da matéria orgânica e atributos relacionados. Rev Bras Cienc Solo 29(5): 777-788.

CURI N, SILVA SHG, POGGERE GC & MENEZES MD. 2017. Mapeamento de Solos e Magnetismo no Campus da UFLA Como Traçadores Ambientais. 1st ed. Lavras: UFLA, 147 p.

DUDA BM, WEINDORF DC, CHAKRABORTY S, LI B, MAN T, PAULETTE L & DEB S. 2017. Soil characterization across catenas via advanced proximal sensors. Geoderma 298: 78-91. URL https://www.sciencedirect.com/science/article/pii/S0016706116308849.

ELBANA TA, MAGDI SELIM H, AKRAMI N, NEWMAN A, SHAHEEN SM & RINKLEBE J. 2018. Freundlich sorption parameters for cadmium, copper, nickel, lead, and zinc for different soils: Influence of kinetics. Geoderma 324: 80-88. URL https://www.sciencedirect.com/science/article/pii/S0016706117321468.

FANG H, CUI Z, HE G, HUANG L & CHEN M. 2017. Phosphorus adsorption onto clay minerals and iron oxide with consideration of heterogeneous particle morphology. Sci Total Environ 605-606: 357-367. URL https://www.sciencedirect.com/science/article/pii/S0048969717312330.

FERREIRA E, DANTAS AAA, OLIVEIRA A & MACHADO RV. 2013. Land use and land cover dynamics on the campus of Federal University of Lavras from 1964 to 2009. Cerne 19(1): 35-42. URL https://www.scielo.br/j/cerne/a/wbG5kh4G3h4k6rfvKVKgQqf/?lang=pt.

GREVE M, GREVE M, KHEIR R, BØCHER P, LARSEN R & MCCLOY K. 2010. Comparing Decision Tree Modeling and Indicator Kriging for Mapping the Extent of Organic Soils in Denmark. In: Digital Soil Mapping 357: 267-280. doi:10.1007/978-90-481-8863-5_22.

GUPTA VV & GERMIDA JJ. 2015. Soil aggregation: Influence on microbial biomass and implications for biological processes. Soil Biol Biochem 80: A3-A9. URL https://www.sciencedirect.com/science/article/pii/S0038071714003010.

HAN T, CAI A, LIU K, HUANG J, WANG B, LI D, QASWAR M, FENG G & ZHANG H. 2019. The links between potassium availability and soil exchangeable calcium, magnesium, and aluminum are mediated by lime in acidic soil. J Soils Sediments 19(3): 1382-1392. doi:https://doi.org/10.1007/s11368-018-2145-6.

HARTEMINK A, ZHANG Y, BOCKHEIM J, CURI N, SILVA S, GRAUER-GRAY J, LOWE D & KRASILNIKOV P. 2020. Chapter Three - Soil horizon variation: A review. In: Sparks DL (Ed), Advances in Agronomy 160: 125-185. Academic Press. URL https://www.sciencedirect.com/science/article/pii/S0065211319301087.

HEDLEY C. 2015. The role of precision agriculture for improved nutrient management on farms. J Sci Food Agric 95(1): 12-19. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jsfa.6734.

HEUNG B, BULMER CE & SCHMIDT MG. 2014. Predictive soil parent material mapping at a regional-scale: A Random Forest approach. Geoderma 214-215: 141-154. URL https://www.sciencedirect.com/science/article/pii/S0016706113003443.

HONG S, PIAO S, CHEN A, LIU Y, LIU L, PENG S, SARDANS J, SUN Y, PEÑUELAS J & ZENG H. 2018. Afforestation neutralizes soil pH. Nat Commun 9(1): 520.

HU B, CHEN S, HU J, XIA F, XU J, LI Y & SHI Z. 2017. Application of portable XRF and VNIR sensors for rapid assessment of soil heavy metal pollution. PLoS ONE 12(2): e0172438.

IMTIAZ M, RIZWAN MS, XIONG S, LI H, ASHRAF M, SHAHZAD SM, SHAHZAD M, RIZWAN M & TU S. 2015. Vanadium, recent advancements and research prospects: A review. Environ Int 80: 79-88. URL https://www.sciencedirect.com/science/article/pii/S0160412015000835.

