University of São Paulo
"Luiz de Queiroz" College of Agriculture

Proximal and remote sensing to soil mineralogy assessment

Nícolas Augusto Rosin

Dissertation presented to obtain the degreee of Master in Science. Area: Soil and Plant Nutrition

Piracicaba
2022

Nícolas Augusto Rosin
Agronomist

Proximal and remote sensing to soil mineralogy assessment

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **JOSÉ ALEXANDRE MELO DEMATTÊ**

Dissertation presented to obtain the degreee of Master in Science. Area: Soil and Plant Nutrition

Piracicaba
2022

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Rosin, Nícolas Augusto

Proximal and remote sensing to soil mineralogy assessment/Nícolas Augusto Rosin. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2022.

77 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. Mapeamento digital de solos 2. Pedometria 3. Espectroscopia de solos 4. pXRF I. Título

CONTENTS

RESUMO

Sensoriamento remoto e próximo para caracterização da mineralogia do solo

A mineralogia é a engrenagem dos processos do solo, desempenhando um papel fundamental em questões relevantes para a humanidade. Porém, o acesso às análises mineralógicas é difícil devido à dificuldade de aquisição pelos métodos tradicionais e formas alternativas de alcançá-las devem ser exploradas. Esta tese foi dividida em dois capítulos que visaram: 1) Compreender os fundamentos das interações da energia na informação do pXRF com ênfase nas formas de ferro, umidade e matéria orgânica do solo para uso na ciência do solo e 2) Mapear as abundâncias dos minerais predominantes para todo o território Brasileiro, em superfície e subsuperfície. Para atingir o primeiro objetivo, três tratamentos de dissolução seletiva foram aplicados para remover: (i) matéria orgânica do solo (−MOS), ii) MOS e formas de ferro pouco cristalinas (−o), iii) MOS e as formas de ferro pouco cristalinas e também as formas bem cristalinas de ferro (−d). Um tratamento adicional iv) incluindo a adição de água (+ W) também foi realizado. O pXRF foi capaz de detectar alterações pelos tratamentos de dissolução seletiva e distribuição granulométrica do solo. Os teores de caulinita, gibbsita, Fe_2O_3, Al_2O_3, SiO_2, TiO_2 e MnO foram quantificados com acurácia satisfatória ($0,61 < R^2 < 0,97$). Fontes de incerteza, principalmente a umidade do solo, devem ser consideradas nas análises. A compreensão dos fundamentos da interação da energia com a matriz da amostra na faixa de raios X é o ponto de partida para a caracterização do solo por meio de pXRF. Para atingir o segundo objetivo, a Biblioteca de Espectral Solos do Brasil (BESB) com dados espectrais no Vis-NIR-SWIR foi utilizada para acessar a abundância de hematita (Hem), goethita (Gt), caulinita (Kt) e gibbsita (Gbs) em amostras de solo do Brasil. Os atributos do terreno (TA) e uma imagem sintética do solo (SySI) com pixel de solo exposto de imagens multitemporais do Landsat (1984 a 2020) foram usados como preditores. Uma nova abordagem foi realizada a fim de obter uma imagem de solo exposto para todo o território brasileiro. O modelo *Random Forest* foi utilizado na predição espacial para obtenção dos mapas minerais e sua incerteza por procedimento de *bootstrapping*. O Hem apresentou os modelos mais acurados com R^2 variando de 0,48 a 0,56, seguido por Gbs (0,42 a 0,44), Kt (0,20 a 0,31) e Gt (0,16 a 0,26). A abordagem proposta foi capaz de revelar a distribuição espacial da abundância relativa de minerais para o território brasileiro. Os mapas minerais estavam de acordo com mapas legados de geologia e pedologia e também com as condições de clima e terreno. A abordagem proposta é um método eficiente para obter informações de mineralogia para grandes áreas.

Palavras-chave: Mapemento digital de solos, Pedometria, Espectroscopia de solos, pXRF

ABSTRACT

Proximal and remote sensing to soil mineralogy assessment

The mineralogy is the gear of soil processes, playing a fundamental role in relevant issues for humanity. However, access to mineralogical analyses is difficult due the difficulty of acquisition through traditional methods and alternative forms to reach it must be explored. This thesis was divided in two chapters that aimed: 1) To understand the fundamental interactions of the energy on pXRF information with emphasis on iron forms, moisture and SOM for use on soil science and 2) To map the abundances of major soil mineralogical components for the whole Brazilian territory at the surface and subsurface. In order to reach the fist objective, three selective dissolution treatments were applied to remove: (i) soil organic matter (−SOM), ii) SOM and poorly crystalline iron forms (−o), iii) SOM and poorly crystalline plus well crystalline iron forms (−d). One additional treatment iv) including water addition (+W) was also carried out. The pXRF was able to detect changes caused by the selective dissolution treatments and soil particle size distribution. The kaolinite, gibbsite, Fe_2O_3, Al_2O_3, SiO_2, TiO_2 and MnO contents were quantified with satisfactory accuracy ($0.61 < R^2 < 0.97$). Sources of uncertainty, mainly soil moisture, must be considered. The understanding of the fundamentals of energy interaction with the sample matrix in the X-ray range is the starting point for characterizing the soil through pXRF. In order to reach the second objective, The Brazilian Spectral Library (BSSL) with Vis-NIR-SWIR spectral data, was used to assess the relative amounts of hematite (Hem), goethite (Gt), kaolinite (Kt) and gibbsite (Gbs) in soil samples from Brazil. Terrain attributes (TA) and a synthetic soil image (SySI) with bare soil pixel from multitemporal Landsat images (1984 to 2020) were used as predictors. A novel approach was performed in order to obtain a bare soil image for the whole Brazilian territory. The model Random Forest (RF) was used for spatial prediction to obtain the mineral maps and their uncertainty by bootstrapping procedure. The Hem presented the more accurate results in RF models with R^2 ranging from 0.48 to 0.56, followed by Gbs (0.42 to 0.44), Kt (0.20 to 0.31) and Gt (0.16 to 0.26). The proposed approach was able to reveal the spatial distribution of the relative abundance of minerals for the Brazilian territory. The mineral maps were in accordance with geology and soil legacy maps and also with the climate and terrain conditions. The approach proposed is an efficient method to obtain mineralogy information for large areas.

Keywords: Digital soil mapping, Pedometric, Soil spectroscopy, pXRF

1. GENERAL INTRODUCTION

The physical, chemical and biological soil processes are essential for the humanity due their relationship to food, fiber and wood supply, climate, flood and water regulation, cultural issues and supporting for nutrient cycling and soil formation (Adhikari et al., 2014; Jónsson and Davíðsdóttir, 2016). Soil is also the largest terrestrial reservoir of C (Stockmann et al., 2013). In this way, the optimization of soil management practices contributes to food security and sustainability (Shah and Wu, 2019) and the knowledge about soil attributes is essential to determine their potentials, limitations and make management decisions (Bouma, 2020).

The mineralogy is the gear of all soil processes. Clay minerals and iron and aluminum oxide minerals are related to cation exchange capacity, aggregation and structure, organic matter dynamic, water retention, adsorption of phosphorus, among others (Gérard, 2016; Gilkes and Prakongkep, 2016; Santos et al., 2017; Zhao et al., 2017). However, the traditional methods to mineralogical characterization, such as X-ray diffraction (XRD) are laborious, expensive, time consuming, and only provide qualitative or semi-quantitative information (Chipera and Bish, 2001; De Souza Bahia et al., 2015; Viscarra Rossel, 2011). Because of it, the mineralogical analysis always was restricted for scientific research level and not commercially available, being few used for the comuns users, such as farmers.

In this context, alternative methods for soil mineralogy characterization come necessary and the proximal and remote sensing techniques can be employed as an efficient tool to reach it (De Souza Bahia et al., 2015; Silva et al., 2021; Viscarra Rossel, 2011). The portable X-ray fluorescence (pXRF) spectrometry sensors access the total elemental content of samples, making possible a simultaneous multi-elemental characterization, with good accuracy and capacity to detect substantial concentration, being easy to use and able to perform analyses directly in the field (Silva et al., 2021). While, the diffuse reflectance spectroscopy, in the Vis-NIR-SWIR range, has specific wavelengths related to the electronic and vibrational transition caused by clay minerals and iron and aluminum oxide minerals (Fang et al., 2018). The spectroscopy in this range is a rapid and environmental friendly method for soil attributes estimation (Soriano-Disla et al., 2014). These specific wavelengths can be used to access the abundances of each mineral by equations or models and it can be spatialized by digital soil mapping procedures plus remote sensed data (Mendes et al., 2021; Poppiel et al., 2020; Silva et al., 2020; Viscarra Rossel, 2011; Viscarra Rossel et al., 2010).

The evolution of the understanding of pXRF is fundamental to continue the study of soil mineralogy due the close relation between the total elemental contents of soil sample and its mineralogy. On the other hand, there is few spatial information about the soil mineralogy and the information obteined from Vis-NIR-SWIR data can be spatialized by digital soil mapping procedures in order to reach it (Viscarra-Rossel et al., 2011). Mineralogy maps for Brazilian territory are almost non-existent.

This thesis was divided in two scientific papers (chapters) that aimed:

1) To understand the fundamental interactions of the energy on pXRF information with emphasis on iron forms, moisture and SOM for use on soil science;

2) To map the abundances of major soil mineralogical components for the whole Brazilian territory at the surface and subsurface;

References

Adhikari, K., Hartemink, A.E., Minasny, B., Bou Kheir, R., Greve, M.B., Greve, M.H., 2014. Digital Mapping of Soil Organic Carbon Contents and Stocks in Denmark. PLoS One 9, e105519. https://doi.org/10.1371/journal.pone.0105519

Bouma, J., 2020. Contributing pedological expertise towards achieving the United Nations Sustainable Development Goals. Geoderma 375, 114508. https://doi.org/10.1016/j.geoderma.2020.114508

Chipera, S.J., Bish, D.L., 2001. Baseline Studies of the Clay Minerals Society Source Clays Powder X-ray Diffraction Analyses. Clays Clay Miner. 49, 398–409. https://doi.org/10.1346/CCMN.2001.0490507

De Souza Bahia, A.S.R., Marques, J., Siqueira, D.S., 2015. Procedures using diffuse reflectance spectroscopy for estimating hematite and goethite in Oxisols of São Paulo, Brazil. Geoderma Reg. https://doi.org/10.1016/j.geodrs.2015.04.006

Fang, Q., Hong, H., Zhao, L., Kukolich, S., Yin, K., Wang, C., 2018. Visible and Near-Infrared Reflectance Spectroscopy for Investigating Soil Mineralogy: A Review. J. Spectrosc. 2018, 1–14. https://doi.org/10.1155/2018/3168974

Gérard, F., 2016. Clay minerals, iron/aluminum oxides, and their contribution to phosphate sorption in soils — A myth revisited. Geoderma 262, 213–226. https://doi.org/10.1016/j.geoderma.2015.08.036

Gilkes, R.J., Prakongkep, N., 2016. How the unique properties of soil kaolin affect the fertility of tropical soils. Appl. Clay Sci. 131, 100–106. https://doi.org/10.1016/j.clay.2016.01.007

Jónsson, J.Ö.G., Davíðsdóttir, B., 2016. Classification and valuation of soil ecosystem services. Agric. Syst. 145, 24–38. https://doi.org/10.1016/j.agsy.2016.02.010

Mendes, W. de S., Demattê, J.A.M., Bonfatti, B.R., Resende, M.E.B., Campos, L.R., Costa, A.C.S. da, 2021. A novel framework to estimate soil mineralogy using soil spectroscopy. Appl. Geochemistry 127. https://doi.org/10.1016/j.apgeochem.2021.104909

Poppiel, R.R., Lacerda, M.P.C., Rizzo, R., Safanelli, J.L., Bonfatti, B.R., Silvero, N.E.Q., Demattê, J.A.M., 2020. Soil color and mineralogy mapping using proximal and remote sensing in Midwest Brazil. Remote Sens. 12, 1–30. https://doi.org/10.3390/rs12071197

Santos, P.G. dos, Almeida, J.A. de, Sequinatto, L., 2017. Mineralogy of the Clay Fraction and Chemical Properties of Soils Developed from Sedimentary Lithologies of Pirambóia, Sanga-the-Cabral and Guará Geological Formations in Southern Brazil. Rev. Bras. Ciência do Solo 41. https://doi.org/10.1590/18069657rbcs20160344

Shah, F., Wu, W., 2019. Soil and Crop Management Strategies to Ensure Higher Crop Productivity within Sustainable Environments. Sustainability 11, 1485. https://doi.org/10.3390/su11051485

Silva, L.S., Marques Júnior, J., Barrón, V., Gomes, R.P., Teixeira, D.D.B., Siqueira, D.S., Vasconcelos, V., 2020. Spatial variability of iron oxides in soils from Brazilian sandstone and basalt. Catena. https://doi.org/10.1016/j.catena.2019.104258

Silva, S.H.G., Ribeiro, B.T., Guerra, M.B.B., de Carvalho, H.W.P., Lopes, G., Carvalho, G.S., Guilherme, L.R.G., Resende, M., Mancini, M., Curi, N., Rafael, R.B.A., Cardelli, V., Cocco, S., Corti, G., Chakraborty, S., Li, B., Weindorf, D.C., 2021. pXRF in tropical soils: Methodology, applications, achievements and challenges. https://doi.org/10.1016/bs.agron.2020.12.001

Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., MacDonald, L.M., McLaughlin, M.J., 2014. The performance of visible, near and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49, 139–186. https://doi.org/10.1080/05704928.2013.811081

Stockmann, U., Adams, M.A., Crawford, J.W., Field, D.J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A.B., Courcelles, V. de R. de, Singh, K., Wheeler, I., Abbott, L., Angers, D.A., Baldock, J., Bird, M., Brookes, P.C., Chenu, C., Jastrow, J.D., Lal, R., Lehmann, J., O'Donnell, A.G., Parton, W.J., Whitehead, D., Zimmermann, M., 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric. Ecosyst. Environ. 164, 80–99. https://doi.org/10.1016/j.agee.2012.10.001

Viscarra Rossel, R.A., 2011. Fine-resolution multiscale mapping of clay minerals in Australian soils measured with near infrared spectra. J. Geophys. Res. 116, F04023. https://doi.org/10.1029/2011JF001977

Viscarra Rossel, R.A., Bui, E.N., De Caritat, P., McKenzie, N.J., 2010. Mapping iron oxides and the color of Australian soil using visible-near-infrared reflectance spectra. J. Geophys. Res. Earth Surf. 115, 1–13. https://doi.org/10.1029/2009JF001645

Zhao, J., Chen, S., Hu, R., Li, Y., 2017. Aggregate stability and size distribution of red soils under different land uses integrally regulated by soil organic matter, and iron and aluminum oxides. Soil Tillage Res. 167, 73–79. https://doi.org/10.1016/j.still.2016.11.007

2. THE FUNDAMENTAL OF THE EFFECTS OF WATER, ORGANIC MATTER, AND IRON FORMS ON THE pXRF INFORMATION IN SOIL ANALYSES

ABSTRACT

Portable X-ray fluorescence (pXRF) has great potential for numerous applications in soil science. However, the basic knowledge about the effects of soil properties on pXRF spectra are still poorly studied, which may lead users to biased interpretations of mathematical models. The present study aimed to evaluate the outcomes of moisture, soil organic matter content (SOM), and iron forms on pXRF data. The work was conducted with seventeen soil samples from the central region of São Paulo state (Brazil). Three selective dissolution treatments were applied to remove: (i) soil organic matter (−SOM), ii) SOM and poorly crystalline iron forms (−o), iii) SOM and poorly crystalline plus well crystalline iron forms (−d). One additional treatment iv) including water addition (+W) was also carried out. The effects of treatments were evaluated for sandy and clayey samples. Soil particle size distribution and elemental content affected the bremsstrahlung and characteristic peaks counts. In +W, there was a generalized decrease in counts mainly for the light elements (magnesium, aluminum and silicon). Regarding the selective dissolution procedures, alterations were verified, reflecting mainly the removal power of reagents. Generally, the most pronounced alterations occurred for −d and moderate alterations for −SOM and −o. The pXRF data showed high correlation with particle size distribution and mineralogy attributes. The kaolinite, gibbsite, Fe_2O_3, Al_2O_3, SiO_2, TiO_2 and MnO contents were quantified with satisfactory accuracy ($0.61 < R^2 < 0.97$). The pXRF was able to detect changes caused by the selective dissolution treatments and soil particle size distribution. Sources of uncertainty, mainly soil moisture, must be considered. The understanding of the fundamentals of energy interaction with the sample matrix in the X-ray range is the starting point for characterizing the soil through pXRF.

Keywords: Proximal sensing; Soil mineralogy; Chemometrics; Tropical soils

Published as: Rosin, N.A., Demattê, J.A.M., Leite, M.C.A., Carvalho, H.W.P, Costa, A.C., Greschuk, L.T., Curi, N., Silva, S.H.G. The fundamental of the effects of water, organic matter, and iron forms on the pXRF information in soil analyses. Catena 105868. https://doi.org/10.1016/j.catena.2021.105868

2.1. Introduction

The optimization of soil management practices contributes to high crop productivity and, consequently, to food security and sustainability (Shah and Wu, 2019). Knowledge about soil attributes is essential to determine their potentials, limitations and make management decisions (Bouma, 2020). However, soil characterization is predominantly carried out via wet chemistry and particle size distribution analyses, with mineralogical analyses being less used by users, even in recent studies using Proximal Sensing (PS) (Afriyie et al., 2020, Allo et al., 2020, Andrade et al., 2020a, Andrade et al., 2020b, Barthès et al., 2020, Benedet et al., 2020, Biney et al., 2020, Costa et al., 2020, Gomez et al., 2020, Ma et al., 2021, Mammadov et al., 2020, Xu et al., 2020). Mineralogical analyses are essential for understanding the dynamics of water and nutrients in soils (da Costa et al., 2020, Souza et al., 2017), especially in tropical environments, where Oxisols cover approximately 40% of the Brazilian territory (Anjos et al., 2012) and almost 50% of Sub–Saharan Africa territory (Dewitte et al., 2013). For instance, effects of the contents of iron and aluminum oxide minerals on soil aggregation and consequently on their permeability were demonstrated by Pinheiro-Dick and Schwertmann, 1996, Lima and Anderson, 1997, Ferreira et al., 1999, Muggler et al., 2007.

In this respect, the focus of the laboratories for users at international level is in soil fertility attributes. This is due to several issues, both historical and, mainly, the difficulties in analyzing soil mineralogy (Kunze and Dixon, 1986). Initially, the mineralogy was analyzed via X-ray diffractometers that assess soil crystalline minerals (Whittig, 1965). However, such technique, which is the basis for mineralogical studies, has limitations such as sample preparation and equipment that is difficult to access and use (Chipera and Bish, 2001, McManus, 1991). This approach has led such analyses to be restricted to the research level, and whose users focus on interpretation of data analyses, in detriment to the use of information by common users (i.e., farmers). Hence, the insertion of soil mineralogy as an essential piece of information for the agronomic community is not as ubiquitous as it could be. With the advent of numerous technologies, such as remote sensing, PS and computational processing and interpretation, this problem tends to decrease (Demattê et al., 2004). In this context, it has been shown that spectral bands at the mid-infrared (Mid-IR) and visible, near infrared, and shortwave infrared (Vis-NIR-SWIR) can be employed to predict a large number of soil attributes (Soriano-Disla et al., 2014).

The Vis-NIR-SWIR and Mid-IR ranges are mainly related to molecular vibrations (Soriano-Disla et al., 2014). However, due to the low energy intensity, they do not reach the atomic level (Soriano-Disla et al., 2014), impacting on the non-detection of the total content of each element. Therefore, within the wide range of applications of PS, special attention should be paid to the bands with the highest energy which have greater strength to cause atomic changes and consequently be detected (Silva et al., 2021, Weindorf et al., 2014). The X-ray fluorescence (XRF) spectrometry relies on electronic transitions at the atomic level, which brings elemental specificity (Silva et al., 2021, Weindorf et al., 2014). The elemental identification and quantification are possible because the rate of fluorescence emission is directly proportional to the concentration of each element (Weindorf et al., 2014). The portable X-ray fluorescence (pXRF) spectrometry sensors make possible a simultaneous multi-elemental characterization, with good accuracy and capacity to detect substantial concentration, being easy to use and able to perform analyses directly in the field (Silva et al., 2021, Zhu et al., 2011).

For soils from tropical environments, there are accurate results for prediction of the pseudo-total content (determined by incomplete digestion procedures) of iron (R^2 of the best regression/model = 0.98), chromium (0.91), manganese (0.72), calcium (0.84), copper (0.81), nickel (0.86), titanium (0.99), vanadium (0.72), phosphorus (0.97), among others (Silva et al., 2019, Silva et al., 2020). Santana et al. (2018) predicted the soil pseudo total content of iron with high accuracy in laboratory (R^2 = 0.96) and field (R^2 = 0.91) conditions. The pXRF can also be used, by inference, to predict clay (R^2 of the best model = 0.91), sand (0.85), soil organic matter (SOM) (0.72), total nitrogen (0.60), cation exchange capacity (CEC) (0.75), and exchangeable/available content of calcium (0.91), cupper (0.61), magnesium (0.85), manganese (0.86) and potassium (0.90) (Andrade et al., 2020a, Andrade et al., 2020b, Tavares et al., 2020, de Lima et al., 2019).

However, certain soil properties and analytical conditions interfere in the interaction between pXRF energy and the sample matrix (Ge et al., 2005, Ribeiro et al., 2017, Sahraoui and Hachicha, 2017). Understanding the relationships between the main properties of the soil, such as mineralogy, SOM and water, can bring light to the interpretations of the results obtained by the pXRF. In fact, most soil studies focus on the quantification of elements and not on the fundamentals. Observing the literature, there are no studies showing the effect of different forms of iron on the pXRF spectra. How to really understand mineralogy without this basic knowledge? Scientific foundations on soil science becomes necessary in view of the large production of soil data

without new knowledge production (Wadoux et al., 2021). Therefore, such investigations are needed especially in tropical soils, since their unique structure is composed of microaggregates, which in turn is influenced manly by the clay mineralogy (especially gibbsite), iron forms and SOM. In such soils, structure is more important than texture in terms of physico-hydrical behavior (Silva et al., 2021).

The evolution of the understanding of pXRF for soil analyses is fundamental to continue the study of soil mineralogy. The hypothesis is that pXRF has sufficient energy to identify iron forms, thus assisting in soil mineralogy understanding. Also, it is expected that moisture causes interference and must be considered on the data evaluation. Thus, the objective of this study was to understand the fundamental interactions of the energy on pXRF information with emphasis on iron forms, moisture and SOM for use on soil science.

2.2 Material and methods

2.2.1 Study area and sample collection

This study was conducted with seventeen soil samples from the central region of São Paulo state (Brazil), between the coordinates 22°45′ − 22°70′ S and 48°60′ −49°47′ W, covering the municipalities of Bauru, Lençóis Paulista and Macatuba, emphasizing that such soil samples represent the major soils of the region (1.702 Km^2) (Instituto Brasileiro de Geografia e Estatística, 2020). The soils of the study region are geologically derived from basic eruptive rocks of the Serra Geral (Early Cretaceous) formation, which overlap the Botucatu (Jurassic-Cretaceous) formation. The latter consists of fine to coarse sandstones. In this region, geological material from the Bauru group is also observed, which is superimposed on the rocks of the Serra Geral (Late Cretaceous) formation (Demattê et al., 2007, Fernandes et al., 2010). The samples were collected at 0–20, 40–60 and 80–100 cm depths, from 13 sampled representative soil profiles, according to Empresa Brasileira de Pesquisa Agropecuária (1995). The samples were from Rhodic Hapludox, Typic Hapludox, Typic Quartzipsamments and Typic Paleudalf soils (Table 1) (Soil Survey Staff, 2014). The flowchart of the methodology can be found in Fig. 1 and more detailed information about the study sites were provided by Demattê et al. (2007) and Silvero et al. (2020).

Table 1: Description of soil samples.

Soil profile	Sample	Soil Type	Depth (cm)	Texture
P1	RH1 (B)	Rhodic Hapludox (RH)	40-60	Clayey
P2	RH2 (A)		0-20	Clayey
	RH2 (B)		40-60	Clayey
P3	TE1 (B)	Typic Eutrustox (TE)	40-60	Clayey
P4	TE2 (B)		40-60	Clayey
P5	TH1 (A)	Typic Hapludox (TH)	0-20	Clayey
	TH1 (B)		40-60	Clayey
P6	TH2 (A)		0-20	Sandy
	TH2 (B)		40-60	Clayey
P7	TH3 (B)		40-60	Sandy
P8	TH4 (B)		40-60	Clayey
	TH5 (A)		0-20	Sandy
P9	TH5 (B)		40-60	Clayey
P10	TP1 (C)	Typic Paleudalf (TP)	80-100	Sandy
P11	TP2 (B)		40-60	Sandy
P12	TQ1 (B)	Typic Quartzpament (TQ)	40-60	Sandy
P13	TQ2 (B)		40-60	Sandy

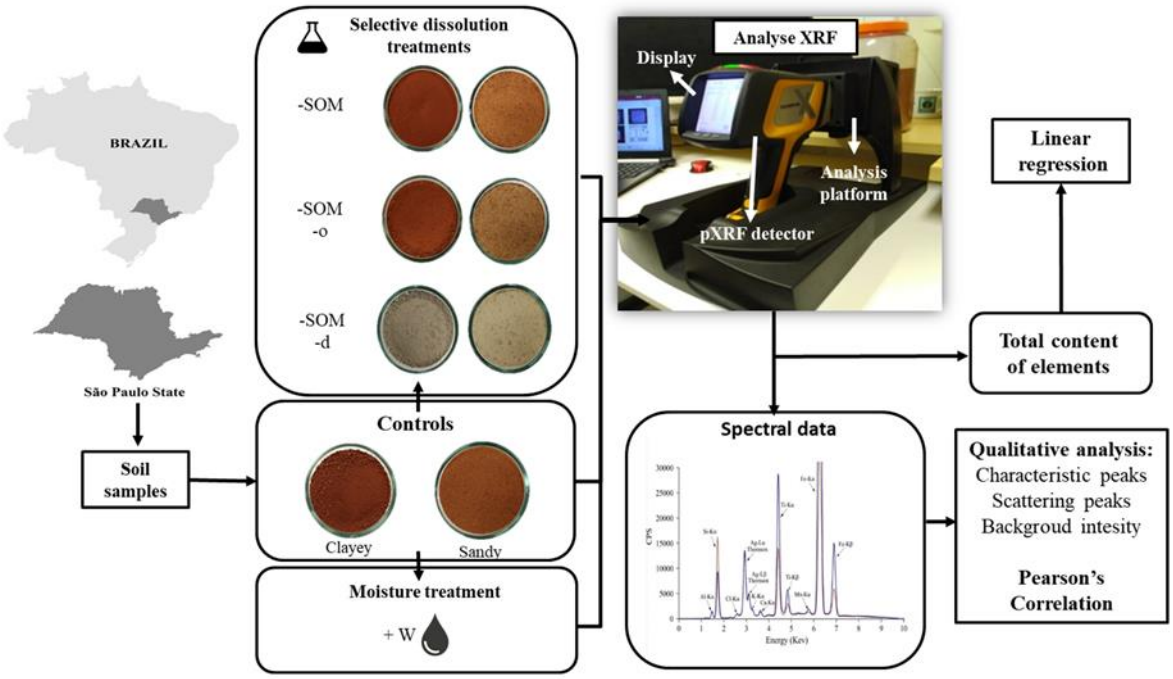

Fig. 1. Flowchart of methodology. -SOM = soil organic matter removal, -o = SOM and poorly crystalline iron removal, -d = SOM and poorly crystalline plus free crystalline iron removal; +W = water addition.

2.2.2 Laboratorial analyses

For laboratory analyses, the soil samples were air-dried, ground, and passed through a 2 mm sieve. The soil chemical analyses performed included: (a) clay, sand and silt contents (particle size distribution), using the pipette method (NaOH - dispersant solution) (Teixeira et al. 2017); (b) soil organic matter content (SOM) determined by the Walkley-Black oxidation method (Walkley and Black, 1934); (c) pH in water, soil acidity (exchangeable/available aluminum and hydrogen) and exchangeable/available calcium, magnesium, potassium and phosphorus contents were determined in addition to the sum of bases (SB), cations exchange capacity (CEC), saturation of bases (V%) and saturation of aluminum (m%) calculated according to van Raij et al. (2001); (d) pseudo total iron (Fe_2O_3), aluminum (Al_2O_3) and silicon oxides (SiO_2) contents were determined by sulfuric acid digestion (Empresa Brasileira de Pesquisa Agropecuária, 1997); (e) Ki (SiO_2/Al_2O_3) and Kr [$SiO_2/(Al_2O_3 + Fe_2O_3)$] were calculated using Fe_2O_3, Al_2O_3 and SiO_2 contents divided by their molar weight; (f) poorly crystalline iron (Feo) content was determined after selective dissolution with ammonium oxalate solution (10 mL at 0.2 M) (McKeague and Day, 1966); (g) poorly crystalline plus well crystalline iron (Fed) were obtained after selective dissolution with sodium citrate-dithionite-bicarbonate solution (Holmgren, 1967); (h) kaolinite (Kt) and gibbsite (Gb) contents were determined by differential thermal analysis (Dixon et al., 1989, McKeague and Day, 1966); (i) Silicon and aluminum poorly crystalline materials were determined after selective dissolution with boiled 0.5 N KOH (Si + Al); (j) other clay minerals (V + M) were calculated by the equation: other clay minerals (%) = 100 – (Kt + Gb + iron oxides).

2.2.3 Treatments

The soil samples without treatment were called control. Three selective dissolution treatments and one water addition treatment were performed in order to verify spectral changes (Fig. 1). The selective removal procedures were carried out according to Jackson (1969) and more detailed information can be found in Demattê and Garcia (1999): (i) SOM removal treatment (−SOM): first, 40 mL of H_2O_2 (30%) were mixed with soil samples (20 g), promoting the cold reaction. After, the hot reaction was carried out by adding sodium acetate, followed by shaking and centrifugation at 1400 rpm. The samples were then washed with ethanol. (ii) SOM and poorly crystalline iron removal treatment (−o): after SOM removal, the poorly crystalline iron was removed by acid ammonium oxalate (0.2 M) in the dark. The samples were then washed with ethanol. (iii) SOM and poorly crystalline plus well crystalline iron removal treatment (−d): after SOM removal, 40 mL of citrate solution (0.3 M) and sodium bicarbonate (1 M) were added, followed by heating and 1 g of sodium dithionite was added. After mixing, 10 mL of acetone, and 10 mL of sodium chloride saturated solution were added. The centrifugation was carried out at 1400 rpm and the supernatant material was discarded. Detailed information about the removal effect of each treatment can be found in Table A.1. Besides, before pXRF measurements, a treatment with water addition was performed: (iv) Water addition treatment (+W): 0.5 mL of pure water was added in a small amount to soil samples (5 g). This amount of water/soil was adequate to evaluate the influence of water in Mid-IR in a previous study (Silvero et al., 2020).

