Journal of Environmental Management 375 (2025) 124155

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Online analysis of Amazon's soils through reflectance spectroscopy and cloud computing can support policies and the sustainable development

Jean Jesus Macedo Novais [a], Borges Marfrann Dias Melo [a], Afrânio Ferreira Neves Junior [b], Raimundo Humberto Cavalcante Lima [c], Renato Epifânio de Souza [d], Valdinar Ferreira Melo [e], Eufran Ferreira do Amaral [f], Nikolaos Tziolas [g], José A.M. Demattê [a,*]

[a] Geotechnologies in Soil Sciences Research Group – GeoCiS, Department of Soil Science, Luiz de Queiroz College of Agriculture – Esalq, University of São Paulo – USP, Piracicaba, São Paulo, Brazil
[b] Federal University of Amazonas (UFAM), Distance Education Center (CED), Manaus, Amazonas, Brazil
[c] Federal University of Amazonas, Institute of Exact Sciences, Department of Geosciences, Manaus, Amazonas, Brazil
[d] Federal Institute of Education, Science and Technology of Acre (IFAC), Cruzeiro do Sul Campus, Acre, Brazil
[e] Federal University of Roraima, Center for Agricultural Sciences, Department of Soils and Agricultural Engineering, Roraima, Brazil
[f] Brazilian Agricultural Research Corporation (EMBRAPA), Embrapa Acre, Rio Branco, Acre, Brazil
[g] Southwest Florida Research and Education Center, Department of Soil, Water and Ecosystem Sciences, Institute of Food and Ecosystem Sciences, University of Florida, FL, United States

A R T I C L E I N F O

Handling Editor: Lixiao Zhang

Keywords:
Remote sensing
Hyperspectral data
Soil analysis
Machine learning
Soil security
Soil health
Soil mapping
Food and environment security

A B S T R A C T

Analyzing soil in large and remote areas such as the Amazon River Basin (ARB) is unviable when it is entirely performed by wet labs using traditional methods due to the scarcity of labs and the significant workforce requirements, increasing costs, time, and waste. Remote sensing, combined with cloud computing, enhances soil analysis by modeling soil from spectral data and overcoming the limitations of traditional methods. We verified the potential of soil spectroscopy in conjunction with cloud-based computing to predict soil organic carbon (SOC) and particle size (sand, silt, and clay) content from the Amazon region. To this end, we request physicochemical attribute values determined by wet laboratory analyses of 211 soil samples from the ARB. These samples were submitted to spectroscopy Vis-NIR-SWIR in the laboratory. Two approaches modeled the soil attributes: M-I) cloud-computing-based using the Brazilian Soil Spectral Service (BraSpecS) platform, and M-II) computing-based in an offline environment using R programming language. Both methods used the Cubist machine learning algorithm for modeling. The coefficient of determination (R^2), mean absolute error (MAE) and root mean squared error (RMSE) served as criteria for performance assessment. The soil attributes prediction was highly consistent, considering the measured and predicted by both approaches M-I and M-II. The M-II outperformed the M-I in predicting both particle size and SOC. For clay content, the offline model achieved an R^2 of 0.85, with an MAE of 86.16 g kg^{-1} and RMSE of 111.73 g kg^{-1}, while the online model had an R^2 of 0.70, MAE of 111.73 g kg^{-1}, and RMSE of 144.19 g kg^{-1}. For SOC, the offline model also showed better performance, with an R^2 of 0.81, MAE of 3.42 g kg^{-1}, and RMSE of 4.57 g kg^{-1}, compared to an R^2 of 0.72, MAE of 3.66 g kg^{-1}, and RMSE of 5.53 g kg^{-1} for the M-I. Both modeling methods demonstrated the power of reflectance spectroscopy and cloud computing to survey soils in remote and large areas such as ARB. The synergetic use of these techniques can support policies and sustainable development.

* Corresponding author.
E-mail addresses: jeannovais@usp.br, jjnagron@gmail.com (J.J.M. Novais), borgesmelo@usp.br (B.M.D. Melo), anevesjr@ufam.edu.br (A.F. Neves Junior), humbertoclima@ufam.edu.br (R.H.C. Lima), renato.souza@ifac.edu.br (R.E. de Souza), valdinar@yahoo.com.br (V.F. Melo), eufran.amaral@embrapa.br (E.F. do Amaral), ntziolas@ufl.edu (N. Tziolas), jamdemat@usp.br (J.A.M. Demattê).

https://doi.org/10.1016/j.jenvman.2025.124155
Received 18 September 2024; Received in revised form 29 December 2024; Accepted 13 January 2025
Available online 22 January 2025

1. Introduction

Further to the vast biodiversity, the Amazon River Basin (ARB) encompasses a rich pedodiversity originated by specific conditions that favor forming unique soil types with differential characteristics compared to other soil classes in Brazil (Schaefer et al., 2023). Amazon's soils are under threat due to the sensitivity of anthropogenic actions to environmental changes, exhibiting significant variation in attributes, including texture, depth, fertility, organic matter (OM), and pH. These properties are associated with climatic and specific pedomorphological conditions for their formation (McBratney and Minasny, 2018).

The Amazon Rainforest covers much of the ARB's soils (Schaefer et al., 2023). The cost and time required, the workforce, and the potential for significant waste generation make soil surveys of large and remote areas using traditional methods unprofitable (Sousa Junior et al., 2011). The few existing soil analysis laboratories would not meet the demand for soil data in the region, making it even more difficult and delaying obtaining accurate information about the soils. Thus, technologies that enhance the efficiency of this process are crucial for environmental protection, and understanding the relationships between anthropogenic interventions and this non-renewable natural resource (Novais et al., 2024). In this context, remote sensing and cloud computing can be essential for providing data on soil promptly available for several purposes (Demattê et al., 2022). Reflectance spectroscopy (RS) technologies are emerging as a valuable complement to classical soil attribute analysis by wet laboratory methods. RS offers fast, cost-effective surveys and mapping with reduced environmental impact (Sousa Junior et al., 2011; Araújo et al., 2015; Pinheiro et al., 2017; Liu et al., 2019; Nocita et al., 2015; Viscarra Rossel et al., 2022; Santos et al., 2023; Pavão et al., 2024).

The spectral behavior of soils enables the modeling of soil attributes that can be validated with a few routine analyses (Nanni and Demattê, 2006; Novais et al., 2024). Due to this technical and scientific advancement, cloud computing has arisen as an efficient technique that enhances the prediction of soil attributes on a large scale via RS (Poppiel et al., 2022). The Brazilian Soil Spectral Service (BraSpecS, http://besbbr.com.br/) represents a pioneering example of these systems. The platform allows users to input Vis-NIR-SWIR or mid-infrared (MIR) spectra, which are used to predict online soil attribute contents and the corresponding prediction metrics (Demattê et al., 2022). This approach takes soil spectroscopy further than the scientific stage. Gathering real-time soil analysis over large areas, such as the Amazon region, is an impactful approach for the community.

This application enhances soil surveying and mapping, as exemplified by its integration at the National Soil Program – PronaSolos (Polidoro et al., 2016, 2021), an effort that aims to map the soils from the 8.5 million km^2 of the Brazilian territory at fine scale-up to 2048. Therefore, the application of RS and modeling cloud computing supports several UN Sustainable Development Goals (SDGs), including Climate Action (SDG 13), Life on Land (SDG 15), and Zero Hunger (SDG 2) (United Nations, 2015). This approach enhances soil health monitoring (Adak et al., 2023), aiding in sustainable land management (Herranz-Luque et al., 2024), biodiversity conservation (Rocchini et al., 2016), and climate change mitigation. So, it promotes climate-resilient agriculture, contributing to food and soil security (Novais et al., 2024).

Enabling the broad distribution of critical environmental data empowers decision-makers to support sustainable development in Amazon. Indeed, Amazon has many difficulties in soil evaluation such as a lower number of laboratories per area and population compared to densely populated areas, inaccessible areas, logistical challenges and lack of security for the team. Despite the scarce laboratories to assist, the difficulty of reaching the areas is huge. We may underline that, during fieldwork, scientists should not have time to collect soil samples to bring back to laboratories. So, a real cloud time analysis would make this faster and it could be linked with field hand sensors such as those performed by Ben-Dor et al. (2024).

