UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS AGRÁRIAS
DEPARTAMENTO DE CIÊNCIAS DO SOLO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO

SHARON GOMES RIBEIRO

REFLECTANCE SPECTROSCOPY AND REMOTE SENSING FOR PREDICTION OF SOIL ORGANIC CARBON AND TOTAL NITROGEN ACROSS DIFFERENT AGRICULTURAL CONTEXTS

FORTALEZA, CE
2025

SHARON GOMES RIBEIRO

REFLECTANCE SPECTROSCOPY AND REMOTE SENSING FOR PREDICTION OF SOIL ORGANIC CARBON AND TOTAL NITROGEN ACROSS DIFFERENT AGRICULTURAL CONTEXTS

Tese apresentada ao Programa de Pós-Graduação em Ciência do Solo da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutora em Ciência do Solo. Área de concentração: Pedologia.

Orientador: Prof. Adunias dos Santos Teixeira, Ph.D

FORTALEZA

2025

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R372r Ribeiro, Sharon Gomes.
Reflectance Spectroscopy and Remote Sensing for prediction of Soil Organic Carbon and Total Nitrogen across different agricultural contexts. / Sharon Gomes Ribeiro. – 2025.
131 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-Graduação em Ciência do Solo, Fortaleza, 2025.
Orientação: Prof. Dr. Adunias dos Santos Teixeira.

1. hyperspectral imaging. 2. organic attribute monitoring. 3. multivariate modeling. I. Título.
CDD 631.4

SHARON GOMES RIBEIRO

REFLECTANCE SPECTROSCOPY AND REMOTE SENSING IN PREDICTION OF SOIL ORGANIC CARBON AND TOTAL NITROGEN ACROSS DIFFERENT AGRICULTURAL CONTEXTS

Tese apresentada ao Programa de Pós-Graduação em Ciência do Solo da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutora em Ciência do Solo. Área de concentração: Pedologia.

Aprovada em: 16/06/2025.

BANCA EXAMINADORA

_____
Prof. Adunias dos Santos Teixeira, PhD. (Orientador)
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Mirian Cristina Gomes Costa
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Pedro Henrique Augusto Medeiros
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

_____
Daniele Zaccaria, PhD.
University of California, Davis (UCDavis)

_____
Elia Scudiero, PhD.
University of California, Riverside (UCR)

_____
Mae Culumber, PhD.
University of California, Agriculture and Natural Resources (UCANR)

A Deus.

Aos meus pais, Raimunda e Valdivino.

ACKNOWLEDGMENTS

First, I thank God and the positive energies that have guided me to this moment, granting me strength, persistence, and resilience to face challenges that my former self could never have imagined overcoming.

To my dear parents, Raimunda Gomes and Valdivino Miguel, and to my grandmother Joana Maria, I express my deepest gratitude for your unconditional love, affection, trust, and constant partnership, as well as your immense dedication to providing me with the best education and always seeing me reach the best place I could be.

Special thanks to my cousin Melissa Gomes, who inspired me to follow the paths of Agronomy and life, being a fundamental presence throughout my journey.

To my partner, Lucas de Lima Farias, I offer my immense gratitude for encouraging me to constantly challenge myself and grow, for believing in me even in the hardest moments, for being my strength when I doubted myself, and for walking side by side along this path.

To my "Agronomy's gift" — Jesimiel Viana, Idayane Souza, Rubens Zimmermann, Júlia Queiros, and Leonardo Formiga — thank you for always being present, making academic and personal challenges lighter and more bearable.

To the friendships that blossomed during graduate school, especially Eveline Menezes, Gustavo Andrade, Gilbenes Bezerra, Felipe Hermínio, Márcio Regys, Felipe Gonçalves, and José Felipe, I express my appreciation for your constant support, companionship during difficult times, and for contributing in countless ways to make this work a reality.

To my advisor since my master's degree, Professor Adunias Teixeira, I am deeply grateful for the trust you employed in me, for your valuable teachings, for the knowledge shared, and above all, for your patience and unwavering support throughout my academic development.

To the evaluation committee in Brazil, Professor Mirian Costa and Professor Pedro Medeiros, I thank you for your dedication, patience, and valuable contributions that positively influenced my professional and personal growth.

To the evaluation committee in California — Daniele Zaccaria, Elia Scudiero, and Mae Culumber — I sincerely thank you for your support, patience, and guidance during my

stay in United States. You made my international experience richer, more welcoming, and transformative, greatly contributing to my development.

To everyone in the “Zac Lab” team, I am thankful not only for welcoming me and making my six-month stay in Davis pleasing and productive but also for your generosity and support that went beyond the professional way, becoming a true network of friendship and encouragement.

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the PDSE (Programa Doutorado-Sanduiche no Exterior) for funding the international academic mobility through public notice number 30/2023, which enabled me to develop the third chapter of this research and experience the international opportunity in the United States of America.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Por fim, agradeço pelo meu idioma materno, o bom português, que foi meu refúgio nos momentos longe do Brasil e o conforto que me conectou com minhas raízes mesmo estando a quase 10.000 km de distância.

ABSTRACT

Techniques for monitoring soil attributes have rapidly evolved, driven by the demand for more efficient, sustainable, and large-scale applicable methods. Soil organic carbon (SOC) and total nitrogen (TN) are essential components for ecosystem functioning. Although traditional laboratory methods are precise, they require considerable time, resources, and labor. Reflectance spectroscopy combined with statistical modeling emerges as a promising tool, allowing estimation of soil attributes from spectral data. This study evaluated the predictive performance of reflectance spectroscopy through three approaches: i) using the hyperspectral satellite EnMAP; ii) with the proximal ASD FieldSpec® sensor directly in the soil surface; and iii) in the laboratory with dried soil samples. The research was conducted in two regions: the Lower Acaraú Irrigation District (R1) in northeastern Brazil, and California's Central Valley (R2), USA. Surface soil samples were collected from the first centimeter in R1 and from 0–10 cm in R2. Laboratory spectral analysis was performed on both sets in darkroom after the oven drying. SOC and TN quantification was done by titration in R1 and by elemental analyzer in R2. Predictive models were developed for each regional set and for a combined global set (R1&R2), using raw reflectance, first derivative, and continuum-removed spectra. The predictive performance of EnMAP and in situ FieldSpec sensor was assessed exclusively in R1, using images obtained by EnMAP. Near the imaging dates, spectral analysis with FieldSpec was conducted at the point corresponding to the central coordinate of the pixel representing exposed soil in the EnMAP image prior to sampling, enabling direct comparison between spectral responses from the orbital and proximal sensors. Results indicate that the EnMAP satellite showed good correlations with spectral data collected by multispectral sensors, with R^2 between 0.50 and 0.86 for individual bands and NDVI. SOC and TN prediction by EnMAP was comparable to that of the FieldSpec sensor in the field, with adj. $R^2 = 0.75$ for SOC and 0.61 for TN. Accuracy improved for SOC > 11.6 g/kg and TN > 0.55 g/kg, reaching adj. $R^2 = 0.91$ and 0.82, respectively. Statistical modeling confirmed the efficiency of reflectance spectroscopy to predict SOC and TN in both local contexts and the global model, regardless of the laboratory methods used. The first derivative transformation was most effective in producing reliable models, especially for R2 using PLSR, with adj. $R^2 = 0.91$ and 0.92 for SOC and RPD of 3.34 and 3.49 for TN. The global R1&R2 set also showed good predictive performance with PLSR (adj. $R^2 = 0.82$ and 0.90; RPD = 2.09 and 2.86 for SOC and TN, respectively). These results reinforce the potential of reflectance spectroscopy and hyperspectral remote sensing for precise and efficient monitoring of organic soil attributes,

offering viable alternatives for diverse agricultural realities and promoting advances in sustainable soil management practices.

Keywords: hyperspectral imaging; organic attribute monitoring; multivariate modeling

RESUMO

As técnicas para monitoramento de atributos do solo vêm evoluindo rapidamente, impulsionadas pela demanda por métodos mais eficientes, sustentáveis e aplicáveis em larga escala. O carbono orgânico do solo (SOC) e o nitrogênio total (TN) são componentes essenciais para o funcionamento dos ecossistemas. Embora métodos laboratoriais tradicionais sejam precisos, eles demandam tempo, recursos e mão de obra consideráveis. A espectroscopia de reflectância, combinada com modelagem estatística, surge como aliada promissora, permitindo estimar atributos do solo a partir de dados espectrais. Este estudo avaliou o desempenho preditivo por espectroscopia de reflectância por três abordagens: *i*) com o satélite hiperespectral EnMAP; *ii*) com o sensor proximal ASD FieldSpec® em campo e *iii*) em laboratório, com amostras de solo secas. A pesquisa foi conduzida em duas regiões: o Distrito de Irrigação do Baixo Acaraú (R1), no nordeste do Brasil, e o Vale Central da Califórnia (R2), EUA. Amostras superficiais foram coletadas no primeiro centímetro do solo em R1 e de 0-10 cm em R2. A análise espectral em laboratório foi realizada em ambos os conjuntos após secagem em estufa. A quantificação de SOC e TN ocorreu por titulação em R1 e por analisador elementar em R2. Modelos preditivos foram desenvolvidos para cada conjunto regional e para um conjunto global combinado (R1&R2), utilizando reflectância bruta, primeira derivada e espectro com contínuo removido. A avaliação da performance preditiva do EnMAP e do sensor FieldSpec in situ foi realizada exclusivamente em R1, utilizando imagens obtidas pelo EnMAP. Próximo às datas do imageamento, foi feita análise espectral com o FieldSpec no ponto correspondente à coordenada central do pixel da imagem referente ao solo exposto, antes da coleta das amostras, permitindo a comparação direta entre as respostas espectrais obtidas pelo sensor orbital e pelo sensor proximal in situ. Os resultados indicam que o satélite EnMAP apresentou boas correlações com dados espectrais coletados por sensores multiespectrais, com R^2 entre 0,50 e 0,86 para bandas individuais e NDVI. A predição de SOC e TN pelo EnMAP foi comparável à do sensor FieldSpec em campo, com $R^2_{aj.}$ = 0,75 para SOC e 0,61 para TN. A acurácia aumentou para SOC > 11,6 g/kg e TN > 0,55 g/kg, alcançando $R^2_{aj.}$ = 0,91 e 0,82, respectivamente. A modelagem estatística confirmou a eficiência da espectroscopia de reflectância para predizer SOC e TN em ambos os contextos locais e no modelo global, independentemente dos métodos laboratoriais usados. A transformação por primeira derivada foi a mais eficaz para gerar modelos confiáveis, especialmente para R2 na PLSR, com $R^2_{aj.}$ = 0,91 e 0,92 para SOC e RPD de 3,34 e 3,49 para TN. O conjunto global R1&R2 também mostrou bom desempenho preditivo na PLSR ($R^2_{aj.}$ = 0,82 e 0,90; RPD = 2,09 e 2,86 para SOC e TN, respectivamente). Esses

resultados reforçam o potencial da espectroscopia de reflectância e do sensoriamento remoto hiperespectral para monitoramento preciso e eficiente de atributos orgânicos do solo, oferecendo alternativas viáveis para diferentes realidades agrícolas e promovendo avanços em práticas sustentáveis de manejo do solo.

Palavras-chave: imageamento hiperespectral; monitoramento de atributos orgânicos; modelagem multivariada

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CCV	California's Central Valley
DIBAU	*Distrito de Irrigação do Baixo Acaraú* (Lower Acaraú Irrigation District)
NIR	Near Infrared
Nm	Nanometer
MLR	Multiple Linear Regression
MPPE	Mean Percentage Prediction Error
PLSR	Partial Least Square Regression
RMSE	Root Mean Square Error
RPD	Ratio of Deviation Performance
SOC	Soil Organic Carbon
SWIR	Short-Wave Infrared
TN	Total Nitrogen
VIS	Visible

SUMMARY

1 GENERAL INTRODUCTION

Organic matter added to the soil is primarily responsible for the availability of organic carbon and nitrogen (both mineral and organic) in the soil environment. These elements constitute a significant structural part of organic compounds and serve as important energy sources for microbial biomass during decomposition.

Quantification of C and N in soil is carried out using classical and accurate laboratory methods. Organic carbon is typically determined by wet digestion with potassium dichromate and sulfuric acid, involving external heating followed by titration with sulfate (Yeomans; Bremner, 1988). Total nitrogen is measured by digestion in acidic medium at high temperature, followed by distillation with sodium hydroxide, boric acid, and hydrochloric acid, and subsequent titration (Kjeldahl, 1883). Besides wet digestion, both elements can also be quantified by dry combustion at temperatures above 1000ºC using an elemental analyzer, which measures C and N as CO_2 and NO_2. This technique is rapid, precise, and clean, albeit requiring substantial maintenance (Farina et al., 1991; Even et al., 2025).

Although accurate, traditional chemical analyses for determining soil organic carbon (SOC) and total nitrogen (TN) require extensive sampling and multiple repetitions. Furthermore, they consume large amounts of harmful reagents, demand considerable time to obtain results, and require substantial human and financial resources, which can impose significant costs on laboratories with limited workforce. To overcome labor challenges and reduce costs associated with analyses and frequent equipment maintenance, methodologies based on reflectance spectroscopy and remote sensing have emerged as complementary approaches to classical laboratory methods.

The study of soil spectral behavior, combined with multivariate statistical methods and mathematical modeling, allows for the estimation of chemical, physical, and biological soil attributes from the spectral response of a sample to incident light. Reflectance factors, with or without mathematical transformations, at each wavelength of the electromagnetic spectrum, function as predictor variables X. Selecting and combining the most influential variables X enables the estimation of SOC and TN contents in the sample.

Given that each environment has specific soil characteristics, it is necessary to adapt prediction models to the most similar conditions possible. Therefore, regional and global models using spectral data from diverse samples have been evaluated for estimating soil

attributes across various environmental contexts (Mishra et al., 2020; Ye et al., 2021; Zhou et al., 2021).

By analyzing the spectral signature of a soil sample in the 350 to 2500 nm range, obtained via hyperspectral sensors, it is possible to extract significant information about its composition and formation. This information is reflected in variations in albedo and the identification of absorption features within the spectral curves.

Recent studies have applied reflectance data obtained from laboratory sensors such as the ASD FieldSpec® (350–2500 nm) to estimate organic attributes in various soils under different land uses (Gholizadeh et al., 2023; Ribeiro et al., 2023; Alsaleh et al., 2025). In addition to laboratory or proximal sensors, orbital sensors have also been employed with regression-based mathematical modeling to predict soil attributes at the surface level (Guo et al., 2020; Avdan et al., 2021; Amoli et al., 2024).

In 2022, the German Aerospace Center launched the hyperspectral satellite EnMAP (Environmental Mapping and Analysis Program) into Earth orbit, a scientific mission dedicated to high-resolution environmental monitoring. The sensor onboard EnMAP provides spectral information across 224 bands ranging from 418 nm to 2445 nm, representing a significant advancement for large-scale environmental analysis. This technology can generate valuable information from hyperspectral imagery and may be considered an alternative or complementary tool to reflectance spectroscopy performed by proximal sensors (Chabrillat et al., 2024).

Accordingly, this research was structured into three main phases: i) validation of the EnMAP sensor for soil monitoring through comparison of its spectral response with established multispectral sensors; ii) evaluation of the performance of EnMAP and the proximal ASD FieldSpec® sensor in predicting organic carbon and total nitrogen in soils under different land uses; and iii) development of regional and global models for estimating SOC and TN on soil surfaces in northeastern Brazil and California’s Central Valley.

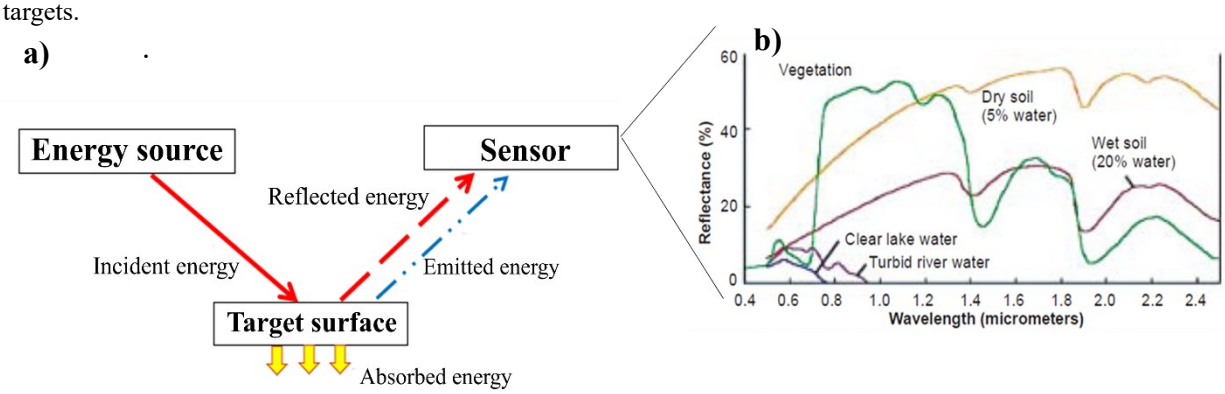
2 LITERATURE REVIEW

2.1 Remote Sensing

Remote Sensing (RS) is a technique based on acquiring data from a given surface without direct physical contact between the sensor and the target. This characteristic makes RS a valuable tool for assessing biophysical and biochemical properties of various materials through spectral information derived from their interaction with electromagnetic radiation (Jensen, 2011; Ribeiro et al., 2021; Légaré et al., 2022).

According to Ribeiro (2021), the process begins with an energy source, which can be natural, such as the Sun, or artificial, that irradiates the target. Sensors capture the radiation reflected from the target's surface and convert it into spectral data capable of distinguishing and characterizing the materials comprising the analyzed surface (Figure 1a).

Figure 1. Components of Remote Sensing. a) Schematic diagram of RS operation; b) Spectral curves of different targets.



Source: a) Adapted from Ribeiro, S.G. (2021); b) JIANG, C.Y.H. (2013)

Remote Sensing employs reflectance spectroscopy to study the relationship between the energy reflected by the target and the incident energy at specific wavelengths, thereby enabling the identification of materials based on their unique spectral signatures (Terra; Demattê; Rossel, 2018; Lafranchi; Cruz; Rocha, 2020). Figure 1b illustrates the classical approach to data acquisition, graphically representing reflectance curves across wavelengths ranging from 400 to 2500 nm, which are used to differentiate various targets

2.2 Remote Sensing for Soil Characterization

When analyzing the spectral behavior of soil samples, it is essential to consider the intrinsic components and interactions within the soil structure, since the spectral characterization is determined by the combined spectral signature of the constituent particles, considering their concentration and size (Gholizadeh et al., 2023; Ribeiro et al., 2023).

Spectral information can be acquired through imaging sensors – which produce two-dimensional images processed after collecting electromagnetic radiation over the area of interest – or non-imaging sensors such as radiometers and spectroradiometers, which generate numerical datasets representing the spectral behavior of the target in graphical form (Naue et al., 2011; Lafranchi; Cruz; Rocha, 2020). The use of remote sensing to evaluate soil samples primarily relies on components that influence spectral reflectance in the visible (400–760 nm), near-infrared (760–1200 nm), and shortwave infrared (1200–2500 nm) spectral regions (Demattê et al., 2019; Ribeiro et al., 2021).

Several studies report that soil reflectance is directly affected by moisture, surface roughness, texture, organic matter, and the presence of clay minerals and iron oxides, which exhibit specific spectral features within the 400–2500 nm range (Nowkandeh; Noroozi; Homaee, 2018; Pearlshtien; Ben-Dor, 2020; Gholizadeh et al., 2023). According to Curcio et al. (2013), variations in reflectance as a function of wavelength, combined with information from spectral libraries and multivariate statistical analysis, enable interpretation and understanding of the nature and concentration of molecules present in the analyzed sample.

Therefore, reflectance spectroscopy can be considered an effective alternative methodology, providing satisfactory results for qualitative and quantitative characterization of soil samples using data obtained from either orbital or proximal sensors (Vasava et al., 2019; Pudelko; Chodak, 2020).

2.3 Hyperspectral Remote Sensing

High spectral resolution sensors, such as hyperspectral sensors, cover numerous continuous bands within narrow wavelength intervals, typically ranging from 1 to 10 nm (Collin et al., 2019; Oliveira et al., 2020). This capability enables the detection of very subtle variations in the spectral signatures of targets.

Researchers have employed portable spectroradiometers with 3nm spectral resolutions to analyze soil spectral behavior, detecting subtle features associated with spectrally active properties such as texture, color, moisture, and organic matter content through numerical spectral data (Angelopoulou et al., 2020; Pearlshtien; Ben-Dor, 2020; Gholizadeh et al., 2023).

Using hyperspectral data has enhanced the utility of remote sensing by enabling simultaneous analysis of multiple soil attributes from a single spectral reading across the 350

to 2500 nm range, allowing for rapid, cost-effective analyses without the need for chemical reagents (Demattê et al., 2019).

Reflectance spectroscopy within the 350–2500 nm range has been successfully applied to quantify soil attributes correlated with variations in spectral features. Essential nutrients for crop development, such as total nitrogen and organic carbon, have been effectively quantified using predictive models derived from the spectral response of surface samples (Seema et al., 2020; Ribeiro et al., 2021; Gholizadeh et al., 2023; Cambou et al., 2022), representing a suitable, non-invasive quantitative technique.

Building on these spectral analysis capabilities, hyperspectral imaging by satellite sensors further advances soil characterization by capturing multiple spectral wavelengths simultaneously and nearly continuously. This capability allows hyperspectral sensors to provide enhanced visualization of subtle spectral features present in materials on the Earth's surface (Naue et al., 2011; Alsaleh et al., 2025). Moreover, as noted by Guo et al. (2020), hyperspectral images obtained from orbital or airborne sensors utilize sophisticated technologies to acquire extensive spectral information, thereby serving as valuable tools for digital soil mapping in arable lands.

Hyperspectral imaging, also referred to as imaging spectroscopy, is an innovative technique within the field of remote sensing. In this approach, each pixel of the captured scene contains continuous spectral information across the 400 to 2500 nm range — a critical portion of the electromagnetic spectrum where reflected solar radiation dominates and several of the most significant absorption features relevant to soil analysis occur (Meneses; Almeida; Baptista, 2019; Chabrillat et al., 2022).

The Hyperion hyperspectral sensor was the first one to operate in orbit. It had 198 spectral bands calibrated spanning 400 to 2400 nm – visible to shortwave infrared region. However, this sensor was discontinued in the beginning of 2017, beginning a period of studies on the potential of hyperspectral imaging spectroscopy (Silva et al., 2009; Mutanga; Van Aardt; Kumar, 2010; Zhang; Gui-Cheng, 2020).

Launched in 2019, the PRISMA satellite (*PRecursore IperSpettrale della Missione Applicativa*) has also been utilized to acquire hyperspectral imagery. It features a hyperspectral sensor with 66 bands in the VNIR range (400–1010 nm) and 171 bands covering the NIR-SWIR region, from 920 to 2505 nm (Oliveira et al., 2020; Amoli et al., 2024).

2.4 Environmental Mapping and Analysis Program (EnMAP)

In 2022, the German Aerospace Center launched the EnMAP (Environmental Mapping and Analysis Program) mission, which carries the hyperspectral Earth observation satellite with same name. This mission, primarily scientific in nature, is managed by the German Research Centre for Geosciences (*Helmholtz-Zentrum für Geoforschung* – GFZ) and aims to study environmental changes, investigate ecosystem responses to human activities, and monitor the global management of natural resources (Chabrillat et al., 2022). Consequently, EnMAP's hyperspectral measurements provide valuable data to address critical scientific questions across various Soil Science applications.

The satellite is equipped with a hyperspectral imaging scanner that records reflected solar radiation from the Earth's surface across the 420 to 2450 nm range, spanning 246 wavelengths. For surface reflectance data products (Level 2A), 224 contiguous spectral bands are provided, with an average bandwidth of 6.5 nm in the VNIR region (420–990 nm) and 10 nm in the SWIR region (990–2450 nm) (Guanter et al., 2015; Storch et al., 2023).

The hyperspectral sensor aboard EnMAP offers a swath width of 30 km and has a spatial resolution of 30 meters, meaning each pixel corresponds to a 900 m² area on the ground. Its temporal resolution is 27 days at nadir observation—when the sensor is perpendicular to the surface—and up to 4 days for off-nadir imaging at angles up to ±30º (Beamish et al., 2022; Chabrillat et al., 2022). Off-nadir imaging enables capturing areas outside the nadir field of view, thus reducing revisit time.

With a signal-to-noise ratio of ≥ 400:1 in the VNIR range and ≥ 170:1 in the SWIR spectrum, EnMAP proves to be an efficient tool for imaging spectroscopy in soil analysis, frequently providing detailed views of subtle absorption features in soil spectral characteristics (Storch et al., 2023; Chabrillat et al., 2024).

By delivering detailed spectral responses, the sensor significantly enhances the capacity of discrimination among different soil components, making analyses more refined and reliable. This positions EnMAP as an auxiliar tool for researchers and professionals seeking to perform robust, high-resolution quantitative analyses in soil studies and its attributes (Guanter et al., 2015; Vanguri; Laneve; Hościło, 2024).

However, due to the sensor's ability to image at both nadir and off-nadir angles, scenes captured at varying viewing angles exhibit pixel-level geometric distortions that affect

spectral responses, which may be underestimated (images at +30º) or overestimated (images at -30º) compared to nadir imagery (Beamish et al., 2022). This occurs because at +30º off-nadir, the sensor's viewing angle is opposite to the primary solar incidence, causing increased shading and reduced spectral response. Conversely, at -30º off-nadir, the viewing angle aligns with solar incidence, resulting in an overestimated albedo due to surface tilt. Therefore, the closer the sensor is to nadir, the lower the spectral and geometric distortions (Chabrillat et al., 2022).

Such issues primarily affect heterogeneous areas, while regions with exposed soil and uniform topography tend to be minimally influenced by sensor position (Beamish et al., 2022). According to Borel and Gerstl (1994), vegetated areas present greater challenges also due to variable spectral mixing patterns at different viewing angles.

Despite the challenges posed by spectral distortions at varying viewing angles, EnMAP's ability to deliver detailed soil surface information highlights its crucial role in advancing Soil Science research and promoting sustainable management of natural resources.

2.5 Assessment of Soil Attributes Using Hyperspectral Remote Sensing

The reflectance characteristics of soils result from the interaction of electromagnetic radiation with intrinsic components present in a soil sample, whether superficial or not (Ribeiro et al., 2021). In addition to organic and mineral composition, other factors such as moisture content, particle size, and their arrangement are fundamental in determining the spectral properties of soils (Demattê et al., 2019; Lafranchi; Cruz; Rocha, 2020).