2.2.4. pXRF analyses

The analyses were carried out in a portable X-ray fluorescence (pXRF) spectrometer Olympus Delta Professional (Olympus Corporation, Waltham, MA, USA), with two excitation modes (EM). The first EM employed 40 keV, 91.1 μA, and is equipped with a 2 mm aluminum filter, which is most suitable to quantify the following elements: vanadium, chromium, iron, cobalt, nickel, copper, zinc, tungsten, mercury, arsenic, lead, bismuth, rubidium, uranium, strontium, zirconium, yttrium, aurum, thorium, niobium and, molybdenum and secondarily: titanium and manganese. The second EM employed 10 keV and 80.5 μA, which improves the signal of light elements (mainly magnesium, aluminum and silicon) and quantifies the following elements: magnesium, aluminum, silicon, phosphorus, sulfur, chlorine, calcium, titanium and manganese. About 5–15 g of the samples were placed in a 20 μm polyethylene bag and submitted to analysis in a platform with protection for X-rays emission. The pXRF Delta Professional is furnished with a 50 keV silver X-ray anode and a silicon drift detector, with 2048 channels. The precision for quantified elements can be seen in Table A.2.

2.2.5 Spectral processing and quantitative analyses

For spectral descriptive analyses, the samples were stratified into two classes: sandy (≤250 g kg^{-1} of clay) and clayey (>250 g kg^{-1} of clay) (Demattê and Demattê, 2009) (Table 1). The EM1 in the range from 4 to 40 keV and the EM2 from 0 to 8 keV average spectra for each textural class were assessed. The counts rate of (i) characteristic peaks of each element (ii) scattering peaks (Compton and Thomsom) and (iii) background intensity were evaluated. The objective of qualitative analyses was to evaluate visually the spectral effects caused by particle size distribution and selective dissolution treatments. The total content of copper, zinc, magnesium, aluminum, silicon, titanium, iron and manganese were determined using the Geochem mode. The Geochem mode determines the elemental contents based on the count rate of characteristic peaks of each element, and has

a Compton normalization method to correct the matrix effects and operates in the two EM mentioned above (Stockmann et al., 2016). The average contents for each texture class and treatments were presented together with the raw spectral information.

In order to identify the relationships between the pXRF channels and soil attributes, Pearson's correlation analysis was carried out between spectral data and soil attributes of control samples. Other papers, such as Silvero et al. (2020), also used the Pearson Correlation to evaluate the relation between soil attributes and spectral data. With the objective of eliminating the blank zones (regions in which the pXRF detector did not register signals), the region from 0 to 7.25 keV from EM2 and the region from 7.25 to 24 keV for EM1 were selected for Pearson's correlation. Linear Regression (LR) analysis was carried out between the elemental contents from pXRF Geochem mode (iron, silicon, aluminum, titanium and manganese) and the soil mineralogical attributes. Also, a linear equation was generated for Kt, Gb, SiO_2, Al_2O_3, TiO_2, MnO and Fe_2O_3 contents. The parameter of accuracy evaluation employed were the coefficient of determination (R^2), with significance value (p) of 0.005, and the root mean square error (RMSE). Statistical analyses were performed using the R software, version 4.0 (R Development Core Team, 2020).

2.3 Results

2.3.1 Descriptive analyses of laboratory data

The soil particle size distribution and chemical attributes had large variation between samples (Fig. 2). The soil particle size contents varied from 30 to 860 g kg^{-1} and from 80 to 870 g kg^{-1}, for sand and clay contents, respectively (Fig. 2a). The SOM content varied from 9 to 21 g kg^{-1} (Fig. 2b). Expressive variations were also verified for the chemical attributes (pH, calcium, magnesium, potassium, phosphorus, aluminum, hydrogen, SB, CEC, V% and m%) (Fig. 2c). Briefly, the CEC showed values ranging from 2.40 to 34.30 cmolc kg^{-1}.

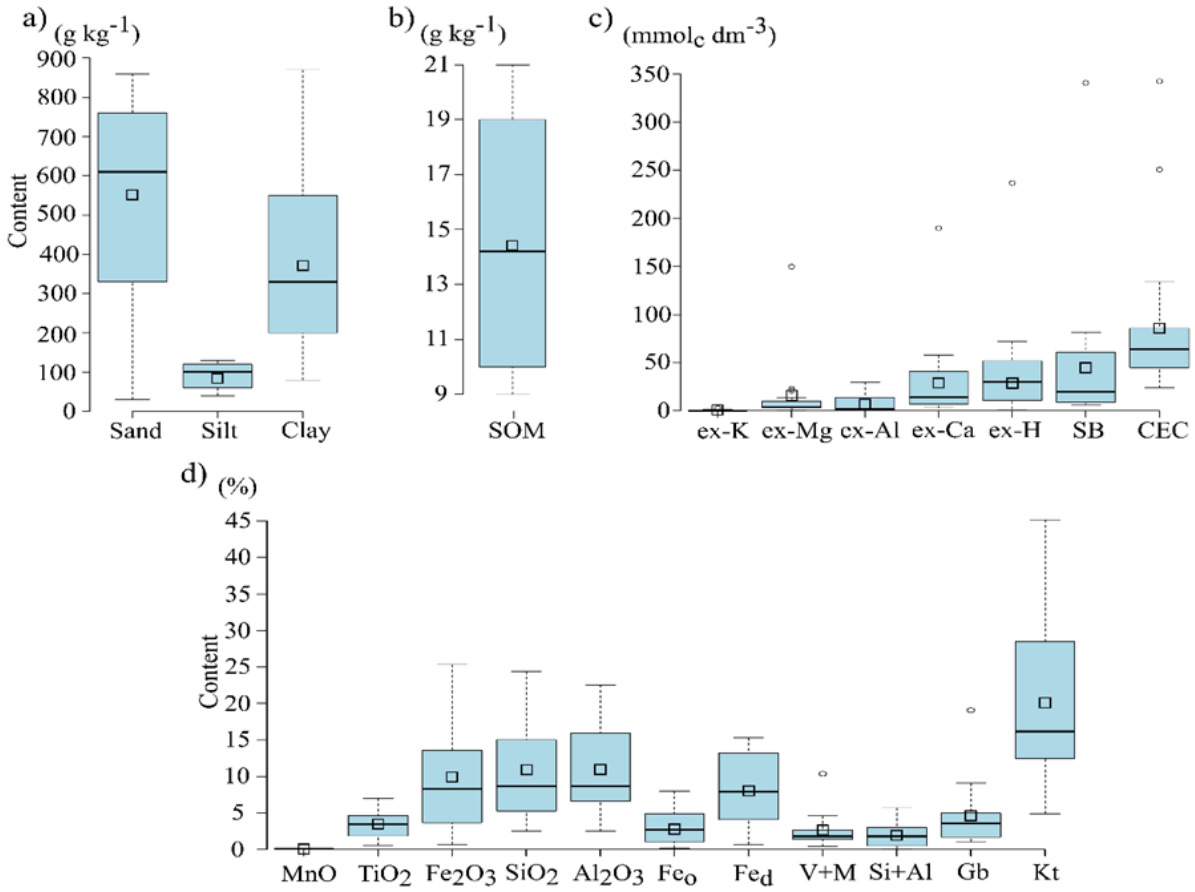

Fig. 2. Boxplots of soil particle size distribution (a), soil organic matter (SOM) (b), chemical (c) and mineralogical (d) attributes. SOM = soil organic matter; ex-: exchangeable/available form; SB = sum of bases; CEC = cations exchange capacity at pH 7; Feo = poorly crystalline iron forms; Fed = poorly plus well crystalline iron forms; Kt = kaolinite; Gb = gibbsite; Si + Al = silicon and aluminum poorly crystalline material; V + M = other clay minerals.

Table 2. Descriptive statistics of soil particle size distribution and chemical attributes for sandy and clayey samples.

	Sand	Silt	Clay	pH H_2O	SOM	ex-P	ex-K	ex-Ca	ex-Mg	ex-Al	ex-H	SB	CEC	V%	m%
	Sandy samples														
	.....g kg^{-1}.....				g kg^{-1}	Ug kg^{-1}	..... $cmol_c$ kg^{-1}							.....%.....	
Mean	769	61	170	5.5	12.6	6.6	0.09	3.56	2.36	0.77	2.29	6.10	8.30	49	29
Min	660	40	80	4.4	9.0	3.0	0.02	0.50	0.10	0.00	0.20	0.60	2.40	11	0
Max	860	100	240	7.2	21.0	21.0	0.31	19.00	15.00	3.00	6.20	34.10	34.30	99	77
SD	78	23	66	0.9	4.2	6.5	0.10	6.82	5.58	1.14	2.10	12.40	11.58	34	35
	Clayey samples														
	.....g kg^{-1}.....				g kg^{-1}	Ug kg^{-1}	..... $cmol_c$ kg^{-1}							.....%.....	
Mean	392	104	504	5.3	15.9	5.6	0.05	2.71	0.91	0.73	5.64	3.66	9.30	45	25
Min	30	40	270	4.5	10.0	1.0	0.01	0.30	0.20	0.00	0.20	0.60	4.50	6	0
Max	620	130	870	6.6	21.0	23.0	0.15	5.80	2.30	2.80	23.70	8.20	25.10	97	75
SD	204	31	192	0.7	4.0	6.8	0.05	2.13	0.77	0.99	6.67	2.87	6.13	33	32

Where: Min = minimum value; Max = maximum value; SD= standard deviation; SOM = soil organic matter; ex-: exchangeable/available forms SB = sum of bases; CEC = cations exchange capacity; V% = bases saturation; m% = aluminum saturation.

Table 3. Descriptive statistics of mineralogical analyses and weathering indexes for sandy and clayey samples.

	S/C	Ki	Kr	Kt	Gb	V+M	Si+Al	SiO_2	Al_2O_3	TiO_2	MnO	Fe_2O_3	Fe_d	Fe_o
	Sandy samples													
	..... %													
Mean	0.41	1.68	1.32	10.0	2.5	1.4	1.4	5.7	5.9	2.1	0.03	3.4	3.1	1.2
Min	0.20	0.83	0.63	4.9	1.2	4.0	1.0	2.5	2.5	6.0	0.01	7.0	7.0	2.0
Max	0.75	2.36	1.79	13.4	5.0	2.4	3.7	10.0	7.4	5.6	0.05	8.6	5.3	3.6
	SD	0.51	0.41	3.4	1.6	7.0	1.5	2.5	18	17	0.01	2.7	1.9	1.1
	Clayey samples													
	..... %													
Mean	0.22	1.68	1.02	27.3	6.1	3.3	2.2	14.4	14.2	4.5	0.09	14.3	11.7	4.4
Min	0.11	1.32	0.68	14.0	1.1	1.4	3.0	4.0	4.6	3.1	0.05	8.1	7.5	1.7
Max	0.44	2.11	1.30	45.2	19.1	10.4	5.7	24.4	22.6	7.0	0.13	25.4	15.3	8.0
SD	0.10	0.27	0.19	9.6	5.3	2.7	1.6	7.1	6.0	1.3	0.03	6.3	3.1	2.0

Where: Min = minimum value; Max = maximum value; SD = standard deviation; S/C = silt/clay; Ki = Ki index; Kr = Kr index; Fe_o = poorly crystalline iron forms; Fe_d = poorly crystalline and free crystalline iron forms; Kt = kaolinite; Gb: gibbsite; V+M = other clay minerals; Si+Al = silicon and aluminum poorly crystalline material.

In the mineralogical analyses, it was verified great differences among samples for some attributes, mainly for Kt, Fe_2O_3, Fed and SiO_2 contents (Fig. 2d). The Kt, Gb, V+M and Si+Al contents presented values from 4.9 to 45.2, 1.1 to 19.1, 0.4 to 10.4 and 0.1 to 5.7%, respectively. These results are in accordance with the differences observed in the Al_2O_3, SiO_2, TiO_2, MnO and Fe_2O_3 values that ranged from 2.5 to 24.4, 25 to 25.5, 0.6 to 7.0, 0.01 to 0.13 and 0.7 to 25.4%, respectively. The Feo and Fed contents showed values ranging from 0.2 to 8.0 and from 0.7 to 15.3 g kg^{-1}, respectively.

Seven samples were allocated into sandy class and 10 into clayey class (Table 1). The sandy samples were generally from the less weathered soil classes (TQ and TP) and also from the A and C horizons of TH. Conversely, the clayey samples corresponded to more weathered soils classes (TH, RH and TE). The chemical attributes showed some variation with textural class (Table 2). Briefly, there were increases for clayey samples of 3.3 g kg^{-1} and 1.0 cmolc kg^{-1} for SOM and CEC, respectively. The soil particle size distribution had greater variations than chemical attributes. The values for sandy samples varied from 80 to 270 g kg^{-1}, with mean of 170 ± 66 g kg^{-1} and from 660 to 860 g kg^{-1}, with mean of 769 ± 78 g kg^{-1}, for clay and sand contents, respectively. For clayey samples, the soil particle size distribution varied from 270 to 870 g kg^{-1}, with mean of 504 ± 192 g kg^{-1} and from 30 to 620 g kg^{-1}, with mean of 392 ± 204 g kg^{-1}, for sand and clay contents, respectively (Table 2).

There was an average increase of 10.9, 8.3, 5.7, 2.4 and 0.06% for Fe_2O_3, Al_2O_3, SiO_2, TiO_2 and MnO, respectively for clayey samples when compared with sandy samples (Table 3). The other mineralogical attributes also showed higher values for clayey samples, with Kt and Gb increasing by 17.3 and 3.6%, respectively. Samples from the highly weathered soil profiles (TH, TE and RH) presented higher iron oxides, aluminum oxides and clay contents (Table 1, Table 3). Generally, the soil weathering indices presented the greatest values associated to sandy samples (except for Ki).

2.3.2. Treatments and X-ray configuration effects

In EM2, one can observe that both background radiation and characteristic peaks increased compared to EM1 (Fig. 3, Fig. 4, Fig. 5). When we look at the mean spectra of sandy and clayey samples, it was possible to verify that soil particle size distribution differed between them in all peaks of the pXRF patterns (Fig. 3). The 15 to 35 keV region is quite influenced by soil particle size distribution, with an increase of background counts. In this region, there are scattering peaks (Compton and Thomson) (Kα and Kβ), that also have count rate variation. The zirconium-Kα count rate was not influenced by soil particle size distribution. Conversely, copper and zinc contents presented differences by textural classes, with contents above the limit of quantification only for clayey samples with average of 0.015 and 0.008%, respectively.

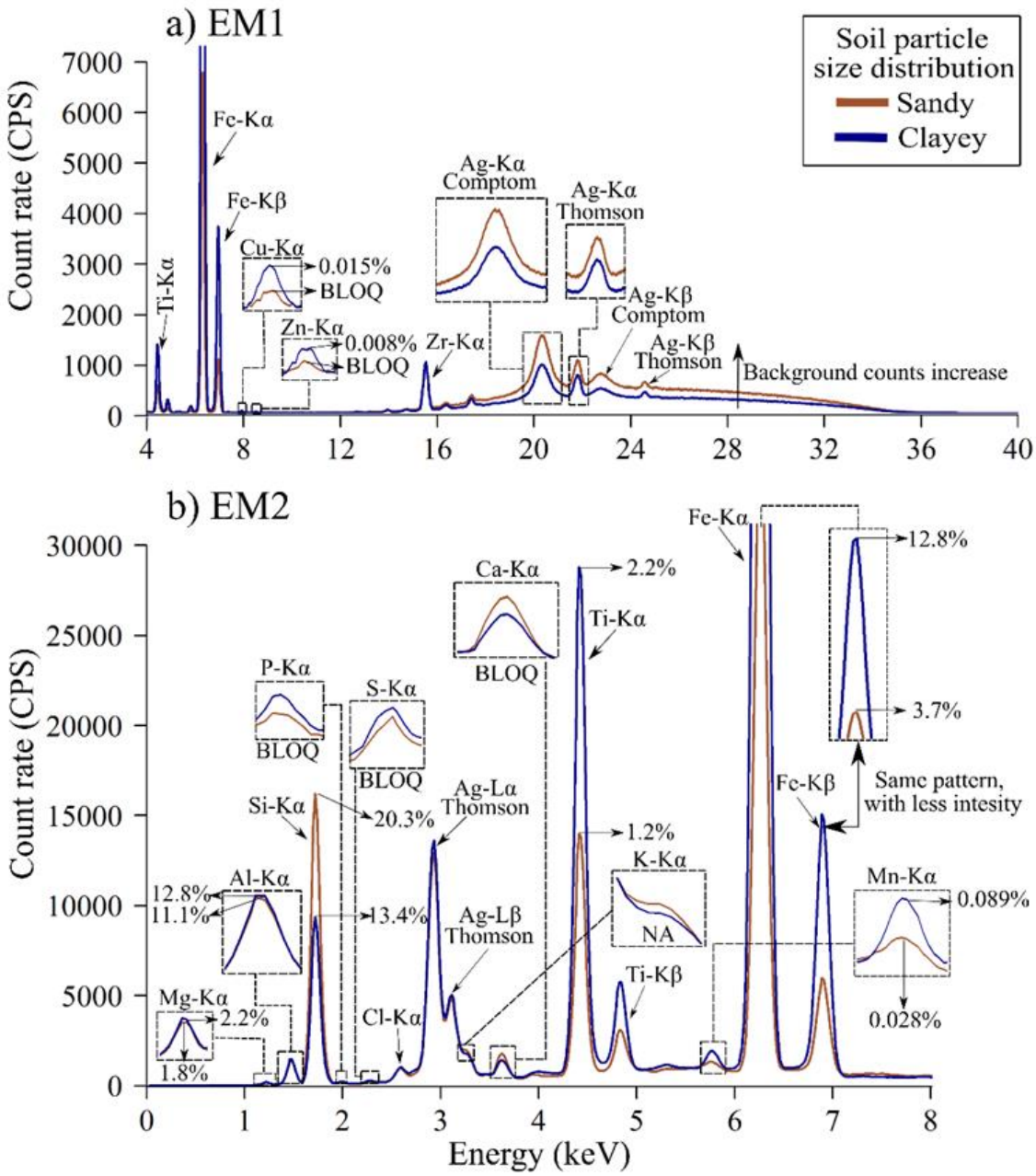

Fig. 3. Mean spectra for sandy and clayey soil samples. EM1 = excitation mode 1; EM2 = excitation mode 2; NA = non-analyzed; BLOQ = below the limit of detection; CPS = counts per second. The percentage corresponds to the concentration of the element calculated by the Geochem mode.

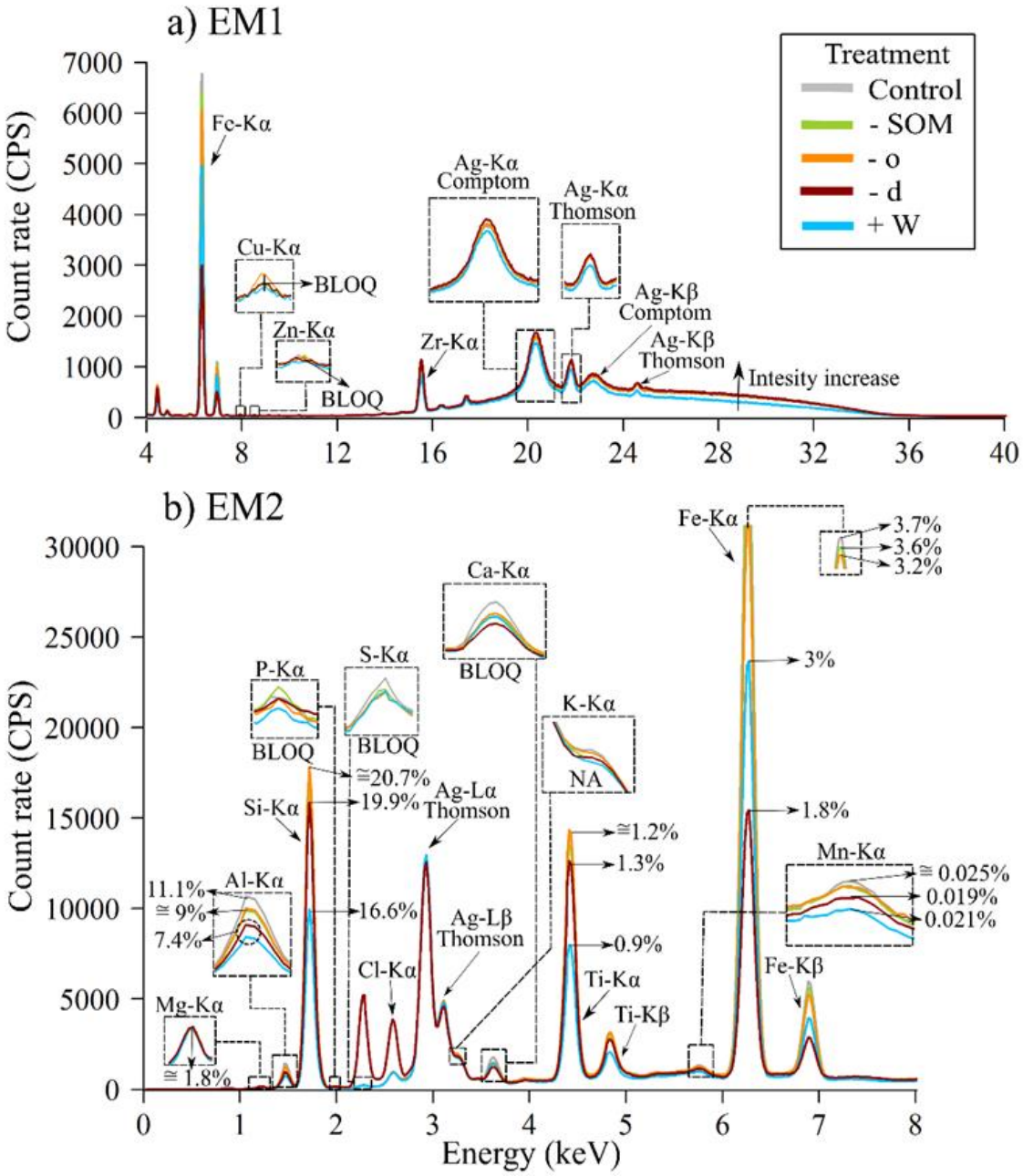

Fig. 4. Mean spectra of treatments for sandy soil samples. EM1 = excitation mode 1; EM2 = excitation mode 2; −SOM = soil organic matter removal; −o = SOM and poorly crystalline iron removal; −d: SOM and poorly crystalline plus well crystalline iron removal; +W: water addition; NA = non-analyzed; BLOQ = below the limit of quantification; CPS = counts per second. The percentage corresponds to the concentration of the element calculated by the Geochem mode.

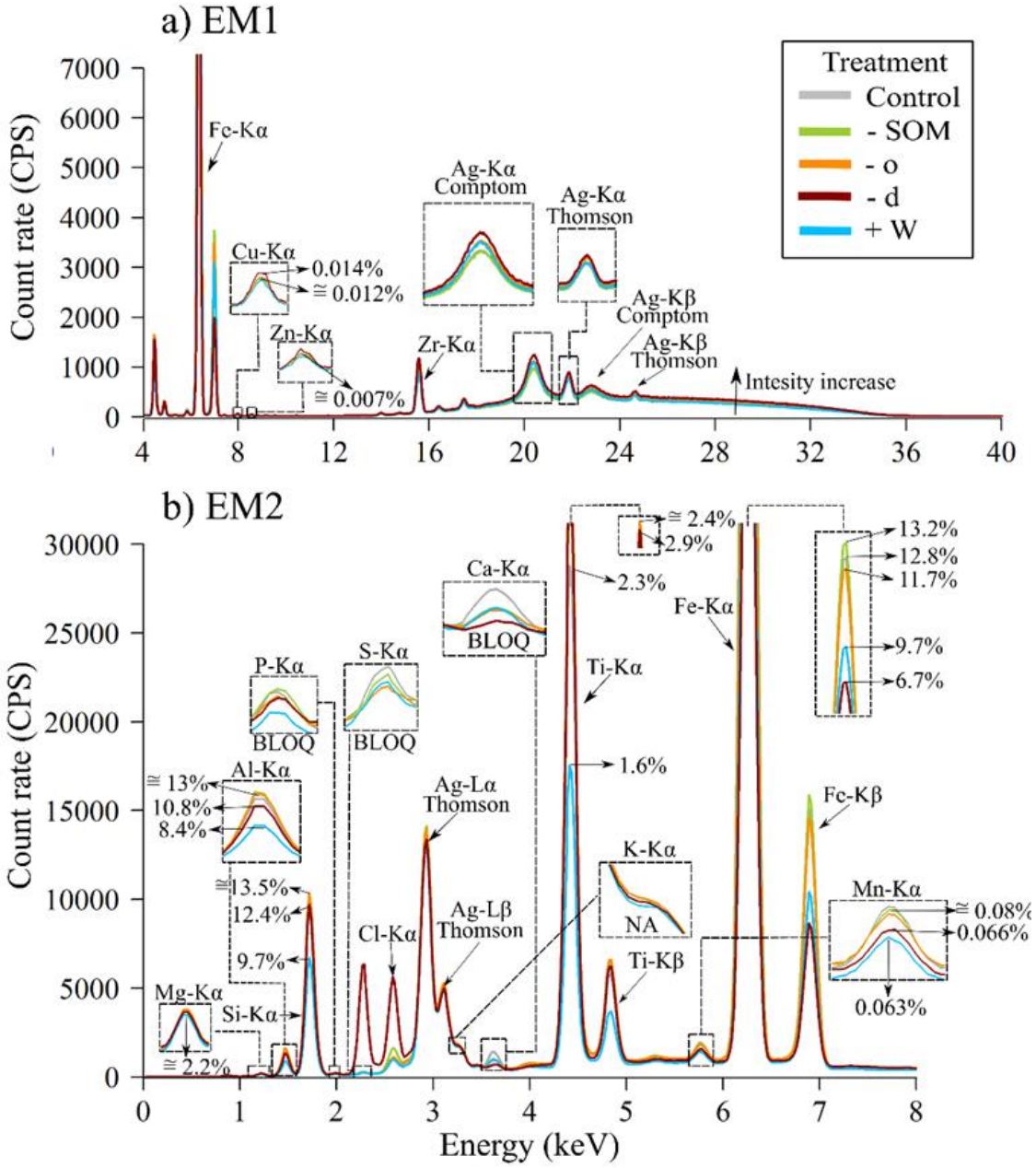

Fig. 5. Mean spectra of treatments for soil clayey samples. EM1 = excitation mode 1; EM2 = excitation mode 2; −SOM = organic matter removal; −o: OM and poorly crystalline iron removal; −d: SOM, poorly crystalline plus well crystalline iron removal; +W: water addition; NA = non-analyzed; BLOQ = below the limit of quantification; CPS = counts per second. The percentage corresponds to the concentration of the element calculated by the Geochem mode.

In EM2, the greatest peaks and elemental concentrations were observed in the mean spectra for clayey samples (Fig. 3). Differences in contents of 9.1, 1.7, 1.0, 0.4 and 0.061% were observed between sandy and clay samples for iron, aluminum, titanium, magnesium and manganese, respectively. Conversely, silicon had greater content in sandy samples, with a difference of 6.9? %. Phosphorus, sulfur, calcium and potassium (−Kα) peaks also showed differences; however, they did not reach the limit of quantification. Calcium and potassium (−Kα) counts were reduced for clayey samples; while phosphorus, and sulfur (−Kα) increased.

The +W treatment had a generalized decrease on the count rate and contents, and distinct alterations for −SOM, −o, and –d treatments were also verified (Fig. 4). The +W treatment caused a major impact on the

pXRF spectra. In EM1 mode, for copper-Kα, only −o treatment varied, when compared with the control, while zinc and zirconium (−Kα) showed stability (Fig. 4a). The background and scattering peaks kept patterns for the treatments, showing greater count rate increase for −d, small increases for −o and −SOM and decrease for +W, when compared to the control.

For EM2 mode (Fig. 4b), large differences of count rates can be observed for aluminum, silicon, calcium, titanium, manganese and iron (−Kα), smaller for phosphorus-Kα and silver Thomson and potassium and magnesium (−Kα) remained stable. Magnesium content was 1.8% for all treatments. Aluminum content was greater for the control (11.1%) and decreased to 9% for −SOM and −o, and to 7.4% for −d and +W treatments, respectively. Silicon content showed stability for −o and −SOM, when compared to the control, with content of 20.7%, while −d and +W reduced it to 19.9% and 16.6%, respectively. Titanium content in the control samples was 1.2%, while the +W treatment reduced it to 0.9% and the −d increased it to 1.3%, while −SOM and −o treatments remained stable. Manganese content remained stable for −SOM and −o treatments, when compared with the control, with values of 0.025%, while +W and −d treatments reduced it to 0.021% and 0.019%, respectively. Iron in the control, was present at 3.7% and decreased for −SOM, −o, +W and −d treatments, with determined values of 3.6%, 3.2%, 3% and 1.8 %, respectively. Titanium and manganese contents in −d treatment increased, despite the count rate decreased when compared with the control and +W treatments, respectively. The calcium-Kα count rate was the most affected by −SOM, and a great increase in sulfur and chlorine (−Kα) peaks was verified for −d treatment.

For EM1 mode, differently from sandy soil samples, one can observe that +W increased the background, Compton, and Thomson peaks (Fig. 5). Copper showed greater content for −d treatment (0.012%), with values of 0.014% greater than other treatments, while zinc contents remained stable (0.007%).

For EM2 mode (Fig. 5b), larger differences of count rates were observed for aluminum, silicon, phosphorus, calcium, titanium, manganese and iron (−Kα), while minor differences were observed for potassium and magnesium (−Kα) and silver Thomson. Magnesium content presented values of 2.2% for all treatments. Aluminum content showed stability for −SOM and −o when compared with the control (13%), decreased to 8.2% for −d and to 8.4% for +W treatments. Silicon content also showed stability for −o and −SOM, when compared to the control (13.5%), while −d and +W treatments reduced it to 12.4% and 9.7%, respectively. Titanium in control samples presented a value of 2.3%, the +W treatment reduced its content to 1.6%, while, the SOM, −o, and −d treatments increased it to 2.3, 2.4, and 2.9%, respectively. Despite the increase in titanium content in −d treatment, the count rate decreased. Manganese content remained stable for −SOM and −o treatments, when compared with the control (0.08%), while −d and +W treatments reduced it to 0.066% and 0.063%, respectively. Compared to control, −SOM increased iron content from 12.8% to 13.2%, while −o, +W, and −d treatments decreased it to 11.7%, 9.7%, and 6.7%, respectively. Similarly to sandy samples, the calcium-Kα counts were the most affected by −SOM, and a great increase in sulfur and chlorine (−Kα) peaks were verified for −d treatment.

2.3.3 Pearson's correlation

The Pearson’s correlation coefficients between the pXRF spectra of the control samples and soil attributes (Fig. 6) showed strong correlation with the soil particle size distribution across the spectrum. Between 1.7 and 7.25 keV, it were observed mostly positive correlation coefficients between clay and silt contents and the

pXRF spectra. The exception to this pattern was observed close to the 1.8 keV range (corresponding to silicon-Kα peak), where a negative correlation was observed with the clay fraction content. The silt and SOM contents followed a similar pattern to the clay content, but with weaker correlations. Between 15 and 24 keV, an opposite pattern was observed for clay and sand contents, showing stronger negative and positive correlations, respectively.