INDRASUMUNAR A, MENZIES NW & DART PJ. 2012. Calcium affects the competitiveness of acid-sensitive and acid-tolerant strains of Bradyrhizobium japonicum

in nodulating and fixing nitrogen with two soybean cultivars in acid soil. Soil Biol Biochem 46: 115-122. URL https://www.sciencedirect.com/science/article/pii/S0038071711004093.

KHORSHIDI M & LU N. 2017. Quantification of Exchangeable Cations Using Soil Water Retention Curve. J Geotech Geoenviron Eng 143(9): 04017057.

LAWSON LS, MCCOMB JQ, DONG R, HAN FX, ROGER C, ARSLAN Z & YU H. 2016. Binding, fractionation, and distribution of Cs, Co, and Sr in a US coastal soil under saturated and field capacity moisture regimes. J Soils Sediments 16(2): 497-508.

LIAW A & WIENER M. 2018. randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4.6–14. URL https://cran.r-project.org/web/packages/randomForest/randomForest.pdf. Accessed Jul. 2, 2021.

LOPES A & GUILHERME LG. 2016. Chapter One - A Career Perspective on Soil Management in the Cerrado Region of Brazil. In: Sparks DL (Ed), Advances in Agronomy 137: 1-71. Academic Press. URL https://www.sciencedirect.com/science/article/pii/S0065211315300043.

MANCINI M, WEINDORF DC, SILVA SHG, CHAKRABORTY S, DOS SANTOS TEIXEIRA AF, GUILHERME LRG & CURI N. 2019. Parent material distribution mapping from tropical soils data via machine learning and portable X-ray fluorescence (pXRF) spectrometry in Brazil. Geoderma 354: 113885. URL https://www.sciencedirect.com/science/article/pii/S0016706119307888.

MCLEAN EO, HEDDLESON MR, BARTLETT RJ & HOLOWAYCHUK N. 1958. Aluminum in Soils: I. Extraction Methods and Magnitudes in Clays and Ohio Soils. Soil Sci Soc Am J 22(5): 382-387. URL https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj1958.03615995002200050005x.

MEHLICH A. 1984. Mehlich 3 soil test extractant: A modification of Mehlich 2 extractant. Comm Soil Sci Plant Anal 15(12): 1409-1416. doi:10.1080/00103628409367568.

MUKHOPADHYAY S, CHAKRABORTY S, BHADORIA P, LI B & WEINDORF DC. 2020. Assessment of heavy metal and soil organic carbon by portable X-ray fluorescence spectrometry and NixPro™ sensor in landfill soils of India. Geoderma Reg 20: e00249. URL https://www.sciencedirect.com/science/article/pii/S2352009419302482.

MULLA DJ. 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosyst Eng 114(4): 358-371. URL https://www.sciencedirect.com/science/article/pii/S1537511012001419. Special Issue: Sensing Technologies for Sustainable Agriculture.

QU M, CHEN J, LI W, ZHANG C, WAN M, HUANG B & ZHAO Y. 2019. Correction of in-situ portable X-ray fluorescence (PXRF) data of soil heavy metal for enhancing spatial prediction. Environ Pollut 254: 112993. URL https://www.sciencedirect.com/science/article/pii/S0269749119322766.

RESENDE M, CURI N, REZENDE SB, CORRÊA GF & KER JC. 2014. Pedologia: Base para distinção de ambientes. 6th ed. Lavras: UFLA, 378 p.

RIBEIRO BT, LIMA JM, GUILHERME LRG & JULIÃO LGF. 2010. Lead sorption and leaching from an Inceptisol sample amended with sugarcane vinasse. Sci Agric 67(4): 441-447. URL https://www.scielo.br/j/sa/a/QB9x56Yzc9mXC4DmwmDxVfR/?lang=en.

RIBEIRO BT, SILVA SHG, SILVA EA & GUILHERME LRG. 2017. Portable X-ray fluorescence (pXRF) applications in tropical Soil Science. Ciênc Agrotec 41(3): 245-254. URL https://www.scielo.br/j/cagro/a/hdNkqprycBmk4SrwdqQMMfN/abstract/?lang=en.

SARKAR I, KHAN MZ & HANIF M. 2019. Soil organic fractions in cultivated and uncultivated soils of Costal area in Bangladesh. J Agric Chem Environ 8(3): 129-144.