Reflectance in the visible, near infrared, and shortwave infrared spectra depends directly on soil composition, including organic matter, clay minerals, primary minerals, oxides, hydroxides, and moisture (Jensen, 2011). These absorptions occur due to electronic transitions of atoms and molecular vibrations involving stretching and bending of atomic groups that form molecules and crystals (Rossel; Behrens, 2010; Stenberg et al., 2010).

Soil attributes considered optically active are those capable of detectably interacting with electromagnetic radiation, being responsible for the processes of light absorption and scattering. Among these components are color, organic matter, moisture, iron oxides and mineral composition (Terra; Demattê; Rossel, 2018).

2.5.1 Role of Organic Matter, Soil Texture, and Minerals in Soil Reflectance

Soil organic matter (SOM) is the primary component of the soil carbon cycle, containing approximately 58% of the element within its structure. Nitrogen, on the other hand, is present as a constituent of proteins, nucleic acids, chitin, cell walls of decomposing organisms, and animal excrements (Hu et al., 2019; Chen et al., 2020; Jiang et al., 2024).

Spectrally, SOM reduces the overall soil reflectance between 400 and 2500 nm due to its high energy-absorbing capacity, masking important attribute features in the visible spectrum and exhibiting characteristic features in the mid-infrared region (Jensen, 2011; Pearlshtien; Ben-Dor, 2020). Specific vibrations can also be detected at wavelengths in the SWIR region related to CH (1414, 1700–1800, and 2300 nm), OH (1440, 1900, and 2200 nm), and NH (1100–1200 nm, 1800 nm, and 2000–2100 nm) interactions (Stenberg et al., 2010; Alsaleh et al., 2025; Reyes-Rojas et al., 2025). Additionally, SOM influences the color of the soil matrix, acting as a pigment and darkening soils proportionally to its concentration (Pearlshtien; Ben-Dor, 2020; Gholizadeh et al., 2023).

Soil texture, although indirectly, also affects reflectance. The arrangement and size of sand, silt, and clay particles influence surface roughness and aggregate orientation, impacting reflectance (Meneses; Almeida; Baptista, 2019). Clayey soils generally exhibit lower reflectance (reduced albedo) compared to sandy soils due to higher aggregation and energy absorption. Within sandy soils, larger quartz particles cause lower reflectance than finer particles (Jensen, 2011; Terra; Demattê; Rossel, 2018; Pudelko; Chodak, 2020).

Clay minerals exhibit specific absorptions in spectral regions, mainly at 1400, 1900, and 2200 nm, with emphasis on montmorillonite, illite, and kaolinite (Demattê et al., 2019). The intimate relationship between organic matter and texture explains why clayey soils tend to retain more organic compounds, thus influencing the spectral detection of total organic carbon and total nitrogen (Stenberg et al., 2010; Laamrani et al., 2019; Ribeiro et al., 2023).

2.5.2 Influence of Iron Oxides and Interactions with Organic Matter

Iron oxides and hydroxides are optically active components in the visible and near-infrared regions of the electromagnetic spectrum, occurring in both crystalline and amorphous forms. Hematite and goethite are primarily responsible for the coloration of the soil matrix, significantly influencing the visible region (Girão et al., 2014; Nowkandeh; Noroozi; Homaee,

2018). Hematite imparts a reddish hue to the soil, with an absorption peak near 530 nm, whereas goethite exhibits absorption around 480 nm (Jensen, 2011).

The interaction between organic matter and iron oxides plays a crucial role in stabilizing organic matter in the soil, directly affecting the spectral determination of carbon and nitrogen. Several studies demonstrate positive correlations between iron concentration and total organic carbon in various soils, indicating a preferential association between organic matter and iron oxides (Hong et al., 2019; Pearlshtien; Ben-Dor, 2020; Ribeiro et al., 2021).

However, compounds containing total organic carbon and total nitrogen may not display evident signals in reflectance spectra visible directly, unlike organic matter and iron oxides. Therefore, mathematical transformations are necessary to highlight spectral variations at different wavelengths.

2.6 Organic Carbon and Total Nitrogen in Soil

In agriculture, a significant portion of soil carbon originates from plant photosynthesis, where plants contribute with organic matter at the end of their growth cycle, serving as substrate for soil microbiota activity (Hogberg; Read, 2006). The decomposition of this organic matter by soil biota plays a key role in maintaining atmospheric CO_2 balance and soil carbon storage (Amelse, 2020).

Globally, soils store up to four times more carbon than vegetation biomass and up to three times more than the atmosphere. Changes in soil organic carbon (SOC) stocks are thus crucial for agricultural management due to their significant effects on soil-biosphere interactions (Zhang et al., 2015; Hu et al., 2019; Gocke et al., 2023).

SOC content is closely linked to soil quality attributes, regulating physical, chemical, and biological properties such as increasing porosity, water retention, and limiting pesticide and herbicide movement, thereby reducing groundwater contamination (Nowkandeh; Noroozi; Homaee, 2018). Studies show that SOC sequestration depends on soil texture, aggregation, and management practices, and can be reduced by 60-75% when natural ecosystems are converted to agricultural use (Lal, 2004; Madhavan et al., 2017; Kukuļs et al., 2019).

Ribeiro et al. (2019) emphasize the need to evaluate SOC together with nitrogen content, as both serve as energy sources for microbial biomass and participate in nutrient cycling and storage. Together, organic carbon and nitrogen are key indicators of soil fertility,

structure, and quality (Reda et al., 2019; Zhang; Su; Yang, 2018), with interdependent variations (Deng et al., 2020).

Nitrogen is the fourth most abundant element in cellular biomass and makes up the majority of Earth's atmosphere as N_2, which most organisms cannot use directly (Stein; Klotz, 2016). In soil, N exists in organic and mineral forms, with ammonium (NH_4^+) and nitrate (NO_3^-) being critical for plant nutrition due to their assimilation (Fowler et al., 2013). Total N accumulation in soil is slow and depends on inputs (organic matter, geological sources, fixation) and outputs (plant uptake, leaching, volatilization, denitrification) (Marty et al., 2017).

N availability varies with the type of organic matter and carbon limitation for microbes, as soil microbiota decomposes organic matter to obtain energy, assimilating C and N for biomass growth (Li; Zhang; Wu, 2021; Shakoor et al., 2023). Microbial growth increases N demand, potentially limiting availability to plants (immobilization). When N availability meets microbial demand, excess N is mineralized and becomes plant-available (Fowler et al., 2013; Villamil et al., 2015).

The carbon:nitrogen ratio (C:N) is a critical soil quality indicator reflecting the concentration and structure of organic matter (Deng et al., 2020). Organic materials with C:N > 30 contain relatively more carbon, slowing decomposition and promoting microbial N immobilization. Lower C:N ratios favor mineralization and N availability to crops (Villamil et al., 2015) Excess soil N can lead to losses via volatilization and leaching. C:N ratios vary widely due to climate, vegetation, soil type, and depth, typically ranging from 8-10 (low) to 15-20 (high) (Zinn; Marrenjo; Silva, 2018; Deng et al., 2020; Zhou et al., 2021).

Therefore, spatial quantification of SOC and total N using reflectance spectroscopy is highly relevant for precision management in agricultural systems, especially in regions with natural limitations of these elements, such as semi-arid zones (Wang; Yang; Bai, 2015).

2.7 Methods for analyzing Organic Carbon and Nitrogen in soils

The analysis of soil organic carbon (SOC) and total nitrogen (TN) is fundamental for understanding soil quality, agricultural management, and addressing environmental challenges. Classical methodologies for determining organic carbon primarily involve wet digestion, notably the method developed by Walkley and Black (1934), and dry combustion through loss on ignition (Ben-Dor; Banin, 1989).

In wet digestion described by Walkley and Black (1934), carbon is oxidized by a potassium dichromate solution in concentrated sulfuric acid, capable of oxidizing approximately 76% of the carbon present in the sample. To optimize the process, a digestion block heater maintaining a constant temperature of 170°C for 30 minutes is currently used, ensuring almost 100% carbon oxidation (Yeomans; Bremner, 1988). After digestion, the excess dichromate is titrated to quantify organic carbon. Despite its efficiency, the use of dichromate generates toxic waste requiring proper disposal (Nowkandeh; Noroozi; Homaee, 2018).

SOC determination by loss on ignition involves controlled combustion of the sample at high temperatures (typically between 400 and 600°C) to remove organic matter, with mass loss associated with carbon content (Salehi et al., 2011; Gerenfes; Giorgis; Negasa, 2022). Although widely used historically, this method is increasingly replaced by faster and automated techniques due to its slower analytical speed.

Regarding total nitrogen analysis, the classical method developed by Kjeldahl (1883) involves digestion of the sample in concentrated sulfuric acid with catalysts, followed by distillation and titration for nitrogen quantification. Modern adaptations reduce the use of toxic catalysts such as mercury, employing digestion with mixtures and sulfuric acid instead.

Technological advances have enabled simultaneous analysis of SOC and TN using elemental analyzers. In this process, samples undergo combustion at approximately 1000°C, converting carbon to CO_2, hydrogen to H_2O, and nitrogen to NO_2, which are separated by gas chromatography and detected by thermal conductivity. This technique reduces the use of chemical reagents and analysis time (Farina et al., 1991; Gazulla et al., 2012; Even et al., 2025).

Additionally, reflectance spectroscopy has been applied to simultaneously estimate carbon and nitrogen in soils, covering spectral ranges from visible to mid-infrared (Stenberg et al., 2010; Reyes-Rojas et al., 2025). This approach estimates element content based on spectral responses of samples using multivariate statistical techniques, decreasing the need for extensive sampling and hazardous reagents (Gholizadeh et al., 2023).

Looking forward, the application of mathematical approaches in soil science promises to revolutionize prediction and monitoring of soil attributes. Algorithms trained to identify patterns relating to remote sensing data, soil samples, and environmental parameters, allowing the construction of precise predictive models for SOC and TN. These innovations are expected to increase efficiency, accuracy, and scope of analyses, contributing to more sustainable agricultural practices.

3 HYPOTHESES

This work is divided into three sections that are based on the following hypotheses:

i) The EnMAP hyperspectral satellite shows significant potential to enhance agricultural monitoring and assessment of soil organic carbon and total nitrogen, offering the opportunity to operate within a multiplatform system alongside established multispectral sensors.

ii) EnMAP is a valuable tool for estimating soil organic carbon and total nitrogen, demonstrating predictive capabilities comparable to field-based sensor ASD FieldSpec® (350-2500 nm).

iii) Predictive models using reflectance spectroscopy offer accurate estimates of soil organic carbon and total nitrogen across various land uses and geographical areas, being efficiently validated, regardless of the soil sampling method or the attribute quantification technique used.

4 OBJECTIVES

The general objectives of this research, based on the relevance and feasibility of the use of remote sensing for environmental monitoring, considering the three sections addressed in this work, were defined as follows:

I. Validate the EnMAP spectral response band by band with Sentinel-2 and PlanetScope sensors and identify wich EnMAP bands that most effectively contribute to the NDVI index.

II. Evaluate the performance of the EnMAP hyperspectral sensor in quantifying soil organic carbon and total nitrogen, comparing its predictive ability with that of the ASD FieldSpec® sensor (350-2500 nm) under field conditions.

III. Develop predictive models for SOC and TN in irrigated areas in Brazil and the United States, with different land uses and distinct quantification methodologies, while identifying the most significant wavelengths for detecting these elements in soil samples.

5 ENMAP FOR LAND USE MONITORING: A VALIDATION WITH SENTINEL-2 AND PLANETSCOPE

ABSTRACT

The use of remote sensing techniques in soil assessment is based on analyzing the components that influence spectral reflectance in the regions of the electromagnetic spectrum between 400-2500 nm. In 2022, the German hyperspectral satellite EnMAP (418–2445 nm) was launched as a tool for detailed environmental monitoring. This article reports findings from a study aimed at evaluating the spectral responses of the EnMAP orbital hyperspectral sensor in comparison to the Sentinel-2 and PlanetScope multispectral sensors, and to identify the spectral regions of EnMAP that best correspond to the NDVI calculated from the two multispectral sensors. The points of interest were within the Lower Acaraú Irrigation District in Ceará, Brazil. Images from the EnMAP, Sentinel-2 and PlanetScope sensors were collected in June and July of 2023. This study evaluated the similarity between the sensors using dispersion analysis and NDVI calculations based on statistical accuracy metrics. The results showed that EnMAP has a high degree of similarity with Sentinel-2 and PlanetScope in the visible spectrum. The NDVI calculated from EnMAP showed a high correlation with the values obtained from the multispectral sensors when combining the 653 nm (Red) and 895 nm (NIR) bands for Sentinel-2 ($R^2 = 0.83$; RMSE = 0.04; MAE = 0.01) and the 653 nm (Red) and 887 nm (NIR) bands for PlanetScope ($R^2 = 0.86$; RMSE = 0.06; MAE = 0.02). EnMAP proves to be a valuable tool for environmental monitoring, offering spectral information that exceeds the capabilities of traditional multispectral sensors, with potential for advanced applications beyond NDVI analysis.

Key words: remote sensing; hyperspectral sensor; spectral reflectance; soil

RESUMO

O uso de técnicas de sensoriamento remoto na avaliação do solo baseia-se na análise dos componentes que influenciam a reflectância espectral nas regiões do espectro eletromagnético entre 400 e 2500 nm. Em 2022, foi lançado o satélite hiperespectral alemão EnMAP (418–2445 nm) como uma ferramenta para monitoramento ambiental detalhado. Este artigo apresenta os resultados de um estudo que teve como objetivo validar as respostas espectrais do sensor hiperespectral orbital EnMAP de acordo com os sensores multiespectrais Sentinel-2 e PlanetScope, bem como identificar as regiões espectrais do EnMAP que melhor correspondem ao NDVI calculado a partir dos dois sensores multiespectrais. Os pontos de interesse situavam-se no Distrito de Irrigação do Baixo Acaraú, no Ceará, Brasil. Imagens dos sensores EnMAP, Sentinel-2 e PlanetScope foram coletadas nos meses de junho e julho de 2023. O estudo avaliou a similaridade entre os sensores por meio de análise de dispersão e cálculos do NDVI baseados em métricas estatísticas de precisão. Os resultados indicaram que o EnMAP apresenta alto grau de similaridade com o Sentinel-2 e PlanetScope no espectro visível. O NDVI calculado a partir do EnMAP mostrou forte correlação com os valores obtidos pelos sensores multiespectrais ao combinar as bandas 653 nm (RED) e 895 nm (NIR) para o Sentinel-2 ($R^2 = 0,83$; RMSE = 0,04; MAE = 0,01) e as bandas 653 nm (RED) e 887 nm (NIR) para o PlanetScope ($R^2 = 0,86$; RMSE = 0,06; MAE = 0,02). O EnMAP se mostra uma ferramenta valiosa para o monitoramento ambiental, oferecendo informações espectrais que superam as capacidades dos sensores multiespectrais tradicionais, com potencial para aplicações avançadas além da análise do NDVI.

Palavras-chave: sensoriamento remoto; sensor hiperespectral; reflectância espectral; solo

5.1 INTRODUCTION

The use of remote sensing techniques in soil assessment is based mainly on analysis of the components that influence the spectral reflectance of the target of interest in regions of the visible spectrum – VIS (400-760 nm), near infrared – NIR (760-1200 nm) and shortwave infrared – SWIR (1200-2500 nm) (Viscarra Rossel et al., 2006).

High spectral resolution sensors, such as hyperspectral ones, provide adequate coverage of multiple continuous narrow bands (from 1 to 10 nm), making it possible to detect subtle variations in the spectral signature of a target (Mutanga et al., 2010). A hyperspectral image records information across several bands that are acquired simultaneously and almost continuously in the 400 to 2500 nm range, enabling better detection of spectral features present in elements on the Earth's surface (Naue et al., 2010).

Hyperion was the first hyperspectral sensor to operate in orbit and is the most widely used in studies of hyperspectral imaging. It had 242 spectral bands between the VNIR and SWIR ranges (400 to 2400 nm), 198 of which had been properly calibrated (Silva et al., 2009; Zhang; Gui-Cheng, 2020). However, use of the sensor was discontinued in 2017, paving the way for a new generation of hyperspectral orbital sensors. The PRISMA (*PRecursore IperSpettrale della Missione Applicativa*) satellite has been used to acquire hyperspectral images since its launch in 2019. It has a 237-band hyperspectral sensor operating between 400-2505 nm (Oliveira et al., 2020).

In 2022, the German Environmental Mapping and Analysis Program (EnMAP) mission launched the hyperspectral satellite of the same name for research and monitoring of terrestrial and aquatic ecosystems. The EnMAP satellite has a hyperspectral sensor that covers wavelengths between 420-2450 nm. For Earth surface reflectance products (Level 2A), 224 contiguous spectral bands are provided at an average interval of 6.5 nm in the VNIR (420 nm - 1000 nm) and of 10 nm in the SWIR (900 nm - 2450nm), with a spatial resolution of 30 m and swath width of 30 km (Guanter et al., 2015; Beamish et al., 2022; Chabrillat et al., 2024).

As an alternative to hyperspectral sensors, there is a wide variety of multispectral imaging platforms that offer discrete spectral bands at wider intervals, of the order of hundreds of nanometres (Oliveira et al., 2020). These platforms offer remote sensing images with varied spatial, temporal and spectral resolutions. This allows soil attributes to be evaluated with different levels of precision and is a crucial tool for environmental mapping and monitoring (Demattê; Terra, 2014; Poppiel et al., 2019; Chabrillat et al., 2019).

In recent years, various studies have used multispectral imaging platforms to monitor and investigate the spectral behaviour of soil surfaces under different types of land use and cover (Avdan et al., 2021; Guo et al., 2021; Minhoni et al., 2021; Vanguri et al., 2024). Among the main platforms adopted, the following stand out: i) Landsat-8 and 9, with 11 spectral bands in the VNIR-SWIR range, with a spatial resolution of 30 m and 16-day revisit period; ii) Sentinel-2, with 10 bands in the VNIR-SWIR, a spatial resolution of 10-20 m, and a five-day revisit period; and iii) PlanetScope, which covers eight bands in the VNIR, at a spatial resolution of 3 m and a 24-hour revisit period.

The PlanetScope and Sentinel-2 platforms are increasingly being used for comparing spatial and temporal resolution, due to the possibility of better target visualisation and a higher revisit frequency, in addition to the high level of agreement, and similarity of their responses in certain spectral bands (Minhoni et al., 2021; Issaoui et al., 2022; Zagajewski et al., 2024).

The application of orbital remote sensing has been increasingly recognized as an important tool for monitoring soil attributes, especially as traditional methods require more time and involve high costs, restricting the area that can be analysed (Chabrillat et al., 2019). In this context, both multispectral and hyperspectral technology can provide valuable information at different scales and levels of detail.

Thus, the use of a multiplatform system in environmental monitoring is crucial to optimize the analysis and capture of precise environmental data, enabling the integration of different sensors with varying temporal resolutions to assess the same target more comprehensively.

To help distinguish terrestrial targets, satellite images usually undergo post-processing techniques that employ spectral indices. These indices identify pixels with reflectance values consistent with the object of interest based on spectral similarity, making it easier to classify different materials on the surface (Jensen, 2009; Pereira et al., 2016; Guo et al., 2020).

The Normalized Difference Vegetation Index (NDVI), developed by Rouse et al. (1973), represents the normalized difference between the spectral response in the red and near infrared bands. This index is considered very effective in discriminating between vegetation and soil in satellite images, since pixels of exposed soil have far lower NDVI values than those of vegetation (Demattê et al., 2018).

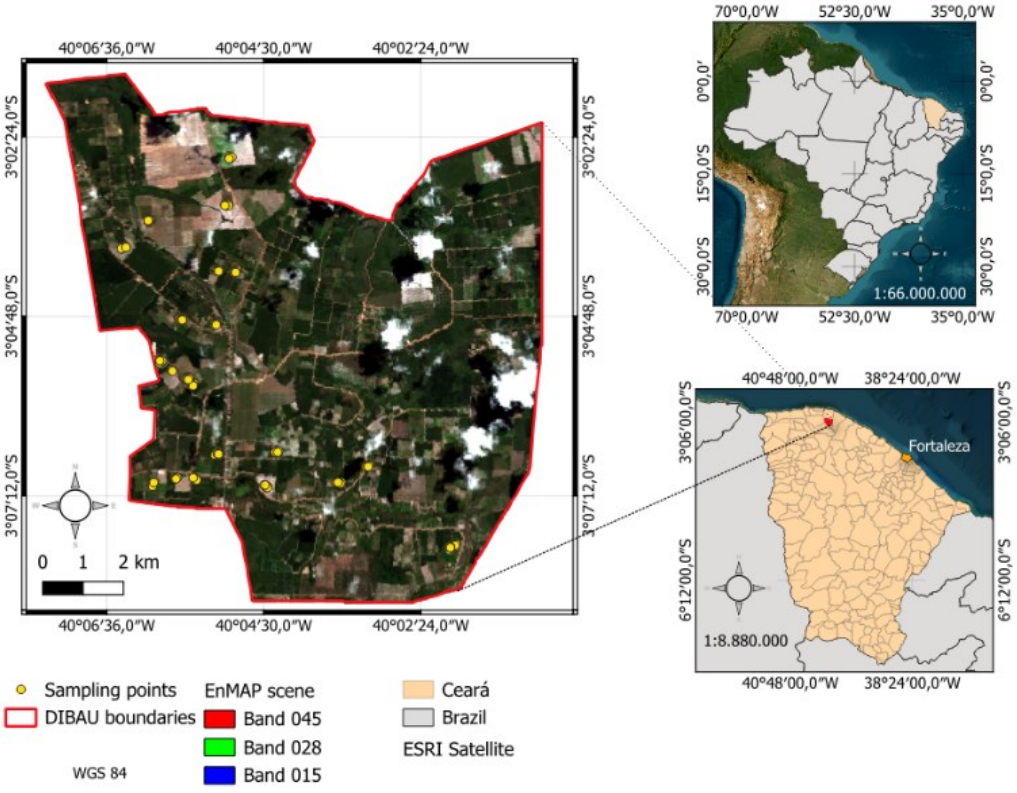
Based on the premise that the EnMAP sensor, with its 30 m spatial resolution, can produce spectral responses that are comparable to, and complementary with, the widely used multispectral sensors of higher spatial resolution, such as Sentinel-2 and PlanetScope, the aim of this study was to: *i*) validate the spectral response of EnMAP by comparing it band-by-band with the spectral responses of Sentinel-2 and PlanetScope; *ii*) identify the EnMAP bands that can most effectively contribute to the NDVI index, commonly used in agriculture; and *iii*) evaluate the potential of EnMAP as a complementary tool in operational scenarios, leveraging its high spectral resolution.

5.2 MATERIAL AND METHODS

5.2.1 Area of interest

The target points of interest for the present study were imaged in the Lower Acaraú River Irrigation District (*Distrito de Irrigação do Baixo Acaraú* - DIBAU), located near the northern coast of the State of Ceará, Brazil, at a distance by road of 220 km from the capital, Fortaleza (Figure 2).

Figure 2. Location map. Satellite image source: EnMAP data ©DLR [2023].



According to Köppen classification, the climate in the region is characterised as Aw type, tropical rainy, with an average annual rainfall of 900 mm and an average annual

temperature of 28.1°C. The predominant soils in the area are generally deep, well-drained, of medium or medium/light texture and highly permeable.

5.2.2 Acquisition of the EnMAP, Sentinel-2 and PlanetScope images

Level-2A images from the EnMAP sensor with atmospheric correction were acquired from the online platform (EnMAP.org, 2022) for the dates shown in Table 1.

Images from the Sentinel-2A MSI sensor (Level 1C) with atmospheric correction and surface reflectance were also obtained close to the EnMAP acquisition dates. Scenes from the PlanetScope satellite, with eight spectral bands and surface reflectance, were similarly acquired, filtered for cloudiness of up to 20%. Due to the high cloud cover in the study area, only two clear scenes were captured by each satellite (Table 1).

Table 1. Imaging period for the EnMAP, Sentinel-2 and PlanetScope scenes used in the study during 2023.

EnMAP	Sentinel-2	PlanetScope
19 June	25 June	21 June
24 July	30 July	28 July

A total of 130 points of interest were selected from pixels of exposed soil or low vegetation, based on the EnMAP RGB false-color NIR image using the 75-47-30 bands. Field verification was conducted on dates close to the EnMAP imaging to confirm land use practices in the selected pixels.

The spectral curves of the EnMAP images were collected from the pixels of the points of interest, the numerical data being converted to surface reflectance values using the image metadata file. The data were extracted using the QGIS 3.28 software with the EnMAP-Box 3.13 plugin (Jakimow et al., 2023).

5.2.3 Image processing

To evaluate and compare the reflectance captured by satellites sensors, the images were submitted to spatial and spectral resampling. Given that the EnMAP images have a spatial resolution of 30 m, the Sentinel-2 (10 to 20 m) and PlanetScope (3 m) images were spatially adjusted to 30 m. In addition, the spectral information from the EnMAP hyperspectral sensor was resampled to fit the Sentinel-2 (11 bands) and PlanetScope (8 bands). These processes were carried out using the QGis 3.28 software, with the aid of the EnMAP-Box add-on version 3.13 (Jakimow et al., 2023).

Spectral responses were compared, band by band, between EnMAP and Sentinel-2, and between EnMAP and PlanetScope. This comparison was performed as a spectral validation of the hyperspectral sensor against already consolidated multispectral sensors to assess similarities and differences in their responses.

A similarity assessment of the sensor responses was also carried out based on the NDVI values for the points of interest in each of the images. The index was calculated as per Equation 1 for both the resampled and original reflectance data. For the post-resampled data, the NDVI was calculated using the Red and NIR bands of each of the multispectral sensors.

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (1)$$

Where, ρ_{NIR} is the target reflectance value in the near-infrared region (Sentinel-2A = 832.8 nm; Sentinel-2B = 832.9 nm; PlanetScope = 865 nm) and ρ_{Red} is the target reflectance value in the red region (Sentinel-2A = 664.6 nm; Sentinel-2B = 664.9 nm; PlanetScope = 665 nm).

In a second comparative approach, in addition to using the resampled spectrum in the NDVI calculation, all the original EnMAP wavelengths in the Red and NIR regions were tested against the original NDVI from Sentinel-2 and PlanetScope, thereby applying successive normalised difference equations to approximate the indices obtained by EnMAP to those calculated by the multispectral sensors.

The EnMAP bands used in the tests were selected based on their central wavelength of each sensor, as recommended by Tao et al. (2021) and Zagajewski et al. (2024).

In the third test of NDVI correspondence, EnMAP spectral bands 38 and 64 were used with the EnMAP-Box add-on, as recommended by Jakimow et al. (2023). These bands correspond, respectively, to the 605 nm (Red) and 778 nm (NIR) regions of the images analysed in this study.