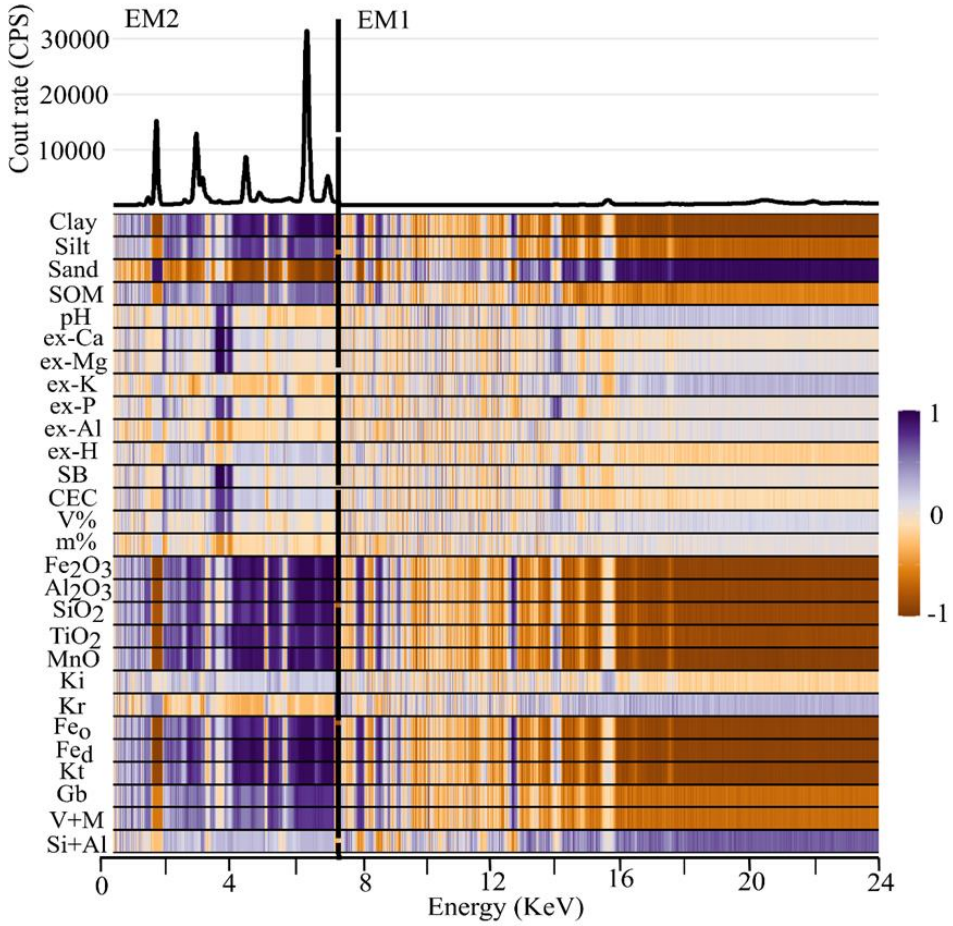

Fig. 6. Pearson's correlation between pXRF spectral data and soil attributes. EM1 = excitation mode 1; EM2 = excitation mode 2; SOM = soil organic matter; ex- = exchangeable/available forms; SB = sum of bases; CEC = cations exchange capacity at pH 7; V% = base saturation; m% = aluminum saturation; Ki = Ki index; Kr = Kr index; Fe_o = poorly crystalline iron forms; Fe_d = poorly crystalline and well crystalline iron forms; Kt = kaolinite; Gb: gibbsite; V + M: other clay minerals; Si + Al: silicon and aluminum poorly crystalline materials. The EM2 spectra range from 0 to 7.25 keV and the EM1 spectra range from 7.25 to 24 keV.

Regarding chemical attributes, the correlations, in general, were not significant across the spectra. However, close to 3.7 keV, corresponding to calcium (Fig. 6), significant correlations were observed with the attributes exchangeable/available calcium, magnesium and phosphorus, pH, SB, CEC and V%, in addition to the negative correlation with exchangeable/avaliable aluminum and hydrogen contents and m%.

The total Fe_2O_3, Al_2O_3, SiO_2, TiO_2, MnO, Feo and Fed and Kt contents showed positive correlations between 1.7 and 7.25 keV (Fig. 6). However, alike the clay content, close to 1.8 keV region, it was observed a negative correlation. Mostly negative correlations were also observed between 15 and 24 keV. The same behavior was observed for Gb and V + M%, but with smaller correlations and partially followed by Si + Al

content, that presented moderate positive correlation between 15 and 24 keV. The Ki and Kr weathering indices did not present significant correlations with the spectrum.

2.3.4 Linear regression (LR)

Generally, the LR between elements determined by the Geochem mode of pXRF and soil attributes showed high accuracy (Table 4). For the oxides from sulfuric acid digestion, the most accurate equations were found for MnO, Fe_2O_3, Al_2O_3 TiO_2, and SiO_2 contents, with R^2 of, 0.97, 0.95, 0.75, 0.75 and 0.64 and RMSE of 0.006, 2.108, 3.001, 0.910 and 4.122%, respectively. For Kt, the Fe from pXRF reached the highest R^2 and the lowest RMSE (0.89 and 3.635%) values, followed by silicon (0.78 and 5.258%) and aluminum (0.58 and 7.258%). For Gb, the Fe content from pXRF reached the highest R^2 and the lowest RMSE (0.61 and 2.741%) values, followed by aluminum (0.53 and 2.953%) and silicon (0.31 and 3.605%).

Table 4. Linear regression (LR) equations between pXRF elements content and soil mineralogical attributes.

Soil attribute (%)	pXRF element content (%)	Equation	R^2	RMSE (%)
Kt	Fe	**Kt** = 1.8543 x **Fe** +3.3682	0.89	3.635
	Al	**Kt =** 3.5965 x **Al** –23.544	0.58	7.258
	Si	**Kt** = -2.1463 x **Si** + 53.84	0.78	5.285
Gb	Fe	**Gb =** 0.5908 x **Fe** –0.7457	0.61	2.741
	Al	**Gb** = 1.3476 x **Al** – 11.771	0.53	2.953
	Si	**Gb** = -0.5271 x **Si** + 12.875	0.31	3.605
SiO_2	Si	**SiO_2** = -1.1989 x **Si** + 29.648	0.64	4.122
Al_2O_3	Al	**Al_2O_3** = 2.2248 x **Al** –16.288	0.75	3.001
TiO_2	Ti	**TiO_2** = 2.4711 x **Ti** –0.9443	0.75	0.910
MnO	Mn	**MnO** = 1.0134 x **Mn** + 0.00026	0.97	0.006
Fe_2O_3	Fe	**Fe_2O_3** = 1.2414 x **Fe** -1.4394	0.95	2.108

Where: Kt = kaolinite; Gb = gibbsite; RMSE = root mean square error. The pXRF elements were determined by the Geochem mode

2.4 Discussion

2.4.1 Descriptive analyses of laboratory data

The variability found in the soil attributes (Fig. 2, Table 1, Table 2) reflects the mineralogy among the studied soils. Highly weathered-leached soils present lower Ki and Kr values due to greater Kt, Gb, and well crystalline iron oxide minerals (Schaefer et al., 2008). The Oxisols (RH, TH and TE) dominated by low-activity clays have lower natural fertility due to the predominance of 1:1 clay minerals and iron and aluminum oxide minerals (Schaefer et al., 2008, Soil Survey Staff, 2014). Otherwise, TQ has a greater amount of quartz minerals (Soil Survey Staff, 2014). The mineralogy attributes influenced the physical and chemical soil attributes (Fig. 2, Table 2, Table 3). The 2:1 (vermiculite and or smectite), the 1:1 (kaolinite) clay minerals and the iron and aluminum oxide minerals, such as hematite and Gb, have a fundamental role in the physicochemical properties of the soil, such as structure, porosity and CEC (Barré et al., 2014).

The stratification of studied soil samples by soil particle size distribution, also separated them according to degree of weathering (Table 3). The RH, TH and TE soils have greater clay content, while TQ is composed mainly of sand-sized particles (Soil Survey Staff, 2014). The clay content also affects the water holding capacity, the SOM and several elements sorption (Ferreira et al., 1999). In Fig. A1, it is presented the

total content of aluminum, silicon, titanium and iron from pXRF for each sample showing the variation of soil chemical composition according to the weathering degree.

2.4.2. Treatments and X-ray configuration effects

2.4.2.1. Effects of X-ray tube configuration and Geochem mode

The EM1 is the most suited for the detection of elements after vanadium in the Periodic Table, while the EM2 is recommended for elements between magnesium and titanium. This different suitability is a consequence of the continuous radiation (bremsstrahlung) which, together with the characteristic anode peaks, excite the elements present in the sample. This profile is largely influenced by the X-ray voltage tube, which in turn, modifies both the excitation efficiency and the count rate (Tavares et al., 2020). The generalized count rate increase in EM2 (Fig. 4, Fig. 5, Fig. 6) is a consequence of the absence of the primary filter. Conversely, under EM1 mode the primary filter suppresses the beam flux on the sample and hence the count rate of the X-ray fluorescence phenomenon. However, this dumping compensated the gains on the detection limits since the decreasing of the continuous radiation facilitates the differentiation between signal and background.

The variation of quantified contents was smaller than the peaks variation and sometimes the count rate decrease was followed by an increase in some element content (like titanium) (Fig. 4, Fig. 5, Fig. 6). This is very common in XRF technique because of the use of equations for matrix effect correction (Gallhofer and Lottermoser, 2018, Rousseau, 2006). In this case, the Geochem mode uses a Compton correction to improve the determinations.

2.4.2.2 Effects of soil particle size distribution

The ability of pXRF to provide information on soil mineralogy and particle size distribution is evidenced by the spectral peaks of different textural classes (Fig. 3). For most elements, the count rate of pXRF peaks followed the same pattern as that obtained by sulfuric acid digestion (Table 3 and Fig. 3). A remarkable exception was found for SiO_2 and silicon-Kα that did not show this pattern since the sulfuric acid digestion does not extract most of the sandy fraction (Castro et al., 1984).

Since the sandy soils of this study are less weathered, the greater silicon peak and lower aluminum, iron, manganese and titanium peaks (Fig. 3) can be explained by the weathering processes. The sand fraction is composed mainly of silicon, which is present in the form of quartz minerals (Soil Survey Staff, 2014). Conversely, in the clay fraction of tropical soils (like Oxisols), there are great amounts of minerals rich in iron, titanium and aluminum (Schaefer et al., 2008). According to Macías and Camps-Arbestain (2020), the susceptibility to weathering of minerals can be accessed through the strength of the oxygen bound to the most mobile elements (Parker index) (Parker, 1970) and the mobility sequence is Ca, Na > Mg > K > Si > Al, Fe. Hence, this parameter can be promptly assessed by X-ray fluorescence spectrometry.

In highly weathered soils, basic cations, such as calcium, potassium, sodium and magnesium were almost completely lost due to leaching processes (Macías and Camps-Arbestain, 2020). However, normally, clayey soils have greater SOM contents due to the stabilization of smaller particles, mainly iron oxide minerals in acidic soils (Hassink, 1997, Lima and Anderson, 1997). These factors affect the sorption of several elements, mostly as outer-sphere complexes (Lima and Anderson, 1997, Schaefer et al., 2008). This behavior might explain the count rate of peaks from several elements, with decrease (calcium and potassium (−Kα)) or increase

(manganese, phosphorus, sulfur, cooper and zinc(−Kα)) in counts for clayey samples (Fig. 3). Sometimes, the peaks variation did not agree with the exchangeable/available contents (phosphorus and magnesium (−Kα)) (Table 2), which can be explained by the several factors that affect the plant accessible reservoirs, as management practices and fertilizer inputs.

The increase of background and scattering peaks for sandy soils (Fig. 3) is a consequence of matrix effects, caused by the diversity of elements forming the sample. The Thomson peak corresponds to the elastic scattering of X-rays, whereas the Compton peak is the inelastic scattering. The count rate of both peaks depends on the sample thickness and average atomic number. Since the measurements were carried under the infinitely thick regime, the first parameter can be neglected. However, one must expect the latter parameter to be highly affected by particle size distribution, since sandy soils tend to contain more light elements than clay soils. Iron and silicon play a major role in the intensity of the Compton peak and background radiation, making them more intense in lighter (sandier) matrixes and less intense in heavier (clayer) matrixes (Gallhofer and Lottermoser, 2018, Tavares et al., 2020). The great influence of particle size distribution in pXRF, related to elemental composition changes and matrix effects, pave the way to the detection of several soil properties.

2.4.2.3 Effects of soil moisture

The free water in pores and adsorbed to the surface of the soil constituent materials explains the generalized reduction of count rate (Fig. 4, Fig. 5). According to Ge et al. (2005), there are two types of water in soil, rock and sediments: (i) structural water, which is constant because it is part of material structure and (ii) free water in pores. The influence of moisture in pXRF is explained by two factors: (i) when the water content increases, there is a decrease in count rate of the characteristic peaks by decrease in the X-rays signal and (ii) water has the ability to interfere in the counts of scattered primary X-rays from the sources (Ge et al., 2005). The attenuation of X-rays by water is higher than that of air due to greater density of the latter compared to the former. As a consequence, the greater the sample moisture, the lower will be the net peak areas of characteristic X-rays that constitute the sample, resulting in lower precision, accuracy and detection limits (Bastos et al., 2012, Sahraoui and Hachicha, 2017, Santana et al., 2019). Therefore, the elemental peaks in wet soils are generally lower than in dry soil as observed between control and +W treatments (Fig. 4, Fig. 5).

The scattering peaks reduction more pronounced in sandy samples (Fig. 4, Fig. 5) evidences the greater effect that water caused in light elements than in heavy ones, mainly by reduction of silicon signal (Potts and Webb, 1992). The silicon content is greater in sandy samples than in clayey ones, which explains the matrix effect to be more significant. According to Bastos et al. (2012), moisture affects the accuracy especially for elements with an atomic number smaller than 30, such as magnesium, aluminum and silicon. Besides that, the Compton normalization present in the Geochem mode partially attenuated the effects of moisture in quantified contents, manly for heavy elements (titanium, iron and manganese). However, it was not enough to obtain results similar to the dry sample. In this way, the soil measurements in field condition is not recommended without an efficient method for moisture effects correction. For example, Bastos et al. (2012) proposed a correction method for titanium, iron and zirconium quantification based on an independent parameter: the background of titanium-Kα. Also, it is possible convert pXRF results obtained in the field to those achieved under laboratory conditions (Dijair et al., 2020).

2.4.2.4 Effects of soil organic matter

Soil organic matter (SOM) is mainly composed of carbon, hydrogen and oxygen, hence its average atomic number is small (Leifeld et al., 2020). Usually, X-ray fluorescence is not suitable for elements of atomic number below magnesium due to the low fluorescence yield of such elements combined to the degree of attenuation of the radiation. Conversely, low atomic number elements contribute to increase the Compton scattering. Preliminary measurements, not shown here, revealed that Compton scattering alone was insensitive to organic matter ranging from 1.5 up to 3.5%, whereas its intensity started to increase above 5% (Fig. A.2). Thus, lower SOM contents in soil samples do not imply in accuracy decreased.

The removal of SOM can, in principle, increase the overall concentration of the remaining elements (Fig. 4, Fig. 5). However, since the samples presented low concentration of SOM, i.e. around 14 g kg^{-1} (Table 2), such effect was not expected to be pronounced. The decrease of concentration observed for most elements in sandy soils might be a result of leaching promoted by the reagents used to oxidize the SOM. This was especially remarkable for calcium, which interacts via outer-sphere complex with soil components.

In clayey samples, the CEC depends mainly on SOM and clay fraction. Although, sandy soils contain less SOM, the CEC is even more dependent on it (Asadu and Chibuike, 2015). Therefore, the removal effect was more pronounced in sandy samples (Fig. 4). It is possible to observe a count rate increase of some peaks, mainly for clayey samples (magnesium, aluminum, phosphorus, titanium and iron (−Kα)) (Fig. 4, Fig. 5). These elements (except for phosphorus- Kα) are structural constituents of soil minerals and the hydrogen peroxide has a small effect on them (Table A.1). The phosphorus behavior can be explained by the formation of inner-sphere complexes with the ferrol (FeOH) and aluminol (AlOH)) surface functional complex groups formed on the broken edges of iron and aluminum oxide minerals in acid soils (Smithson, 1999). Another factor related to counts increase of some elements can be the annulation of matrix effect caused by the elements removed altogether with SOM.

2.4.2.5 Effects of iron forms

The iron forms include silicate-free iron oxides that involve well crystalline, poorly crystalline and non-crystalline oxides (Table A.1). These forms are affected differently by iron dissolution treatments, due to different extraction actions, either by sodium citrate-dithionite-bicarbonate or by ammonium oxalate acid in the dark (Table A.1). The ammonium oxalate acid extraction is a selective dissolution method for poorly crystalline iron forms (Rennert, 2019), which occur in small contents in highly weathered soils (Schaefer et al., 2008). The high complexing power of the oxalate anion and the effect of the ammonium oxalate solution are responsible for extracting ferrihydrite and small amounts of organically bound iron (Table A.1). In −o treatment, it was verified a small reduction in iron-Kα counts (Fig. 5, Fig. 6), indicating that the iron of the soils was associated mainly with well crystalline oxides. Inda et al. (2014) elucidate that in aerobic pedo-environments like tropical soils, iron oxide minerals exhibited high stability and are able to persist in the soil for long periods, for this reason, they tend to have low poorly crystalline iron content.

In sandy soils, the decrease of aluminum and manganese contents (Fig. 4, Fig. 5) suggests that the −o treatment removed not only iron, but also aluminum and manganese, in poorly crystalline forms (Table A.1). However, for clayey soils the signal of aluminum increased, which might be a consequence of modification of matrix composition due to a more significant reduction in iron content. Even though the −o treatment can affect

silicon content by removing extractable forms, represented by silicon from poorly crystalline soil components (Table A.1), no decrease was observed in silicon-Kα intensity (Fig. 4, Fig. 5). This can occur due to the low content of poorly crystalline silicon in the samples and matrix changes.

Conversely, the −d treatment showed decrease on several peaks signal in sandy and clayey samples compared to the control (Fig. 4, Fig. 5). This decrease is expected because sodium citrate- dithionite-bicarbonate dissolves selectively all forms of pedogenic iron oxide minerals (Table A.1). This extraction allows one to determine, in addition to iron, the content of other elements that can isomorphically replace iron within the iron oxide minerals structure (Table A.1). The dissolution of poorly crystalline aluminum oxides by sodium citrate-dithionite-bicarbonate explains the aluminum-Kα peak intensity reduction. The decrease in the silicon-Kα intensity (Fig. 4, Fig. 5) is related to the poorly crystalline silica compounds (Table A.1). Besides, the sulfur and chlorine count rates increased for −d treatment are due to the use of ammonium dithionite and sodium chloride reagents. The spectral changes due to selective dissolution treatments (−o and −d) demonstrate the capacity of pXRF to capture changes in soil chemical composition and corroborate the elucidation of soil pedogenic processes.

2.4.3 Pearson's correlation

The results (Fig. 6) are in agreement with previous studies that show correlation between soil particle size distribution and some elements. Zhu et al. (2011) described positive correlation between soil clay fraction and titanium, chromium, iron, cobalt, copper, zinc, calcium, arsenic, lead, manganese, and potassium. Conversely, negative correlation coefficients were observed between the soil sandy fraction and these same elements. Corroborating this, Benedet et al., 2020, Silva et al., 2018 described the positive coefficients with the same elements and others like chlorine, nickel and aluminum, in addition to the silicon, which were inversely related with the soil clay fraction (Fig. 4).

Positive correlations were found between the spectral region around 3.6 keV, which corresponds to calcium peak, with SOM, pH, SB, CEC, V%, and also with exchangeable/available calcium, magnesium, potassium and phosphorus. Sharma et al., 2014, Zhu and Weindorf, 2009 also reported positive correlations between calcium peak and pH. Likewise, Sharma et al. (2015) were also successful in obtaining CEC values, using Ca content from pXRF. Such positive correlation is a result of the chemical role played by calcium, which affects, in tropical cultivated soils, pH, CEC, V%, and magnesium concentration.

There are few studies showing correlations between pXRF data and mineralogical attributes. Similar to the present study, de Lima et al. (2019) described positive correlation between clay fraction and FeO_2 and Al_2O_3, in addition to inverse relationship with SiO_2. Thus, the negative correlation between mineralogical attributes and those obtained by sulfuric acid digestion related with the spectrum feature corresponding to silicon peak (1.8 keV) (Fig. 6) is justified, and can reflect the weathering-leaching processes. It is important to note that, in the present study, silicon signal was positively correlated to sand fraction content (Fig. 6). However, it was not positively correlated with SiO_2 determined by the sulfuric acid digestion. This disagreement is due to the difference between the methods of analyses. The quartz in the sand fraction is not dissolved by H_2SO_4 (Mehra and Jackson, 2013, Castro et al., 1984), while pXRF is able to capture the total silicon, including the sand fraction (mainly composed of quartz mineral). This disagreement is accentuated by the great sand content in part of the samples (Fig. 1).

Between 15 and 24 keV, it were observed mostly negative correlations between the studied attributes and the spectra (Fig. 6). This pattern can be explained by the inverse relationship between scattering peaks and silicon content, as described above.

2.4.4. Linear regression

Previous studies also presented equations, such as those in Table 4, which were able to predict the total and pseudo-total content of titanium, silicon, iron, and aluminum. In tropical soils, Silva et al. (2020) found equations with high accuracy for Fe_2O_3 and TiO_2, using linear regression. While, Silva et al. (2019) found accurate results for Fe_2O_3 and MnO. O'Rourke et al. (2016) reported results for iron, manganese, titanium, and aluminum total contents and for other several elements. As a general trend, the results for iron predictions presented higher accuracy than for lighter elements, like silicon and aluminum (O'Rourke et al., 2016, Santana et al., 2018, Silva et al., 2020).

The high accuracy for MnO and Fe_2O_3 (Table 4) is related to the great molar mass of iron and manganese (Lindgren, 2006) and the high potential for assessing such oxides by the sulfuric acid digestion (Castro et al., 1984).

Santana et al. (2018) were able to determine Fe_2O_3 under laboratory and field conditions. Additionally, the equations for TiO_2 and MnO achieved great accuracy in laboratory condition. However, the SiO_2 and Al_2O_3 presented less accurate predictions. Silva et al. (2020) found low accuracy results for Al_2O_3 and SiO_2 with linear regression, but high accuracy with multiple linear regression, with inclusion of other elements and soil particle distribution data in the models. In this case, the low accuracy for SiO_2 can be related to the inefficient extraction of SiO_2 by sulfuric acid digestion in the sand fraction (Castro et al., 1984). Besides that, in this study, SiO_2 delivered by sulfuric acid digestion was inversely proportional to silicon content from pXRF.

The pXRF was able to estimate Kt content and the best results were obtained using iron and aluminum as independent variables (Table 3). Kt is a 1:1 clay mineral, composed of aluminum and silicon and it is the most common phyllosilicate mineral found in the soils of the humid tropical and subtropical climates (Schaefer et al., 2008). In this way, Jozanikohan et al. (2016) found accurate results for clay minerals, mainly using as predictor variables the iron and aluminum content from pXRF, and also using multivariate statistics, with several elements.

For Gb, the pXRF also provided accurate results (Table 4). Although Gb is basically formed by layers of aluminum (octahedra), the LR showed less adjustment for aluminum in relation to iron for this mineral. It is explained by the indirect relation between the iron and Gb content, which occurs since soils rich in iron also tend to have great Gb contents, driven by leaching processes (Schaefer et al., 2008). Besides that, the iron has greater molar mass than aluminum, promoting more accurate determinations by pXRF.

The pXRF was able to access the particle size distribution modifications, water addition and dissolution procedures to remove SOM, poorly crystalline iron forms and well crystalline iron forms in soil samples.

2.5 Conclusions

The soil particle size distribution implied in changes by different elemental composition of samples and matrix effects caused by it, which is related to soil mineralogy and weathering processes and pave the way for determination of several soil attributes.

The presence of lower contents of SOM in the soil samples do not imply mandatorily in accuracy decrease. However, alterations in pXRF data were detected by removal of SOM with hydrogen peroxide (i.e. diminution of calcium peak).

The water addition reduced the signal of the elements, which must be considered if measurements are to be carried out under field conditions, and methods to correct the soil moisture effects should be used. Such effects were more pronounced for light elements, such as aluminum and silicon.

The well crystalline iron removal by sodium citrate-dithionite-bicarbonate promoted high decrease in iron content and other elements also were removed. Conversely, the acid ammonium oxalate promoted only moderate alterations in pXRF data due the low amount of poorly crystalline iron oxides in tropical soils.

The pXRF demonstrated to be a very useful tool for mineralogical and chemical studies, delivering satisfactory results for Kt, Gb, Fe_2O_3, Al_2O_3, SiO_2, TiO_2 and MnO contents estimation by LR. Besides that, the qualitative and Person's correlation analyses showed relationships between the pXRF data and both soil particle size distribution and soil fertility attributes.

The understanding of the fundamentals of interactions between pXRF energy with the sample matrix is the starting point for characterizing the soil mineralogy through this technique.

References

Afriyie, E., Verdoodt, A., Mouazen, A.M., 2020. Estimation of aggregate stability of some soils in the loam belt of Belgium using mid-infrared spectroscopy. Sci. Total Environ. 744, 140727. https://doi.org/10.1016/j.scitotenv.2020.140727.

Allo, M., Todoroff, P., Jameux, M., Stern, M., Paulin, L., Albrecht, A., 2020. Prediction of tropical volcanic soil organic carbon stocks by visible-near- and mid-infrared spectroscopy. Catena 189, 104452. https://doi.org/10.1016/j.catena.2020.104452.

Andrade, R., Faria, W.M., Silva, S.H.G., Chakraborty, S., Weindorf, D.C., Mesquita, L.F., Guilherme, L.R.G., Curi, N., 2020a. Prediction of soil fertility via portable X-ray fluorescence (pXRF) spectrometry and soil texture in the Brazilian Coastal Plains. Geoderma 357, 113960. https://doi.org/10.1016/j.geoderma.2019.113960.

Andrade, R., Silva, S.H.G., Weindorf, D.C., Chakraborty, S., Faria, W.M., Mesquita, L.F., Guilherme, L.R.G., Curi, N., 2020b. 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.

Anjos, L.H.C, Jacomine, P.T.K., Santos, H.G., Oliveira, V.A., Oliveira, J.P, 2012. Sistema brasileiro de classificação de solos. Pedologia - Fundamentos, first ed. SBCS, Viçosa, pp. 303–343.

Asadu, C.L.A., Chibuike, G.U., 2015. Contributions of organic matter, clay and silt to the effective CEC of soils of different land use history. Adv. Nat. Appl. Sci. 9, 110–115.

Barré, P., Fernandez-Ugalde, O., Virto, I., Velde, B., Chenu, C., 2014. Impact of phyllosilicate mineralogy on organic carbon stabilization in soils: Incomplete knowledge and exciting prospects. Geoderma 235–236, 382–395. https://doi.org/ 10.1016/j.geoderma.2014.07.029.

Barthès, B.G., Kouakoua, E., Coll, P., Clairotte, M., Moulin, P., Saby, N.P.A., Le Cadre, E., Etayo, A., Chevallier, T., 2020. Improvement in spectral library-based quantification of soil properties using representative spiking and local calibration – the case of soil inorganic carbon prediction by mid-infrared spectroscopy. Geoderma 369, 114272. https://doi.org/10.1016/j.geoderma.2020.114272.

Bastos, R.O., Melquiades, F.L., Biasi, G.E.V., 2012. Correction for the effect of soil moisture on in situ XRF analysis using low-energy background. X-Ray Spectrom. 41 (5), 304–307. https://doi.org/10.1002/xrs.v41.510.1002/xrs.2397.

Benedet, L., Faria, W.M., Silva, S.H.G., Mancini, M., Demattˆe, J.A.M., Guilherme, L.R.G., Curi, N., 2020. Soil texture prediction using portable X-ray fluorescence spectrometry and visible near-infrared diffuse reflectance spectroscopy. Geoderma 376, 114553. https://doi.org/10.1016/j.geoderma.2020.114553.

Biney, J.K.M., Boruvka, L., Agyeman, P.C., N˘eeme˘ecek, K., Klement, A., 2020. Comparison of field and laboratory wet soil spectra in the vis-NIR range for soil organic carbon prediction in the absence of laboratory dry measurements. Remote Sens. 12, 1–16. https://doi.org/10.3390/RS12183082.

Bouma, J., 2020. Contributing pedological expertise towards achieving the United Nations Sustainable Development Goals. Geoderma 375, 114508. https://doi.org/ 10.1016/j.geoderma.2020.114508.

Brown, G., 1953. The occurrence of lepidocrocite in some British soils. J. Soil Sci. 4 (2), 220–228. https://doi.org/10.1111/j.1365-2389.1953.tb00655.x.

Castro, F.J.O., Antunes, F.S., Polivanov, H., Santos, R.D., Gama, J.R.N.F., 1984. Perspectivas do uso do ataque sulfúrico e da dissolução alcalina para a análise mineralógica expedita de solos com B textural e B latossólico, Empresa Brasileira de Pesquisa Agropecuária: Boletim de Pesquisa no 34. EMBRAPA, Rio de Janeiro. Chipera, S.J., Bish, D.L., 2001.

Baseline studies of the clay minerals society source clays powder X-ray diffraction analyses. Clays Clay Miner. 49, 398–409. https://doi.org/ 10.1346/CCMN.2001.0490507.

Cornell, R.M., Schwertmann, U., 2003. The Iron Oxides: Structure, Properties, Reactions, Occurences and Uses, second ed. John Wiley & Sons.

Costa, J.J.F., Giasson, E., da Silva, E.B., Coblinski, J.A., Tiecher, T., 2020. Use of color parameters in the grouping of soil samples produces more accurate predictions of soil texture and soil organic carbon. Comput. Electron. Agric. 177, 105710. https:// doi.org/10.1016/j.compag.2020.105710.

Demattê, J.A.M., Campos, R.C., Alves, M.C., Fiorio, P.R., Nanni, M.R., 2004. Contribution of the chemical and mineralogical properties of sandy-loam tropical soils to the cation exchange capacity. Rev. Bras. Cienc. do Solo 44, 1–18. https://doi. org/10.36783/18069657rbcs20200019

Demattê, J.A.M., Campos, R.C., Alves, M.C., Fiorio, P.R., Nanni, M.R., 2004. Visible-NIR reflectance: A new approach on soil evaluation. Geoderma 121 (1-2), 95–112. https://doi.org/10.1016/j. geoderma.2003.09.012.

da Costa, A.C.S., Junior, I.G.de.S., Canton, L.C., Gil, L.G., Figueiredo, R., 2020. Contribution of the chemical and mineralogical properties of sandy-loam tropical soils to the cation exchange capacity. Rev. Bras. Cienc. do Solo 44, 1–18. https://doi.org/10.36783/18069657rbcs20200019.

Demattˆe, J.A.M., Garcia, G.J., 1999. Alteration of soil properties through a weathering sequence as evaluated by spectral reflectance. Soil Sci. Soc. Am. J. 63 (2), 327–342. https://doi.org/10.2136/sssaj1999.03615995006300020010x.

Demattê, J.A.M., Nanni, M.R., Formaggio, A.R., Epiphanio, J.C.N., 2007. Spectral reflectance for the mineralogical evaluation of Brazilian low clay activity soils. Int. J. Remote Sens. 28 (20), 4537–4559. https://doi.org/10.1080/01431160701250408.

Demattê, J.L.I., Demattê J.A.M., 2009. Ambientes de produção como estratégia de manejo na cultura da cana-de-açucar. International Plant Nutrition Institute, 27. [WWW Document]. URL http://www.ipni.net/publication/ia-brasil.nsf/0/7DA98D7CA7D3007E83257A900012F0DF/$FILE/Jornal-127.pdf (accessed 05.10.21).

