LUIZA MARIA PEREIRA PIERANGELI

PREDICTION OF SOIL ATTRIBUTES VIA PXRF SPECTROMETRY, MAGNETIC SUSCEPTIBILITY, AND TERRAIN ATTRIBUTES IN A HIGHLY HETEROGENEOUS TROPICAL AREA

LAVRAS – MG
2020

LUIZA MARIA PEREIRA PIERANGELI

PREDICTION OF SOIL ATTRIBUTES VIA PXRF SPECTROMETRY, MAGNETIC SUSCEPTIBILITY, AND TERRAIN ATTRIBUTES IN A HIGHLY HETEROGENEOUS TROPICAL AREA

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Ciência do Solo, área de concentração em Recursos Ambientais e Uso da Terra, para a obtenção do título de Mestre.

Prof. Dr. Sérgio Henrique Godinho Silva
Orientador

LAVRAS – MG
2020

LUIZA MARIA PEREIRA PIERANGELI

PREDICTION OF SOIL ATTRIBUTES VIA PXRF SPECTROMETRY, MAGNETIC SUSCEPTIBILITY, AND TERRAIN ATTRIBUTES IN A HIGHLY HETEROGENEOUS TROPICAL AREA

> Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Ciência do Solo, área de concentração em Recursos Ambientais e Uso da Terra, para a obtenção do título de Mestre.

APROVADO em 31 de janeiro de 2020.
Dr. Sérgio Henrique Godinho Silva   UFLA
Dr. Nilton Curi   UFLA
Dr. Julierme Zimmer Barbosa   Instituto Federal do Sudeste de MG

Prof. Dr. Sérgio Henrique Godinho Silva
Orientador

LAVRAS – MG
2020
AGRADECIMENTOS

ABSTRACT

Digital elevation models (DEM) and their derived variables, terrain attributes (TA), are commonly used in soil mapping. The use of proximal sensors, such as portable X-ray fluorescence spectrometer (pXRF) and susceptibilimeter, which determines magnetic susceptibility (MS), provides additional information that has improved the results obtained using only TAs. This work is composed of two chapters, whose studies were conducted at the Palmital Experimental Farm, belonging to the Federal University of Lavras (UFLA). The chapters are related to the use of proximal sensors in conjunction with TA for the prediction of physical and chemical attributes of soils. The first chapter contemplates the use of two proximal sensors, pXRF and MS, together with TA for the prediction of clay, silt, and sand contents through the random forest algorithm. The second chapter discusses the use of pXRF and MS in conjunction with TA in predicting available contents of B, Cu, Fe, Mn, and Zn. The maps were generated for the Palmital farm and validated for each predicted attribute, comparing the efficiency of each model. For the prediction of clay, silt, and sand, all models used the information acquired by pXRF in the final models. On the other hand, for the prediction of B and Zn, only the TA information was sufficient to achieve satisfactory R^2 values. Clay and sand showed moderate accuracy, while silt showed low accuracy. For the prediction of chemical attributes, Cu, Fe, Mn, and Zn presented high to moderate accuracy. However, B reached low accuracy. This shows that pXRF is a powerful tool to assist in the accurate prediction of some soil attributes in a punctual and spatial way, contributing to the digital soil mapping.

Keywords: Micronutrients. Granulometric Fractions. Random Forest. Proximal Sensing. Pedometric.

RESUMO

Modelos digitais de elevação (MDE) e seus derivativos, atributos de terreno (AT), são comumente utilizados no mapeamento de solos. O uso de sensores proximais, como espectrômetro de florescência de raios-X portátil (pXRF) e suscetibilímetro, que determina a susceptibilidade magnética (SM), fornece informações adicionais que têm melhorado os resultados obtidos utilizando apenas ATs. Esta dissertação é composta por dois capítulos, cujo estudo foi realizado na Fazenda experimental Palmital, pertencente a Universidade Federal de Lavras (UFLA). Os capítulos estão relacionados ao uso de sensores proximais em conjunto a AT na predição de atributos físicos e químicos do solo. O primeiro capítulo contempla o uso de dois sensores proximais, pXRF e SM, em conjunto com AT para a predição de argila, silte e areia através do algoritmo random forest. O segundo capítulo aborda o uso do pXRF e SM em conjunto com AT na predição dos teores disponíveis de B, Cu, Fe, Mn e Zn. Os mapas foram gerados para a fazenda Palmital e validados para cada atributo predito, comparando-se a eficiência de cada modelo. Para a predição de argila, silte e areia todos os modelos utilizaram as informações adquiridas pelo pXRF nos modelos finais. Porém, para a predição de B e Zn, apenas as informações de AT foram suficientes para alcançar valores de R^2 satisfatórios. Argila e areia apresentaram moderada acurácia enquanto silte apresentou baixa acurácia. Já para a predição dos atributos químicos, Cu, Fe, Mn e Zn apresentaram entre alta a moderada acurácia, entretanto B alcançou baixa acurácia. Isso mostra que o pXRF é uma ferramenta poderosa para auxiliar na predição acurada de alguns atributos do solo de forma pontual e espacial, contribuindo para o mapeamento digital de solos.

Palavras-chave: Micronutrientes. Frações Granulométricas. Random Forest. Sensores Próximos. Pedometria.

SUMÁRIO

FIRST PART
DIGITAL SOIL MAPPING AND RANDOM FOREST

1   INTRODUCTION

Digital soil mapping (DSM) is a fundamental activity for detailed spatial information on soils, which is essential for urban, agricultural, and environmental planning (BONFATTI et al., 2018; DHARUMARAJAN; HEGDE; SINGH, 2017; KESKIN; GRUNWALD, 2018; MCBRATNEY; MENDONÇA SANTOS; MINASNY, 2003; PRASUHN et al., 2013). Most of the soil maps found in Brazil are at small scales (MENDONÇA-SANTOS; SANTOS, 2007; MENEZES et al., 2016; SANTOS et al., 2014a; SILVA et al., 2016a), which makes it challenging to use them for specific purposes such as agricultural management (OBADE, 2019), soil and water conservation (SANTOS et al., 2014a; SÖDERSTRÖM et al., 2016), among others.

DSM is the union of three factors: the intake in the form of field and laboratory observational methods, the process used in terms of spatial and non-spatial soil inference systems, and the output in the form of spatial soil information systems (LAGACHERIE; MCBRATNEY, 2006; MINASNY; MCBRATNEY, 2016). Furthermore, DSM uses a range of proximal and remote sensing tools to gather and analyze data with the aid of powerful algorithms to model and predict soil data (CHAKRABORTY et al., 2019b; MENEZES et al., 2016; PELEGRINO et al., 2019; SILVA et al., 2016b, 2017a).

The use of proximal and remote sensors combined with DSM allows an fast evaluation of soil characteristics at low cost and without residues production (SILVA et al., 2017a). Some examples are the portable X-ray fluorescence (pXRF) scanners, used to identify and quantify chemical elements present in different materials, including soil (RIBEIRO et al., 2017; SILVA et al., 2016a), and the susceptibilimeter, which quantify the magnetic susceptibility (MS) of different materials (CAMARGO et al., 2018; CERVI et al., 2014; SILVA et al., 2016a; SIQUEIRA et al., 2015). The pXRF sensors are able to detect many elements of the Periodic Table since each element has unique fluorescence energy. pXRF has the advantage of assessing total elemental contents in the soil in a non-destructive manner (CHAKRABORTY et al., 2019b; SILVA et al., 2016a; STOCKMANN et al., 2016b). MS is a useful technique for predicting magnetic soil oxides since different iron forms and dynamics reflect different soil-forming factors and processes (MAHER, 1986; SILVA et al., 2016a). Also, some elements can be affected by iron oxides (CAMARGO et al., 2018) such as As, P, and Zn.

The use of pXRF and MS information can improve the prediction and modeling of soil attributes. Thus, this dissertation was divided into two chapters, the study of which was carried out at Fazenda Experimental Palmital, belonging to the Federal University of Lavras (UFLA). The chapters are related to the use of proximal sensors in conjunction with TA in predicting the physical and chemical attributes of the soil. The first chapter contemplates the use of two proximal sensors, pXRF, and MS, together with TA for the prediction of clay, silt, and sand through the random forest algorithm. The second chapter discusses the use of pXRF and MS in combination with TA in predicting the available levels of B, Cu, Fe, Mn, and Zn.

2 THEORETICAL REFERENCE

2.1 Soil mapping

Soil surveys are fundamental sources of information for land use determination, such as agricultural management, soil conservation, environmental protection, among others (AMADO et al., 2009; BESKOW et al., 2009; SAYÃO et al., 2018; WEINDORF; BAKR; ZHU, 2014). Nevertheless, in Brazil, most existing maps were made at small scales (GIASSON et al., 2011; LAGACHERIE; MCBRATNEY, 2006; MCBRATNEY; MENDONÇA SANTOS; MINASNY, 2003; MENDONÇA-SANTOS; SANTOS, 2007; MENEZES et al., 2013; MINASNY; MCBRATNEY, 2016; SILVA et al., 2016a, 2016b, 2016c). Detailed and semi-detailed soil surveys are in general available in small areas to support local-specific agricultural and environmental projects (MENDONÇA-SANTOS; SANTOS, 2007), e.g., at the level of watersheds (MENEZES et al., 2016; SANTOS et al., 2014b; SILVA et al., 2014, 2016b, 2016c).

The lack of investment coupled with the vast area of the country with difficult access by road (SILVA et al., 2016a), combined with a shortage of materials and humans' resources and insufficient institutional support compelled the country to opt for the execution of generalized small-scale surveys (MENDONÇA-SANTOS; SANTOS, 2007), although more recently, the National Soils Program (PRONASOLOS) has been created to enhance Brazilian soils information. Therefore, the introduction of digital technologies have provided new opportunities to predict soil properties and classes (CAMARGO et al., 2018; CHAKRABORTY et al., 2019b; MENEZES et al., 2013; PELEGRINO et al., 2019; SILVA et al., 2016b; STOCKMANN et al., 2016a; TERRA; DEMATTÊ; ROSSEL, 2018), by means of remote and proximal sensing, increasing of computer processing speed, management of spatial data, and quantitative methods to describe soil patterns and processes

(CHAKRABORTY et al., 2019a; GRUNWALD, 2009; MENEZES et al., 2016; PELEGRINO et al., 2019; SILVA et al., 2014, 2017a, 2018a; TEIXEIRA et al., 2018b). In short, digital soil mapping (DSM) may improve the overall process of soil mapping since DSM is based on the generation of information systems, which allows establishing mathematical relationships between environmental and soil variables and, as a result, to predict the spatial distribution of soil classes and properties (GIASSON et al., 2013; MENEZES et al., 2013; SILVA et al., 2014).

HUDSON (1992) pointed out that soil survey is a science-based on the fact that soils are natural bodies distributed in a predictable way and in response to a systematic interaction of environmental factors. Soil classes and properties are ordinarily and spatially distributed in a foreseeable pattern due to the soil-landscape relationship. This relationship is the answer to the water's movement throughout the landscape, which goes through the relief outlining it. Thus, DSM uses digital elevation model (DEM) and its derivatives, e.g., slope, terrain curvatures, topographical wetness index (WI), aspect, etc. to predict how soils are distributed in the landscape (MCBRATNEY; MENDONÇA SANTOS; MINASNY, 2003; MENEZES et al., 2016; PRASUHN et al., 2013; SANTOS et al., 2014b; SILVA et al., 2016a, 2014, 2016b; TESKE; GIASSON; BAGATINI, 2014).

According to COELHO and GIASSON (2010), studying different soils classes in Rio Grande do Sul, Brazil, from the set of variables used based on DEM, the ones which best explained the relationship between the landscape and the spatial distribution of soil classes in the area were: slope, profile curvature, elevation, plan curvature, and wetness index. Although DEMs are widely used, the relationship between relief and soil occurrence may vary in different physiographical areas; thus, it is necessary to test the models by quantifying their ability to predict soil classes in an environmentally diverse situation (GIASSON et al., 2013). GRUNWALD (2009) reported that most studies on digital soil mapping and modeling focused on external drivers such as climate, vegetation, and land use that modulate soil organic carbon, total phosphorus, and other soil properties to be predicted. However, to incorporate intrinsic soil properties, i.e., mineralogy, structure, soil aggregation, etc., to further improve the prediction power of those models, chemical analyses are usually required, and those may be expensive, time-consuming, and generate chemical residues.

Even though DTMs have been used worldwide as adequate predictors of soil classes and properties, recent studies are searching for new tools associated with soil attributes, especially those concerning chemical features (SILVA et al., 2016a). Nowadays a range of

sensors are being used in soil surveys to investigate a soil profile and varying soil attributes (HARTEMINK, 2015; HARTEMINK; MINASNY, 2014), e.g., magnetic susceptibility (MS), portable X-ray fluorescence (pXRF) spectrometer, ground penetrating radar (GPR), electrical resistivity (ER), the cone penetrometer or diffuse reflectance spectroscopy (DRS).

Portable X-ray fluorescence (pXRF) spectrometry is a rapid, proximal scanning technology that allows for total elemental quantification for elements from Mg to U in the Periodic Table in a few seconds (RIBEIRO et al., 2017; WEINDORF et al., 2012). Besides, pXRF is a non-invasive, environmentally friendly, and low-cost alternative compared to laboratory methods. With the aid of pXRF, elements that are less commonly evaluated in soil science can be easily and rapidly detected (SILVA et al., 2018b, 2018a). According to STOCKMANN et al. (2016), pXRF results have been used to study a number of soil properties. In these studies, the total elemental concentrations are used to infer a range of soil properties using multiple linear regression and other types of models. In another research line, WEINDORF et al. (2012) concluded that pXRF could help to unveil the weathering degree of the soil materials and to assist in distinguishing between two different parent materials. Focusing on soil genesis, changes in elemental contents from A to Cr horizons, due to weathering, were analyzed with pXRF in the sand, silt, and clay fractions, supporting detailed X-ray diffractometry analyses of the mineralogy of an Inceptisol and enhancing soil characterization of chemical properties within the soil profile (SILVA et al., 2018a).

pXRF was able to identify the presence or absence of Fe in light-colored subsoil mottles in situations in which it is hard to figure out if the Fe form comes from reduction, depletion, or remnant from parent material (WEINDORF et al., 2012). STOCKMANN et al. (2016), investigating the pedogenetic pathways of three different soil profiles, concluded that pXRF analyses might add considerable value to in-field soil description for the objective assessment of soil formation pathways. Although most of such studies were performed in temperate regions, in Brazil, a tropical country with different soils, the use of pXRF for in-field or in-lab soil characterization is increasing, and more studies are needed to further evaluate the effects of different tropical soil conditions (RIBEIRO et al., 2017; SILVA et al., 2016a).

Regarding other sensors, the susceptibilimeter has also been widely used for being capable of determining the magnetic susceptibility (MS) of materials (CERVI et al., 2014). CAMARGO et al. (2018) observed that MS was an adequate predictor of the presence of potentially toxic elements, with predictions of Co, Ni, and Zn contents presenting R^2 of 0.85,

0.66, and 0.87, respectively. MS was more efficient in the identification of areas with different patterns of pedogenetic variability than the hue determined by DRS for Oxisols under parent material transition in São Paulo (SIQUEIRA et al., 2015).

Finally, SILVA et al. (2016b), working with proximal sensing and digital terrain models (DTM) to map soils, concluded that the use of pXRF and MS data helped create more detailed soil maps. Therefore, the use of DSM tools combined with different sensors can be a powerful ally in the process of soil mapping in Brazil, creating opportunities to improve the scales of the maps already existing at lower costs.

2.2 Random forest

The use of prediction models to help predict or simulate real events are essential in DSM (MCBRATNEY; MENDONÇA SANTOS; MINASNY, 2003). Several methods are currently available, and they can estimate soil classes and soil properties at unsampled locations. Hence the choice of the best algorithm is of utmost importance.

Random Forest (RF) algorithm is one of the most commonly used machine learning techniques. It is an improvement of decision trees since it grows several trees to generate a final prediction (BREIMAN et al., 1984). RF can work for both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap Aggregation (bagging). Bagging involves training each decision tree on a different data sample where sampling is done with replacement (BREIMAN, 1996). The idea is to combine multiple decision trees in determining the final output rather than relying on an individual decision tree.

As a regression algorithm, PELEGRINO et al. (2019) used RF to model and predict the available contents of micronutrients from pXRF in addition to terrain attributes and parent material information with high accuracy. The combination of the different information allowed an accurate prediction through the RF algorithm. In another work comparing stepwise multiple linear regression (SMLR) and RF, the RF algorithm was able to predict soil fertility properties with more accuracy than SMLR (SILVA et al., 2017a).

REFERENCES

ABREU, C. A. DE; LOPES, A. S.; SANTOS, G. C. G. DOS. Micronutrientes. In: **Fertilidade do Solo**. Viçosa, MG: SBCS, 2007. p. 1017.

AMADO, T. J. C. et al. Chemical and physical attributes of oxisols and their relation with irrigated corn and common bean yields. **Revista Brasileira de Ciência do Solo**, v. 33, n. 4, p. 831–843, 2009.