For the fourth and final test of NDVI similarity, EnMAP wavelengths close to the central values of the Red and NIR bands of the Sentinel-2 (Table 2) and PlanetScope sensors (Table 3).

Table 2. EnMAP spectral bands based on the regions of the Sentinel-2 sensors. The highlighted EnMAP wavelengths correspond to the values closest to the central wavelength of each sensor.

	Wavelengths (nm)		
Region	**Sentinel-2A (centre)**	**Sentinel-2B (centre)**	**EnMAP**
Aerosol	432.7- 452.7 (442.7)	432.2 - 452.2 (442.2)	434, 439, **444**, 449
Blue	460.7 - 524.7 (492.7)	459.8 - 524.8 (492.3)	463, 468, 473, 477, 482, 487, **491**, 496, 501, 506, 510, 515, 520, 525
Green	542.3 - 577.3 (559.8)	541.4 - 576.4 (558.9)	540, 545, 550, 555, **561**, 566, 571, 577
Red	649.6 - 679.6 (664.6)	649.4 - 680.4 (664.9)	653, 660, **666**, 673, 679
Red Edge 1	697.1 - 711.1 (704.1)	696.3 - 711.3 (703.8)	699, **706**, 713
Red Edge 2	733.5 - 747.5 (740.5)	732.1 - 746.1 (739.1)	734, **741**
Red Edge 3	772.8 - 792.8 (779.7)	769.7 - 789.7 (782.8)	**778**, 786
NIR	773.8 - 891.8 (832.8)	775.4 - 890.4 (832.9)	786, 793, 801, 808, 816, 824, 832, **839**, 847, 855, 863, 871, 879, 887, 895
Red Edge 4	854.7 - 874.7 (864.7)	854.0 - 874.0 (864)	855, **863**, 871, 879
SWIR 1	1569.7 - 1657.7 (1613.7)	1563.9 - 1656.9 (1610.4)	1564, 1575, 1586, 1597, **1609**, 1620, 1631, 1642, 1653
SWIR 2	2112.9 - 2291.9 (2202.4)	2095.2 - 2276.2 (2185.7)	2104, 2113, 2122, 2130, 2139, 2148, 2156, 2165, 2174, 2182, **2191**, 2199, 2207, 2216, 2224, 2232, 2241, 2249, 2257, 2265, 2273, 2282, 2290

Adapted from European Space Agency (2025)

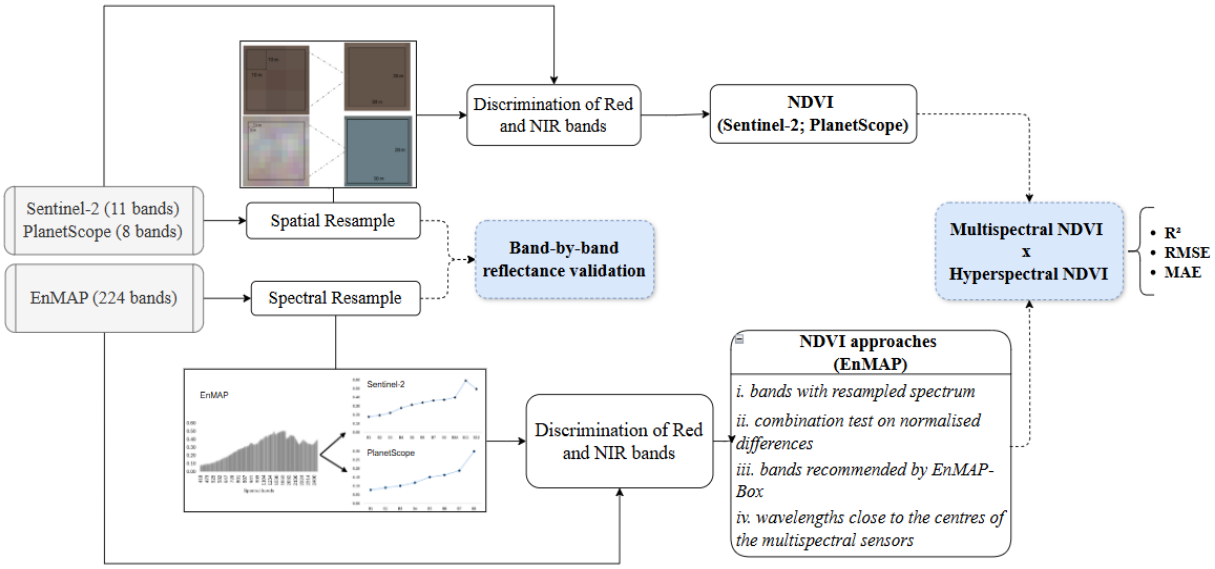
Table 3. EnMAP spectral bands based on the regions of the PlanetScope. The highlighted EnMAP wavelengths correspond to the values closest to the central wavelength of the sensor.

	Spectral bands (nm)	
Region	**PlanetScope (centre)**	**EnMAP**
Coastal Blue	431 – 452 (443)	434, 439, **444**, 449
Blue	465 – 515 (490)	468, 473, 477, 482, 487, **491**, 496, 501, 506, 510, 515
Green I	513 - 549 (531)	515, 520, 525, **530**, 535, 540, 545, 550
Green	547 – 583 (565)	550, 555, 561, **566**, 571, 577, 582
Yellow	600 - 620 (610)	599, 605, **611**, 616, 622
Red	650 – 680 (665)	653, 660, **666**, 673, 679
Red Edge	697 – 713 (705)	699, **706**, 713
NIR	845 – 885 (865)	839, 847, 855, **863**, 871, 879, 887

Adapted from Planet Labs Inc. (2025)

Figure 3 outlines the workflow carried out in this research, highlighting the comparative band by band validation between the sensors and the approaches used to analyse the best method for calculating the NDVI using EnMAP.

Figure 3. Workflow for validation analyses of band-by-band reflectance and NDVI values between EnMAP and the multispectral sensors.



The similarity of the NDVI calculated by EnMAP to the NDVI acquired by the multispectral sensors was evaluated using the Coefficient of Determination (Equation 2), Mean Absolute Error (Equation 3) and Root Mean Square Error (Equation 4).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^{N}\left|\hat{Y}_I - Yi\right|}{N} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2}{N}} \quad (4)$$

Where, Y_i and $\bar{Y}$ are, respectively, the NDVI measured by the multispectral sensors for the i-th observation and the observed mean value; $\hat{Y}_i$ is the NDVI calculated from the EnMAP data for the i-th observation; N is the number of observations.

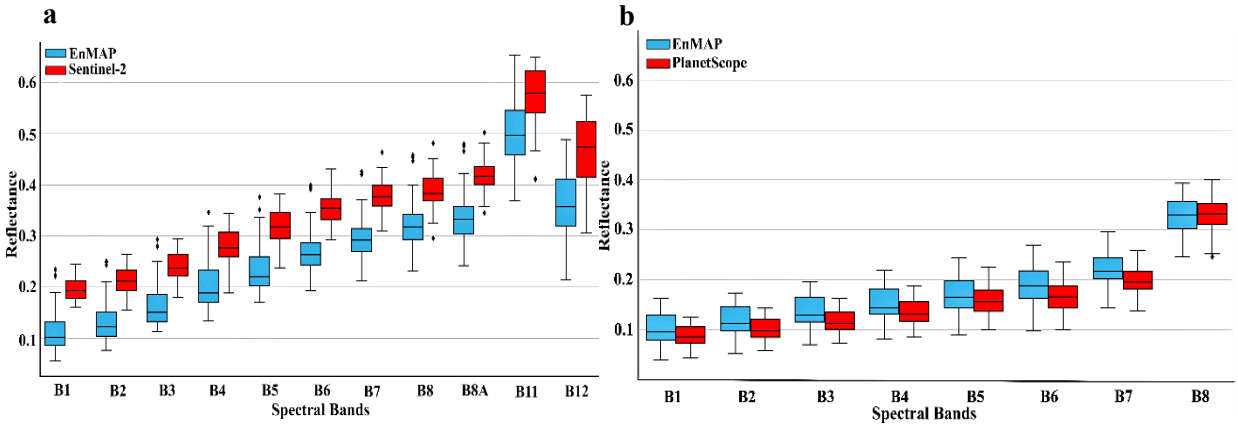
5.3 RESULTS

5.3.1 Spectral resampling between EnMAP and the Sentinel-2 and PlanetScope sensors

After resampling, the sensors showed similar responses. When resampled to the spectral resolution of Sentinel-2, EnMAP presented reflectance values lower than the 11-band multispectral sensor (Figure 4a). On the other hand, when resampled to the 8-band resolution of PlanetScope (Figure 4b), EnMAP showed higher similarity with the spectral behavior of the

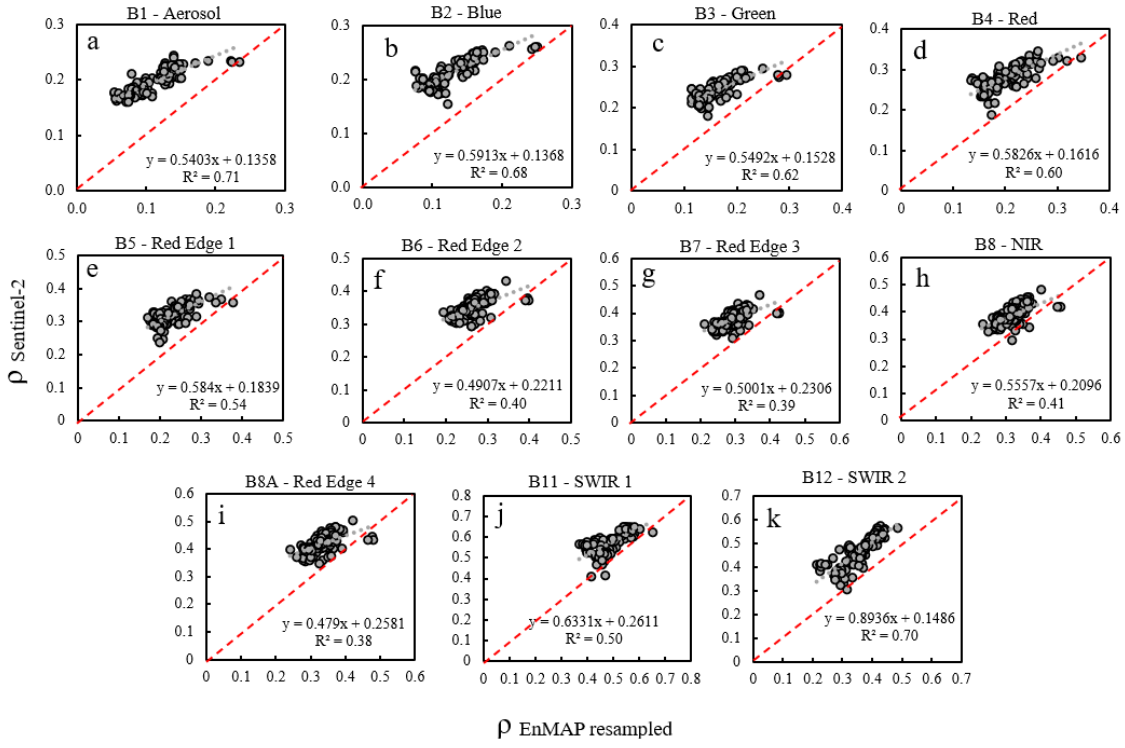
ground surface, recording values higher than the reflectance obtained with the multispectral sensor.

Figure 4. Spectral behavior of the soil captured by EnMAP resampled for the (a) Sentinel-2 and (b) PlanetScope sensor bands.



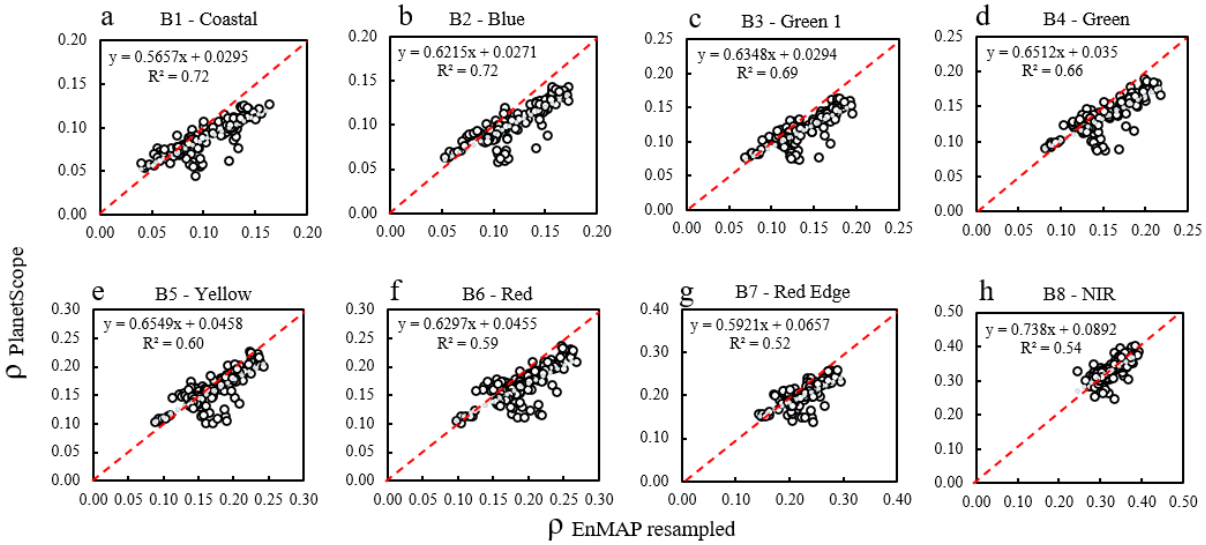
The validation analysis of the spectral bands revealed a strong similarity between EnMAP and Sentinel-2 in B1 (aerosols, Figure 5a), B2 to B4 (visible, Figures 5b to 5d), and B12 (SWIR, Figure 5k). The coefficient of determination (R^2) varied from 0.60 to 0.71, indicating some dispersion of the data for these bands.

Figure 5. Band-by-band reflectance validation after resampling EnMAP to Sentinel-2 spectral resolution.



In the validation analysis between the resampled EnMAP and PlanetScope, the results are similar to those observed when comparing EnMAP with Sentinel-2 (Figures 6a to 6f).

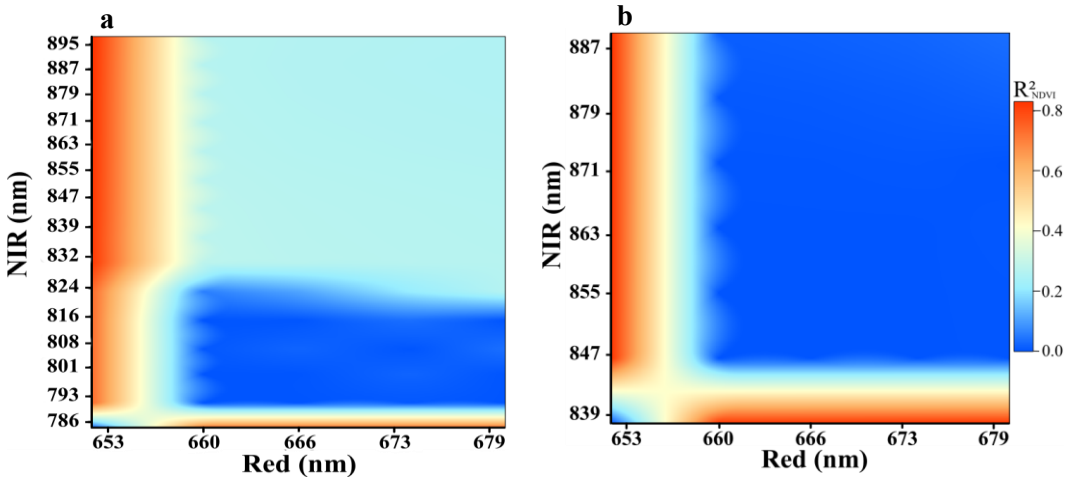
Figure 6. Band-by-band reflectance validation after resampling EnMAP to PlanetScope spectral resolution.



5.3.2 Assessment of the Normalized Difference Vegetation Index (NDVI)

The results of NDVI calculation using the EnMAP bands corresponding to the Red and NIR regions of Sentinel-2 and PlanetScope are shown in Figures 7a and 7b, respectively. Figure 7a shows that combining the 653 nm band with EnMAP NIR wavelengths between 793-895 nm produced indices that were highly compatible with the NDVI calculated by Sentinel-2 ($R^2 > 0.60$).

Figure 7. R^2 values obtained from the normalized difference tests using the Red and NIR bands of EnMAP to approximate the NDVI calculated by (a) Sentinel-2 and (b) PlanetScope.

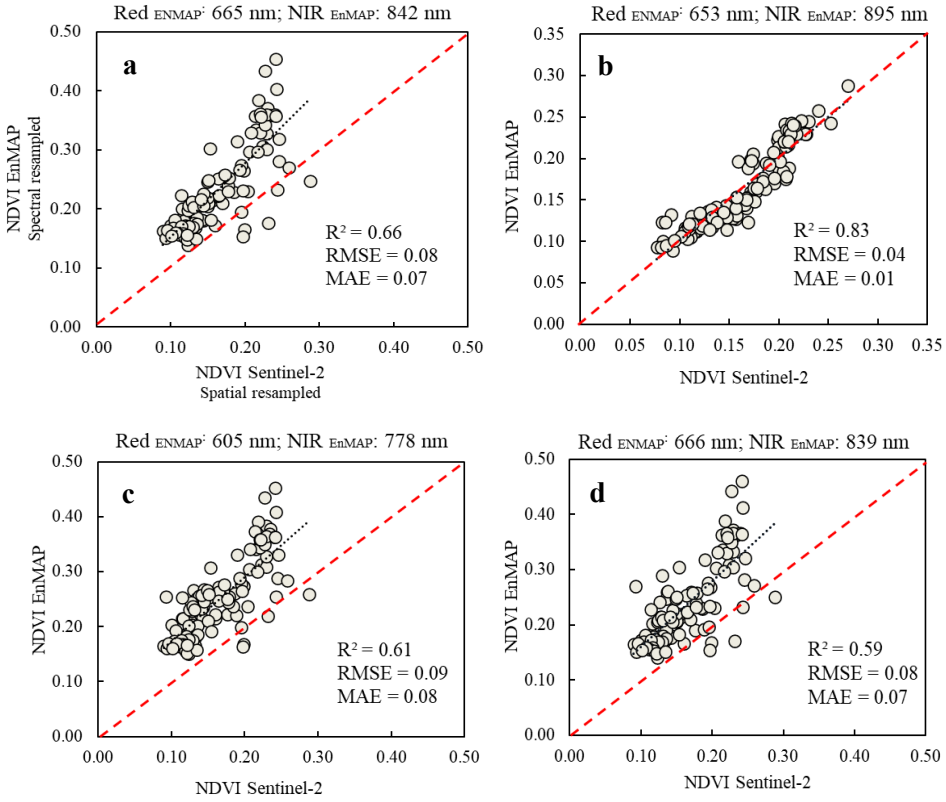


In Figure 7b, combining the 653 nm band with wavelengths between 847-887 nm, and 660–679 nm with 839 nm, showed high similarity with the NDVI from PlanetScope (R^2 > 0.60).

The tests showed that the EnMAP bands that best match the NDVI from Sentinel-2 use reflectance at 653 nm and 895 nm. This combination results in R^2 = 0.83, RMSE = 0.04 and MAE = 0.01 (Figure 8b). The second-best match was achieved with a resampled dataset using the same bands on both sensors (665 nm and 842 nm) (Figure 8a), where the NDVI calculated with EnMAP showed R^2 = 0.66, RMSE = 0.08 and MAE = 0.07 in relation to Sentinel-2.

The NDVI calculations using bands recommended by EnMAP-Box (Figure 8c) and bands close to those represented by the Red and NIR from Sentinel-2 (Figure 8d) also showed high similarity with the original NDVI from Sentinel-2 despite these approaches showing slightly lower, albeit still reliable, accuracy metrics.

Figure 8. Validation analysis of NDVI values from Sentinel-2 and EnMAP, where the NDVI for EnMAP was calculated using: (a) resampled spectrum; (b) optimal bands identified through normalized difference analysis; (c) bands recommended by EnMAP-Box; (d) hyperspectral bands closest to those of Sentinel-2.

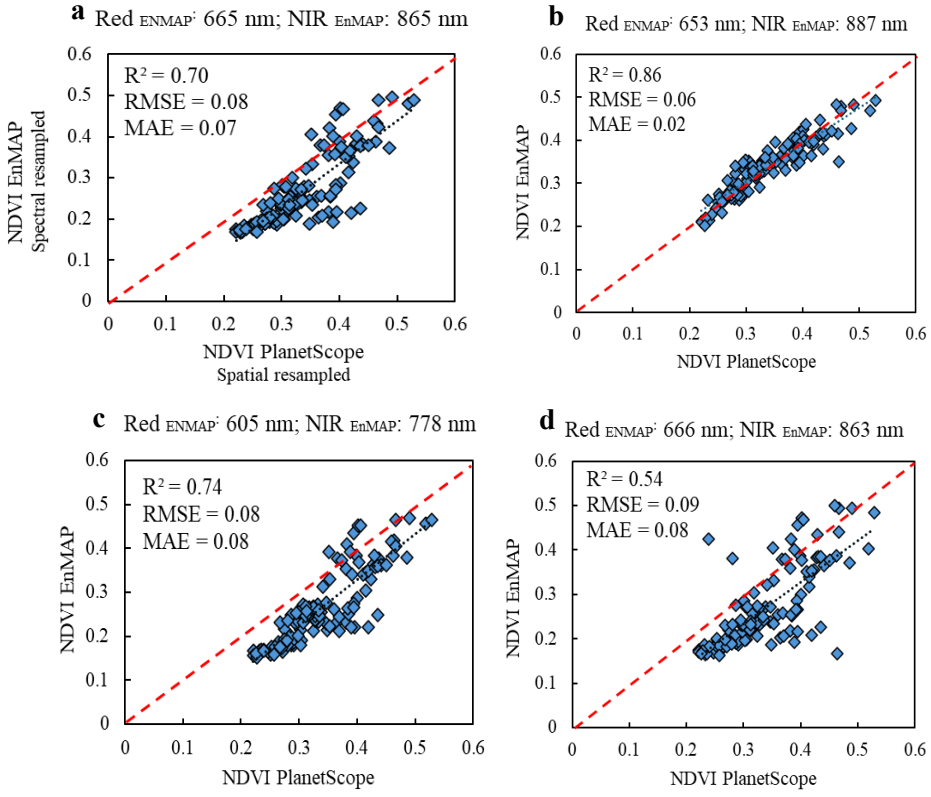


In the comparison with the NDVI from PlanetScope, the combination of EnMAP bands, identified by means of normalized difference tests at 653 nm and 887 nm, presented the best approximation, with an $R^2 = 0.86$, and RMSE = 0.06 and MAE = 0.02 (Figure 9b).

The second-best result was obtained using the bands recommended by EnMAP-Box, at 605 nm and 778 nm. This approach gave $R^2 = 0.74$, RMSE = 0.08 and MAE = 0.08, as shown in Figure 9c. The use of resampled EnMAP data, employing the PlanetScope bands, also showed good NDVI approximation, with $R^2 = 0.70$, RMSE = 0.08 and MAE = 0.07 (Figure 9a).

In contrast, the least efficient approximation method for calculating the NDVI used bands at 666 nm and 863 nm, close to the spectral centers of the PlanetScope Red and NIR bands, resulting in $R^2 = 0.54$, RMSE = 0.09 and MAE = 0.08 (Figure 9d).

Figure 9. Validation analysis of NDVI values from PlanetScope and EnMAP, where the NDVI for EnMAP was calculated using: (a) resampled spectrum; (b) optimal bands identified through normalized difference analysis; (c) bands recommended by EnMAP-Box; (d) hyperspectral bands closest to those of PlanetScope.



These results highlight the importance of selecting specific bands for calculating spectral indices when using hyperspectral sensors such as EnMAP.

5.4 DISCUSSION

5.4.1 Spectral analysis between EnMAP and Sentinel-2

When EnMAP is resampled to the spectral resolution of Sentinel-2, the difference in reflectance values can reach up to 10% (Figure 4a), with EnMAP showing the lowest albedo. This suggests that resampling can harmonize reflectance values in specific bands, in line with what reported by Poppiel et al. (2019), who noted that two sensors can have similar responses following convolution of their spectral data.

The similarity between the resampled images from different sensors is evident in terms of spectral behavior for the points on exposed soil, with a steady increase in albedo and a reflectance peak at B11 (1610 nm), corresponding to the SWIR 1 region. Similar behavior was obtained by Deodoro et al. (2021) and Guo et al. (2021), who worked with the MSI sensor of the Sentinel-2 satellite to detect pixels of exposed soil.

The results for scattering shown in Figures 5a to 5d show a strong correlation between the reflectance obtained with Sentinel-2 and with EnMAP (R^2 between 0.60 and 0.71). This is because the spectral bands in the visible region tend to be more stable to atmospheric interference and abrupt variations in reflectance compared to the longer-wavelength bands (Kaufman; Sendra, 1988).

Lo and Yeung (2002) and Vermote et al. (2016) explained that radiation scattering dominates at shorter wavelengths, while energy absorption by water vapor is more significant at wavelengths from the near infrared onwards, making the region after the visible spectrum more sensitive to variations in reflectance and noise.

The results of the present research agree with the abovementioned studies, since the R^2 in band B1 (443 nm) is 0.71, possibly due to the greater stability of both sensors for atmospheric variations. However, this value reduces significantly in the NIR region ($R^2 = 0.41$), reaching 0.38 at 865 nm, the Red Edge region in the Sentinel-2 MSI sensor. This suggests that compatibility between EnMAP and Sentinel-2 in this region was favored by the stability of the visible bands.

Various authors (Viscarra Rossel et al., 2006; Wang et al., 2008; Lange et al., 2017; Ren et al., 2022) pointed out that the NIR region of the electromagnetic spectrum is strongly influenced by soil moisture and the vegetation structure. These variables can appear different

to each sensor, which may explain the lack of agreement between the EnMAP and Sentinel-2 data in this region of the spectrum.

The lower correlation between the sensors in the 705 nm to 865 nm range (Red Edge and NIR) can therefore be attributed to the greater sensitivity of this region to the vegetation structures and to subtle differences in the level of soil moisture (Ren et al., 2022), which can affect the effectiveness of capturing spectral characteristics following spectral and spatial resampling.

In turn, the spectral bands in the SWIR range (1610 nm and 2190 nm) show R^2 values of 0.50 and 0.70, respectively (Figures 5j and 5k). Since these wavelengths are considered essential for detecting soil moisture and the mineral characteristics of the soil (Chabrillat et al., 2019; Poppiel et al., 2019), the higher correlation between the EnMAP sensor and the Sentinel 2 MSI sensor for this region suggests that, even with resampling, there is reasonable agreement between the properties of the exposed soil detected by these sensors.