Dewitte, O., Jones, A., Spaargaren, O., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Mich´eli, E., Montanarella, L., Thiombiano, L., Van Ranst, E., Yemefack, M., Zougmore, R., 2013. Harmonisation of the soil map of africa at the continental scale. Geoderma 211–212, 138–153. https://doi.org/10.1016/j.geoderma.2013.07.007.

Dijair, T.S.B., Silva, F.M., Teixeira, A.F.D.S., Silva, S.H.G., Guilherme, L.R.G., Curi, N., 2020. Correcting field determination of elemental contents in soils via portable X-ray fluorescence spectrometry. Ciˆencia e Agrotecnologia 44. https://doi.org/10.1590/ 1413-7054202044002420.

Dixon, J.D., Weed, S.b., 1989. Minerals in Soil Environments, SSSA Book Series. Soil Science Society of America, Madison. https://doi.org/10.2136/sssabookser1.2ed.

dos Anjos, L.H.C., Franzmeier, D.P., Schulze, D.G., 1995. Formation of soils with plinthite on a toposequence in Maranh o State, Brazil. Geoderma 64 (3–4), 257–279. https:// doi.org/10.1016/0016-7061(94)00022-3.

Drosdoff, M., Miles, E.F., 1983. Action of hydrogen peroxide on weathered mica. Soil Sci. 46 (5), 391–395. https://doi.org/10.1097/00010694-193811000-00004.

Empresa Brasileira de Pesquisa Agropecuária, 1995. Procedimentos Normativos de An´alises de Solos. EMBRAPA, Brasília. Empresa Brasileira de Pesquisa Agropecu´aria, 1997. Manual de Métodos de Análises de Solo, 2ª edição. EMBRAPA, Rio de Janeiro.

Farmer, V.C., Mitchell, B.D., 1993. Occurrence of oxalates in soil clays following hydrogen peroxide treatment. Soil Sci. 96 (4), 221–229. https://doi.org/10.1097/ 00010694-196310000-00001.

Fernandes, A.J., Maldaner, C.H., Sobrinho, J.M.A., Pressinotti, M.M.N., Wahnfried, I., 2010. Estratigrafia dos Derrames de Basaltos da Formação Serra Geral (Ribeirão Preto - SP) Baseada na Geologia Física, Petrografia e Geoquímica. Rev. do Inst. Geociências – USP 10 (2), 73–99.

Ferreira, M.M., Fernandes, B., Curi, N., 1999. Influência da mineralogia da fração argila nas propriedades físicas de latossolos da região sudeste do Brasil. Rev. Bras. Ciência do Solo 23 (3), 515–524. https://doi.org/10.1590/S0100-06831999000300004.

Fitzpatrick, R.W., Schwertmann, U., 1982. Al-substituted goethite-An indicator of pedogenic and other weathering environments in South Africa. Geoderma 27 (4), 335–347. https://doi.org/10.1016/0016-7061(82)90022-2.

Gallhofer, D., Lottermoser, B., 2018. The influence of spectral interferences on critical element determination with portable X-ray fluorescence (pXRF). Minerals 8 (8), 320. https://doi.org/10.3390/min8080320.

Ge, L., Lai, W., Lin, Y., 2005. Influence of and correction for moisture in rocks, soils and sediments on in situ XRF analysis. X-Ray Spectrom. 34 (1), 28–34. https://doi.org/10.1002/(ISSN)1097-453910.1002/xrs.v34:110.1002/xrs.782.

Gomez, C., Chevallier, T., Moulin, P., Bouferra, I., Hmaidi, K., Arrouays, D., Jolivet, C., Barth`es, B.G., 2020. Prediction of soil organic and inorganic carbon concentrations in Tunisian samples by mid-infrared reflectance spectroscopy using a French national library. Geoderma 375, 114469. https://doi.org/10.1016/j.geoderma.2020.114469.

Hallmark, C.T., Wilding, L.P., Smeck, E., 1982. Silicon. In: Page, A.L., Miller, R.H., Kenny, D.R. (Eds.), Methods of soil analysis: Part 2. Soil Science Society of America, Madison, pp. 263–273. Harada, Y., Inoko, A., 1977. The oxidation products formed from: Soil organic matter by hydrogen peroxide treatment. Soil Sci. Plant Nutr. 23 (4), 513–521. https://doi.org/10.1080/00380768.1977.10433069.

Hassink, J., 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. Plant Soil 191, 77–87. https://doi.org/10.1023/A: 1004213929699.

Holmgren, G.G.S., 1967. A rapid citrate-dithionite extractable iron procedure. Soil Sci. Soc. Am. J. 31 (2), 210–211. https://doi.org/10.2136/sssaj1967.03615995003100020020x.

Inda, A.V., Tomasi, C.A., de Oliveira, J.S., Fink, J.R., 2014. ´Oxidos de ferro e ´area superficial de Latossolo subtropical sob campo e floresta nativa. Cienc. Rural 44, 289–292. https://doi.org/10.1590/S0103-84782014000200015.

Instituto Brasileiro de Geografia e Estatística, 2020. Cidades e Estados [WWW Document]. URL https://cidades.ibge.gov.br (accessed 12.20.20).

Jackson, M.L., 1969. Soil Chemical Analysis: Advanced Course. University of Wisconsin, Madison. Jozanikohan, G., Sahabi, F., Norouzi, G.H., Memarian, H., Moshiri, B., 2016. Quantitative analysis of the clay minerals in the Shurijeh Reservoir Formation using combined X-ray analytical techniques. Russ. Geol. Geophys. 57, 1048–1063. https://doi.org/10.1016/j.rgg.2016.06.005.

Kunze, G.W., Dixon, J.B., 1986. Pretreatment for Mineralogical Analysis, in: Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods, second ed. pp. 91–100.

Leifeld, J., Klein, K., Wüst-Galley, C., 2020. Soil organic matter stoichiometry as indicator for peatland degradation. Sci. Rep. 10, 1–9. https://doi.org/10.1038/s41598-020-64275-y.

de Lima, T.M., 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.

Lima, J.M., Anderson, S.J., 1997. Aggregation and aggregate size effects on extractable iron and aluminum in two hapludoxs. Soil Sci. Soc. Am. J. 61 (3), 965–970. https://doi.org/10.2136/sssaj1997.03615995006100030036x.

Lindgren, E.S., 2006. Energy dispersive, X-ray fluorescence analysis. In: Encyclopedia of Analytical Chemistry. John Wiley & Sons Ltd, Chichester, UK, pp. 71–114. https://doi.org/10.1002/9780470027318.a6806.

Ma, Y., Minasny, B., McBratney, A., Poggio, L., Fajardo, M., 2021. Predicting soil properties in 3D: Should depth be a covariate? Geoderma 383, 114794. https://doi.org/10.1016/j.geoderma.2020.114794.

Macías, F., Camps-Arbestain, M., 2020. A biogeochemical view of the world reference base soil classification system: Homage to Ward Chesworth, first ed. In: Advances in Agronomy. Elsevier Inc. https://doi.org/10.1016/bs.agron.2019.11.002.

Mammadov, E., Denk, M., Riedel, F., Lewinska, K., Ka´zmierowski, C., Glaesser, C., 2020. Visible and near-infrared reflectance spectroscopy for assessment of soil properties in the caucasus mountains. Azerbaijan. Commun. Soil Sci. Plant Anal. 51 (16), 2111–2136. https://doi.org/10.1080/00103624.2020.1820027.

McKeague, J.A., Brydon, J.E., Milles, N.M., 1971. Differentiation of Forms of Extractable Iron and Aluminum in Soils. Proc. Soil Sci. Soc. Am. Proc. 35 (1), 33–38.https://doi org/10.2136/sssaj1971.03615995003500010016x.

McKeague, J.A., Day, J.H., 1966. Dithionite- and oxalate-extractable Fe and Al as aids in differentiating various classes of soils. Can. J. Soil Sci. 46 (1), 13–22. https://doi.org/10.4141/cjss66-003.

McManus, D.A., 1991. Suggestions for authors whose manuscripts include quantitative clay mineral analysis by X-ray diffraction. Mar. Geol. 98 (1), 1–5. https://doi.org/10.1016/0025-3227(91)90030-8.

Mehra, O.P., Jackson, M.L., 2013. Iron oxide removal from soils and clays by a dithionite–citrate system buffered with sodium bicarbonate. Clays Clay Miner. 317–27 https://doi.org/10.1016/B978-0-08-09235-5.50026-7.

Mitchell, B.C., Farmer, V.C., McHardy, J., 1964. Amorphous Inorganic Materials In Soils. Adv. Agron. 16 (C), 327–383. https://doi.org/10.1016/S0065-2113(08)60028-0.

Muggler, C.C., Buurman, P., van Doesburg, J.D.J., 2007. Weathering trends and parent material characteristics of polygenetic oxisols from Minas Gerais, Brazil: I. Mineralogy. Geoderma 138 (1-2), 39–48. https://doi.org/10.1016/j.geoderma.2006.10.008.

O'Rourke, S.M., Minasny, B., Holden, N.M., McBratney, A.B., 2016. Synergistic use of Vis-NIR, MIR, and XRF spectroscopy for the determination of soil geochemistry. Soil Sci. Soc. Am. J. 80 (4), 888–899. https://doi.org/10.2136/sssaj2015.10.0361.

Parfitt, R.L., Henmi, T., 1982. Comparison of an oxalate-extraction methon and an infrared spectroscopic method for determining allophane in soll clays. Soil Sci. Plant Nutr. 28 (2), 183–190. https://doi.org/10.1080/00380768.1982.10432435.

Parker, A., 1970. An index of weathering for silicate rocks. Geol. Mag. 107 (6), 501–504. https://doi.org/10.1017/S0016756800058581.

Pinheiro-Dick, D., Schwertmann, U., 1996. Microaggregates from oxisols and inceptisols: dispersion through selective dissolutions and physicochemical treatments. Geoderma 74 (1-2), 49–63. https://doi.org/10.1016/S0016-7061(96)00047-X.

Potts, P.J., Webb, P.C., 1992. X-ray fluorescence spectrometry Philip. In: Hall, G.H.M. (Ed.), Geoanalysis, 44. Elsevier Science Publishers, Amsterdam, pp. 251–296.

R Development Core Team, 2020. The R Project for Statistical Computing [WWW Document]. URL https://www.r-project.org (accessed 12.12.20).

van Raij, B., Andrade, J.C., Cantarella, H., Quaggio, J.A., 2001. Análise química para avaliação da fertilidade de solos tropicais, Campinas: Instituto Agronômico. IAC, Campinas.

Rennert, T., 2019. Wet-chemical extractions to characterise pedogenic Al and Fe species- a critical review. Soil Res. 57, 1–16. https://doi.org/10.1071/SR18299.

Ribeiro, B.T., Silva, S.H.G., Silva, E.A., Guilherme, L.R.G., 2017. Aplicações da fluorescência de raios-X portátil (pXRF) na Ciência do Solo tropical. Cienc. e Agrotecnologia 41, 245–254. https://doi.org/10.1590/1413-70542017413000117.

Rousseau, R.M., 2006. Corrections for matrix effects in X-ray fluorescence analysis-A tutorial. Spectrochim. Acta – Part B At. Spectrosc. 61 (7), 759–777. https://doi.org/ 10.1016/j.sab.2006.06.014.

Sahraoui, H., Hachicha, M., 2017. Effect of soil moisture on trace elements concentrations. J. Fundam. Appl. Sci. 9, 468–484. Santana, M.L.T., Carvalho, G.S., Guilherme, L.R.G., Curi, N., Ribeiro, B.T., 2019. Elemental concentration via portable x-ray fluorescence spectrometry: assessing the impact of water content. Cienc. e Agrotecnologia 43. https://doi.org/10.1590/1413-7054201943029218.

Santana, M.L.T., Ribeiro, B.T., Silva, S.H.G., Poggere, G.C., Guilherme, L.R.G., Curi, N., 2018. Conditions affecting oxide quantification in unknown tropical soils via handheld X-ray fluorescence spectrometer. Soil Res. 56, 648–655. https://doi.org/10.1071/SR18099.

Schaefer, C.E.G.R., Fabris, J.D., Ker, J.C., 2008. Minerals in the clay fraction of Brazilian Latosols (Oxisols): a review. Clay Miner. 43 (1), 137–154. https://doi.org/10.1180/ claymin.2008.043.1.11.Schwertmann, U., 1964.

Differenzierung der Eisenoxide des Bodens durch Extraktion mit Ammoniumoxalat-L¨osung. Zeitschrift für Pflanzenern¨ahrung, Düngung 105 (3), 194–202. https://doi.org/10.1002/jpln.3591050303.

Schwertmann, U., Carlson, L., 1994. Aluminum Influence on Iron Oxides: XVII. Unit-Cell Parameters and Aluminum Substitution of Natural Goethites. Soil Sci. Soc. Am. J. 58 (1), 256–261. https://doi.org/10.2136/sssaj1994.03615995005800010039x.

Schwertmann, U., Taylor, R., 1989. Iron Oxides. In: Dixon, J., Weed, S. (Eds.), Minerals in Soil Environments, second ed.1. Soil Science Society of America, Madison, pp. 379–438. Shah, F., Wu, W., 2019.

Soil and crop management strategies to ensure higher crop productivity within sustainable environments. Sustainability 11, 1485. https://doi. org/10.3390/su11051485.

Sharma, A., Weindorf, D.C., Man, T., Aldabaa, A.A.A., Chakraborty, S., 2014. Characterizing soils via portable X-ray fluorescence spectrometer: 3. Soil reaction (pH). Geoderma 232–234, 141–147. https://doi.org/10.1016/j.geoderma.2014.05.005.

Sharma, A., Weindorf, D.C., Wang, D.D., Chakraborty, S., 2015. Characterizing soils via portable X-ray fluorescence spectrometer: 4. Cation exchange capacity (CEC). Geoderma 239, 130–134. https://doi.org/10.1016/j.geoderma.2014.10.001

Silva, E.A., Weindorf, D.C., Silva, S.H.G., Ribeiro, B.T., Poggere, G.C., Carvalho, T.S., Gonçalves, M.G.M., Guilherme, L.R.G., Curi, N., 2019. Advances in tropical soil characterization via portable X-ray fluorescence spectrometry. Pedosphere 29 (4), 468–482. https://doi.org/10.1016/S1002-0160(19)60815-5.

Silva, S.H.G., Hartemink, A.E., Teixeira, A.F.dos S., Inda, A.V., Guilherme, L.R.G.., Curi, N., 2018. Soil weathering analysis using a portable X-ray fluorescence (pXRF)spectrometer in an Inceptisol from the Brazilian Cerrado. Appl. Clay Sci. 162, 27–37. https://doi.org/10.1016/j.clay.2018.05.028.

Silva, S.H.G., Ribeiro, B.T., Guerra, M.B.B., de Carvalho, H.W.P., Lopes, G., Carvalho, G. S., Guilherme, L.R.G., Resende, M., Mancini, M., Curi, N., Rafael, R.B.A., Cardelli, V., Cocco, S., Corti, G., Chakraborty, S., Li, B., Weindorf, D.C., 2021. pXRF in tropical soils: Methodology, applications, achievements and challenges. https://doi.org/10 .1016/bs.agron.2020.12.001.

Silva, S.H.G., Silva, E.A., Poggere, G.C., Junior, A.L.P., Gonçalves, M.G.M., Guilherme, L. R.G., Curi, N., 2020. Modeling and prediction of sulfuric acid digestion analyses data from pxrf spectrometry. Sci. Agric. 77 https://doi.org/10.1590/1678-992x-2018-0132.

Silvero, N.E.Q.., Di Raimo, L.A.D.L., Pereira, G.S., Magalh˜aes, L.P.de., Terra, F.da.S., Dassan, M.A.A., Salazar, D.F.U., Demattˆe, Jos´e.A.M., 2020. Effects of water, organic matter, and iron forms in mid-IR spectra of soils: assessments from laboratory to satellite-simulated data. Geoderma 375, 114480. https://doi.org/10.1016/j. geoderma.2020.114480.

Smithson, P., 1999. Special issue on phosphorus availability, uptake and cycling in tropical agroforestry. Agroforestry Forum. 1–68. Soil Survey Staff, 2014. Keys to Soil Taxonomy, twelfth ed. United States Department of Agriculture, Washington.

Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., Macdonald, L.M., McLaughlin, M. J., 2014. The performance of visible, near and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49 (2), 139–186. https://doi.org/10.1080/05704928.2013.811081.

Souza, I.F., Archanjo, B.S., Hurtarte, L.C.C., Oliveros, M.E., Gouvea, C.P., Lidizio, L.R., Achete, C.A., Schaefer, C.E.R., Silva, I.R., 2017. Al-/Fe-(hydr)oxides–organic carbon associations in Oxisols—from ecosystems to submicron scales. Catena 154, 63–72. https://doi.org/10.1016/j.catena.2017.02.017.

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. https://doi.org/10.1016/j.catena.2016.01.007.

Tavares, T.R., Molin, J.P., Nunes, L.C., Alves, E.E.N., Melquiades, F.L., de Carvalho, H.W. P., Mouazen, A.M., 2020. Effect of x-ray tube configuration on measurement of key soil fertility attributes with XRF. Remote Sens. 12 (6), 963. https://doi.org/10.3390/ rs12060963.

Teixeira, P.C., et al., 2017. Manual de m´etodos de an´alise de solo. EMBRAPA, Brasília, p. 212. Trolard, F., Bourrie, G., Jeanroy, E., Herbillon, A.J., Martin, H., 1995. Trace metals in natural iron oxides from laterites: A study using selective kinetic extraction. Geochim. Cosmochim. Acta 59 (7), 1285–1297. https://doi.org/10.1016/0016-7037 (95)00043-Y.

Wadoux, A.M.J.-C., Rom´an-Dobarco, M., McBratney, A.B., 2021. Perspectives on data- driven soil research. Eur. J. Soil Sci. 72 (4), 1675–1689. https://doi.org/10.1111/ ejss.v72.410.1111/ejss.13071. Walkley, A., Black, I.A., 1934. An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci. 37 (1), 29–38. https://doi.org/10.1097/00010694-193401000-00003.

Weaver, R.M., Syers, J.K., Jackson, L., 1968. Determination of Silica in Citrate- Bicarbonate-Dithionite Extracts of Soils. Soil Sci. Soc. Am. J. 32 (4), 497–501. https://doi.org/10.2136/sssaj1968.03615995003200040023x.

Weindorf, D.C., Bakr, N., Zhu, Y., 2014. Advances in portable X-ray fluorescence (PXRF) for environmental, pedological, and agronomic applications. Adv. Agron. https:// doi.org/10.1016/B978-0-12-02139-2.00001-9.

Whittig, L.D., 1965. X-ray diffraction techniques for mineral identification and mineralogical composition. In: Black, C.A. (Ed.), Methods of Soil Analysis: Part 1 Physical and Mineralogical Properties, Including Statistics of Measurement and Sampling. American Society of Agronomy, Madison, pp. 671–698. https://doi.org/10.2134/agronmonogr9.1.c49.

Xu, H., Xu, D., Chen, S., Ma, W., Shi, Z., 2020. Rapid determination of soil class based on visible-near infrared, mid-infrared spectroscopy and data fusion. Remote Sens. 12 (9), 1512. https://doi.org/10.3390/rs12091512.

Zhu, Y., Weindorf, D.C., 2009. Determination of soil calcium using field portable x-ray fluorescence. Soil Sci. 174, 151–155. https://doi.org/10.1097/ SS.0b013e31819c6e1b.

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. https://doi.org/10.1016/j.geoderma.2011.08.010.

Appendix

Table A.1. Removal effects of Sodium citrate-bicarbonate dithionite, Ammonium oxalate and Hydrogen peroxide on forms of iron, aluminum and silicon of soils.

Extraction	Sodium citrate-dithionite-bicarbonate	Acid ammonium oxalate	Hydrogen peroxide
Iron	Well crystalline, poorly crystalline, and non-crystalline oxide minerals, such as hematite, goethite, maghemite, lepidocrocite and ferrihydrite (Schwertmann and Carlson, 1994; dos Anjos et al., 1995; Trolard et al., 1995).	Poorly crystalline iron oxide minerals, such as ferrihydrite and small amounts of organically bound iron (Schwertmann, 1959, 1964; Cornell and Schwertmann, 2003).	a) Removal of macro and micro nutrients present or adsorbed to soil organic matter. b) Calcium oxalate formation in some clayey soil (Brown, 1953; Farmer and Mitchell, 1963) c) Poorly crystalline minerals are subject to attack or alteration (Mitchell et al., 1964). d) Formation of water soluble chelated oxalates of aluminum and iron, and water insoluble chelated oxalates, on the surface of clay minerals. (Farmer and Mitchell, 1963). e) Dissolution of sesquioxides (Harada and Inoko, 1977). f) Partial dissolution of MnO_2 (Jackson, 1969). g) Exfoliation of mica (Drosdoff and Miles, 1938).
Aluminum and Manganese	Elements that replace iron in the structure of iron oxide minerals. For aluminum, it can be affected by the partial dissolution of poorly crystalline aluminum oxides during extractions of kaolinite and gibbsite (Fitzpatrick and Schwertmann, 1982; Schwertmann and Taylor, 1989).	Poorly crystalline forms aluminum complexed with soil organic matter (McKeague et al., 1971; Parfitt and Henmi, 1982).	
Silicon	The removal of $Si(OH)_4$ is attributed to the lower silica retaining soil remaining capacity of the minerals (Weaver et al., 1968).	Silicon from poorly crystalline soil minerals (Hallmark et al., 1982).	

Table A.2. Precision average for the measured samples

Element	Precision (%) *
Mg	± 0.3504
Al	± 0.1500
Si	± 0.0921
Ti	± 0.0251
Mn	± 0.0028
Fe	± 0.0466
Cu	± 0.0008
Zn	± 0.0005

*Average for all samples with contents above the limit of quantification informed by equipment

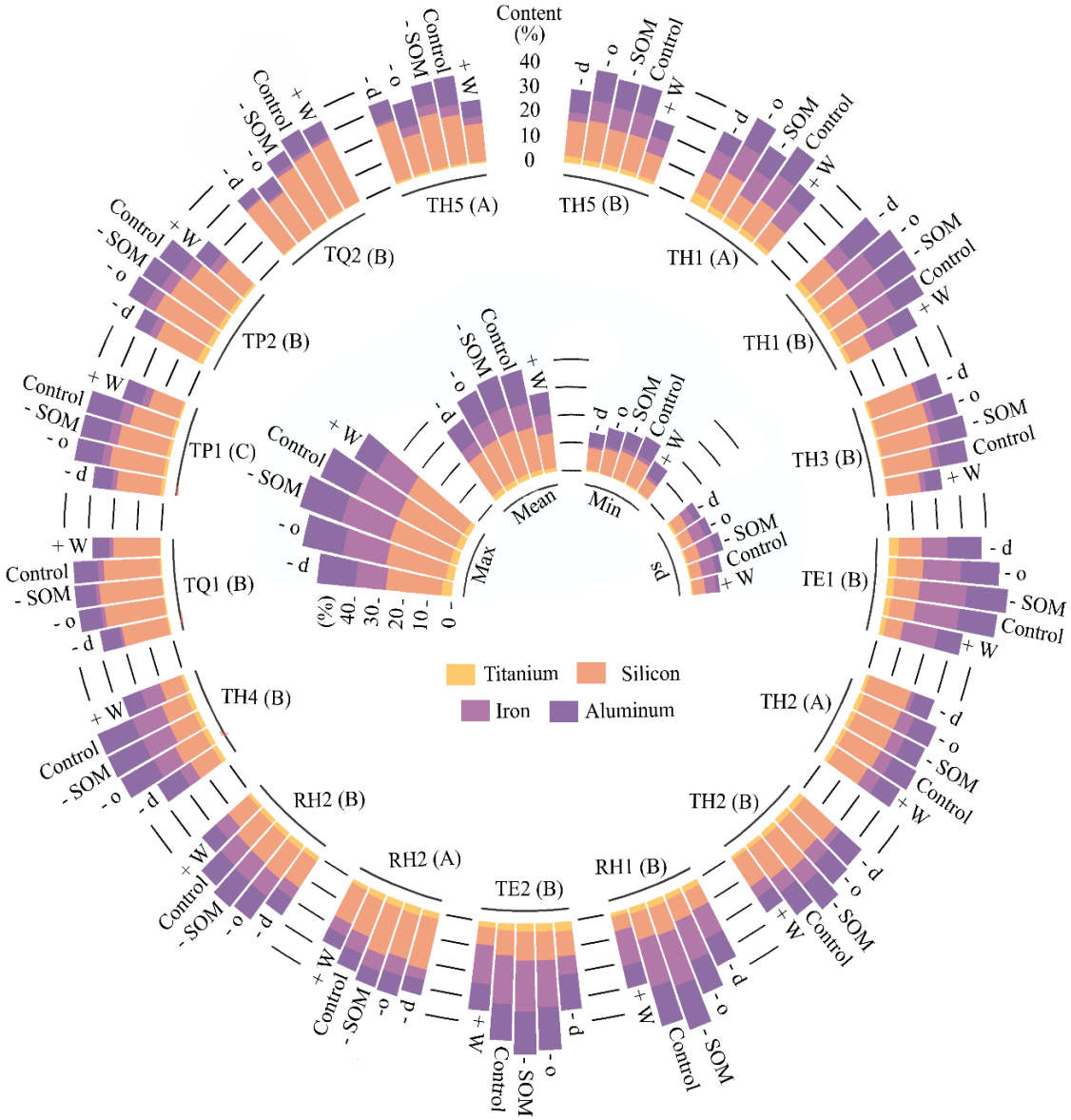

Fig. A.1. Titanium, silicon, iron and aluminum total contents determined by XRF Delta Olympus internal model for the treatments. -SOM = Organic matter removal; -o: SOM and poorly crystalline iron forms removal; -d: OM and poorly crystalline plus free crystalline iron forms removal; +W: Water addition; Max = maximum; Min = minimum; SD = standard deviation; TH = Typic Hapludox; RH = Rhodic Hapludox; TE = Typic Eutrustox; TP Typic Paleudalf; TQ = Typic Quartzpament.

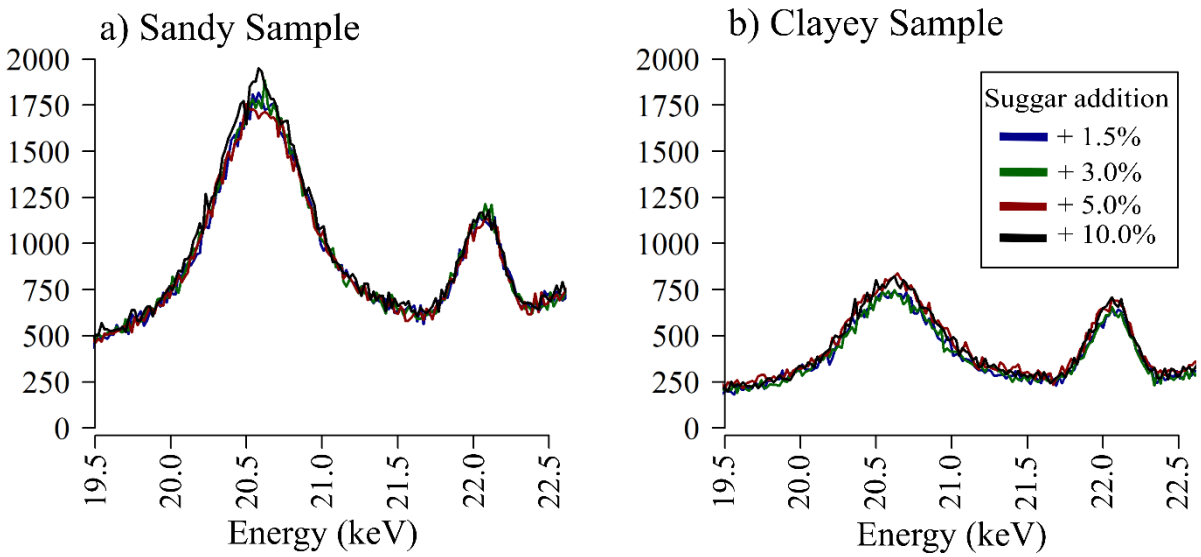

Fig. A.2. Compton and Thomson behavior with sugar addition for sandy (a) and clayey (b) sample. The sandy sample had 50 g kg^{-1} of clay and 1 g kg^{-1} of soil organic matter (SOM), while the clayey sample had 739 g kg^{-1} of clay and 34 g kg^{-1} of SOM.

3. MAPPING BRAZILIAN SOIL MINERALOGY USING PROXIMAL AND REMOTE SENSING DATA

ABSTRACT

The mineralogy is the gear of soil processes, playing a fundamental role in relevant issues for humanity, such as food and fuel supply, and climate and water regulating. Nevertheless, the mineralogical analyses are few used for soil characterization, due the difficulty of acquisition data through traditional methods and low commercial availability. On the other hand, the abundance of soil minerals can be accessed by soil spectroscopy (Vis-NIR-SWIR) and spatialized by digital soil mapping frameworks. This work aims to map the abundances of major soil mineralogical components for the whole Brazilian territory at the surface and subsurface. The Brazilian Spectral Library (BSSL) with Vis-NIR-SWIR spectral data, was used to assess the relative amounts of hematite (Hem), goethite (Gt), kaolinite (Kt) and gibbsite (Gbs) in soil samples from Brazil. Terrain attributes (TA) and a synthetic soil image (SySI) with bare soil pixel from multitemporal Landsat images (1984 to 2020) were used as predictors. A novel approach was performed in order to obtain a bare soil image for the whole Brazilian territory. The model Random Forest (RF) was used for spatial prediction to obtain the mineral maps and their uncertainty by bootstrapping procedure. The mineral maps were compared with geology and soil legacy maps and also climate and terrain conditions at national, regional and farm levels. The major iron and aluminum oxide minerals and Kt were successfully identified and its abundances quantified by Vis-NIR-SWIR reflectance from BSSL. The Hem presented the more accurate results in RF models with R^2 ranging from 0.48 to 0.56, followed by Gbs (0.42 to 0.44), Kt (0.20 to 0.31) and Gt (0.16 to 0.26). The proposed approach was able to reveal the spatial distribution of the relative abundance of minerals for the Brazilian territory. The mineral maps were in accordance with geology and soil legacy maps and also with the climate and terrain conditions. The approach proposed for this paper is an efficient method to obtain mineralogy information for large areas.

Keywords: Remote Sensing; soil spectroscopy; chemometrics; tropical soils.

3.1 Introduction

The mineralogy is the gear of soil processes, playing a fundamental role in several relevant issues for humanity, such as food and fuel production and regulation of climate and water. It influences the soil fertility, plant growth, CO_2 sequestration, climate changes mitigation, dynamics of pesticides, and contaminants and land degradation (Viscarra Rossel, 2011). Clay minerals and sesquioxides, such as hematite (Hem), goethite (Gt), kaolinite (Kt) and gibbsite (Gbs) are related to cation exchange capacity, aggregation and structure, organic matter dynamic, water retention, adsorption of phosphorus, among others (de Oliveira et al., 2020; Gilkes and Prakongkep, 2016; Hassink, 1997; Heuvelink et al., 2021; Santos et al., 2017; Zhao et al., 2017). The soil mineralogy is also key to soil genesis and geochemistry processes (Macías and Camps-Arbestain, 2020).