Given the above, we assume that the synergistic application of soil sensing at an extensive dataset located in cloud computing can provide accurate data on the soil of large and remote areas, informing society toward conserving sustainable exploitation of soil resources. Therefore, this study aims to demonstrate both the utility of RS for accurate soil characterization and the feasibility of real-time online analysis, emphasizing its potential for remote areas. The offline model demonstrates superior precision; however, it necessitates observations for training, while the online model does not require such observations. The RS technique has high potential to support of policies and sustainable development for the Amazon region.

2. Material and methods

2.1. Description of the physical environment

The ARB encompasses an area of approximately 7 million km^2 in South America, making it the largest river basin in the world. Only in Brazil, it extends for 4 million km^2 (42% of the national territory) and 57% of basin boundaries, followed by Peru (17%), Bolivia (11%), Colombia (5.8%), Ecuador (2.2%), Venezuela (0.8%), Guyana (0.2%), and Suriname (0.2%). Of the total area, 5.5 million km^2 is covered by tropical forest, namely the Amazon Rainforest. These forest arrangements form a large biome that covers territories belonging to nine nations, in which Brazil holds 60% of this vegetation, with Peru holding 12%, and smaller portions of the rainforest are located in Colombia, Ecuador, Bolivia, Venezuela, Guyana, Suriname, and French Guiana (Hoorn and Wesselingh, 2010).

The ARB, underlain by ancient Precambrian rocks primarily consisting of granites and metamorphic formations, features the Guyana and Brazilian shields with the Amazon Craton between them, filled with Paleozoic to Cenozoic sediments (Smith, 1982). Geomorphologically, the Amazon Basin features vast floodplains, extensive river networks, and numerous wetlands, characterized by its flat topography with elevations typically below 500 m and extensive alluvial plains formed by sediment deposition from the Amazon River and its tributaries (Hoorn and Wesselingh, 2010). Orellana Segovia et al. (2020) state that Ferralsols and Lixisols predominate in ARB. These soils are highly weathered and nutrient-poor due to intense tropical weathering. Also, there is the occurrence of Plinthosols, Gleisols, Luvisols, and Histosols, which are soils with less pedogenic evolution levels yet support dense vegetation through efficient nutrient cycling (Irion, 1978; IUSS Working Group WRB, 2022).

The ARB presents a humid tropical climate with high temperatures and heavy rainfall throughout the year, ranging from 1500 to 3000 mm annually, which supports the dense rainforest dominating the region (Sombroek, 2001; Marengo, 2006). The vegetation includes the world's largest tropical rainforest, comprising various ecosystems such as dryland, floodplain, and black-water flooding "igapó" forests, supporting unparalleled biodiversity with many species remaining undocumented (Gentry, 1992).

2.2. Proceedings

This section presents the material and methods used to assess soil analysis methods in ARB by reflectance spectroscopy, cloud computing, and traditional modeling as support to policies. Fig. 1 depicts the overall workflow comprising five steps, which collectively facilitate a comprehensive understanding of the dynamics in question. The methodology begins with users inputting their ARB spectra into the BraSpecS platform, thereby obtaining soil attributes (particle size and soil organic carbon). Subsequently, a comparison is conducted between the aforementioned soil attributes and those predicted by the offline Cubist Model. Based on the results, the final stage of the sequence consists of a discussion about the usability of spectroscopy and cloud computing in the context of policies designed to benefit society.

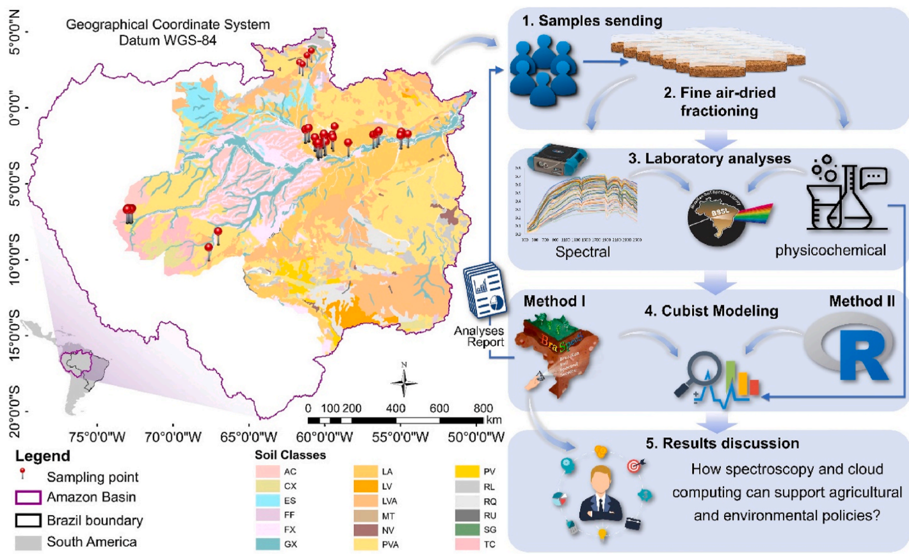

Fig. 1. Location of the sampling points overlayed to the soil map and workflow of applied methodology. Brazilian soil map (Santos et al., 2011, 2018) with corresponding classes at the IUSS Working Group WRB (2022): AC (Alisol), CX (Haplic Cambisol), ES (Podzol), FF (Petric Plinthosol), FX (Haplic Plinthosol), GX (Haplic Gleisol), LA (Xantic Ferralsol), LV (Rhodic Ferralsol), LVA (Haplic Ferralsol), MT (Chernozem), NV (Rhodic Nitisol), PV (Rhodic Lixisol), PVA (Haplic Lixisol), RL (Leptsol), RQ (Arenosol), RU (Fluvisol), SG (Solonetz) e TC (Luvisol).

2.2.1. *Soil analysis and data processing*

Firstly, a request was made for users from ARB to send soil samples for soil analysis and modeling purposes. The collection points were distributed across the ARB, encompassing the Brazilian states of Acre, Amazonas, Pará, Roraima, and Rondônia, as illustrated in Fig. 1. It is noteworthy that among the soil data observed and utilized in this article, 27 analyses correspond to selected subsurface horizons provided by the coordination of the XV Soil Correlation and Classification Meeting held by Embrapa in the study region. The samples were collected from both the surface and subsurface layers, with sampling depths varying according to soil class. The samples were then subjected to standardized laboratory analysis to determine their physicochemical and spectral attributes at the fine earth air-dried fraction (particles < 2 mm) (Soil Survey Soil Survey Staff, 2017). A soil database was then compiled, comprising 211 sample points, which included data on hyperspectral analysis, particle size, and SOC content.

2.2.2. *Vis-NIR-SWIR reflectance spectroscopy*

Following, we used the internal protocol for spectral analyses to read soil samples from the Remote Sensing Laboratory of the Luiz de Queiroz College of Agriculture (Esalq/USP) at the University of São Paulo, Brazil (Poppiel et al., 2022). These samples were received in the air-dried fine soil fraction, i.e., air-dried, crushed, and sieved to a >2 mm mesh. The samples were arranged in Petri dishes and submitted to spectral readings in a controlled laboratory environment (black room) with artificial light, as described by Poppiel et al. (2022). The equipment used for the spectral readings of the samples, FieldSpec 3 Pro© spectroradiometer (ASD, 2019), covers a wide range of wavelengths (from 350 to 2500 nm) segmented into 2150 bands and interpolated every 1 nm. We used the protocol developed by Ben-Dor et al. (2015) for spectral measurements. The spectral curves were morphologically interpreted in terms of shape and main features according to Demattê et al. (2014) recommendations. The reflectance data compiled a representative SSL of ARB, enabling the modeling online and offline.

2.2.3. *Online spectral modeling – M-I*

The first method occurred in the BraSpecS system, an online platform that models soil hyperspectral data. This cloud-based prediction uses the Cubist model, a machine learning algorithm that enables efficient dataset analysis to reduce processing time (Demattê et al., 2022). This platform works jointly with a database from the Brazilian Soil Spectral Library (BSSL, Demattê et al., 2019), which has approximately forty thousand soil sample results with Vis-NIR-SWIR and MIR spectra attached to physicochemical values.

Formerly, BraSpecS platform was validated by Demattê et al. (2022) using a 70/30 split of calibration and validation datasets using k-fold cross-validation for soil properties such as clay and SOC. The performance of the model was evaluated using the following metrics: R^2, root mean square error (RMSE), and Mean Absolute Error (MAE). Blind testing involved 500 Brazilian users, and international datasets from 65 countries were used to evaluate predictions across diverse regions. A comparative analysis was conducted among three modeling approaches: BraSpecS, local ExCSSL, and global GSSL. This analysis demonstrated the system's reliability and the advantages of localized spectral models (Novais et al., 2024).