The similarity of parameters derived from EnMAP and Sentinel-2 in the RGB and SWIR 2 bands shows that, even with a six-day difference in the acquisition of the imaged scenes, the hyperspectral sensor can be used together with the multispectral sensor to obtain robust spectral indices.

5.4.2 Spectral analysis between EnMAP and PlanetScope

Figure 4b shows that EnMAP presents similar spectral behavior to PlanetScope after resampling, with subtle differences across the spectrum. Compatibility between the sensors is shown by the maximum difference of 3% in reflectance values in the Red and Red Edge bands (B6 and B7). These results highlight the potential of EnMAP as a complementary tool to PlanetScope: despite the slight differences in reflectance, the overall spectral compatibility suggests they can be used together for soil analysis.

Since PlanetScope can provide detailed images with high temporal frequency (daily) and EnMAP offers detailed information on spectral characteristics, the results show that the ability to combine different reflectance capture capabilities enables a robust and continuous analysis of the same target, maximizing both spatial and temporal coverage and providing a more comprehensive view of environmental processes over time.

Purnamasari et al. (2021) showed that hyperspectral sensors are better at capturing information from the red edge region, which can result in variations when adapting this information to a lower spectral resolution such as that of PlanetScope.

As with the convolution for Sentinel-2, the spectral resampling of EnMAP for the PlanetScope bands showed a high degree of agreement in the visible region. The high correlation coefficients, 0.72 in B1 (coastal, 443 nm) to 0.66 in B4 (green, 565 nm), indicate that the resampled EnMAP data are suitable for comparing with PlanetScope (Figures 6a to 6d).

The B5 (yellow, 610nm) and B6 (red, 665nm) bands showed $R^2 = 0.60$ and 0.59 respectively (Figures 6e and 6f), suggesting moderate agreement between the EnMAP and PlanetScope sensors and high consistency between the spectral responses. However, these bands present a slightly weaker correlation when compared to the blue and green bands, which can be attributed to a slight loss in capture stability (Kaufman; Sendra, 1988).

On the other hand, the longer-wavelength bands, B7 (Red Edge, 705 nm) and B8 (NIR, 865 nm), showed poor agreement between sensors, with $R^2 = 0.52$ and 0.54, respectively (Figures 6g and 6h). Despite the lower correlation with the other bands, these spectral regions showed higher compatibility between EnMAP and PlanetScope than between EnMAP and Sentinel-2 (Figures 5g and 5h).

This is possibly related to the differences in data acquisition geometry, spatial resolution and imaging time between the multispectral sensors and EnMAP. According to Huang et al. (2020), these variables are important factors in the variability found in comparative analyses between reflectance values from different sensors.

According to the authors, acquisition geometry refers to the viewing angle of the sensors when capturing the light reflected by the Earth's surface, which can affect the amount of reflected light that is recorded by the sensor. According to Beamish et al., 2022, this mainly affects heterogeneous areas that have a significant mix of vegetation and exposed soil, where variations in the capture angle can have an impact on reflectance.

The low level of correspondence between EnMAP and PlanetScope in B7 (Figure 6g) and B8 (Figure 6h) is due to the sensitivity of the Red Edge band to the vegetation structure, while the NIR band responds to variations in soil moisture (Wang et al., 2008; Ren et al., 2022; Xu et al., 2022). Thus, the delay in imaging time between EnMAP and PlanetScope proved to be a crucial factor for the differences found during the resampling process.

Nevertheless, the spectral resampling of EnMAP to align with Sentinel-2 and PlanetScope validated the reflectance values of the EnMAP sensor in the regions covered by these platforms even when the images were acquired three to six days apart.

This validation demonstrates the reliability of EnMAP, highlighting its potential as a promising tool for land use monitoring. By offering high spectral resolution, EnMAP can provide valuable insights for more accurate and comprehensive environmental assessments in short-term surveys.

5.4.3 Comparison of NDVI values

In general, the NDVI calculated with Sentinel-2 data presented values between 0.10 and 0.30 (Figures 8a through 8d), in line with results from studies conducted by Bezerra et al. (2018) and Poppiel et al. (2019), which showed NDVI values ≤ 0.30 typically discriminating field conditions between exposed soil and sparse vegetation.

On the other hand, the NDVI calculated with PlanetScope data resulted in higher values ($0.20 < NDVI \leq 0.55$), as shown in Figures 9a to 9d. In this case, the NDVI is possibly associated with the influence of soil moisture at the time of image acquisition, since, according to West et al. (2018), the dynamics of NDVI are affected by soil moisture, especially in areas where vegetation is scarce.

During visits to the areas of interest, it was observed that farmers typically activated irrigation in the early morning hours, and, in some locations, the soil remained moist during the satellite's passage, which may have further influenced the NDVI values.

The use of NDVI with Sentinel-2 for monitoring bare soils was also employed by Minhoni et al. (2021), whose results were promising for the detection of surface soil organic carbon (SOC), with NDVI values ranging from 0.048 to 0.45.

The scatter plots in Figures 8a to 8d and Figures 9a to 9d show close agreement between the NDVI values calculated using data from EnMAP and those obtained using data from the Sentinel-2 and PlanetScope sensors. However, each approach presents subtle differences in the coefficients of determination and errors, highlighting the impact of band selection on calculating the index.

In studies using simulated EnMAP data, Dotzler et al. (2015) found strong similarities between the hyperspectral sensor and Sentinel-2 when producing spectral indices in forests, even at different spatial resolutions. The same behavior was obtained in this study

when comparing the NDVI calculated with actual EnMAP data after resampling for Sentinel-2, with $R^2 = 0.66$, RMSE = 0.08 and MAE = 0.07 (Figure 8a).

The effectiveness of EnMAP in replicating the most widely used vegetation index for environmental monitoring is highlighted when calculating the NDVI at wavelengths of 653 and 895 nm – a combination obtained with the normalized difference test (Figure 7a). As shown in Figure 8b, EnMAP-based NDVI is highly compatible with that calculated by Sentinel-2, with $R^2 = 0.83$ and minimal errors (RMSE = 0.04 and MAE = 0.01). These results corroborate the findings of Bezerra et al. (2018), showing that using the red band affects the accuracy of the NDVI. Bands positioned closer to the vegetation spectrum show more robust correlations.

Vanguri et al. (2024) also showed promising results integrating scenes from EnMAP and Sentinel-2 in studies using vegetation indices to map forest species and machine learning to classify them, underlining the potential of combining data from these sensors for environmental studies.

Lange et al. (2017) indicated that, although multispectral data are sufficient for environmental analyses using NDVI, hyperspectral sensors offer additional information that is crucial to validating the results. This highlights the importance of spectral resampling to ensure similarity between platforms and expand the applicability for environmental indices.

Data acquisition geometry, spatial resolution, and imaging time are important factors when comparing and combining orbital sensors, as they can produce conflicting results in calculating NDVI values, even when using spectral bands that are close to each other (Bezerra et al., 2018; Huang et al., 2020). In the present study, variations in the value of the NDVI can be partly explained by the six-day difference between the Sentinel-2 and EnMAP images and the position of the orbital sensors when acquiring the images.

According to Beamish et al. (2022), scenes produced at different angles suffer geometric distortion at the pixel level, affecting the spectral response. This should be considered when using EnMAP, since the sensor can acquire scenes with off-nadir angles that vary between ±30º.

Bezerra et al. (2018) and West et al. (2018) reported that there is a positive correlation between soil moisture and the NDVI from Sentinel-2. Overestimation of the index by EnMAP, seen in approaches i, iii and iv (Figures 8a, 8c and 8d, respectively) is possibly due to the Sentinel-2 images capturing variations in the NDVI in 10 m x 10 m areas (corresponding

to the spatial resolution of the RGB-NIR bands) that contained dry or exposed soil when the image was acquired.

On the other hand, the EnMAP images captured the region under more humid conditions, when there were fewer variations in each 30m x 30m pixel area. The coarser spatial resolution may result in the spectral blending of individual pixels, thereby diluting the signal from sparse vegetation or exposed/dry soil in some locations (West et al., 2018; Beamish et al., 2022).

Unlike the results from Sentinel-2, the NDVI values obtained from data acquired by EnMAP were underestimated compared to those from PlanetScope (Figures 9a to 9d). Since the NDVI values varied between 0.20 and 0.55, it can be assumed that the level of moisture in the study area was high when captured by the multispectral sensor, with results that were consistent with those of the EnMAP sensor, albeit differing from those from Sentinel-2.

Despite the spectral differences noticed between images acquired by sensors in the individual bands (Figures 6a to 6h), the overall correspondence of the NDVI was satisfactory when calculated using the resampled PlanetScope and EnMAP data ($R^2 = 0.70$), which shows a high level of similarity between the sensors, as observed in Figure 9a.

Such correspondence suggests that the data derived from different sensors may be interchangeable for some applications. Xu et al. (2022) showed that mulch alters the spectral response in the red-NIR region, reducing the accuracy of the signal as the amount of mulch on the soil surface increases. For this reason, spectral resampling is an essential technique for ensuring accurate NDVI determinations that are comparable between imaging platforms over time in any area.

Despite individual bands do not show perfect correlation, Légaré et al. (2022) found that the NDVI can still capture important spectral dynamics, reinforcing the idea that the relationship between the Red and NIR bands can remain consistent, even if the individual correlations are weaker.

As with the comparison with Sentinel-2, the best match was obtained when testing original normalized difference EnMAP bands in the Red and NIR region. Figure 9b shows that when the 653 nm (red) and 887 nm (NIR) EnMAP bands are used to calculate the NDVI of exposed soil, the index values are very similar to those obtained using the red and NIR bands of PlanetScope, with $R^2 = 0.86$, RMSE = 0.06 and MAE = 0.02.

Since the bands near 900 nm are broader in the NIR region, it can capture a wider range of information about the target than can narrower bands, resulting in better discrimination of the spectral characteristics (Gomes et al., 2021). The results of this study underline how the choice of spectral bands can significantly impact the NDVI results, as also reported by Assmann et al. (2019). Therefore, even if the bands are close, the spectral response may be different, thus affecting the results.

Using the bands recommended by the EnMAP-Box documentation also proved to be a reliable and complementary approach to the response of PlanetScope, with $R^2 = 0.74$, and RMSE and MAE = 0.08, which underlines the compatibility of the hyperspectral sensor with multispectral sensors, even with the differences in spatial resolution.

The high spatial resolution of PlanetScope is essential for capturing spatial heterogeneity in the environment and is fundamental for consistent responses that are sensitive to spatial changes (Zagajewski et al., 2024). However, when calculating the NDVI with the EnMAP spectral bands closest to those of PlanetScope, the correspondence proved to be unsatisfactory ($R^2 = 0.54$: RMSE = 0.09; MAE = 0.08). This may be related to the differences in spatial resolution and viewing angle between the sensors, affecting data acquisition and reducing the efficiency of the estimates.

Since the spatial resolution of EnMAP (30 m) is lower than that of PlanetScope (3 m), direct comparison of spectral responses and indices may be limited, especially when using bands close to those used by the multispectral sensor to calculate the NDVI, as the response of the pixels is different.

Therefore, any comparison between the data from different sensors should consider not only spectral proximity, but also the specific characteristics of each sensor and its respective calibrations, and is essential for ensuring more accurate analyses, as discussed by West et al. (2018), Huang et al. (2020) and Zagajewski et al. (2024).

In general, the results from the research presented in this article agree with Bezerra et al. (2018), who found that despite spectral and spatial differences the NDVI can show similar values and consistent statistical metrics when comparing different sensors. However, selecting the appropriate bands for the calculation is a decisive factor in the accuracy of the measurements.

The results obtained here represent an initial step for future studies with EnMAP. Its excellent spectral response on surfaces of different soils, even with a spatial resolution of 30 m, is crucial for advancing more in-depth research, such as statistical modeling for the prediction of soil attributes. With its ability to provide detailed and high spectral resolution data, EnMAP opens new possibilities for environmental monitoring and the understanding of natural processes on larger scales and with higher precision.

5.5 CONCLUSIONS

This research demonstrates that EnMAP is a valuable data source for environmental monitoring, effectively validating its spectral responses against established multispectral sensors with higher spatial resolution. Consequently, EnMAP serves as an efficient tool within a multiplatform system combining Sentinel-2 and PlanetScope, enhancing environmental monitoring by providing both high spatial and spectral resolution.

EnMAP provides high-resolution spectral information that surpasses the capabilities of traditional multispectral sensors, making it especially useful in spectral index studies such as NDVI. This is particularly true for distinguishing soil from vegetation, particularly in the 653 nm and 895 nm bands when used in conjunction with Sentinel-2, and in the 653 nm and 887 nm bands when combined with PlanetScope.

Beyond the comparisons provided in this research, the potential of EnMAP lies in its ability to provide deeper insights for more advanced applications in environmental research and monitoring.

6 PREDICTION OF SOIL ORGANIC CARBON AND TOTAL NITROGEN ON BARE SOIL GROUND SURFACES USING ENMAP IMAGERY AND IN SITU REFLECTANCE SPECTROSCOPY

ABSTRACT

Analyzing the dynamics of soil organic carbon (SOC) and total nitrogen (TN) is essential to understanding agronomic and environmental issues. Hyperspectral remote sensing proved to be a promising tool for the rapid and non-destructive analysis of soil attributes, and the EnMAP satellite sensor is seen as a significant advancement for large-scale environmental monitoring studies using remote sensing. This research aimed to evaluate the performance of EnMAP in predicting SOC and TN in bare soil ground surface, as compared with observations obtained in situ by spectroscopy. Ninety-five samples of topsoil (0–1 cm) were collected in the Lower Acaraú Irrigated Perimeter, northeastern Brazil. The investigated areas were selected using EnMAP images, and in situ spectral readings were taken with ASD FieldSpec® spectroradiometer (350-2500 nm). The spectral data were processed and the most relevant bands for SOC and TN were determined using a stepwise selection, then followed by fitting multiple linear regression models using cross-validation. The models showed good predictive performance for both sensors. EnMAP stood out over the SOC range between 11.6–23.6 gkg^{-1} (Adj.R2 = 0.91; RPD = 6.03), whereas FieldSpec showed better performance to the SOC range between 9.2-11.6 g kg^{-1} and higher than16 g kg^{-1} (Adj.R2 = 0.83; RPD = 2.66). Both sensing techniques performed well in predicting SOC using SWIR wavelengths and TN > 0.550 g kg^{-1}, showing equivalent results: Adj.R2 = 0.85 and RPD = 2.64 for FieldSpec using Visible and SWIR bands, and Adj.R2 = 0.82 and RPD = 2.43 for EnMAP using SWIR bands. The predictive ability of EnMAP proved been similarly effective as that of the high-spectral resolution hyperspectral sensor, performing efficiently in modelling SOC and TN at the field scale and reducing the need for expensive and time-consuming laboratory analyses.

Keywords: soil monitoring; hyperspectral remote sensing; proximal sensing; organic matter

RESUMO

Analisar a dinâmica do carbono orgânico do solo (SOC) e do nitrogênio total (TN) é essencial para compreender questões agronômicas e ambientais. O sensoriamento remoto hiperespectral mostrou-se uma ferramenta promissora para a análise rápida e não destrutiva dos atributos do solo, e o sensor do satélite EnMAP é considerado um avanço significativo para estudos de monitoramento ambiental em larga escala utilizando sensoriamento remoto. Esta pesquisa teve como objetivo avaliar o desempenho do EnMAP na predição de SOC e TN em solo descoberto, comparando com as observações obtidas in situ por espectroscopia. Noventa e cinco amostras superficiais de solo (0-1 cm) foram coletadas no Perímetro Irrigado do Baixo Acaraú, no nordeste do Brasil. As áreas investigadas foram selecionadas utilizando imagens do EnMAP, e leituras espectrais in situ foram realizadas com o espectrorradiômetro ASD FieldSpec® (350-2500 nm). Os dados espectrais foram processados e as bandas mais relevantes para SOC e TN foram determinadas por meio de uma seleção sequencial, seguida do ajuste de modelos de regressão linear múltipla utilizando validação cruzada. Os modelos apresentaram bom desempenho preditivo para ambos os sensores. O EnMAP se destacou na faixa de SOC entre 11,6-23,6 g kg^{-1} ($R^2_{adj.}$ = 0,91; RPD = 6,03), enquanto o FieldSpec mostrou melhor desempenho na faixa de SOC entre 9,2-11,6 g kg^{-1} e acima de 16 g kg^{-1} ($R^2_{adj.}$ = 0,83; RPD = 2,66). Ambas as técnicas de sensoriamento apresentaram bom desempenho na predição de SOC utilizando comprimentos de onda SWIR e TN > 0,550 g kg^{-1}, mostrando resultados equivalentes: $R^2_{adj.}$ = 0,85 e RPD = 2,64 para o FieldSpec utilizando bandas Visíveis e SWIR, e $R^2_{adj.}$ = 0,82 e RPD = 2,43 para o EnMAP utilizando bandas SWIR. A capacidade preditiva do EnMAP provou ser igualmente eficaz à do sensor hiperespectral de alta resolução espectral, demonstrando eficiência na modelagem de SOC e TN em escala de campo e reduzindo a necessidade de análises laboratoriais caras e demoradas.

Palavras-chave: monitoramento do solo; sensoriamento remoto hiperespectral; sensoriamento proximal; matéria orgânica

6.1 INTRODUCTION

The quantification of Soil Organic Carbon (SOC) and Total Nitrogen (TN) is crucial for appraising soil quality for agricultural landscapes and mitigation of environmental degradation. Approximately 58% of organic matter comprises carbon, making it the most important constituent of the soil carbon cycle (Raheb et al., 2017; Hu et al., 2019; Ribeiro et al., 2021). Nitrogen is present in the structure of organic materials, making up proteins, nucleic acids, chitin, and the cell walls of decomposing organisms. Nitrogen can be incorporated into the soil as urea, via fertilizers or animal excrement and manure (Chen et al., 2020; Jiang et al., 2024).

As the demand for sustainable soil management practices grows, research into analytical methods becomes even more important. SOC is generally determined using dry combustion (Ben-Dor; Banin, 1989) or wet digestion (Walkley; Black, 1934), while TN is quantified by digesting the organic material in acid medium at high temperature, followed by distillation and titration with hydrochloric acid (Kjeldahl, 1883), all conducted in laboratory conditions.

To reduce costs and frequent equipment maintenance, remote sensing has emerged as a viable alternative to traditional methods thanks to the fast response and its capacity to derive many different soil parameters from just one measurement (Xiaoju et al., 2021; Amoli et al., 2024). Studying the spectral behavior of the soil enables to analyze properties with non-invasive techniques, based on the relationship between the incident energy and the energy reflected by the surface at the same wavelength (Lanfranchi et al., 2021).

Spectroradiometers, such as the ASD FieldSpec®, with a 1 to 3 nm interval in the 350-2500 nm range, have been used to analyze soil spectral curves and detect subtle features associated with spectrally active properties, such as texture, color, moisture, and organic matter content (Poppiel et al., 2019; Reda et al., 2019). Such devices allow detailed spectral information to be obtained from small soil samples, offering a more accurate analysis of the characteristics of the soil (Andrade et al., 2022).

Reflectance spectroscopy in the 350-2500 nm range has been successfully applied to the analysis of elements that affect the development of agricultural crops, which can be efficiently quantified by predictive models based on the spectral responses of the soil (Ribeiro et al., 2021; Cambou et al., 2022).

In addition to point-wise laboratory and field spectroscopy, spaceborne imaging spectroscopy has emerged as a promising solution for estimating soil attributes from space. Hyperion was the first orbital hyperspectral sensor used for studies in the period 2000 to 2017, with 198 calibrated spectral bands from 400 to 2400 nm (Zhang; Gui-Cheng, 2020). Following the discontinuation of Hyperion, new hyperspectral satellite missions were developed, such as PRISMA, launched in 2019 by the Agenzia Spaziale Italiana (ASI), with 237 spectral bands between 400-2505 nm (Castaldi et al., 2015; Amoli et al., 2024).

In 2022, the German EnMAP satellite (Environmental Mapping and Analysis Program) was launched, with its mission focusing on high-resolution imaging spectroscopy. EnMAP provides spectral information in 224 bands between 418-2445.5 nm, with scenes in 30-km swath, a 30-m spatial resolution and off-nadir viewing angle from ± 30º. The main instrument of the sensor is the Hyper-Spectral Imager (HSI), which provides high spectral, geometric, and radiometric accuracy. In the HSI, the visible and near-infrared (VNIR) region (418-993 nm, in intervals of 6.4 nm) and short-wave infrared (SWIR) region (902-2445.5 nm in intervals of 10 nm) operate individually and can be assessed separately or jointly (Steinberg et al., 2016; Storch et al., 2023)

The potential of EnMAP has been investigated to improve environmental monitoring by providing valuable large-scale information on soil properties and conditions, and to complement laboratory reflectance spectroscopy data (Chabrillat et al., 2024). Before the launch of EnMAP satellite in 2022, studies used simulated EnMAP data generated by the EnMAP End-to-End Simulator (Segl et al., 2012) tool and have given satisfactory results in assessing levels of clay, iron oxides and organic carbon at the surface (Guanter et al., 2015; Steinberg et al., 2016).

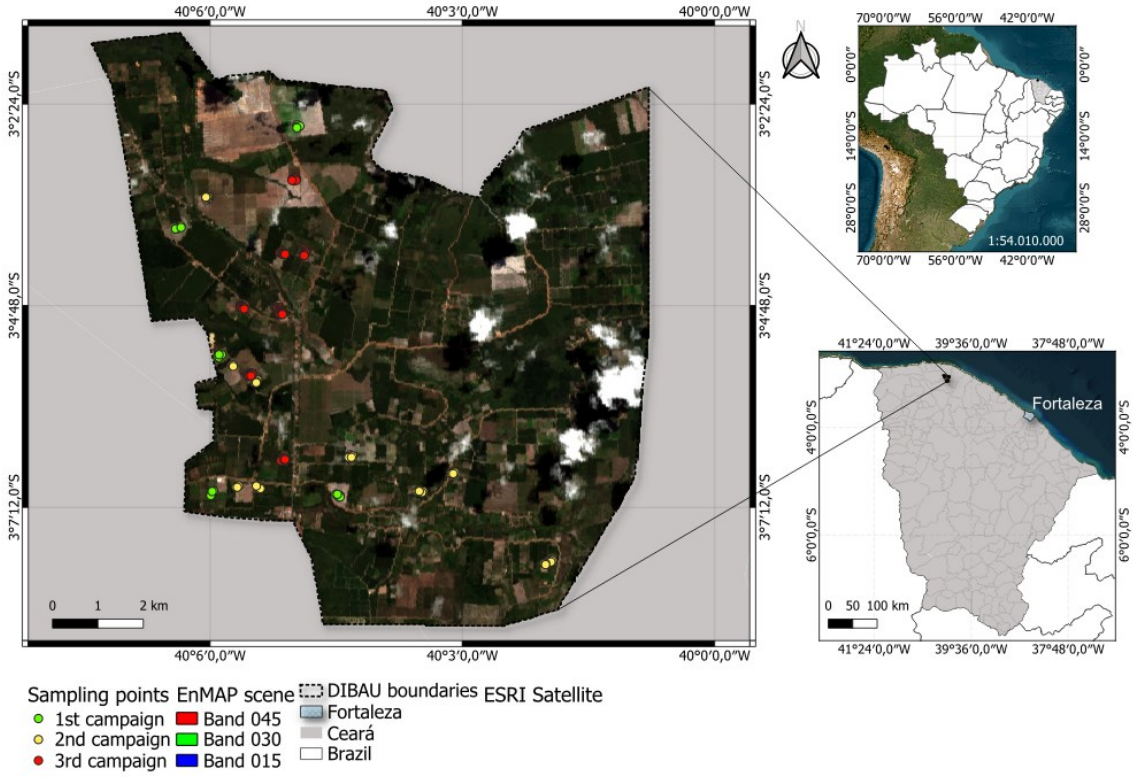
In this context, the research presented in this article was based on the hypothesis that the EnMAP orbital hyperspectral sensor is an auxiliary tool for providing accurate information that allows quantifying soil organic carbon and total nitrogen in different types of soil and landscapes, similar to those provided by the ASD FieldSpec® hyperspectral sensor in field campaigns. The main objective is to evaluate the performance of EnMAP as a tool for estimating SOC and TN, comparing its predictive ability with that of FieldSpec (350-2500 nm) under field conditions.

6.2 MATERIAL AND METHODS

6.2.1 The study Area

Ground surface samples of different classes of soil were collected in the Lower Acaraú River Irrigation District (DIBAU), located in the state of Ceará, northeastern Brazil (Figure 10). The irrigation district encompasses approximately 570 agricultural production farms and covers an area of 11,000 ha. According to the Köppen classification, the climate in the region is characterized as Aw, tropical rainy, with an average annual rainfall of 900 mm and average annual temperature of 28.1°C.

Figure 10. Location of the Study Area and Sampling Sites. Satellite image source: EnMAP data ©DLR [2023].



Thirteen agricultural sites were accessed for soil sampling, nine of which had fully uncovered soil ready for reseeding crops, while the others had fruit crops in the early stages of development, namely coconut, cashew, watermelon and acerola. There were small quantities of grass straw, burning residue or manure on the surface of each of the sampled plots, which were used by farmers as sources of organic matter for the soil (Appendix A).

Three different field campaigns were conducted, in which a total of 95 samples were taken from the surface layer of the soil (0-1 cm).

6.2.2 Planning of Field Sampling and Spectral Data Acquisition

The EnMAP images were acquired on demand through the platform's online portal (EnMAP.org) at L2A (land-surface reflectance) processing level, based on the following criteria: viewing angle as close as possible to the nadir and cloud cover conditions below 30%, ensuring that the area of interest was imaged with maximum clear sky conditions.

The images acquired were used to select the sample sites, giving priority to areas with exposed soil with minimal interference from vegetation or clouds. The sampling campaigns were scheduled to coincide with the dates closest to the imagery acquisition, reducing the influence of any environmental changes between the date of the sensor imaging and the field sampling (Table 4).

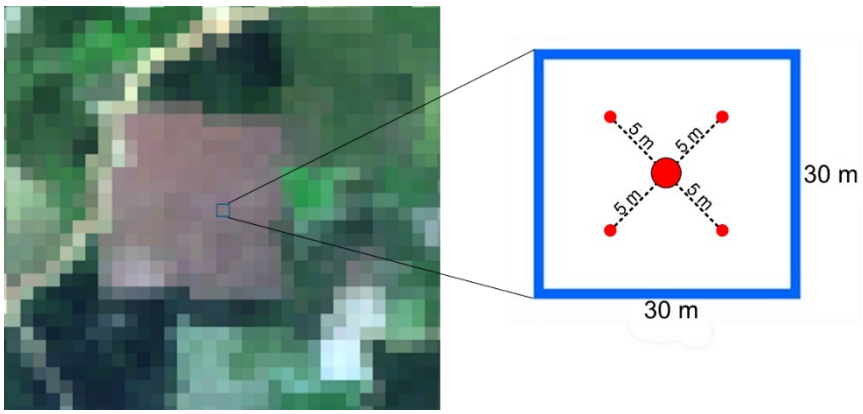
Table 4. Dates of the field campaigns, number of sampling spots, and corresponding EnMAP scenes.