Soil is the result of the interaction of the forming factors: climate, organisms, relief, parent material and time (Jenny, 1941). Several factors command the processes of soil formation and determine the soil mineralogy composition (Heimsath et al., 2012). The extensive Brazilian territory has great geological (Gómez et al., 2019), climate (Alvares et al., 2013), topography and biodiversity (Gomes et al., 2019) variability that results in innumerous soil types, with distinct properties, including mineralogy (Instituto Brasiliero de Geografia e Estatística, 2021; Santos et al., 2018). The maps of soil classes available for Brazilian territory is at inappropriate scales or restricted for some regions (Lepsch, 2013; Nolasco de Carvalho et al., 2015). For example, in the north region, where the Amazon rainforest is located, soil maps with semi-detailed scale are almost non-existent (Mendonça-Santos and dos Santos, 2006). Mineralogical maps covering all Brazilian territory are still non-existent.

Traditional methods to mineralogical characterization by X-ray diffraction (XRD) and X-ray powder diffraction (XRPD) are laborious, expensive, time consuming, and only provide qualitative or semi quantitative information (Chipera and Bish, 2001; De Souza Bahia et al., 2015; Fang et al., 2018; Kunze and Dixon, 1986; McManus, 1991; Viscarra Rossel, 2011; Whittig, 1965) Consequently, the mineralogical evaluation historically was restricted to scientific research level, being less available to other users, such as farmers. Additionally, these analyzes are not commercially available and are falling into disuse. Obtaining mineralogical data for large territorial extensions is even more challenging. The traditional techniques, due the costs of data acquisition, are not applicable to obtain data on a large scale (Mendes et al., 2021) and because it, the available information is punctual or limited to small areas (Viscarra Rossel, 2011).

In this context, an alternative method for soil mineralogy assessment becomes necessary (Fang et al., 2018; Viscarra Rossel, 2011). The proximal sensing techniques, specifically the diffuse reflectance spectroscopy in the visible, near infrared and shortwave infrared (Vis-NIR-SWIR) ranges provide a rapid, non-destructive and environmentally friendly method for soil characterization (Nocita et al., 2015; Soriano-Disla et al., 2014). The 350-2500 nm range is used for estimation of several soil attributes (Soriano-Disla et al., 2014; Viscarra Rossel et al., 2006), such as soil particle size distribution (Coblinski et al., 2020; Silva et al., 2019), soil organic carbon (Barthès et al., 2019; Moura-Bueno et al., 2020) and soil chemical attributes (Vaudour et al., 2018; Wadoux et al., 2019; Zhao et al., 2021). According to Bahia et al. (2015), while it takes more than 55 hours per sample to obtain the iron oxide minerals contents by conventional method, with Vis-NIR-SWIR spectroscopy, it can be reached in 20 minutes.

The Vis-NIR-SWIR range has specific wavelengths related to the electronic and vibrational transition caused by the interaction of energy with minerals in the soils (Fang et al., 2018). The use of spectral pre-processing, such as the Kubelka-Munk function combined with the second derivative by Savitzy-Golay method can be used to enhance these spectral features (Barrón and Torrent, 1986; Scheinost, 1998) and improve their estimations. In general, the intensity of the band amplitude calculated from derivative values between minima and maxima at specific absorption feature is proportional to the mineral abundance in the sample (Kosmas et al., 1984; Mendes et al., 2021).

Several studies focused on the understanding of the fundamentals of the soil mineralogy influence in Vis-NIR-SWIR spectra (Barrón and Torrent, 1986; Clark et al., 1990; Madeira-Neto et al., 1995; Scheinost, 1998). Additionally, some researchers developed equations and models to quantify the minerals content based on amplitude from this range and XRD data (Canton et al., 2021; De Souza Bahia et al., 2015; Fernandes et al., 2020; Madeira-Neto et al., 1995; Mendes et al., 2021; Silva et al., 2020). Finally, spatial predictions were carried out for mineral amplitudes/indexes (Fernandes et al., 2020; Mendes et al., 2021; Poppiel et al., 2020; Ramos et al., 2020; Viscarra Rossel et al., 2010; Viscarra Rossel, 2011) or mineral contents (Silva et al., 2020). Thus, spectroscopy becomes an important tool to obtain quantitative and spatial information of soil mineralogy.

Using proximal sensing and geostatistic techniques, such as kriging, the relative abundance of iron oxides (Hem and Gt) in the soil from Vis-NIR-SWIR were accessed and mapped for all Australia territory (Viscarra Rossel et al., 2010). Environmental covariables representing the soil formation factors were used to digitally map Kt, illite and smectite in Australian territory (Viscarra Rossel, 2011). For the west of São Paulo State region in Brazil, Silva et al., (2020) mapped the content of Hem and Gt and Fernandes et al. (2020) the Kt/(Kt+Gbs) ratio. Meanwhile, Ramos et al. (2020) mapped the Hem/(Hem+Gt) ratio for the Rio Grande do Sul

State. Environmental covariates, such as terrain attributes and synthetic soil image (SySI) with bare soil pixels, from multi-temporal satellite data, also can be used as predictors for soil mineralogy mapping (Mendes et al., 2021; Poppiel et al., 2020). Poppiel et al. (2020) mapped the relative abundance of Hem, Gt, Kt, Gbs and 2:1 minerals for Goiás State in Brazil. Moreover, Mendes et al. (2021) mapped Hem, Gt, Kt, Gbs and several 2:1 minerals at regional scale in Brazil.

The mineralogy has been neglected in analyses for soil characterization and its impact on several agricultural and environmental issues (Viscarra Rossel, 2011). Brazil, as a country with an important agribusiness sector and considered the center of several environmental issues at global level, information on soil mineralogy spatial variability needs to be better known. And thus, this study aims to map the abundances of major soil mineralogical components for the whole Brazilian territory at the surface and subsurface. The abundances of soil minerals such as Hem, Gt, Kt and Gbs can be identified by Vis-NIR-SWIR spectra (Fang et al., 2018) and that the interaction of electromagnetic energy with soil matrix fundamentals are the same at laboratory and satellite levels, respecting the proportions of scale (Bellinaso et al., 2021). The hypothesis is that the soil mineral abundances can be accessed by punctual information from Vis-NIR-SWIR spectra and spatialized by SySI and terrain attributes derived from remote sensed data.

3.2 Materials and Methods

3.2.1 Study area and soil spectral library

The study area covered Brazilian territory, with approximately 8.8 million of km^2 (Fig. 1), comprising a high diversity of climates, geologies and soil types. The climate is divided into tropical (81%), dry (5%) and subtropical (14%) zones, according to the Koppen classification system (Alvares et al., 2013) (Fig. 1). The geology is complex with several types of metamorphic igneous and sedimentary rocks (Gómez et al., 2019) (Fig. A1a). The main soil classes of the country are Latossolos (Ferralsols) and Argissolos (Acrisols/Lixisols/Alisols), which comprises more than 60% of territory (Instituto Brasiliero de Geografia e Estatística, 2021) (Fig. A1b). Brazil also has a great biodiversity with six distinct biomas, being that the Amazonia (49.29%) and Cerrado (22%) cover most part of the territory (Gomes et al., 2019).

A legacy database with spectral data in Vis-NIR-SWIR range (350 to 2500 nm) from The Brazilian Soil Spectral Library (BSSL) (Demattê et al., 2019), accessible at https://besbbr.com.br, was used in this study. The BSSL was constructed with the support of more than 40 Brazilian researchers and has more than 50,000 samples with spectral data in the Vis-NIR-SWIR range. However, only the samples with exact coordinates obtained by a Global System Position (GPS) equipment were used in this study, totalizing 30,334 soil samples (Fig. 1). The major part of BSSL observations were from 0-0.2 (10,306), 0.4-0.6 (7,676) and 0.8-1m (7,697) layers and the remaining (7,264) from other depths or soil profiles with irregular sampling.

For spectral analysis, the samples were dried, ground, sieved with 2 mm mesh and distributed in a Petri dish. A FieldSpec 3 spectroradiometer (Analytical Spectral Devices, Boulder, CO) was used to obtain the reflectance of soil samples with resolution of 1 nm from 350 to 700 nm, 3 nm from 700 to 1,400 nm and 10 nm from 1,400 to 2,500 nm. The output data was resampled to 1 nm (totalizing 2151 channels). The light source was two 50 W halogen lamps positioned 35 cm from the samples, with zenith angle of 30º and 90º between them. The reflected energy was captured by a fiber-optic cable allocated 8 cm away from the sample. The

spectroradiometer was calibrated with a Spectralon standard every 20 min. Each spectrum was the mean from 100 readings and two replicates (changing the position of the petri dish).

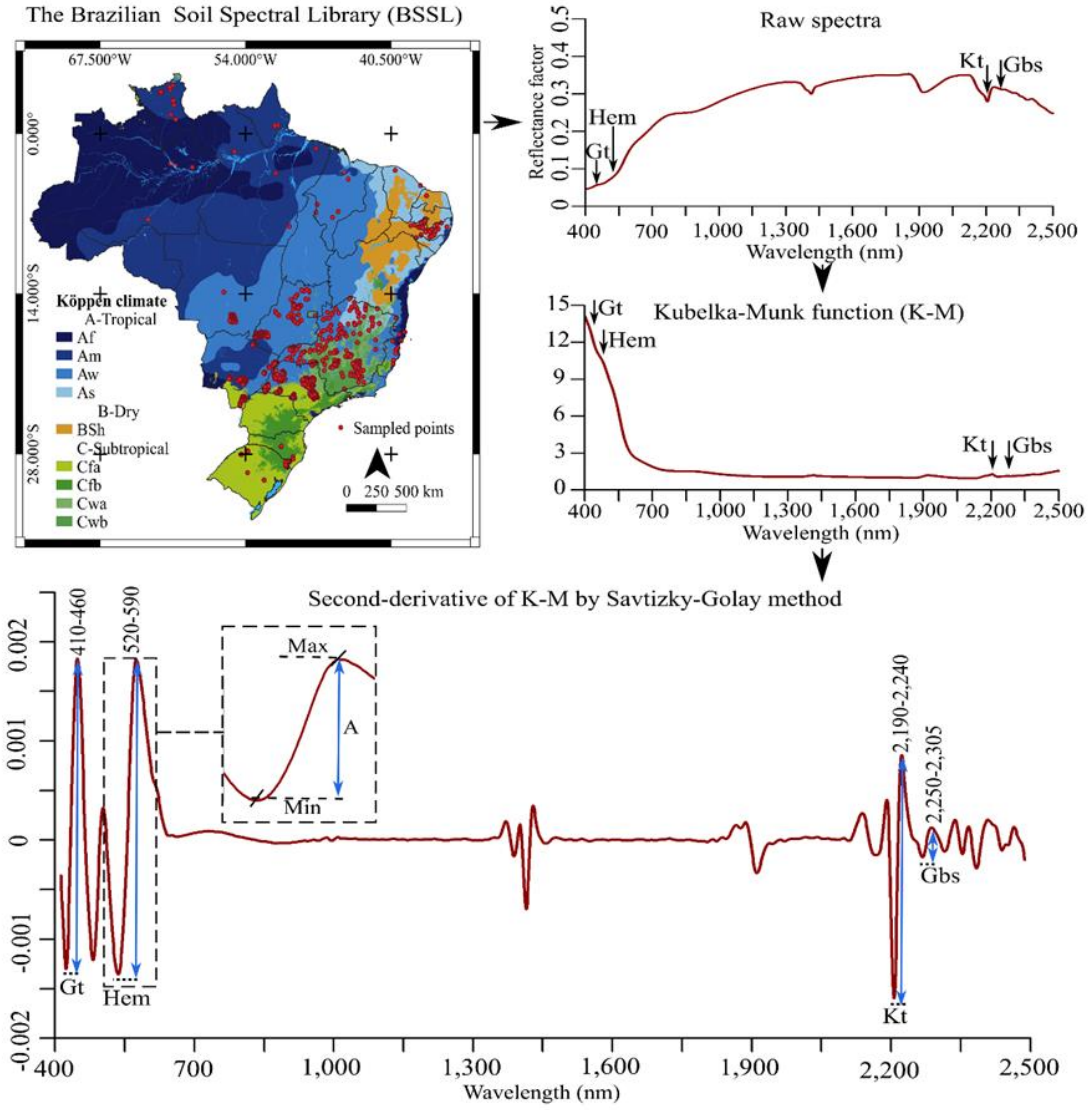

Figure 1: The Brazilian Soil Spectral Library (BSSL) points for 0-0.2m depth and flowchart of mineral amplitude calculation procedure. A= Tropical zone, without dry season (Af), monsoon (Am), dry winter (Aw) or dry summer (As); B= Dry zone with semi-arid climate and low latitude and altitude (BSh) C: Subtropical zone with oceanic climate, without dry season, with hot summer (Cfa) or with temperate summer (Cfb) or subtropical zone with dry winter and hot summer (Cwa) or temperate summer (CWB); Gt = goethite; Hem = hematite; Kt = kaolinite; Gbs = gibbsite.

3.2.2 Calculation of relative abundance of minerals

The raw reflectance spectra were transformed to absorbance using the Kubelka-Munk function (KM) [(KM = $(1-R)^2/2R$)], where R is the reflectance (Barrón and Torrent, 1986) (Fig. 1). Since the KM spectra had overlapping bands, we calculated their 2º derivative (SD) using the Savitzky-Golay method (Savitzky and Golay, 1964) to resolve and enhance the spectral features of interest (Poppiel et al., 2020). The intervals related to Gt, Hem, Kt and Gbs were defined in the SD spectra based on literature (Table 1). Thus, the amplitude of minerals

(AM) was calculated [AM = Maxλ - Minλ], where Maxλ and Minλ are the maximum (positive) and minimum (negative) values into the specific ranges, respectively.

In order to normalize the data for spatial predictions, the AGt, AHem, AKt and AGbs were submitted to depth harmonization by spline interpolation in the GSIF R package (Hengl and MacMillan, 2019). The output data was divided into five layers stratified by 0.2 m from surface up to 1m depth (0-0.2, 0.2-0.4, 0.4-0.6, 0.6-0.8 and 0.8-1m). The study, before the spline procedure, was conducted with 9,600 observation points varied from 0-0.2m layer (about 1.09 x 10^{-3} points per km^2) to 7,300 for the 0.8-1m layer (about 8.29 x 10^{-4} points per Km^2) (Table 2).

Table 1: Selected bands for amplitude of minerals (AM) calculation.

Soil mineral	Minimum band in literature (nm)	Maximum band in literature (nm)	Selected range (nm)	Reference
Goethite	~415	~455	410-460	Scheinost et al. (1998)
Hematite	~535	~580	520-590	Scheinost et al. (1998)
Kaolinite	2205	2225	2,190-2,240	Clark et al. (1990) Madeira-Neto et al. (1995)
Gibbsite	2265	2295	2,250-2,305	Clark et al. (1990) Poppiel et al. (2020)

3.2.3 Environmental covariates

3.2.3.1 Syntetich soil image (SySI) and terrain attributes (TA)

The environmental covariates used were a synthetic soil image (SySI) and terrain attributes (TA), at 30m of spatial resolution (Fig. 2). We assumed that is possible to infer about the undersurface using the soil reflectance available for surface, being that Mendes et al. (2021), highlighted the use of SySI as a predictor for AM mapping at soil depths until 1m. Thus, the SySI was used as covariate also for undersurface layers.

The SySI, was obtained using the collection of Landsat images, from 1984 to 2020 (Landsat 5, 6, 7 and 8), through the GEOS3 method (Demattê et al., 2020, 2018), in Google Earth Engine (GEE) (Gorelick et al., 2017). The GEOS3 employs multitemporal satellite images and spectral indices to select only bare soil pixels and creates a synthetic image with the median reflectance of bare soil pixels over time. The SySI is directly related to spectral data obtained at the laboratory level and was successfully used to map soil attributes (Bellinaso et al., 2021; Demattê et al., 2018; Fongaro et al., 2018; Silvero et al., 2021). The SySI had the same number of bands of the satellite used as input data for GEOS3, in this case (Landsat), it had 6 bands: blue (450–520 nm), green (520–600 nm), red (630–690 nm), NIR (760–900 nm), SWIR1 (1550–1750 nm) and SWIR2 (2080–2350 nm). The GEOS3 uses the soil spectral trend and the Normalized Difference Vegetation Index (NDVI) and Normalized Burn Ratio 2 (NBR2) to create the soil mask (Demattê et al., 2018). The bare soil pixels were selected when there was an increase in reflectance from blue to SWIR1, and NDVI and NBR2 between -0.15 to 0.25 and -0.15 to 0.15, respectively. The SySI pixels are the median of all bare soil pixels detected along the time series images.

Other than that, a total of 13 TA were derived from the digital elevation model (DEM) of Advanced Land Observing Satellite (ALOS) (Japan Aerospace Exploration Agency: ALOS Research and Application Project, 2021), using the package terrain analysis in Google Earth Engine (TAGEE) (Safanelli et al., 2020a).

The SySI and TA values were extracted for the observation points, in a spreadsheet containing the AM values. The spatial predictors were selected based on the Spearman correlation with AM. The correlation analyses were performed for all depth combinations in conjunction (without stratification by layer). The criteria for selection of covariable for modelling was correlation > 0.10 or < -0.10. In this way, the covariates selected were all bands from SySI: blue, red, green, NIR, SWIR1 and SWIR2 and four TA from DEM: elevation, slope, maximum curvature and hillshade. The results of Spearman correlation analyses can be found in Table A1.

3.2.3.2 Synthetic soil image with full coverage (SySIc)

The GEOS3 showed to be a powerful tool for spatial prediction of soil attributes (Poppiel et al., 2021, 2020; Safanelli et al. 2020b, 2021ab). However, the method is useful only in locations with naturally exposed soil or exposed by anthropic activities (i.e. soil tillage). For the Brazilian territory, which have vast areas with native vegetation (i.e. Amazon rainforest), the SySI only covers about 30% of the total area. In this way, in order to obtain continuous maps of soil mineralogy for the entire Brazilian territory, a synthetic soil image with full coverage (SySIc) was created using environmental covariates and machine learning (Fig. 2).

The TA selected were the same as mentioned in section 3.2.3.1 (Elevation, Slope, Maximal Curvature and Hillshade). Besides that, the bands blue, green, red, NIR, SWIR1 and SWIR2 obtained from a natural vegetation mosaic produced with the mean of Landsat images of the year 1984 also was used as covariates. It was used an available mosaic generated for Hengl et al. (2018), obtained in the GEE, with the spatial resolution of 1km and after resampled to 30m.

A random sampling was realized in SySI, TA and natural vegetation mosaic, limited by the extent of SySI and totaling 85.740 points in Brazilian territory. The correlation between the covariates (vegetation mosaic and TA) and SySI was obtained as topic 3.2.3.1 and are also available in table A.1. The modelling was performed using each SySI band as dependent variable and the mosaic bands and TA as independent variables, totalizing six models. The modelling procedure and model evaluation (Table A.2) were performed, as mentioned in section 3.2.4. The bootstrap predictions were performed for each band as mentioned below (section 3.2.6).

Finally, the SySI original bands were overlapped to the six bare soil predicted bands and the merged were united in an unique raster file, originating the SYSIc. In this way, in the final image, the pixels with soil exposition during the time remained with the original information, while the “gaps” were filled with the predicted bands. This procedure aimed to obtain an image with full coverage to use as covariates for minerals mapping. Besides that, we preserved the real reflectance of pixels with bare soil detected by GEOS3, instead of the use of only predicted reflectance, because, the prediction processes had uncertainties associated that can be propagated to the AM prediction.

3.2.4 Prediction models for soil minerals

The selected environmental covariates were used as independent variables to obtain prediction models for AGt, AHem, AKt and AGbs using the Random Forest algorithm (RF) (Breiman, 2001). We used the scikit-learn algorithm in Python (Pedregosa et al., 2011) to implement a bootstrapping procedure for training and testing the RF algorithm, as described by Safanelli et al. (2021ab). The calibration set was composed of bootstrapped samples, while the other samples, not used for calibration, were used exclusively for validation.

The number of samples used in calibration were 63% of the total dataset, according to Raschka (2018). The bootstrapping was performed up to 500 times, which is similar to each forest size (FS) tested.

A grid search procedure was performed in order to select the best hyperparameters combination, looking to reduce the overfitting possibilities. It was tested diferentes values for FS (30, 60, 100, 200 and 500 trees), number of predictors in tree splits (nRP) (3, 5, 8 and 10) and minimum number of observations at leaves (minSL) (10, 20, 30, 40, 50, 100, 200 and 500). The optimal combination (FS, nPR and minSL) with the least root mean square error (RMSE) for the validation set was selected for each mineral. However, when the RMSE difference between the hyperparameters combination were less than $1x10^{-5}$, the combination with less FS value was chosen, to make the spatial prediction faster, in view of the largest territory to map and the great influence of FS in processing time.

The accuracy of developed models were evaluated by coefficient of determination (R^2), RMSE, ratio of the performance to interquartile distance (RPIQ = (Q3-Q1)/RMSE), where Q1 and Q2 are 1º and 3º quartiles.

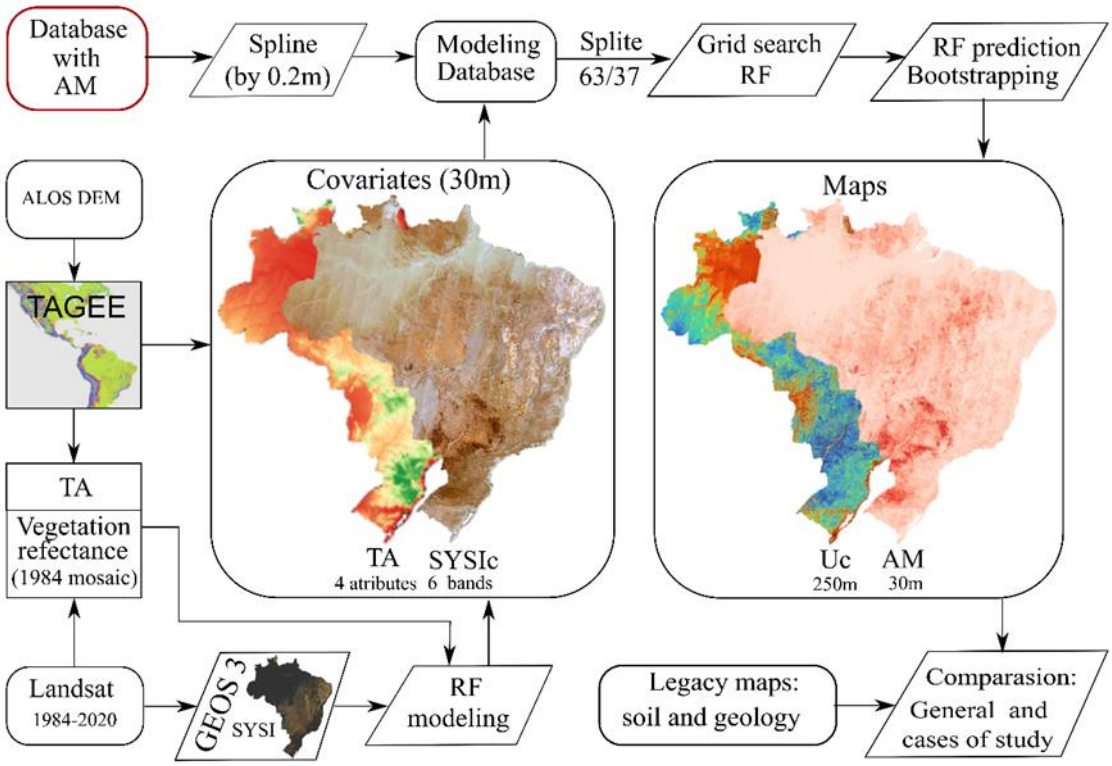

Figure 2: Flowchart of spatialization procedures. AM = amplitude of minerals; DEM = Digital elevation model; TA= Terrain attributes; SySI = synthetic soil image; SySIc = synthetic soil image full coverage; Uc = uncertainty; RF = Random Forest.

3.2.5 Spatial prediction

The spatial prediction was performed by bootstrapping routine (Efron and Tibshirani, 1993) in GEE (Gorelick et al., 2017), in order to obtain the maps of AM and their uncertainties. The models were fitted with TA and SySI (to increase the accuracy) and the spatial prediction was performed using TA and SySIc, in order to obtain mineral maps covering all the territory. This step was more demanding in time for processing, thus, the number of bootstraps was fixed in 50.

The final AM maps were obtained by the mean of 50 bootstrap predictions, with 30m of spatial resolution. The uncertainty maps were obtained by the coefficient of variation (CV= SD/mean×100), where SD is the standard deviation, with spatial resolution of 250m. The CV was chosen because permits the comparation between attributes with different magnitudes. The pixels with high CV values had more uncertain predictions, while the pixels with low CV values had less uncertain predictions (Poppiel et al., 2021).

3.2.6 Interpretation, indexes and spatial validation

For the description of AM variation in Brazilian territory, the predicted maps were stratified in tree equal percentiles: 1) less occurrence = AM value into the inferior percentile; 2) moderate occurrence = sample with AM value into the intermediary percentile; 3) high occurrence = sample with AM value into the superior percentile. The limits of percentiles varied for each mineral due the differences of magnitude. The uncertainty maps were also stratified in three classes: 1) CV equal or less than 12%; 2) CV value between 12 and 24%; 3) CV value equal or higher than 24%.

Indexes are calculated from the AM maps in order to demonstrate where each mineral is predominant. The AM maps were normalized between 0 and 100 (AM_n) using the equation: $AM_n = (AMx100)/AMmax$, where AMmax is the higher AM of all samples for each mineral. After that, the A $Hem_n/(Hem_n+Gt_n)$ and $Kt_n/(Kt_n+Gbs_n)$ ratio maps were calculated.

In order to evaluate the distribution of iron oxides (Hem and Gt), Kt and Gbs, the AM_n maps passed by a second normalization procedure (AM_{n2}), using the following equations: $Hem+Gt_{n2} = (Hm_n+Gt_n)/(Hm_n+Gt_n+Kt_n+Gbs_n)$, $Kt_{n2}= Kt_n/(Hm_n+Gt_n+Kt_n+Gbs_n)$ and $Gbs_{n2}= Gbs_n/(Hm_n+Gt_n+Kt_n+Gbs_n)$. Finally, a spatial ternary plot was made, with red band = iron oxides ($Hem_{n2}+Gt_{n2}$), green = Kt_{n2} and blue = Gbs_{n2}.

For interpretation of results and spatial validation the AM and indexes maps were visually compared with 1.5.000.000 geologic (Gómez et al., 2019) and 1:1.000.000 soil (Instituto Brasiliero de Geografia e Estatística, 2021) maps and also with climate (Alvares et al., 2013), elevation and slope (Japan Aerospace Exploration Agency: ALOS Research and Application Project, 2021) patterns. The 1;1,000.000 soil map is derived from the RADAMBRASIL project and has a cartographic base of 1.250.000, however the soil information is compatible with the 1.1.000.000 scale (Instituto Brasiliero de Geografia e Estatística, 2021).

The following case studies also were carried out followed by visual comparison of: 1) 1.5.000.000 geologic (Gómez et al., 2019) and 1:1.000.000 soil (Instituto Brasiliero de Geografia e Estatística, 2021) maps, AHem, AGt and AHem/(AHem+AGt) ratio maps for Rio Grande do Sul (RS), Santa Catarina (SC), Paraná (PR), São Paulo (SP) and Mato Grosso do Sul (MS) States, with zoom in the west of PR. 2) 1:1.000.000 soil map (Instituto Brasiliero de Geografia e Estatística, 2021), AHem and AHem/(AHem+AGt) ratio maps for the west of Pará State (PA). 3) 1:1.000.000 soil map (RADAMBRASIL, 1983), AHem and AGbs maps for the shout of Goiás State (GO). 4) 0-0.2m layer and 0.8-1m layer of AHem and AKt for west of São Paulo State (SP), with zoom at farm level. 5) 0-0.2m layer and 0.8-1m layer of AGt for the southwest of Amazonas State (AM) and of AHem for south of GO (farm level). 6) 1.100.000 soil map (adpted from Oliveira and Prado, 1989), AHem and AGbs for Piracicaba-SP region. 7) 1.10.000 soil map (Souza, 2020), AHem and AGbs maps at farm level in Piracicaba-SP. 8) 1.5.000 soil map (Demattê et al., 2004) and AHem at farm level in Piracicaba-SP.

3.3. Results

3.3.1 Descriptive statistics of relative abundance of minerals

The AM, which reflects the relative abundance of minerals in soil samples, increased with depth. The iron oxides, AHem and AGt showed mean values of 360, 411, 439, 454, 463 ($x10^{-6}$) and 527, 575, 605, 670, 615 ($x10^{-6}$) for 0.0-0.2, 0.2-0.4, 0.4-0.6, 0.6-0.8 and 0.8-1.0m, respectively. For the same depths, AKt and AGbs showed mean values of 196 ,240, 274, 291 and 301 ($x10^{-6}$) and 42, 46, 50, 50 and 51 ($x10^{-6}$), respectively. The AGbs values showed stability above 0.40m while a more pronounced increase in deep layers was verified for AKt, followed by AGt.

The AM showed pronounced variations for the sampled points (Table 2). The highest variabilities were verified for AGbs, with CV ranging from 114.80 to 127.83%. The iron oxides showed intermediary variability between the minerals, with CV of AHem and AGt ranging from 65.75 to 71.28% and from 67.44 to 82.04%, respectively. The AKt showed the least variations with CV ranging from 46.54 to 60.13%. The high CV were verified for 0.0-0.2m depth, with AM values showing less variations in deep layers.

Table 2: Descriptive statistics for amplitude or minerals (AM) and Random Forest (RF) hyperparameters and accuracy.