SHARMA A, WEINDORF DC, MAN T, ALDABAA AAA & CHAKRABORTY S. 2014. Characterizing soils via portable X-ray fluorescence spectrometer: 3. Soil reaction (pH). Geoderma 232-234: 141-147. URL https://www.sciencedirect.com/science/article/pii/S0016706114002018.

SHARMA A, WEINDORF DC, WANG D & CHAKRABORTY S. 2015. Characterizing soils via portable X-ray fluorescence spectrometer: 4. Cation exchange capacity (CEC). Geoderma 239-240: 130-134. URL https://www.sciencedirect.com/science/article/pii/S0016706114003681.

SILVA EA, WEINDORF DC, SILVA SH, RIBEIRO BT, POGGERE GC, CARVALHO TS, GONÇALVES MG, GUILHERME LR & CURI N. 2019. Advances in Tropical Soil Characterization via Portable X-Ray Fluorescence Spectrometry. Pedosphere 29(4): 468-482. URL https://www.sciencedirect.com/science/article/pii/S1002016019608155.

SILVA SHG, HARTEMINK AE, DOS SANTOS TEIXEIRA AF, INDA AV, GUILHERME LRG & CURI N. 2018a. Soil weathering analysis using a portable X-ray fluorescence (PXRF) spectrometer in an Inceptisol from the Brazilian Cerrado. Appl Clay Sci 162: 27-37. URL https://www.sciencedirect.com/science/article/pii/S0169131718302412.

SILVA SHG, SILVA EA, POGGERE GC, GUILHERME LRG & CURI N. 2018b. Tropical soils characterization at low cost and time using portable X-ray fluorescence spectrometer (pXRF): Effects of different sample preparation methods. Ciênc Agrotec 42(1): 80-92. URL https://www.scielo.br/j/cagro/a/3kZZPw6CsQqr6cNC6fVwg8z/?lang=en.

SILVA SHG, TEIXEIRA AFS, MENEZES MD, GUILHERME LRG, MOREIRA FMS & CURI N. 2017. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). Ciênc Agrotec 41(6): 648-664. URL https://www.scielo.br/j/cagro/a/TGDkNHPxRLNkHV3mRvzP7FS/abstract/?lang=en.

SILVA SHG ET AL. 2020. Soil texture prediction in tropical soils: A portable X-ray fluorescence spectrometry approach. Geoderma 362: 114136. URL https://www.sciencedirect.com/science/article/pii/S0016706119301612.

SILVA SHG ET AL. 2021. pXRF in tropical soils: Methodology, applications, achievements and challenges. In: Sparks DL (Ed), Advances in Agronomy 167: 1-62. Academic Press. URL https://www.sciencedirect.com/science/article/pii/S006521132030105X.

SMIČIKLAS I, DIMOVIĆ S, JOVIĆ M, MILENKOVIĆ A & ŠLJIVIĆ IVANOVIĆ M. 2015. Evaluation study of cobalt(II) and strontium(II) sorption–desorption behavior for selection of soil remediation technology. Int J Environ Sci Technol 12(12): 3853-3862.

SUN W, ZHANG X, SUN X, SUN Y & CEN Y. 2018. Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals. Geoderma 327: 25-35. URL https://www.sciencedirect.com/science/article/pii/S0016706117310510.

TABACHNICK BG & FIDELL LS. 1996. Using Multivariat Statistics. 3rd ed. Northridge: HarperCollins College, 880 p.

TEIXEIRA AFS, WEINDORF DC, SILVA SHG, GUILHERME LRG & CURI N. 2018. Portable X-ray fluorescence (pXRF) spectrometry applied to the prediction of chemical attributes in Inceptisols under different land uses. Ciênc Agrotec 42(5): 501-512. URL https://www.scielo.br/j/cagro/a/gSJDQjpNzWxWrFRMRHQwgNv/?lang=en.

TRIBERTI L, NASTRI A & BALDONI G. 2016. Long-term effects of crop rotation, manure and mineral fertilisation on carbon sequestration and soil fertility. Eur J Agron 74: 47-55. URL https://www.sciencedirect.com/science/article/pii/S1161030115300721.

VÅGEN TG, WINOWIECKI LA, TONDOH JE, DESTA LT & GUMBRICHT T. 2016. Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. Geoderma 263: 216-225. URL https://www.sciencedirect.com/science/article/pii/S0016706115300082.