Campaign	Sampling dates (2023)	Spots	EnMAP scene dates (2023)
First	23–25 June	18	19 June
Second	19–21 July	32	24 July
Third	30 July–1 August	45	24 July

During the second campaign, no EnMAP images with favorable sky conditions (cloud-free) were available for the study area. As such, the closest available scene to the actual field survey and conditions was used.

The sampling points were defined based on the geographic coordinates of the center of the pixel that represented exposed soil. The selection was made from EnMAP images in NIR false-color RGB, using a combination of the bands 75-47-30. Each observation consisted of five individual samples, the first taken at the central point previously defined by the coordinates of the pixel of interest, and the others positioned five meters along the diagonal (Figure 11).

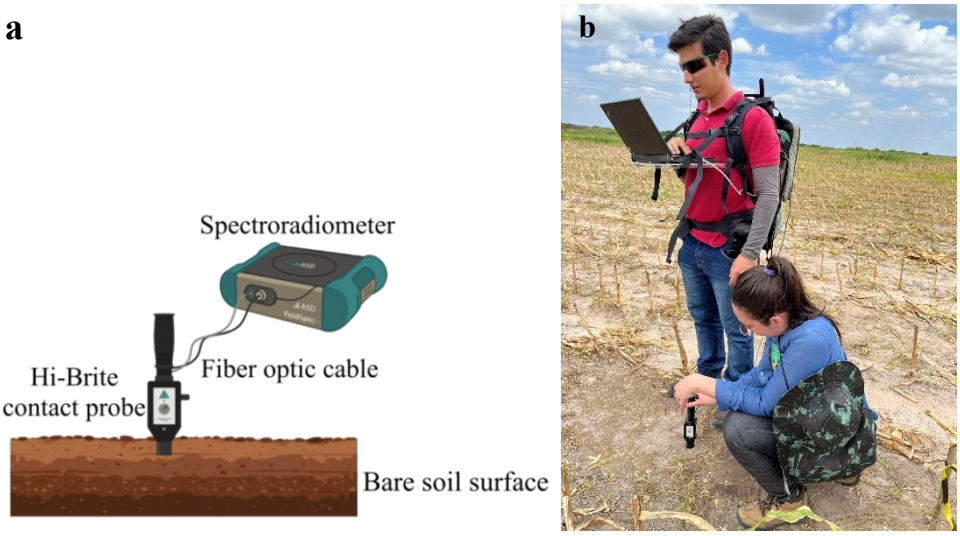
Figure 11. Soil sampling based on the location of exposed soil pixels.
Satellite image source: EnMAP data ©DLR [2023].



The five samples were then mixed for each plot to create a representative overall sample for one image pixel. Considering that the spatial resolution of EnMAP is 30 m (900 m^2), this distribution was planned to avoid sampling along the edges of the pixel, ensuring less influence from mixed targets.

Spectral information from the soil surface was acquired *in situ* using a Hi-Brite probe and a portable ASD FieldSpec® PRO FR 3 spectroradiometer (350-2500 nm) (Figure 12a), prior to collecting the soil samples directly in the field. The spectral analysis was conducted at the central coordinate of the pixel in the EnMAP images and at random locations around it (Figure 12b). Three readings were taken at each point, each with ten automatic repetitions, giving a total of 30 readings. The average of the spectral signatures of each sample represented the reflectance behavior of the soil between 350-2500 nm.

Figure 12. a) Scheme for field spectral data acquisition using the ASD FieldSpec® PRO FR 3 (350–2500 nm); b) Photo of hyperspectral data collection on the soil surface.



The sensor was standardized at the start of each reading using a Spectralon plate, which simulates a Lambertian surface, reflecting 100% of the incident energy in all directions. The Hi-Brite probe, recommended for use in the field, including in the shade, has an internal light source and rubber edges that isolate the area being assessed.

6.2.3 Chemical analysis of the soil organic carbon and total nitrogen

Soil samples were air-dried, crushed and sieved through a 2mm mesh. A part was macerated to determine the SOC and TN in the Pedology Laboratory of the Soil Science Department, at the Centre for Agricultural Sciences of the Federal University of Ceará, Brazil.

The SOC was quantified using the wet oxidation method proposed by Walkley and Black (1934) and Yeomans and Bremner (1988). The total nitrogen was determined by the classic method of digestion in concentrated H_2SO_4 and a digestion mixture (potassium sulphate + copper sulphate), as described by Kjeldahl (1883).

6.2.4 Statistical analysis, selection of the spectral variables and multiple linear regressions

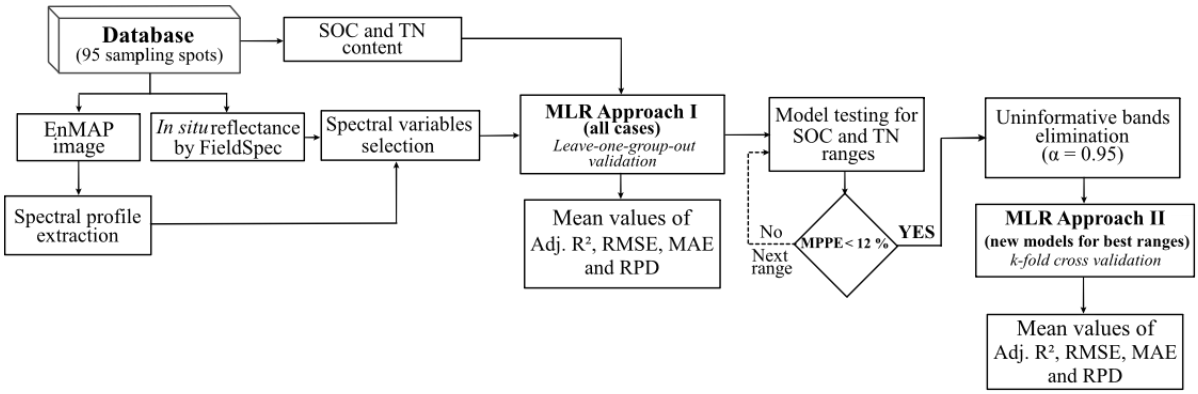
Descriptive statistics were used to assess the variation in SOC and TN in the soil samples (mean, median, maximum and minimum values), as well as a data frequency analysis and the Kruskal-Wallis test to compare mean values at 5% for each set of samples using the 'agricolae' package (Mendiburu, 2023) of the RStudio software (R Core Team, 2024).

The most influential spectral bands were selected for the spectra obtained by the EnMAP and FieldSpec hyperspectral sensors, with the aim of identifying the wavelengths that best explain the variation in the SOC and TN contents of the samples. The EnMAP dataset contains 224 spectral variables, while FieldSpec has 2150, making it necessary to reduce the dimensionality of the data. For this purpose, the stepwise method was implemented in the RStudio software (R Core Team, 2024) using the 'MASS' package (Venables; Ripley, 2002), which allowed identifying the most influential spectral variables for estimating SOC and TN to be identified, thereby ensuring simpler and more-robust model outputs.

Multiple Linear Regression (MLR) models were built separately for each sensor using the selected spectral variables. The models were initially fitted considering the entire dataset, covering the minimum and maximum levels of SOC and TN (Approach I, Figure 13). The "Leave-one-group-out" validation was used, in which each test group corresponded to samples from the same agricultural plot, tested alternately until all the groups had been validated.

After the adjustment and initial validation, Approach II (Figure 13) was applied, in which the models were refined considering the SOC and TN content that showed the best predictions based on the lowest Mean Percentage Prediction Error (MPPE). A limit of 12% was defined as the maximum MPPE to maintain sufficient observations in the model without losing predictive performance. In this step, only the most significant spectral variables among those previously selected were used to fit the new models (p-value < 0.05). For this approach, the k-fold cross-validation technique was used, where k was equal to 4.

Figure 13. Flowchart of methodological approaches used for development and validation of predictive models for Soil Organic Carbon (SOC) and Total Nitrogen (TN).



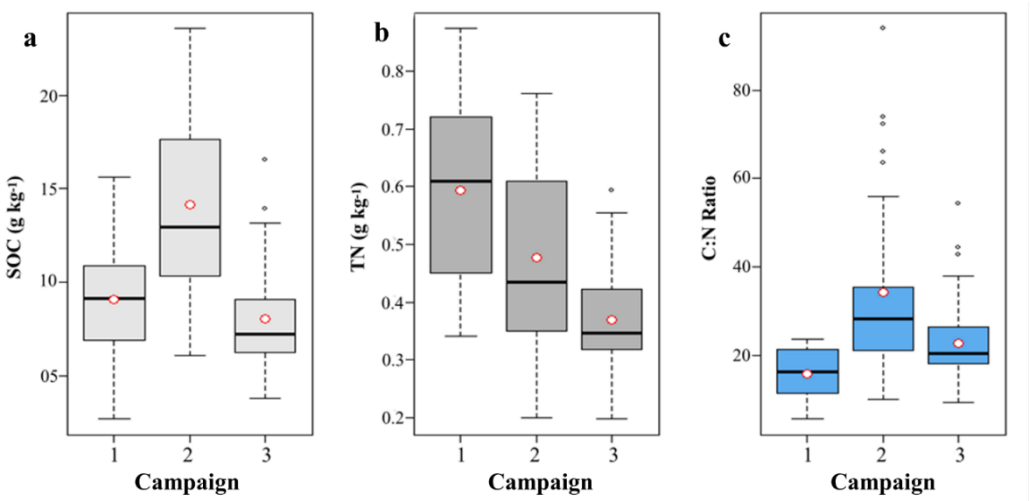
All the adjustments and validation were carried out as indicated in the Figure 4, evaluating predictive performance using the Adjusted R-squared (Adj.R^2), Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Ratio of Performance to Deviation – RPD, given by the ratio of the standard deviation of the observed SOC or TN content by the RMSE.

6.3 RESULTS

6.3.1 Descriptive Statistical Analysis

The exploratory analysis of the SOC and TN contents across the irrigated perimeter revealed variations between the samples. The SOC had the highest content in the second set of samples, with an average of 14.49 g kg^{-1} (Figure 14a), while the highest TN content was recorded in the first set of samples, with an average of 0.62 g kg^{-1} (Figure 14b).

Figure 14. Boxplot showing the variation in (a) soil organic carbon content, (b) total nitrogen, and (c) C:N ratio observed for each sampling campaign.



These variations were reflected in the average C:N ratio of the samples, with lower values in Sample 1 (mean value = 15.7) and higher values in Sample 2 (mean value = 34.56) (Figure 14c). Although the average organic carbon content did not vary between the first and third campaigns, both differed from the second campaign, which had a higher SOC content (Table 5), as shown by the Kruskal-Wallis test at a level of 95%. The TN content and the C:N ratio showed significant variations between the three campaigns (Table 5).

Table 5. Descriptive statistics and mean difference analysis of the SOC and TN content and C:N ratio across the sampling campaigns.

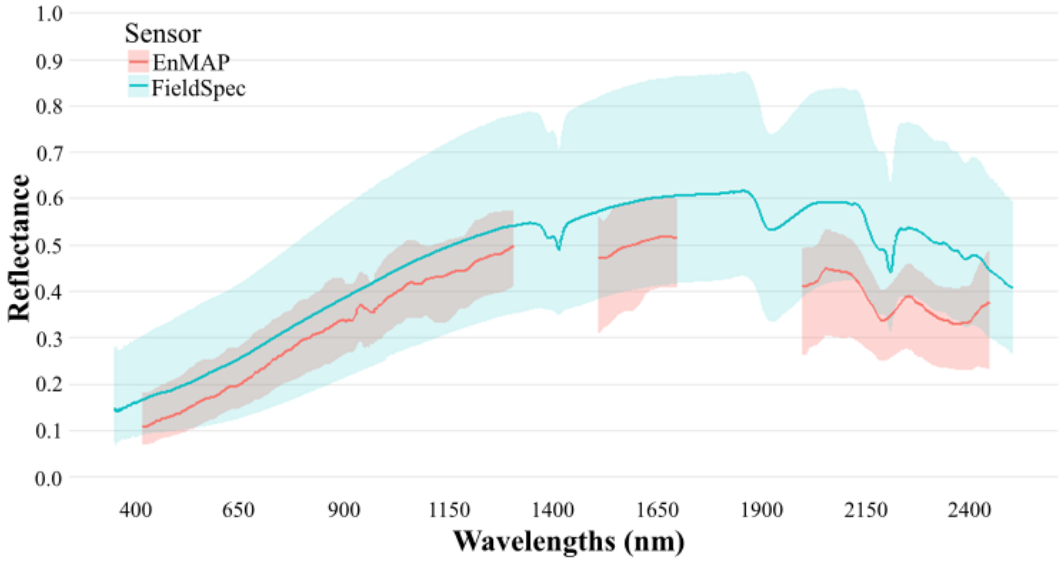
Campaign	SOC				TN				C:N ratio			
	max	min	μ	CV	max	min	μ	CV	max	min	μ	CV
	(g kg^{-1})				(g kg^{-1})							
First	16.24	2.67	9.08^{b}	40.6%	0.88	0.341	0.59^{a}	29.0%	23.65	5.46	15.67^{c}	35.9%
Second	23.60	6.06	14.14^{a}	36.6%	0.76	0.200	0.48^{b}	32.7%	93.95	9.97	34.16^{a}	58.2%
Third	16.51	3.79	8.04^{b}	36.1%	0.59	0.197	0.37^{c}	23.4%	54.33	9.24	22.56^{b}	39.8%

Mean values (μ) followed by the same letter do not differ significantly (α = 0.95).

6.3.2 Spectral Analysis: EnMAP versus FieldSpec

The EnMAP sensor produces 224 spectral bands at Level-2A (bottom-of-atmosphere [BOA] reflectance), therefore it provides similar information to that obtained in the field with the FieldSpec spectroradiometer, which operates at 2150 wavelengths. The differences between the sensors are most apparent in the 1300-1490 nm, 1706-1980 nm and 2200-2300 nm ranges (Figure 15).

Figure 15. Average spectral response of the soil acquired by the EnMAP sensor and the FieldSpec in situ spectroradiometer. The shaded area indicates the variation between the minimum and maximum reflectance values observed with each sensor.

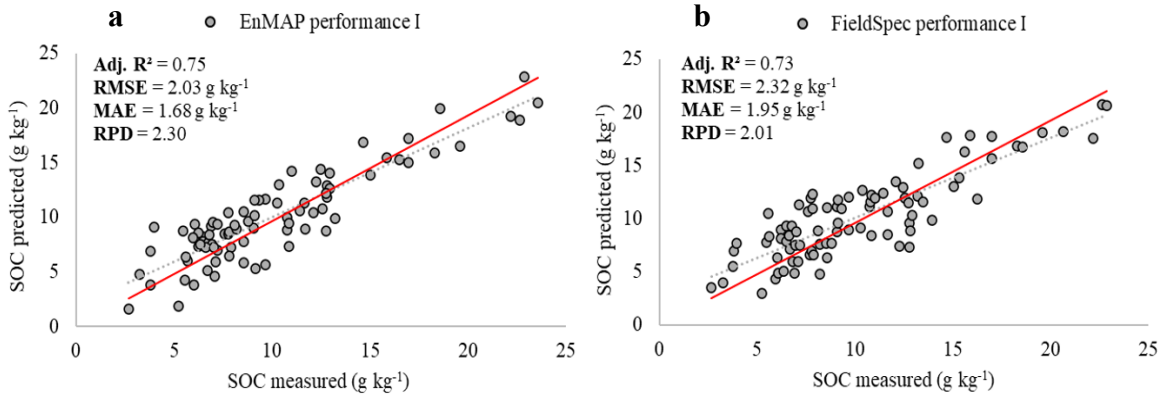


6.3.3 Regression Models using Hyperspectral Data

6.3.3.1 Prediction of Soil Organic Carbon – Approach I

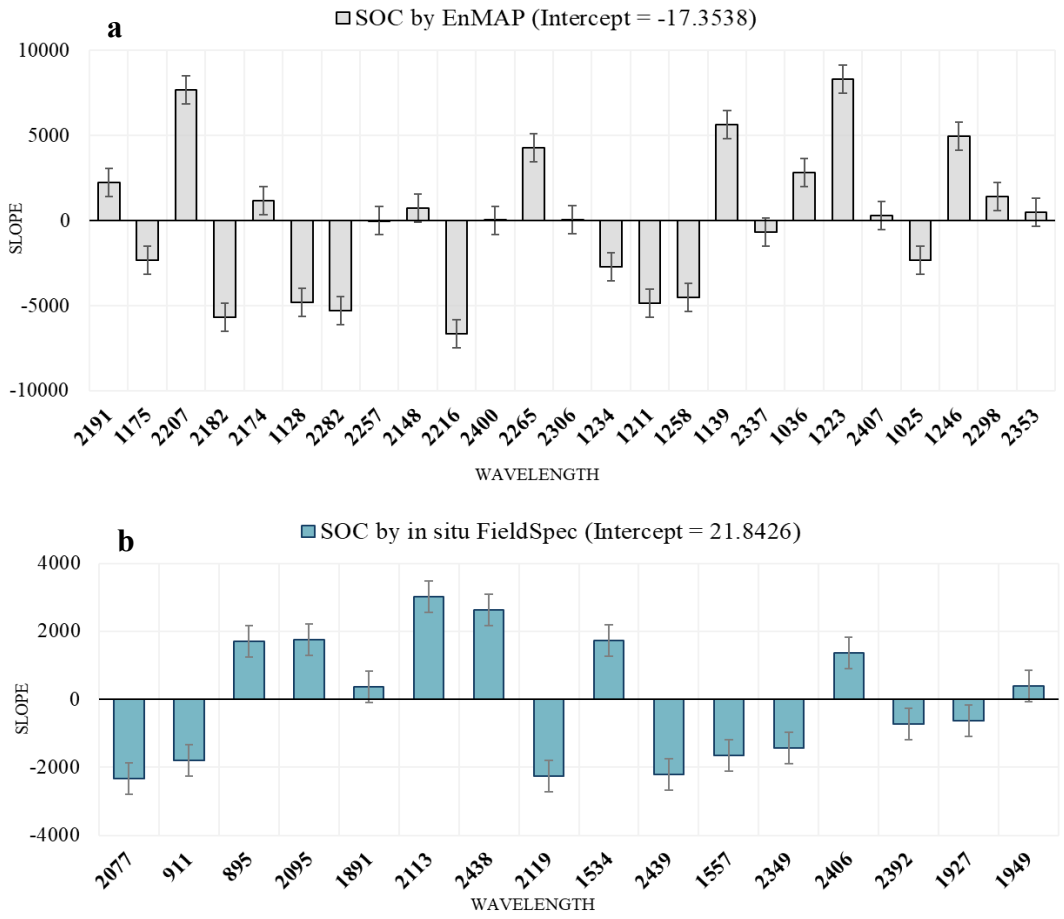
The models developed to estimate SOC content from the complete set of samples (ranging from 2.67 to 23.6 g kg^{-1}) showed high performance and robustness using both spectral sensors. The reflectance data from EnMAP, with Adj.R^2 = 0.75, RMSE = 2.03 g kg^{-1}, MAE = 1.68 g kg^{-1} and RPD = 2.30 (Figure 16a), were very accurate in predicting SOC with Approach I. Similarly, the model using the FieldSpec *in situ* spectral data obtained Adj.R^2 = 0.73, RMSE = 2.32 g kg^{-1}, MAE = 1.95 g kg^{-1} and RPD = 2.01 (Figure 16b), showing a slightly poorer performance, although still reliable for practical applications.

Figure 16. Performance of the regression models developed using spectral data from (a) EnMAP and (b) FieldSpec *in situ* for SOC prediction using Approach I.



The EnMAP model used 25 spectral bands from the NIR-SWIR region (1100-2400 nm) to estimate SOC with Approach I (Figure 17a). The model using FieldSpec data employed 16 wavelengths as predictor variables (Figure 17b), two in the visible region (895 nm and 911 nm) and the others in the SWIR (1530-2440 nm).

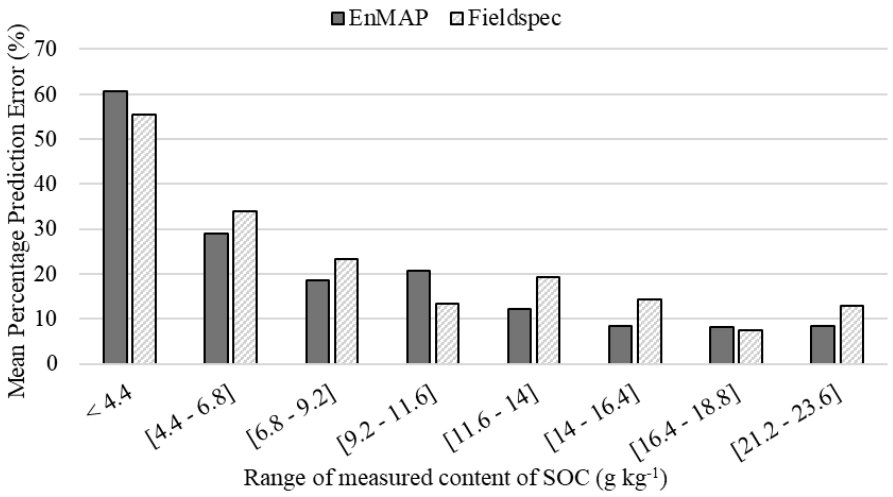
Figure 17. Model parameters for predicting Soil Organic Carbon (SOC) with Approach I, developed using spectral data from a) the EnMAP sensor and b) the FieldSpec spectroradiometer.



Although the models are generally reliable, the SOC ranges between 2.67 and 23.6 g kg^{-1} presented different values for MPPE for each sensor (Figure 18); it can be observed that MPPE is higher in the lowest SOC range (< 4.4 g kg^{-1}) and, that as the SOC values increase, the mean percentage error gradually decreases, stabilizing between 10-20% for ranges higher than 9.2 g kg^{-1}.

The results show that SOC values were best estimated by the EnMAP spectral data in the 11.6-23.6 g kg^{-1} range, with a maximum MPPE of 11.95%, whereas the *in situ* FieldSpec reflectance data showed the best performance for the SOC content between 9.2 and 11.6 g kg^{-1} and from 16.4-23.6 g kg^{-1}, with an MPPE ranging from 7.5% to 12%.

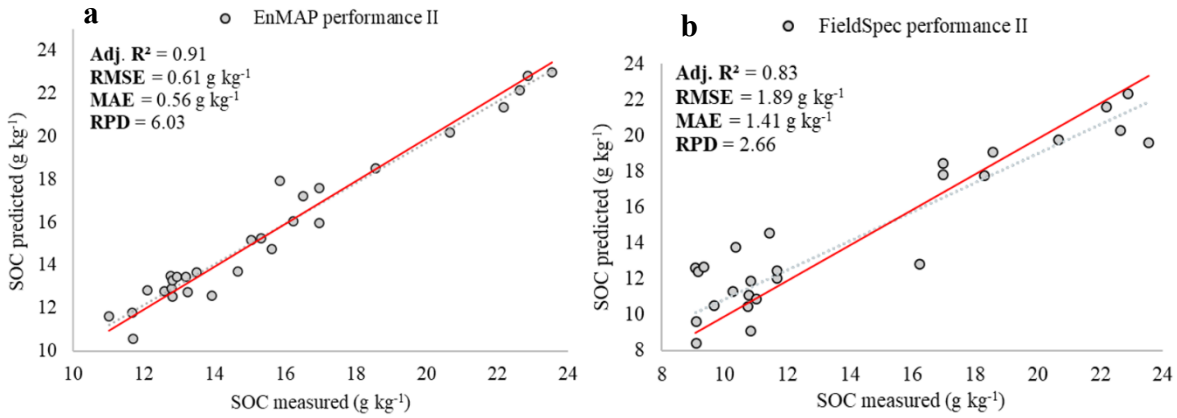
Figure 18. Mean Percentage Prediction Error (MPPE) of SOC for each measured content range using spectral data from EnMAP and FieldSpec.



6.3.3.2 Prediction of Soil Organic Carbon – Approach II

The EnMAP reflectance data yielded more robust models for predicting SOC between 11.6 and 23.6 g kg^{-1} (Figure 19a), with Adj.R^2 = 0.91, RMSE = 0.61 g kg^{-1}, MAE = 0.56 g kg^{-1} and RPD = 6.03, outperforming the model applied to the full set. On the other hand, FieldSpec performed relatively worse than EnMAP, but also showed an improvement when the prediction was limited to the range between 9.2 and 11.6 g kg^{-1} and to SOC content higher than 16 g kg^{-1} (Figure 19b), with Adj.R^2 = 0.83, RMSE = 1.89 g kg^{-1}, MAE = 1.41 g kg^{-1} and RPD = 2.66.

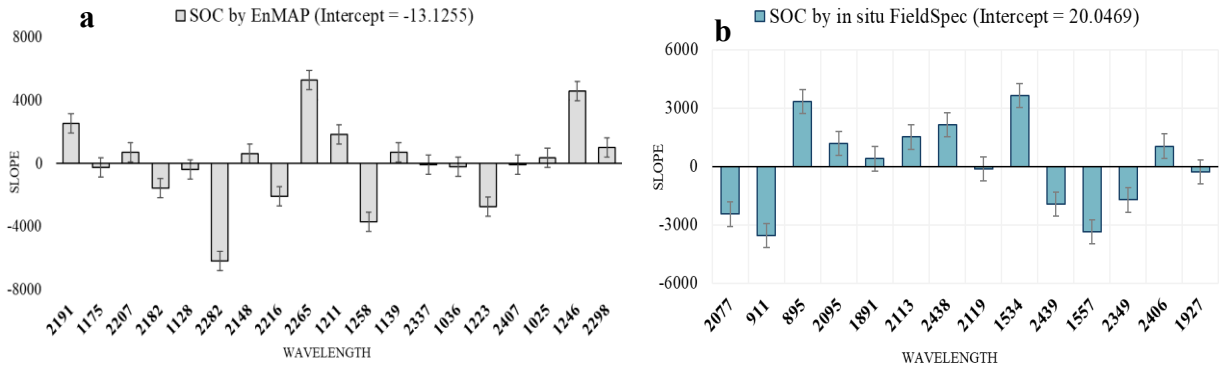
Figure 19. Performance of the models developed using spectral data from (a) EnMAP and (b) FieldSpec in situ for SOC prediction within the defined ranges of Approach II.