ANDRADE, R. et al. Assessing models for prediction of some soil chemical properties from portable X-ray fluorescence (pXRF) spectrometry data in Brazilian Coastal Plains. **Geoderma**, v. 357, n. August 2019, p. 113957, jan. 2020.

AYOUBI, S.; JABBARI, M.; KHADEMI, H. Multiple linear modeling between soil properties, magnetic susceptibility and heavy metals in various land uses. **Modeling Earth Systems and Environment**, v. 4, n. 2, p. 579–589, 2018.

BAVER, L. D.; GARDNER, W. H.; GARDNER, W. R. **Soil Physics**. 5th ed ed. New York, NY, USA: John Wiley & Sons, 1972.

BESKOW, S. et al. Soil erosion prediction in the Grande River Basin, Brazil using distributed modeling. **Catena**, v. 79, p. 49–59, 2009.

BONFATTI, B. R. et al. A mechanistic model to predict soil thickness in a valley area of Rio Grande do Sul, Brazil. **Geoderma**, v. 309, n. September 2017, p. 17–31, 2018.

BOUYOUCOS, G. J. Directions for making mechanical analyses of soils by the hydrometer method. **Soil Science**, v. 42, n. 3, 1936.

BREIMAN, L. et al. **Classification and Regression Trees**. 1. ed. New York, US: Chapman and Hall(Wadsworth, Inc), 1984.

BREIMAN, L. Bagging predictors. **Machine Learning**, n. 421, p. 123–140, 1996.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

BRUNGARD, C. W. et al. Machine learning for predicting soil classes in three semi-arid landscapes. **Geoderma**, v. 239–240, p. 68–83, fev. 2015.

CAMARGO, L. A. et al. Predicting potentially toxic elements in tropical soils from iron oxides, magnetic susceptibility and diffuse reflectance spectra. **Catena**, v. 165, n. March, p. 503–515, 2018.

CERVI, E. C. et al. Magnetic susceptibility and the spatial variability of heavy metals in soils developed on basalt. **Journal of Applied Geophysics**, v. 111, p. 377–383, 2014.

CHAGAS, C. DA S. et al. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. **CATENA**, v. 139, p. 232–240, abr. 2016.

CHAKRABORTY, S. et al. External parameter orthogonalisation of Eastern Eurpoean VisNIR-DRS soil spectra. **Geoderma**, v. 337, p. 65–75, 2019a.

CHAKRABORTY, S. et al. Use of portable X-ray fluorescence spectrometry for classifying soils from different land use land cover systems in India. **Geoderma**, v. 338, n. May 2018, p. 5–13, mar. 2019b.

COELHO, F. F.; GIASSON, E. Comparação de métodos para mapeamento digital de solos com utilização de sistema de informação geográfica. **Ciência Rural**, v. 40, p. 2099–2106, 2010.

CONRAD, O. et al. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. **Geosci. Model Dev.**, v. 8, n. 7, p. 1991–2007, 7 jul. 2015.

DANTAS, A. A. A.; CARVALHO, L. G. DE; FERREIRA, E. Climatic classification and tendencies in Lavras region, MG. **Ciência e Agrotecnologia**, v. 31, n. 6, p. 1862–1866, 2007.

DE SOUZA BAHIA, A. S. R. et al. Prediction and Mapping of Soil Attributes using Diffuse Reflectance Spectroscopy and Magnetic Susceptibility. **Soil Science Society of America Journal**, v. 81, n. 6, p. 1450, 2017.

DEARING, J. A. Environmental Magnetic Susceptibility: using the Bartington MS2 System.

Chi Pub, 1994.

DHARUMARAJAN, S.; HEGDE, R.; SINGH, S. K. Spatial prediction of major soil properties using Random Forest techniques - A case study in semi-arid tropics of South India. **Geoderma Regional**, v. 10, n. April, p. 154–162, 2017.

DOS SANTOS TEIXEIRA, A. F. et al. Tropical soil pH and sorption complex prediction via portable X-ray fluorescence spectrometry. **Geoderma**, v. 361, n. October 2019, p. 114132, mar. 2020.

DUDA, B. M. et al. Soil characterization across catenas via advanced proximal sensors. **Geoderma**, v. 298, p. 78–91, 2017.

EMBRAPA. **Manual de Métodos de Análise de Solo**. 3ª edição ed. Brasília, DF: Embrapa, 2017.

FAGERIA, N. K.; STONE, L. F. Micronutrient deficiency problems in South America. **Micronutrient Deficiencies in Global Crop Production**, p. 245–266, 2008.

GANASRI, B. P.; RAMESH, H. Assessment of soil erosion by RUSLE model using remote sensing and GIS - A case study of Nethravathi Basin. **Geoscience Frontiers**, v. 7, n. 6, p. 953–961, nov. 2016.

GEE, G. W.; BAUDER, J. W. Particle-size Analysis1. In: **Methods of Soil Analysis: Part 1—Physical and Mineralogical Methods**. SSSA Book Series SV - 5.1. Madison, WI: Soil Science Society of America, American Society of Agronomy, 1986. p. 383–411.

GIASSON, E. et al. Decision trees for digital soil mapping on subtropical basaltic steeplands. **Scientia Agricola**, v. 68, p. 167–174, 2011.

GIASSON, E. et al. Avaliação de cinco algoritmos de árvores de decisão e três tipos de modelos digitais de elevação para mapeamento digital de solos a nível semidetalhado na Bacia do Lageado Grande, RS, Brasil. **Ciência Rural**, v. 43, p. 1967–1973, 2013.

GRUNWALD, S. Multi-criteria characterization of recent digital soil mapping and modeling approaches. **Geoderma**, v. 152, p. 195–207, 2009.

HARTEMINK, A. E. New Tools for Pedologists: Digital Soil Morphometrics. **Soil Horizons**, v. 56, p. 1–2, 2015.

HARTEMINK, A. E.; MINASNY, B. Towards digital soil morphometrics. **Geoderma**, v. 230–231, p. 305–317, 2014.

HENGL, T. et al. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. **Nutrient Cycling in Agroecosystems**, v. 109, n. 1, p. 77–102, 2 set. 2017.

HUDSON, B. D. The Soil Survey as Paradigm-based Science. **Soil Science Society of America Journal**, v. 56, n. 3, p. 836, 1992.

JACKSON, M. L. Soil chemical analysis prentice Hall. **Inc., Englewood Cliffs, NJ**, v. 498, 1958.

JENNY, H. Factors of soil formation. 281 pp. **New York**, 1941.

KÄMPF, N.; MARQUES, J. J.; CURI, N. Mineralogia de Solos Brasileiros. In: **Pedologia Fundamentos**. Viçosa, MG: SBCS, 2012. p. 343.

KESKIN, H.; GRUNWALD, S. Regression kriging as a workhorse in the digital soil mapper's toolbox. **Geoderma**, v. 326, n. April, p. 22–41, 2018.

KUHN, M. Building predictive models in R using the caret package. **Journal of statistical software**, v. 28, n. 5, p. 1–26, 2008.

LAGACHERIE, P.; MCBRATNEY, A. B. Chapter 1 Spatial Soil Information Systems and Spatial Soil Inference Systems: Perspectives for Digital Soil Mapping. In: **Developments in Soil Science**. [s.l: s.n.]. v. 31p. 3–22.

LEE, S.; WOLBERG, G.; SHIN, S. Y. Scattered data interpolation with multilevel B-splines. **IEEE transactions on visualization and computer graphics**, v. 3, n. 3, p. 228–244, 1997.

LIAW, A.; WIENER, M. Classification and Regression by randomForest. **R News**, v. 2, n. 3,

p. 18–22, 2002.

LIMA, T. M. DE et al. Elemental analysis of Cerrado agricultural soils via portable X-ray fluorescence spectrometry: Inferences for soil fertility assessment. **Geoderma**, v. 353, n. June, p. 264–272, 2019.

LOPES, A. S.; GUILHERME, L. R. G. A career perspective on soil management in the Cerrado region of Brazil. In: **Advances in Agronomy**. [s.l.] Elsevier, 2016. v. 137p. 1–72.

MAHER, B. A. Characterisation of soils by mineral magnetic measurements. **Physics of the Earth and Planetary Interiors**, v. 42, n. 1–2, p. 76–92, 1986.

MANCINI, M. et al. Tracing tropical soil parent material analysis via portable X-ray fluorescence (pXRF) spectrometry in Brazilian Cerrado. **Geoderma**, v. 337, n. October 2018, p. 718–728, 2019.

MARQUES JR., J. et al. Magnetic susceptibility and diffuse reflectance spectroscopy to characterize the spatial variability of soil properties in a Brazilian Haplustalf. **Geoderma**, v. 219–220, p. 63–71, maio 2014.

MCBRATNEY, A. B.; MENDONÇA SANTOS, M. L.; MINASNY, B. On digital soil mapping. **Geoderma**, v. 117, p. 3–52, 2003.

MEHLICH, A. **Determination of P, Ca, Mg, K, Na and NH4**. Raleigh, NC, USA: University of North Carolina, 1953.

MENDONÇA-SANTOS, M. L.; SANTOS, H. G. The State of the Art of Brazilian Soil Mapping and Prospects for Digital Soil Mapping. **Developments in Soil Science**, v. 31, p. 39–54, 2007.

MENEZES, M. D. DE et al. Digital soil mapping approach based on fuzzy logic and field expert knowledge. **Ciência e Agrotecnologia**, v. 37, p. 287–298, 2013.

MENEZES, M. D. DE et al. Spatial prediction of soil properties in two contrasting physiographic regions in Brazil. **Scientia Agricola**, v. 73, n. 3, p. 274–285, 2016.

MINASNY, B.; MCBRATNEY, A. B. Digital soil mapping: A brief history and some lessons. **Geoderma**, v. 264, p. 301–311, 2016.

OBADE, V. DE P. Integrating management information with soil quality dynamics to monitor agricultural productivity. **Science of the Total Environment**, v. 651, p. 2036–2043, 2019.

PELEGRINO, M. H. P. et al. Synthesis of proximal sensing, terrain analysis, and parent material information for available micronutrient prediction in tropical soils. **Precision Agriculture**, v. 20, n. 4, p. 746–766, 18 ago. 2019.

POGGERE, G. C. et al. Maghemite quantification and magnetic signature of Brazilian soils with contrasting parent materials. **Applied Clay Science**, v. 161, p. 385–394, 2018.

PRASUHN, V. et al. A high-resolution soil erosion risk map of Switzerland as strategic policy support system. **Land Use Policy**, v. 32, p. 281–291, 2013.

QUEMÉNEUR, J. J. G. et al. Geologia da Folha Lavras 1:100.000. In: **Geologia e Recursos Minerais do Sudeste Brasileiro**. 1ed. ed. Belo Horizonte, Minas Gerais: Companhia Mineradora de Minas Gerais, Secretaria de Desenvolvimento Econômico, 2003. p. 259–319.

RAWAL, A. et al. Determination of base saturation percentage in agricultural soils via portable X-ray fluorescence spectrometer. **Geoderma**, v. 338, n. November 2018, p. 375–382, mar. 2019a.

RAWAL, A. et al. Determination of base saturation percentage in agricultural soils via portable X-ray fluorescence spectrometer. **Geoderma**, v. 338, n. December 2018, p. 375–382, mar. 2019b.

RIBEIRO, B. T. et al. Portable X-ray fluorescence (pXRF) applications in tropical Soil Science. **Ciência e Agrotecnologia**, v. 41, p. 245–254, 2017.

ROBINSON, G. W. A new method for the mechanical analysis of soils and other dispersions. **The Journal of Agricultural Science**, v. 12, n. 3, p. 306–321, 27 jul. 1922.

ROSEMARY, F. et al. Exploring the spatial variability of soil properties in an Alfisol soil catena. **CATENA**, v. 150, p. 53–61, mar. 2017.

SANTOS, W. J. R. DOS et al. Detailed soil survey of an experimental watershed representative of the Brazilian Coastal Plains and its practical application. **Ciência e Agrotecnologia**, 2014a.

SANTOS, W. J. R. DOS et al. Detailed soil survey of an experimental watershed representative of the Brazilian Coastal Plains and its practical application. **Ciência e Agrotecnologia**, v. 38, n. 1, p. 50–60, 2014b.

SAYÃO, V. M. et al. Satellite land surface temperature and reflectance related with soil attributes. **Geoderma**, v. 325, n. September 2017, p. 125–140, 2018.

SCUDIERO, E.; SKAGGS, T. H.; CORWIN, D. L. Regional-scale soil salinity assessment using Landsat ETM + canopy reflectance. **Remote Sensing of Environment**, v. 169, p. 335–343, nov. 2015.

SHAHBAZI, F. et al. Evaluating the spatial and vertical distribution of agriculturally important nutrients — nitrogen, phosphorous and boron — in North West Iran. **Catena**, v. 173, n. December 2017, p. 71–82, 2019.

SILVA, S. et al. Proximal Sensing and Digital Terrain Models Applied to Digital Soil Mapping and Modeling of Brazilian Latosols (Oxisols). **Remote Sensing**, v. 8, n. 8, p. 614, 25 jul. 2016a.

SILVA, S. H. G. et al. A Technique for Low Cost Soil Mapping and Validation Using Expert Knowledge on a Watershed in Minas Gerais, Brazil. **Soil Science Society of America Journal**, v. 78, p. 1310–1319, 2014.

SILVA, S. H. G. et al. Retrieving pedologist's mental model from existing soil map and comparing data mining tools for refining a larger area map under similar environmental conditions in Southeastern Brazil. **Geoderma**, v. 267, p. 65–77, 2016b.

SILVA, S. H. G. et al. Geomorphometric tool associated with soil types and properties spatial variability at watersheds under tropical conditions. **Scientia Agricola**, v. 73, n. 4, p. 363–370, 2016c.

SILVA, S. H. G. et al. Proximal sensing and digital terrain models applied to digital soil mapping and modeling of Brazilian Latosols (Oxisols). **Remote Sensing**, v. 8, n. 8, p. 614, 2016d.

SILVA, S. H. G. et al. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). **Ciência e Agrotecnologia**, v. 41, n. 6, p. 648–664, 2017a.

SILVA, S. H. G. et al. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). **Ciência e Agrotecnologia**, v. 41, n. 6, p. 648–664, 2017b.

SILVA, S. H. G. et al. Soil weathering analysis using a portable X-ray fluorescence (PXRF) spectrometer in an Inceptisol from the Brazilian Cerrado. **Applied Clay Science**, v. 162, n. June, p. 27–37, 2018a.

SILVA, S. H. G. et al. Tropical soils characterization at low cost and time using portable X-ray fluorescence spectrometer (pXRF): Effects of different sample preparation methods. **Ciência e Agrotecnologia**, v. 42, n. 1, p. 80–92, 2018b.

SILVA, S. H. G. et al. Soil texture prediction in tropical soils: A portable X-ray fluorescence spectrometry approach. **Geoderma**, v. 362, n. September 2019, p. 114136, mar. 2020.

SIQUEIRA, D. S. et al. Detailed mapping unit design based on soil-landscape relation and spatial variability of magnetic susceptibility and soil color. **Catena**, v. 135, p. 149–162, 2015.

SÖDERSTRÖM, M. et al. Sensor mapping of Amazonian Dark Earths in deforested croplands. **Geoderma**, v. 281, p. 58–68, 2016.

STOCKMANN, U. et al. Utilizing portable X-ray fluorescence spectrometry for in-field

investigation of pedogenesis. **Catena**, v. 139, p. 220–231, 2016a.

STOCKMANN, U. et al. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. **CATENA**, v. 139, p. 220–231, abr. 2016b.

TEAM, R. D. C. **R: A language and environment for statistical computing**R foundation for statistical computing Vienna, Austria, , 2009.

TEIXEIRA, A. F. DOS S. et al. Portable x-ray fluorescence (pXRF) spectrometry applied to the prediction of chemical attributes in inceptisols under different land use. **Ciencia e Agrotecnologia**, v. 42, n. 5, p. 501–512, 2018a.

TEIXEIRA, D. D. B. et al. Mapping units based on spatial uncertainty of magnetic susceptibility and clay content. **Catena**, v. 164, n. December 2017, p. 79–87, 2018b.

TEIXEIRA, P. C. et al. **Manual de metodos de analises**. 3ed rev. e ed. Brasília, DF: Embrapa: [s.n.].

TERRA, F. S.; DEMATTÊ, J. A. M.; ROSSEL, R. A. V. Proximal spectral sensing in pedological assessments: vis–NIR spectra for soil classification based on weathering and pedogenesis. **Geoderma**, v. 318, p. 123–136, 2018.

TESKE, R.; GIASSON, E.; BAGATINI, T. Comparação do uso de modelos digitais de elevação em mapeamento digital de solos em Dois Irmãos, RS, Brasil. **Revista Brasileira de Ciencia do Solo**, v. 38, n. 5, p. 1367–1376, 2014.

WEINDORF, D. C. et al. Characterizing soils via portable x-ray fluorescence spectrometer: 2. Spodic and Albic horizons. **Geoderma**, v. 189–190, p. 268–277, 2012.

WEINDORF, D. C.; BAKR, N.; ZHU, Y. Advances in Portable X-ray Fluorescence (PXRF) for Environmental, Pedological, and Agronomic Applications. In: **Advances in Agronomy**. [s.l.] Academic Press, 2014. v. 128p. 1–45.