WALKLEY A & BLACK IA. 1934. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci 37(1): 29-38.

WAN M, HU W, QU M, LI W, ZHANG C, KANG J, HONG Y, CHEN Y & HUANG B. 2020. Rapid estimation of soil cation exchange capacity through sensor data fusion of portable XRF spectrometry and Vis-NIR spectroscopy. Geoderma 363: 114163. URL https://www.sciencedirect.com/science/article/pii/S0016706119314739.

WEIHRAUCH C & OPP C. 2018. Ecologically relevant phosphorus pools in soils and their dynamics: The story so far. Geoderma 325: 183-194. URL https://www.sciencedirect.com/science/article/pii/S0016706117314076.

WEINDORF DC, BAKR N & ZHU Y. 2014. Chapter One - Advances in Portable X-ray Fluorescence (PXRF) for Environmental, Pedological, and Agronomic Applications. In: Sparks DL (Ed), Advances in Agronomy 128: 1-45. Academic Press. URL https://www.sciencedirect.com/science/article/pii/B9780128021392000019.

WEINDORF DC, CHAKRABORTY S, HERRERO J, LI B, CASTAÑEDA C & CHOUDHURY A. 2016. Simultaneous assessment of key properties of arid soil by combined PXRF and Vis–NIR data. Eur J Soil Sci 67(2): 173-183. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ejss.12320.

WEINDORF DC, HERRERO J, CASTAÑEDA C, BAKR N & SWANHART S. 2013. Direct Soil Gypsum Quantification via Portable X-Ray Fluorescence Spectrometry. Soil Sci Soc Am J 77(6): 2071-2077. URL https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj2013.05.0170.

WIESMEIER M, BARTHOLD F, BLANK B & KÖGEL-KNABNER I. 2011. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. Plant Soil 430(1): 7-24.

ZHANG Y & HARTEMINK AE. 2020. Data fusion of vis–NIR and PXRF spectra to predict soil physical and chemical properties. Eur J Soil Sci 71(3): 316-333. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ejss.12875.

ZHU Y, WEINDORF DC & ZHANG W. 2011. Characterizing soils using a portable X-ray fluorescence spectrometer: 1. Soil texture. Geoderma 167-168: 167-177. URL

https://www.sciencedirect.com/science/article/pii/S0016706111002527.

How to cite
BENEDET L, NILSSON MS, SILVA SHG, PELEGRINO MHP, MANCINI M, MENEZES MD, GUILERME LRG & CURI N. 2021. X-ray fluorescence spectrometry applied to digital mapping of soil fertility attributes in tropical region with elevated spatial variability. An Acad Bras Cienc 93: e20200646. DOI 10.1590/0001-3765202120200646.

Manuscript received on April 28, 2020;
accepted for publication on September 21, 2020

LUCAS BENEDET
https://orcid.org/0000-0002-0560-8790

MATHEUS S. NILSSON
https://orcid.org/0000-0001-8703-2336

SÉRGIO HENRIQUE G. SILVA
https://orcid.org/0000-0003-2750-5976

MARCELO H.P. PELEGRINO
https://orcid.org/0000-0002-0963-1189

MARCELO MANCINI
https://orcid.org/0000-0003-4118-7943

MICHELE D. DE MENEZES
https://orcid.org/0000-0001-7052-8327

LUIZ ROBERTO G. GUILHERME
https://orcid.org/0000-0002-5387-6028

NILTON CURI
https://orcid.org/0000-0002-2604-0866

Universidade Federal de Lavras, Departamento de Ciência do Solo, Caixa Postal 3037, 37200-900 Lavras, MG, Brazil

Correspondence to: **Nilton Curi**
E-mail: niltcuri@ufla.br

Author contributions

Lucas Benedet wrote the manuscript, Matheus Sterzo Nilsson, Marcelo Henrique Procópio Pelegrino and Marcelo Macnini collected the soil samples, built the prediction models and ran the statistical analyses, Sérgio Henrique Godinho Silva idealized and supervised the research, collected the soil samples and contributed to the discussion and review of the text, Michele Duarte de Menezes collected the soil samples and contributed to the text review, Nilton Curi and Luiz Roberto Guimarães Guilherme supervised the research and contributed to the text review.