Mineral Amplitude	Layer (m)	Descriptive statistic						RF hyperparameters			Calibration set results			Validation set results		
		n	Min ($x10^{-6}$)	Max ($x10^{-6}$)	Mean ($x10^{-6}$)	SD ($x10^{-6}$)	CV (%)	FS	nPR	minSL	R^2	RMSE ($x10^{-6}$)	RPIQ	R^2	RMSE ($x10^{-6}$)	RPIQ
Hematite (AHem)	0.0 - 0.2	9,645	3.4	1246	360	257	71.28	100	10	100	0.62	158	2.55	0.56	171	2.34
	0.2 – 0.4	8,216	2.2	1247	411	260	63.41	30	7	100	0.62	161	2.60	0.56	173	2.41
	0.4 – 0.6	8,011	3.5	1191	439	267	60.91	60	10	100	0.59	171	2.51	0.52	186	2.31
	0.6 – 0.8	7,411	1.4	1184	454	259	57.09	30	10	100	0.59	166	2.51	0.51	180	2.31
	0.8 – 1.0	7,371	2.5	1188	463	263	56.75	30	10	100	0.56	174	2.40	0.48	189	2.23
Goethite (AGt)	0.0 - 0.2	9,644	7.5	2975	527	433	82.04	30	10	200	0.31	361	1.13	0.25	372	1.10
	0.2 – 0.4	8,220	7.8	2874	575	425	73.99	200	10	200	0.32	351	1.24	0.26	366	1.19
	0.4 – 0.6	8,015	5.6	2987	604	452	74.77	100	10	200	0.27	388	1.15	0.20	403	1.10
	0.6 – 0.8	7,408	8.2	2951	610	411	67.44	60	3	200	0.25	355	1.19	0.20	368	1.16
	0.8 – 1.0	7,368	0.0	2980	615	439	71.43	60	10	200	0.22	387	1.08	0.16	403	1.03
Kaolinite (AKt)	0.0 - 0.2	9,645	3.2	1194	196	118	60.13	30	10	200	0.34	95	1.53	0.29	100	1.47
	0.2 – 0.4	8,216	6.2	1133	240	127	52.9	30	10	200	0.37	101	1.57	0.31	105	1.50
	0.4 – 0.6	8,011	4.9	1363	274	141	51.68	30	10	200	0.32	117	1.46	0.25	122	1.40
	0.6 – 0.8	7,411	0.3	1183	291	135	46.54	200	5	200	0.30	113	1.43	0.25	118	1.37
	0.8 – 1.0	7,731	3.4	1197	301	140	46.66	30	10	500	0.22	124	1.34	0.20	125	1.32
Gibbsite (AGbs)	0.0 - 0.2	9,567	0.0	300	42	54	127.83	30	7	50	0.56	36	1.16	0.42	41	1.03
	0.2 – 0.4	8,074	0.0	300	46	55	120.48	30	10	100	0.51	39	1.20	0.43	42	1.11
	0.4 – 0.6	7,868	0.4	300	50	60	120.32	30	10	100	0.51	42	1.20	0.43	46	1.11
	0.6 – 0.8	7,224	0.3	298	50	57	114.8	30	10	50	0.57	37	1.35	0.44	43	1.18
	0.8 – 1.0	7,184	0.7	298	51	60	116.9	60	10	50	0.56	40	1.30	0.41	45	1.12

Where: Min = minimum value; Max= maximum value; SD = standard deviation; CV = coefficient of variation; RF = Random Forest; FS = Forest size; nPR = number of predictors; minSL = minimum samples to leaf; R^2 = coefficient of determination; RMSE = root mean square error; RPIQ = ratio of the performance to interquartile distance.

3.3.2 Performance of AM models

The modelling procedure reached satisfactory results for predicting the relative abundance of soil minerals. The best results were found in superficial layers, decreasing the accuracy with increasing depth (Table 2). The more accurate results were reached for AHm with R^2 for validation set varying between 0.48 and 0.56. The RMSE ranged from $171x10^{-6}$ to $189x10^{-6}$ and RPIQ from 2.23 to 2.34. The AGb reached R^2 values ranging between 0.41 and 0.44, with RMSE ranging from $41x10^{-6}$ to $46x10^{-6}$ and RPIQ from 1.03 to 1.18 in the validation set. The less accurate results were found for AKt and AGt with R^2 values from 0.20 and 0.29 and from 0.16 to 0.25 in the validation data set, respectively. For Kt, the RMSE ranged from $100x10^{-6}$ to $125x10^{-6}$ and the RPIQ from 1.32 to 1.50, while for Gt, the RMSE ranged from $372x10^{-6}$ to $402x10^{-6}$ and RPIQ from 1.03 to 1.19.

3.3.3 Mineralogy and uncertainty maps

The maps of iron oxide minerals along the Brazilian territory revealed localized areas with higher occurrence of Hem and Gt in 0-0.2m depth, in the midwest, southwest and south, (Fig. 3). In this layer, 87, 11 and 2.1% of the area have low (equal or less than $270x10^{-6}$), moderate (between $270x10^{-6}$ and $540x10^{-6}$) and high (equal or greater than $540x10^{-6}$) AHem occurrence, respectively. In the same way, 65, 32 and 3% of area had low (equal or less than $350x10^{-6}$), moderate (between $350x10^{-6}$ and $700x10^{-6}$) and high (equal or greater than $700x10^{-6}$) AGt occurrence, respectively.

For the 0.8-1m depth, there was a generalized increase in the area with greater occurrence of Hem and Gt when compared with the superficial layer, with this increase being more pronounced for Gt. In this layer, 71, 22 and 7% of area for AHem and 1, 89 and 10% of area for AGt showed values into the mentioned ranges, respectively. Thus, almost all of the Brazilian territory has considerable occurrence of Gt in depth.

Only a small part of the territory has a high occurrence of Gbs (Fig. 3e). The distribution at 0.8-1m depth showed that the Brazilian territory has 86, 6 and 8% of the area with low (equal or less than $6x10^{-6}$), moderate (between $65x10^{-6}$ and $130x10^{-6}$) and high (equal or higher than $130x10^{-6}$) AGbs occurrence, respectively. The Kt occurs in significant amounts in Brazilian soils (Fig. 3f). The map showed that 0, 24 and 75% of the area have low (equal or less than $135x10^{-6}$), moderate (between $135x10^{-6}$ and $270x10^{-6}$) and high (equal or higher than $270x10^{-6}$) AKt occurrence, respectively. The Gbs and Kt amounts increased in the 0.8-1m layer in relation to the 0-0.2m layer. For 0-0-2m depth, 90, 10 and 0% of area have AGbs and 23, 72, 5% of area have Akt into the mentioned ranges (maps non shown).

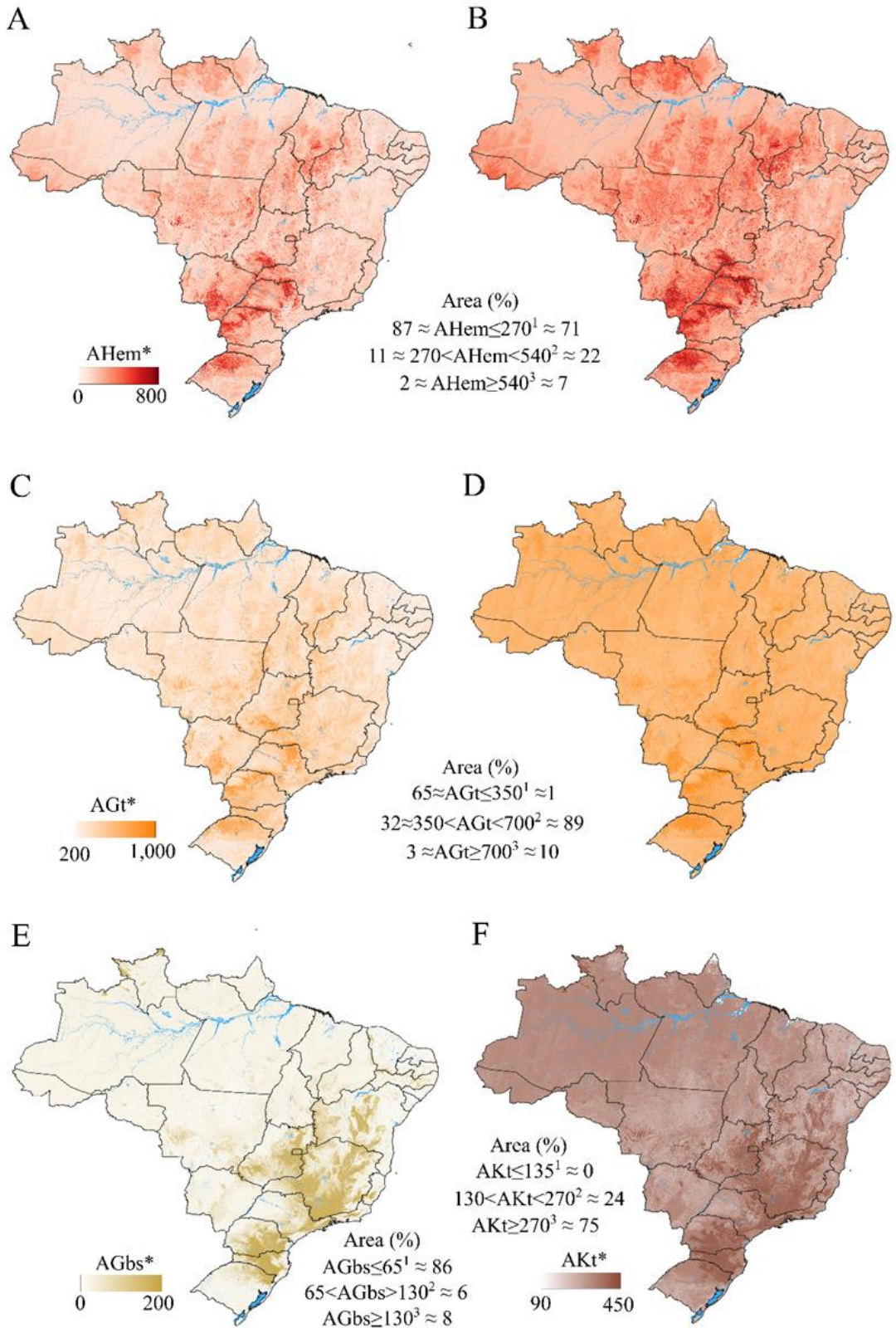

Figure 3: Maps of hematite amplitude (AHem) for 0-0.2m (A) and 0.8-1m (B), goethite amplitude (AGt) for 0-0.2m (C) and 0.8-1m (D), gibbsite amplitude (AGbs) for 0.8-1m (E) and kaolinite amplitude (AKt) for 0.8-1m (F) depths. $*10^{-6}$ scale; [1]low occurrence; [2]moderate occurrence; [3]high occurrence.

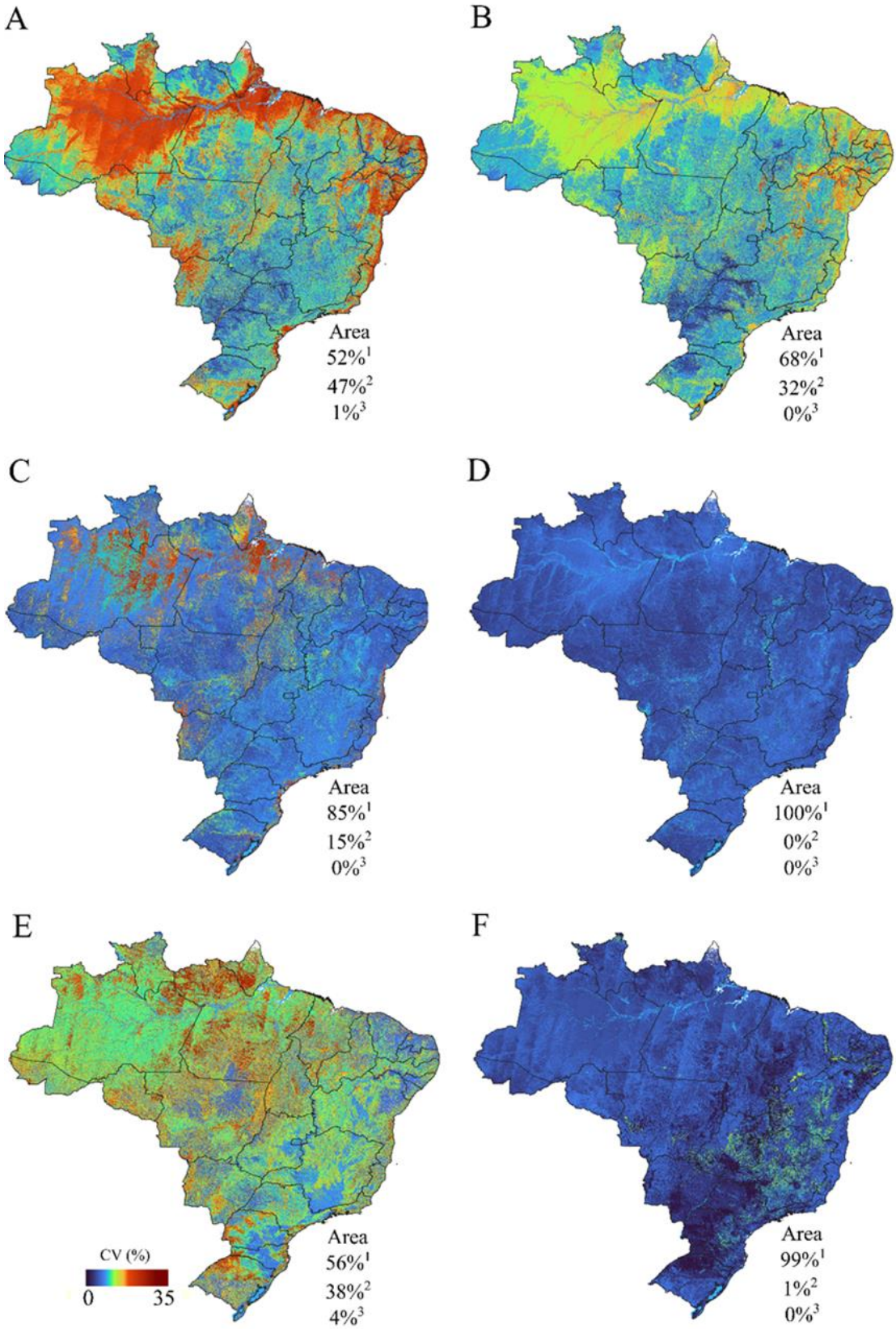

Figure 4: Uncertainty maps of hematite for 0-0.2m (A) and 0.8-1m (B), goethite for 0-0.2m (C) and 0.8-1m (D), gibbsite for 0.8-1m (E) and kaolinite for 0.8-1m (F) depths. CV= coefficient of variation. [1]CV<=12%; [2]12%<CV<24%; [3]CV=> 24%.

In general, the areas with high amounts of minerals showed the lowest coefficient of variations (CV) in the 50 bootstrap predictions (Fig. 3 and Fig. 4). The uncertainty maps for AHem showed 52, 47 and 1% of area in 0-0.2m depth with CV equal or less than 12%, between 12 and 24% and equal or higher than 24%, respectively (Fig. 4ab). For 0-8-1m depth, 68, 32 and 0% of area were into the mentioned ranges. The CV for AGt predictions were less than AHem, the layer 0-0.2m showed 85, 15 and 0% of area and the layer 0-8-1m showed 100, 0 and 0% of area with CV values into the mentioned ranges (Fig. 4cd). The AGbs and AKt maps showed 56, 38 and 4% and 99, 1, 0% of area into the mentioned ranges, respectively (Fig. 4ef).

3.3.4 Indexes

The Hem/(Hem+Gt) ratio showed predominance for Gt in Brazilian territory with similar patterns in 0-0.2m and 0.8-1m layers (Fig. 5ab). However, there were some exemptions, such as the higher values for some locations in the west of São Paulo state and lower values in the Amazon basin for 0.8-1m than 0-0.2m depths. The Kt/(Kt+Gt) ratio showed predominance of Kt in relation to Gbs in the territory, with Gbs occurrence concentrated in the south and southwest regions (Fig. 6c). The ternary plot followed the patterns of each individual mineral map (Fig. 4) and showed the predominance of Kt in Brazilian soils (Fig.4 and Fig. 6d). The occurrence of iron oxides is mainly located in some places in the south, southwest and midwest, while Gbs predominates mainly in the south and southwest regions (Fig. 6d).

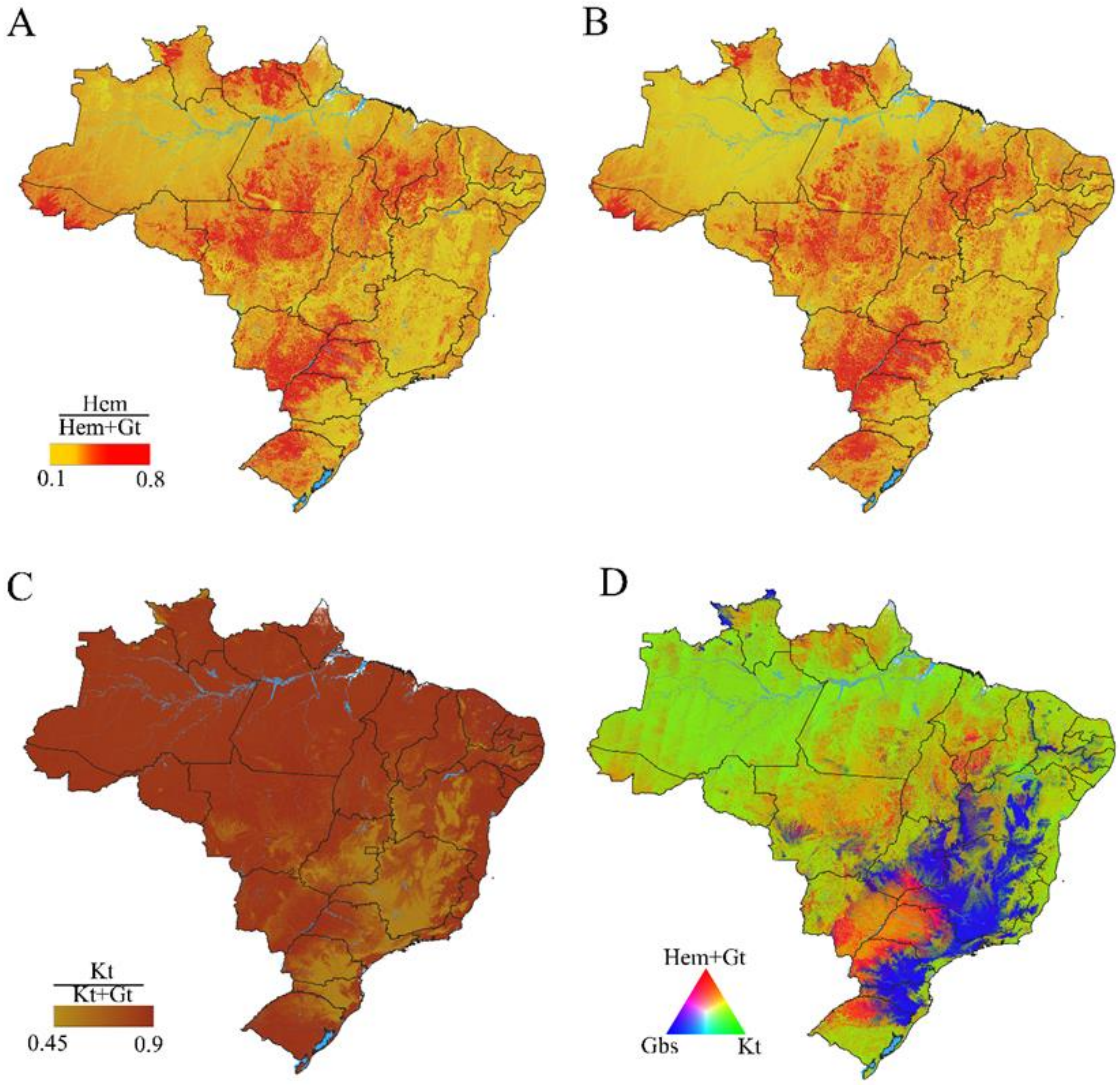

Figure 5: Hem/(Hem+Gt) ratio for 0-0.2m (A) and 0.8-1m (B) layers, Kt/(Kt+Gt) ratio (C) and ternary plot (D) for 0.8-1m layer. Hem= hematite; Gt = goethite; Kt=kaolinite; Gbs = gibbsite.

3.3.5 Study cases

The areas with higher iron oxides (Hem and Gt) amounts are related to basalt rocks and more weathered soil classes, such as red (hue equal to 2.5YR or more redish) Ferralsols (FR) and Nitisols (NT), covering part of Rio Grande do Sul (RS) (north), Paraná (PR) (west) and Mato Grosso do Sul (MS) (south) States (Fig. 6a). In the west of PR, it is possible to see the high amounts of Hem and Gt for FR and NT then Acrisols (AC), Lixisols (LX) and Alisols (AL). The west of Pará (PA) presented moderated Hem amount and high Hem/(Hem+Gt) ratio (equal or greater than 0.47) for red-yellow (hue equal to 5YR) AC, LX and FE and low Hem and Hem/(Hem+Gt) for yellow (hue equal to 7.5YR or more yellowish) AC, LC, FR and Gleysol (GL) (Fig. 6b). The south of Goiás (GO), presented moderate Hem amounts in red and clayey (clay content equal or more than 350 g kg^{-1}) FR, high in red, clayey and ferric (Fe_2O_3 equal or more of 18 g kg^{-1}) FR and low content in Arenosols (AR) and other soil classes. The Gbs amount in GO was also high in red and clayey FR, however restricted to flat and elevated areas (Fig A2cb).

The areas of FR in the west of SP showed a low increase of Hem and Kt from 0-0.2m to 0.8-1m, while the areas of AC, LX and AL presented a high increase (Fig. 7a). In the zoom at farm level also it is possible to visualize the increase of Hem and Kt deep in an area of AL. The areas of yellow FR in the south-west of Amazonas (AM) State had an significant increase in AGt amount from 0-0.2m to 0.8-1m, while for red and ferric FR the Hem amount is similar in the topsoil and 1m of depth (Fig. 7b).

In the study case in Piracicaba region, the more weathered and clayey soil classes, such as Ferralsols, showed high values of Hem and moderate values of Gbs, while the less weathered and sandy (clay content less than 350 g kg^{-1}) soil classes showed low values for both minerals (Fig. 8a). An area (farm) showed high Hem and moderated Gbs amounts for red and ferric FR and NT derived from basalt rocks, while the other soils classes from sandstone and siltstone, such as GL, Plinthosols (PT) Cambisols (CM) and Leptosols (LP), red AC and red-yellow AC, showed low values for both minerals (Fig. 8b). Another area (farm), showed high Hem amount for red and ferric FR derived from basalt and low Hem amount for red AC, red-yellow AC, CM, LP and GL from sandstones and other rocks (Fig. 8c).

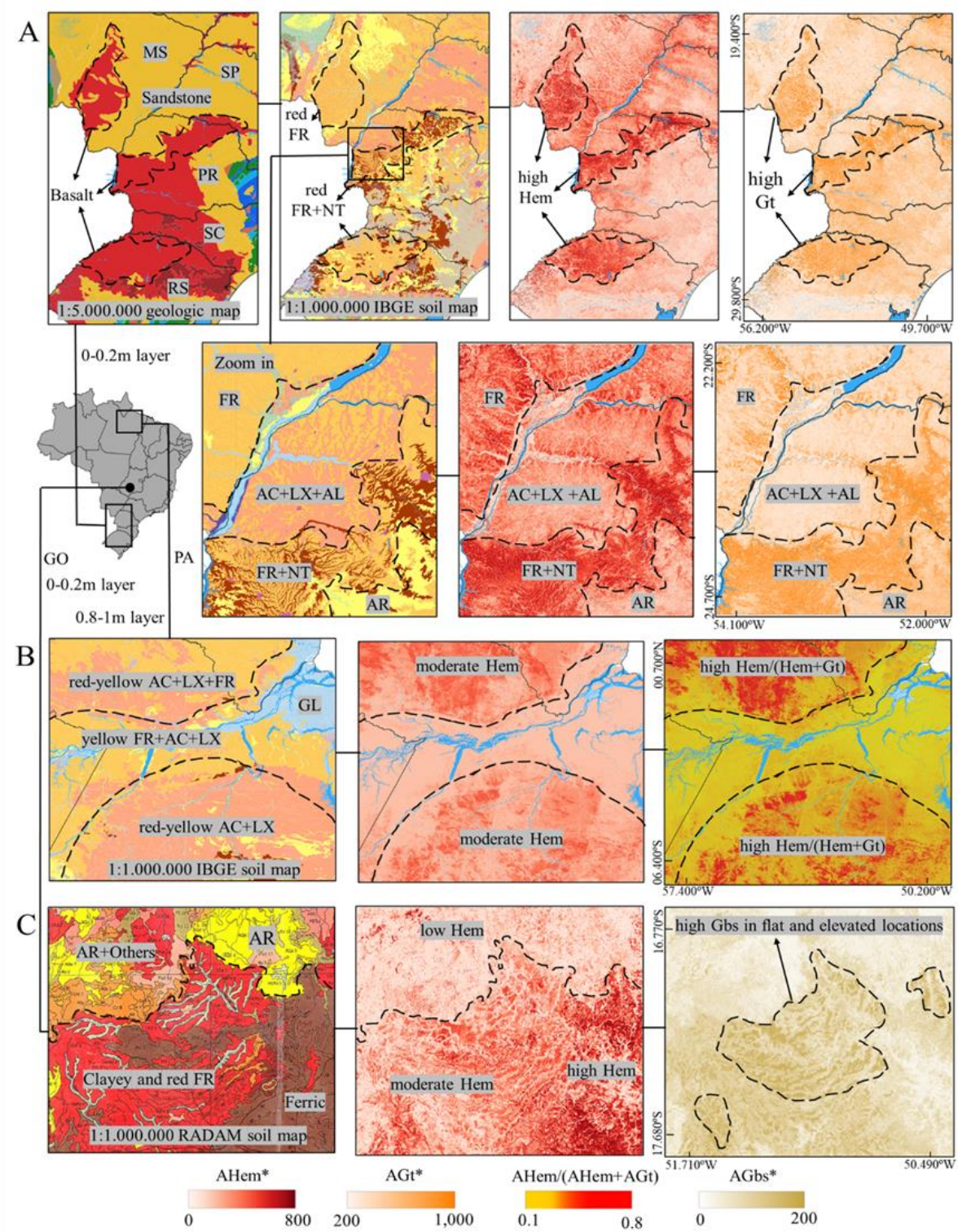

Figure 6: Study cases for hematite (AHem) and goethite (AGt) amplitudes, covering Rio Grande do Sul (RS), Santa Catarina (SC), Paraná (PR), São Paulo (SP) and Mato Grosso do Sul (MS) States, with zoom in the west of PR (A), for AHem and AHem/(AHem+AGt) ratio covering the northeast of Pará State (PA) (B) and for AHem and Gbs amplitude (AGbs) covering the south of Goiás State (GO) (C). FR = Ferralsol; NT= Nitisol; AC=Acrisol; LX=Lixisol; AL=Alisol. The geologic (adapted), the first and the second soil maps were from Gómez et al. (2019), Instituto Brasileiro de Geografia e Estatística (2021) and RADAMBRASIL (1983), respectively. Red soils have hue equal to 2.5YR or more reddish and red-yellow soils have hue equal to 5YR. Ferric soils had Fe_2O_3 equal or more of 18 g kg^{-1} and clayey soil had clay equal or more of 350 g kg^{-1}. $*10^{-6}$ scale.

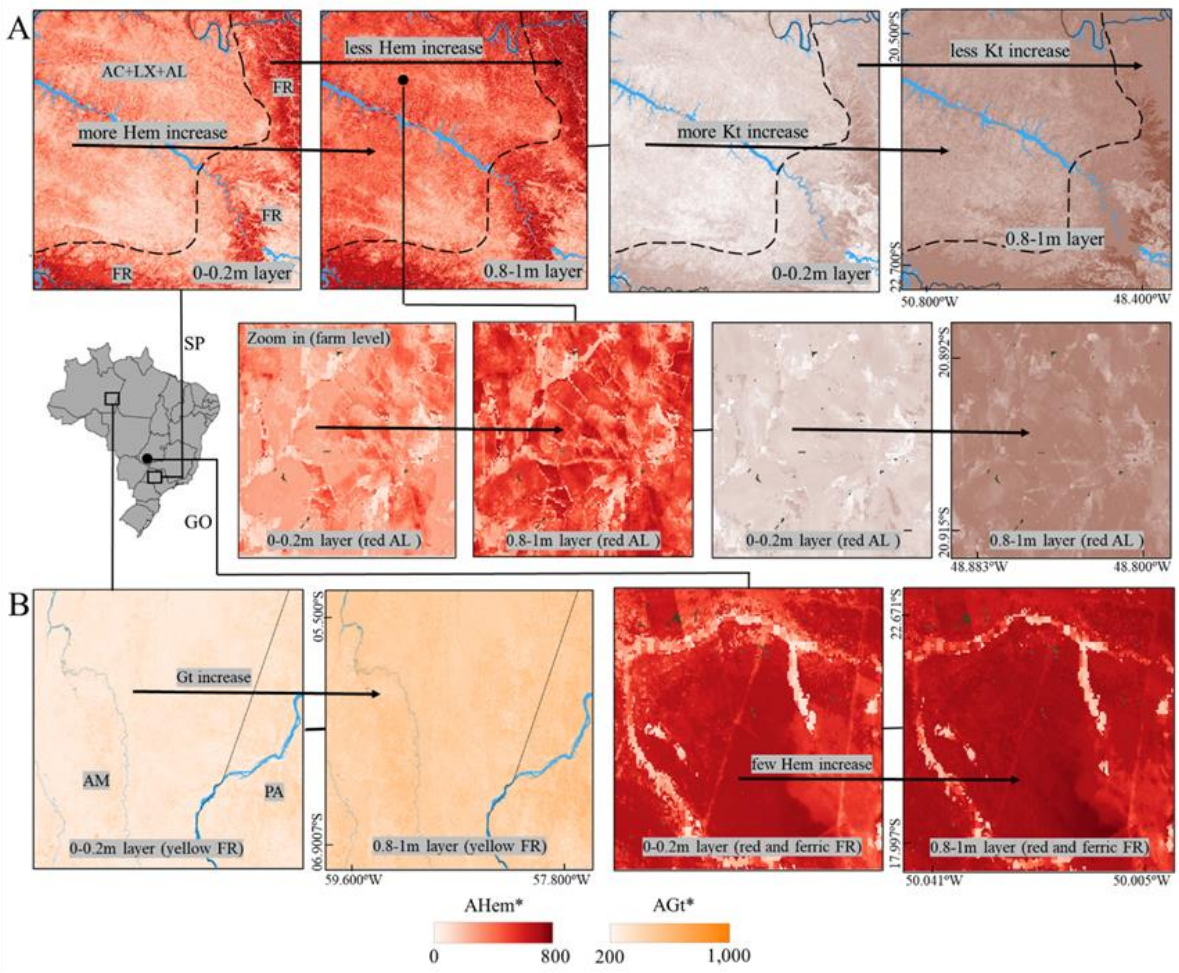

Figure 7: Case of studies in 0-0.2m and 0.8-1m layers for hematite (AHem) and kaolinite (AKt) amplitudes covering the west São Paulo (SP) State, with zoom at farm level (A), for goethite amplitude (AGt) in the south-west of Amazonas (AM) and hematite amplitude (AHem) in the south of Goiás States (B). FR = Ferralsol; AC=Acrisol; LX=Lixisol; AL=Alisol. Red soils have hue equal to 2.5YR or more reddish, red-yellow soils have hue equal to 5YR and yellow soils have hue equal to 7.5YR or more yellowish. Ferric soils had Fe_2O_3 equal or more of 18 g kg^{-1}. $*10^{-6}$ scale.

