

Integrating GIS and remote sensing for soil attributes mapping in degraded pastures of the Brazilian Cerrado

Rômulo Oliveira Louzada^{a,*}, Ivan Bergier^b, Édson Luis Bolfe^b, Jayme Garcia Arnal Barbedo^b

^a Instituto de Meio Ambiente de Mato Grosso do Sul, Campo Grande, MS, Brazil

^b Embrapa Digital Agriculture, Campinas, SP, Brazil

ARTICLE INFO

Keywords:

Multisensor integration
Random Forest
Soil texture
Subsurface soil layers
Statistical summarization

ABSTRACT

Remote sensing is a crucial tool for soil assessment, yet soil complexity and sensor limitations hinder accurate analysis. This study integrates active and passive remote sensing data with Machine Learning (ML) methods to predict the physicochemical properties of degraded sandy soils in the Brazilian Cerrado. The 1197 ha area was divided into management zones. Soil samples were collected from each management zone at 0–0.2 m and 0.2–0.4 m depths. The samples were then bulked ($n = 99$) and analyzed for texture (clay, silt, sand), pH, soil organic matter (SOM), cation exchange capacity (CEC), effective CEC (ECEC), base saturation (V), and macro- and micronutrients (e.g., Ca, Mg, K, Fe, Mn). Composite samples from management zones, were matched with 128 orbital variables from Sentinel-1, Sentinel-2 (2023), and ALOS-PALSAR-1. The variables include spectral bands, vegetation and soil indices, gray-level co-occurrence matrices (GLCM), backscatter coefficients, polarimetric decompositions, and topographic indices. A key innovation was evaluating statistical metrics beyond the mean—such as medians, sums, and variances—within MZs. The models were processed using Random Forest (RF), with variable selection assessed via the Boruta algorithm. The tested approaches included (T1) RF with mean-based variables, (T2) RF + Boruta, (T3) RF with the highest correlation metrics, and (T4) RF + Boruta with correlation-based metrics. Results showed that Boruta-enhanced models (T2 + T4) improved performance in 89 % of cases. Correlation-based metrics (T3/T4) were more effective in 72 % of models than mean-based approaches (T1/T2). The best models demonstrated high accuracy for clay ($R^2 = 0.81$; RMSE = 25.2 %), CEC (0.73; 23.6 %), silt (0.71; 44.7 %), and K (0.62; 44.3 %) in the 0–0.2 m layer. In the 0.2–0.4 m layer, top-performing attributes included clay ($R^2 = 0.86$; RMSE = 19.1 %), sand (0.78; 10.6 %), silt (0.76; 39.3 %), and SOM (0.68; 21 %). Elevation and GLCM metrics emerged as key predictors across depths. These findings highlight the effectiveness of integrating diverse remote sensing data with ML for soil attributes mapping, particularly for clay and CEC.

1. Introduction

Remote sensing (RS) has become an indispensable tool for assessing soil physical and chemical properties, providing cost-effective, large-scale insights into landscape variability (Diaz-Gonzalez et al., 2022; Wang et al., 2023). Its applications extend to optimizing field sampling, evaluating soil degradation, and analyzing land cover, all critical for environmental monitoring and sustainable agricultural management (Bégué et al., 2018; Radočaj et al., 2024). Achieving reliable results, however, requires integrating diverse predictor variables that capture the complexities of soil attributes across the electromagnetic spectrum (Yüzügüllü et al., 2024). The patterns captured by RS are often

challenging to interpret and model without advanced machine learning (ML) algorithms (Padarian et al., 2020).

ML algorithms are particularly effective in uncovering complex relationships between soil attributes and environmental variables, improving the precision of soil characterization and mapping (e.g., Khanal et al., 2018; Emadi et al., 2020; Folorunso et al., 2023). Among these, Random Forest (RF) excels at managing non-linear relationships, handling complex datasets, and reducing overfitting risks (Pichler and Hartig, 2023). Integrating RF with variables selection techniques, such as Boruta, further enhances predictive accuracy by isolating the most relevant predictors (Moradpour et al., 2023; Bouslihim et al., 2024). While soil properties are often inferred using average values of orbital

* Corresponding author.

E-mail address: romullo_louzada@yahoo.com.br (R.O. Louzada).

<https://doi.org/10.1016/j.soilad.2025.100044>

Received 7 January 2025; Received in revised form 23 March 2025; Accepted 23 March 2025

Available online 27 March 2025

2950-2896/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

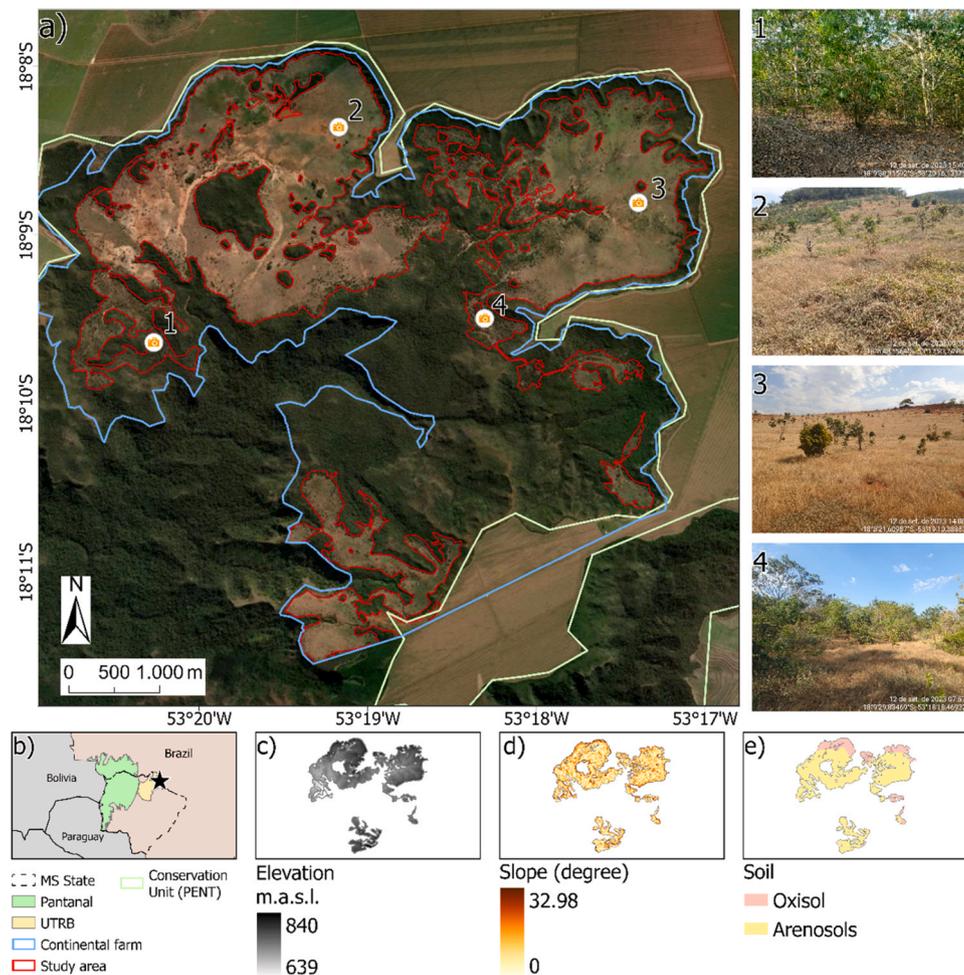


Fig. 1. Map of the study area (a) within the Continental farm, State of Mato Grosso do Sul. The area belongs to the Conservation Unit of the Nascentes do Rio Taquari State Park, within the Upper Taquari River Basin (UTRB) (b). The other three maps are elevation (c), slope (d), and soil types (e). Numbers 1–4 represent some photos of georeferenced landscapes throughout the study area.

variables, this approach may oversimplify field variability. Exploring alternative statistical summarizations—such as medians, ranges, and variances—can provide deeper insights into soil heterogeneity (Argento et al., 2021; Kerry et al., 2024).

RS provides the foundation for identifying spatial patterns, ML offers a robust analytical framework, and Geographic Information System (GIS) tools integrate these insights into spatially explicit maps, enabling the analysis of spatial relationships (Sangeetha et al., 2024). Spatial soil data is as crucial as, if not more than, numerical data. For example, identifying fertility gradients, erosion-prone areas, and compacted zones enables targeted and efficient land management practices (Techen et al., 2020). Consequently, a significant body of work has incorporated soil attributes into GIS and RS data layers to spatially model variability and enable predictions (Khanal et al., 2018; Yüzügüllü et al., 2020; Singh and Sarma, 2023).

In digital agriculture, soil zonation is fundamental for defining fertilization strategies. Management zones (MZs)—homogeneous spatial units based on soil and landscape characteristics—have gained prominence for optimizing field operations and agrochemical applications. MZs allow practices tailored to localized soil properties, such as texture and cation exchange capacity (Nawar et al., 2017; Belal et al., 2021). The integration of RS and GIS has significantly improved the delineation of MZs by capturing spatial and temporal variability in soil and vegetation attributes, guiding data-driven decision-making (Georgi et al., 2018). Various clustering techniques, including k-means, fuzzy c-means, and self-organizing maps, have been employed to define MZs, often

using spectral indices such as NDVI and OSaVI. These techniques demonstrate varying effectiveness depending on regional and dataset characteristics (Song et al., 2009; Javadi et al., 2022). However, the inherent variability of soils and landscapes challenges identifying a universally optimal clustering approach (Sosa et al., 2021; Navarro et al., 2023).

Compared to agriculture, livestock farming generally pays less attention to soil attributes. The use of MZs for fertilization planning in pastures remains uncommon, particularly in the Brazilian Cerrado—the country's primary livestock production biome (Oldoni et al., 2025). Several factors contribute to this, including the financial constraints of livestock farmers, fluctuations in exchange rates affecting livestock prices, and the low return on investment in soil management (Pereira et al., 2024). As a result, approximately 60 % of Brazil's pastures—totaling 109.77 million hectares—exhibit some degree of degradation (Bolfé et al., 2024). Studies highlight the ongoing decline in soil quality due to inadequate management practices such as insufficient pH correction and fertilization, the use of low-fertility seeds, and overgrazing (Locatelli et al., 2023). Poor management or land abandonment not only devalues pastures from an economic perspective but also triggers passive vegetation regeneration, often dominated by invasive species (Feltran-Barbieri and Féres, 2021). In both cases, RS data, particularly those utilizing vegetation indices, offer valuable tools for assessing pasture degradation (Oliveira et al., 2020) and monitoring the reestablishment of native Cerrado vegetation (Lewis et al., 2022; Louzada et al., 2023).

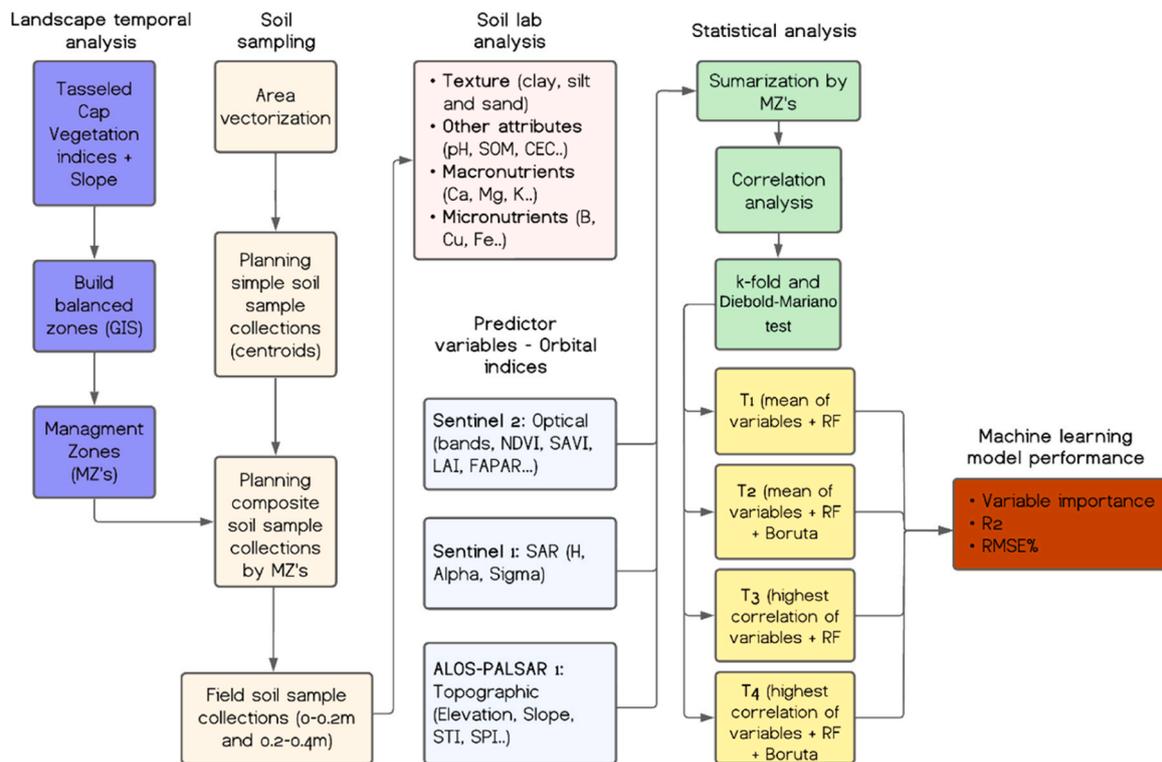


Fig. 2. Flowchart of the methodology.