In practical terms, when a user inserts an unknown spectrum, or an SSL, in the BraSpecS, up to 15 predicted soil attributes return with their respective metrics in a few minutes (Demattê et al., 2022). Thus, we loaded the SSL compiled in the system and selected sand, silt, clay and SOC attributes, which were submitted to Pearson Correlation Coefficient (PCC) for verifying the adequacy of this data to achieve the aim. That means, spectra were entered via the internet into BraSpecS and the results were compared with the field truth ones. Afterward, we compared the results with soil attributes measured for validation. This step enabled us to discuss the different soil spectral patterns that affect the modeling results.

2.2.4. *Offline spectral modeling – M-II*

The modeling process for predicting particle size and SOC content in the ARB region was performed using the local SSL and observed data in the R-Cran environment (R Core Team, 2024) by the Cubist model. The Cubist regression algorithm was selected for its superior ability to handle non-linear relationships, as evidenced by its higher predictive accuracy in predicting soil attributes (Hengl et al., 2018). The enhanced performance of the Cubist algorithm relative to linear models, such as Partial Least-Squares Regression (PLSR), can be attributed to the integration of decision trees with localized linear regression, a distinctive feature of the Cubist algorithm (Demattê et al., 2022). This integration enhances the algorithm's robustness and overall effectiveness in modeling complex soil attribute data.

We adhered to established statistical parameters and metrics consistent with those used in BraSpecS, ensuring methodological symmetry. The training and validation techniques followed a 70/30 split, where 70% of the data was allocated for model training and 30% for performance testing. A k-fold cross-validation was recursed to ensure the robustness of our results. The model is trained 10 times, using 9 subsets for training and 1 for testing at each run. This was followed by calibration and cross-validation for all possible combinations of 10-fold, and identification of the optimal combination based on statistical performance measures. The performance is evaluated as the mean of the results from the 10 iterations, providing a robust and reliable estimate of the model's performance (Hengl et al., 2018). This approach employed the caret package for data partitioning and model training, while *doParallel* facilitated parallel processing to enhance computational efficiency.

The model's performance was assessed using cross-validation techniques provided by *caret*. The importance of spectral variables was visualized with *ggplot2* and *viridis*. The Coefficient of Determination (R^2), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were calculated using a custom *goof* function. This methodology supports robust and reliable model validation by aligning with established practices and ensuring result reproducibility.

2.2.5. *Validation*

We determined the R^2 (Steel and Torrie, 1960), MAE and RMSE (Kenney and Keeping, 1951) as assessment criteria due to their complementary roles in evaluating the performance and accuracy of predictive models. The first one, R^2 (Equation (1)) quantifies the strength of the linear relationship between predicted and observed values. A given R^2 value near 1 suggests a more robust correlation, indicating that the model effectively captures the variability in the data. So, we established an R^2 of 0.7 between the data observed in the laboratory and those modeled via BraSpecS as a feasibility criterion to support policies, particularly the PronaSolos (Polidoro et al., 2021).

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

Where n is the number of observations; yi is the actual value of the dependent variable for the *i*th observation; is the predicted value of the dependent variable from the regression model for the *i*th observation; is the mean of the actual values yi. According to Miller (2017), the MAE is an error measure used to evaluate the accuracy of prediction models, especially in regression analyses. Technically, MAE calculates the average of the absolute differences between the values predicted by the model and the observed values. The formula is shown in Equation (2). In addition, RMSE (Equation (3)) measures the mean magnitude of the errors between predicted and observed values, providing insight into the model's prediction accuracy. Lower RMSE values indicate a model with fewer errors and better predictive performance.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}. \qquad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad \text{Equation (3)}$$

Where is the actual value of the dependent variable for the *i*th observation; is the predicted value of the dependent variable for the *i*th observation; n is the number of observations.

Using both metrics offers a comprehensive assessment, with MAE and RMSE focusing on prediction accuracy and R^2 on correlation strength, ensuring similarity of the methods and robustness to the models' performance evaluation. The manuscript presents two modeling approaches, online (M-I) and offline (M-II), highlighting the advantages of the online method without directly comparing the two. Both models used a 70/30 data split for training and validation, with M-I incorporating spectral and field data, while M-II relied solely on spectral data. Validation for both methods involved recursive k-fold cross-validation to ensure robustness, iteratively repeated 10 times. Performance assessment using consistent metrics support the maintaining of standardized and reproducible evaluation practices. This approach underscores the complementary nature of the methods while emphasizing the online model's benefits. Finally, backscattering charts from the *ggplot2* package represented the dispersion of predicted data from both methods for modeling performance visualization.

3. Results and discussion

3.1. *Physicochemical characteristics and spectral patterns of amazonian soils*

Amazonian soils form complex arrangements influenced by geological, biological, and climatic factors and play a vital role in the rainforest's sustainability (Orellana Segovia et al., 2020). Fig. 2 shows the data through violin plots, integrating the functionality of boxplots with the distributional insight of histograms, and making them effective for visualizing underlying distributions, and comparing multiple distributions simultaneously. Our data revealed that most of the samples were sandy (mean 650 g kg^{-1} of sand, 50 g kg^{-1} of Silt and 300 g kg^{-1} of Clay), while the SOC content ranged from 10 to 150 g kg^{-1}. The results were consistent with the literature (i.e., Schaefer et al., 2023) and influenced both spectral behavior and modeling performance.

3.2. *Spectral patterns*

The spectral signatures revealed by diffuse ER exhibited features characteristic of tropical soils (Fig. 3). Spectral data revealed consistency with soil classes such as Ferralsols and Lixisols. These associations are supported by distinct spectral features in the VNIR-SWIR range, aiding soil classification efforts. The spectral curves are an expression of the composition of the soil concerning the absorption or reflection of light and have supported to understand the physical and chemical properties of the soil. Among soils with a lower degree of pedogenetic evolution, such as Arenosols, Luvisols and Fluvisols, quartz minerals increased the albedo of the curves between 350 and 900 nm. The predominance of sandy texture, rich in quartz minerals, lower OM and water content causes high albedo in the curve, which positions it at third quartile, above 0.45. While the spectral patterns of hydromorphic soils, such as Plinthosols and Gleisols were positioned in the first quartile of the graph (0.35). This behavior results from the clayey texture, OM accumulation and oxides that downgrade the reflectance (Demattê et al., 2014).

The median curve (Fig. 3) shows the representative behavior of tropical climates. According to Novais et al. (2024), these soils are characterized by shaping the spectral curve to present features soils, such as high absorption peaks at 1400, 1950, and 2200 nm caused by residual water. These curves stood out for their high reflectance and peaks characteristic of the presence of kaolinite in absorption at 1390 and 2190 nm. In addition, characteristic traits of the presence of gibbsite appear in 2265 nm (Fig. 3). Features at the end of the spectral curve, between 2200 and 2500 nm, were also observed, indicating the presence of 2:1 minerals such as illite and montmorillonite, which suggest fertile soils.

Correlating observed soil classes with spectral features in the VNIR-SWIR region offers valuable insights into the relationship between soil properties and their spectral responses (Novais et al., 2023, 2024). Distinct mineralogical and organic compositions exhibited by soil classes such as Ferralsols, Acrisols, and Podzols influence their reflectance and absorption patterns across these spectral regions (Viscarra Rossel

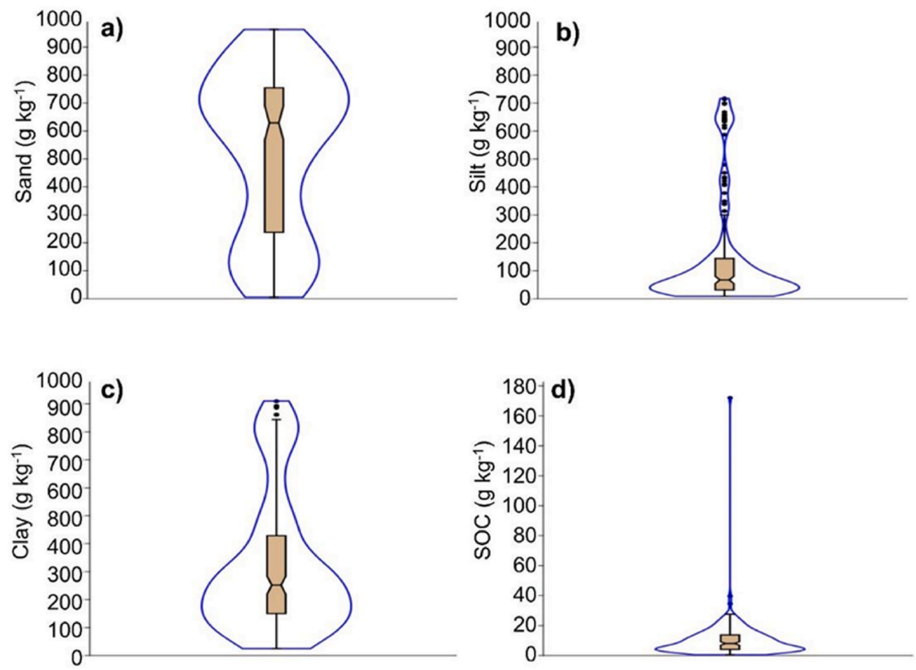

Fig. 2. Observed soil attributes data violin plots of measured soil attributes of the study area. (a) sand, (b) silt, (c) clay, and (d) soil organic carbon. The x-axis represents the soil attribute distribution based on statistical parameters.