WEINDORF, D. C.; CHAKRABORTY, S. Portable X-ray Fluorescence Spectrometry Analysis of Soils. **Methods of Soil Analysis**, v. 1, n. 1, p. 0, 2016.

ZERAATPISHEH, M. et al. Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. **Geoderma**, v. 338, n. January 2018, p. 445–452, mar. 2019.

ZHANG, Y.; HARTEMINK, A. E. Soil horizon delineation using vis-NIR and pXRF data. **CATENA**, v. 180, n. April, p. 298–308, set. 2019.

SECOND PART – ARTICLES

3 ARTICLE 1. PREDICTION OF SOIL FRACTIONS USING DIFFERENT PROXIMAL SENSORS AND RANDOM FOREST

*Article prepared according to the rules of Pedosphere (preliminary version).

Luiza Maria Pereira Pierangeli; Sérgio Henrique Godinho Silva; Michele Duarte de Menezes; João José Marques; Luiz Roberto Guimarães Guilherme; David C. Weindorf; Nilton Curi

ABSTRACT

Soil texture is related to many other physical, chemical, and biological properties. Knowing its spatial distribution is essential because it contributes to decisions on agricultural management and environmental sustainability. Digital Elevation Models (DEM) are used continuously to represent terrain topography, enabling the generation of terrain attributes (TA). Moreover, proximal sensors, such as portable X-ray fluorescence spectrometry (pXRF), have also been an ally in digital soil mapping. The objective of this work was to apply random forest algorithm to model and predict clay, silt, and sand contents via magnetic susceptibility (MS) and pXRF information in addition to TA information. A total of 39 soil samples were collected from A and B horizons, totaling 78 samples, and analyzed by a susceptibilimeter and pXRF in addition to laboratory analyses to determine clay, silt, and sand contents. Seventeen TAs were generated from DEM. These data were divided into six datasets: TA; TA+pXRF; TA+MS; TA+MS+pXRF; MS+pXRF; pXRF. The samples were divided into A and B horizons separately and A+B horizon combined. Models were validated through root mean square error (RMSE) and the coefficient of determination (R^2) to determine the performance of the models. Finally, the best models were spatialized to the entire study area. The combination of pXRF information with MS and TA improved the prediction of silt in the A and B horizons. Sand content in A and B horizon and clay content in the B horizon were better predicted using only pXRF information. MS information, in combination with pXRF, allowed a better prediction of clay content in the A horizon. The use of pXRF allowed for better predictions of soil fractions contents than TA.

Keywords: Digital soil mapping; pXRF; Magnetic Susceptibility; Random Forest; Tropical soils.

3.1 INTRODUCTION

Particle size distribution (PSD) is a fundamental soil property. It is related to other physical, chemical, biological, and mineralogical properties, as well as water retention capacity, porosity, and others (Rosemary *et al.*, 2017). Besides, soil texture is necessary for several hydrological, environmental, and climate risk assessment models (Ganasri and Ramesh, 2016; Scudiero *et al.*, 2015). The methods currently used to quantify soil texture are costly, time-consuming, laborious, and non-environmentally friendly, as they make use of chemical reagents. The most used methods are the pipette (Robinson, 1922) and the hydrometer methods (Bouyoucos, 1936).

Understanding the spatial variability of the PSD ensures more sustainable soil use and management and soil conservation. This variability is related to soil forming factors such as parent material, climate, and relief, and the weathering degree. Therefore, the use of digital soil mapping (DSM) is an alternative for predicting how the PSD varies across the landscape. The DSM relates the soil attributes to other continuous, easily accessible environmental variables, such as terrain attributes (TA) obtained from the digital elevation model (DEM) (Grunwald, 2009; McBratney *et al.*, 2003; Minasny and Mcbratney, 2016). Furthermore, proximal sensors such as portable X-ray fluorescence (pXRF) spectrometer, susceptibilimeters, visible and near-infrared (Vis-NIR) spectroscopy combined with TA has been recently investigated and delivered accurate soil mapping results (Hartemink and Minasny, 2014; Silva *et al.*, 2016d).

PXRF is used to quantify the total elemental content in different types of materials, as each element produces specific fluorescence energy (Ribeiro *et al.*, 2017; Weindorf *et al.*, 2014). Another equipment is the susceptibilimeter, which quantifies magnetic susceptibility (MS) of different materials (Dearing, 1994). Besides, both techniques are non-invasive, non-destructive, and can be used in field and lab conditions. Also, such information is acquired in a fast, inexpensive, and environmentally friendly way.

Several studies have been conducted on spatial variation of soil properties using DSM and proximal sensing. Pelegrino *et al.* (2018) predicted the available contents of Fe, Cu, Mn, and Zn from pXRF data along with TA derived from DEM with 5 and 10 m resolutions. Stockmann *et al.* (2016) studied pedogenesis pathways and parent materials of three different soil types in Australia utilizing pXRF. Zhang and Hartemink (2019) evaluated the use of Vis-NIR and pXRF for delineating soil horizons and if moisture can affect the measurements in Wisconsin, USA. Marques *et al.* (2014) used MS and diffuse reflectance spectroscopy as

predictive variables in the characterization of physical and chemical properties of soil profiles in Brazil. Silva et al. (2016) used TA, MS and pXRF data to predict soil classes and PSD.

Several studies have used machine learning to model and predict soil variability, correlating soil properties with environmental variables (Brungard *et al.*, 2015; Hengl *et al.*, 2017; Silva *et al.*, 2017b; Zeraatpisheh *et al.*, 2019). Among the algorithms used, Random Forest (RF) (Breiman, 2001) has proven to be a robust algorithm suitable for soil-related studies. RF can fit non-linear relationships, has low bias and variance, has high performance in predictions, and can determine the importance of variables (Chagas *et al.*, 2016; Chakraborty *et al.*, 2019b; Rawal *et al.*, 2019a). Therefore, this study aims to evaluate the importance of TA, pXRF and MS to produce maps of clay, sand, and silt contents for soil A and B horizons in an experimental farm in Brazil, based on the RF.

3.2 MATERIALS AND METHODS

3.2.1 Study Area

The study was conducted at the Palmital farm (Fig. 1), located in Ijaci county, Minas Gerais, Brazil, between UTM longitudes 506,793 and 508,882 m and latitudes 7,659,470 and 7,660,685 m zone 23K, datum WGS84. This study area is ~117 ha, featuring native forest (Cerrado forest), eucalyptus, mahogany and pinus plantations, agricultural crops, e.g. soybeans, corn, and beans, and pasture. The climate is Cwa (C: Humid subtropical; w: dry winter; a: hot summer) according to Köppen classification, the average annual temperature is 21 °C, and the average annual rainfall is 1500 mm (Dantas *et al.*, 2007).

The altitude in the study area ranges from 793 to 868 m. The geology information of the area was obtained from a 1:100,000 scale map (MAPA GEÓLOGICO – FOLHA LAVRAS, 2003), encompassing gneisses TTG, metamorphosed limestone and phyllite-limestone, alluvial sediment from quaternary.
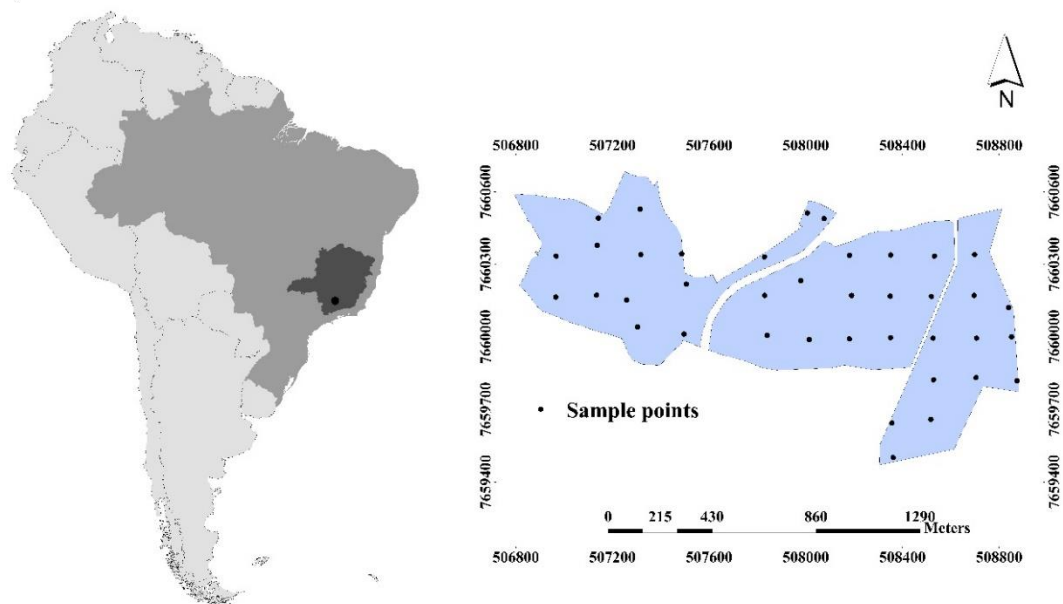
Fig. 1. Study area and sampling sites of Palmital farm, Minas Gerais state, Brazil.

3.2.2 Soil Sampling and Laboratory Analyses

A total of 39 sampling sites (small trenches) were excavated throughout the area, following a regular-grid design with 173 m between sampling places, reaching a sampling density of 1 sampling site per 3.5 ha. In total, 78 samples (39 – A horizon; 39 – B horizon) were collected. Soil samples were air-dried, ground and passed through a 2 mm sieve. Clay, silt, and sand were determined by the pipette method (Baver *et al.*, 1972; Gee and Bauder, 1986; Teixeira *et al.*, 2017).

A Vanta series pXRF (Olympus, Waltham, MA, USA) spectrometer was used to obtain the total contents of the diverse elements on the soil samples. The pXRF features a Rh X-ray tube operated at 8–50 kV as the excitation source. Scans were performed in Geochem Mode with scanning time set to 30 s per beam; the instrument scans via two beams in sequence such that one whole scan was completed in 60 s, according to Weindorf and Chakraborty (2016). Prior to scanning, the instrument was calibrated with a factory calibration alloy coin. The accuracy of the equipment was evaluated through scanning standard reference materials 2710a and 2711a certified by the National Institute of Standards and Technology (NIST). The recovery values (value obtained by pXRF /certified value of the reference material *100) for the elements identified in all samples and used in this work follow (2710a/2711a): Ag (0/0); Al (61/61); As (106/143); Ca (65/88); Cd (0/91); Cu (102/101); Fe (92/96); Hg (273/175); K (83/85); Mn (93/96); Mo (0/0); Ni (0/115); P (63/52);

25

Pb (102/108); Rb (0/0); S (0/0); Se (0/0); Si (57/61); Sr (96/96); Th (0/0); Ti (90/94); U (99/0); V (0/93); y (0/); Zn (100/106).

For the analyses of MS, a Bartington MS2B susceptibilimeter (Dearing, 1994) was used, and 10 g of air-dried 2-mm sieved soil were used. Data were measured at low frequency (0.47 kHz).

3.2.3 Terrrain Variables

A DEM of 12.5 m resolution was obtained through the ALOS PALSAR Global Radar Imagery (https://search.asf.alaska.edu/#/). The TA derivative from the DEM were created in SAGA GIS software (CONRAD et al., 2015), including: slope, aspect, cross-sectional and longitudinal curvatures (CsC and LC, respectively), convergence index (CI), flow accumulation (flow), topographic wetness index (TWI), LS factor (lsf), channel network base level (CNBL), vertical distance to channel network (VDCN), valley depth (VD), relative slope position (RSP), catchment area (CA), modified catchment area (MCA), closed depressions (CD), catchment slope (CS) and, SAGA wetness index (SWI). The TA and the land use of the area, in addition to MS and pXRF data, were used in random forest models and adjusted to predict clay, silt, and sand.

3.2.4 Modeling and Accuracy

RF models were adjusted for the samples per horizon (39 samples for A and 39 samples for B) and also by combining A + B horizons (78 samples), all considering the use information of the area, and divided in the following datasets: (i) TA, (ii) TA + pXRF, (iii) TA + MS, (iv) TA + pXRF + MS, (v) MS + pXRF, and (vi) pXRF. No transformations were made on the data sets, and 70% of the samples were randomly selected for the calibration set, while the remaining 30% were used as the validation set. The RF models were created in R software (version 3.6.1) (R Development Core Team, 2009) through "caret" package (Kuhn, 2008) performed with the following parameters: number of trees of the model (ntrees)=1000, node size=5, and number of variables used in each tree (mtry)=3 adjusted using the calibration set, and the model performance was evaluated on the validation set. The performance of each model was calculated via coefficient of determination (R^2) and root mean square error (RMSE) (Eq. 1).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

where n: number of observations, X_{model}: estimated value by the model, X_{obs}: measured value by chemical analysis.

Also, the RF model determines the importance of each variable, the increase in percentage of the mean square error (MSE) of predictions as each variable is removed from the set of predictor variables (%IncMSE) was calculated using the caret package. For each tree, the prediction error on test is recorded (MSE). Then the same is done after permuting each predictor variable. The difference between the two is then averaged over all trees, and normalized by the standard deviation of the differences (Liaw and Wiener, 2002). The higher the difference is, the more important the variable.

The best models for clay, silt, and sand were used in the creations of maps to the entire area. For the soil mapping procedure, TA information was continuously available for the entire study area, but variables obtained from pXRF and MS data at the 39 sampled sites had to be extrapolated to the entire area. The pXRF and MS data were modeled across the entire area through multilevel B-spline interpolation (Lee *et al.*, 1997). These maps were validated through calculations of R^2 and RMSE.

3.3 RESULTS AND DISCUSSIONS

3.3.1 Characterization of soils according to texture and pXRF data

A textural triangle including the samples, according to the Brazilian Textural Classification Chart (EMBRAPA, 2017), is shown in Fig. 2. The triangle shows that the samples were distributed mostly in the clay category. The mean clay, sand, and silt contents (%) were 48, 34, 18, respectively (Table 1). Clay fractions ranged from 27 to 74%, sand ranged from 9 to 55%, while silt ranged from 1 to 36%. The maximum silt content found was 36%, since most Brazilian soils tend to present low-level of this particle size fraction. Silva et al. (2020) analyzed samples across Brazil and found a wide range of soil textural classes; most of the soils analyzed were rich in sand or clay.

The coefficient of variation (CV) in most cases was equal or greater than 20%, except for the clay fraction in B horizon, which indicates the heterogeneity of the sample sets. The silt fraction had the highest CV, indicating that the soil in the area has different levels of weathered. This variability in the results is likely because of the different parent material that these soils derived from.

The high clay contents in the soils are mostly related to weathering. Most minerals in the clay fraction of Brazilian soils are kaolinite or Fe- and Al-oxides such as hematite, goethite, and gibbsite (Kämpf *et al.*, 2012). The latter is more commonly found in extremely weathered soils. In smaller proportions, maghemite, are also found in the clay fraction, this Fe-oxide is related to magnetic susceptibility in soils (Kämpf *et al.*, 2012; Poggere *et al.*, 2018).
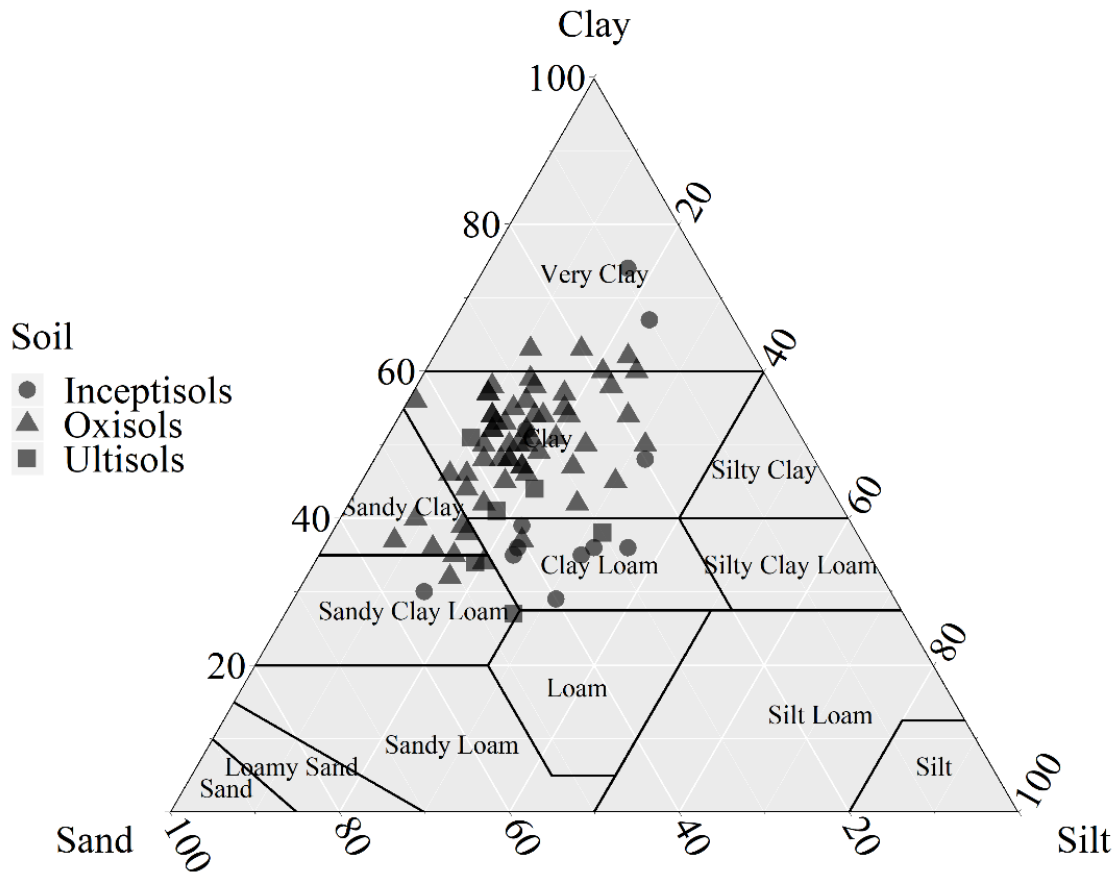
Fig. 2. Soil texture of the collected samples plotted by soil order based on Embrapa textural classification chart.