Given the context of soil assessments using RS data, we present a case study of a degraded pasture undergoing passive vegetation restoration in the Brazilian Cerrado, specifically in the Central-West region. This large-scale restoration initiative, known as the “Sementes do Taquari” project, aims to implement soil management and conservation measures, facilitate passive vegetation recovery, and plant Cerrado tree species across approximately 1200 ha. Our assessment began with integrating a GIS-based approach to delineate management zones (MZs) using RS data, which guided soil sampling at two depths. We then collected Sentinel-1/2 variables and topographic features to test Random Forest (RF) models, evaluating variations in accuracy by applying the Boruta algorithm to predict soil physical and chemical properties within the MZs. This study aims to identify the most predictive orbital variables and statistical summaries, offering actionable insights for sustainable land management, soil restoration, and carbon sequestration efforts in areas undergoing early-stage vegetation recovery.

2. Materials and methods

2.1. Study area

The study area spans 1197 ha and was delineated using the Esri Vivid Maxar base map sensor dated 11/07/2023, with a spatial resolution of 0.31 m. Located within the Nascentes do Rio Taquari State Park Conservation Unit, it lies between latitudes 18°07' S and 18°11' S and longitudes 53°16' W and 53°20' W, in the northeastern region of Mato Grosso do Sul, Brazil (Fig. 1A). Known as “Continental Farm,” the area was acquired by the state government in 2022 after over four decades of extensive livestock farming. It is managed for soil and water conservation and ecological restoration under the “Sementes do Taquari” project, which serves as a proof of concept for rehabilitating degraded pastures and ravines. Detailed project information is available at [<https://storymaps.arcgis.com/stories/ef35a57241114b789de6725ea7f795bf>].

Situated in the Cerrado biome, the study area is part of the Upper Taquari River Basin (UTRB), a key source of the Pantanal drainage system. The region features plateaus and depressions, with a tropical

climate characterized by a dry winter season (Köppen classification) and an average annual rainfall of 1450 mm. Elevation range from 639 to 840 m above sea level (Fig. 1C), with slopes classified as smooth (3–8 %, covering 43 % of the area) and smooth-wavy (8–20 %, covering 57 % of the area) (Fig. 1D). The soils are predominantly arenosols (83 %), with dystroferric dark red oxisols comprising the remaining 17 %, primarily at the toes of slopes (Fig. 1E).

2.2. Flowchart of methodology

The methodology is summarized in Fig. 2 and encompasses soil collection, laboratory analysis, image processing to delineate the MZs, response variables, and statistical analysis using machine learning (ML).

2.3. Soil samples map

The mapping of soil samples began by dividing the study area into 1 ha × 1 ha polygons using the grid index tool in ArcGIS®PRO, generating 1238 polygons. Subsequently, these polygons' centroids were extracted, configuring our virtual sampling points (n = 1238) to guide field collections (see the virtual points map in Fig. S1). The following process was to generate the MZs, given the need to aggregate single samples into composite samples in regions with homogeneous characteristics. This technique is helpful in digital soil maps to reduce operational costs (collection and segregation) and especially laboratory analysis costs without losing variability (Flowers et al., 2005).

To create the MZs, we used the *build balance zones* tool in ArcGIS®PRO, which uses a genetic algorithm to create spatially contiguous zones (Patel and Padhiyar, 2010). In this method, we chose the criterion establishing the number of zones and the attribute target. For the number of zones, we set n = 120, with approximately 10 unique sampling points in each MZ.

In the attribute target section, the software allows the selection of layers (rasters or shapefiles) with weights that represent their importance for creating the zones. In our study, we included four Tasseled Cap Vegetation Indices (TC_VEG) (Bannari et al., 1995) from Sentinel 2 A

Table 1
List of physical attributes and chemical properties of the soil laboratory analysis.

Category	Attribute/element	Unit
Texture	Clay	%
	Silt	%
	Sand	%
Other attributes	pH	-
	Soil Organic Matter (SOM)	g/dm ³
	Cation Exchange Capacity (CEC)	cmol / dm ³
	Effective Cation Exchange Capacity (ECEC)	cmol / dm ³
	Saturation basis (V)	%
Macronutrients	Calcium (Ca)	cmol / dm ³
	Magnesium (Mg)	cmol / dm ³
	Potassium (K)	mg/dm ³
	Phosphorus (P)	mg/dm ³
	Sulfur (S)	mg/dm ³
Micronutrients	Boron (B)	mg/dm ³
	Cuprum (Cu)	mg/dm ³
	Iron (Fe)	mg/dm ³
	Manganese (Mn)	mg/dm ³
	Zinc (Zn)	mg/dm ³

Level 2 imagery (scene T22KBE), on the dates 2022–09–17, 2022–12–01, 2023–03–26, and 2023–05–15 (Fig. S2 to S5). In addition, we included the slope of the DEM obtained by the ALOS-PALSAR 1 sensor (12.5 m resolution) as the fifth layer (Fig. S6). The slope was weighted twice as much as the TC_VEG indices to account for its influence on soil variability (Lin et al., 2005).

2.4. Soil samples collection

Soil samples were collected by Coperplan Inc. (Dourados, Mato Grosso do Sul, Brazil) in August 2023 at 0–0.2 m and 0.2–0.4 m depths. The procedure consisted of a semi-mechanized collection using a drill and a motorcycle for transportation between MZs.

At the time of collection, some points selected in the original 120 MZs became inaccessible due to vegetation regeneration or obstacles such as erosions (gullies branches), requiring adjustments to the sampling plan (final map in Fig. S7). Therefore, these MZs were excluded from the final sampling, which consisted of 1016 points of single soil samples grouped into 99 MZs or composite samples. On average, we obtained 10.26 single samples per MZ, as demonstrated in the histogram in Fig. S8. The samples within a MZ were mixed, and the composite sample was stored in a clean plastic bag. Some georeferenced photographs are presented in Fig. S10, and the geographic coordinates of the single collection points are presented in Table S1.

2.5. Soil laboratory analysis

Laboratory analyses were conducted at Sinergia Laboratório de Análises Agronômicas Inc. using standardized methods outlined by EMBRAPA (Teixeira et al., 2017). The granulometric analysis involved the separation of the sand, clay, and silt fractions. The pH was measured one hour after mixing 10 g of soil with 25 ML of CaCl₂ 0.01 mol L⁻¹. The soil organic matter (SOM) content was determined by the difference in weight after drying in an oven and subsequent incineration in a muffle furnace at 600°C. The soil's phosphorus (P) content was extracted using the Mehlich-1 solution (0.05 mol L⁻¹ HCl and 0.0125 mol L⁻¹ H₂SO₄) and quantified by ultraviolet spectrometry. Exchangeable potassium (K⁺) was determined by flame spectrophotometry. For the exchangeable bases calcium (Ca²⁺), magnesium (Mg²⁺), and aluminum (Al³⁺), extraction was performed with KCl 1 mol L⁻¹, followed by volumetric analysis. The concentration of H⁺ cations was measured using 1 mol L⁻¹ calcium acetate at pH 7. Both exchangeable bases were used to calculate the cation exchange capacity (CEC) by the sum of the elements K⁺, Ca²⁺, Mg²⁺, Al³⁺, and H⁺, and the effective cation exchange capacity (ECEC) excluding the H⁺ cation. The basis saturation (V%) was calculated by the formula: $((K^+ + Ca^{2+} + Mg^{2+}) \times 100) \div CEC$. Finally,

sulfur (S) was calculated using the turbidimetric method with barium chloride (BaCl₂) in macronutrients.

The micronutrients Fe, Mn, Zn, and Cu were analyzed by flame atomic absorption spectrometry, following extraction with Mehlich-1 solution. Boron (B) was extracted from 20 g of soil, subjected to boiling water for 5 minutes, added 0.1 mol L⁻¹ CaCl₂, and filtered through filter paper before analysis.

2.6. Orbital variables

Soil attributes were evaluated using 128 orbital variables categorized into optical, synthetic aperture radar (SAR), and topographic indices.

2.6.1. Optical

Optical data, including 118 variables, were derived from Sentinel 2 A imagery (T22KBE, 08/28/2023) using SNAP 9.0 and ArcGIS®PRO (Table 2). In addition to 10 bands of Sentinel 2, we selected soil and vegetation indices including the Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Transformed Soil Adjusted Vegetation Index (TSAVI), Modified Soil Adjusted Vegetation Index (MSAVI). The soil radiometric indices of Brightness Index (BI), Color Index (CI), and Redness Index (RI). Tasseled cap brightness (TC_BRI), Tasseled Cap Soil Brightness (TC_SOIL_BRI), Tasseled Cap Vegetation (TC_VEG), and Tasseled Cap Wetness (TC_WET) as the Tasseled Cap indices. Principal component analysis (PCA) and biophysical indices of Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Fraction of Vegetation Cover (FCOVER), and Leaf Area Index (LAI). Geology indices, including FERROUS_IRON, FERROUS_OXID, and FERROUS_SILICATES, were also included. In addition, we include gray-level co-occurrence matrices (GLCM) or texture variables for each band, being angular second momentum (ASM), contrast (CON), dissimilarity (DIS), entropy (ENT), correlation (CORR), MEAN, variance (VAR), homogeneity (HOM), and maximum probability (MAX). For the optical texture variables, we resampled to 10 m resolution (B2), with a 9 × 9 window size and angles of 0°, 45°, 90°, and 135° in the SNAP 9.0 software.

2.6.2. Synthetic Aperture Radar (SAR)

The SAR variables (n = 4) were processed in the SNAP 9.0 software and divided into two groups (Table 3), where H (entropy) and Alpha represent the polarimetric decomposition and backscattering (sigma) into VV (SIG_VV) and VH (SIG_VH) polarizations, completing the intensity group. The processing sequences followed the routines that Diniz et al. (2020) and Louzada et al. (2023) developed. In summary, for polarimetry, the process started with deburst, then multi-look, H-Alpha dual-pol decomposition, and terrain correction. In backscattering, the orbit was applied, followed by the thermal noise removal phase, Sigma 0 calibration, deburst, multi-look, application of the speckle filter, and finally, the terrain correction.

Polarimetry H: p_i expresses the appearance probability for each contribution; Backscattering intensity: DN_i is the digital number of the pixel i ; A is an absolute calibration constant.

2.6.3. Topographic indices

In addition to orbital variables related to soil and vegetation and radar-derived variables, terrain attributes can significantly enhance the modeling of soil properties (see examples in Schillaci et al., 2017; Hateffard et al., 2019). In this study, six terrain variables were computed (Table 4), beginning with Elevation, derived from the Digital Elevation Model (DEM) of the ALOS-PALSAR 1 sensor (dated 02/16/2011), obtained from the Alaska Satellite Facility website (<https://search.asf.alaska.edu/>). The second variable, Slope (in degrees), was processed in ArcGIS®PRO and served as the foundation for deriving four additional indices: Stream Power Index (SPI), Sediment Transport Index (STI), Terrain Roughness Index (TRI), and Topographic Wetness Index (TWI).

Table 2
Optical variables.