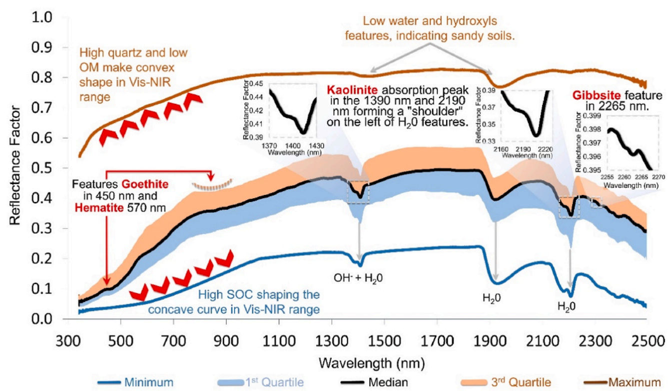

Fig. 3. Characterization of the soil spectral library in the visible and near-infrared (Vis-NIR) to shortwave infrared (SWIR) regions exhibiting a range of reflectance values, from the maximum to the minimum and highlighting the main spectral signatures features.

et al., 2022). For instance, clay-rich soils, such as Ferralsols, exhibited distinct absorption features related to hydroxyl groups in clay minerals, while soils with high OM, such as Gleisols, demonstrate reduced reflectance in the visible range. The presence of iron oxides, which are abundant in Ferralsols, Lixisols, Acrisols and Plinthosols, frequently results in characteristic absorption features in the VNIR-SWIR region due to electronic transitions and charge transfer processes. These features are mainly caused by hematite and goethite in 450 and 570 nm highlighted in red in Fig. 3. The analysis of these spectral characteristics facilitates the association of specific soil classes with their corresponding spectral features, thereby enhancing the accuracy of soil mapping and classification (Novais et al., 2023. This correlation also supports the development of predictive models, enabling the utilization of remote sensing techniques for large-scale soil monitoring and sustainable land management, as stated by Sousa Junior et al. (2011).

3.3. Spectral capacity in modeling soil attributes

Spectral data are apt to represent field truth because VNIR-SWIR wavelengths cover multiple ranges and enable the modeling of soil attributes. Thus, the Pearson Correlation Coefficient (PCC) presented different values throughout the wavelengths (Fig. 4a). This illustration presents the correlation between particle size and SOC throughout the spectral curve, highlighting the various regions with negative correlation, which means that when clay and SOC increase, the albedo decreases (Novais et al., 2024). On the one hand, the entire spectral range studied presented a positive PCC with sand and silt content, indicating a directly proportional relationship as these attributes increase, the albedo also increases. On the other hand, clay and SOC demonstrated the opposite behavior with a negative PCC, signifying that these soil attributes reduce the albedo.

Fig. 4b-e complements the visualization of the relationship between soil attributes and spectral ranges. This representation enables us to observe the importance score wavelength for each soil attribute studied for modeling via the Cubist algorithm, considering the M-II which employed conjointly spectral and soil attributes. In Fig. 4b and c it is possible to note the better range of the visible bands highlighted sand and silt, while the SWIR spectral range was the most important covariate to model the clay (Fig. 4d) and SOC content (Fig. 4e).

3.4. Soil characterization

Upland Amazon soils are highly weathered, with acidic pH, low mobile element content, and reduced cation exchange capacity compared to other Brazilian biomes and tropical rainforests worldwide (Souza et al., 2018). According to Schaefer et al. (2023), in Amazon also occurs fertile soils because the formation of some of these soils involves a dynamic process on sedimentary rocks that is the product of alluvial transport with removal and deposition of mineral and organic material, which causes physical, chemical, and biological changes. In drylands, most soils are highly weathered, a process in which the primary minerals are transformed into Fe and Al sesquioxides, as well as hydrated silicates in different degrees of weathering depending on the pedoenvironmental conditions, which give rise to colloidal clays (Novais et al., 2024).

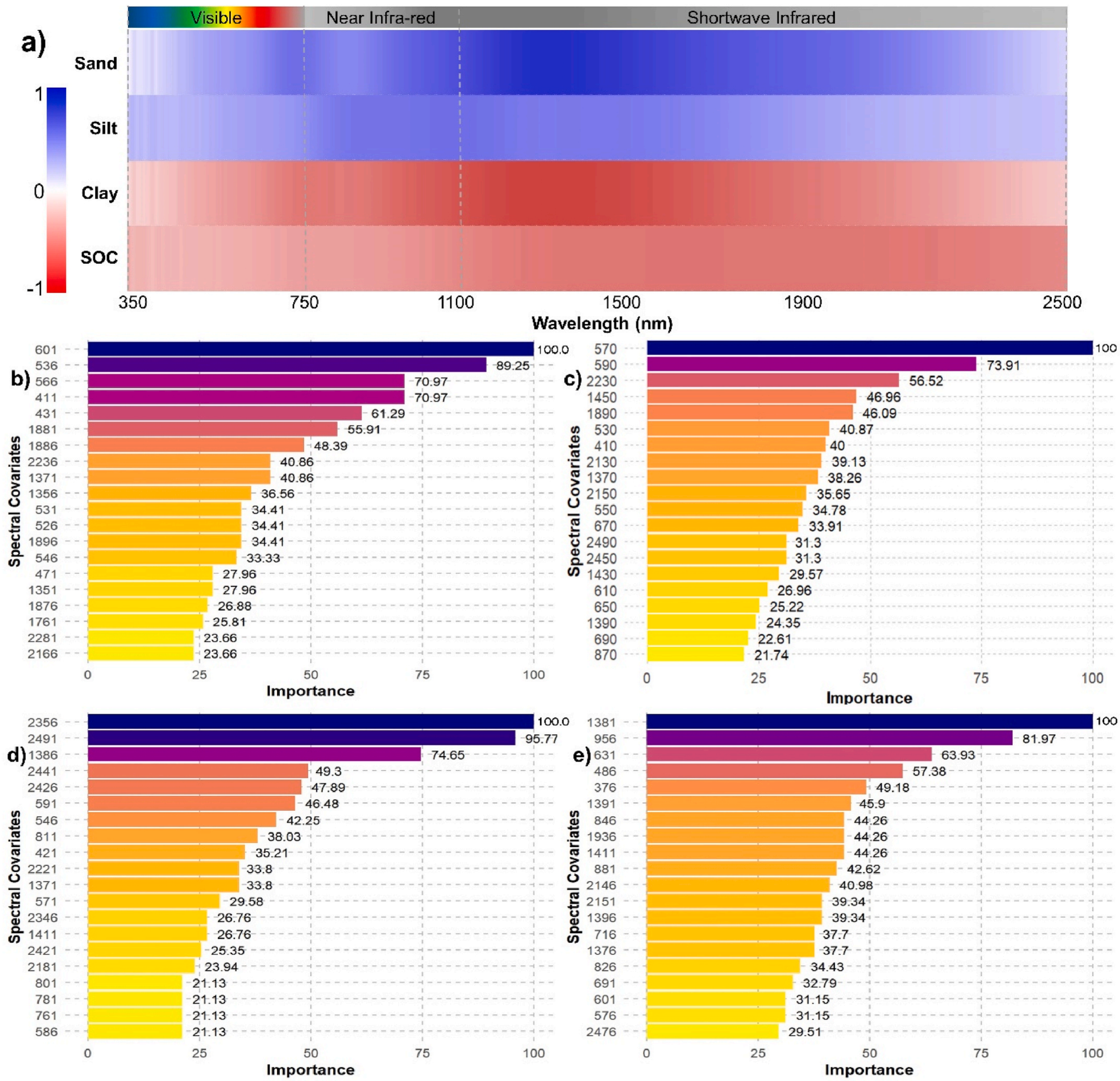

Fig. 4. Spectral capacity in model soil attributes. (a) correlogram throughout the VNIR-SWIR ranges for particle size and soil organic carbon (SOC); Importance ranking of spectral bands for modeling sand (b), silt (c), clay (d), and SOC (e) by cubist model.