Despite the relatively small area and the similarity of parent material and land use for all soils, the soil orders found in the area presented somewhat similar pXRF information (Table 2 and Table 3). The highest contents of Al and Fe were found in the Oxisols since they are more weathered than the other soils in the area. Ultisols showed an apparent change between the sandier surface horizon and the more clayey subsurface horizon. This is evident in the high contents of Al, associated with kaolinite in the B horizon, and the high contents of SiO_2, associated with quartz in the A horizon.

The Ca contents are higher in the surface horizon when compared with the subsurface horizon. Liming, a common practice in the tropical soils, mostly for superficial acidity

correction (Lopes and Guilherme, 2016), may leave residues in the soil, causing these differences in Ca content among the horizons. Soils derived from different parent materials showed the same tendencies in Lavras, Brazil (Pelegrino *et al.*, 2018).

The highest contents of Al, Ca, K, Mn, P, S, and Si were found in the A horizon, while the higher contents of Fe were found in the B horizon. These findings were similar to the ones reported by Teixeira *et al.* (2020), although different from the results found by Andrade *et al.* (2020). Among the elements found in tropical soils (Mancini *et al.*, 2019), the higher contents of Ni, Pb, Rb, Sr, V, and Y were found in the B horizon, while Sr was found in higher contents in the A horizon of Ultisols. Magnetic susceptibility (MS) content presented higher values in the Oxisols, and this is likely because of the residual concentration of Fe in these soils, especially in the form of magnetite and maghemite, generally found in the sand and clay fractions, respectively (Kämpf *et al.*, 2012).

Table 1. Summary statistics of particle soil distribution of the A and B horizons of soils from Fazenda Palmital in Brazil

Soil Fraction (%)	N	Min	Max	Mean	SD	CV	Range
				%			
Horizon		--------------------------------Clay--------------------------------					
A+B	78	27	74	48	10	20	47
A	39	27	67	46	9	20	40
B	39	30	74	50	9	19	44
		--------------------------------Silt--------------------------------					
A+B	78	1	36	18	7	38	35
A	39	1	36	19	8	41	36
B	39	9	32	17	6	33	23
		--------------------------------Sand--------------------------------					
A+B	78	9	55	34	10	29	46
A	39	10	55	35	9	26	45
B	39	9	55	33	10	32	46

Min: Minimum; Max: Maximum; SD: standard deviation; CV: coefficient of variation.

Table 2. Summary statistics of pXRF data (mg kg^{-1}) and MS ($\times 10^{-6}$ m^3 kg^{-1}) analyzing A horizons of soils from Palmital farm in Brazil.

Parameter	Al	As	Ca	Cr	Cu	Fe	K	Mn	Ni	P	Pb	Rb	S	Si	Sr	Ti	V	Y	Zn	MS
								Inceptisol (6)												
Min	44781	8	0	56	26	34571	6571	137	30	0	18	51	122	76276	27	6052	57	12	48	3
Max	78254	13	2293	123	40	48249	15245	1450	40	1045	37	112	562	101482	56	7203	91	18	284	15
Mean	58347	10	1231	83	32	40056	11013	512	34	584	25	79	316	89768	42	6507	78	17	110	8
SD	12098	2	769	25	5	5007	3489	539	4	394	7	23	143	9950	13	447	11	2	87	4
CV (%)	21	20	63	30	16	13	32	105	10	67	27	29	45	11	32	7	15	14	79	49
								Oxisol (30)												
Min	50815	7	0	59	16	26817	0	68	16	214	9	7	0	41149	10	7042	36	9	28	2
Max	105232	38	3001	112	46	68014	12540	740	42	1602	27	80	635	111766	65	10970	98	22	212	57
Mean	77454	14	1015	79	26	45104	2973	178	28	679	18	27	226	63066	26	8917	66	14	52	15
SD	14592	6	806	12	6	9903	3498	143	6	300	6	22	114	15956	15	1061	14	3	33	12
CV (%)	19	39	79	16	24	22	118	80	24	44	32	84	50	25	57	12	21	20	63	80
								Ultisol (3)												
Min	45997	6	0	66	20	24863	6462	181	18	135	15	55	114	83710	18	6333	59	10	42	3
Max	54944	9	3742	150	59	43687	13421	385	38	1475	18	99	332	144397	74	7418	74	12	85	10
Mean	50497	8	1458	96	37	35224	10286	288	26	674	17	75	248	108045	41	6821	67	11	61	6
SD	4474	2	2003	47	20	9554	3530	102	10	707	2	22	117	32079	29	551	8	1	22	3
CV (%)	9	20	137	48	53	27	34	36	40	105	9	30	47	30	71	8	11	11	36	54

Min: Minimum; Max: Maximum; SD: standard deviation; CV: coefficient of variation.

Table 3. Summary statistics of pXRF data (mg kg^{-1}) and MS ($\times10^{-6}$ m^3 kg^{-1}) analyzing A horizons of soils from Palmital farm in Brazil.

Parameter	Al	As	Ca	Cr	Cu	Fe	K	Mn	Ni	P	Pb	Rb	S	Si	Sr	Ti	V	Y	Zn	MS
Inceptisol (6)																				
Min	50848	10	0	78	26	38211	5629	81	32	74	20	60	0	78155	21	5144	52	11	47	6
Max	74688	17	0	143	44	62605	15194	739	55	463	32	122	74	116214	60	6983	92	24	81	19
Mean	60105	12	0	102	30	47486	10738	320	45	270	24	84	38	91759	41	6400	80	18	66	9
SD	8769	3	0	30	7	8764	3402	280	7	134	4	22	32	15347	18	675	14	5	12	5
CV (%)	15	22	0	29	23	18	32	88	16	50	18	27	84	17	43	11	18	29	18	61
Oxisol (30)																				
Min	49650	8	0	52	17	27261	0	64	14	145	6	5	0	39146	8	6076	41	10	22	2
Max	98170	24	2018	150	40	75936	9596	947	44	777	31	80	233	85217	81	10829	94	24	65	75
Mean	74079	15	152	79	26	48393	2457	149	26	386	17	26	101	57422	24	9002	66	15	38	17
SD	11090	4	440	20	7	11651	2984	164	7	157	7	23	66	14903	19	1164	14	3	13	16
CV (%)	15	24	290	25	25	24	121	110	28	41	37	88	66	26	77	13	22	18	33	93
Ultisol (3)																				
Min	54997	8	0	81	25	37452	5426	101	28	104	18	45	0	87211	17	5443	73	7	38	5
Max	65825	11	0	142	38	50015	11535	243	49	245	21	102	62	98199	38	6546	88	14	59	17
Mean	59857	10	0	106	30	44879	8401	156	42	166	20	69	21	91561	28	5977	78	11	46	10
SD	5498	2	0	32	7	6587	3058	76	12	72	2	29	36	5840	11	552	8	4	12	6
CV (%)	9	16	0	30	22	15	36	49	29	43	9	42	173	6	37	9	11	33	25	65

Min: Minimum; Max: Maximum; SD: standard deviation; CV: coefficient of variation.

3.3.2 Model performance to predicting clay, silt, and sand contents

Table 4 shows the performance of the models using data from A and B horizons, separately and combined (A+B), considering the different inputs on random forest models. In general, the models for prediction of PSD in the B horizon performed better than the models for the A horizon. The models for A and B horizons combined always presented higher R^2 and smaller RMSE values compared to the A horizon, however for B horizon, this same pattern was not observed.

Table 4. Root mean square error (RMSE) and R^2 of random forest models to clay, silt, and sand in soils from Palmital farm, Brazil.

Model		TA	TA+pXRF	TA+MS	TA+MS+pXRF	MS+pXRF	pXRF
Parameter[1]		RMSE	RMSE	RMSE	RMSE	RMSE	RMSE
	A	8.45	7,65	8,45	7.72	7,53	7,50
Clay	B	9.18	5,91	8,99	5.77	5,18	4,98
	A+B	7.92	6,48	7,93	6.51	6,16	6,19
	A	6.99	6,64	6,80	6.40	6,12	6,20
Silt	B	5.67	5,51	5,68	5.53	5,61	5,59
	A+B	5.82	5,08	5,79	5.04	5,51	5,49
	A	9.86	8,37	9,93	8.32	8,38	8,37
Sand	B	11.08	8,36	10,90	8.51	7,50	7,09
	A+B	8.91	7,47	9,00	7.46	7,26	7,26

[1]root mean square error

Analyzing Fig. 3, which shows the R^2 values from the RF models, it is noticed that the best values were found with the models using pXRF data. Overall, when using pXRF data, the predictive power of the models increased. The most noticeable observation in R^2 and RMSE occurred in B horizon for clay content (TA: R^2 = 0.01 and RMSE = 9.18; pXRF: R^2 = 0.67

and RMSE = 4.98). These values indicate the potential of using pXRF to provide variables to increase the predictive power of models of soil texture in tropical regions. For clay in A and B horizons separately, the prediction models achieved maximum R^2 of 0.30 and 0.67, respectively. For predicting silt in A and B horizon and for A+B horizons combined, the highest obtained R^2 were 0.25, 0.31, respectively. However, for the combined horizons, the silt prediction model achieved a R^2 of 0.36. For sand in A and B horizons and A+B horizons combined, the highest obtained R^2 were 0.19, 0.68, and 0.55, respectively.

For predicting clay in the A horizon, the MS+pXRF model with A+B data was selected, and for the B horizon, the pXRF model with B horizon data was chosen. For silt prediction in the A and B horizons, the TA+MS+pXRF model with A+B data was selected for both. For predicting sand in the A horizon, the pXRF model with A+B data was chosen, and for the B horizon, the pXRF model with B data was selected.

Remarkably, all datasets provided optimal models by utilizing pXRF information alone or in combination with MS and TA information. The best sand prediction model used only pXRF information to predictions in A and B horizons. Only silt used TA information, maybe due to silty soils occurrence in steep areas, thus correlating with some TA. Silva *et al.* (2020) attempted to predict clay, silt, and sand from pXRF in Brazil using 1565 soils samples. The best predictions were delivered by support vector machine and RF, reaching, respectively, R^2 of 0.83 and 0.83 for clay, 0.70 and 0.75 for silt, and 0.87 and 0.84 for sand. The difference in the models' performance between those results and those found in these work can be attributed to the larger dataset used to produce a more robust model than the dataset used in the current work (Silva *et al.*, 2017; Zhu *et al.*, 2011).
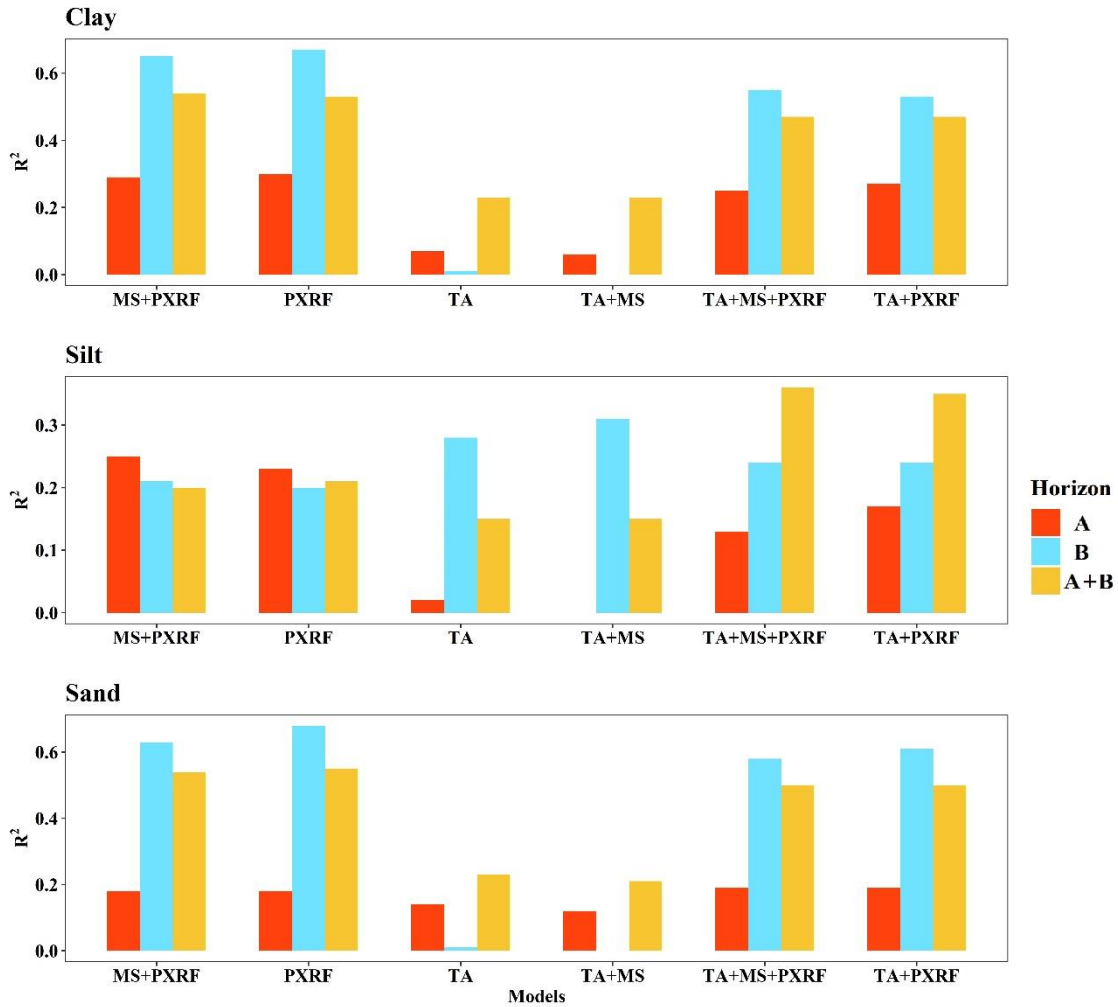
Fig. 3. Determination coefficient (R^2) of clay, silt, and sand prediction models in A horizon, B horizon, and in A+B horizons combined in the Brazilian soils.

3.3.3 Variables Importance

The variable importance calculated for the best models to predict PSD is shown in Fig. 4. The percentage of increment of Mean Square Error (%IncMSE) is an important and significant measure of the relative importance of one independent variable (Liaw and Wiener, 2002). The higher the %IncMSE value, the more important the variable to the prediction model.