Category	Variable	Equation	Reference
Spectral bands	B2, B3, B4, B5, B6, B7, B8, B8A, B11, and B12	-	-
Soil and vegetation indices	NDVI	$(B8-B4) / (B8 + B4)$	Rouse et al., 1973
	SAVI	$(1 + L) * (B8-B4) / (B8 + B4 + L)$	Huete (1988)
	MSAVI	$\frac{2 * B8 + 1 - \sqrt{(2 * B8 + 1)^2 - 8 * (B8 - B4)}}{2}$	Qi et al. (1994)
	TSAVI	$s * (B8 - s * B4 - a) / (s * B8 + B4 - a * s + X * (1 + s * s))$	Baret and Guyot (1991)
Soil radiometric indices	BI	$\sqrt{B4^2 + B3^2} / 2$	Mathieu et al. (1998)
	CI	$(B4-B3) / (B4 + B3)$	Escadafal (1989)
	RI	$(B4)^2 / (B3)^2$	Barron and Torrent (1986)
Tasseled Cap indices	TC_BRI	$0.3037 * B2 + 0.2793 * B3 + 0.4743 * B4 + 0.5585 * B8 + 0.1863 * B12$	Crist and Cicone (1984)
	TC_SOIL_BRI	$0.332 * B3 + 0.603 * B4 + 0.675 * B6 + 0.262 * B8A$	Bannari et al. (1995)
	TC_VEG	$2848 * B2 - 0.2435 * B3 - 0.5436 * B4 + 0.7243 * B8 + 0.0840 * B11 - 0.1800 * B12$	Bannari et al. (1995)
	TC_WET	$0.1509 * B2 + 0.1973 * B3 + 0.3279 * B4 + 0.3406 * B8 - 0.7112 * B11 - 0.4572 * B12$	Crist and Cicone (1984)
Principal Components Analysis*	PCA	Calculation method in the SNAP 9.0. All bands were considered after resampling to 10 m spatial resolution.	-
Biophysical indices*	FAPAR	Calculation method in the SNAP 9.0. All bands were considered after resampling to 10 m spatial resolution.	Liang (2007)
	FCOVER		Kallel et al. (2007)
Geology indices	LAI		Price (1993)
	FERROUS_IRON	$B11/B8$	Rowan and Mars (2003)
	FERROUS_OXID	$B4/B2$	Henrich et al. (2011)
Texture*	FERROUS_SILICATES	$B11/B12$	Henrich et al. (2011)
	Band_ASM	$\sum_i \sum_j \{p(i,j)\}^2$	Haralick et al. (1973)
	Band_CON	$\sum_{n=0}^{N_k-1} n^2 \left\{ \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} p(i,j) \right\}, i-j = n$	
	Band_DIS	$\sum_{n=1}^{N_k-1} n \left\{ \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} p(i,j)^2 \right\}, i-j = n$	
	Band_ENT	$-\sum_i \sum_j p(i,j) \log(p(i,j))$	
	Band_CORR	$\frac{\sum_i \sum_j (i,j) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	
	Band_MEAN	$\sum_{i=2}^{2N_k} (ip_{(x+y)}(i))$	
	Band_VAR	$\sum_i \sum_j (i - \mu)^2 * p(i,j)$	
Band_HOM	$\sum_i \sum_j \frac{p(i,j)}{1 + (i-j)^2}$		
Band_MAX	$MAX_{i,j} p(i,j)$		

Terms of the above equations: B2 (blue), B3 (green), B4 (red), B5 (red edge), B8 (near infrared), B6, B7 and B8A (vegetation red edge), B11 (short infrared wave 1) and B12 (short infrared wave 2); SAVI: L is the adjustment factor; TSAVI: a is the soil line intercept; s is the soil line slope; X is the adjustment factor to minimize soil noise; Texture: $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the means and standard deviations for the row- and column-marginal probabilities of the co-occurrence matrix composed of $p(i, j)$. * Pre-processing by resampling for 10 m pixels for bands B5, B6, B7, B8A, B11 and B12.

Table 3
Radar variables.

Polarimetry	H	$-\sum_{i=1}^3 p_i \log_3 p_i \cdot p_i = \frac{\lambda_i}{\sum_{j=1}^3 \lambda_j}$	Moreira et al. (2013)
	Alpha	$\arccos(e_{rl})$	
Intensity	SIG_VV	DN_i^2 / A_i^2	Diniz et al. (2020)
	SIG_VH		

2.7. Statistical analysis

The response variables were summarized because the sampling units comprised polygons formed by aggregating individual soil samples. Initially, the first statistical moment (mean or average) was selected as the key descriptive statistic due to its recurrence in studies of soil

Table 4
Terrain variables.

Topographic indices	Elevation	-	-
	Slope	-	-
	SPI	$A_s \tan \beta$	Wilson and Gallant (2000)
	STI	$\left(\frac{A_s}{22.13}\right)^{0.6} \left(\frac{\sin \beta}{0.0896}\right)^{1.3}$	
	TRI	$Z \left[\sum (x_{i,j} - x_{00})^2 \right]^{0.5}$	
	TWI	$\ln \left(\frac{A_s}{\tan \beta} \right)$	

SPI: A_s is the catchment area, and β is the local slope; TRI: Z is the elevation at the central cell.

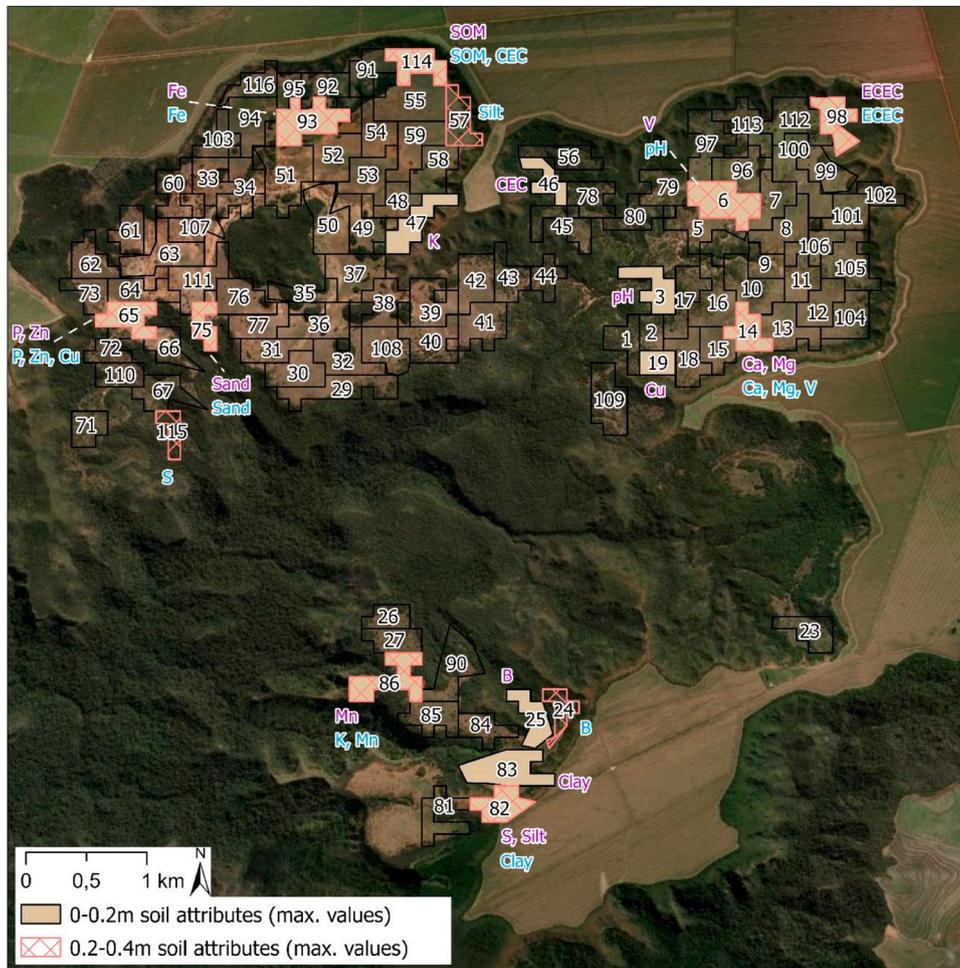


Fig. 3. Spatial distribution of the 99 Management Zones (MZs) and the maximum value per layer for each soil attribute. The purple and blue colors indicate the elements with maximum values for the 0–0.2 m and 0.2–0.4 m layers, respectively.

properties or crop data using remote sensing (RS) techniques (Pôças et al., 2015; Tunçay et al., 2021; Burns et al., 2022). In addition to the mean, six other descriptive statistics were evaluated: maximum (MAX), minimum (MIN), range (interval), median, standard deviation (SD), and the 90th percentile (PCT90), which represents values within a 90 % confidence interval. These seven metrics were derived using the zonal statistics as a table tool in ArcGIS®PRO. Subsequently, Spearman’s correlation analysis was performed to evaluate the relationship between soil attributes and statistical metrics for each orbital variable.

The analytical tests were divided into four scenarios: T1, using the mean of orbital variables as predictors in the Random Forest (RF) model (see Section 2.8); T2, incorporating the mean alongside the joint application of the RF model and the Boruta algorithm; T3, employing the statistical metric with the highest correlation per orbital variable as input for the RF model; and T4, integrating the highest correlated metric between RF and Boruta.

2.8. Random forest regression

Here, we used the Random Forest (RF) model for its ability to capture complex relationships and robustness to outliers and autocollinearity (Sheykhoumousa et al., 2020). Several works have used this machine learning method in soil nutritional assessments with RS (Taghizadeh-Mehrjardi et al., 2021; Siqueira et al., 2024). Model processing was performed in R Software for Statistical Computing v.4.3.3 (R Core Team, 2013) with the randomForest and caret packages. 70 % of the samples (n = 69) were spared for training and 30 % (n = 30) for

validation (testing) following other studies applying RF models (see examples in Wang et al., 2022; Amankulova et al., 2024; Fu et al., 2024). The RF models were fitted using a fixed *ntree* of 500 and an *mtry* ranging from 1 to 11, representing the integer closest to the square root of the maximum number of variables (n = 128). The analysis of the optimal *mtry* was performed by k-fold cross-validation, considering 10 folds. To assess overfitting, we also used the Diebold-Mariano test, which analyzes whether there is a significant difference (p > 0.05) between the root mean square error (RMSE) of the training and test.

To evaluate variable importance, we considered the increase in mean squared error (%IncMSE) and the increase in node purity (IncNodePurity). Model performance was assessed using the coefficient of determination (R²) and percentage root mean square error (RMSE%), according to Eqs. (1) and (2), respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{1}$$

$$RMSE\% = \left(\frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\bar{y}} \right) * 100 \tag{2}$$

where y_i are the observed values, \hat{y}_i are the values predicted by the model, \bar{y} is the average of the observed values, and n refers to the number of observations.

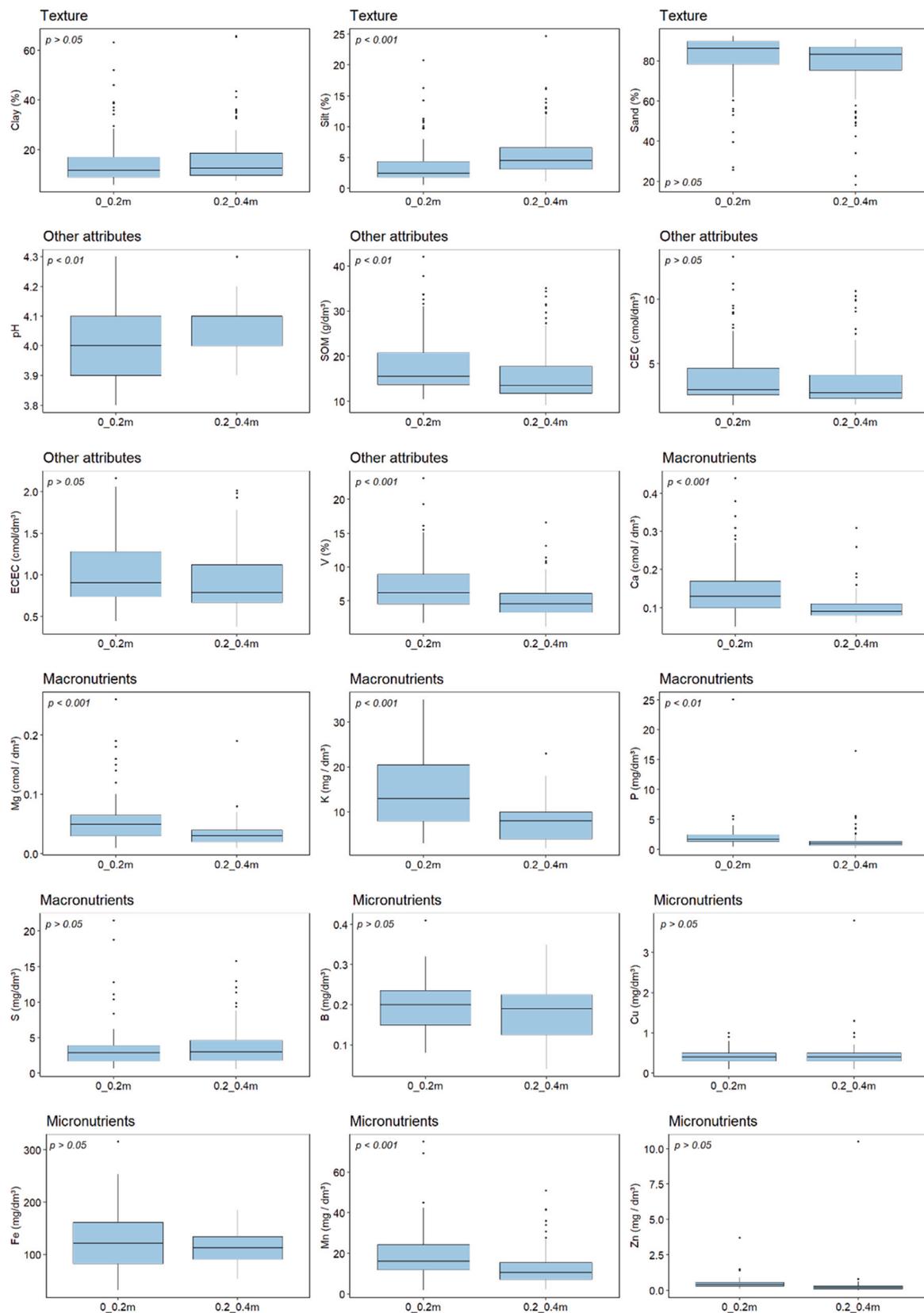


Fig. 4. Box plots of laboratory soil analyses in layers 0–0.2 m and 0.2–0.4 m with the p-value results in the Kruskal-Wallis and Mann-Whitney paired tests.