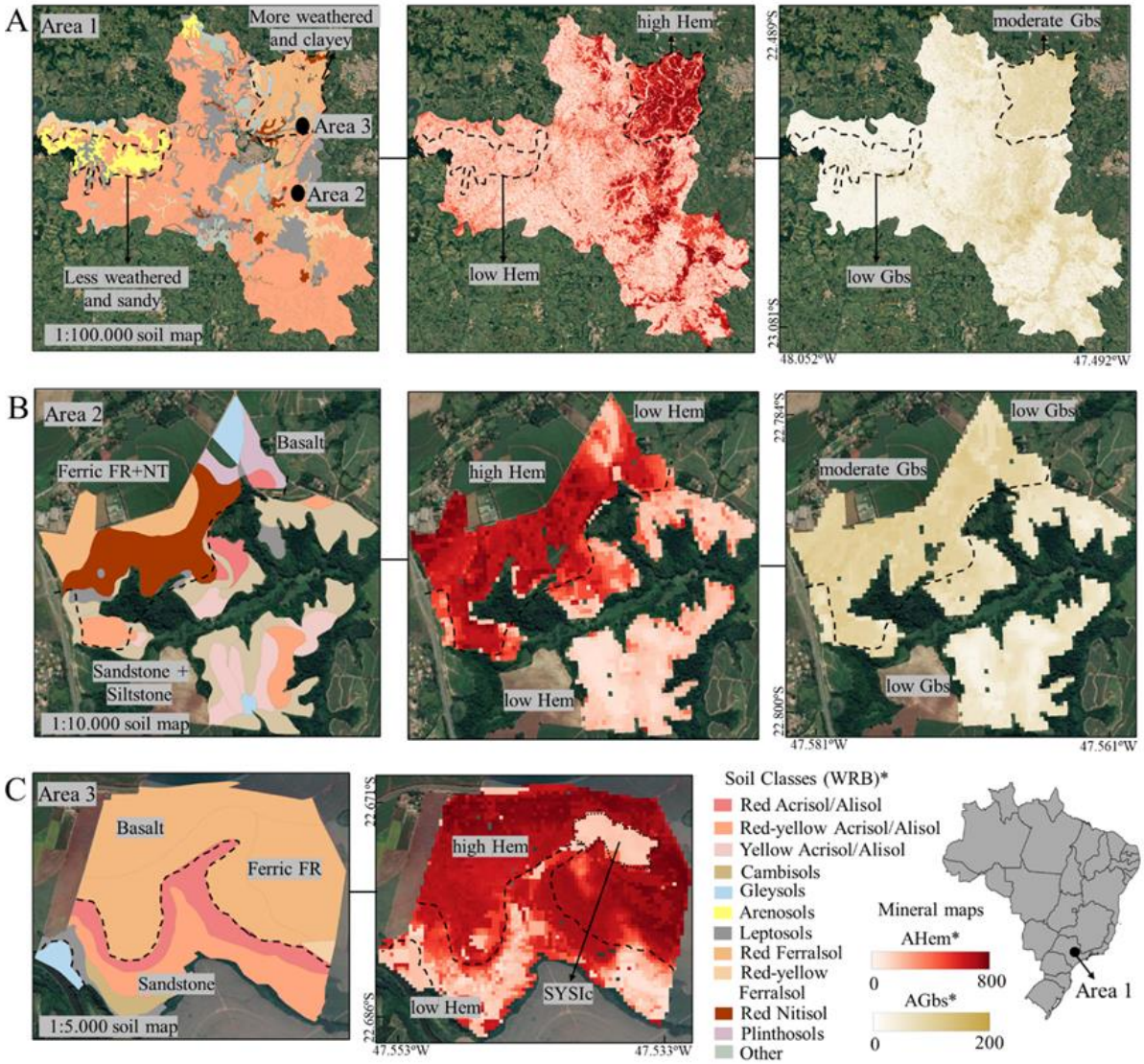

Figure 8: Case of studies in 0.8-1m layer for hematite (AHem) and gibbsite (AGt) amplitudes in regional (Piracicaba-São Paulo region) (A) and farm (B) (C) levels. The 1:100.000 (adapted), 1:10.000 and 1:5.000 soil maps were from Olivera and Prado (1989), Souza et al., (2020) and Demattê et al. (2004), respectively. Red soils have hue equal to 2.5YR or more reddish, red-yellow soils have hue equal to 5YR and yellow soils had hue equal to 7.5YR or more yellowish. Ferric soils had Fe_2O_3 equal or more of 18 g kg^{-1}. *10-6 scale. SySIc = Soil synthetic image with full coverage.

3.4. Discussion

3.4.1 Descriptive statistics of relative abundance of minerals

The reported values for AM (Table 2, Fig. 3) was in accordance with the literature (Mendes et al., 2021; Poppiel et al., 2020a). These studies in Brazilian soils also found increased values for Gt, Hem, Kt and Gbs with the depth increase (Mendes et al., 2021; Poppiel et al., 2020a). In fact, mineral amounts tend to increase, by the clay, silicon, iron and aluminum content increases in depth (horizon B) by pedogenetic processes, such as eluviation and illuviation (Blume and Schwertmann, 1969; Buol et al., 2011; Carroll, 1953; Macías and Camps-Arbestain, 2020). The soil observations from BSSL represents the variability of Brazilian soils, however the most samples were from agricultural land with manly FR, AC and LX soil classes (Demattê et al., 2019). These soils cover more than 60% of the Brazilian territory and are characterized as highly weathered soils, with intense

base lixiviation and predominance of Kt and iron and aluminum oxides (Macías and Camps-Arbestain, 2020; Schaefer et al., 2008).

The Hem and Gbs showed high variations being the occurrence related with specific conditions of relief, climate, geology and soil (Table 2, Fig. 1, Fig. 3abcd, Fig. A1 and Fig. A2) (Schaefer et al., 2008). The Kt and Gt showed less variability due the more occurrence in Brazilian soils and stability in the environment (Table 2 and Fig. 3de) (Fink et al., 2016; Schaefer et al., 2008). The more variability found in the 0-0.2m layer can be explained by the soil types of Brazil (Table 2) (Fig. A1b). Brazil is composed of soils with homogeneity along the profiles, such as FR and PT and also with textural gradients promoted by the translocation of fine particles, such as AC, LX, AL and NT (Fig A1b) (Santos et al., 2018). The soil classes with homogeneous profiles have similar amounts of minerals in surface and subsurface, while the soil classes with textural gradients have accumulation of minerals in B textural horizon (Lelong et al., 1976; Santos et al., 2018).

3.4.2 Mineral modelling and accuracy of maps

3.4.2.1 Predictive models

The more accurate results for the superficial layer than deep layers is explained by the use of SySI as covariate for DSM (Table 2), which presented higher correlation than TA, with the AM (Table A1). It occurs because the SySI is a direct measurement of topsoil reflectance, revealed by temporal satellite images and data mining procedures (Demattê et al., 2018). The SySI can be used to predict soil attributes in soil subsurface layers, with a degree of accuracy decrease (Mendes et al., 2019). Mendes et al. (2021), highlighted the use of SySI as a predictor for AM mapping at soil depths until 1m.

The more accurate results obtained for AHem modelling were in accordance with the literature (Table 2). Other authors also found more accurate results for AHem than other minerals, using several statistical methods for spatial modelling and punctual content estimation (Canton et al., 2021; De Souza Bahia et al., 2015; Mendes et al., 2021; Poppiel et al., 2020a; Sellitto et al., 2009; Silva et al., 2020). On the other hand, the AGt modelling showed the worst results. In the literature the results for AGt are more variable than AHem, being almost always lower (Canton et al., 2021; Mendes et al., 2021; Poppiel et al., 2020a; Sellitto et al., 2009; Silva et al., 2020). The electronic transitions of Hem in the 535-580 nm and of Gt in 415-455nm range are well separated from each other, permitting the distinction of iron forms and detection of low contents in soil samples (Scheinost, 1998). The absorption feature of Hem and Gt affects also the reflectance of soils obtained by Landsat satellites mainly in the visible bands (Madeira-Neto et al., 1997), that were used as covariates in this study. The AHem and AGt showed highest and moderate correlations in the Vis range, respectively (Table 1A). The Hem has a deeper and well-defined absorption feature than Gt (Demattê and Garcia, 1999; Kosmas et al., 1984). The substitution of iron by aluminum in the Gt structure can reach until 40% in Brazilian FR and can be a cause of the low accuracy for Gt, by the less stability in the absorption feature (Kosmas et al., 1984; Poppiel et al., 2020a; Schaefer et al., 2008; Scheinost, 1998).

The second better result for mineral modelling was obtained for AGbs followed by AKt (Table 2). The AGbs and AKt determination by spectroscopy has been less explored than iron oxides, with works showing variables results from several statistical methods, normally with more accuracy for AGbs than AKt (Fernandes et al., 2020; Madeira-Neto et al., 1995; Mendes et al., 2021; Poppiel et al., 2020b; Viscarra Rossel, 2011); Low correlations with SySI were verified for AKt and AGbs than AHem and AGt (Table A1). The groups OH and

metal-OH show vibrational activity in the 2,200 to 2,300 region, being this region related beyond AKt and AGbs, to halloysite and 2:1 minerals, such as montmorillonites, muscovite and illite (Clark et al., 1990; Dufréchou et al., 2015). Because of it, this region cannot be used to detect the Al-OH groups in all cases (Clark et al., 1990). The halloysite and montmorillonite (near 2,200nm) have similar absorption features than AKt (Clark et al., 1990; Goetz et al., 2009) and it can explain the low accuracy for AKt modelling.

There are few studies using environmental covariates and DSM procedures to soil mineral mapping at fine resolution, such as the described above, that can be directly compared with statistics parameters reported in table 2. Poppiel et al., (2020) mapped an 851,000 km^2 area with spatial resolution of 30m and found R^2 of 0.71, 0.72, 0.72 for AHem, 0.45, 0.45, 0.24 for AGt, 0.47, 0.55, 0,59 for AKt and 0.55, 0.64, 0.65 for AGbs, for 0-0.2, 0.2-0.6 and 0.6-1m layers, using TA, a synthetic vegetation image, SySI, and climate covariates. Poppiel et al., (2020) also used the RF algorithm, however the models were validated by 10-fold cross validation, which tends to overestimate the accuracy parameters when compared to other split methods (Volkan Bilgili et al., 2010), such as bootstrapping. Viscarra-Rossel et al., (2011) used the Cubist algorithm and Landsat bands, TA, climate, geological and gamma ray data to predict AKt, with resolution of 90m for Australian territory (7,7 million of km^2) and reached R^2 ranging from 0.50 to 0.53 for 0-0.2m layer and from 0.45 to 0.48 for 0.6-0.8m layers. In a regional study area (2,274 km^2) located in Brazil, Mendes et al., (2021) obtained mineral maps with 30m of resolution using the SySI and RF algorithm. The authors reported R^2 of 0.54, 0.17, 0.62 for AHem, 0.16, 0.10, 0.24 for AGt, 0.32, 0.00, 0,38 for AKt and 0.17, 0.09, 0.62 for AGbs, for 0-0.2, 0.4-0.6 and 0.8-1m layers.

3.4.4.2 Spatial uncertainty and quality of maps

Generally, the AM maps obtained were coherent with legacy maps, such as soil and geology, in national, regional and farm level (Fig. 3, Fig.6, Fig. 7, Fig. 8 and Fig. A1). However, some errors and low quality of maps in some locations, such as north of Brazil (Amazon rainforest) (Fig. 3 and 4) can be observed. These errors may be associated with to manly factors: 1) low density of soil observations and 2) the use of the predicted bands of SySIc (locations without any bare soil pixel from 1984 to 2020, as was explained in the topic 3.2.3.2). The coefficient of variation (CV) for AHem and AGbs prediction was high for areas with low predicted mean values (Fig. 4abe and Fig. 3abe). It can be an indicator of overestimation of predicted values in these areas as reported by Safanelli et al. (2021a) for clay content mapping. The less accurate models (AGt and AKt) showed low CV values and it was not possible to see variation along the mapped territory (in the scale visualized), except for AGt in 0-0.2m layer (Fig. 4cde). The uncertainty found is in accordance with other papers that used DSM and environmental covariates to predict soil attributes (Poggio et al., 2021; Poppiel et al., 2021; Safanelli et al., 2021a)

The large territory mapped had low density of soil observation (Table 2) (ranging from about 8.29 x 10-4 to 1.09 x 10^{-3} points per km^2) and the soil observations were concentrated in the southwest and middle-west region, with few sampled points in the south, northwest and north regions (Fig 1.). It occurred because of the use of legacy data from BSSL, which had the samples located mainly in agricultural areas (Demattê et al., 2019). Similar situations were related by Liu et al., (2012), that mapped the soil particle size distribution for China territory and Poppiel et al. (2020) in mineralogy mapping described above.

As mentioned, the SySI is a directed measurement of topsoil reflectance, while the SySIc has the original SySI plus the predicted bands in the location without soil exposure, that have uncertainties included. The

"gaps" of SySI were predicted using a Landsat mosaic of vegetation (with low spatial resolution) and TA with relatively low accuracy (Table A2). The longitudinal lines perceived in the SySIc (Fig. A1b) is related to the clouds, that interfere in the Landsat imagens that compose the vegetation mosaic used as covariate and causes the exclusion of some entire images. These lines can be perceived also in the AM maps (Fig. 3). Some areas with only predicted reflectance values, showed more doubtful AM values in predicted maps, such as in the mountainous regions along the coast of the Atlantic Ocean and Amazon rainforest (Fig. 3, Fig. 5, Fig 7b). In a study case at farm level, it is possible to see an area with probably error due the use of predicted reflectance (Fig. 8c).

3.4.3 Spatial and depth variation of minerals

3.4.3.1 Iron oxides

The geologic played a fundamental role on iron oxide minerals amounts distribution that can be perceived by visual analysis in the AM maps for Brazil territory (Fig. 3abcd; Fig. A1a) and also in studies of cases at a detailed level (Fig. 7a; Fig. 8bc). The volcanic mafic rocks, such as basalt, are rich in iron and magnesium and originate soil with great amounts of iron oxide minerals (Schaefer et al. 2008; Goulart et al, 1998; Long et al., 2011). In fact, among the soil forming factors, the parent material had the greater influence in mineralogy (Kiely, 1991). In the study case for RS, SC, PR, SP and MS it is possible perceive the great influence of geology in the middle-east (in MS State) and geology plus climate in the south of Brazil (RS and PR States) in iron oxides amounts, mainly Hem (Fig 1. Fig. 6a) (Schaefer et al., 2008). The Hem occurrence is related mainly with more weathered and red soils (Viscarra Rossel et al., 2010), such as FR and NT classes (Fig. 3ab, Fig A1b, Fig. 6) and also with soils characterized as clayey or/and ferric soils (Fig. 6c, Fig.7b and Fig. 8).

The climate and terrain also influenced the iron oxide minerals distribution (Fig. 1, Fig. 3abcd and Fig. A2 bc). The Hem high amounts occurred specifically in some locations under tropical and subtropical zones, with less amounts in dry environments. Indeed, occurrences of iron oxides are characteristic of warmer and humid climates (Long et al., 2011; Macías and Camps-Arbestain, 2020; Schaefer et al., 2008). The Hem occurrence is limited by drained conditions, provided by the terrain, being absent in flat areas with poor drainage, such as the Pantanal located in the west of MS and the Amazon basin (north of Brazil), while, the Gt ocurres also under limited drainage conditions (Fig. 4abcd, Fig. 6b, Fig. 7b, Fig. A2cd). The Gt can occur on imperfect drainage conditions, by reduction of iron oxide minerals by anaerobic microorganisms (Breemen and Buurman, 2002; Macedo and Bryant, 1989; Macías and Camps-Arbestain, 2020). The Gt is more stable than Hem, being present in most variable environments and more predominant than Hem (Macedo and Bryant, 1989; Resende et al., 1986; Schaefer et al., 2008).

As mentioned, soils with homogeneous profiles means similar distribution at depth, as the case of AHem and AGt distribution in FR originated from basalt rocks (Fig. 3abcd, Fig.7a Fig. A1a). For example, it can be perceived in the west of SP and south of GO (Fig. 7). Conversely, soil types with textural gradients had iron oxides increases at deep layers along the territory. The AHem increase was punctual, that can be associated with the occurrence of red AC, LX and AL and can be perceived in the west of SP, for example (Fig. 7a)

The more generalized occurrence verified for Gt than Hem in the 0.8-1m layer, can be related for the intermediary weathered soil types, such as red-yellow and yellow AC, LC and AL, that normally have more abundance of Gt in the subsurface (Fig. A1b) as related by Aquino et al. (2016). It is possible to visualize a more

pronounced increase for Gt in the Amazon region (north of Brazil), which has predominance of soils with matiz 5YR or more yellowish (Fig. 4cd Fig. A1). However, areas with yellow FR also showed Gt amount increased with depth, while it did not occurred for Hem in red RF (Fig. 7b). A probability explanation for this phenomenon is the masking of iron oxides by soil organic matter in the 0-0.2m layer (Croft et al., 2012; Heller Pearlshtien and Ben-Dor, 2020). This effect is more pronounced for Gt than Hem, due to the pigmentation power of Hem, directly affecting the spectral response in the visible region. Additionally the Gt has higher affinity for organic matter, compared to Hem (Guzman et al., 1994).

The variation of Hem/(Hem+Gt) ratio verified in Fig. 5ab represents the proportional amounts of Hem and Gt being directly related with soil color. The Hem is responsible for the red color of soils even in small amounts, while the Gt had yellow pigmentation (Kosmas, 1984; Sheinost et al., 1998). A mixture of 1 part of Hem for 4 part of Gt is sufficient to confer a hue of 5YR. (Kosmas, 1984). In this way, the Hem is almost absent in yellow soils, while small amounts of Hem are sufficient to mask the Gt and give red color for soils (Schertaman, 1993; Schaefer et al. 2008; Kosmas, 1984). In southern Brazil the Hem/Gt+Hem is driven mainly by climatic conditions, while in the central region the parent material, climatic conditions and drainage control (Schaefer et al. 2008).

It is possible to see high values of Hem/(Gt+Hem) ratio in locations with varied parent material and climate conditions, except in dry zones (northeast) and subtropical zones with temperate summer (south) (Fig. 1, Fig. A1a and Fig. 5ab). The terrain had great influence, with higher values of Hem/(Gt+Hem) ratio for locations with middle elevation and slope (Fig. 5ab and Fig. A2cd). The areas with flat relief promote the Gt occurrence by the poor drainage, while areas with high slope are unstable and because of it, without conditions for intense weathering and high amounts of Hem occurrence (Breemen and Buurman, 2002; Macías and Camps-Arbestain, 2020). In the PA state, it is possible to see high Hem/(Hem+Gt) values related to red-yellow AC, LX and FR and low values related to yellow FR, AC, LC and GL (Fig. 6b). Generally, there was no variations of Hem/(Gt+Hem) ratio with depth, being the Amazon basin (north of Brazil) an example of exception due the increase of Gt amount related to the occurrence of soils with hue 5YR or more yellow in the subsurface layer or the masking of Gt in the surface layer as mentioned above. Finally, the ternary plot showed the occurrence of iron oxides mainly associated with the parent material, with predominance in soils from volcanic mafic rocks (Fig. 5d and Fig. A2a).

3.4.3.2 Gibbsite

The occurrence of Gbs in Brazilian territory was related to the position with high altitude (>500m) and felsic plutonic and metamorphic rocks, such as granite and gneisses (Fig. 3e Fig. 6c and Fig. A1a). In fact, although Gbs can occur in mostly varied parent material, Buol et al. (2011) related the propensity of this kind of rocks to originate soil with high amounts of Gbs. An example of high occurrence of Gbs in flat and elevated locations was the south of GO (Fig 6c). Poppiel et al., (2020) also verified this behavior for Gbs using a combined proximal and remote sensing approach in the middle-west of Brazil. Studies of cases at regional and farm level showed more occurrence of high amounts of Gbs in the most weathered soil classes (Fig. 8ab). In fact, several studies related the Gbs occurrence with the altitude and more weathered soils in Brazil (Gomes et al., 2004; Poppiel et al., 2020a; Reatto et al., 2008; Schaefer et al., 2008).

Moderate amounts of Gbs were verified in soils derived from volcanic mafic rocks, associated with iron oxides (Fig. 3abcde, Fig. A1a). The Gbs occurrence were associated with several kinds of soils, such as highly weathered FR, intermediary, such as AC and LX and less weathered such as CX (Fig. 3e, Fig. 6e and Fig. A1). This point partially contrasts the literature that attributes more amounts of Gbs to highly weathered soils from mafic rocks, due the small amounts of silicon that increases the Gbs formation (Schaefer et al., 2008). In fact, the high Gbs amounts in less weathered soils, mainly in the mountainous regions along the coast of the Atlantic Ocean can be wrong. On the other hand, the Gbs occurrence is not related exclusively to high weathered soil classes formed under humid tropical climates (Macías Vazquez, 1981). Two processes that occur under free drainage conditions, low silica activity and few bases contents can explain the occurrence of Gbs in soil, the neo formation in the initial phases of Al-silicates weathering and the intense weathering process with dissolution of Kt (Macías Vazquez, 1981; Schaefer et al., 2008; Tardy et al., 1973) In the same way, the Gbs abundance is extremely variable in Brazilian soils, even in FR (Schaefer et al., 2008).

According to the literature, the Gbs can increase with depth, keep stable or decrease (Buol et al., 2011; Macedo and Bryant, 1989), being verified the least increases when compared with the other minerals (Fig. 4e) (map for 0-0.2m layer not shown). The Kt/(Kt+Gbs) ratio and the ternary plot showed the predominance of Gbs in regions with higher altitude, mainly associated with felsic rocks (Fig. 2a, Fig. 6cd and Fig. A1c).

3.4.3.3 Kaolinite

The Kt mineral showed considerable amounts in all soils types from several parent material and presented a generalized increase with depth, reaching significant amounts in almost all territory (Fig. 3f) (map for 0-0.2m layer not shown). In fact, the literature does mention a generalized occurrence of Kt in Brazilian soils, being present mainly in FR, NT, AC, LX and PT (Schwertmann and Kämpf, 1985). The development of Kt minerals can occur in several conditions, mainly in humid and warmer climates, with free drainage conditions, low pH and non-excessive Si leaching (Schaefer et al., 2008). The regions with moderate values have sedimentary parent material or dry climate conditions, such as some regions in the northeast of Brazil (Fig. 1, Fig. 2a and Fig. 4f). In the west of SP, it is possible to visualize the increase of Kt abundance from topsoil to 0.8-1m depth due the illuviation and eluviation in AC, LX and AL and the low variation with depth in FR (Fig. 7a). The Kt/(Kt+Gbs) ratio and the ternary plot showed the predominance of Kt in the national territory under several conditions (Fig 1, Fig 2, Fig 6cd and Fig. A1cd).

3.5. Conclusion

The abundances of major soil mineralogical components: goethite (Gt), hematite (Hem), kaolinite (Kt) and gibbsite (Gbs) were successfully identified and quantified by Vis-NIR-SWIR reflectance in samples from Brazilian Soil Spectral Library (BSSL).

These soil minerals presented a significant correlation with spectral data of a synthetic soil image (SySI) with bare soil pixels obtained from temporal Landsat images and clay content. The Hem presented the more accurate results in spatial prediction with R^2 ranging from 0.48 to 0.56, followed by Gbs (0.42 to 0.44), Kt (0.20 to 0.31) and Gt (0.16 to 0.26).

The spatial distribution of minerals was predicted for the entire Brazilian territory. For it, a novel framework was implemented to obtain a continues bare soil reflectance image.

The mineral maps obtained were in accordance with geology and soil legacy maps and also with the climate and terrain conditions at the national, regional and farm level.

The proposed approach revealed the distribution of mineral abundances in the Brazilian territory and consists of an efficient method to obtain mineralogy information for large areas.

References

Alvares, C.A., Stape, J.L., Sentelhas, P.C., de Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Zeitschrift 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Aquino, R.E. de, Marques, J., Campos, M.C.C., Oliveira, I.A. de, Bahia, A.S.R. de S., Santos, L.A.C. dos, 2016. Characteristics of color and iron oxides of clay fraction in Archeological Dark Earth in Apuí region, southern Amazonas. Geoderma 262, 35–44. https://doi.org/10.1016/j.geoderma.2015.07.010

Barrón, V., Torrent, J., 1986. Use of the Kubelka-Munk theory to study the influence of iron oxides on soil colour. J. Soil Sci. 37, 499–510. https://doi.org/10.1111/j.1365-2389.1986.tb00382.x

Barthès, B.G., Kouakoua, E., Clairotte, M., Lallemand, J., Chapuis-Lardy, L., Rabenarivo, M., Roussel, S., 2019. Performance comparison between a miniaturized and a conventional near infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen. Geoderma 338, 422–429. https://doi.org/10.1016/j.geoderma.2018.12.031

Bellinaso, H., Silvero, N.E.Q., Ruiz, L.F.C., Accorsi Amorim, M.T., Rosin, N.A., Mendes, W. de S., Sousa, G.P.B. de, Sepulveda, L.M.A., Queiroz, L.G. de, Nanni, M.R., Demattê, J.A.M., 2021. Clay content prediction using spectra data collected from the ground to space platforms in a smallholder tropical area. Geoderma 399, 115116. https://doi.org/10.1016/j.geoderma.2021.115116

Blume, H.P., Schwertmann, U., 1969. Genetic Evaluation of Profile Distribution of Aluminum, Iron, and Manganese Oxides. Soil Sci. Soc. Am. J. 33, 438.

Breemen, N. V., Buurman, P., 2002. Soil Formation, 2nd ed. Springer, Amsterdam.

Breiman, L., 2001. Random Forests. Mach. Learn. 5–32. https://doi.org/https://doi.org/10.1023/A:1010933404324

Buol, S.W., Southard, R.J., Graham, R.C., McDaniel, P.A., 2011. Soil Genesis and Classification, 6ª. ed. John Wiley & Sons, Ltd, Chichester, UK.

Canton, L.C., Souza Júnior, I.G. de, Silva, L.S., Marques Júnior, J., Costa, A.C.S. da, 2021. Identification and quantification of iron oxides by diffuse reflectance spectroscopy with Praying Mantis accessory and integration sphere. Catena. https://doi.org/10.1016/j.catena.2020.104899

Carroll, D., 1953. Clay Minerals in a Limestone Soil Profile1. Clays Clay Miner. 2, 171–182. https://doi.org/10.1346/CCMN.1953.0020115

Chipera, S.J., Bish, D.L., 2001. Baseline Studies of the Clay Minerals Society Source Clays Powder X-ray Diffraction Analyses. Clays Clay Miner. 49, 398–409. https://doi.org/10.1346/CCMN.2001.0490507

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. J. Geophys. Res. 95. https://doi.org/10.1029/jb095ib08p12653

Coblinski, J.A., Giasson, É., Demattê, J.A.M., Dotto, A.C., Costa, J.J.F., Vašát, R., 2020. Prediction of soil texture classes through different wavelength regions of reflectance spectroscopy at various soil depths. Catena 189. https://doi.org/10.1016/j.catena.2020.104485

Croft, H., Kuhn, N.J., Anderson, K., 2012. On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. CATENA 94, 64–74. https://doi.org/10.1016/j.catena.2012.01.001

de Oliveira, J.S., Inda, A.V., Barrón, V., Torrent, J., Tiecher, T., de Oliveira Camargo, F.A., 2020. Soil properties governing phosphorus adsorption in soils of Southern Brazil. Geoderma Reg. https://doi.org/10.1016/j.geodrs.2020.e00318

De Souza Bahia, A.S.R., Marques, J., Siqueira, D.S., 2015. Procedures using diffuse reflectance spectroscopy for estimating hematite and goethite in Oxisols of São Paulo, Brazil. Geoderma Reg. https://doi.org/10.1016/j.geodrs.2015.04.006

Demattê, J.A.., Campos, R.C., Alves, M.C., Fiorio, P.R., Nanni, M.R., 2004. Visible–NIR reflectance: a new approach on soil evaluation. Geoderma 121, 95–112. https://doi.org/10.1016/j.geoderma.2003.09.012

Demattê, J.A.M., Dotto, A.C., Paiva, A.F.S., Sato, M. V., Dalmolin, R.S.D., de Araújo, M. do S.B., da Silva, E.B., Nanni, M.R., ten Caten, A., Noronha, N.C., Lacerda, M.P.C., de Araújo Filho, J.C., Rizzo, R., Bellinaso, H., Francelino, M.R., Schaefer, C.E.G.R., Vicente, L.E., dos Santos, U.J., de Sá Barretto Sampaio, E. V., Menezes, R.S.C., de Souza, J.J.L.L., Abrahão, W.A.P., Coelho, R.M., Grego, C.R., Lani, J.L., Fernandes, A.R., Gonçalves, D.A.M., Silva, S.H.G., de Menezes, M.D., Curi, N., Couto, E.G., dos Anjos, L.H.C., Ceddia, M.B., Pinheiro, É.F.M., Grunwald, S., Vasques, G.M., Marques Júnior, J., da Silva, A.J., Barreto, M.C. d. V., Nóbrega, G.N., da Silva, M.Z., de Souza, S.F., Valladares, G.S., Viana, J.H.M., da Silva Terra, F., Horák-Terra, I., Fiorio, P.R., da Silva, R.C., Frade Júnior, E.F., Lima, R.H.C., Alba, J.M.F., de Souza Junior, V.S., Brefin, M.D.L.M.S., Ruivo, M.D.L.P., Ferreira, T.O., Brait, M.A., Caetano, N.R., Bringhenti, I., de Sousa Mendes, W., Safanelli, J.L., Guimarães, C.C.B., Poppiel, R.R., e Souza, A.B., Quesada, C.A., do Couto, H.T.Z., 2019. The Brazilian Soil Spectral Library (BSSL): A general view, application and challenges. Geoderma 354, 113793. https://doi.org/10.1016/j.geoderma.2019.05.043

Demattê, J.A.M., Fongaro, C.T., Rizzo, R., Safanelli, J.L., 2018. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. Remote Sens. Environ. 212, 161–175. https://doi.org/10.1016/j.rse.2018.04.047

Demattê, J.A.M., Garcia, G.J., 1999. Alteration of Soil Properties through a Weathering Sequence as Evaluated by Spectral Reflectance. Soil Sci. Soc. Am. J. 63, 327–342. https://doi.org/10.2136/sssaj1999.03615995006300020010x

Demattê, J.A.M., Safanelli, J.L., Poppiel, R.R., Rizzo, R., Silvero, N.E.Q., Mendes, W. de S., Bonfatti, B.R., Dotto, A.C., Salazar, D.F.U., Mello, F.A. de O., Paiva, A.F. da S., Souza, A.B., Santos, N.V. dos, Maria Nascimento, C., Mello, D.C. de, Bellinaso, H., Gonzaga Neto, L., Amorim, M.T.A., Resende, M.E.B. de, Vieira, J. da S., Queiroz, L.G. de, Gallo, B.C., Sayão, V.M., Lisboa, C.J. da S., 2020. Bare Earth's Surface Spectra as a Proxy for Soil Resource Monitoring. Sci. Rep. 10, 1–11. https://doi.org/10.1038/s41598-020-61408-1

Dufréchou, G., Grandjean, G., Bourguignon, A., 2015. Geometrical analysis of laboratory soil spectra in the short-wave infrared domain: Clay composition and estimation of the swelling potential. Geoderma 243–244, 92–107. https://doi.org/10.1016/j.geoderma.2014.12.014

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, Londres.