The combination of the six datasets [i: TA; ii: TA+pXRF; iii: TA+MS; iv: TA+MS+pXRF; v: MS+pXRF; vi: pXRF] and the data for A, B, or A+B horizons resulted in a total of 54 models for the prediction of the soil particle size fractions. In general, Al, As, Ti, Pb, and Sr were the five most important variables in all soil horizons studied. These variables

may be related These variables can also be related to the weathering of tropical soils, and is related to the parent material of the studied soils (Mancini *et al.*, 2019).
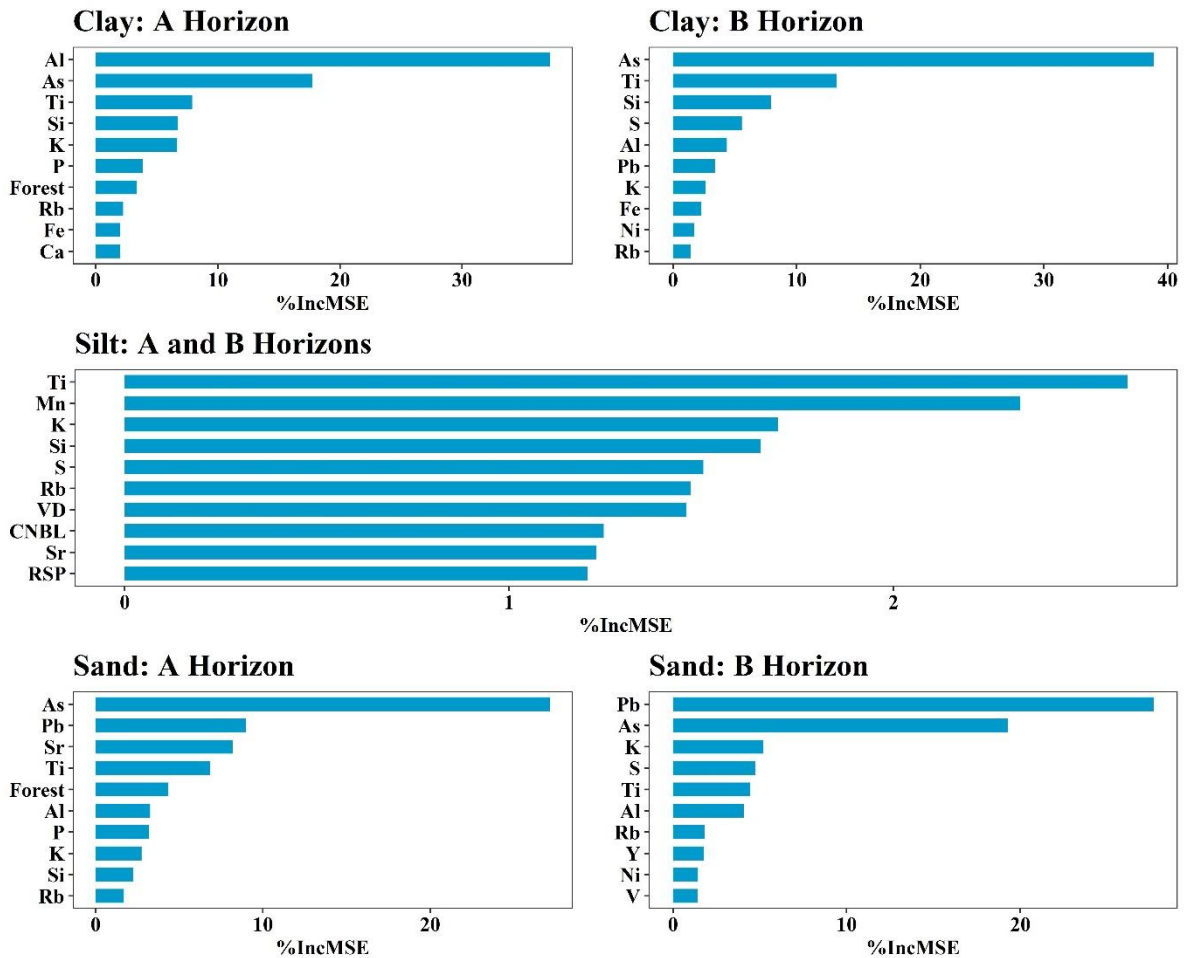

Fig. 4. Plots of relative importance of variables in RF model clay, silt, and sand for the best prediction models in Brazilian soils. VD, valley depth; CNBL, channel network base level; RSP, relative slope position.

3.3.4 Spatial predictions of clay, silt, and sand contents

From the best models determined in the previous section, spatial predictions of clay, silt and sand contents were generated for the whole study area (Fig. 5) based on TA and spatialized MS and the elemental contents obtained by pXRF. The clay content varied from 31 to 61 % in the A horizon, whereas the variation in the B horizon was from 36 to 65 %. The silt content had a variation from 8 to 30 % in the A horizon, while the B horizon content ranged from 12 to 29 %. Furthermore, the sand content ranged from 19 to 48 % in the A horizon and 18 to 48 % in the B horizon.

The maps generated for both horizons showed great similarity between the content of the soil particle size fractions, since most of the area consists of Oxisols. Oxisols are in an

advanced weathering stage, evolved as a result of vigorous transformations of the parent material (Kämpf *et al.*, 2012), and in consequence, they have little differentiation between the subsuperficial horizons.
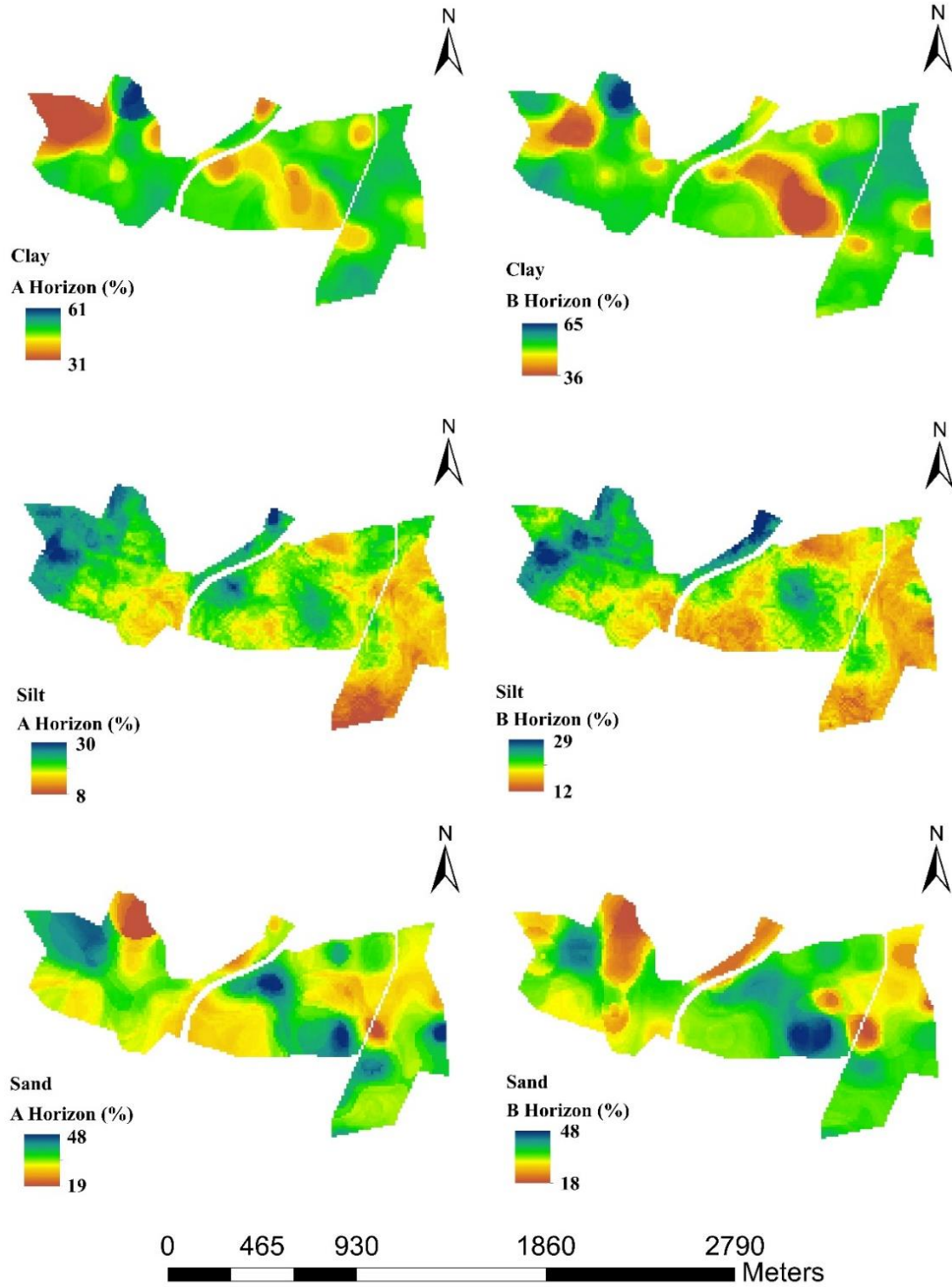
Fig. 5. Maps of clay, silt, and sand contents in A and B horizon obtained from the best models of random forest for the study area in Brazilian soils.

3   CONCLUSIONS

The elemental contents reported by pXRF provided an adequate characterization of the soils in this study. In general, the combination of MS, pXRF, and TA information allowed for satisfactory predictions of clay, silt, and sand contents with the use of random forest algorithm in tropical soils. The use of pXRF information was able to predict soil fractions adequately, and when associated with MS and TA information optimized models' performance. The work shows the potential to use pXRF as a source of variables to help spatial prediction of soil properties rapidly, at low cost and without generating residues. Thus, for the region of study, the use of proximal sensors and remote sensing is recommended for digital mapping and modeling.

REFERENCES

Andrade R, Silva S H G, Weindorf D C, Chakraborty S, Faria W M, Mesquita L F, Guilherme L R G, Curi N. 2020. Assessing models for prediction of some soil chemical properties from portable X-ray fluorescence (pXRF) spectrometry data in Brazilian Coastal Plains. *Geoderma* **357**: 113957.

Baver L D, Gardner W H, Gardner W R. 1972. Soil Physics. 5th Edn. John Wiley & Sons. New York.

Bouyoucos G J. 1936. Directions for making mechanical analyses of soils by the hydrometer method. *Soil Science*. **42**: 225-230

Breiman L. 2001. Random Forests. *Mach. Learn*. **45**, 5–32.

Brungard C W, Boettinger J L, Duniway M C, Wills S A, Edwards T C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma* **239–240**: 68–83.

Chagas C da S, de Carvalho Junior W, Bhering S B, Calderano Filho, B. 2016. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. *Catena*. **139**: 232–240.

Chakraborty, S, Li B, Weindorf D C, Deb S, Acree A, De P, Panda P. 2019. Use of portable X-ray fluorescence spectrometry for classifying soils from different land use land cover systems in India. *Geoderma*. **338**: 5–13.

Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L, Wehberg J, Wichmann V, Böhner J. 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev*. **8**: 1991–2007.

Dantas A A A, de Carvalho L G, Ferreira, E. 2007. Climatic classification and tendencies in Lavras region, MG. *Ciência e Agrotecnologia* **31**: 1862–1866.

Dearing J A. 1994. Environmental Magnetic Susceptibility. *In* Bartington Instruments. Witney, Oxon, Englang.

Ganasri B P, Ramesh H. 2016. Assessment of soil erosion by RUSLE model using remote sensing and GIS - A case study of Nethravathi Basin. *Geosci. Front*. **7**. 953–961.

Gee G W, Or D. 2002. Particle size analysis. *In* Dane J H, Topp G C (eds.) Methods of Soil Analysis. Part 4. Physical Methods. SSSA, Madison. pp. 255–293.

Grunwald S. 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma*. **152**: 195–207.

Hartemink A E, Minasny B. 2014. Towards digital soil morphometrics. *Geoderma*. **230–231**: 305–317.

Hengl T, Leenaars J G B, Shepherd K D, Walsh M G, Heuvelink G B M, Mamo T, Tilahun H, Berkhout E, Cooper M, Fegraus E, Wheeler I, Kwabena N A. 2017. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. *Nutr. Cycl. Agroecosystems*. **109**: 77–102.

Kämpf N, Marques J J, Curi N. 2012. Mineralogy of Brazilian soils. *In* Ker J C, Curi N, Schaefer C E G R, Vidal-Torrado P (eds.) Pedology—Fundamentals (in Portuguese). Sociedade Brasileira de Ciência do Solo, Viçosa. pp. 81–146.

Kuhn, M., 2008. Building predictive models in R using the caret package. *J. Stat. Softw*. **28**: 1–26.

Lee S, Wolberg G, Shin S Y. 1997. Scattered data interpolation with multilevel B-splines. *IEEE Trans. Vis. Comput. Graph.* **3**: 228–244.

Liaw A, Wiener M. 2002. Classification and Regression by randomForest. *R News*. **2**: 18–22.

Lopes A S, Guilherme L R G. 2016. A career perspective on soil management in the Cerrado region of Brazil. *Adv in Agron*. **137**: 1–72.

Mancini M, Weindorf D C, Chakraborty S, Silva S H G, Teixeira A F S, Guilherme L R G,

Curi N. 2019. Tracing tropical soil parent material analysis via portable X-ray fluorescence (pXRF) spectrometry in Brazilian Cerrado. *Geoderma.* **337**: 718–728.

Marques Jr J, Siqueira D S, Camargo L A, Teixeira D D B, Barrón V, Torrent J. 2014. Magnetic susceptibility and diffuse reflectance spectroscopy to characterize the spatial variability of soil properties in a Brazilian Haplustalf. *Geoderma.* **219–220**: 63–71.

McBratney A B, Santos, M L M, Minasny B. 2003. On digital soil mapping, *Geoderma.* **117**: 3-52.

Minasny B, McBratney A B. 2016. Digital soil mapping: A brief history and some lessons. *Geoderma.* **264**: 301–311.

Pelegrino M H P, Weindorf D C, Silva S H G, de Menezes M D, Poggere G C, Guilherme L R G, Curi N. 2019. Synthesis of proximal sensing, terrain analysis, and parent material information for available micronutrient prediction in tropical soils. *Precis. Agric.* **20**. 746–766.

Poggere G C, Inda A V, Barrón V, Kämpf N, de Brito A D B, Barbosa J Z, Curi N. 2018. Maghemite quantification and magnetic signature of Brazilian soils with contrasting parent materials. *Appl. Clay Sci.* **161**: 385–394.

Quéméneur J J G, Ribeiro A, Paciullo F V P, Heilbron M, Trouw R A J, Valença J G, Noce C M. 2003. Geological map-Folha Lavras scale 1:100000 (in Portuguese). COMIG, Belo Horizonte, MG, Brazil. Available online at http://www.portalgeologia.com.br/index.php/mapa/ (verified on November 19 , 2019).

R Development Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Rawal A, Chakraborty S, Li B, Lewis K, Godoy M, Paulette L, Weindorf DC. 2019. Determination of base saturation percentage in agricultural soils via portable X-ray fluorescence spectrometer. *Geoderma.* **338**: 375–382.

Ribeiro B T, Silva S H G, Silva E A, Guilherme L R G. 2017. Portable X-ray fluorescence (pXRF) applications in tropical Soil Science. *Ciência e Agrotecnologia.* **41**: 245–254.

Robinson G W. 1922. A new method for the mechanical analysis of soils and other dispersions. *J. Agric. Sci.* **12**: 306–321.

Rosemary F, Vitharana U W A, Indraratne S P, Weerasooriya R, Mishra U. 2017. Exploring the spatial variability of soil properties in an Alfisol soil catena. *Catena.* **150**: 53–61.

Scudiero E, Skaggs T H, Corwin D L, 2015. Regional-scale soil salinity assessment using Landsat ETM + canopy reflectance. *Remote Sens. Environ.* **169**: 335–343.

Silva S H G, Poggere G C, de Menezes M D, Carvalho G S, Guilherme L R G, Curi N. 2016. Proximal sensing and digital terrain models applied to digital soil mapping and modeling of Brazilian Latosols (Oxisols). *Remote Sens.* **8**: 614-635.

Silva S H G, Teixeira A F S, de Menezes M D, Guilherme L R G, Moreira F M S, Curi N. 2017. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). *Ciência e Agrotecnologia.* **41**: 648–664.

Silva S H G, Weindorf D C, Pinto L C, Faria W M, Acerbi Junior F W, Gomide L R, de Mello J M, de Pádua Junior A L, de Souza I A, Teixeira A F S, Guilherme L R G, Curi N. 2020. Soil texture prediction in tropical soils: A portable X-ray fluorescence spectrometry approach. *Geoderma.* **362**: 114136.

Stockmann U, Cattle S R, Minasny B, McBratney A B. 2016. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. *Catena.* **139**: 220–231.

Teixeira P C, Donagemma G K, Fontana A, Teixeira W G. 2017. Manual methods of soil analysis. 3th Edn (in Portuguese). Brasília, DF. pp 574.

Teixeira A F S, Pelegrino, M H P, Faria, W M, Silva S H G, Gonçalves M G M, Acerbi Júnior, F W , Rezende Gomide, L., Linares Pádua Júnior, A., de Souza, I.A., Chakraborty, S., Weindorf, D.C., Roberto Guimarães Guilherme, L., Curi, N., 2020. Tropical soil pH and sorption complex prediction via portable X-ray fluorescence spectrometry. *Geoderma*. **361**: 114132.

Weindorf D C, Bakr N, Zhu Y. 2014. Advances in Portable X-ray Fluorescence (PXRF) for Environmental, Pedological, and Agronomic Applications. *Adv. Agron*. **128**: 1–45.

Weindorf D C, Chakraborty S. 2016. Portable X-ray Fluorescence Spectrometry Analysis of Soils. *In* Hirmas D (ed.) Methods of Soil Analysis. Soil Science Society of America, Madison. pp. 1–8.

Zeraatpisheh M, Ayoubi S, Jafari A, Tajik S, Finke P. 2019. Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. *Geoderma*. **338**: 445–452.

Zhang Y, Hartemink A E. 2019. Soil horizon delineation using vis-NIR and pXRF data. *Catena*. **180**: 298–308.

Zhu Y, Weindorf D C, Zhang W. 2011. Characterizing soils using a portable X-ray fluorescence spectrometer: 1. Soil texture. *Geoderma*. **167–168**: 167–177.

4 ARTIGO 2. AVAILABLE MICRONUTRIENTS PREDICTION IN TROPICAL SOILS VIA PROXIMAL SENSING AND TERRAIN ANALYSIS

*Article prepared according to the rules of Geoderma Regional (preliminary version).