Beyond the results of the RF model using all variables, further optimization was pursued through the Boruta algorithm in R—a feature selection method designed to identify the most relevant predictors (Zixian et al., 2021). This analysis compares the original variables

against random variables, known as "shadow features" generated by the model. Variables were classified into three categories based on their importance: (1) Accepted, if the importance of the original variable exceeds that of the most important shadow feature; (2) Rejected, if the

Table 5
Spearman correlation analyses of RS variables with soil attributes at each depth.

Soil parameter	0–0.2 m				0.2–0.4 m			
	Correlation with the mean	Highest correlation obtained	Statistics with the highest correlation obtained	Frequency (%)	Correlation with the mean	Highest correlation obtained	Statistics with the highest correlation obtained	Frequency (%)
Clay	−0.07	0.24	SD	22	−0.08	0.25	MAX	20
Sand	0.07	0.24	SD	22	0.08	0.26	SD	20
Silt	−0.05	0.21	SD	24	−0.08	0.26	MEDIAN	21
pH	0.06	0.25	SUM	47	0.03	0.18	PCT90	26
SOM	−0.09	0.25	MAX	29	−0.11	0.29	MAX	30
CEC	−0.09	0.27	MAX, MIN and RANGE	19	−0.07	0.22	MAX	27
ECEC	−0.08	0.24	RANGE	18	−0.1	0.28	MAX	26
V	−0.08	0.24	RANGE	18	0.04	0.17	RANGE	23
Ca	0.02	0.2	SUM	55	0.01	0.16	MIN	30
Mg	−0.04	0.19	SUM	23	−0.05	0.21	MIN	27
K	−0.08	0.26	PCT90 and MEDIAN	21	−0.08	0.26	MEDIAN	24
P	0.08	0.24	MAX	23	0.07	0.21	MIN	22
S	−0.02	0.1	MEDIAN	20	0.03	0.12	MEDIAN	27
B	−0.04	0.14	RANGE	24	−0.04	0.13	MEDIAN	22
Cu	−0.04	0.19	PCT90	22	−0.03	0.19	MEDIAN and SD	19
Fe	−0.07	0.24	SD	30	−0.03	0.16	SUM	36
Mn	0	0.16	SUM	52	0	0.17	SUM	47
Zn	−0.02	0.16	SUM	20	0	0.18	MEDIAN	31

variable's importance is significantly lower than the maximum shadow feature or comparable to the minimum shadow; and (3) Provisional, if the model cannot definitively classify the variable after the specified number of iterations (Kursa and Rudnicki, 2010). Typically, only variables classified as accepted by Boruta are included in the final RF model.

3. Results

3.1. Soil fertility analysis results

Fig. 3 shows the spatial distribution of the 99 MZs and the maximum values for each soil attribute across layers. Most of the elements presented maximum values in the MZs on the edge of the study area, coinciding with the most clayey and highest altitude soils, such as units 14, 98, 114, 97, 53, 82, and 83. Furthermore, many elements were associated in which the highest values were found in the same MZ, for example, for Fe (MZ=93), SOM (114), ECEC (98), Ca and Mg (14), and P and Zn (65).

The laboratory results of the samples by MZ are presented in the box plots of Fig. 4 (see the results of the descriptive statistics in Table S2). Half of the attributes ($n = 9$) showed no statistical difference between soil layers ($p > 0.05$), including clay and sand, CEC, ECEC, and the micronutrients B, Cu, Fe, and Zn. On the other hand, the pH, SOM, and P ($p < 0.01$), and silt, Ca, Mg, K, V, and Mn showed even more significant variation ($p < 0.001$). In 78 % of the analyses, the median of the surface layer 0–0.2 m was higher than the subsurface layer 0.2–0.4 m, with excess clay and silt in the texture. Also, pH and S. Furthermore, some characteristics draw attention, such as the high percentage of sand in both soil profiles had a high sand content (79.34 ± 14.04 %) and low values for pH (4.04 ± 0.1), CEC (3.85 ± 2.26 cmol/dm³), ECEC (0.98 ± 0.38 cmol/dm³), and V (6.07 ± 3.34 %).

3.2. Remote sensing variables

The Spearman correlation analysis between soil attributes and orbital variables showed that the mean statistic was not the best predictor for most soil attributes (see all Spearman correlation results from Tables S3 to S37). The SUM and SD metrics showed the strongest correlations with soil properties overall, whereas the MEDIAN and MAX metrics were more effective in the 0.2–0.4 m soil layer (Table 5).

The application of the Boruta algorithm significantly improved variable selection. For instance, as shown in Fig. 5, the number of variables

was reduced from 128 to $n = 6$ confirmed variables in T2 and $n = 10$ in T4 for clay across both soil depths. While most variables were rejected (red), a select few were confirmed (green), highlighting the effectiveness of Boruta in narrowing down key predictors.

Most variables selected in each test presented importance below $RMSE\% < 10$ (see full results in Tables S39 and S40). In Fig. 6, we show only the most relevant ones with $RMSE\% > 10$, highlighting the significant influence of the Elevation variable on the granulometric components (clay-silt-sand), in addition to K, SOM, and CEC. In the background, the B12_CON in Ca, Mg, and other textural techniques such as variance, mean, and correlation of bands 8, 8 A, and 11 for the attributes Ca, CEC, and ECEC.

Diebold-Mariano tests indicated that there was a significant difference in RMSE ($p < 0.05$) between the training and testing data only for the micronutrients Cu, Mn, and Zn in the 0–0.2 m layer in tests T3 and T4 (full results in Table S41). This result suggests possible overfitting for these elements.

The RF models enhanced by Boruta (T2 and T4) outperformed the standalone RF models (T1 and T3), showing higher R^2 values and lower $RMSE\%$ (see Table 6). The only exception was for Fe in T3 at the 0–0.2 m layer, where Boruta did not improve model performance. Similarly, in the 0.2–0.4 m layer, 15 out of 18 soil attributes showed performance gains with the method.

The best performances per layer were clay in T2 at 0–0.2 m, reaching $R^2 = 0.81$ and $RMSE\% = 25.19$ %, and the same attribute in 0.2–0.4 m in the T4 test with $R^2 = 0.86$ and $RMSE\% = 19.15$ %. Other promising results in the surface layer were Silt (T4; $R^2 = 0.71$, $RMSE\% = 44.73$), CEC (T2; $R^2 = 0.73$, $RMSE\% = 23.63$) and K (T4; $R^2 = 0.62$, $RMSE\% = 44.3$). In the subsurface layer again Silt (T2; $R^2 = 0.76$, $RMSE\% = 39.32$) and CEC (T4; $R^2 = 0.64$, $RMSE\% = 32.11$), in addition to Sand (T4; $R^2 = 0.78$, $RMSE\% = 10.59$) and SOM (T4; $R^2 = 0.68$, $RMSE\% = 21.01$).

Box plots of R^2 and $RMSE\%$ distributions (Fig. 7) showed significantly higher predictive accuracy for R^2 in T4 compared to T1 in the 0–0.2 m layer ($p < 0.001$), based on the Kruskal-Wallis test and Mann-Whitney pairwise comparisons at both depths. However, $RMSE\%$ values were not statistically different ($p > 0.05$) between tests and layers. The figure also shows that the attributes clay, sand, silt, K, and CEC, which demonstrated promising performance for R^2 , had moderate $RMSE\%$ values between the first and third quartiles for tests T2 and T4 at both depths.

Comparative maps of actual field data and the best model predictions for clay are shown in Fig. 8. The spatial contrasts in clay percentage are

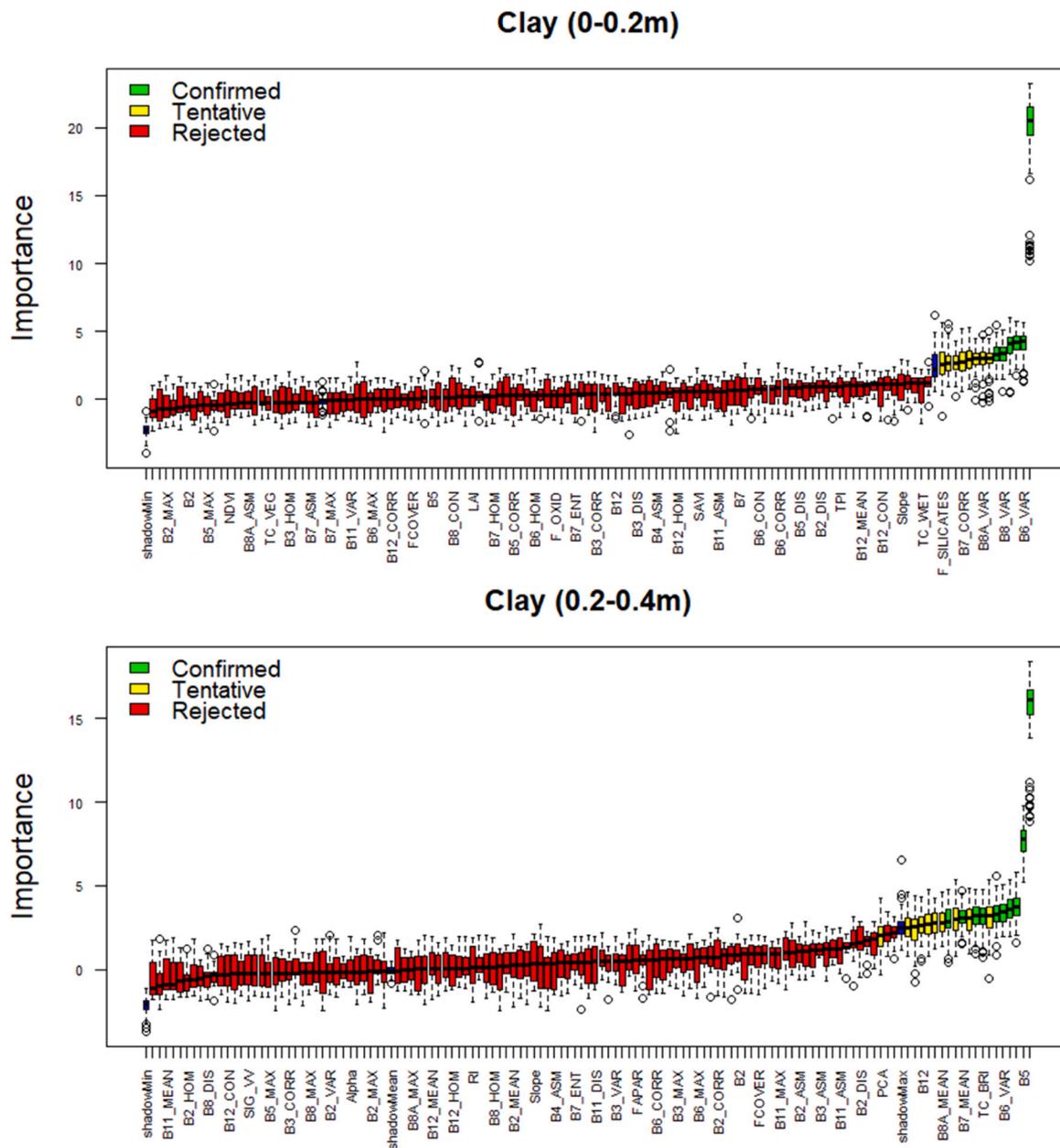


Fig. 5. Examples of variable importance graph generated by the Boruta algorithm for the parameter clay in test T2 (0–0.2 m) and T4 (0.2–0.4 m). The blue color of the graph refers to the limits (shadows) for the algorithm’s rejection/acceptance of variables. F_SILICATES is the abbreviation for FERROUS_SILICATES, and for all orbital variables, see the list in Section 2.6.

subtle; however, the predicted maps for test T2 in the 0–0.2 m layer (Fig. 8b) and test T4 in the 0.2–0.4 m layer (Fig. 8d) indicate that the models were more conservative, showing slightly lower variation in clay across the MZs. This effect was particularly noticeable along the edges of the study area, where the original soil has already demonstrated a higher proportion of this granulometric fraction.

4. Discussion

4.1. Key findings from remote sensing

Direct soil responses from orbital remote sensing present a methodological challenge due to the influence of surface cover (Balsamo et al., 2018). In this study, conducted in a passive restoration environment of Cerrado vegetation, we achieved high predictive performance, with R² values exceeding 0.8 and moderate RMSE% (<25 %) for clay content in

both layers. Additionally, we obtained promising results for Silt, Sand, CEC, K, and SOM. We also observed that the Boruta algorithm enhanced the predictive performance of RF models, aligning with findings from similar studies (Keskin et al., 2019; Peng et al., 2023). However, our correlation analysis of descriptive statistical measures proved essential for achieving the best results. Tests T3 and T4 outperformed T1 and T2 in 72 % of the models, indicating that summarization using the mean may not fully capture the predictive potential of remote sensing data.