Pinheiro et al. (2017) observed that the RS technique makes it possible to identify several soil attributes. The dynamics of OM in the Amazon reveal the diversity of environments since it can originate waste in different degrees of decomposition, such as animals, excreta, roots, trunks, leaves, and flowers (Pinheiro et al., 2017). This OM converts to inorganic combinations, such as ammonium (NH_4^+), phosphate ($H_2PO_4^-$), and sulfate (SO_4^{2-}). The polyphenolic compounds in senescent leaves also affect the speed at which waste decomposes. Araújo et al. (2015), when determining the properties of the "Terra-Preta-de-Índio" (Anthroposols) of the Amazon employing RE, state that this can be a viable alternative to evaluate and model the spatial distribution of key properties of these soils compared to traditional soil analyses. Santos et al. (2023), when characterizing Amazonian soils utilizing RE, detected specific pedogenetic processes such as melanization (darkening of the profile by OM), xanthization (yellowing of the profile by Iron and Aluminum Oxides), and lessivage (transition of clay from surface to subsurface horizon).

3.5. Modeling performances

Both procedures successfully modeled particle size (sand, silt, and clay) and SOC content based on the spectra as predictor covariates. Table 1 presents the summary statistics of observed and predicted data by M-I and M-II. The offline method achieved means of particle size similar to observed mean values, while this parameter was 6.7% lower than field data compared to the online method. These results show the higher predictive capacity of both models (Novais et al., 2024).

Fig. 5a–d illustrates the scattering observed and predicted data by M-

Table 1
Descriptive statistics from observed and predicted data of methods I and II.

	Observed				Method 1 (online)				Method 2 (offline)			
	Sand	Silt	Clay	SOC	Sand	Silt	Clay	SOC	Sand	Silt	Clay	SOC
Mean	530	149	321	8	494	142	363	8	524	159	317	7
Std. Error	22	14	18	1	16	9	13	1	36	26	29	1
Median	629	68	252	6	482	100	347	7	619	71	263	6
Mode	724	26	151	4	712	58	613	2	8	316	454	2
Std. Deviation	292	195	246	10	218	120	178	9	286	190	234	5
S. Variance	85,203	37,841	60,281	110	47,383	14,318	31,623	78	81,200	36,350	54,250	25
Kurtosis	−1	2	0	99	−1	3	−1	88	−1	2	1	6
Asymmetry	0	2	1	9	0	2	0	8	0	2	1	2
Interval	957	708	885	129	867	614	708	106	1005	700	908	27
Minimum	6	9	25	0	51	28	27	1	5	12	7	1
Maximum	963	717	910	129	918	641	735	107	1010	712	915	29
Sum	96,455	27,174	58,373	1438	89,950	25,930	66,108	1486	31,480	9482	19,060	418
Count	211	211	211	211	211	211	211	211	524	159	317	7

Std. Standard; S. sample; SOC: Soil Organic Carbon.

I (salmon dots) and M-II (turquoise blue circles) referent to both datasets. In these illustrations, it is possible to notice the spilling of the predicted data from the observed values. The particle size and SOC models performed near the observed data, indicating high R^2 and low errors for both methods. Nevertheless, The M-I clay content values remained more dispersed compared to the observed data. This result suggests a lower R_2 and higher errors compared to the other attributes. The data dispersion was slightly higher for Sand and Clay, while they were a little less spread for Clay, Silt and SOC, more evident in SOC predicted by M-I. Despite M-II values indicating a lower, they still have significant correlation and acceptable errors. These results are indicative of both methods' effectiveness in soil particle size attributes and SOC content prediction based on spectral data.

Modeling performances are exhibited in Fig. 6. This illustration shows the distribution of R^2 for both online (Fig. 6a) and offline (Fig. 6b) methods between observed and predicted data and their respective MAE and RMSE. Both approaches achieved promising results, with an outstanding M-II prediction with R^2 ranging from 0.79 to 0.85 and an RMSE from 4.57 g kg^{-1} of SOC to 144.76 g kg^{-1} of Sand. Comparatively, Fig. 6a shows that the modeling M-I, via BraSpecS, achieved an R^2 of 0.72 for SOC content, while reaching 0.81 in M-II. These results suggest a potential predictive capacity for both methods, whose outcomes can be enhanced by model calibration and more representative sampling (Demattê et al., 2022).

These findings were consistent with most of the studies consulted that proposed to analyze the soil attributes of the study region. For example, Pinheiro et al. (2017) modeled soils in the western Amazon using remote sensing and found mean clay contents of 257 g kg^{-1} with an R^2 of 0.8 and a 62 g kg^{-1} error in their models. These results are similar to those achieved in the present analysis, concerning the methods and sampling size. Complementarily, to increase the result values, cautious optimization in constructing a robust ML model is necessary, integrating algorithms to provide a more effective and accurate SOC prediction method with VNIR spectroscopy (Santos et al., 2023).

In evaluating the effectiveness of Vis-NIR data for predicting soil attributes of Amazon, Pavão et al. (2024) demonstrated different levels of performance in different ML algorithms, indicating their R^2 and RMSE values. These models slightly varied in soil texture components (sand, clay, and silt) modeling. For Sand, the Random Forest model achieved a relatively high R^2 of 0.87 in the superficial horizon, with an RMSE of 9.58. This indicates a strong predictive capability with relatively low error. Similarly, the Support Vector Machine also yielded a higher R^2, achieving 0.87 for this parameter, and an RMSE of 10.63. These authors attested to the high capacity of ML to predict soil attributes from spectra.

Our findings highlight the slight superiority of the M-II model over M-I utilizing Vis-NIR data for soil texture and SOC prediction. Nevertheless, the online prediction model also achieved higher R^2 values and lower RMSE across different soil components underscoring its robustness and reliability. Meanwhile, the M-I model's comparatively lower R^2 and higher RMSE indicate that it also may perform as well-suited for this application, including for other soil attributes (Demattê et al., 2022). This implies that while Vis-NIR data can be valuable for soil properties prediction, the choice for this modeling method can achieve accurate and reliable results.

Showcasing demonstrate the feasibility and benefits of online analysis without requiring prior observations, we highlighted that the offline model has higher precision but requires observations to train, whereas the online model does not. Continuous improvement and validation of higher-precision models can further enhance remote sensing, particularly spectroscopy use in agriculture and soil management policies.

3.6. *Perspectives for the use of data for policies and sustainable development*

Based on the promising results of this research, we can confirm the potential of spectroscopy to support public policies. Integrating cloud-based modeling and spectroscopy in soil analysis significantly advances data-driven policy-making. By providing accurate, timely, and cost-effective soil data, these technologies can inform various policies, from agriculture and environmental conservation to urban planning and climate change effects mitigation (Viscarra Rossel et al., 2022).