Luiza Maria Pereira Pierangeli; Sérgio Henrique Godinho Silva; Michele Duarte de Menezes; João José Marques; David C. Weindorf; Nilton Curi

ABSTRACT

Soil mapping is a crucial activity for detailing spatial information about soils in areas of interest. For this reason, the use of digital elevation models (DEM) is essential in the representation of terrain topography, enabling the generation of terrain attributes (TA) that support the digital soil mapping. Besides, the use of proximal sensors such as portable X-ray fluorescence spectrometry (pXRF) and susceptibilimeter, that provides the magnetic susceptibility (MS) of materials has also been a great ally in digital soil mapping. This work aimed to model and predict the available contents of B, Cu, Fe, Mn, and Zn from pXRF and MS data in addition to terrain attributes (TAs) via random forest algorithm. A total of 78 soil samples were collected from A and B horizons, representing Inceptisols, Oxisols, and Ultisols. The samples were analyzed by MS and pXRF and subjected to laboratory analyses to determine the available contents of B, Cu, Fe, Mn, and Zn. Seventeen TAs were generated from a 12.5 m resolution DEM. These data were divided into six datasets: TA; TA+pXRF; TA+MS; TA+MS+pXRF; MS+pXRF; pXRF. The samples were divided into A and B horizons separately and combined (A+B). RF was used to predict the available contents of B, Cu, Fe, Mn, and Zn, and the models were validated. Root mean square error (RMSE) and coefficient of determination (R^2) were used to determine the performance of the models. Finally, the best models were spatialized to cover the entire study area. The combination of pXRF information with TA improved the prediction of available Cu, Fe, and Mn in the A horizon. Available B and Zn were better predicted using only TA information for both horizons, whereas available Cu and Mn were better predicted using just pXRF information for B horizon. The MS information, in combination with pXRF, allowed a better prediction of available Fe in the B horizon. The combination of MS, pXRF, and TA information allowed satisfactory predictions of available micronutrients, especially for Cu, Fe, and Mn.

Keywords: Digital soil mapping; pXRF; magnetic susceptibility; random forest; tropical soils.

4.1 INTRODUCTION

Micronutrients (B, Cu, Fe, Mn, and Zn) are essential elements for plant growth. These elements (micronutrients) are needed in smaller quantities than macronutrients, but their limitations can affect plant growth. B participates in biological processes in the plant and also constitute the cell wall of plants, while Cu, Fe, Mn, and Zn are constituent and participate in the activation of enzymes; besides, Mn is linked to chlorophyll synthesis (Abreu et al., 2007). Micronutrient deficiencies are a limiting factor for annual crop production in the tropics; the main reason is that weathered soils in this region are acidic and have a low natural content of some micronutrients (Fageria and Stone, 2008).

The method used to quantify micronutrient content in the soil is expensive, time-consuming, and generate chemical waste. In this sense, it is advantageous to apply new instruments that provide elementary information for soil samples quickly, economically, and in a sustainable matter. Some examples are the portable X-ray fluorescence (pXRF) spectrometer, susceptibilimeter, visible, and near-infrared (Vis-NIR) spectroscopy. PXRF can quantify different elements in the soil, as each element produces specific fluorescence energy (Ribeiro et al., 2017; WeindorF et al., 2014). Another technique is the use of a susceptibilimeter to quantify the magnetic susceptibility (MS) of samples. Besides, both methods are non-invasive, non-destructive, and can be used in both field and lab conditions.

Several soil studies have been conducted using pXRF and MS separately, but their combined use still requires investigations. Teixeira et al. (2018) predicted exchangeable Ca^{2+}, V %, and pH in areas cultivated with coffee and eucalyptus through the use of pXRF data. Based only on pXRF information, soil fertility properties were predicted in Cerrado soils (Lima et al., 2019). Andrade et al. (2020), used pXRF information to study and predict chemical soil properties in Brazilian Coastal Plains region. MS information and diffuse reflectance spectra were used in combination to characterize the spatial variability of the properties of an Haplustalf in Brazil (Marques jr. et al., 2014). Naimi and Ayoubi (2013) used MS to estimate metals contents in an industrial site in Iran and proposed the use of MS as an indicator of metal contamination of soils from anthropogenic sources. Multiple linear regression showed that the combination of soil properties and MS could explain the variability of heavy metals in different land uses areas in Iran (Ayoubi et al., 2018).

Recently, the aforementioned sensors have been used to assess the spatial variability of soils in combination or independently with other variables (Duda et al., 2017; Silva et al., 2016d), thus contributing to digital soil mapping (DSM). DSM estimates soil properties through soil-environment relationships, using variables related to the soil forming factors (climate, organisms, parent material, relief, time - Jenny, 1941), but mainly relief. Through the widely available digital elevation models (DEM), a continuous representation of terrain, it is possible to derive other relief-related variables, the so-called terrain attributes (TA), such as slope, curvature, distance to channel network, etc. (Grunwald, 2009; McBratney et al., 2003; Minasny and McBratney, 2016). Thus, the use of proximal and remote sensing combined with TAs has been little explored, but it can become a powerful tool for enhancing predictions of soil classes and attributes (Hartemink and Minasny, 2014; Hengl et al., 2017; Shahbazi et al., 2019; Weindorf et al., 2014).

Therefore, the objective of this study were to: (1) characterize soils with pXRF of an experimental farm in Brazil with diverse land uses, (2) predict plant available contents of B, Cu, Fe, Mn, and Zn and evaluate the importance of MS, pXRF and TA information separately and in association as predictor variables, (3) assess the importance of such variables in micronutrient prediction, and (4) produce spatial predictions of such micronutrients for the A and B horizons of the soils of the study area.

4.2 MATERIALS AND METHODS

4.2.1 Study Area

The study was conducted in Fazenda Palmital (Fig. 1.), located in Ijaci county, Minas Gerais, Brazil, between UTM longitudes 506,793 and 508,882 m and latitudes 7,659,470 and 7,660,685 m zone 23K, datum WGS84. This study area is covers approximately 117 ha, featuring agricultural crops, pasture, native forest, and eucalyptus, mahogany and Pinus plantations. It is located in the Rio Grande watershed region, the altitude in the study area ranges from 793 to 868 m. The geologic and soil parent material information is derived from a 1:100,000 map (MAPA GEÓLOGICO – FOLHA LAVRAS, 2003), and based on this, the parent material of the area are gneisses TTG, metamorphosed limestone and phyllite-limestone, alluvial sediment from quaternary. The regional climate is humid subtropical with dry winters and hot summers (Köppen Cwa), with mean annual temperature of 21 ºC, and mean precipitation of 1500 mm (Dantas et al., 2007).
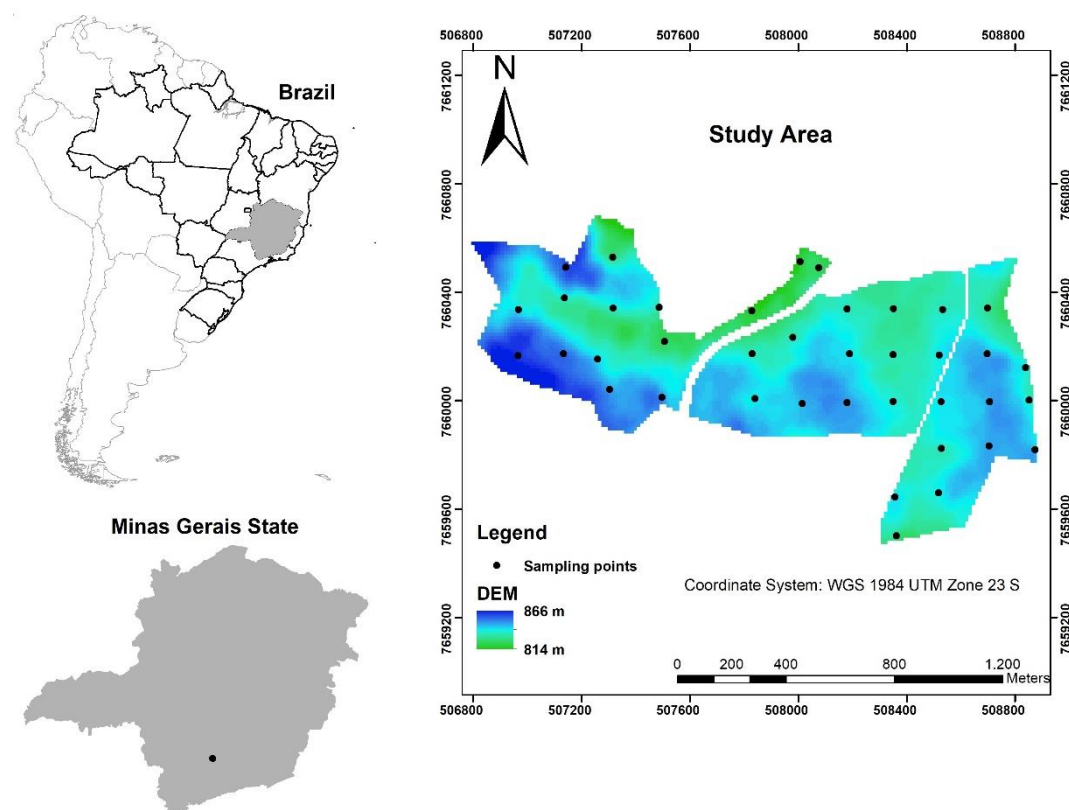
Fig. 1. Study area and sampling sites of Fazenda Palmital, MG, Brazil.

4.2.3 Soil Sampling and Laboratory Analyses

A total of 39 sites were sampled in A and B horizons following a regular-grid design with 173 m of distance between sampling places, totalizing 78 samples. The soil samples were collected in July 2018, in the winter season. Soil samples were air-dried, ground and passed through a 2 mm sieve. Samples were then subjected to laboratory analyses for the determination of available B, Cu, Fe, Mn, and Zn with Mehlich-1 solution (Mehlich, 1953). Quantification was made via atomic absorption spectrophotometry on an AAS 800 (Perkin Elmer, Waltham, MA, USA) (Jackson, 1958). And B was extracted with hot water and determined by Azomethine H colorimetric method.

The MS was determined in a low frequency (0.47 kHz) using 10 g of air-dried and sieved soil in a Bartington MS2B susceptibility meter (Dearing, 1994).

Elemental analyses were performed using a Vanta series pXRF (Olympus, Waltham, MA, USA) spectrometer to characterize all the soil samples. The instrument scans via two beams in sequence such that one whole scan was completed in 60 s and is integrated with

GeoChem software (Weindorf and Chakraborty, 2016). This pXRF features a Rh X-ray tube operated at 8–50 kV as the excitation source, detecting elements ranging from Mg to U.

Before scanning, the instrument was calibrated with a factory calibration alloy coin. The accuracy of the equipment was evaluated through scanning standard reference materials 2710a and 2711a certified by the National Institute of Standards and Technology (NIST). The recovery values for the elements identified in all samples and used in this work follow (2710a/2711a): Ag (0/0); Al (0.61/0.61); As (1.06/1.43); Ca (0.65/0.88); Cd (0/0.91); Cu (1.02/1.01); Fe (0.92/ 0.96); Hg (2.73/1.75); K (0.83/0.85); Mn (0.93/0.96); Mo (0/0); Ni (0/1.15); P (0.63/0.52); Pb (1.02/ 1.08); Rb (0/0); S (0/0); Se (0/0); Si (0.57/0.61); Sr (0.96/0.96); Th (0/0); Ti (0.90/ 0.94); U (0.99/0); V (0/0.93); y (0/); Zn (1.00/1.06).

4.2.3 Terrain variables

Topography is one of the main forming factors, therefore a DEM with 12.5 m resolution from the ALOS PALSAR Global Radar Imagery was downloaded (https://search.asf.alaska.edu/#/). From this DEM, 17 TA were selected using and SAGA GIS, including: slope, aspect, cross-sectional and longitudinal (CsC and LC, respectively) curvature, convergence index (CI), flow accumulation (flow), topographic wetness index (TWI), LS factor (lsf), channel network base level (CNBL), vertical distance to channel network (VDCN), valley depth (VD), relative slope position (RSP), catchment area (CA), modified catchment area (MCA), closed depressions (CD), catchment slope (CS) and, SAGA wetness index (SWI). The extraction of the DEM values from each terrain attribute and altitude through geo-referenced sampling sites along with the elemental data from pXRF and TA data for each sampling site were adjusted to predict available B, Cu, Fe, Mn, and Zn through a random forest model.

4.2.4 Modeling and validation of the predictions

Soil samples were randomly separated into modeling and validation data sets, respectively, consisting of 70% and 30% of the total data. Also, the samples were subdivided and modeled in two ways: i) specific models, according to the two horizon, with n = 39 for each depth, with 27 samples for modeling and 12 for validation; and ii) general model, including all samples (n = 78, 54 for modeling and 24 for validation).

In order to adjust the models for predicting available B, Cu, Fe, Mn, and Zn contents pXRF, MS, and TA information were divided in sex datasets: (i) TA, (ii) TA + pXRF, (iii) TA + MS, (iv) TA + pXRF + MS, (v) MS + pXRF, and (vi) pXRF.

The RF analysis was performed in R software (3.6.1) (R Development Core Team, 2009), using the caret package (Kuhn, 2008), with the following parameters established: number of trees of the model (ntrees) = 1000, number of variables in each node (nodesize) = 5, and number of variables used in each tree (mtry) = 3. Also, the RF model determines the importance of each variable. This tool is a powerful and significant measure of the relative importance of one independent variable (Liaw and Wiener, 2002). The higher the %IncMSE value, the more critical the variable to the prediction model.

The performance of each model was calculated from the validation samples using the indices: coefficient of determination (R^2), RMSE (root mean square error) (Eq. 1). The RMSE was normalized (nRMSE) to compare variables at different scales (Eq. 2):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

(1)

$$nRMSE = \frac{RMSE}{\sigma}$$

(2)

where n: number of observations, X_{model}: estimated value by the model, X_{obs}: measured value by chemical analysis, σ: standard deviation.

The best models for B, Cu, Fe, Mn, and Zn were used in the creations of maps to the entire area. For the soil mapping procedure, TA information was continuously available for the entire study area, but variables obtained from pXRF and MS data at the 39 sampled sites had to be extrapolated to the entire area. The pXRF and MS data were modeled across the entire area through multilevel B-spline interpolation (Lee et al., 1997). These maps were validated through calculations of R^2, RMSE.

4.3 RESULTS AND DISCUSSIONS

4.3.1 Characterization of soils properties

The descriptive statistics of micronutrient available contents for A and B horizons shows the high variability of these elements in the area (Table 1), demonstrated by the ranges

of minimum, maximum, mean, and coefficient of variation (CV) values. This high variability of the data was caused by the diverse land uses, soil management, including areas with and without fertilizer application, soil classes and parent materials. Such variability can generate more robust models for different conditions, since the samples represent a wide range of values of the analyzed properties, such as Cu from 0.10 to 4.08 mg kg^{-1}, Fe from 14.09 to 343.02 mg kg^{-1}, Mn from 1.08 to 180.55 mg kg^{-1}, and the Zn from 0.10 to 86.63 mg kg^{-1}. However, B had a small variation, ranging from 0.02 to 0.22 mg kg^{-1}.

Part of the studied area has been cultivated with agricultural crops and forest farming. Therefore, soil management that includes the addition of fertilizers explains the higher values of Cu, Mn, and Zn in the A horizon that has more OM helping in the complexation of these nutrients. Similar results were also found in soils of the Brazilian Coastal Plains region, where the micronutrient total contents were also higher in A horizon (Andrade et al., 2020).

Table 1: Summary statistics of available B, Cu, Fe, Mn, and Zn information (mg kg^{-1}) analyses of A and B horizons, and A+B horizons of soils in Brazil.

Micronutrient	Soil horizon	n	Min[a]	Max[b]	Mean	SD[c]	CV[d]
				mg kg^{-1}			%
B	A	39	0.03	0.22	0.11	0.05	45.45
	B	39	0.02	0.22	0.08	0.05	62.50
	A+B	78	0.02	0.22	0.10	0.05	50.00
Cu	A	39	0.18	4.08	0.84	0.73	86.90
	B	39	0.10	2.57	0.71	0.52	73.24
	A+B	78	0.10	4.08	0.78	0.64	82.05
Fe	A	39	17.99	230.24	63.89	52.89	82.78
	B	39	14.09	343.02	49.47	57.91	117.06
	A+B	78	14.09	343.02	56.68	55.57	98.04
Mn	A	39	8.35	180.55	32.95	33.02	100.21
	B	39	1.08	72.46	10.01	12.71	126.97
	A+B	78	1.08	180.55	21.48	27.40	127.56
Zn	A	39	0.49	86.63	7.87	13.47	171.16
	B	39	0.10	4.79	1.05	1.09	103.81
	A+B	78	0.10	86.63	4.46	10.09	226.23

[a]minimum, [b]maximum, [c]standard deviation, [d]coefficient of variation.

4.3.2 Characterization of soils by pXRF

Characterization of soils between horizons by pXRF is showed in Table 2. The main elements present in these soils was Al, followed by Si and Fe. Similar to the results found by Lima et al. (2019) studying Cerrado soils. The Ca contents ranged from 0 to 3742 mg kg^{-1} (mean 600 mg kg^{-1}) and the highest values were found on the surface horizon. Liming and gypsum application are common practice in the tropical soils (Lopes and Guilherme, 2016). The higher content of Al, Mn, P, and Si were higher in the A horizon while Fe was higher in B horizon. Teixeira et al. (2020) observed similar tendencies working in tropical soils.