Our findings suggest that alternative statistical measures to the mean, such as the maximum, standard deviation, and median (applied in the T3 and T4 tests), often exhibit stronger correlations with soil properties analyzed in the laboratory. However, studies assessing the impact of statistical summarization methods for homogeneous soil regions on model performance remain scarce. For instance, Zhu et al. (2024) found that the geometric median provided greater robustness against outliers in soil organic carbon (SOC) models compared to the univariate mean

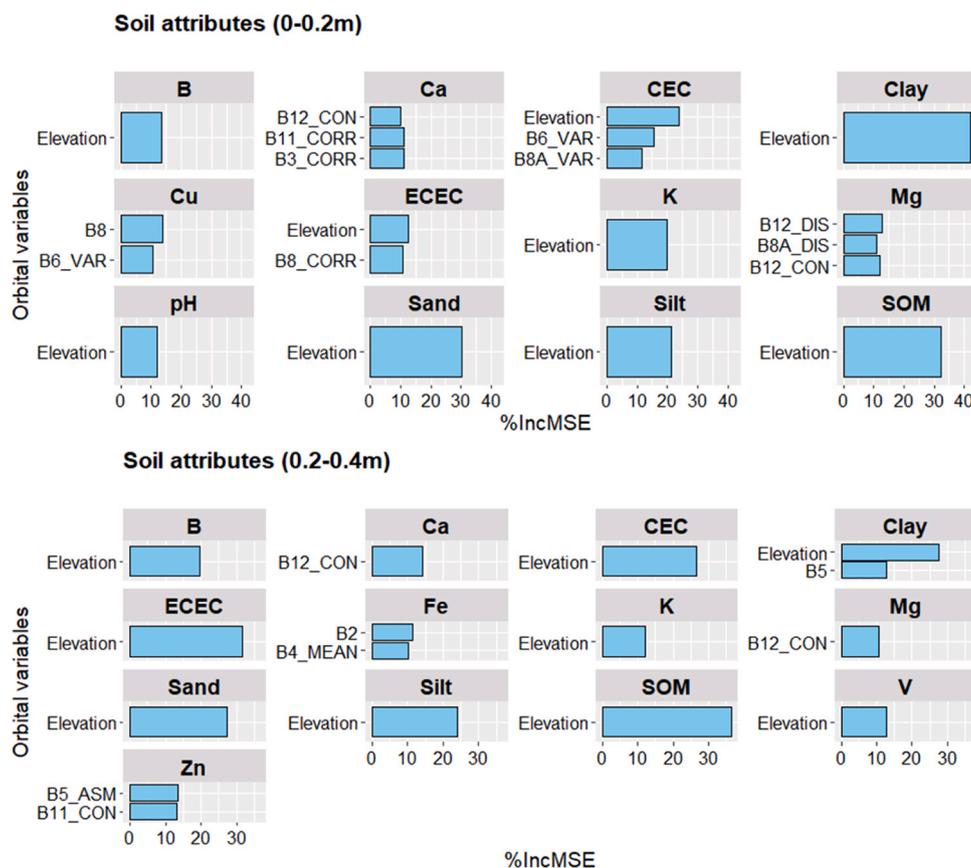


Fig. 6. Selection of variables with Mean Square Error Increment (%IncMSE) greater than 10% in importance in RF and Boruta in the best-performing tests, divided by soil attribute and depth. For abbreviations of orbital variables, see the list in [Section 2.6](#).

and median. These approaches are particularly valuable in agricultural systems, where soil attributes' spatial and temporal variability directly affects crop yields and resource use efficiency (Sishodia et al., 2020).

Aggregating soil samples in homogeneous regions is a common agricultural practice for the rational use of agrochemicals. Research in this field has explored clustering methods to delineate zones using RS data, including the k-means algorithm (Zeraatpisheh et al., 2022; Muniyammal VEDIAPPAN et al., 2025), density-based spatial clustering of applications with noise (DBSCAN) (Javadi et al., 2022), and the partitioning around medoids (PAM) method (Cammarano et al., 2020). Here, we introduce another option to this set of techniques: the Build Balanced Zones tool in ArcGIS®PRO. This GIS/RS tool remains unexplored mainly (see examples in other fields of knowledge in Nakai et al., 2021; DeBruin and Khani, 2024) and is novel in soil science. Our study processed polygons based on variable weights or importance, integrating spatial and temporal factors, including historical TC_VEG indices and terrain slope.

Comparing model performance by depth, we found that in the 0–0.2 m layer, the average R^2 and RMSE% were 0.42 and 37.37%, respectively, slightly higher than the 0.2–0.4 m layer, which had values of 0.40 and 51.44%. This suggests that the predictive power of orbital variables is likely linked to the direct response of soil reflectance in the surface layer (<20 cm), whereas deeper layers exhibit indirect responses influenced by surface cover (Sharma et al., 2022; Abdulraheem et al., 2023). Consequently, most multispectral sensor studies have focused primarily on topsoil assessments (e.g., Ma et al., 2021; Tunçay et al., 2021; Charishma et al., 2024). Research on parameters such as salinity content (Li et al., 2022), root zone soil moisture (Bartels et al., 2021), and SOM/SOC estimation (Reis et al., 2021; Bócoli et al., 2025) across different soil depths indicates that surface layers generally yield more accurate modeling results. However, some studies have reported good

predictive performance for SOC even at depths greater than 20 cm (Poppiel et al., 2019; Ren et al., 2020). In this context, our study further explored this potential by evaluating direct and indirect assessments using multispectral RS, SAR, and terrain variables across a broad range of soil attributes in the two main layers.

Regarding the variables, we highlight that the Sentinel-1 group did not provide a predictive advantage, contradicting expectations that C-band penetration would enhance soil property predictions (El Hajj et al., 2019). On the other hand, Elevation played a significant role in the models, with %IncMSE values exceeding 20% for clay, sand, silt, SOM, CEC, and ECEC. The strong influence of altitude on soil attribute modeling aligns with previous studies (Mosleh et al., 2016; Hengl et al., 2021; Zang et al., 2024), as it is a fundamental factor in pedogenesis. However, in this case, elevation may also be linked to clay-rich and more fertile soils found at the foothills of escarpments, which mark the beginning of the study area (see Fig. 1e for details). Another key group in the models consisted of GLCM metrics, including mean, variance, and contrast. These texture-based indicators are frequently applied in soil assessments (Duan and Zhang, 2021; Haoyuan et al., 2023) due to their ability to capture subtle patterns related to soil structure and texture. In our study, these techniques were particularly valuable for modeling texture fractions and attributes such as Ca, Mg, K, and CEC.

4.2. Implications for environmental evaluations and economic activities

The soils in the study area, typical of the Cerrado biome, are characterized by high leaching and weathering, with sand dominating the profile (Oliveira et al., 2023). Remote mapping of textural composition allows for precise soil management. These supporting practices improve water infiltration and reduce surface runoff, allowing for an even more accurate assessment of pasture degradation (Bolfe et al., 2024).

Table 6
Summary of model performances by tests and depth.

Test and soil layer	Performance	Clay	Silt	Sand	pH	SOM	CEC	ECEC	V	Ca	Mg	K	P	S	B	Cu	Fe	Mn	Zn
T1	R ²	0.06	0.39	0.37	0.1	0.21	0.01	0.18	0	0.15	0.09	0.09	0	0.13	0.02	0.02	0.2	0.06	0.08
0-0.2 m	RMSE%	57.76	71.85	14.36	2.33	25.99	58.19	38.15	52.82	52.47	48.67	60.53	65.53	108.03	30.87	28.09	32.41	71.55	60.72
T2	R ²	0.81	0.69	0.54	0.42	0.34	0.73	0.19	0.1	0.28	0.38	0.36	0.02	0.46	0.25	0.06	0.15	0.13	0.16
0-0.2 m	RMSE%	25.19	48.54	12.09	1.86	24	23.63	38.22	48.54	48.54	37.53	48.78	52.81	102.74	27.76	29.7	37.82	70.25	57.05
T3	R ²	0.4	0.34	0.37	0.24	0.27	0.03	0.3	0.06	0	0.2	0.35	0.12	0.03	0.15	0.07	0.34	0.01	0.05
0-0.2 m	RMSE%	47.22	73.12	14.36	2.18	24.92	54.82	35.67	50.74	57.87	47.74	57.81	44.48	113.57	28.9	27.94	29.76	73.27	61.78
T4	R ²	0.77	0.71	0.6	0.4	0.56	0.46	0.38	0.16	0.31	0.23	0.62	0.18	0.55	0.25	0.23	0.28	0.05	0.25
0-0.2 m	RMSE%	28.79	44.73	11.23	1.89	18.85	36.93	32.55	46.91	47.91	43.41	44.3	38.7	91.17	27.76	25.65	32.13	73.13	54.59
T1	R ²	0.16	0.18	0.44	0.25	0.26	0	0.18	0.04	0.04	0.13	0.18	0.02	0.03	0.07	0.02	0.11	0.03	0.02
0.2-0.4 m	RMSE%	49.82	66.25	15.58	2.29	29.63	62.55	43.24	42.08	51.91	51.62	58.04	68.25	82.28	32.39	38.53	23.57	78.09	67.14
T2	R ²	0.8	0.76	0.72	0.41	0.38	0.38	0.28	0.41	0.01	0.39	0.15	0.07	-	0.32	-	0.1	0.34	-
0.2-0.4 m	RMSE%	25.18	39.32	10.44	1.9	27.45	42.54	40.28	33.31	52.54	58.75	59.25	76.1	-	27.88	-	25.02	67.66	-
T3	R ²	0.45	0.49	0.44	0.44	0.34	0.01	0.37	0.1	0.13	0.17	0.31	0.17	0.09	0.08	0	0.12	0.08	0.01
0.2-0.4 m	RMSE%	41.03	55.81	15.08	2.09	28	59.12	239	40.8	51.43	52.19	54.09	52.66	41.38	32.13	34.8	23.45	76.17	64.87
T4	R ²	0.86	0.71	0.78	0.4	0.68	0.64	0.38	0.18	0.28	0.19	0.36	0.18	0.2	0.24	-	0.19	0.34	0.04
0.2-0.4 m	RMSE%	19.15	40.76	10.59	1.94	21.01	32.11	37.68	41.58	48.33	48.29	49.89	49.89	75.48	30.28	-	24.13	67.39	324.4

Furthermore, layer variations of 0.2–0.4 m, clay, or silt detected via RS may indicate abrupt textural transitions, potentially increasing the risk of subsurface erosion (Hosseinalizadeh et al., 2019).

The developed framework has broad applications in precision agriculture and native vegetation restoration, enabling targeted fertilizer applications and reducing nutrient losses through leaching (Shojaei et al., 2022). By leveraging models and maps, it is possible to identify areas with higher clay content, allowing for selecting sites with more significant potential for fertilizer response (Donagemma et al., 2016). Mapping CEC is equally essential, as this parameter influences the frequency and dosage of nitrogen- and potassium-based fertilizers (Yahaya et al., 2023) and provides insights into pH correction needs (Albuquerque et al., 2024).

4.3. Limitations, implications for the study area, and Future Studies

One limitation of this study was the inability to isolate exposed soil pixels, as outlined in Gasmí et al. (2021) and Silvero et al. (2023). Consequently, RS data captured spatial variations within composite samples, including vegetation cover. Future research should incorporate pure pixel masks to separate soil from vegetation effects (Zepp et al., 2021). This approach could optimize sample collection efforts, especially when combined with hyperspectral imagery for more precise modeling (Guo et al., 2021). Additionally, integrating data from unmanned aerial vehicles (UAVs) and ground-based sensors could further enhance soil property assessments (Zhang et al., 2023). LIDAR data also offers critical value in assessing topographic variations, which influence soil texture and erosion patterns (Frizzle et al., 2021).

Exploring other variables selection techniques, such as Recursive Feature Elimination (RFE), Forward Recursive Feature Selection (FRFS), or Modified Greedy Feature Selection (MGFS), in conjunction with these sensor technologies, could further enhance model accuracy (Liu et al., 2021; Zhang et al., 2023).

In future studies using orbital remote sensing, we increasingly expect pixels to reflect vegetation cover rather than bare soil. In this process of active (planting) and passive (natural regeneration) restoration of the Cerrado, orbital variables may track the transition from predominantly grassland vegetation to a shrub-dominated stratum in the medium term and potentially to a forested layer in sites with higher natural fertility and clay content (Souza et al., 2021). To increase the correlation of soil properties with vegetation growth, it will be necessary to select MZs that have a minimum number of simple soil samples ($n > 10$). Experiments on fertilization at planting and its effects on soil and vegetation may also be detected through RS data. New sensors, vegetation indices, and machine-learning techniques can also be incorporated to monitor the restoration process.

5. Conclusions

This study analyzed a regenerating Cerrado vegetation area in Brazil to assess the relationships between 18 physicochemical soil attributes at two depths, using machine learning predictive models based on 128 orbital remote sensing variables. The key findings can be summarized as follows: 1) the ArcGIS®PRO *Build Balanced Zones* tool is an effective method to create soil MZs; 2) previous correlation analyses between summary measurements in polygons, fields, or MZs have positive effects on the performance of predictive models; 3) RF and Boruta ensure better performance in soil modeling via RS data; 4) clay, silt, sand, CEC and K obtained the best results in both layers; 5) SAR variables added little, but Elevation and GLCM metrics were fundamental. This study demonstrates the value of remote soil assessments, offering practical insights for land restoration, precision agriculture, and ecosystem service evaluations.