Fang, Q., Hong, H., Zhao, L., Kukolich, S., Yin, K., Wang, C., 2018. Visible and Near-Infrared Reflectance Spectroscopy for Investigating Soil Mineralogy: A Review. J. Spectrosc. 2018, 1–14. https://doi.org/10.1155/2018/3168974

Fernandes, K., Marques Júnior, J., Bahia, A.S.R. de S., Demattê, J.A.M., Ribon, A.A., 2020. Landscape-scale spatial variability of kaolinite-gibbsite ratio in tropical soils detected by diffuse reflectance spectroscopy. Catena. https://doi.org/10.1016/j.catena.2020.104795

Fink, J.R., Inda, A.V., Tiecher, T., Barrón, V., 2016. Iron oxides and organic matter on soil phosphorus availability. Ciência e Agrotecnologia 40, 369–379. https://doi.org/10.1590/1413-70542016404023016

Fongaro, C., Demattê, J., Rizzo, R., Lucas Safanelli, J., Mendes, W., Dotto, A., Vicente, L., Franceschini, M., Ustin, S., 2018. Improvement of Clay and Sand Quantification Based on a Novel Approach with a Focus on Multispectral Satellite Images. Remote Sens. 10, 1555. https://doi.org/10.3390/rs10101555

Gilkes, R.J., Prakongkep, N., 2016. How the unique properties of soil kaolin affect the fertility of tropical soils. Appl. Clay Sci. 131, 100–106. https://doi.org/10.1016/j.clay.2016.01.007

Goetz, A.F.H., Curtiss, B., Shiley, D.A., 2009. Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. Miner. Eng. 22, 490–499. https://doi.org/10.1016/j.mineng.2008.12.013

Gomes, L.C., Faria, R.M., de Souza, E., Veloso, G.V., Schaefer, C.E.G.R., Filho, E.I.F., 2019. Modelling and mapping soil organic carbon stocks in Brazil. Geoderma 340, 337–350. https://doi.org/10.1016/j.geoderma.2019.01.007

Gomes, J.B. V., Curi, N., Motta, P.E.F., Ker, J.C., Marques, J.J.G.S.M., Schulze, D.G., 2004. Análise de componentes principais de atributos físicos, químicos e mineralógicos de solos do bioma cerrado. Rev. Bras. Ciência do Solo 28, 137–153. https://doi.org/10.1590/S0100-06832004000100014

Gómez, J., Schobbenhaus, C., Montes, N.E., 2019. Geological Map of South America 2019. Scale 1:5 000 000. https://doi.org/https://doi.org/10.32685/10.143.2019.929

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Guzman, G., Alcantara, E., Barron, V., Torrent, J., 1994. Phytoavailability of phosphate adsorbed on ferrihydrite, hematite, and goethite. Plant Soil 159, 219–225. https://doi.org/10.1007/BF00009284

Hassink, J., 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. Plant Soil 191, 77–87. https://doi.org/10.1023/A:1004213929699

Heimsath, A.M., DiBiase, R.A., Whipple, K.X., 2012. Soil production limits and the transition to bedrock-dominated landscapes. Nat. Geosci. 5, 210–214. https://doi.org/10.1038/ngeo1380

Heller Pearlshtien, D., Ben-Dor, E., 2020. Effect of Organic Matter Content on the Spectral Signature of Iron Oxides across the VIS–NIR Spectral Region in Artificial Mixtures: An Example from a Red Soil from Israel. Remote Sens. 12, 1960. https://doi.org/10.3390/rs12121960

Hengl, T., MacMillan, R.A., 2019. Predictive Soil Mapping with R [WWW Document]. URL https://soilmapper.org/ (accessed 7.7.21).

Hengl, T., Walsh, M.G., Sanderman, J., Wheeler, I., Harrison, S.P., Prentice, I.C., 2018. Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential. PeerJ 6, e5457. https://doi.org/10.7717/peerj.5457

Heuvelink, G.B.M., Angelini, M.E., Poggio, L., Bai, Z., Batjes, N.H., Bosch, R., Bossio, D., Estella, S., Lehmann, J., Olmedo, G.F., Sanderman, J., 2021. Machine learning in space and time for modelling soil organic carbon change. Eur. J. Soil Sci. 72, 1607–1623. https://doi.org/10.1111/ejss.12998

Instituto Brasiliero de Geografia e Estatística, 2021. Mapa de Solos do Brasil [WWW Document]. URL https://www.ibge.gov.br/geociencias/downloads-geociencias.html (accessed 9.10.21).

Japan Aerospace Exploration Agency: ALOS Research and Application Project, 2021. Advanced Land Observing Project [WWW Document]. URL https://www.eorc.jaxa.jp/ALOS/en/index_e.htm (accessed 7.15.21).

Jenny, H., 1941. Factors of Soil Formation: A System of Quantitative Pedology. Dover Publications, Nova York.

Kosmas, C.S., Curi, N., Bryant, R.B., Franzmeier, D.P., 1984. Characterization of Iron Oxide Minerals by Second-Derivative Visible Spectroscopy. Soil Sci. Soc. Am. J. 48, 401–405. https://doi.org/10.2136/sssaj1984.03615995004800020036x

Kunze, G.W., Dixon, J.B., 1986. Pretreatment for Mineralogical Analysis, in: Methods of Soil Analysis: Part 1 Physical and Mineralogical Methods, Second Edition. pp. 91–100.

Lelong, F., Tardy, Y., Grandin, G., Trescases, J.J., Boulange, B., 1976. Pedogenesis, chemical weathering and processes of formation of some supergene ore deposits, in: Wolf, K.H. (Ed.), Handbook of Strata-Bound and Stratiform Ore Deposits. Elsevier Scientific Publishing, Amsterdam, p. 81.

Lepsch, I.F., 2013. Status of Soil Surveys and Demand for Soil Series Descriptions in Brazil. Soil Horizons 54. https://doi.org/10.2136/sh2013-54-2-gc

Liu, F., Geng, X., Zhu, A.-X., Fraser, W., Waddell, A., 2012. Soil texture mapping over low relief areas using land surface feedback dynamic patterns extracted from MODIS. Geoderma 171–172, 44–52. https://doi.org/10.1016/j.geoderma.2011.05.007

Long, X., Ji, J., Balsam, W., 2011. Rainfall-dependent transformations of iron oxides in a tropical saprolite transect of Hainan Island, South China: Spectral and magnetic measurements. J. Geophys. Res. 116, F03015. https://doi.org/10.1029/2010JF001712

Macedo, J., Bryant, R.B., 1989. Preferential Microbial Reduction of Hematite Over Goethite in a Brazilian Oxisol. Soil Sci. Soc. Am. J. 53, 1114–1118. https://doi.org/10.2136/sssaj1989.03615995005300040022x

Macías, F., Camps-Arbestain, M., 2020. A biogeochemical view of the world reference base soil classification system: Homage to Ward Chesworth, 1st ed, Advances in Agronomy. Elsevier Inc. https://doi.org/10.1016/bs.agron.2019.11.002

Macías Vazquez, F., 1981. Formation of gibbsite in soils and saprolites of temperate-humid zones. Clay Miner. 16, 43–52. https://doi.org/10.1180/claymin.1981.016.1.03

Madeira-Neto, J., Bedidi, A., Cervelle, B., Pouget, M., Flay, N., 1997. Visible spectrometric indices of hematite (Hm) and goethite (Gt) content in lateritic soils: The application of a Thematic Mapper (TM) image for soil-mapping in Brasilia, Brazil. Int. J. Remote Sens. 18, 2835–2852. https://doi.org/10.1080/014311697217369

Madeira-Neto, J., Bedidi, A., Pouget, M., Cervelle, B., Flay, N., 1995. Spectral (MIR) determination of kaolinite and gibbsite contents in lateritic soils. Comptes Rendus - Acad. des Sci. Ser. II Sci. la Terre des Planetes 321, 119–128.

McManus, D.A., 1991. Suggestions for authors whose manuscripts include quantitative clay mineral analysis by X-ray diffraction. Mar. Geol. 98, 1–5. https://doi.org/10.1016/0025-3227(91)90030-8

Mendes, W. de S., Demattê, J.A.M., Bonfatti, B.R., Resende, M.E.B., Campos, L.R., Costa, A.C.S. da, 2021. A novel framework to estimate soil mineralogy using soil spectroscopy. Appl. Geochemistry 127. https://doi.org/10.1016/j.apgeochem.2021.104909

Mendes, W. de S., Medeiros Neto, L.G., Demattê, J.A.M., Gallo, B.C., Rizzo, R., Safanelli, J.L., Fongaro, C.T., 2019. Is it possible to map subsurface soil attributes by satellite spectral transfer models? Geoderma 343, 269–279. https://doi.org/10.1016/j.geoderma.2019.01.025

Mendonça-Santos, M.L., dos Santos, H.G., 2006. Chapter 3 The State of the Art of Brazilian Soil Mapping and Prospects for Digital Soil Mapping. pp. 39–601. https://doi.org/10.1016/S0166-2481(06)31003-3

Moura-Bueno, J.M., Dalmolin, R.S.D., Horst-Heinen, T.Z., ten Caten, A., Vasques, G.M., Dotto, A.C., Grunwald, S., 2020. When does stratification of a subtropical soil spectral library improve predictions of soil organic carbon content? Sci. Total Environ. 737, 139895. https://doi.org/10.1016/j.scitotenv.2020.139895

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Dor, E. Ben, Brown, D.J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J.A.M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai, H., Soriano-Disla, J.M., Shepherd, K.D., Stenberg, B., Towett, E.K., Vargas, R., Wetterlind, J., 2015. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. Adv. Agron. 132, 139–159. https://doi.org/10.1016/bs.agron.2015.02.002

Nolasco de Carvalho, C.C., Nunes, F.C., Homem Antunes, M.A., Nolasco, M.C., 2015. Soil Surveys in Brazil and Perspectives in Soil Digital Mapping. Soil Horizons 56. https://doi.org/10.2136/sh14-01-0002

Oliveira, J.B., Prado, H., 1989. Carta pedológica de Piracicaba: Escala 1:100.000.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Bertrand, T., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240. https://doi.org/10.5194/soil-7-217-2021

Poppiel, R.R., Demattê, J.A.M., Rosin, N.A., Campos, L.R., Tayebi, M., Bonfatti, B.R., Ayoubi, S., Tajik, S., Afshar, F.A., Jafari, A., Hamzehpour, N., Taghizadeh-Mehrjardi, R., Ostovari, Y., Asgari, N., Naimi, S., Nabiollahi, K., Fathizad, H., Zeraatpisheh, M., Javaheri, F., Doustaky, M., Naderi, M., Dehghani, S., Atash, S., Farshadirad, A., Mirzaee, S., Shahriari, A., Ghorbani, M., Rahmati, M., 2021. High resolution middle eastern soil attributes mapping via open data and cloud computing. Geoderma 385, 114890. https://doi.org/10.1016/j.geoderma.2020.114890

Poppiel, R.R., Lacerda, M.P.C., Rizzo, R., Safanelli, J.L., Bonfatti, B.R., Silvero, N.E.Q., Demattê, J.A.M., 2020a. Soil Color and Mineralogy Mapping Using Proximal and Remote Sensing in Midwest Brazil. Remote Sens. 12, 1197. https://doi.org/10.3390/rs12071197

Poppiel, R.R., Lacerda, M.P.C., Rizzo, R., Safanelli, J.L., Bonfatti, B.R., Silvero, N.E.Q., Demattê, J.A.M., 2020b. Soil color and mineralogy mapping using proximal and remote sensing in Midwest Brazil. Remote Sens. 12, 1–30. https://doi.org/10.3390/rs12071197

RADAMBRASIL, 1983. Mapa Exploratório de Solos: Goiânia Folha SE 22.

Ramos, P.V., Inda, A.V., Barrón, V., Siqueira, D.S., Marques Júnior, J., Teixeira, D.D.B., 2020. Color in subtropical brazilian soils as determined with a Munsell chart and by diffuse reflectance spectroscopy. Catena 193, 104609. https://doi.org/10.1016/j.catena.2020.104609

Raschka, 2018. Model evaluation, model selection, and algorithm selection in machine learning [WWW Document]. arXiv. URL https://arxiv.org/abs/1811.12808 (accessed 8.15.21).

Reatto, A., Bruand, A., de Souza Martins, E., Muller, F., da Silva, E.M., Carvalho, O.A. de, Brossard, M., 2008. Variation of the kaolinite and gibbsite content at regional and local scale in Latosols of the Brazilian Central Plateau. Comptes Rendus Geosci. 340, 741–748. https://doi.org/10.1016/j.crte.2008.07.006

Resende, M., Allan, J., Coey, J.M.D., 1986. The magnetic soils of Brazil. Earth Planet. Sci. Lett. 78, 322–326. https://doi.org/10.1016/0012-821X(86)90071-3

Rosin, N.A., Demattê, J.A.M., Leite, M.C.A., de Carvalho, H.W.P., Costa, A.C., Greschuk, L.T., Curi, N., Silva, S.H.G., 2021. The fundamental of the effects of water, organic matter, and iron forms on the pXRF information in soil analyses. CATENA 105868. https://doi.org/10.1016/j.catena.2021.105868

Safanelli, J.L., Poppiel, R., Ruiz, L., Bonfatti, B., Mello, F., Rizzo, R., Demattê, J., 2020a. Terrain Analysis in Google Earth Engine: A Method Adapted for High-Performance Global-Scale Analysis. ISPRS Int. J. Geo-Information 9, 400. https://doi.org/10.3390/ijgi9060400

Safanelli, J.L., Chabrillat, S., Ben-Dor, E., Demattê, J.A.M., 2020b. Multispectral Models from Bare Soil Composites for Mapping Topsoil Properties over Europe. Remote Sens. 12, 1369. https://doi.org/10.3390/rs12091369

Safanelli, J. L., Demattê, J.A.M., Chabrillat, S., Poppiel, R.R., Rizzo, R., Dotto, A.C., Silvero, N.E.Q., Mendes, W. de S., Bonfatti, B.R., Ruiz, L.F.C., ten Caten, A., Dalmolin, R.S.D., 2021a. Leveraging the application of Earth observation data for mapping cropland soils in Brazil. Geoderma 396, 115042. https://doi.org/10.1016/j.geoderma.2021.115042

Safanelli, J. L., Demattê, J.A.M., Santos, N.V. dos, Rosas, J.T.F., Silvero, N.E.Q., Bonfatti, B.R., Mendes, W. de S., 2021b. Fine-scale soil mapping with Earth Observation data: a multiple geographic level comparison. Rev. Bras. Ciência do Solo 45. https://doi.org/10.36783/18069657rbcs20210080Santos, P.G. dos, Almeida, J.A.

de, Sequinatto, L., 2017. Mineralogy of the Clay Fraction and Chemical Properties of Soils Developed from Sedimentary Lithologies of Pirambóia, Sanga-the-Cabral and Guará Geological Formations in Southern Brazil. Rev. Bras. Ciência do Solo 41. https://doi.org/10.1590/18069657rbcs20160344

Santos, H.G., Jacomine, P.K.T., Anjos, L.H.C., Oliveira, V.A., Lumbreras, J.F., Coelho, M.R., Almeida, J.A., Araújo Filho, J.C., Oliveira, J.B., Cunha, T.J.F., 2018. Brazilian Soil Classification System, 5 ed. EMBRAPA, Brasilia.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047

Schaefer, C.E.G.R., Fabris, J.D., Ker, J.C., 2008. Minerals in the clay fraction of Brazilian Latosols (Oxisols): a review. Clay Miner. 43, 137–154. https://doi.org/10.1180/claymin.2008.043.1.11

Scheinost, A.C., 1998. Use and Limitations of Second-Derivative Diffuse Reflectance Spectroscopy in the Visible to Near-Infrared Range to Identify and Quantify Fe Oxide Minerals in Soils. Clays Clay Miner. 46, 528–536. https://doi.org/10.1346/CCMN.1998.0460506

Schwertmann, U., Kämpf, N., 1985. Properties of goethite and hematite in kaolinitic soils of southern and central Brazil. Soil Sci. 139, 344–350.

Sellitto, V.M., Fernandes, R.B.A., Barrón, V., Colombo, C., 2009. Comparing two different spectroscopic techniques for the characterization of soil iron oxides: Diffuse versus bi-directional reflectance. Geoderma 149, 2–9. https://doi.org/10.1016/j.geoderma.2008.11.020

Silva, E.B., Giasson, É., Dotto, A.C., Caten, A. ten, Demattê, J.A.M., Bacic, I.L.Z., Veiga, M. da, 2019. A Regional Legacy Soil Dataset for Prediction of Sand and Clay Content with Vis-Nir-Swir, in Southern Brazil. Rev. Bras. Ciência do Solo 43. https://doi.org/10.1590/18069657rbcs20180174

Silva, L.S., Marques Júnior, J., Barrón, V., Gomes, R.P., Teixeira, D.D.B., Siqueira, D.S., Vasconcelos, V., 2020. Spatial variability of iron oxides in soils from Brazilian sandstone and basalt. Catena. https://doi.org/10.1016/j.catena.2019.104258

Silvero, N.E.Q., Demattê, J.A.M., Amorim, M.T.A., Santos, N.V. dos, Rizzo, R., Safanelli, J.L., Poppiel, R.R., Mendes, W. de S., Bonfatti, B.R., 2021. Soil variability and quantification based on Sentinel-2 and Landsat-8 bare soil images: A comparison. Remote Sens. Environ. 252, 112117. https://doi.org/10.1016/j.rse.2020.112117

Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., MacDonald, L.M., McLaughlin, M.J., 2014. The performance of visible, near and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49, 139–186. https://doi.org/10.1080/05704928.2013.811081

Souza, A.B., 2020. Proximal and remote sensing on the soil processes: from punctual to spatial approaches. Universaty of São Paulo.

Tardy, Y., Bocquier, G., Paquet, H., Millot, G., 1973. Formation of clay from granite and its distribution in relation to climate and topography. Geoderma 10, 271–284. https://doi.org/10.1016/0016-7061(73)90002-5

Vaudour, E., Cerovic, Z., Ebengo, D., Latouche, G., 2018. Predicting Key Agronomic Soil Properties with UV-Vis Fluorescence Measurements Combined with Vis-NIR-SWIR Reflectance Spectroscopy: A Farm-Scale Study in a Mediterranean Viticultural Agroecosystem. Sensors 18, 1157. https://doi.org/10.3390/s18041157

Viscarra Rossel, R.A., 2011. Fine-resolution multiscale mapping of clay minerals in Australian soils measured with near infrared spectra. J. Geophys. Res. 116, F04023. https://doi.org/10.1029/2011JF001977

Viscarra Rossel, R.A., Bui, E.N., De Caritat, P., McKenzie, N.J., 2010. Mapping iron oxides and the color of Australian soil using visible-near-infrared reflectance spectra. J. Geophys. Res. Earth Surf. 115, 1–13. https://doi.org/10.1029/2009JF001645

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131, 59–75. https://doi.org/10.1016/j.geoderma.2005.03.007

Volkan Bilgili, A., van Es, H.M., Akbas, F., Durak, A., Hively, W.D., 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. J. Arid Environ. 74, 229–238. https://doi.org/10.1016/j.jaridenv.2009.08.011

Wadoux, A.M.J.-C., Padarian, J., Minasny, B., 2019. Multi-source data integration for soil mapping using deep learning. SOIL 5, 107–119. https://doi.org/10.5194/soil-5-107-2019

Whittig, L.D., 1965. X-Ray Diffraction Techniques for Mineral Identification and Mineralogical Composition, in: Black, C.A. (Ed.), Methods of Soil Analysis: Part 1 Physical and Mineralogical Properties, Including Statistics of Measurement and Sampling. American Society of Agronomy, Madison, pp. 671–698. https://doi.org/10.2134/agronmonogr9.1.c49

Zhao, D., Arshad, M., Li, N., Triantafilis, J., 2021. Predicting soil physical and chemical properties using vis-NIR in Australian cotton areas. CATENA 196, 104938. https://doi.org/10.1016/j.catena.2020.104938

Zhao, J., Chen, S., Hu, R., Li, Y., 2017. Aggregate stability and size distribution of red soils under different land uses integrally regulated by soil organic matter, and iron and aluminum oxides. Soil Tillage Res. 167, 73–79. https://doi.org/10.1016/j.still.2016.11.007

Appendix

Table A1: Spearman correlation between the Amplitude of minerals (AM), terrain attributes (TA), Synthetic soil image (SySI) bands and Landsat mosaic of 1984.

Covariate	Amplitude of minerals				SySI bands					
	AHem	AGt	AKt	AGbs	blue	green	red	NIR	SWIR1	SWIR2
Elevation	**-0.19**	0.07	0.02	**0.43**	**-0.23**	**-0.20**	**-0.10**	**-0.10**	-0.07	-0.09
Slope	**-0.12**	-0.02	0.01	0.05	**-0.18**	**-0.14**	**-0.20**	**-0.07**	-0.08	**-0.15**
Aspect	-0.02	-0.01	-0.02	-0.06	-0.03	0.00	-0.08	0.00	0.00	-0.03
Hillshade	**0.12**	0.00	0.02	0.00	**0.19**	**0.20**	**0.18**	**0.14**	**0.16**	**0.19**
Northness	-0.03	-0.06	-0.04	0.00	0.03	0.05	0.02	0.07	0.07	0.05
Eastness	0.02	0.02	0.01	0.07	0.03	0.01	0.09	0.01	0.00	0.03
Horizontal curvature	-0.07	-0.05	-0.07	-0.09	-0.04	-0.05	-0.03	-0.03	-0.02	-0.02
Vertical curvature	-0.01	-0.02	0.01	-0.02	-0.06	-0.06	-0.02	-0.02	-0.01	0.00
Mean curvature	-0.05	-0.05	-0.04	-0.07	-0.06	-0.06	-0.02	-0.02	-0.01	0.00
Minimal curvature	0.07	0.01	0.02	0.00	0.02	0.00	0.07	0.01	0.04	0.09
Maximal curvature	**-0.14**	-0.08	-0.08	**-0.12**	-0.06	-0.06	-0.02	-0.02	-0.01	0.00
Gaussian curvature	0.00	-0.01	0.00	-0.03	0.00	0.00	0.01	0.00	0.00	0.01
Shape index	0.04	0.00	0.01	-0.01	-0.08	-0.07	-0.02	-0.02	-0.01	0.00
SySI blue	**-0.63**	**-0.40**	**-0.36**	**-0.38**						
SySI green	**-0.60**	**-0.40**	**-0.36**	**-0.40**						
SySI red	**-0.47**	**-0.37**	**-0.33**	**-0.42**						
SySI NIR	**-0.48**	**-0.39**	**-0.36**	**-0.44**						
SySI SWIR1	**-0.51**	**-0.43**	**-0.36**	**-0.42**						
SySI SWIR2	**-0.49**	**-0.42**	**-0.36**	**-0.43**						
Mosaic blue					**0.33**	**0.29**	**0.28**	**0.18**	**0.18**	**0.20**
Mosaic green					**0.26**	**0.24**	**0.26**	**0.16**	**0.14**	**0.17**
Mosaic red					**0.18**	**0.16**	**0.21**	**0.11**	**0.11**	**0.12**
Mosaic NIR					0.05	**0.11**	**0.18**	**0.23**	**0.19**	**0.22**
Mosaic SWIR1					**0.26**	**0.24**	**0.30**	**0.29**	**0.34**	**0.32**
Mosaic SWIR2					**0.20**	**0.17**	**0.24**	**0.20**	**0.24**	**0.25**

Where: AHem = hematite amplitude; AGt = goethite amplitude; AKt = kaolinite amplitude; AGbs = gibssite amplitude. The bold values were the correlation >0.10 and <-0.10.

Table A2: Accuracy parameters for Synthetic soil image with full coverage (SySIc) modelling.

SySI band	Calibration set			Validation set		
	R^2	RMSE	RPIQ	R^2	RMSE	RPIQ
blue	0.35	174.2	1.54	0.28	181.6	1.48
green	0.36	202.6	1.55	0.29	211.3	1.49
red	0.32	229.1	1.52	0.25	239.2	1.49
NIR	0.32	333.6	1.62	0.25	348.7	1.56
SWIR1	0.46	511.0	1.72	0.36	557.3	1.58
SWIR2	0.38	429.3	1.49	0.32	450.1	1.42

Where: R^2 = coefficient of determination; RMSE = root mean square error; RPIQ = ratio of the performance to interquartile distance.

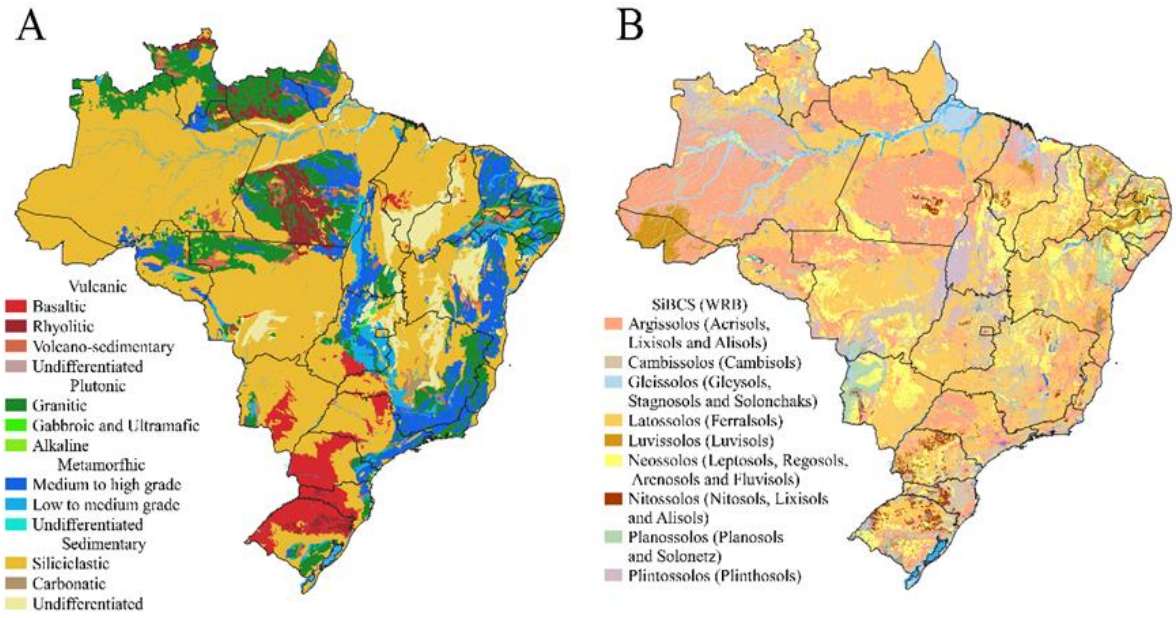

Figure A1: Geologic (adapted from Gómez et al. 2019) (A) and soil (Instituto Brasileiro de Geografia e Estátistica 2021) (B) maps for of Brazilian territory.

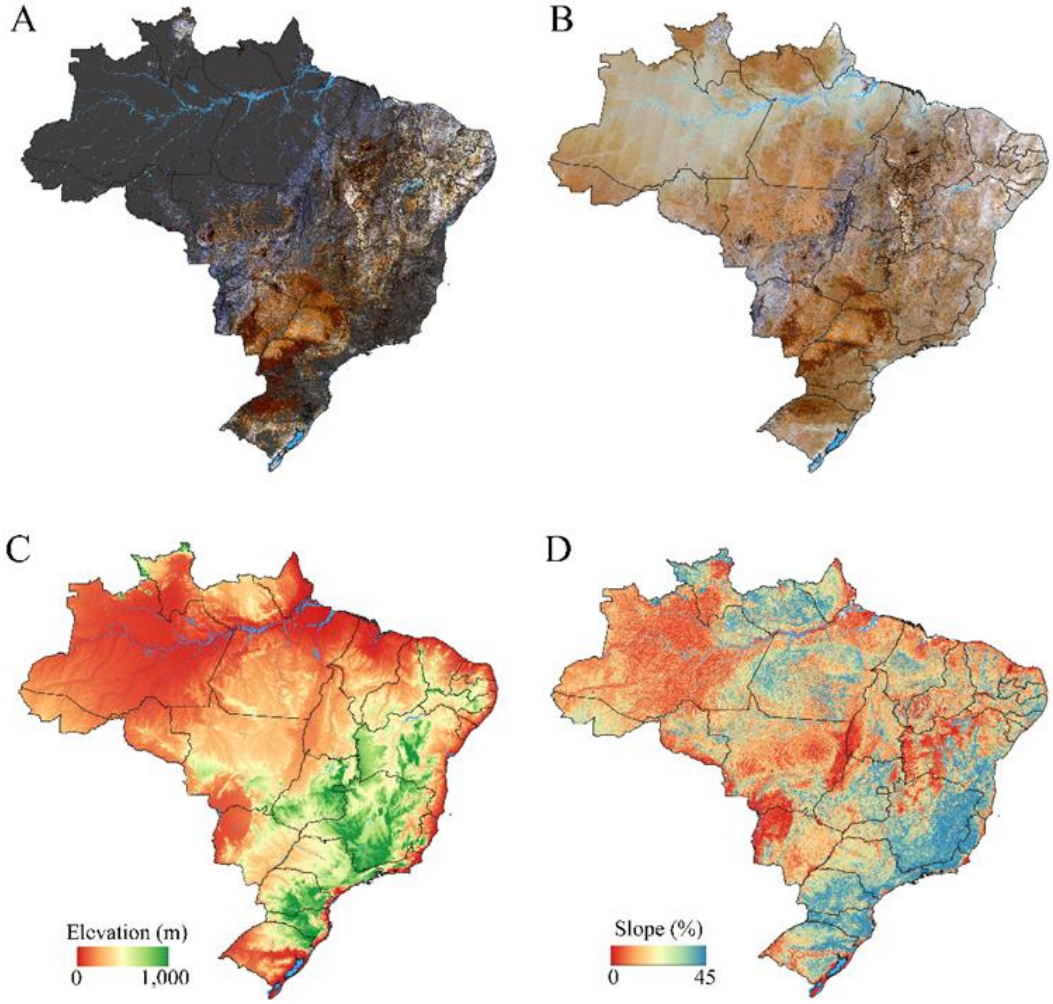

Figure A2: Soil synthetic image (SySI) (A), soil synthetic image with full coverage (SySIc), elevation (B) and slope (C). The SySI and SySIc are presented in true color (3,2,1) RGB composition.

4. FINAL REMARKS

The mineralogy always was neglected in analyses for soil characterization. In the present thesis we successfully explored to potential of proximal and remote sensing to reach it. In the first chapter, it was showed the close relation between the pXRF information and soil mineralogy and the potential of this technique to reach several soil attributes beyond the elemental composition. In the second chapter, it was developed a proximal and remote sensing integrated framework to map the mineralogy in the whole Brazilian territory. The Vis-NIR-SWIR spectra was used to calculate the abundances of soil minerals and the remote sensed data to spatialize it via digital soil mapping (DSM).

Briefly, the main conclusions were:

1) The soil particle size distribution, water and iron forms implied in changes imp XRF information by different elemental composition of samples and matrix effects caused by it, which is related to soil mineralogy and weathering processes and pave the way for determination of several soil attributes.

2) The pXRF demonstrated to be a very useful tool for mineralogical and chemical studies, delivering satisfactory results for Kt, Gb, Fe_2O_3, Al_2O_3, SiO_2, TiO_2 and MnO contents estimation by LR. Besides that, the qualitative and Person's correlation analyses showed relationships between the pXRF data and both soil particle size distribution and soil fertility attributes.

3) The understanding of the fundamentals of interactions between pXRF energy with the sample matrix is the starting point for characterizing the soil mineralogy through this technique.

4) The relative abundance of soil minerals presented significant correlation with spectral data of a synthetic soil image (SySI) with bare soil pixels obtained from temporal Landsat images.

5) The spatial distribution of minerals was accessed for the entire Brazilian territory. The hematite (Hem) presented the more accurate results in spatial prediction with R^2 ranging from 0.48 to 0.56, followed by gibbsite (Gbs) (0.42 to 0.44), Kt (0.20 to 0.31) and goethite (Gt) (0.16 to 0.26). The mineral maps obtained were in accordance with geology and soil legacy maps and also with the climate and terrain conditions at national, regional and farm level.

6) The use of Vis-NIR-SWIR data combined with DSM was able to reveal the distribution of mineral amounts in the Brazilian territory and consists of an efficient method to obtain mineralogy information for large areas.