When restricting the data set to the best-predicted SOC ranges, the regression models were refined by filtering the wavelengths with the best statistical significance (*p*-value < 0.05) for the new data set. This reduction resulted in an optimized predictive model for

EnMAP, with 19 wavelengths capable of estimating the SOC content in the 11.6 to 23.6 g kg^{-1} range, achieving Adj.R^2 = 0.91 (Figure 20a). The most influential bands were around 1200 nm and 2200 nm. For FieldSpec, the best-performing regression model included 14 significant wavelengths (Adj.R^2 = 0.83), with the greatest influence at 895 nm, 911 nm, 1534 nm, 1557 nm, 2077 nm and in the 2400 nm region (Figure 20b).

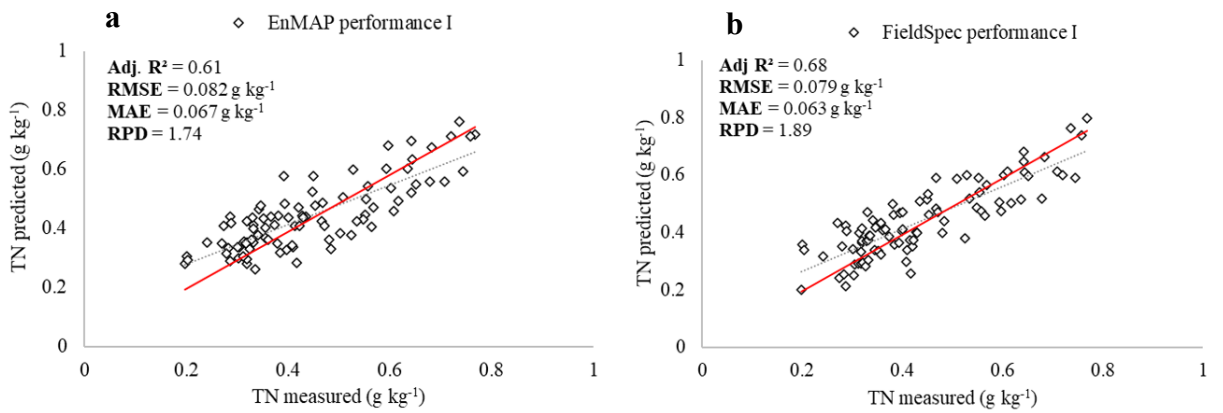
Figure 20. Model parameters for predicting Soil Organic Carbon (SOC) with Approach II, developed using spectral data from a) EnMAP and b) FieldSpec.



6.3.3.3 Prediction of Total Nitrogen – Approach I

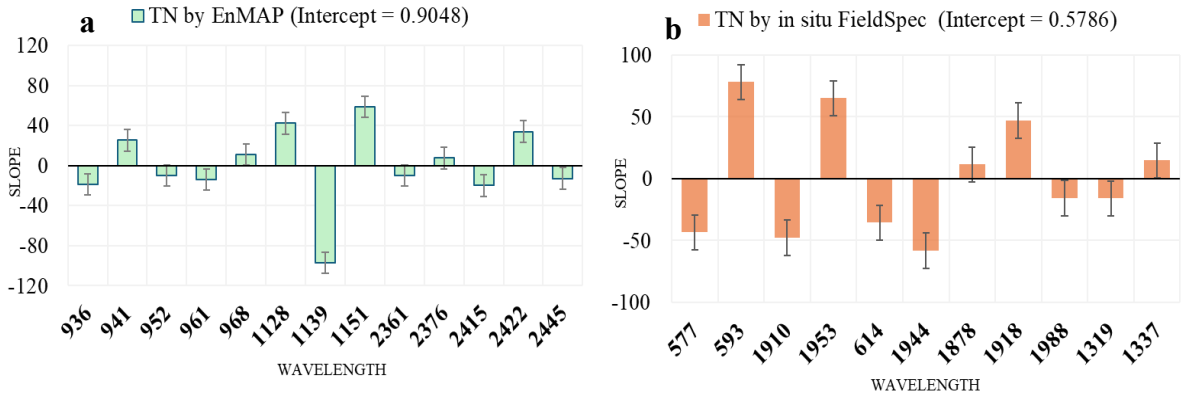
The model developed using EnMAP data presented Adj.R^2 = 0.61, RMSE = 0.082 g kg^{-1}, MAE = 0.067 g kg^{-1} and RPD = 1.74 (Figure 21a), demonstrating ability for reliable predictions, but requiring some adjustment to properly estimate TN levels from 0.197 to 0.880 g kg^{-1}. The model generated from the FieldSpec data achieved Adj.R^2 = 0.68, RMSE = 0.079 g kg^{-1}, MAE = 0.063 g kg^{-1} and RPD = 1.89 (Figure 21b), revealing a slightly better performance than the EnMAP model, but still requiring adjustments to improve the prediction ability.

Figure 21. Performance of the predictive models developed using spectral data from (a) the EnMAP sensor and (b) FieldSpec for predicting TN in the 0.197–0.880 g kg^{-1} range.



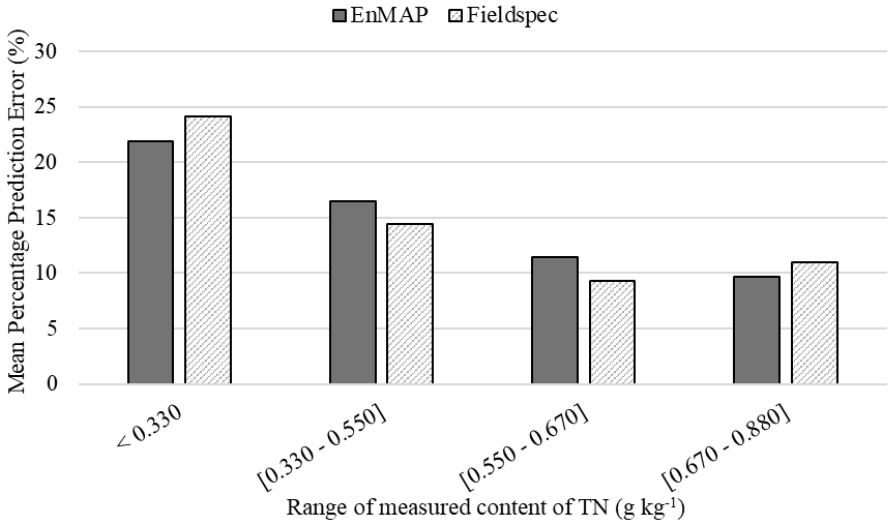
With stepwise selection, fewer spectral bands were considered for predicting TN at the ground surface (Figure 22a) compared to the predictive models for SOC. In the model based on the EnMAP spectral data, thirteen wavelengths were chosen in the NIR-SWIR region: between 936-968 nm, around 1100 nm and between 2361-2445 nm. On the other hand, the model using the *in situ* FieldSpec data presented eleven influential spectral bands distributed over the visible region (577 nm, 593 nm and 614 nm) and in the 1900 nm region (Figure 22b).

Figure 22. Parameters of TN predictive models for Approach I developed using spectral data from the a) EnMAP and b) FieldSpec sensors.



The results show that both EnMAP and FieldSpec were able to reliably predict a TN content between 0.197 and 0.880 g kg^{-1} for ground surfaces of different soils, although it is necessary to adjust the dataset. As with the SOC prediction, the models built using spectral data from EnMAP and FieldSpec showed the highest MPPE for a lower TN content (< 0.330 g kg^{-1}). In this range, the errors reached 22.4% with EnMAP and 24.10% with FieldSpec (Figure 23).

Figure 23. Mean Percentage Error of prediction (MPPE) of TN for each measured content range using spectral data from EnMAP and FieldSpec.

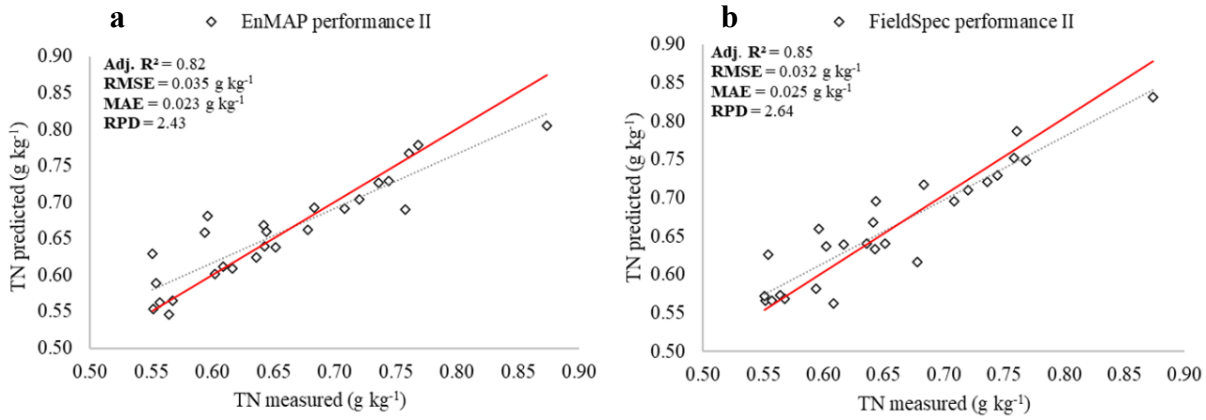


With both sensors, the prediction error gradually decreased as the TN content of the soil samples increased. The most accurate estimates for EnMAP were observed between 0.55 and 0.88 g kg^{-1}, with an MPPE of 11.45% for the lower range and 9.83% for the upper range. Similarly, the model based on FieldSpec showed an MPPE of 9.40% and 11% for the upper and lower ranges, respectively.

6.3.3.4 Prediction of Total Nitrogen – Approach II

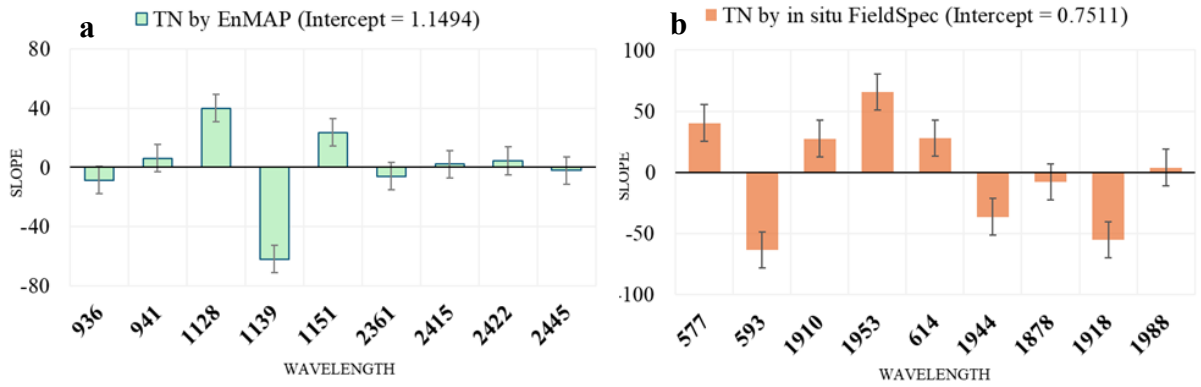
When adjusting the dataset for the 0.55 to 0.88 g kg^{-1} range of TN, the predictive models provided satisfactory performance using spectral data from both sensors. The model based on EnMAP obtained Adj.R^2 = 0.82, RMSE = 0.035 g kg^{-1}, MAE = 0.023 g.kg^{-1} and an RPD of 2.43, indicating excellent predictive ability at the determined levels (Figure 24a). These metrics were similar to those obtained using the FieldSpec spectral data (Figure 24b), although with a slightly better performance (Adj.R^2 = 0.85, RMSE = 0.032 g kg^{-1}, MAE = 0.025 g kg^{-1} and RPD = 2.64). It is worth noting that the regression parameters for all the models tested are provided in Appendix C

Figure 24. Performance of the models developed using spectral data from (a) EnMAP and (b) FieldSpec in situ for predicting TN in the ranges defined for Approach II.



Both models showed good accuracy and robustness in predicting the TN content, with low dispersion around the line of best fit (red) and satisfactory accuracy metrics. For the 0.55 to 0.88 g kg^{-1} range of TN, the models reduced the number of significant wavelengths for the prediction (p-value < 0.05). For the EnMAP model, nine spectral bands were kept as predictor variables (Figure 25a), with the most influential being 1128 nm, 1139 nm and 1151 nm. For FieldSpec, the significant spectral bands were concentrated in the visible spectrum (577-614 nm) and in the SWIR (1878 nm, 1918 nm, 1944 nm and 1953 nm), also with a total of nine spectral bands that are relevant for predicting TN from reflectance data (Figure 25b).

Figure 25. Parameters of TN predictive models for Approach II developed using spectral data from a) EnMAP and b) FieldSpec.



6.4 DISCUSSION

6.4.1 Quantitative distribution of SOC and TN observed during sampling.

The variability in the SOC and TN content and in the C:N ratio was influenced by the differences in land use at the sample sites in the irrigated perimeter. This is because, regardless of the sample set, all the collected samples were of a sandy texture, with a sand content higher than 88% and clay content ranging from 1.8% to 6.0% (Appendix B). The texture of the soil surface may therefore not have been a determinant factor in the accumulation of organic properties.

The second set of samples (second campaign) revealed the highest average SOC content, with 14.49 g kg^{-1} (Figure 14a), a result corroborated by the mean-value comparison test (Table 5), which can be attributed to the type of material used as a source of organic carbon. This set was collected in bare-ground areas with a low level of residue accumulation, the principal residue being remnants of earlier corn crops and burning residue (Appendix A). During the sampling period, the temperature varied between 26ºC and 32.5ºC, which may also have affected the decomposition of any organic matter that was added, even in small quantities.

Studies showed that applying or incorporating straw in sandy soils is an effective way of increasing the soil carbon stock, since these soils have a rapid and strong response to SOC accumulation. Straw beneath the surface contributes significantly to an increase in SOC levels, especially under irrigated conditions in regions with elevated temperatures (Wang et al., 2021; Gocke et al., 2023; Schjønning, 2023).

The results found in this study are in accordance with the work of Hu et al. (2019), who found an increase in carbon storage in semi-arid areas with soil under irrigation following the application of corn straw between crops.

In contrast to SOC, the average TN content was higher during the first sampling campaign (0.62 g kg^{-1}), with a gradual reduction in subsequent collections (Figure 14b). This result was expected since there was a predominance of legume residue on the soil surface at this time.

Furthermore, nitrogen fertilization is a common and frequent practice among producers in the corporate areas, which contributed to the increase in nitrogen at the surface at the time of soil sampling. Urea applied to the soil can be converted into different forms of nitrogen that are detectable by TN analyses, which in turn have a direct effect on quantifying the element (Bolinder et al., 2020; Chen et al., 2020; Jiang et al., 2024).

Fluctuations in the SOC and TN content during the sampling campaigns had a direct impact on the C:N ratio for each collection (Figure 14c). The second campaign showed the highest value for the C:N ratio (34.56), reflecting the relative increase in SOC at the expense of nitrogen, influenced by the excess of corn straw and charcoal on the surface, as reported by Zhang et al. (2015).

In addition to materials rich in organic carbon that accumulated on the surface, some of the areas accessed during the second sampling campaign also contained significant amounts of seasoned cattle manure, which may help increase TN in soil samples (Chen et al. 2020). Organic residue, such as cattle manure or legume leaves, have a lower C:N ratio and consequently decompose rapidly, which helps to mineralize the N, making it available and reducing the C:N ratio in the soil (Villamil et al., 2015).

6.4.2 Spectral Analysis

The EnMAP hyperspectral sensor, with 224 surface reflectance bands, demonstrated its ability to capture important spectral characteristics of the soil, even though it has a lower spectral resolution than FieldSpec, with 2150 pieces of spectral information (Figure 15).

The main differences between the sensors occur in the 1300-1490 nm and 1706-1980 nm ranges, known as ‘bad bands’ due to the noisy influence of atmospheric water. However, excluding this noise in processing the EnMAP data minimizes the atmospheric effects

and improves the quality of the usable data (Guanter et al., 2015; Fruth et al., 2018). EnMAP generally performs in a similar way to FieldSpec in capturing the average spectral response of exposed soil, except in areas of higher atmospheric interference.

Peaks around 1900 nm are often associated with OH groups present in 2:1 clay mineral and the NH amide functional group, also found in organic compounds and in surface-applied urea (Angelopoulou et al., 2019; Cambou et al., 2022). The absorption in this region is clearly detected by FieldSpec. However, with EnMAP, this band is excluded when processing images using surface reflectance data, limiting its ability to capture information specific to these wavelengths (Salazar; Coffman, 2020).

Absorption peaks around 2200 nm and 2300 nm, important for detecting clay minerals and carbonates (Bishop et al., 2008; Zhou; Wang, 2017; Monteiro et al., 2024), are detected by both sensors. However, the higher spectral resolution of FieldSpec allows these peaks to be identified with better definition, while with EnMAP, they appear smooth due to the larger width of the spectral bands.

The smoothed information from EnMAP can, however, be highlighted using spectral transformation techniques such as first- and second-order derivatives and continuum removal, which are able to highlight subtle variations that are impossible to see in the reflectance spectrum.

The high spectral resolution of FieldSpec is also reflected in the greater variability of the captured data, highlighting its sensitivity to variations in the soil components and environmental conditions. In contrast, EnMAP presents lower variability, indicating a more stable response. Despite the less-detailed responses of EnMAP compared to FieldSpec, Chabrillat et al. (2024) showed that the requirements of reflectance spectroscopy for detecting different soil attributes were met by that sensor two years after its launch, in its first exploratory results.

6.4.3 Prediction of Soil Organic Carbon

The reflectance data provided by EnMAP performed well in predicting the SOC from 2.67 to 23.6 g kg^{-1}, with Adj.R^2 = 0.75, RMSE = 2.03 g kg^{-1}, MAE = 1.68 g kg^{-1} and RPD = 2.30 (Figure 16a). Similarly, the spectral information obtained by FieldSpec directly in the field was able to estimate the SOC content, although with an Adj.R^2 = 0.73, RMSE = 2.32 g kg^{-1}, MAE = 1.95 g kg^{-1} and RPD = 2.01. This similarity is likely because the image capture and

sampling occurred only a few days apart, providing this research with a robust database of both spectral and chemical information from each spot.

Equivalent results were achieved by Gholizadeh et al. (2022), who developed predictive models with $R^2 = 0.74$ in samples of litter and $R^2 = 0.77$ in samples of topsoil. The authors obtained the spectral readings in the laboratory under controlled conditions and using a contact probe, which suggests that acquiring spectral data in the field using a probe – as in the present research – also can generate efficient predictive models.

Reflectance spectroscopy is highly effective at estimating soil properties when the RPD of the predictive model is higher than 2.0. If $1.4 < RPD < 2.0$, the model is still usable, but requires adjustment, while RPD values < 1.4 indicate poor predictive performance (Chang et al., 2001; Bellon-Maurel et al., 2010). Based on this criterion, the spectral data acquired by both EnMAP and FieldSpec resulted in excellent predictive models, capable of estimating SOC levels from the full range of 2.67 to 23.6 g kg^{-1}.

Carvalho et al. (2022) used remote sensing to predict SOC in sediment from reservoirs in the semi-arid region of Brazil, with results similar to those presented in this study. They used the FieldSpec sensor in the laboratory with a halogen lamp to model predictions using partial least squares regression (PLSR). Pudelko and Chodak (2020) estimated SOC levels between 2.9 and 25.2 g kg^{-1} in soils of different textures, applying PLSR models and artificial neural networks. Their results presented RPD values > 2.0 and RMSE values between 2.18 and 2.25 g kg^{-1}.

Mzid et al. (2022) and Amoli et al. (2023) also found similar results when investigating the use of the PRISMA sensor in estimating SOC using complex regression methods, with both studies achieving $R^2 \approx 0.85$. These findings suggest that multiple linear regression models, such as those employed in this study, can be just as effective as more-complex predictive models.

Results of the EnMAP predictions are also consistent with those of Steinberg et al. (2016), who used simulated orbital sensor data through the EnMAP End-to-End Simulation tool (Segl et al., 2012) to estimate SOC using regression models based on reflectance. The authors classified EnMAP as a reliable predictive tool, with performance metrics of $R^2 = 0.67$, RMSE = 2.8 g kg^{-1} and RPD = 1.7.

Similarly, Castaldi et al. (2016) evaluated predictive SOC models using simulated spectral data from EnMAP but found lower accuracy in practical applications (Adj.R^2 = 0.25-0.67, RMSE = 0.20-0.48 g kg^{-1} and RPD = 1.17-1.80). These studies show that EnMAP is currently robust in acquiring spectral information at field scale and can be used as a reliable forecasting tool.

The EnMAP satellite sensor, which was designed for environmental and agricultural applications, proved being capable of feeding data to build an effective regression model (Adj.R^2 = 0.75) for estimating SOC using 25 bands, mostly in the SWIR region. The importance of this spectral region for EnMAP studies is corroborated by research of Castaldi et al. (2016) and Mzid et al. (2022), who emphasized the importance of the shortwave infrared in predicting soil properties. On the other hand, FieldSpec used in the field made it possible to develop a regression model with 16 spectral bands between 895 nm and 2440 nm, obtaining Adj.R^2 = 0.73. The reduced number of bands required for SOC prediction suggests that FieldSpec, by capturing information with less influence from spectral mixing, detects subtle variations in the soil more efficiently, reducing the need for complex mathematical resources when modelling.

Tiwari et al (2015) assessed the importance of spatial resolution in SOC prediction, suggesting that hyperspectral data collected in the field can facilitate the detection of soil characteristics that are important when modelling. In addition, the study suggests that the use of a reduced number of spectral bands, when well selected, can generate simpler and more effective models for identifying local variations in SOC.

In this study, the most influential wavelengths in predicting SOC with Approach I (2.67-23.6 g kg^{-1}) using EnMAP spectral data are at 1128 nm, 1139 nm, 1175 nm, 1211 nm, 1223 nm, 1246 nm, 1258 nm, 2182 nm, 2207 nm, 2265 nm and 2282 nm. This is in line with findings from Stuart (2004), who identified the effect of C-H stretching in aromatic organic compounds on energy absorption in the 1100-1250 nm range. Similarly, Stenberg et al. (2010) emphasized the importance of carbon in functional groups and aromatic compounds in the 1100 nm region, as well as the excitation of aliphatic molecules with carboxyl, which influence absorption at 2270 nm.

More recently, Ribeiro et al. (2023) studied the influence of organic carbon on humic substances on the surface of different soils and also identified absorption bands around 1130 nm and 1220 nm related to the prediction of organic carbon in humic acids and humin.

The use of FieldSpec spectral data enabled the development of predictive models influenced by the following wavelengths: 895 nm, 911 nm, around 2100 nm and in the 2350-2440 nm range. The only two NIR wavelengths highlighted in Approach I may be related to absorption by humic compounds, particularly the C-H functional groups present in the cellulose and lignin of plant residue on the soil surface (Alsaleh et al., 2025; Reyes-Rojas et al., 2025). This behavior can be attributed to the presence of plant material that had not fully decomposed in some of the sampled areas, and whose spectral signature was detected easily by the field spectroradiometer, which captured more-localized information.

Wavelengths around 2300 nm may be influenced by the stretching of carboxylic groups and C=O bonds, which are often associated with the structure of organic matter adsorbed to clay minerals in the soil (Ribeiro et al., 2021; Alsaleh et al., 2025). Stenberg et al. (2010) also associated this region with the interaction of energy and carbohydrates.

Although both sensors generated reliable predictive models, the EnMAP and *in situ* FieldSpec data showed more relevant prediction errors for SOC values < 9 g kg^{-1} (Figure 18). The results indicate that the sensors are more reliable in the higher SOC ranges, while higher errors occur at lower concentrations. This characteristic introduces a considerable challenge for drylands, where the levels of soil organic carbon are naturally low. The levels of SOC in agricultural areas in arid and semi-arid regions are between 3 and 12 g kg^{-1}, while dry sub-humid regions or those needing supplemental irrigation can reach up to 20 g kg^{-1} (Zhang et al., 2015; Raheb et al., 2017; Guoju et al., 2019), as observed in the samples collected across the irrigated perimeter.

At low SOC levels, the variability of spectral data tends to increase, resulting in higher predictive uncertainty. Studies by Lesaignoux et al. (2012) confirm the difficulty of predicting SOC in soils with a low carbon content, since reflectance can be influenced more by moisture and texture than by the presence of organic matter, making the estimate less accurate. On the other hand, a higher SOC content can affect the spectral response in a more consistent and more precise way, by helping to form structures in the soil and enabling mathematical models to better capture the interactions of these structures with the radiation (Herranz-Luque et al., 2024).

When using higher SOC ranges (from 11.6 to 23.6 g kg^{-1}), the predictive capacity of EnMAP improved, with Adj.R^2 = 0.91, RMSE = 0.61 g kg^{-1}, MAE = 0.56 g kg^{-1} and RPD = 6.03 (Figure 19a). This reduction of the dataset led to the selection of nineteen significant

spectral bands, highlighting the wavelengths between 1211 and 1258 nm and the region around 2265 nm and 2282 nm (Figure 20a). The persistence of these bands in the model underlines their relevance in predicting SOC at higher levels, and their importance for modelling.

The presence of the bands at 1200 nm can be explained by the functional groups found in the humic substances of the organic compounds. Furthermore, absorption peaks near 2260 nm are related to the strong association of organic particles with clay minerals. This spectral region is strongly influenced by the energy absorption of minerals such as kaolinite and 2:1 clay (Ribeiro et al. 2021; Gholizadeh et al., 2022). In studies by Laamrani et al (2019), bands at 1230 nm and around 2200 nm were considered influential for SOC detection using reflectance spectroscopy.

The reduction in the SOC dataset also resulted in improved FieldSpec performance in predicting SOC levels between 9.2 and 11.6 g kg^{-1} and between 16.2 and 23.2 g kg^{-1} (Adj.R^2. = 0.83, RMSE = 1.89 g kg^{-1}, MAE = 1.41 g kg^{-1} and RPD = 2.66). Observations with SOC levels of 11.6 to 16.2 g kg-¹ showed average prediction errors of more than 12%, which is why this range was not used to build new models. This result may be related to the higher variability in this range and the reduced number of available samples, which affected the accuracy of the spectral data acquired by the proximal sensor.

FieldSpec has a high spectral resolution, which allows subtle variations in spectral properties to be detected in environments with some degree of heterogeneity (Laamrani et al., 2019). If these variations are due to noise, there may be interference between the spectral response and the chemical variation of the sample. According to Wang et al. (2020), direct measurements in the field can also be affected by soil moisture, which reduces the accuracy of predicting soil properties.