Integrating cloud-based modeling and spectroscopy, particularly in initiatives such as the PronaSolos, holds significant potential for influencing and shaping policy decisions. PronaSolos aims to map the entire Brazilian territory (8.5 million km^2) by 2048, utilizing scales ranging

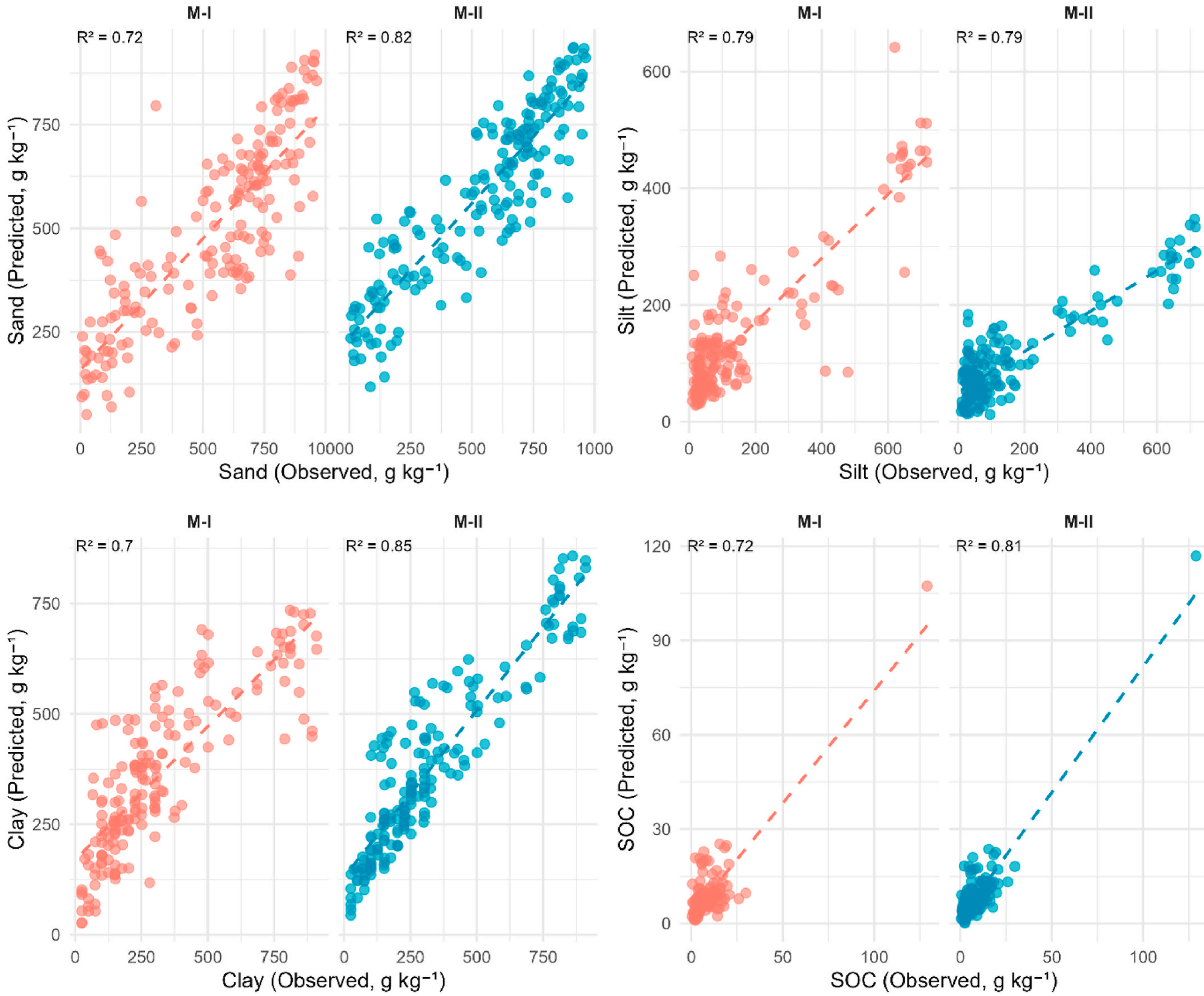

Fig. 5. Backscattering charts of predicted and observed data for particle size (a–c) and soil organic carbon (d) by both methods M-I (online, salmon) and M-II (offline, light blue) from each dataset. Coefficient of determination (R^2); SOC: Soil Organic Carbon.

from 1:25,000 to 1:100,000 (Polidoro et al., 2016). This ambitious program has adopted remote sensing and other advanced methodologies, considered unconventional before the 2000s, but has now become widely accepted due to advancements and improvements in accuracy over the past two decades (Poppiel et al., 2022).

A primary advantage of cloud-based modeling and spectroscopy is the cost-effectiveness and efficiency in soil analysis. Traditional soil analysis methods are expensive and time-consuming. For instance, some wet laboratories' price list indicates that the average cost for soil classification, including sorptive complex, texture, and OM, is about US$ 30 per sample, with a turnaround time of approximately 10–20 days. If we consider the operationalization in Amazon, this should take much more than one month. Given the need for 211 samples, the cost escalates to US $ 6330. Sousa Junior et al. (2011) performed this calculation and estimated a significant cost reduction with soil physicochemical analyses. From these, carbon is a bottleneck for soil classification, soil health and climate change. The current soil analysis uses sulphuric acid and dichromate of potassium. Indeed, inside the $30 indicated for all soil analysis, organic matter is responsible for 70% of this cost. This impacts the main soil attribute necessary for understanding the Amazon situation. Using Braspecs models and new upcoming ones, this factor will drop and give more chances to new scientific studies in the region.

In contrast, remote sensing techniques offer a faster and more economical alternative. Using RS, the exact number of samples can be analyzed in hours or minutes, significantly reducing the costs and time involved. Even when using 10% of traditional analyses to validate the model, the savings in both time and resources are substantial. This efficiency gain is crucial for large-scale initiatives like PronaSolos (Polidoro et al., 2016), enabling more comprehensive and timely soil mapping and analysis.

The data generated from these advanced methodologies have profound implications for policy-making. Accurate and timely soil information can inform various policies, including: 1. Agricultural Policies: Enhanced soil data can lead to improved land management practices, optimizing crop yields and ensuring sustainable agricultural practices. These workings can drive policies focused on food security and sustainable farming. 2. Environmental Policies: Detailed soil maps can help conserve ecosystems and biodiversity. Policies can be developed to protect areas with vulnerable soil types and mitigate the effects of soil erosion and degradation.

Additionally, 3. Urban Planning and Infrastructure: Accurate soil data can inform urban development projects, ensuring that construction

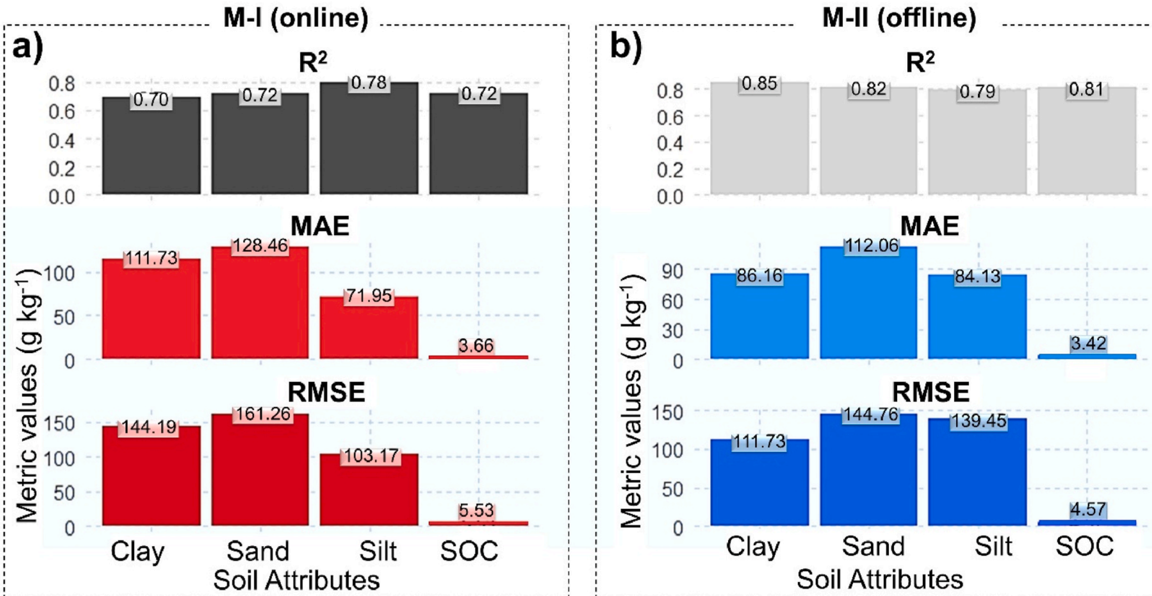

Fig. 6. Coefficient of determination (R^2) and metrics for modeling of particle size and soil organic carbon by modeling methods. a) M–I: Method I (online); b) M-II: Method II (offline); SOC: soil organic carbon; MAE: Mean Absolute Error; RMSE: Root Square Mean Error.

is carried out on suitable land, reducing the risk of structural failures, and enhancing infrastructure longevity; and 4. Climate Change Mitigation: Soil data can play a critical role in understanding and mitigating the impacts of climate change. Policies can be formulated to enhance soil carbon sequestration and manage land use to reduce greenhouse gas emissions. Therefore, the perspectives for the future of policy development will increasingly rely on such advanced methodologies to ensure sustainable and informed decision-making.