Among the other elements found by the pXRF, the higher contents of As, Cr, Ni, V, and Y were found in the B horizon, while Cu, Ti, and Sr were found in higher contents in the A horizon. These results resemble those found by Silva et al. (2017) in Brazilian Cerrado soils, containing higher content of Ti and Sr in the A horizon and V in the B horizon. MS content presented higher values in B horizon, and this is likely because of the lower organic matter contents and higher content of Fe oxides in this horizon, especially in the form of maghemite, since it is the only clay mineral with powerful magnetism in tropical and subtropical soils (Poggere et al., 2018). In addition, the parent material is directly related with the MS in soils.

Table 2: Summary statistics of pXRF data (mg kg^{-1}) and MS ($\times 10^{-6}$ m^3 kg^{-1}) analyzing A horizons of soils from Palmital farm in Brazil.

Parameter	Horizon	Al	As	Ca	Cr	Cu	Fe	K	Mn	Ni	P	Pb	Rb	S	Si	Sr	Ti	V	Y	Zn	MS
Min[a]		44781	6	0	56	16	24863	0	68	16	0	9	7	0	41149	10	6052	36	9	28	2
Max[b]		105232	38	3742	150	59	68014	15245	1450	42	1602	37	112	635	144397	74	10970	98	22	284	57
Mean	A	72441	13	1082	81	28	43567	4772	238	28	664	19	38	242	70634	30	8385	68	14	62	13
SD[c]		16494	5	896	18	8	9604	4765	263	7	342	6	31	120	21794	17	1370	14	3	48	11
CV[d] (%)		23	41	83	22	29	22	100	110	24	51	32	80	49	31	56	16	20	21	77	83
Min[a]		49650	8	0	52	17	27261	0	64	14	74	6	5	0	39146	8	5144	41	7	22	2
Max[b]		98170	24	2018	150	44	75936	15194	947	55	777	32	122	233	116214	81	10829	94	24	81	75
Mean	B	70835	14	117	85	27	47983	4189	176	30	351	19	38	85	65331	27	8369	69	15	43	15
SD[c]		11898	4	390	24	7	10812	4399	187	11	162	6	32	67	20403	19	1578	15	3	16	14
CV[d] (%)		17	26	334	28	25	23	105	106	36	46	34	85	78	31	69	19	22	23	37	95
Min[a]		44781	6	0	52	16	24863	0	64	14	0	6	5	0	39146	8	5144	36	7	22	2
Max[b]		105232	38	3742	150	59	75936	15245	1450	55	1602	37	122	635	144397	81	10970	98	24	284	75
Mean	A+B	71638	13	600	83	27	45775	4480	207	29	507	19	38	164	67982	29	8377	69	14	52	14
SD[c]		14310	5	841	21	7	10400	4565	229	9	309	6	31	124	21142	18	1468	14	3	36	13
CV[d] (%)		20	34	140	25	27	23	102	111	31	61	33	82	76	31	62	18	21	22	70	90

[a]minimum, [b]maximum, [c]standard deviation, [d]coefficient of variation.

4.3.3 Model performance to predicting available contents of B, Cu, Fe, Mn, and Zn

The R² values obtained from the comparison between the observed and estimated values of available B, Cu, Fe, Mn, and Zn generated by RF are presented in Table 3. In general, the prediction models for the combined horizons (A+B) presented higher R^2 and smaller RMSE than the models using the separate horizon data.

Table 3: Root mean square error (RMSE) and R^2 of random forest models for predicting B, Cu, Fe, Mn, and Zn in soils from Palmital farm, Brazil.

Model	Parameter[1]	TA		TA+pXRF		TA+MS		TA+MS+pXRF		MS+pXRF		pXRF	
		RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
	A	0.06	0.01	0.05	0.00	0.06	0.01	0.05	0.00	0.05	0.11	0.05	0.08
B	B	0.05	0.11	0.05	0.14	0.05	0.09	0.04	0.02	0.04	0.00	0.04	0.02
	A+B	0.04	0.37	0.04	0.25	0.04	0.37	0.04	0.25	0.05	0.14	0.05	0.11
	A	0.69	0.18	0.58	0.01	0.63	0.40	0.57	0.00	0.59	0.02	0.58	0.03
Cu	B	0.50	0.02	0.27	0.74	0.51	0.00	0.27	0.73	0.27	0.75	0.24	0.80
	A+B	0.54	0.13	0.42	0.41	0.53	0.12	0.43	0.39	0.47	0.31	0.47	0.31
	A	58.05	0.00	44.53	0.40	52.04	0.18	43.71	0.40	46.63	0.37	47.41	0.37
Fe	B	51.33	0.14	39.26	0.61	50.57	0.07	40.92	0.50	33.69	0.68	34.68	0.63
	A+B	37.14	0.51	34.98	0.59	37.94	0.49	43.64	0.39	35.86	0.57	36.00	0.54
	A	28.36	0.10	12.54	0.86	28.32	0.10	12.46	0.87	11.92	0.85	11.88	0.85
Mn	B	8.94	0.13	6.70	0.49	8.99	0.13	6.80	0.47	6.35	0.55	6.15	0.59
	A+B	21.83	0.29	15.62	0.66	19.48	0.41	15.93	0.66	15.74	0.67	15.70	0.67
	A	9.89	0.00	9.48	0.02	12.06	0.00	8.48	0.02	4.96	0.01	7.72	0.00
Zn	B	0.79	0.02	0.89	0.23	0.72	0.27	0.86	0.25	0.90	0.24	0.88	0.22
	A+B	2.84	0.50	3.34	0.42	3.08	0.31	3.67	0.29	3.18	0.43	4.12	0.21

[1] RMSE: root mean square error.

For available B and Zn, only TA with A+B horizons data combined, the prediction models achieved maximum R^2 of 0.37 and 0.50, respectively. For available Cu, TA+pXRF with A+B horizons data found the optimal predictions for A horizon (R^2=0.41), and pXRF with B horizon data provided the highest R^2 (0.80) for B horizon. Utilizing TA+pXRF with A+B horizons data, Fe prediction obtained a moderate R^2 (0.59) for the models representing A horizon. For available Fe in B horizon, MS+pXRF with B horizon data performed better than the other models, achieving an R^2 of 0.68. For Mn prediction, TA+pXRF+MS with A horizon data found the best R^2 of 0.87 for A horizon, while pXRF with A+B horizons data provided maximum R^2 of 0.67 for B horizon. The use of only TA information did not contribute satisfactorily in most predictions. In general, when pXRF information was added to TA information, a little improvement in the models' performance for A and B horizons was achieved.

In Table 3, the validation errors assessed by RMSE demonstrate that the addition of pXRF information decreased the values of RMSE. For Cu and Mn models, lower RMSE were obtained for all datasets, whereas for Zn, the error of predictions increased in the B horizon. For Fe predictions, A+B horizons data featured higher RMSE in one of the models using pXRF (TA+MS+pXRF). For available B prediction, RMSE values were similar between the datasets, and there was not a significant improvement with the addition of pXRF information.

Fig. 2 presents the variation of normalized RMSE (nRMSE) of the models. Different combinations of dataset and soil information strongly influenced models accuracy. For Zn, the lowest values of nRMSE were found for A+B horizons. Whereas, the highest values were found for B in the A horizon. The best models were selected by observing the highest R^2 and the lowest RMSE values (Table 3). Models containing the best predictions for each available micronutrient are shown in Table 4.

In general, the models had a satisfactory predictive capacity for Fe, Mn, and Zn, although they were unsatisfactory for B and Cu prediction. The results were better for B horizon except for Mn for A horizon. For A horizon, the use of the combined data set presented better results than A data only. For this reason, the use of the combined data set may improve the prediction of attributes. These results also indicate the potential of using pXRF to aid in the prediction of micronutrients in Brazilian soils. Similar results were found in Brazil, where TA, pXRF, and parent material information were used to predicted Cu, Fe, Mn, and Zn (Pelegrino et al., 2019). In the study, the best models to predict Cu, Fe, Mn, and

Zn included pXRF information and reached R^2_{adj} of 0.74, 0.86, 0.71, and 0.28, respectively. The small difference in the performance of predictive models among these Brazilian soils is probably due to the similarity in the properties of the soils studied. However, further studies are needed to develop the best methodology to deliver better results for each soil-environmental condition.

Comparatively, the works developed in Sub-Saharan region were able to predict available B, Cu, Fe, Mn, and Zn with moderate accuracy using GIS information (Hengl et al., 2017). Hengl et al. (2017) found an R^2 of 0.41 for B, 0.54 for Cu, 0.68 for Fe, 0.53 for Mn, and 0.47 for Zn. Shahbazi et al. (2019) predicted B contents (maximum R^2=0.22) in different soil depth and other nutrients in North West Iran using GIS information, thus, presenting a low R^2 similar to the current work. In a different approach, Camargo et al. (2018) predicted potentially toxic elements in tropical soils utilizing magnetic susceptibility (MS) and diffuse reflectance spectra. The best predictive models for Cu, Mn, and Zn semi-total contents had an R^2 of 0.89, 0.77, and 0.55, respectively using MS-calibrated models. Finally, in the present study the combination of pXRF with other information improved the predictions of available micronutrients content. Concerning the use of the datasets, some of the best performances were achieved using the A+B data, which can be assigned to the most extensive dataset used to produce an RF model (Rawal et al., 2019b; Shahbazi et al., 2019; Silva et al., 2017a).
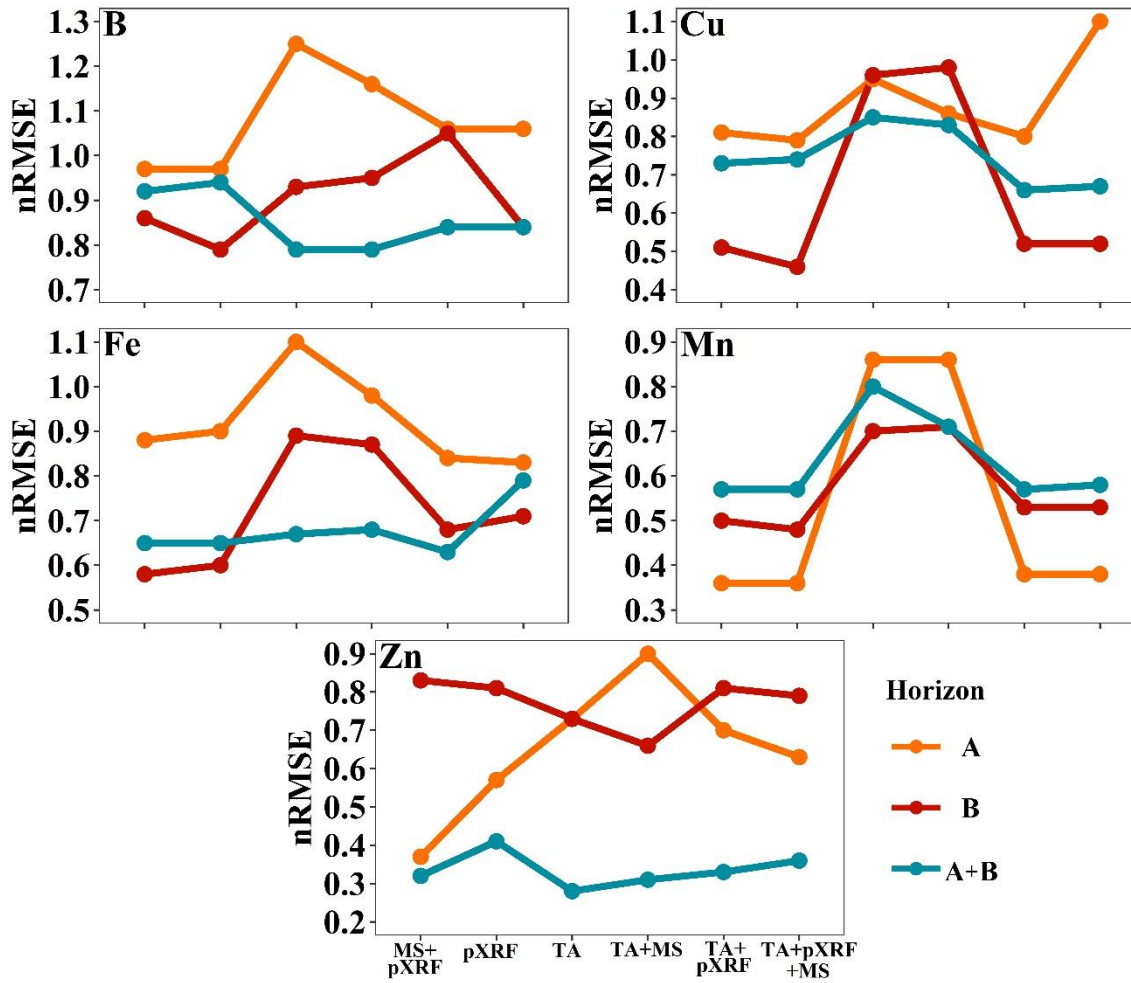
Fig. 2. Normalized root mean square error (nRMSE) of predictions for available B, Cu, Fe, Mn, and Zn using different proximal sensor with random forest of Brazilian soils.

4.3.4 Variables Importance

The best models to predict B, Cu, Fe, Mn, and Zn had their variable importance calculated by the percentage of increment of Mean Square Error (%IncMSE). In order to choose the best models per horizon, the criteria of the highest R^2 and the smallest RMSE were adopted. The five most essential variables per model are shown in Fig 3.
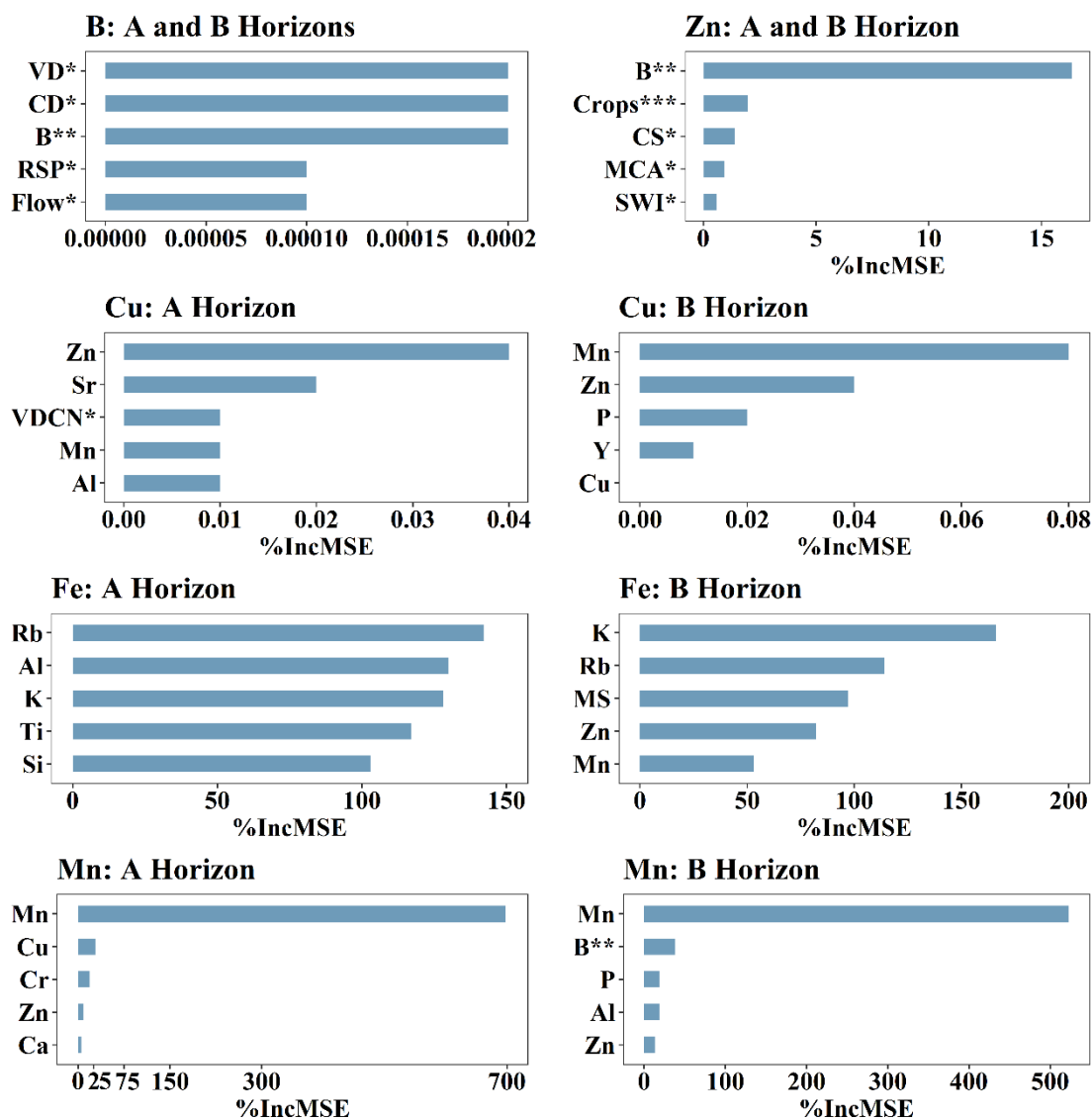
Fig. 3. Plots of relative importance of variables in RF model for the B, Cu, Fe, Mn, and Zn best prediction models in Brazilian soils. * Terrain attributes; ** horizon; *** land use; VD, valley depth; CD, closed depressions; RSP, relative slope position; flow, flow accumulation; CS, catchment slope; MCA, modified catchment area; SWI, SAGA wetness index; VDCN, vertical distance to channel network; MS, magnetic susceptibility; Crops, cultivated area.