The higher spectral variability of the FieldSpec sensor is confirmed when its spectral information is compared with that of the EnMAP sensor. It can be observed that the FieldSpec sensor captured more heterogeneous spectral information, even when captured in the area corresponding to the pixel from the orbital sensor.

To estimate SOC values in the 9.2 to 11.6 g kg^{-1} and 16.2 to 23.2 g kg^{-1} ranges, the model with the FieldSpec data reduced the predictor variables to 14 spectral bands, maintaining the most influential: 895 nm, 911 nm, 1534 nm, 1557 nm, and around 2400 nm. Spectral bands found in the 2400 nm range can be associated with C-H combinations in the structure of organic compounds in the soil, while absorption peaks at 1550 nm are often associated with energy

absorption by carboxylic groups in combination with amine structures (Stuart, 2004; Stenberg et al., 2010).

6.4.4 Prediction of Total Nitrogen (TN)

The TN prediction in Approach I (0.197-0.880 g kg^{-1}) showed a similar performance for EnMAP and FieldSpec, with a slight advantage for the *in situ* reflectance data, which resulted in higher accuracy, with Adj.R^2 = 0.68, RMSE = 0.079 g kg^{-1}, MAE = 0.063 g kg^{-1} and RPD = 1.89. The result can be attributed to the absence of atmospheric interference and the higher spectral resolution, allowing better detection of the bands related to TN. Despite showing slightly poorer performance, EnMAP obtained Adj.R^2 = 0.61, RMSE = 0.08 g kg^{-1}, MAE = 0.067 g kg^{-1} and RPD = 1.74, meaning that both sensors were similarly effective in estimating total nitrogen at the soil surface.

The developed models proved being reliable in predicting TN between 0.197 and 0.880 g kg^{-1}, using both *in situ* and satellite reflectance data (EnMAP). However, the resulting RPD metrics (1.89 and 1.74, respectively) show that, although the models are usable in practice, the data processing strategies need improvement (Bellon-Maurel et al., 2010).

Predictive models for TN, developed by Golizadeh et al. (2022) using FieldSpec spectral data in the laboratory, showed similar performance to that obtained in the present study. The researchers mentioned reported an R^2 between 0.45 and 0.70, with a minimum RMSE of 0.07 g kg^{-1} in soil samples collected at a depth between 2 and 20 cm.

The spectral bands selected in the models reveal the influence by specific regions of the electromagnetic spectrum when predicting TN using Approach I. In the case of EnMAP, the most important bands are centered around 936 nm and 941 nm, and between 1128 and 1151 nm, which are often associated with organic matter and nitrogen compounds in the soil (Stuart et al., 2004; Stenberg et al., 2010).

According to studies conducted by Tahmasbian et al. (2018) using a spectral camera with 462 spectral bands, wavelengths in the 900 nm range are also significant for predicting TN. Gholizadeh et al. (2022) reported that bands around 1150 nm are affected by the absorption of energy by N-H amine molecules.

The FieldSpec-based prediction model accounted contributions from bands at 577 nm and 593 nm, associated with the molecular absorption of chlorophyll by plant materials that had not yet fully decomposed (Sukhova et al., 2019). As in the present research, Misbah et al.

(2024) identified the strong influence of spectral bands between 500-600 nm on TN prediction in different soils, with R^2 = 0.84, RMSE = 0.082 and RPD = 2.53.

Still in respect of the *in situ* FieldSpec data (Figure 22b), bands between 1878 and 1988 nm were influential in predicting TN, as they are usually associated with the vibration of OH groups present in the structure of clay minerals, which may be linked to adsorbed organic materials (Demattê et al., 2019). Furthermore, according to finding reported by Dalal and Henry (1986) and Stenberg et al. (2010), the intervals around 1870 nm are important excitation peaks for molecules with NH and ≡N bonds.

The robustness of the FieldSpec model, a result of the absence of atmospheric interference, is due to the contact probe, which isolates the environment between the target and the sensor, allowing the spectral signatures of the soil to be more accurately captured. However, the performance of EnMAP shows that, as a hyperspectral satellite sensor, it also has potential for regional applications in predicting TN, and can be used for large-scale monitoring, if properly used.

The performance of the predictions across different TN ranges shows that the models are limited in their ability to estimate low concentrations in the soil. The models developed with EnMAP and the *in situ* FieldSpec data are most accurate for values higher than 0.550 g kg^{-1}. Authors such as Dinakaran et al. (2016) and Brunet et al. (2007) note that preparing the soil sample is important for maintaining the homogeneity of the accuracy metrics. Thus, the reduction in accuracy in predicting low TN levels may be related to the heterogeneity of the evaluation medium, given that the spectral reading was conducted directly in the field with the samples that were not pre-processed.

Similar to SOC prediction, estimating low levels of total nitrogen represents a challenge for predictive models, due to the influence of external factors such as moisture, clay content and shading by micro-terrain on the variations in reflectance data (Zhu et al., 2021).

Reconstructing the models for the most predictive intervals resulted in a significant improvement in the performance of the spectral data for estimating TN. The results show improved predictive performance between 0.550-0.880 g kg^{-1}, with Adj.R^2 = 0.82, RMSE = 0.035, MAE = 0.023 and RPD = 2.43 for the EnMAP data and Adj.R^2. = 0.85, RMSE = 0.032, MAE = 0.025 and RPD = 2.64 with the *in situ* FieldSpec data (Figure 24b).

The levels of performance found in the present research corroborate the findings of Tahmasbian et al. (2018), who developed TN predictive models from hyperspectral data using PLSR, with R^2 = 0.86 and RPD = 2.08. Similarly, Reda et al. (2019) applied complex models to estimate the TN content in different soil types, obtaining R^2 = 0.75, RMSE = 1.10 g kg^{-1} and RPD = 2.55 when using reflectance data.

Misbah et al. (2024) developed PLSR models to estimate TN levels > 0.5 g kg^{-1} from PRISMA images, obtaining R^2 = 0.7, RMSE = 0.14 and RPD = 1.48. Dinakaran et al. (2016) estimated TN in dry and sieved soil samples using the FOSS NIRS laboratory spectrometer, and achieved predictions with R^2 = 0.75, RMSE = 0.01 and RPD = 6.98, where the samples analyzed by the authors had levels higher than 1.2 g kg^{-1}. Tahmasbian et al. (2018) confirmed these findings in the laboratory, using images captured by cameras with 462 spectral bands, developing predictive TN models with R^2 = 0.79 and RMSE = 0.03 g kg-1 for levels greater than 0.7 g kg^{-1}.

In the final model developed with data from the EnMAP sensor, wavelengths around 900 nm and 1100 nm were highlighted (Figure 16a). These wavelengths are associated with the excitation of molecules containing NH groups (Stenberg et al., 2010). In the model fitted with the *in situ* FieldSpec data, the visible region (577-614 nm) and the region between 1878 and 1988 nm were kept (Figure 16b); these regions were previously identified as strongly influenced by the vibrations of molecules containing NH and ≡N groups (Dalal; Henry, 1986).

It should be noted that studies using reflectance spectroscopy for estimating TN and SOC are being progressively conducted with improvements, both in the laboratory and using hyperspectral images. The present research highlights the significant results of using EnMAP images and the *in situ* FieldSpec as alternative and/or complementary technologies to more complex prediction methods.

6.5 CONCLUSIONS

This study evaluated the performance of multiple linear regression models for predicting soil organic carbon (SOC) and total nitrogen (TN) using spectral data from the EnMAP hyperspectral satellite and the ASD FieldSpec® proximal sensor *in situ*, across different bare soil areas in northeastern Brazil. Both sensors demonstrated good predictive performance for organic attributes, with MLR models proving as effective as PLSR models used in other studies.

The results revealed that EnMAP has significant potential for large-scale monitoring of SOC and TN. The combination of ideal conditions during the study – low cloud cover, bare soil, sampling and imagery only a few days apart, and an adequate number of samples – ensured the excellent performance of the satellite-based predictive models. These favorable conditions ensured accuracy comparable to *in situ* ASD FieldSpec®.

Furthermore, the research highlighted that calibrating the models for specific SOC and TN ranges can further improve accuracy, especially for lower levels of these attributes. In this way, EnMAP proves to be a practical and efficient alternative for large-scale monitoring, eliminating the need for transporting large and valuable equipment to the field, reducing operational costs, and facilitating data collection in remote areas.

7 PREDICTIONS OF SOIL ORGANIC CARBON AND TOTAL NITROGEN IN AGRICULTURAL CONTEXTS USING REFLECTANCE SPECTROSCOPY: A PREDICTIVE MODELING APPROACH

ABSTRACT

Soil organic carbon (SOC) and total nitrogen (TN) are fundamental for soil fertility and agricultural productivity. While traditional methods are common, remote sensing provides a low-impact, accessible alternative. However, soil types respond differently to remote sensing, requiring specific approaches. This research aimed to develop predictive models for SOC and TN, quantified using distinct methodologies and under different land uses, by spectral information (350-2500 nm). Soil samples were collected from the Lower Acaraú Irrigation Perimeter (Brazil) – R1 – and California's Central Valley (USA) – R2. In R1, SOC and TN were quantified by titration, while in R2, the elemental analyzer was used. Local datasets (R1 and R2) and the generalized dataset (R1&R2) were spectrally analyzed with spectroradiometer in laboratory , and MLR and PLSR regression models were developed with spectral variables selected by Stepwise and Backward methods. Reflectance, first derivative and continuum-removed data were used. The first derivative proved efficient in developing high-accuracy models across all trials. The best performance was observed for R2 samples for SOC (Adj. R^2 = 0.90 and RPD = 3.53), and TN (Adj. R^2 = 0.91 and RPD = 3.67) in MLR, while in PLSR, the values were Adj. R^2 = 0.91 and 0.92, and RPD = 3.34 and 3.49 for SOC and TN, respectively. Also, the generalized approach (R1&R2 dataset) was effective in predicting these attributes using the first derivative of reflectance, regardless of the quantification methodology and soil type, reflecting its broad applicability in different agricultural contexts, enabling consistent and scalable predictions under various edaphoclimatic conditions.

Keywords: environment monitoring; hyperspectral remote sensing; titration; spectral behavior

RESUMO

O carbono orgânico do solo (SOC) e o nitrogênio total (TN) são fundamentais para a fertilidade e produtividade agrícola. Enquanto métodos tradicionais são comumente usados, o sensoriamento remoto oferece uma alternativa eficiente, acessível e de baixo impacto ambiental. No entanto, diferentes tipos de solo respondem de maneira distinta ao sensoriamento remoto. Esta pesquisa objetivou desenvolver modelos preditivos de SOC e TN, quantificados por metodologias distintas e sob diferentes usos da terra, utilizando informações espectrais (350-2500 nm). As amostras de solo foram coletadas no Perímetro Irrigado do Baixo Acaraú (Brasil) – R1 – e no Vale Central da Califórnia (EUA) – R2. Em R1, SOC e TN foram quantificados por titulação, enquanto em R2, foi utilizado o analisador elementar. Os conjuntos de dados locais (R1 e R2) e o conjunto global (R1&R2) foram analisados espectralmente com o ASD FieldSpec®, e modelos de regressão MLR e PLSR foram desenvolvidos com variáveis espectrais selecionadas por *Stepwise* e *Backward.* Foram utilizados dados de reflectância, derivada e contínuo removido. A primeira derivada mostrou-se eficiente para desenvolver modelos de elevada acurácia em todos os ensaios. O melhor desempenho se deu para amostras de R2 para SOC (R^2_{adj} = 0.90 e RPD = 3.53), e para TN (R^2_{adj} = 0.91 e RPD = 3.67) na MLR, enquanto na PLSR, os valores foram R^2_{adj} = 0.91 e 0.92, e RPD = 3.34 e 3.49 para SOC e TN, respectivamente. A abordagem generalista também foi eficaz na predição desses atributos usando a primeira derivada da reflectância, independente da metodologia de quantificação e tipo de solo, refletindo sua ampla aplicação em diferentes contextos agrícolas, permitindo predições consistentes e escaláveis sob diversas condições edafoclimáticas.

Palavras-chave: monitoramento ambiental; sensoriamento remoto hiperespectral; titulação; comportamento espectral

7.1 INTRODUCTION

Soil organic carbon (SOC) and total nitrogen (TN) play essential roles in agriculture, being crucial for soil quality and crop productivity (Schjønning, 2023; Jiang et al., 2024). Accurate analyses are essential for sustainable management and ecosystem conservation. Among the traditional techniques for determining SOC and TN in laboratory, the following are highlighted: *i*) wet oxidation by Walkley and Black (1934) with Yeomans and Bremner (1988) improvements using digestion block and titration; *ii*) the Elemental Analyzer, with sample combustion at temperatures above 1000ºC, quantifying the CO_2 and NO_2 generated by combustion chamber (Farina et al., 1991).

With the increasing demand for fast, efficient, and cost-effective methodologies, there is a growing need for techniques that require fewer resources for maintenance and labor (Amoli et al., 2024). Remote sensing stands out, as it allows for low destructive analysis of soil attributes, providing information about its composition based on the interaction of electromagnetic radiation with the soil's intrinsic components (Ribeiro et al., 2021; Alsaleh et al., 2025). Spectral data obtained from various wavelengths can provide valuable information about carbon and nitrogen levels through molecular interactions of groups such as CH, NH, and OH, in a non-invasive manner (Stenberg et al., 2010; Cambou et al., 2022).

The application of mathematical techniques, such as Multiple Linear Regression (MLR) and Partial Least Squares Regression (PLSR), has been essential in predicting soil attributes from spectral data (Angelopoulou et al., 2020; Ribeiro et al., 2021). MLR represents the relationship between several independent variables and the predicted variable. When multicollinearity among spectral variables is high, the PLSR technique becomes particularly useful, combining principal component analysis with linear regression, maximizing the covariance between independent and dependent variables. Spectral variable selection has also proven to be an effective approach for reducing dimensionality and improving the efficiency of predictive models (Reda et al., 2019; Li et al., 2022).

Since different environments and land uses have exclusive soil characteristics, it is necessary to adjust prediction models to the most similar local conditions (Mishra et al., 2020). Both local and generalized models have been developed to estimate soil attributes, based on spectral data collected from different environmental contexts. Generalized models allow for large-scale analysis, applicable across different regions, while local models provide more

detailed and specific predictions for areas with specific characteristics (Ye et al., 2021; Zhou et al., 2021).

The knowledge and application of different methodologies for determining soil organic attributes are essential for agricultural sciences (Amoli et al., 2024; Alsaleh et al., 2025). Advances in techniques, whether through traditional methods or new approaches based on remote sensing and mathematical modeling, contribute to the accuracy of soil characterization, promoting more sustainable and efficient agricultural practices (Pudelko; Chodak, 2020). This is crucial for the conservation of natural resources and combating soil degradation under various agricultural land uses.

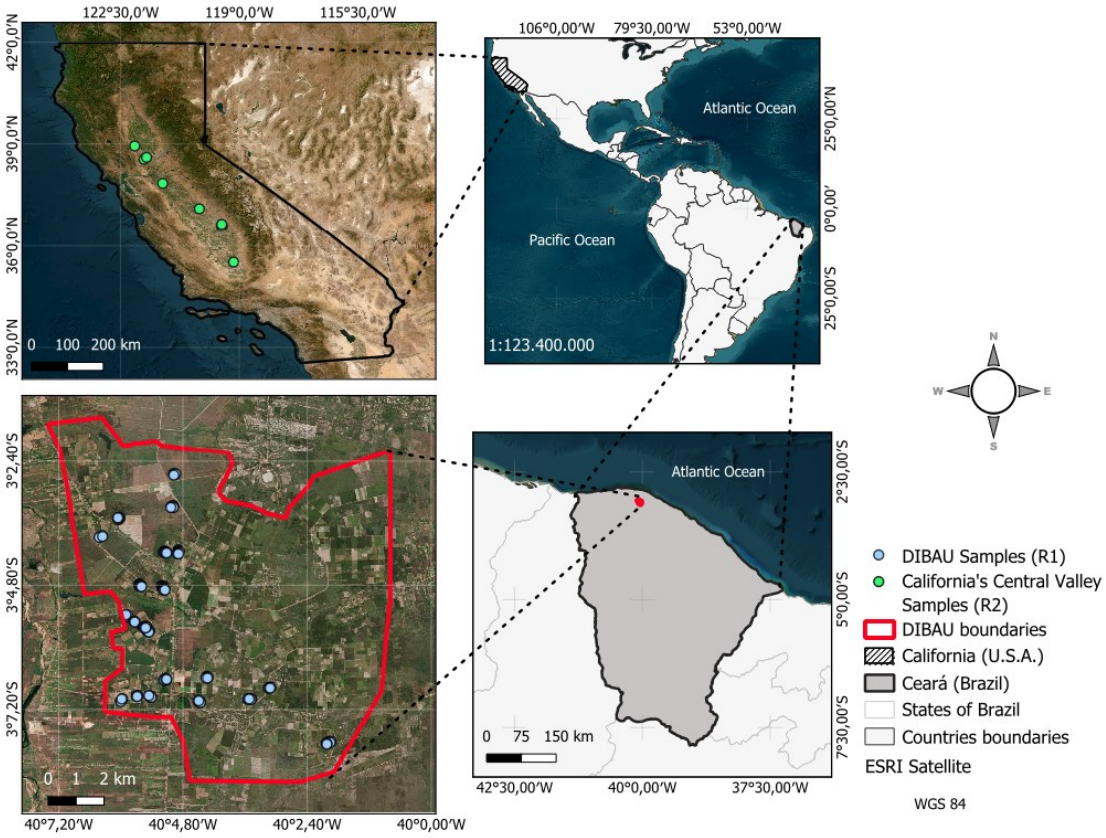
The hypothesis of this study is that predictive models using reflectance spectroscopy offer accurate estimates of soil organic carbon and total nitrogen across various land uses and geographical areas, being efficiently validated, regardless of the soil sampling method or the attribute quantification technique used. Therefore, this research aimed to develop predictive models for SOC and TN in irrigated areas in Brazil and the United States, with different land uses and quantified by distinct methodologies, using reflectance spectroscopy through local and generalized models. Additionally, the study sought to identify the most significant wavelengths for detecting these elements in soil samples.

7.2 MATERIAL AND METHODS

7.2.1 Study Areas

This research was conducted through the sampling of agricultural soils from two distinct regions: the Lower Acaraú Irrigation District – DIBAU (R1), located in the north of State of Ceará, Brazil, with central coordinates at 03°05’11.04’’ S and 40°03’46.08’’ W, and the California’s Central Valley – CCV (R2), U.S.A., centered at 37°19’12.00” N and 119°41’24.00’’ W (Figure 26).

Figure 26. Sampling locations map



The Lower Acaraú Irrigation District spans approximately 11,000 ha and has a tropical wet climate (Aw), according to the Köppen classification. The region experiences a rainy season in the summer and a dry season in the winter, with an average annual temperature of 28.1°C and soils ranging from loam to sandy (ADECE, 2011).

California's Central Valley covers about 155,000 km² and features two distinct climates according to the Köppen classification. To the north, the climate is classified as Csa, characterized by hot, dry summers and cold, wet winters, typical of Mediterranean regions. To the south of the valley, the climate is Bsk, a semi-arid steppe climate with hot to extremely hot and dry summers, and moderate, wet winters (Kauffman, 2003). The soil in this region is naturally saline, with a texture ranging from loam to clay-loam (Jin et al., 2018).

7.2.2 Sampling Logistics and Chemical Analysis

A total of 257 soil samples were collected from irrigated areas with no vegetation cover, with 130 samples from DIBAU (R1) and 127 samples from CCV (R2). The samples were taken from areas with initial stage fruit crops (watermelon, coconut, cashew, and acerola) or

being prepared for new crops in R1, and from almond orchards in R2. The soil sampling methods and determination of SOC and TN are detailed in Table 6.

Table 6. Soil sampling overview and analytical methods used for Soil Organic Carbon (SOC) and Total Nitrogen (TN) determination in DIBAU (R1) and CCV (R2).

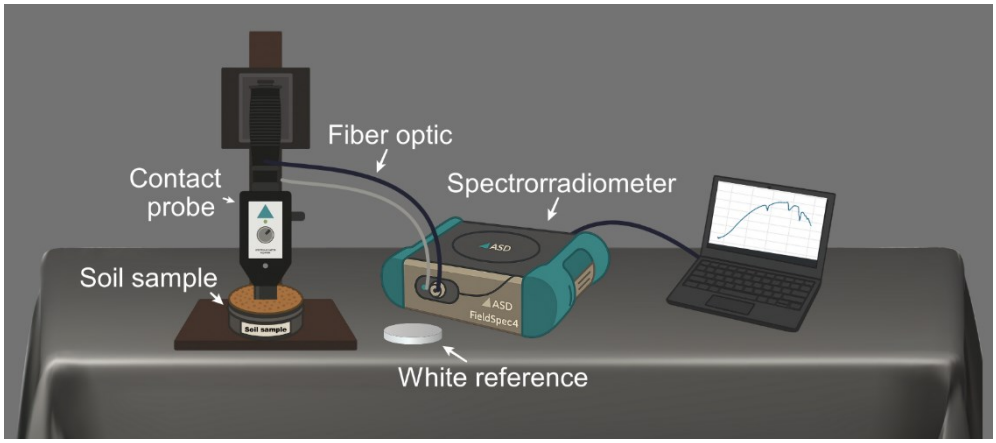
Area of Interest	Samples	Sampling Depth	Land Use	SOC Determination (%)	TN Determination (g/kg)
R1	130	0-1 cm	Prepared for planting and fruit crops in early development	Titration (Walkley; Black, 1934; Yeomans; Bremner, 1988)	Titration (Kjeldahl, 1883)
R2	127	0-10 cm	Established almond orchads	Elemental Analyzer LECO® (Skjemstad; Baldock, 2011)	Elemental Analyzer LECO® (Skjemstad; Baldock, 2011)

The collected samples were air-dried, crushed, sieved through a 2 mm mesh and oven-dried (45º C / 48 h) for spectral analysis, and subsequently ground for chemical analysis. The samples from R1 were analyzed for SOC and TN using traditional titration methods, while the samples from R2 were quantified using an elemental analyzer for organic carbon and nitrogen.

7.2.3 Spectral analysis

T The sample sets were oven-dried at 45°C until reaching constant weight for moisture standardization, and then spectrally analyzed using the ASD FieldSpec® spectroradiometer (350-2500 nm). The R1 samples were analyzed in the darkroom of the Geoprocessing Laboratory at the Center for Agricultural Sciences, Federal University of Ceará, Fortaleza, Brazil. The R2 samples were analyzed in the darkroom of the Department of Land, Air, and Water Resources at the University of California, Davis, USA. For both sets, the spectral data acquisition geometry followed the procedure described in Figure 27.

Figure 27. Geometry of spectral data acquisition for soil samples in a darkroom.



The analyses were performed with the ASD Hi-Brite contact probe at three points on the surface of the samples, which were placed in black polypropylene containers, with ten automatic readings per point (30 spectral readings/sample). Spectral standardization was carried out using a Lambertian white plate, with reflectance close to 100%.

7.2.4 Spectral transformations

To develop the predictive models for SOC and TN in the soil samples, the spectral data obtained in the darkroom were used in three forms: *i*) raw reflectance values (non-transformed); *ii*) reflectance with continuum removal; and *iii*) first derivative of reflectance. The spectral data transformations were performed in RStudio (R Core Team, 2024), using the "prospectr" package (Stevens; Ramirez-Lopez, 2024).

Before the transformations, the spectral data were smoothed via the Savitzky-Golay algorithm, with a 3-window interval (Equation 5), to reduce random noise and prevent distortions in the spectrum, while preserving the shape of the spectrum (Savitzky; Golay, 1964).

$$y_j^* = \frac{1}{N} \sum_{h=-k}^{k} C_h \, y_{j+h} \tag{5}$$

Where: y_j^* is the new smoothed value; C_h represents the smoothing filter coefficients; N is the smoothing window size (N = 3); k is the number of neighboring values on each side of j.

After reflectance smoothing, the continuum removal process was applied to eliminate continuous spectral features and highlight the absorption features of the targets (Deo et al., 2024). The use of the first-order spectral derivative, in turn, aimed to enhance the oscillations detected in the smoothed spectrum from 350-2500 nm, between consecutive wavelengths, serving as an important technique for improving the resolution of overlapping spectra and removing noise that might be interpreted as signals (O'Haver, 1979).

The first derivative spectral transformation ($d\rho\lambda$) results from the variation in reflectance values ($\rho\lambda$) as a function of wavelength (x) at a specific point (i), numerically approximated from a central point between successive spectral bands according Equation 6, described by Rudorff et al. (2007).

$$\frac{d\rho_\lambda}{dx} \cong \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta x} \quad (6)$$

Where: Δx represents the distance between two successive bands ($\Delta x = x_{i+1} - x_{i-1}$); ρ_{i+1} refers to the reflectance factor of the point following *i*, and ρ_{i-1} corresponds to the reflectance factor of the point preceding *i*.

7.2.5 Selection of Significant Variables and Regression Models

The selection of spectral variables was performed for two regional datasets (samples from R1 and R2 separately) and one generalized dataset (R1&R2), with the goal of identifying the wavelengths that best explain the variation in SOC and TN levels for different contexts. To achieve this, the Stepwise, Backward, and Forward algorithms were applied in IBM SPSS Statistics (IBM Corp., 2012), ensuring more robust models with improved dimensionality.

Predictive models for SOC and TN were developed for all three datasets, utilizing the spectral variables most influential in the variation of SOC and TN, with Multiple Linear Regression (MLR) and Partial Least Squares Regression (PLSR). MLR was applied to assess the efficiency of simpler algorithms in predicting SOC and TN from spectral information across different contexts. PLSR, in turn, was used as a robust method, maximizing the correlation between spectral variables and chemical variables, transforming them into latent variables, and optimizing the variability within the predictor set.

All models were validated using leave-one-group-out validation, where each test group corresponded to samples collected spatially close to one another. The accuracy of model validation was measured using metrics such as: Adj. R^2, Root Mean Square Error (RMSE), Ratio of Performance to Deviation (RPD), and Mean Absolute Error (MAE).

7.3 RESULTS

7.3.1 Distribution of SOC and TN in the evaluated areas

From the descriptive statistics, it was observed that the samples collected from R1 showed higher average SOC values (1.102%) compared to those from R2 (0.889%), while the opposite was true for TN levels, where the samples from R1 exhibited lower averages (0.508 g/kg) compared to those from the Central Valley (0.891 g/kg). The combined data (R1&R2) presented an average SOC of 1.004% and TN of 0.684 g/kg (Table 7).