The integration of reflectance spectroscopy and cloud computing in the analysis of Amazonian soils directly aligns with several United Nations Sustainable Development Goals (SDGs), particularly SDG 13 (Climate Action), SDG 15 (Life on Land), and SDG 2 (Zero Hunger) (United Nations, 2015). By enabling precise and large-scale monitoring of soil health, this technological approach supports the sustainable management of land resources, crucial for maintaining biodiversity and combating deforestation in the Amazon (SDG 15). The data generated can guide climate-resilient agricultural practices, contributing to food security (SDG 2) while mitigating the impacts of climate change by promoting carbon sequestration and reducing land degradation (SDG 13). Moreover, the use of cloud computing facilitates the widespread dissemination of this valuable information, empowering policymakers, researchers, and local communities to make informed decisions that foster sustainable development in one of the world's most critical ecosystems.

4. Conclusions

Results demonstrated the effectiveness of reflectance spectroscopy and cloud computing to provide faster, cost-effective, and environmentally friendly soil analyses. The NIR and SWIR regions (350–2500 nm) showed high sensitivity to key soil properties, particularly clay minerals, iron oxides and SOM composition, enabling precise soil characterization. Moreover, the synergistic use of these techniques can support traditional analysis in remote and large areas such as ARB, where wet labs are scarce. This work confirms the correlation between matter and the electromagnetic energy reflected in the Vis-NIR-SWIR band. Thus, this relationship allowed us to characterize the soil samples in terms of composition by evaluating the data collected with a sensor. In this way, we identified the key features of the spectral curves of the soils, which provide modeling of the attributes.

The insertion of SSL in the BraSpecS system resulted in the rapid and consistent analysis of soil samples from the Amazon. This information reiterates the possibility of obtaining soil analysis results via ER and cloud processing reducing the reliance on traditional laboratory methods. However, some of these analyses are important because they feed the data population of the field observations system for model calibration. Our approach highlights the analysis method, especially improving accessibility to analysis through soil spectroscopy and cloud computing. As wet laboratories are scarce in the region, cloud computing and soil spectroscopy can assist the stakeholders in analyzing the soils. In addition, traditional laboratory analyses provide more specific data in case of the need for repeat testing. For successful online prediction, two key factors are fundamental: a – that the user has a spectral sensor for data collection in the field or laboratory; b – that the server where the primary data is hosted covers the patterns of unknown and sufficiently diversified samples. In this case, BraSpecS already has an internal SSL with 49,000 samples from all Brazilian states (and is continuously updating on data and methods, now going towards deep learning).

Therefore, the results are promising, particularly for predicting particle size and SOC. The key to the success of this initiative lies in the population of cloud-loaded analytics (BasSpecS). The use of RE can significantly enhance PronaSolos by improving the efficiency of soil analysis while providing information quickly and with minimal environmental impact. The prospects are that the user can obtain spectral readings in the field, send them via the internet to BraSpecS, and receive the result in real-time. For this, continuous investments are needed to optimize the platform. Shortly, geotechnologies will revolutionize soil study, and when combined with artificial intelligence and the Internet of Things, they will offer promising approaches for advancing sustainability natural resources surveys. In the future, research should concentrate on increasing the size of spectral libraries and using cloud-based platforms to improve scalability and accuracy. This approach is different from conventional methods because it offers real-time analysis, which results in a significant reduction in time and costs.

CRediT authorship contribution statement

Jean Jesus Macedo Novais: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Borges Marfrann Dias Melo:** Writing – review & editing, Writing – original draft, Data curation. **Afrânio Ferreira Neves Junior:** Writing – review & editing, Writing – original draft. **Raimundo Humberto Cavalcante Lima:** Writing – review & editing, Writing – original draft. **Renato Epifânio de Souza:** Writing – review & editing, Writing – original draft. **Valdinar Ferreira Melo:** Writing – review & editing, Writing – original draft. **Eufran Ferreira do Amaral:** Writing – review & editing, Writing – original draft. **Nikolaos Tziolas:** Writing – review & editing, Writing – original draft. **José A.M. Demattê:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank Dr. Lumbreras, J.F., for giving in data on behalf of the XV RCC. The technical support provided by the members of the Geotechnologies in Soil Science group (https://esalqgeocis.wixsite.com/geocis) is greatly appreciated. We would like to thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support in the projects 2022/14935-0, 2021/05129-8 and 88887.840324/2023–00, respectively.

Data availability

Data will be made available on request.

References

Adak, S., Mandal, N., Maity, P.P., Mukhopadhyay, A., Gunturi, A., 2023. Digital soil health: unveiling the potential of soil health monitoring. Food and Scientific Reports 4 (9), 96–103.

Araújo, S.R., et al., 2015. Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy. Geoderma 237–238, 308–317. https://doi.org/10.1016/j.geoderma.2014.09.014. ISSN 0016-7061.

ASD Inc., 2019. ASD Fieldspec® 4: The Industry-Leading Portable Device for Field Spectroscopy, 6th ed.; Analytical Spectral Device Inc.: Boulder, CO, Available in: USA. https://www.malvernpanalytical.com/en/products/product-range/asd-range/fieldspec-range/fieldspec-4-hi-res-ng-spectroradiometer Accessed in September 2024.

Ben Dor, E., Ong, C., Lau, I.C., 2015. Reflectance measurements of soils in the laboratory: standards and protocols. Geoderma 245–246, 112–124. https://doi.org/10.1016/j.geoderma.2015.01.002.

Ben Dor, E., Efrati, B., Amir, O., Francos, N., Shepherd, J., Khosravi, V., Gholizadeh, A., Klement, A., Borůvka, L., 2024. A standard and protocol for in-situ measurement of surface soil reflectance. Geoderma 447, 116920. https://doi.org/10.1016/j.geoderma.2024.116920.

Demattê, J.A.M., et al., 2014. Morphological Interpretation of Reflectance Spectrum (MIRS) using libraries looking towards soil classification. Soils and Plant Nutrition, Sci. agric. 71 (6). https://doi.org/10.1590/0103-9016-2013-0365.

Demattê, J.A.M., Dotto, A.C., Paiva, A.F.S., Sato, M.V., Dalmolin, R.S.D., Araújo, M.S.B. de, Silva, E.B. da, Nanni, M.R., ten Caten, A., Noronha, N.C., Lacerda, M.P.C., Araújo Filho, J.C. de, Rizzo, R., Bellinaso, H., Francelino, M.R., Schaefer, C.E.G.R., Vicente, L.E., dos Santos, U.J., Sampaio, E.V. de Sá B., Menezes, R.S.C., de Souza, J.J.L. de, Abrahão, W.A.P., Coelho, R.M., Grego, C.R., Lani, J.L., Fernandes, A.R., Gonçalves, D.A.M., Silva, S.H.G., Menezes, M.D. de, Curi, N., Couto, E.G., dos Anjos, L.H.C. dos, Ceddia, M.B., Pinheiro, É.F.M., Grunwald, S., Vasques, G.M., Marques Júnior, J., da Silva, A.J., de Vasconcelos Barreto, M.C. de, Nóbrega, G.N., da Silva, M.Z., de Souza, S.F., Valladares, G.S., Viana, J.H.M., Terra, F.S., 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., Mendes, W. de S., Safanelli, J.L., Guimarães, C.C.B., Poppiel, R.R., Souza e Souza, A.B., Quesada, C.A., Zarate do Couto, H.T., 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., et al., 2022. The Brazilian soil spectral Service (BraSpecS): a user-friendly system for global soil spectra communication. Rem. Sens. 14, 740. https://doi.org/10.3390/rs14030740.

Gentry, A.H., 1992. Tropical forest biodiversity: distributional patterns and their conservational significance. Oikos 63 (1), 19–28. https://doi.org/10.2307/3545512.

Herranz-Luque, J.E., Gonzalez-Canales, J., Martín-Sanz, J.P., Antón, O., Moreno-Delafuente, A., Navas-Vázquez, M.J., Ramos-Nieto, R., Bienes, R., García-Díaz, A., Marques, M.J., Blanca, Sastre, 2024. Monitoring land management practices using vis–NIR spectroscopy provides insights into predicting soil organic carbon and limestone levels in agricultural plots. Agronomy 14, 1150. https://doi.org/10.3390/agronomy14061150.

Hengl, T., Nussbaum, M., Wright, Mn, Heuvelink, G.B.M., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ V. https://doi.org/10.7287/peerj.preprints.26693v2.

Hoorn, C., Wesselingh, F.P., 2010. Amazonia: Landscape and Species Evolution. Wiley-Blackwell. https://doi.org/10.1002/9781444306408.fmatter.