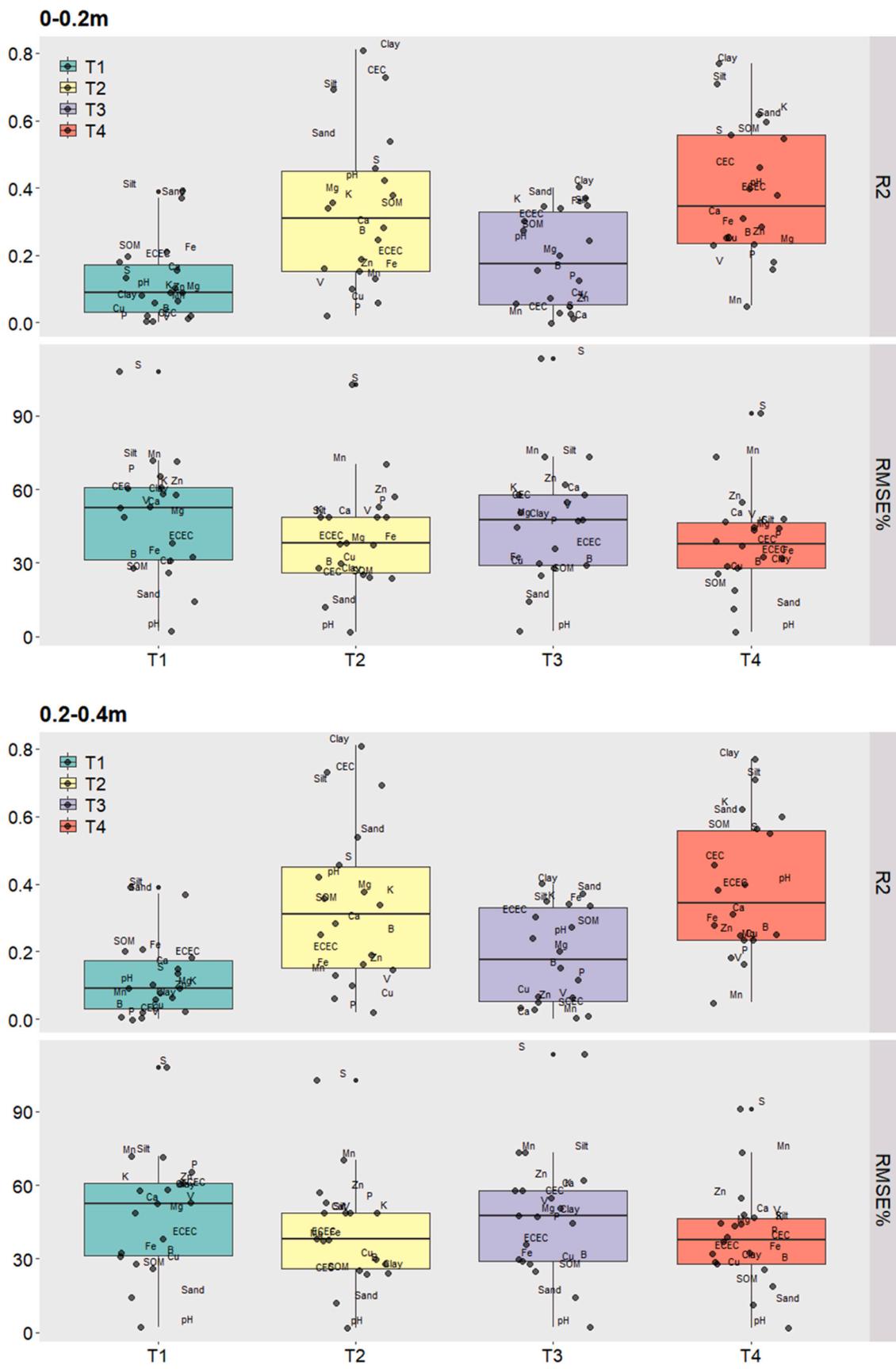


Fig. 7. Box Plot of all R^2 and RMSE% by tests and depth.

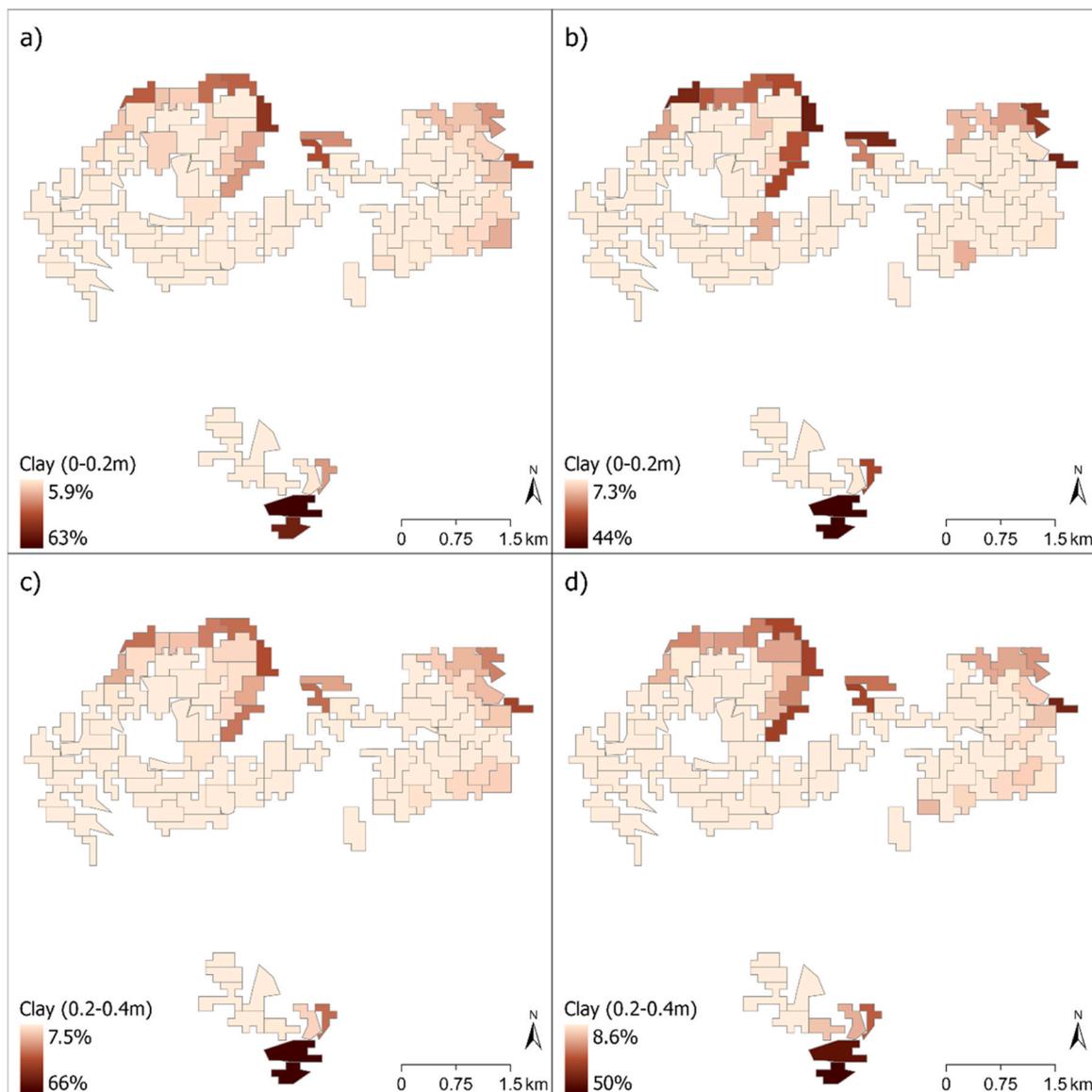


Fig. 8. Comparative maps of actual field data for clay in the 0–0.2 m layer (a) and the 0.2–0.4 m layer (c) are shown alongside the predicted results using RF+Boruta. Predictions were based on the mean of orbital indices in the T2 test (b) and the highest correlation statistics in the T4 test (d).

CRediT authorship contribution statement

Bergier Ivan: Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Louzada Romullo Oliveira:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Barbedo Jayme Garcia Arnal:** Writing – review & editing, Supervision, Formal analysis. **Bolfe Édson Luis:** Writing – original draft, Supervision, Methodology.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests. Romullo Louzada reports financial support was provided by IMASUL. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We want to thank the support of the Environmental Institute of Mato Grosso do Sul, in the person of the President Director, Mr. André Borges, for supporting and financing soil analysis and the environmental inspector, Bruna Gomes de Oliveira, in collecting field photos. We thank the FAPESP grant (São Paulo Research Foundation), process number #2022/09319–9.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.soilad.2025.100044](https://doi.org/10.1016/j.soilad.2025.100044).

Data availability

Data will be made available on request.