In general, the best variables were provided by the pXRF information. These variables may be related to the parent material of tropical soils (Mancini et al., 2019). In the case of the Al, that appeared in some models, it is usually related to the weathering of tropical soils (Kämpf et al., 2012). For B, the most essential variable was closed depression (CD), followed by B horizon, as a covariate factor, and valley depth (VD). For available Cu and Mn, among the five most essential variables were Zn and Mn, both obtained by pXRF. For the available

Zn, the most important variables were B horizon, as a covariate factor, and Crops, meaning that the area was cultivated, followed by catchment slope (CS), a TA information.

For Fe, the most important variables were Rb and K. Although for the B horizon model, MS was classified as the third most important variable. Bahia et al. (2017) found a high correlation between MS and iron extracted with dithionite-citrate-bicarbonate (0.95) and iron extracted with ammonium oxalate (0.87) in Brazilian soils. Fe_d is related to high-crystallinity pedogenetic iron oxides (hematite, maghemite), and Fe_o is related to low-crystallinity pedogenetic iron oxides (ferrihydrite) (Poggere et al., 2018). The use of MS information may have improved Fe prediction in the B horizon, distinguishing different Fe-oxides.

4.3.5 Spatial Predictions of B, Cu, Fe, Mn, and Zn

RF was used in the spatial prediction of the studied properties (Fig. 4) using the best models determined in the previous section. The available B varied from 0.17 to 0.05 mg kg^{-1} in the A horizon and from 0.16 to 0.05 mg kg^{-1} in the B horizon. Available Cu content had a variation of 2.37 to 0.40 mg kg^{-1} in the A horizon, while the B horizon content ranged from 1.85 to 0.26 mg kg^{-1}. The available Fe varied from 164.91 to 31.42 mg kg^{-1} in the A horizon and from 149.35 to 28.66 mg kg^{-1} in the B horizon. Available Mn content had a more extensive variation of 79.83 to 2.32 mg kg^{-1} in the A horizon, while the B horizon content ranged from 134.79 to 12.55 mg kg^{-1}. Furthermore, the Zn content ranged from 58.90 to 1.27 mg kg^{-1} in the A horizon and from 10.20 to 0.31 mg kg^{-1} in the B horizon.

A horizon presented a higher content of available Cu and Zn in the area. This may be related to the association of these elements with soil organic matter, which retains them, in addition to fertilizers application. Conversely, Fe and Mn had a higher content in the B horizon. In highly weathered Brazilian soils rich in Fe- and Mn-oxides, these oxides associate with micronutrients through adsorption mechanisms. The maps generated for both horizons showed very similarity between the concentration for available B, although the accuracy of the model was not satisfactory as presented in a previous section.
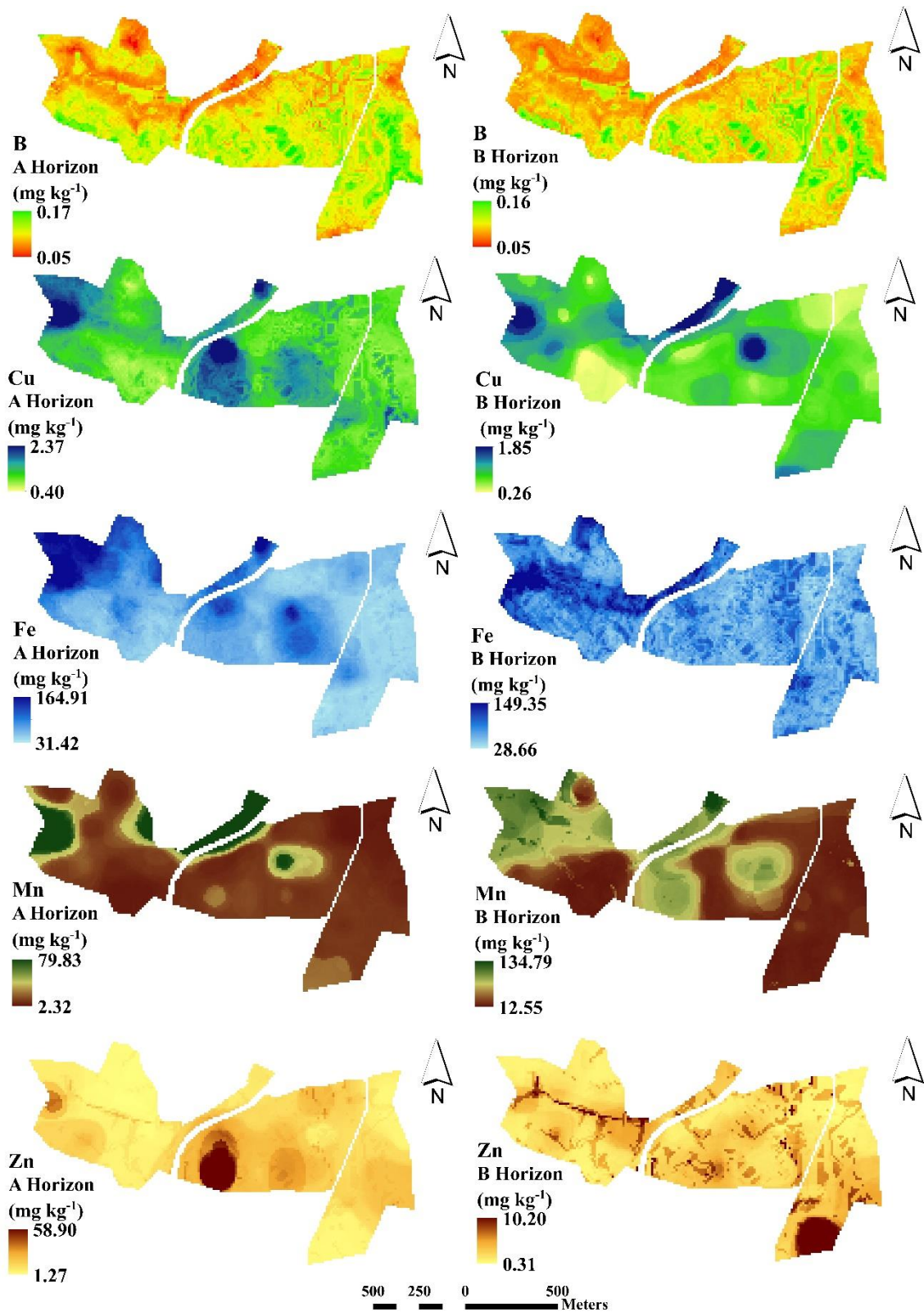
Fig. 4. Maps of available B, Cu, Fe, Mn, and Zn contents in A and B horizon obtained from the best models of random forest for the study area in Brazilian soils.

4.4 CONCLUSIONS

The combination of MS, pXRF, and TA information allowed for satisfactory predictions of available Cu, Fe, Mn, and Zn with the use of random forest algorithms in Brazilian soils. Available B and Zn contents were better predicted using only TA information. Available Fe and Mn contents were the only elements that reached a good prediction using MS information combined with the other information. The use of pXRF information improved models when such data was associated with TA information. Therefore, for the region of study, the combination of proximal and remote sensing data is recommended for digital mapping and modeling of available micronutrient contents, mainly Fe, Mn and Zn.

REFERENCES

Abreu, C.A. de, Lopes, A.S., Santos, G.C.G. dos, 2007. Micronutrientes, in: Fertilidade Do Solo. SBCS, Viçosa, MG, p. 1017.

Andrade, R., Silva, S.H.G., Weindorf, D.C., Chakraborty, S., Faria, W.M., Mesquita, L.F., Guilherme, L.R.G., Curi, N., 2020. Assessing models for prediction of some soil chemical properties from portable X-ray fluorescence (pXRF) spectrometry data in Brazilian Coastal Plains. Geoderma 357, 113957. https://doi.org/10.1016/j.geoderma.2019.113957

Ayoubi, S., Jabbari, M., Khademi, H., 2018. Multiple linear modeling between soil properties, magnetic susceptibility and heavy metals in various land uses. Model. Earth Syst. Environ. 4, 579–589. https://doi.org/10.1007/s40808-018-0442-0

Camargo, L.A., Marques, J., Barrón, V., Alleoni, L.R.F., Pereira, G.T., Teixeira, D.D.B., Bahia, A.S.R. de S., 2018. Predicting potentially toxic elements in tropical soils from iron oxides, magnetic susceptibility and diffuse reflectance spectra. Catena 165, 503–515. https://doi.org/10.1016/j.catena.2018.02.030

Dantas, A.A.A., Carvalho, L.G. De, Ferreira, E., 2007. Climatic classification and tendencies in Lavras region, MG. Ciência e Agrotecnologia 31, 1862–1866. https://doi.org/10.1590/S1413-70542007000600039

de Souza Bahia, A.S.R., Marques, J., La Scala, N., Pellegrino Cerri, C.E., Camargo, L.A., 2017. Prediction and Mapping of Soil Attributes using Diffuse Reflectance Spectroscopy and Magnetic Susceptibility. Soil Sci. Soc. Am. J. 81, 1450. https://doi.org/10.2136/sssaj2017.06.0206

Dearing, J.A., 1994. Environmental Magnetic Susceptibility: using the Bartington MS2 System. Chi Pub.

dos Santos Teixeira, A.F., Henrique Procópio Pelegrino, M., Missina Faria, W., Henrique Godinho Silva, S., Gabriela Marcolino Gonçalves, M., Weimar Acerbi Júnior, F., Rezende Gomide, L., Linares Pádua Júnior, A., de Souza, I.A., Chakraborty, S., Weindorf, D.C., Roberto Guimarães Guilherme, L., Curi, N., 2020. Tropical soil pH and sorption complex prediction via portable X-ray fluorescence spectrometry. Geoderma 361, 114132. https://doi.org/10.1016/j.geoderma.2019.114132

Duda, B.M., Weindorf, D.C., Chakraborty, S., Li, B., Man, T., Paulette, L., Deb, S., 2017. Soil characterization across catenas via advanced proximal sensors. Geoderma 298, 78–91. https://doi.org/10.1016/j.geoderma.2017.03.017

Fageria, N.K., Stone, L.F., 2008. Micronutrient deficiency problems in South America. Micronutr. Defic. Glob. Crop Prod. 245–266. https://doi.org/10.1007/978-1-4020-6860-7_10

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152, 195–207. https://doi.org/10.1016/j.geoderma.2009.06.003

Hartemink, A.E., Minasny, B., 2014. Towards digital soil morphometrics. Geoderma 230–231, 305–317. https://doi.org/10.1016/j.geoderma.2014.03.008

Hengl, T., Leenaars, J.G.B., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B.M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Wheeler, I., Kwabena, N.A., 2017. Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr. Cycl. Agroecosystems 109, 77–102. https://doi.org/10.1007/s10705-017-9870-x

Jackson, M.L., 1958. Soil chemical analysis prentice Hall. Inc., Englewood Cliffs, NJ 498.

Jenny, H., 1941. Factors of soil formation. 281 pp. New York.

Kämpf, N., Marques, J.J., Curi, N., 2012. Mineralogia de Solos Brasileiros., in: Pedologia Fundamentos. SBCS, Viçosa, MG, p. 343.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28, 1–26.

Lee, S., Wolberg, G., Shin, S.Y., 1997. Scattered data interpolation with multilevel B-splines. IEEE Trans. Vis. Comput. Graph. 3, 228–244.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Lima, T.M. de, Weindorf, D.C., Curi, N., Guilherme, L.R.G., Lana, R.M.Q., Ribeiro, B.T., 2019. Elemental analysis of Cerrado agricultural soils via portable X-ray fluorescence spectrometry: Inferences for soil fertility assessment. Geoderma 353, 264–272. https://doi.org/10.1016/j.geoderma.2019.06.045

Lopes, A.S., Guilherme, L.R.G., 2016. A career perspective on soil management in the Cerrado region of Brazil, in: Advances in Agronomy. Elsevier, pp. 1–72.

Mancini, M., Weindorf, D.C., Chakraborty, S., Silva, S.H.G., dos Santos Teixeira, A.F., Guilherme, L.R.G., Curi, N., 2019. Tracing tropical soil parent material analysis via portable X-ray fluorescence (pXRF) spectrometry in Brazilian Cerrado. Geoderma 337, 718–728. https://doi.org/10.1016/j.geoderma.2018.10.026

Marques Jr., J., Siqueira, D.S., Camargo, L.A., Teixeira, D.D.B., Barrón, V., Torrent, J., 2014. Magnetic susceptibility and diffuse reflectance spectroscopy to characterize the spatial variability of soil properties in a Brazilian Haplustalf. Geoderma 219–220, 63–71. https://doi.org/10.1016/j.geoderma.2013.12.007

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4

Mehlich, A., 1953. Determination of P, Ca, Mg, K, Na and NH4. University of North Carolina, Raleigh, NC, USA.

Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. Geoderma 264, 301–311. https://doi.org/10.1016/j.geoderma.2015.07.017

Pelegrino, M.H.P., Weindorf, D.C., Silva, S.H.G., de Menezes, M.D., Poggere, G.C., Guilherme, L.R.G., Curi, N., 2019. Synthesis of proximal sensing, terrain analysis, and parent material information for available micronutrient prediction in tropical soils. Precis. Agric. 20, 746–766. https://doi.org/10.1007/s11119-018-9608-z

Poggere, G.C., Inda, A.V., Barrón, V., Kämpf, N., de Brito, A.D.B., Barbosa, J.Z., Curi, N., 2018. Maghemite quantification and magnetic signature of Brazilian soils with contrasting parent materials. Appl. Clay Sci. 161, 385–394. https://doi.org/10.1016/j.clay.2018.05.014

Queméneur, J.J.G., Ribeiro, A., Paciullo, F.V.P., Heilbron, M., Trouw, R.A.J., Valença, J.G., Noce, C.M., 2003. Geologia da Folha Lavras 1:100.000, in: Geologia e Recursos Minerais Do Sudeste Brasileiro. Companhia Mineradora de Minas Gerais, Secretaria de Desenvolvimento Econômico, Belo Horizonte, Minas Gerais, pp. 259–319.

Rawal, A., Chakraborty, S., Li, B., Lewis, K., Godoy, M., Paulette, L., Weindorf, D.C., 2019. Determination of base saturation percentage in agricultural soils via portable X-ray fluorescence spectrometer. Geoderma 338, 375–382. https://doi.org/10.1016/j.geoderma.2018.12.032

Ribeiro, B.T., Silva, S.H.G., Silva, E.A., Guilherme, L.R.G., 2017. Portable X-ray fluorescence (pXRF) applications in tropical Soil Science. Ciência e Agrotecnologia 41, 245–254. https://doi.org/10.1590/1413-70542017413000117

Shahbazi, F., Hughes, P., McBratney, A.B., Minasny, B., Malone, B.P., 2019. Evaluating the spatial and vertical distribution of agriculturally important nutrients — nitrogen, phosphorous and boron — in North West Iran. Catena 173, 71–82. https://doi.org/10.1016/j.catena.2018.10.005

Silva, S.H.G., Poggere, G.C., Menezes, M.D. de, Carvalho, G.S., Guilherme, L.R.G., Curi, N., 2016. Proximal sensing and digital terrain models applied to digital soil mapping and modeling of Brazilian Latosols (Oxisols). Remote Sens. 8, 614.

Silva, S.H.G., Teixeira, A.F. dos S., Menezes, M.D. de, Guilherme, L.R.G., Moreira, F.M. de S., Curi, N., 2017. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). Ciência e Agrotecnologia 41, 648–664. https://doi.org/10.1590/1413-70542017416010317

Team, R.D.C., 2009. R: A language and environment for statistical computing.

Teixeira, A.F. dos S., Weindorf, D.C., Silva, S.H.G., Guilherme, L.R.G., Curi, N., 2018. Portable x-ray fluorescence (pXRF) spectrometry applied to the prediction of chemical attributes in inceptisols under different land use. Cienc. e Agrotecnologia 42, 501–512. https://doi.org/10.1590/1413-70542018425017518

Weindorf, D.C., Bakr, N., Zhu, Y., 2014. Advances in Portable X-ray Fluorescence (PXRF) for Environmental, Pedological, and Agronomic Applications, in: Advances in Agronomy. Academic Press, pp. 1–45. https://doi.org/10.1016/B978-0-12-802139-2.00001-9

Weindorf, D.C., Chakraborty, S., 2016. Portable X-ray Fluorescence Spectrometry Analysis of Soils. Methods Soil Anal. 1, 0. https://doi.org/10.2136/methods-soil.2015.0033