Irion, G., 1978. Soil infertility in the amazonian rainforest. Naturwissenschaften 65, 515–519. https://doi.org/10.1007/BF00439791.

IUSS Working Group WRB, 2022. World Reference Base for Soil Resources. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps, fourth ed. International Union of Soil Sciences (IUSS), Vienna, Austria, p. 236.

Kenney, J.F., Keeping, E.S., 1951. Mathematics of Statistics, second ed. Van Nostrand, p. 429. 1951.

Liu, S., Shen, H., Chen, S., Zhao, X., Biswas, A., Jia, X., Shi, Z., Fang, J., 2019. Estimating forest soil organic carbon content using vis-NIR spectroscopy: implications for large-scale soil carbon spectroscopic assessment. Geoderma 348, 37–44. https://doi.org/10.1016/j.geoderma.2019.04.003.

Marengo, J.A., 2006. On the hydrological cycle of the Amazon Basin: a historical review and current state-of-the-art. Revista Brasileira de Meteorologia 21 (3).

McBratney, A.B., Minasny, B., 2018. In: Stockmann, U. (Ed.), Pedometrics, 1st ed., XIII. Springer, p. 720 https://doi.org/10.1007/978-3-319-63439-5.

Miller, J.D., 2017. Statistics for Data Science. Leverage the Power of Statistics for Data Analysis, Classification, Regression, Machine Learning, and Neural Networks. Packt Publishing, Birmingham, UK, p. 279.

Nanni, M.R., Demattê, 2006. Spectral reflectance methodology in comparison to traditional soil analysis. Soil Sci. Soc. Am. J. 70, 393–407. https://doi.org/10.2136/sssaj2003.0285.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Ben Dor, E., 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., et al., 2015. Soil spectroscopy: An alternative to wet chemistry for soil monitoring 139–159. https://doi.org/10.1016/bs.agron.2015.02.002.

Novais, J.J., Rosin, N.A., Rosas, J.T., Poppiel, R.R., Dotto, A.C., Paiva, A.F., Bellinaso, H., Albarracín, H.S., Amorim, M.T., Bartsch, B., Vogel, L.G., Mello, D.C., Francelino, M. R., Alves, M.R., Falcioni, R., Demattê, J.A.M., 2024. The Brazilian soil spectral library data opening. Dokuchaev Soil Bulletin 119, 261–305. https://doi.org/10.19047/0136-1694-2024-119-261-305.

Novais, J.J., et al., 2023. VNIR-SWIR spectroscopy, XRD and traditional analyses for pedomorphogeological assessment in a tropical toposequence. AgriEngineering 5, 1581–1598. https://doi.org/10.3390/agriengineering5030098.

Orellana Segovia, J.F., et al., 2020. Características Físico-Químicas Dos Principais Solos Na Amazônia. Embrapa Amapá. https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1129378/1/CPAF-AP-2020-cap-3-Caracteristicas-fisico.pdf.

Pavão, Q.S., Ribeiro, P.G., Maciel, G.P., Silva, S.H.G., Araújo, S.R., Fernandes, A.R., Demattê, J.A.M., Souza Filho, P.W.M., Ramos, S.J., 2024. Texture prediction of natural soils in the Brazilian Amazon through proximal sensors. Geoderma Regional 37, e00813. https://doi.org/10.1016/j.geodrs.2024.e00813.

Pinheiro, E.F.M., et al., 2017. Prediction of soil physical and chemical properties by visible and near-infrared diffuse reflectance spectroscopy in the central Amazon. Rem. Sens. 9, 293. https://doi.org/10.3390/rs9040293.

Polidoro, J.C., Mendonça-Santos, M. de L., Lumbreras, J.F., Coelho, M.R., Carvalho Filho, A. de, Motta, P.E.F. da, Carvalho Junior, W. de, Araujo Filho, J.C. de, Curcio, G.R., Correia, J.R., Martins, E. de S., Spera, S.T., Oliveira, S.R. de M., Bolfe, E.L., Manzatto, C.V., Tosto, S.G., Venturieri, A., Sa, I.B., Oliveira, V.A. de, Shinzato, E., Anjos, L.H.C. dos, Valladares, G.S., Ribeiro, J.L., Medeiros, P.S.C. de, Moreira, F.M. de S., Silva, L. S.L., Sequinatto, L., Aglio, M.L.D., Dart, R. de O., 2016. Programa Nacional de Solos do Brasil (PronaSolos). Embrapa Soils. (Documents 183). Rio de Janeiro. 53p.

Polidoro, J.C., Coelho, M.R., Carvalho Filho, A. de, Lumbreras, J.F., Oliveira, A. P. de, Vasques, G. de M., Macario, C.G. do N., Victoria, D. de C., Bhering, S.B., Freitas, P. L. de, Quartaroli, C.F., Brefin, M. de L.M. S. (Eds.), 2021. Programa Nacional de Levantamento e Interpretação de Solos do Brasil (PronaSolos): diretrizes para implementação. Embrapa Solos. E-book. (Embrapa Solos. Documentos, Rio de Janeiro, 225). Available in: https://ainfo.cnptia.embrapa.br/digital/bitstream/item/226841/1/CNPS-DOC-225-2021.epub, 03 mai. 2024.

Poppiel, R.R., et al., 2022. Bridging the gap between soil spectroscopy and traditional laboratory: insights for routine implementation. Geoderma 425, 116029. https://doi.org/10.1016/j.geoderma.2022.116029. ISSN 0016-7061.

R Core Team, 2024. R: Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.r-project.org/.

Rocchini, D., Boyd, D.S., Féret, J.B., Foody, G.M., He, K.S., Lausch, A., et al., 2016. Satellite remote sensing to monitor species diversity: potential and pitfalls. Remote Sensing in Ecology and Conservation 2 (1), 25–36. https://doi.org/10.1002/rse2.9.

Santos, H.G., et al., 2018. Brazilian soil classification system. 5. In: Rev. & Ampl. Embrapa, Brasília, DF, p. 356. https://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1094003.

Santos, H.G., et al., 2011. O novo mapa de solos do Brasil: legenda atualizada (Embrapa Solos. Documentos, 130). Embrapa Solos. http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/920267.

Santos, L.A.C., et al., 2023. Magnetic and spectral signature of anthropogenic and non-anthropogenic soils to identify pedogenetic processes in Amazonia, Brazil. Regional Geoderm 32, e00622. https://doi.org/10.1016/j.geodrs.2023.e00622. ISSN 2352-0094.

Schaefer, C.E.G.R., et al., 2023. Soils from Brazilian amazonia. In: Schaefer, C.E.G.R. (Ed.), The Soils of Brazil. Springer Nature, Switzerland AG, pp. 85–129. https://doi.org/10.1007/978-3-031-19949-3.

Smith, N., 1982. Rainforest Corridors: the Transamazon Colonization Scheme. University of California Press, p. 268. https://doi.org/10.2307/jj.13167885.

Soil Survey Staff, 2017. Soil Survey Manual (Handbook No. 18). United States Department of Agriculture - USDA. Minor amendments February 2018. https://ncsslabdatamart.sc.egov.usda.gov/labmanuals.aspx

Sombroek, W., 2001. Spatial and temporal patterns of Amazon rainfall. Ambio 30 (7), 388–396. https://doi.org/10.1579/0044-7447-30.7.388.

Souza, J.J.L.L., Fontes, M.P.F., Gilkes, R., Costa, L.M., Oliveira, T.S., 2018. Geochemical signature of Amazonian tropical rainforest soils. Rev Bras Cienc Solo 42, e0170192. https://doi.org/10.1590/18069657rbcs20170192.

Sousa Junior, J.G., Demattê, J.A.M., Araújo, S.R., 2011. Terrestrial and orbital spectral models in the determination of soil attribute contents: potential and costs. Soils and Plant Nutrition 70 (3). https://doi.org/10.1590/S0006-87052011000300017, 2011.

Steel, R.G.D., Torrie, J.H., 1960. Principles and Procedures of Statistics with Special Reference to the Biological Sciences. McGraw Hill, New York, pp. 187–287.

United Nations, 2015. Transforming our world: the 2030 agenda for sustainable development. Resolution Adopted by the General Assembly on 25 September 2015 42809, 1–13. https://doi.org/10.1007/s13398-014-0173-7.2.

Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Chabrillat, S., et al., 2022. Diffuse reflectance spectroscopy for estimating soil properties: a technology for the 21st century. Eur. J. Soil Sci. 73 (4), e13271. https://doi.org/10.1111/ejss.13271.