References

- Abdulraheem, M.I., Zhang, W., Li, S., Moshayedi, A.J., Farooque, A.A., Hu, J., 2023. Advancement of remote sensing for soil measurements and applications: a comprehensive review. *Sustainability* 15 (21). <https://doi.org/10.3390/su152115444>.
- Albuquerque, C.G. d., Gavelaki, F., Matera, H.B., Motta, A.C.V., Prior, S.A., Ercole, T.M., Araújo, E.M., 2024. Relationship between pH and base saturation associated with soil cation exchange capacity in soils of Mato Grosso do Sul, Brazil. *Bragantia* 83, e20230291. <https://doi.org/10.1590/1678-4499.20230291>.
- Amankulova, K., Farmonov, N., Omonov, K., Abdurakhimova, M., Mucsi, L., 2024. Integrating the Sentinel-1, Sentinel-2 and topographic data into soybean yield modelling using machine learning. *Adv. Space Res.* 73 (8), 4052–4066. <https://doi.org/10.1016/j.asr.2024.01.040>.
- Argento, F., Anken, T., Abt, F., Vogelsanger, E., Walter, A., Liebis, F., 2021. Site-specific nitrogen management in winter wheat supported by low-altitude remote sensing and soil data. *Precis. Agric.* 22 (2), 364–386. <https://doi.org/10.1007/s11119-020-09733-3>.
- Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Zeng, X., 2018. Satellite and in situ observations for advancing global earth surface modelling: a review. *Remote Sens.* 10 (12). <https://doi.org/10.3390/rs10122038>.
- Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. *Remote Sens. Rev.* 13 (1–2), 95–120. <https://doi.org/10.1080/02757259509532298>.
- Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35 (2), 161–173. [https://doi.org/10.1016/0034-4257\(91\)90009-U](https://doi.org/10.1016/0034-4257(91)90009-U).
- Barron, V., Torrent, J., 1986. Use of the Kubelka–Munk theory to study the influence of iron oxides on soil colour. *J. Soil Sci.* 37 (4), 499–510. <https://doi.org/10.1111/j.1365-2389.1986.tb00382.x>.
- Bartels, G.K., Castro, N.M. d R., Pedrollo, O., Collares, G.L., 2021. Soil moisture estimation in two layers for a small watershed with neural network models: Assessment of the main factors that affect the results. *CATENA* 207, 105631. <https://doi.org/10.1016/j.catena.2021.105631>.
- Bégué, A., Arvor, D., Bellon, B., Betbeder, J., De Abelleira, D., P. D. Ferraz, R., R. Verón, S., 2018. Remote sensing and cropping practices: a review. *Remote Sens.* 10 (1). <https://doi.org/10.3390/rs10010099>.
- Belal, A.A., El-Ramady, H., Jalhoum, M., Gad, A., Mohamed, E.S., 2021. Precision farming technologies to increase soil and crop productivity. In: Abu-hashim, M., Khebour Allouche, F., Negm, A. (Eds.), *Agro-Environmental Sustainability in MENA Regions*. Springer International Publishing, Cham, pp. 117–154.
- Bócoli, F.A., Ribeiro, D., Mancini, M., de Sousa, L.A., Barbosa, S.M., Serafim, M.E., Silva, S.H.G., 2025. Can environmental variables, high sampling density and machine learning deliver detailed maps of soil organic carbon and carbon stock in tropical regions? *CATENA* 249, 108718. <https://doi.org/10.1016/j.catena.2025.108718>.
- Bolfe, É. L., Victoria, D.D., Sano, E.E., Bayma, G., Massruhá, S.M., de Oliveira, A.F., 2024. Potential for agricultural expansion in degraded pasture lands in Brazil based on geospatial databases. *Land* 13 (2). <https://doi.org/10.3390/land13020200>.
- Bouslih, Y., John, K., Miftah, A., Azmi, R., Aboutayeb, R., Bouasria, A., Hssaini, L., 2024. The effect of covariates on Soil Organic Matter and pH variability: a digital soil mapping approach using random forest model. *Ann. GIS* 30 (2), 215–232. <https://doi.org/10.1080/19475683.2024.2309868>.
- Burns, B.W., Green, V.S., Hashem, A.A., Massey, J.H., Shew, A.M., Adviento-Borbe, M.A.A., Milad, M., 2022. Determining nitrogen deficiencies for maize using various remote sensing indices. *Precis. Agric.* 23 (3), 791–811. <https://doi.org/10.1007/s11119-021-09861-4>.
- Cammarano, D., Zha, H., Wilson, L., Li, Y., Batchelor, W.D., Miao, Y., 2020. A remote sensing-based approach to management zone delineation in small scale farming systems. *Agronomy* 10 (11). <https://doi.org/10.3390/agronomy10111767>.
- Charishma, D.S., Kuligod, V.B., Gundlur, S.S., Potdar, M.P., Doddamani, M.B., Nagaveni, H.C., 2024. Estimation of top soil properties by Sentinel-2 imaging. *Geol., Ecol., Landsc.* 1–10. <https://doi.org/10.1080/24749508.2024.2392920>.
- Crist, E.P., Cicone, R.C., 1984. A physically-based transformation of thematic mapper data—the TM tasseled cap (GE-). *Geosci. Remote Sens., IEEE Trans.* 22 (3), 256–263. <https://doi.org/10.1109/tgrs.1984.350619>.
- DeBruin, H., Khani, A., 2024. Geographic information system-based comprehensive shared micromobility station siting optimization for small urban areas. *Transp. Res. Rec.* 2678 (10), 85–101. <https://doi.org/10.1177/03611981241231799>.
- Diaz-Gonzalez, F.A., Vuelvas, J., Correa, C.A., Vallejo, V.E., Patino, D., 2022. Machine learning and remote sensing techniques applied to estimate soil indicators – Review. *Ecol. Indic.* 135, 108517. <https://doi.org/10.1016/j.ecolind.2021.108517>.
- Diniz, J.M.F. d S., Gama, F.F., Adami, M., 2020. Evaluation of polarimetry and interferometry of sentinel-1A SAR data for land use and land cover of the Brazilian Amazon Region. *Geocarto Int.* 1–19. <https://doi.org/10.1080/10106049.2020.1773544>.
- Donagemma, G.K., Freitas, P.L. d., Balieiro, F. d C., Fontana, A., Spera, S.T., Lumbreiras, J. F., Albuquerque, M.R. d., 2016. Characterization, agricultural potential, and perspectives for the management of light soils in Brazil. *Pesqui. Agropecu. ária Bras.* 51 (9), 1003–1020. <https://doi.org/10.1590/S0100-204X2016000900001>.
- Duan, M., Zhang, X., 2021. Using remote sensing to identify soil types based on multiscale image texture features. *Comput. Electron. Agric.* 187, 106272. <https://doi.org/10.1016/j.compag.2021.106272>.
- El Hajj, M., Baghdadi, N., Bazzi, H., Zribi, M., 2019. Penetration analysis of SAR Signals in the C and L Bands for Wheat, Maize, and Grasslands. *Remote Sens.* 11 (1). <https://doi.org/10.3390/rs11010031>.
- Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A., Scholten, T., 2020. Predicting and Mapping of Soil Organic Carbon Using Machine Learning Algorithms in Northern Iran. *Remote Sensing* 12 (14). <https://doi.org/10.3390/rs12142234>.
- Escadafal, R., 1989. Remote sensing of arid soil surface color with Landsat thematic mapper. *Adv. Space Res.* 9 (1), 159–163. [https://doi.org/10.1016/0273-1177\(89\)90481-X](https://doi.org/10.1016/0273-1177(89)90481-X).
- Feltran-Barbieri, R., Féres, J.G., 2021. Degraded pastures in Brazil: improving livestock production and forest restoration. *R. Soc. Open Sci.* 8 (7), 201854. <https://doi.org/10.1098/rsos.201854>.
- Flowers, M., Weisz, R., White, J.G., 2005. Yield-based management zones and grid sampling strategies: describing soil test and nutrient variability. *Agron. J.* 97 (3), 968–982. <https://doi.org/10.2134/agronj2004.0224>.
- Folorunso, O., Ojo, O., Busari, M., Adebayo, M., Joshua, A., Folorunso, D., Olabanjo, O., 2023. Exploring Machine Learning Models for Soil Nutrient Properties Prediction: A Systematic Review. *Big Data Cogn. Comput.* 7 (2). <https://doi.org/10.3390/bdcc7020113>.
- Frizze, C., Fournier, R.A., Trudel, M., Luther, J.E., 2021. Using the Soil and Water Assessment Tool to develop a LiDAR-based index of the erosion regulation ecosystem service. *J. Hydrol.* 595, 126009. <https://doi.org/10.1016/j.jhydrol.2021.126009>.
- Fu, Y., Li, R., Zhu, Z., Xue, Y., Ding, H., Wang, X., Xia, W., 2024. SCARF: A new algorithm for continuous prediction of biomass dynamics using machine learning and Landsat time series. *Remote Sens. Environ.* 314, 114348. <https://doi.org/10.1016/j.rse.2024.114348>.
- Gasmi, A., Gomez, C., Lagacherie, P., Zouari, H., Laamrani, A., Chehbouni, A., 2021. Mean spectral reflectance from bare soil pixels along a Landsat-TM time series to increase both the prediction accuracy of soil clay content and mapping coverage. *Geoderma* 388, 114864. <https://doi.org/10.1016/j.geoderma.2020.114864>.
- Georgi, C., Spengler, D., Itzerott, S., Kleinschmit, B., 2018. Automatic delineation algorithm for site-specific management zones based on satellite remote sensing data. *Precis. Agric.* 19 (4), 684–707. <https://doi.org/10.1007/s11119-017-9549-y>.
- Guo, L., Sun, X., Fu, P., Shi, T., Dang, L., Chen, Y., Zeng, C., 2021. Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas. *Geoderma* 398, 115118. <https://doi.org/10.1016/j.geoderma.2021.115118>.
- Haoyuan, Y., Ce, C., Ujue, H., Jiangdong, J., Yinwen, C., Ruiqi, D., Zhitao, Z., 2023. Synergistic estimation of soil salinity based on Sentinel-1 image texture and Sentinel-2 salinity spectral indices. *J. Appl. Remote Sens.* 17 (1), 018502. <https://doi.org/10.1117/1.JRS.17.018502>.
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Trans. Syst., Man, Cybern., SMC* 3 (6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>.
- Hatefard, F., Dolati, P., Heidari, A., Zolfaghari, A.A., 2019. Assessing the performance of decision tree and neural network models in mapping soil properties. *J. Mt. Sci.* 16 (8), 1833–1847. <https://doi.org/10.1007/s11629-019-5409-8>.
- Hengl, T., Miller, M.A.E., Krizan, J., Shepherd, K.D., Sila, A., Kilibarda, M., Crouch, J., 2021. African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Sci. Rep.* 11 (1), 6130. <https://doi.org/10.1038/s41598-021-85639-y>.
- Henrich, V., Krauss, G., Götze, C., & Sandow, C. (2011). The IndexDatabase. Hosseinizadeh, M., Kariminejad, N., Rahmati, O., Keesstra, S., Alinejad, M., Mohammadian Behbahani, A., 2019. How can statistical and artificial intelligence approaches predict piping erosion susceptibility? *Sci. Total Environ.* 646, 1554–1566. <https://doi.org/10.1016/j.scitotenv.2018.07.396>.
- Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25 (3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X).
- Javadi, S.H., Guerrero, A., Mouazen, A.M., 2022. Clustering and smoothing pipeline for management zone delineation using proximal and remote sensing. *Sensors* 22 (2). <https://doi.org/10.3390/s22020645>.
- Kalle, A., Le Hégarat-Masclé, S., Otlé, C., Hubert-Moy, L., 2007. Determination of vegetation cover fraction by inversion of a four-parameter model based on isoline parametrization. *Remote Sens. Environ.* 111 (4), 553–566. <https://doi.org/10.1016/j.rse.2007.04.006>.
- Kerry, R., Ingram, B., Oliver, M., Frogbrook, Z., 2024. Soil sampling and sensed ancillary data requirements for soil mapping in precision agriculture II: contour mapping of soil properties with sensed z-score data for comparison with management zone averages. *Precis. Agric.* 25 (3), 1212–1234. <https://doi.org/10.1007/s11119-023-10108-7>.
- Keskin, H., Grunwald, S., Harris, W.G., 2019. Digital mapping of soil carbon fractions with machine learning. *Geoderma* 339, 40–58. <https://doi.org/10.1016/j.geoderma.2018.12.037>.
- Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer, S., 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Comput. Electron. Agric.* 153, 213–225. <https://doi.org/10.1016/j.compag.2018.07.016>.
- Kursa, M.B., Rudnicki, W.R., 2010. Feature selection with the Boruta package. *J. Stat. Softw.* 36, 1–13. <https://doi.org/10.18637/jss.v036.i11>.
- Lewis, K., de V. Barros, F., Cure, M.B., Davies, C.A., Furtado, M.N., Hill, T.C., Rowland, L., 2022. Mapping native and non-native vegetation in the Brazilian Cerrado using freely available satellite products. *Sci. Rep.* 12 (1), 1588. <https://doi.org/10.1038/s41598-022-05332-6>.
- Li, Y., Chang, C., Wang, Z., Zhao, G., 2022. Remote sensing prediction and characteristic analysis of cultivated land salinization in different seasons and multiple soil layers in the coastal area. *Int. J. Appl. Earth Obs. Geoinf.* 111, 102838. <https://doi.org/10.1016/j.jag.2022.102838>.

- Liang, S., 2007. Recent developments in estimating land surface biogeophysical variables from optical remote sensing. *Prog. Phys. Geogr.: Earth Environ.* 31 (5), 501–516. <https://doi.org/10.1177/0309133307084626>.
- Lin, H., Wheeler, D., Bell, J., Wilding, L., 2005. Assessment of soil spatial variability at multiple scales. *Ecol. Model.* 182 (3), 271–290. <https://doi.org/10.1016/j.ecolmodel.2004.04.006>.
- Liu, Y., Qian, J., Yue, H., 2021. Comprehensive Evaluation of Sentinel-2 Red Edge and Shortwave-Infrared Bands to Estimate Soil Moisture. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 7448–7465. <https://doi.org/10.1109/JSTARS.2021.3098513>.
- Locatelli, J.L., de Lima, R.P., Santos, R.S., Cherubin, M.R., Creamer, R.E., Cerri, C.E.P., 2023. Soil Strength and Structural Stability Are Mediated by Soil Organic Matter Composition in Agricultural Expansion Areas of the Brazilian Cerrado Biome. *Agronomy* 13 (1). <https://doi.org/10.3390/agronomy13010071>.
- Louzada, R.O., Reis, L.K., Diniz, J.M.F. d S., Roque, F. d O., Gama, F.F., Bergier, I., 2023. Combining optical and microwave remote sensing for assessing gullies in human-disturbed vegetated landscapes. *CATENA* 228, 107127. <https://doi.org/10.1016/j.catena.2023.107127>.
- Ma, G., Ding, J., Han, L., Zhang, Z., Ran, S., 2021. Digital mapping of soil salinization based on Sentinel-1 and Sentinel-2 data combined with machine learning algorithms. *Reg. Sustain.* 2 (2), 177–188. <https://doi.org/10.1016/j.regsus.2021.06.001>.
- Mathieu, R., Pouget, M., Cerville, B., Escadafal, R., 1998. Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typical soil color of an arid environment. *Remote Sens. Environ.* 66 (1), 17–28. [https://doi.org/10.1016/S0034-4257\(98\)00030-3](https://doi.org/10.1016/S0034-4257(98)00030-3).
- Moradpour, S., Entezari, M., Ayoubi, S., Karimi, A., Naimi, S., 2023. Digital exploration of selected heavy metals using Random Forest and a set of environmental covariates at the watershed scale. *J. Hazard. Mater.* 455, 131609. <https://doi.org/10.1016/j.jhazmat.2023.131609>.
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. *IEEE Geosci. Remote Sens. Mag.* 1 (1), 6–43. <https://doi.org/10.1109/MGRS.2013.2248301>.
- Mosleh, Z., Salehi, M.H., Jafari, A., Borujeni, I.E., Mehnatkesh, A., 2016. The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environ. Monit. Assess.* 188 (3), 195. <https://doi.org/10.1007/s10661-016-5204-8>.
- Muniyammal VEDIAPPAN, D., Godi, A., Golla, B., 2025. Geospatial mapping of soil properties of forest types using the k-means fuzzy clustering approach to delineate site-specific nutrient management zones in Goa, India. *J. Indian Soc. Remote Sens.* <https://doi.org/10.1007/s12524-024-02082-y>.
- Nakai, H., Itatani, T., Horiike, R., 2021. Construction of an evacuee placement model for tsunami shelters considering physical distancing to prevent COVID-19 infection. *Prog. Disaster Sci.* 11, 100183. <https://doi.org/10.1016/j.pdisas.2021.100183>.
- Navarro, R., Wirkus, L., Dubovyk, O., 2023. Spatio-Temporal Assessment of Olive Orchard Intensification in the Saiss Plain (Morocco) Using k-Means and High-Resolution Satellite Data. *Remote Sens.* 15 (1). <https://doi.org/10.3390/rs15010050>.
- Nawar, S., Corstanje, R., Halcro, G., Mulla, D., & Mouazen, A.M., 2017. Chapter Four - Delineation of Soil Management Zones for Variable-Rate Fertilization: A Review. In: D.L. Sparks (Ed.), *Advances in Agronomy* (Vol. 143, pp. 175–245): Academic Press.
- Oldoni, H., Magalhães, P.S.G., Oliveira, A.L.G., Lima, J.P., Figueiredo, G.K.D.A., Moro, E., Amaral, L.R., 2025. Management zones delineation: a proposal to overcome the crop-pasture rotation challenge. *Precis. Agric.* 26 (1), 21. <https://doi.org/10.1007/s11119-024-10214-0>.
- Oliveira, J., Campbell, E.E., Lamparelli, R.A.C., Figueiredo, G.K.D.A., Soares, J.R., Jaiswal, D., Sheehan, J.J., 2020. Choosing pasture maps: an assessment of pasture land classification definitions and a case study of Brazil. *Int. J. Appl. Earth Obs. Geoinf.* 93, 102205. <https://doi.org/10.1016/j.jag.2020.102205>.
- Oliveira, V.A. d, Santos, G.G., Ker, J.C., Couto, E.G., Jacomine, P.K., Corrêa, G.R., Schaefer, C.E.G.R., 2023. Soils of Cerrados, the Brazilian Savannas. In: Schaefer, C.E. G.R. (Ed.), *The Soils of Brazil*. Springer International Publishing, Cham, pp. 129–173.
- Padarian, J., Minasny, B., McBratney, A.B., 2020. Machine learning and soil sciences: a review aided by machine learning tools. *SOIL* 6 (1), 35–52. <https://doi.org/10.5194/soil-6-35-2020>.
- Patel, N., Padhiyar, N., 2010. Alien genetic algorithm for exploration of search space. *AIP Conf. Proc.* 1298 (1), 325–330. <https://doi.org/10.1063/1.3516325>.
- Peng, Y., Wang, T., Xie, S., Liu, Z., Lin, C., Hu, Y., Mao, X., 2023. Estimation of Soil Cations Based on Visible and Near-Infrared Spectroscopy and Machine Learning. *Agriculture* 13 (6). <https://doi.org/10.3390/agriculture13061237>.
- Pereira, M. d A., Bungenstab, D.J., Euclides, V.P.B., Malafaia, G.C., Biscola, P.H.N., Menezes, G.R.O., de Souza, V.F., 2024. From traditionally extensive to sustainably intensive: a review on the path to a sustainable and inclusive beef farming in Brazil. *Animals* 14 (16). <https://doi.org/10.3390/ani14162340>.
- Pichler, M., Hartig, F., 2023. Machine learning and deep learning—a review for ecologists. *Methods Ecol. Evol.* 14 (4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>.
- Pôças, I., Paço, T.A., Paredes, P., Cunha, M., Pereira, L.S., 2015. Estimation of actual crop coefficients using remotely sensed vegetation indices and soil water balance modelled data. *Remote Sens.* 7 (3), 2373–2400. <https://doi.org/10.3390/rs70302373>.
- Poppiel, R.R., Lacerda, M.P.C., Safaneli, J.L., Rizzo, R., Oliveira, M.P., Novais, J.J., Demattê, J.A.M., 2019. Mapping at 30 m resolution of soil attributes at multiple depths in Midwest Brazil. *Remote Sens.* 11 (24). <https://doi.org/10.3390/rs11242905>.
- Price, J.C., 1993. Estimating leaf area index from satellite data. *IEEE Trans. Geosci. Remote Sens.* 31 (3), 727–734. <https://doi.org/10.1109/36.225538>.
- Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. *Remote Sens. Environ.* 48 (2), 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1).
- R Core Team, R., 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.
- Radočaj, D., Gašparović, M., Jurišić, M., 2024. Open remote sensing data in digital soil organic carbon mapping: a review. *Agriculture* 14 (7). <https://doi.org/10.3390/agriculture14071005>.
- Reis, A.S., Rodrigues, M., Alemparte Abrantes dos Santos, G.L., Mayara de Oliveira, K., Furlanetto, R.H., Teixeira Crusiol, L.G., Nanni, M.R., 2021. Detection of soil organic matter using hyperspectral imaging sensor combined with multivariate regression modeling procedures. *Remote Sens. Appl.: Soc. Environ.* 22, 100492. <https://doi.org/10.1016/j.rsase.2021.100492>.
- Ren, Y., Li, X., Mao, D., Wang, Z., Jia, M., Chen, L., 2020. Investigating spatial and vertical patterns of wetland soil organic carbon concentrations in China's Western Songnen Plain by Comparing Different Algorithms. *Sustainability* 12 (3). <https://doi.org/10.3390/su12030932>.
- Rowan, L.C., Mars, J.C., 2003. Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. *Remote Sens. Environ.* 84 (3), 350–366. [https://doi.org/10.1016/S0034-4257\(02\)00127-X](https://doi.org/10.1016/S0034-4257(02)00127-X).
- Sangeetha, C., Moond, V., Rajesh, G., Damor, J.S., Pandey, S.K., Kumar, P., Singh, B., 2024. Remote Sensing and Geographic Information Systems for Precision Agriculture: A Review. *Int. J. Environ. Clim. Change* 14 (2), 287–309. <https://doi.org/10.9734/ijec/2024/v14i23945>.
- Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappiè, M., Märker, M., Saia, S., 2017. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. *Sci. Total Environ.* 601–602, 821–832. <https://doi.org/10.1016/j.scitotenv.2017.05.239>.
- Sharma, R., Mishra, D.R., Levi, M.R., Sutter, L.A., 2022. Remote sensing of surface and subsurface soil organic carbon in tidal wetlands: a review and ideas for future research. *Remote Sens.* 14 (12). <https://doi.org/10.3390/rs14122940>.
- Sheykhouma, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S., 2020. Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 13, 6308–6325. <https://doi.org/10.1109/JSTARS.2020.3026724>.
- Shojaei, M.J., Or, D., Shokri, N., 2022. Localized delivery of liquid fertilizer in coarse-textured soils using foam as carrier. *Transp. Porous Media* 143 (3), 787–795. <https://doi.org/10.1007/s11242-022-01820-5>.
- Rouse JW Jr, Haas RH, Schell JA, Deering DW (1973) Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *Prog. Rep. RSC 1978-1*, Remote Sensing Center, Texas A, M Univ., College Station, nr E73-106393, 93 (NTIS No. E73-106393).
- Silvero, N.E. Q., Demattê, J.A. M., Minasny, B., Rosin, N.A., Nascimento, J.G., Rodríguez Albarraçin, H.S., ... Gómez, A.M. R., 2023. Chapter Three - Sensing technologies for characterizing and monitoring soil functions: A Review. In D.L. Sparks (Ed.), *Advances in Agronomy* (Vol. 177, pp. 125–168): Academic Press.
- Singh, S., Sarma, K., 2023. Exploring soil spatial variability with GIS, remote sensing, and geostatistical approach. *J. Soil, Plant Environ.* 2 (1), 79–99. <https://doi.org/10.56946/jspae.v2i1.186>.
- Siqueira, R.G., Moquedace, C.M., Fernandes-Filho, E.I., Schaefer, C.E.G.R., Francelino, M.R., Sacramento, I.F., Michel, R.F.M., 2024. Modelling and prediction of major soil chemical properties with Random Forest: Machine learning as tool to understand soil-environment relationships in Antarctica. *CATENA* 235, 107677. <https://doi.org/10.1016/j.catena.2023.107677>.
- Sishodia, R.P., Ray, R.L., Singh, S.K., 2020. Applications of remote sensing in precision agriculture: a review. *Remote Sens.* 12 (19). <https://doi.org/10.3390/rs12193136>.
- Song, X., Wang, J., Huang, W., Liu, L., Yan, G., Pu, R., 2009. The delineation of agricultural management zones with high resolution remotely sensed data. *Precis. Agric.* 10 (6), 471–487. <https://doi.org/10.1007/s11119-009-9108-2>.
- Sosa, L., Justel, A., Molina, Í., 2021. Detection of crop hail damage with a machine learning algorithm using time series of remote sensing data. *Agronomy* 11 (10). <https://doi.org/10.3390/agronomy11102078>.
- Souza, G.F. d, Almeida, R.F., Bijos, N.R., Fagg, C.W., Munhoz, C.B.R., 2021. Herbaceous-shrub species composition, diversity and soil attributes in moist grassland, shrub grassland and savanna in Central Brazil. *Braz. J. Bot.* 44 (1), 227–238. <https://doi.org/10.1007/s40415-020-00672-x>.
- Taghizadeh-Mehrjardi, R., Fathizad, H., Ali Hakimzadeh Ardakani, M., Sodaiezhadeh, H., Kerry, R., Heung, B., Scholten, T., 2021. Spatio-temporal analysis of heavy metals in arid soils at the catchment scale using digital soil assessment and a random forest model. *Remote Sens.* 13 (9). <https://doi.org/10.3390/rs13091698>.
- Techen, A.-K., Helming, K., Brüggemann, N., Veldkamp, E., Reinhold-Hurek, B., Lorenz, M., Vogel, H.-J., 2020. Chapter Four - Soil research challenges in response to emerging agricultural soil management practices. In: Sparks, D.L. (Ed.), *Advances in Agronomy*, 161. Academic Press, pp. 179–240.
- Teixeira, P.C., Donagemma, G.K., Fontana, A., Teixeira, W.G., 2017. *Manual de métodos de análise de solo*. Embrapa Bras. flia, DF.
- Tunçay, T., Kılıç, Ş., Dedeoğlu, M., Dengiz, O., Başkan, O., Bayramin, İ., 2021. Assessing soil fertility index based on remote sensing and gis techniques with field validation in a semiarid agricultural ecosystem. *J. Arid Environ.* 190, 104525. <https://doi.org/10.1016/j.jaridenv.2021.104525>.
- Wang, J., Zhen, J., Hu, W., Chen, S., Lizaga, I., Zeraatpisheh, M., Yang, X., 2023. Remote sensing of soil degradation: Progress and perspective. *Int. Soil Water Conserv. Res.* 11 (3), 429–454. <https://doi.org/10.1016/j.iswcr.2023.03.002>.

- Wang, Q., Zhao, L., Wang, M., Wu, J., Zhou, W., Zhang, Q., Deng, M., 2022. A random forest model for drought: monitoring and validation for grassland drought based on multi-source remote sensing data. *Remote Sens.* 14 (19). <https://doi.org/10.3390/rs14194981>.
- Wilson, J.P., Gallant, J.C., 2000. *Terrain Analysis: Principles and Applications*. John Wiley & Sons.
- Yahaya, S.M., Mahmud, A.A., Abdullahi, M., Haruna, A., 2023. Recent advances in the chemistry of nitrogen, phosphorus and potassium as fertilizers in soil: A review. *Pedosphere* 33 (3), 385–406. <https://doi.org/10.1016/j.pedsph.2022.07.012>.
- Yüzügüllü, O., Fajraoui, N., F, L., 2024. Soil texture and pH mapping using remote sensing and support sampling. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 17, 12685–12705. <https://doi.org/10.1109/JSTARS.2024.3422494>.
- Yüzügüllü, O., Lorenz, F., Fröhlich, P., Liebisch, F., 2020. Understanding Fields by Remote Sensing: Soil Zoning and Property Mapping. *Remote Sensing* 12 (7). <https://doi.org/10.3390/rs12071116>.
- Zang, M., Wang, X., Chen, Y., Faramarzi, S.E., 2024. Estimation of soil health in the semi-arid regions of northwestern Iran using digital elevation model and remote sensing data. *Environ. Monit. Assess.* 196 (4), 353. <https://doi.org/10.1007/s10661-024-12527-z>.
- Zepp, S., Jilge, M., Metz-Marconcini, A., Heiden, U., 2021. The influence of vegetation index thresholding on EO-based assessments of exposed soil masks in Germany between 1984 and 2019. *ISPRS J. Photogramm. Remote Sens.* 178, 366–381. <https://doi.org/10.1016/j.isprsjprs.2021.06.015>.
- Zeraatpisheh, M., Bottega, E.L., Bakhshandeh, E., Owliaie, H.R., Taghizadeh-Mehrdadi, R., Kerry, R., Xu, M., 2022. Spatial variability of soil quality within management zones: Homogeneity and purity of delineated zones. *CATENA* 209, 105835. <https://doi.org/10.1016/j.catena.2021.105835>.
- Zhang, Q., Liu, M., Zhang, Y., Mao, D., Li, F., Wu, F., Man, W., 2023. Comparison of Machine Learning Methods for Predicting Soil Total Nitrogen Content Using Landsat-8, Sentinel-1, and Sentinel-2 Images. *Remote Sens.* 15 (11). <https://doi.org/10.3390/rs15112907>.
- Zhu, Y., Qi, L., Wu, Z., Shi, P., 2024. Spectra-based predictive mapping of soil organic carbon in croplands: single-date versus multitemporal bare soil compositing approaches. *Geoderma* 449, 116987. <https://doi.org/10.1016/j.geoderma.2024.116987>.
- Zixian, Z., Xuning, L., Zhixiang, L., Hongqiang, H., 2021. Outburst prediction and influencing factors analysis based on Boruta-Apriori and BO-SVM algorithms. *J. Intell. Fuzzy Syst.* 41, 3201–3218. <https://doi.org/10.3233/JIFS-210466>.