Table 7. Descriptive statistics parameters for content of Soil Organic Carbon (SOC) and Total Nitrogen (TN) for the R1, R2, and R1&R2 datasets

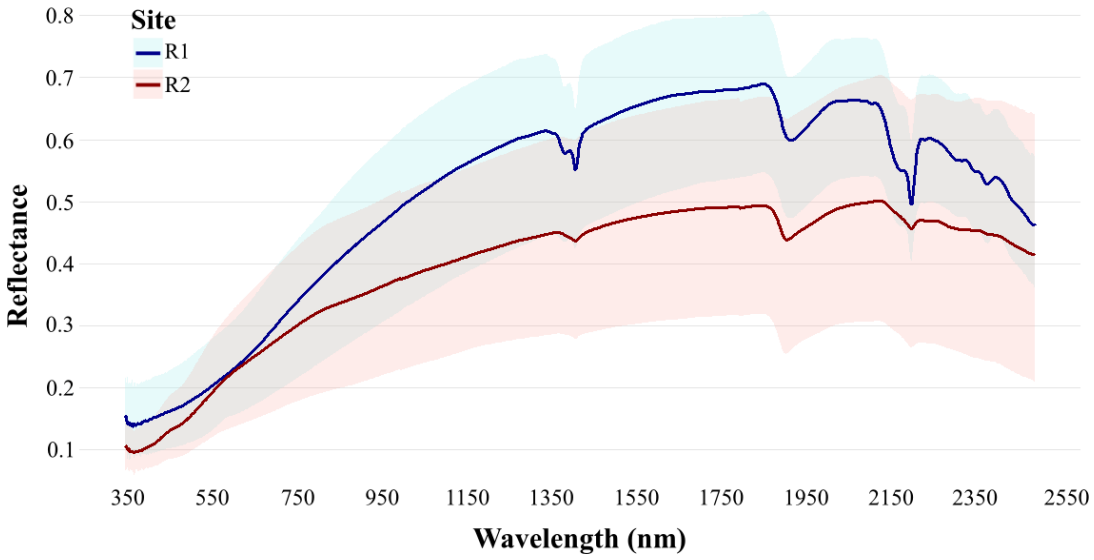
Statistics parameters	R1		R2		R1&R2	
	% SOC	TN g/kg	% SOC	TN g/kg	% SOC	TN g/kg
Mean	1.102	0.508	0.889	0.891	1.004	0.684
Median	1.038	0.460	0.818	0.828	0.881	0.640
Standard deviation	0.496	0.200	0.409	0.370	0.469	0.347
CV (%)	45.00	39.45	46.00	41.48	46.74	50.79
Minimum	0.267	0.197	0.320	0.230	0.267	0.197
Maximum	2.504	1.142	2.320	2.130	2.504	2.130

CV (%): Coefficient of variation

7.3.2 Spectral analysis and reflectance transformations

Through the analysis of the average spectral behavior of surface soil samples from R1 and R2 (Figure 28), it is possible to observe that the samples collected in the California's Central Valley presented a lower albedo compared to those from R1 for the entire spectrum.

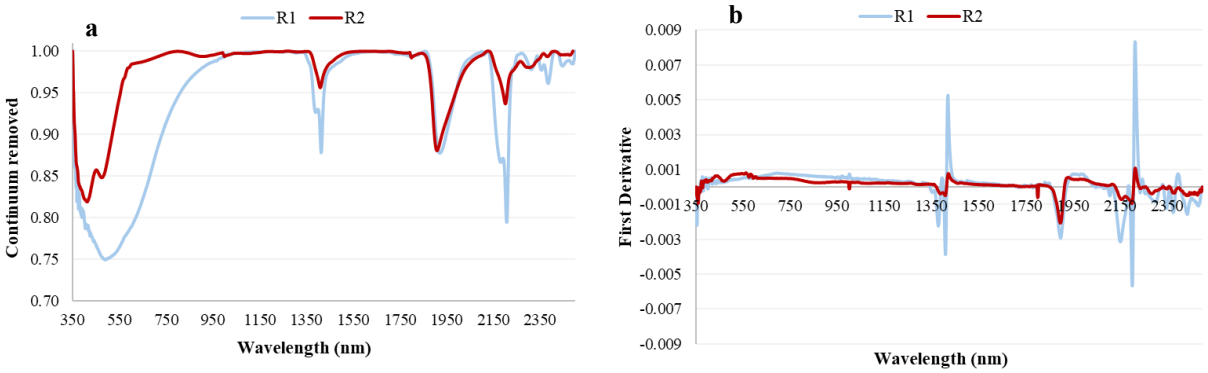
Figure 28. Average spectral behavior (continuous line) and reflectance variations (shaded area) for the samples collected in R1 and R2.



The spectral behavior of the samples collected in R1 mainly shows prominent absorption features from clay and structural moisture around 1400 nm, 1900 nm, and 2200 nm. This is less evident in the samples from R2, which show absorption features characteristic of iron oxides, although subtle, between 500 and 900 nm.

The reflectance variations in the features at 1400 nm, 1900 nm, and 2200 nm were also clearly highlighted by transforming the spectral information to a continuum-removed spectrum (Figure 29a) and to the first derivative (Figure 29b).

Figure 29. Average spectral behavior with reflectance transformed into a) continuum-removed spectrum and b) first derivative, for samples collected from R1 and R2.



7.3.3 Predictive Models

7.3.3.1 Multiple Linear Regression

Reflectance spectral data and their transformations through continuum removal and first derivative demonstrated reliable predictive performance across all datasets, with RPD > 1.59 and minimum Adj. R^2 values of 0.60 and 0.80 in leave-one-group-out validation for SOC and TN prediction, respectively. Both models employed untransformed reflectance as the predictor variable (Table 8).

Table 8. Performance metrics of MLR models for predicting Soil Organic Carbon (SOC) and Total Nitrogen (TN) using reflectance and transformations for R1, R2, and R1&R2 datasets using leave-one-group-out cross-validation.

Dataset	MLR	% SOC			TN (g/kg)		
		Reflectance	CR	1st der	Reflectance	CR	1st der
R1 (130)	Bands selected	23	22	25	23	24	26
	RMSE	0.246	0.245	0.213	0.082	0.082	0.064
	MAE	0.207	0.198	0.172	0.065	0.068	0.050
	RPD	1.67	1.87	2.23	2.23	2.30	2.88
	Adj. R^2	0.64	0.71	0.80	0.80	0.81	0.88
R2 (127)	Bands selected	19	23	21	15	24	21
	RMSE	0.129	0.110	0.106	0.108	0.102	0.100
	MAE	0.106	0.091	0.085	0.091	0.082	0.081
	RPD	2.80	3.47	3.53	2.79	3.62	3.67
	Adj. R^2	0.85	0.90	0.90	0.85	0.91	0.91
R1&R2 (257)	Bands selected	16	24	43	23	37	43
	RMSE	0.239	0.211	0.142	0.123	0.110	0.101
	MAE	0.195	0.172	0.115	0.102	0.084	0.078
	RPD	1.59	1.96	3.12	2.46	3.15	3.44
	Adj. R^2	0.60	0.73	0.89	0.83	0.89	0.90

CR: Continuum-removed; 1st der: First derivative

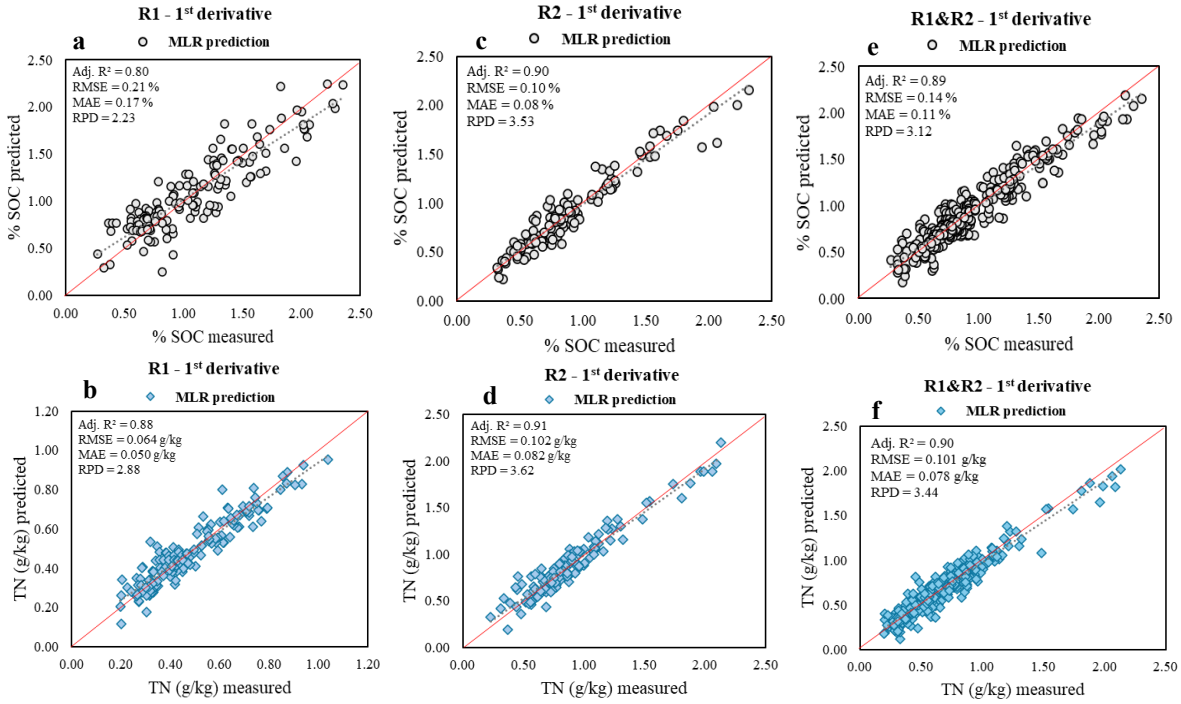
For the R1 sample set, the prediction of % SOC was more robust when using the values of first derivative of reflectance as the predictor variables, with Adj. R^2 = 0.81, RMSE =

0.213, MAE = 0.172, and RPD = 2.23 (Figure 30a). The prediction of TN (g/kg) showed efficient performance with the first derivative of reflectance data, yielding Adj. R^2 = 0.88, RMSE = 0.064, MAE = 0.050, and RPD = 2.88 (Figure 30b).

For the R2 dataset, the model using the first derivative showed RPD = 3.53, Adj. R^2 = 0.90, MAE = 0.085, and RMSE = 0.106 (Figure 30c). The model also predicted TN levels more accurately when using the first derivative of reflectance, with RMSE = 0.100, MAE = 0.081, RPD = 3.67, and Adj. R^2 = 0.91 (Figure 30d).

In the global model R1&R2, the best predictive performance was also achieved using spectral data transformed into the first derivative. For the % SOC model (Figure 30e), the selected wavelengths showed satisfactory predictive performance (Adj. R^2 = 0.89, RPD = 3.12, RMSE = 0.142, and MAE = 0.115). The same was observed for TN prediction, where the model that performed best showed predictive performance with Adj. R^2 = 0.90, RPD = 3.44, RMSE = 0.101, and MAE = 0.078, using the first derivative of reflectance (Figure 30f). It is worth noting that the validation parameters for all the models tested are provided in Appendix E.

Figure 30. Validation of the best MLR models developed for predicting Soil Organic Carbon (SOC) and Total Nitrogen (TN) for the R1, R2, and R1&R2 datasets.



7.3.3.2 Partial Least Squares Regression

Similar to MLR, all PLSR models demonstrated adequate performance in predicting SOC and TN, with most showing RPD > 2.0 and Adj. R^2 > 0.70 (Table 9). The lowest accuracy predictions were observed in models using untransformed reflectance values as predictor variables across all datasets.

Table 9. Performance metrics of PLSR models for predicting Soil Organic Carbon (SOC) and Total Nitrogen (TN) using reflectance and transformations for R1, R2, and R1&R2 datasets using leave-one-group-out cross-validation.

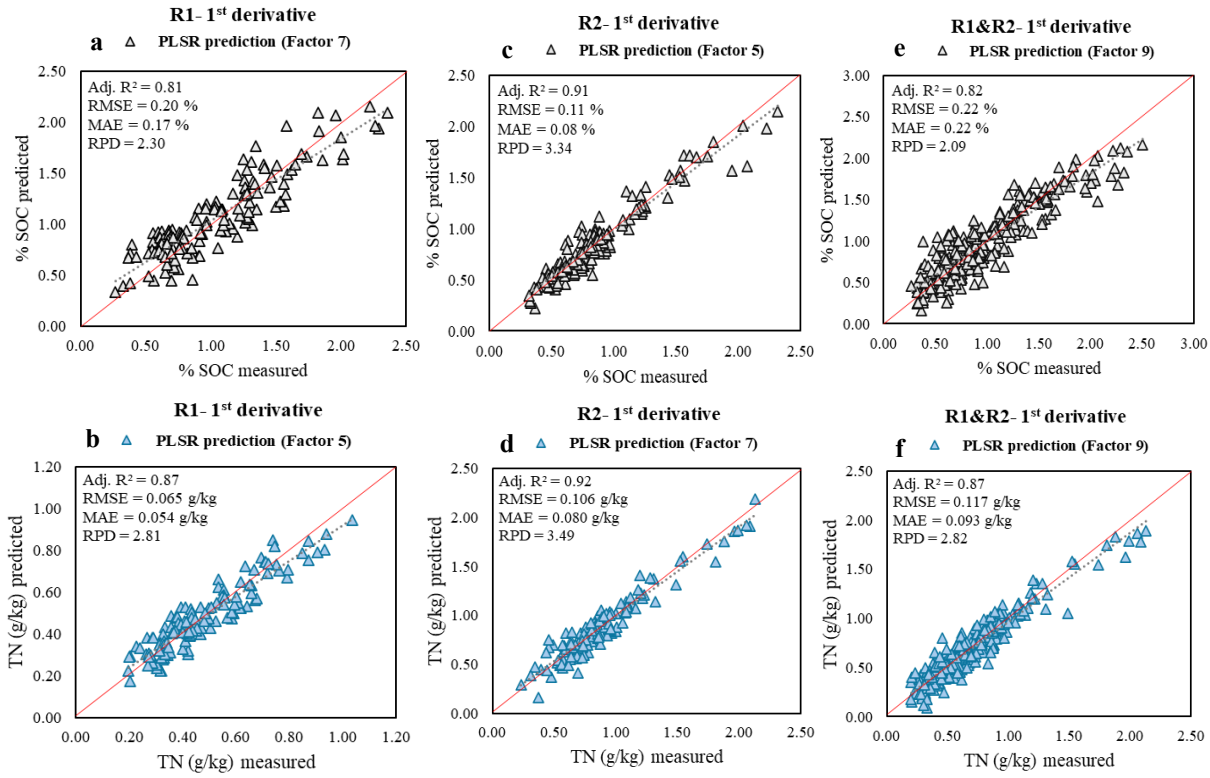
Dataset	PLSR	% SOC			TN (g/kg)		
		Reflectance	CR	1st der	Reflectance	CR	1st der
R1 (130)	Bands selected	23	22	25	23	24	26
	Nº of Factors	8	7	7	8	4	5
	RMSE	0.223	0.219	0.199	0.081	0.075	0.065
	MAE	0.188	0.178	0.168	0.069	0.064	0.054
	RPD	1.68	2.10	2.30	1.97	2.09	2.81
	Adj. R^2	0.66	0.77	0.81	0.74	0.79	0.87
R2 (127)	Bands selected	19	23	21	15	24	21
	Nº of Factors	5	8	5	4	10	7
	RMSE	0.166	0.138	0.114	0.149	0.119	0.106
	MAE	0.133	0.116	0.087	0.116	0.092	0.080
	RPD	2.16	2.77	3.34	2.14	3.02	3.49
	Adj. R^2	0.79	0.87	0.91	0.78	0.90	0.92
R1&R2 (255)	Bands selected	16	24	43	23	37	43
	Nº of Factors	10	13	9	9	13	9
	RMSE	0.261	0.229	0.225	0.128	0.117	0.104
	MAE	0.216	0.188	0.220	0.104	0.093	0.091
	RPD	1.51	1.82	2.09	2.36	2.82	2.86
	Adj. R^2	0.57	0.70	0.82	0.82	0.87	0.90

CR: Continuum-removed; 1st der: First derivative

For the R1 samples, the spectral transformation to the first derivative achieved the best predictive performance for both SOC and TN, using 7 and 5 factors, respectively (Figures 31a and 31b). The predictive model for % SOC showed the highest reliability, with Adj. R^2 = 0.81, RPD = 2.30, and errors of 0.199 (RMSE) and 0.168 (MAE). The TN (g/kg) prediction was even more accurate, with Adj. R^2 = 0.87, RPD = 2.81, RMSE = 0.065, and MAE = 0.054.

The best predictive model for SOC with the R2 samples showed Adj. R^2 = 0.92, RPD = 3.34, RMSE = 0.117, and MAE = 0.087 (Figure 31c), while the TN prediction, also with the first derivative, achieved Adj. R^2 = 0.92, RPD = 3.49, RMSE = 0.106, and MAE = 0.080 (Figure 31d).

Figure 31. Validation of the best PLSR models developed for predicting Soil Organic Carbon (SOC) and Total Nitrogen (TN) for the R1, R2, and R1&R2 datasets.



For the R1&R2 dataset, the SOC predictive model showed intermediate performance, with Adj. R^2 = 0.82, RPD = 2.09, RMSE = 0.225, and MAE = 0.220 using first derivative spectral data (Figure 31e). The TN prediction also showed better performance with the first derivative data, with intermediate metrics between the local models, such as Adj. R^2 = 0.90, RPD = 2.86, RMSE = 0.104, and MAE = 0.091 (Figure 31f).

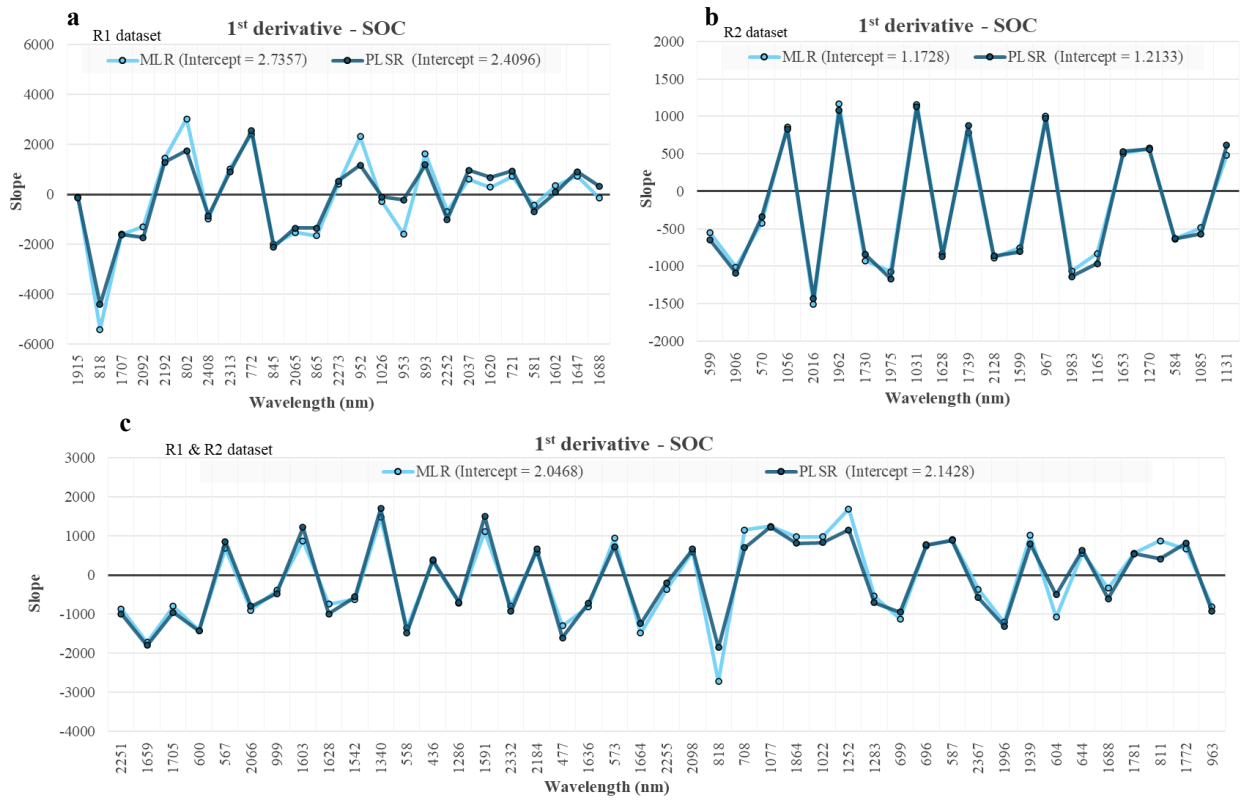
7.3.4 Influential Wavelengths for Predictive Models

The selection of spectral bands with potential influence on the variation of SOC and TN was performed prior to predictive modeling, with the nature of this influence being observed after model analysis.

In the best MLR model for % SOC, using the first derivative of R1 samples, bands at 772 nm, 802 nm, 893 nm, 952 nm, and 2192 nm stood out with strong positive correlations. Bands such as 818 nm, 1707 nm, 2408 nm, and 2092 nm showed negative slopes, suggesting inverse correlations (Figure 32a). In the PLSR model, the same set of bands for % SOC showed direct influence at 772 nm, 802 nm, and 2313 nm, while the 818 nm band exhibited a strong inverse correlation with SOC on the 7th factor.

In the regression models applied to the R2 samples, nine wavelengths stood out as the most influential, exhibiting high slopes with the same sign for both methods: 967 nm, 1031 nm, 1165 nm, 1906 nm, 1962 nm, 1975 nm, 1983 nm, and notably, 2016 nm (Figure 32b). For the global dataset R1&R2, the model also highlighted the 818 nm band, which showed the highest slope even after PLSR regularization (Figure 32c).

Figure 32. Comparison of regression coefficients (slope) for Multiple Linear Regression (MLR) and Partial Least Squares Regression (PLSR) models applied to the first derivative of reflectance for predicting Soil Organic Carbon (SOC) for the datasets (a) R1, (b) R2, and (c) R1 & R2. The intercept values are shown in the legends, indicating the model fits.

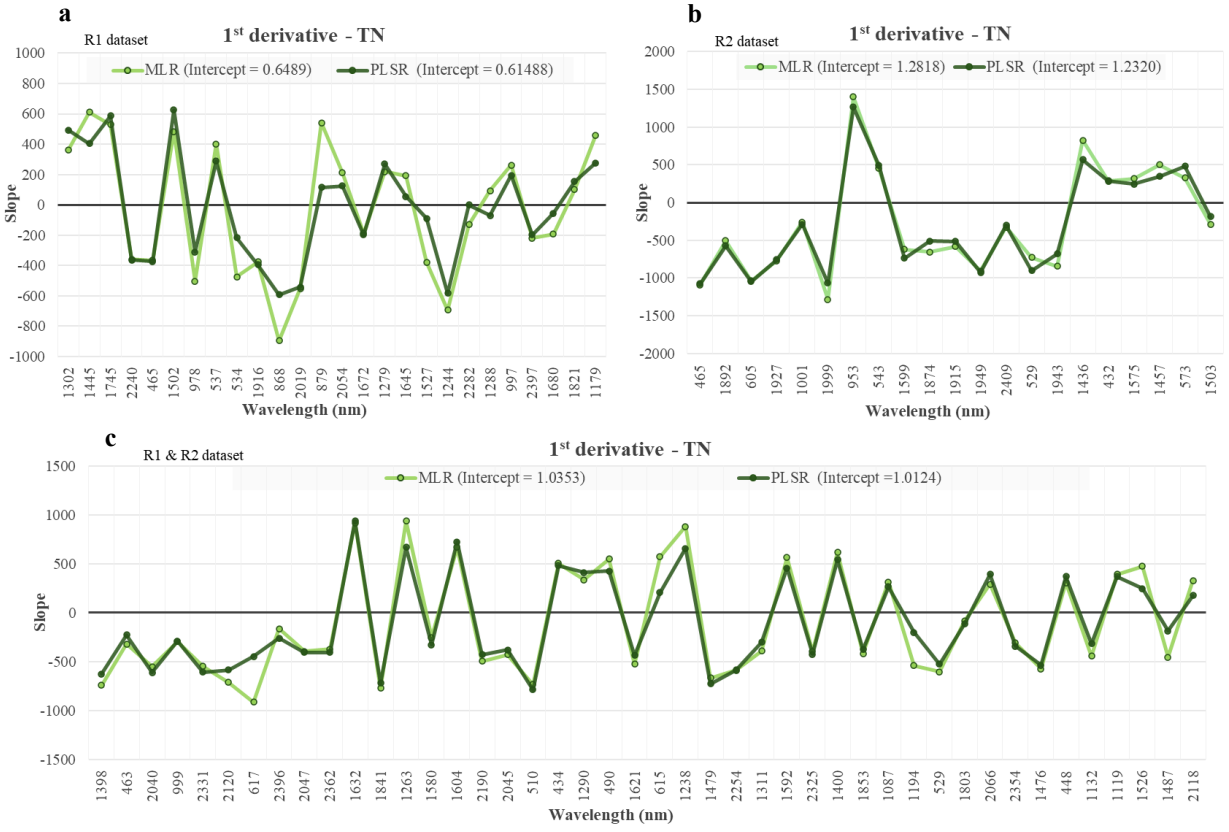


The TN prediction for R1 using MLR and PLSR, in turn, showed numerical discrepancies between the spectral band slopes (Figure 33a). In MLR, the first derivative of reflectance indicated influences at bands 534 nm, 868 nm, 879 nm, 978 nm, 1179 nm, 1244 nm, 1302 nm, 1445 nm, 1502 nm, 1745 nm, and 2019 nm. Bands such as 868 nm and 879 nm had high slopes in MLR but were smoothed in PLSR, which also balanced other extreme ones, improving model robustness and generalization capability.

Due to the similarity in coefficient behavior between MLR and PLSR with the R2 samples, it is possible to observe that the most influential first derivative wavelengths for prediction in both methods are: 465 nm, 529 nm, 543 nm, 605 nm, 953 nm, 1436 nm, 1949 nm, 1943 nm, and 1999 nm (Figure 33b).

In the global model (Figure 33c), the first derivative coefficients show greater consistency than in the isolated R1 models. The R1&R2 dataset showed intermediate performance, effectively representing the variation of the datasets, with the most influential and consistent bands for both methods being: 1440 nm, 1592 nm, 1604 nm, and 1632 nm – with a direct effect – and 510 nm, 1479 nm, 1841 nm, and 2254 nm – with an inverse effect.

Figure 33. Comparison of regression coefficients (slope) for Multiple Linear Regression (MLR) and Partial Least Squares Regression (PLSR) models applied to the first derivative of reflectance for predicting Total Nitrogen (TN) for the datasets (a) R1, (b) R2, and (c) R1&R2. The intercept values are shown in the legends, indicating the model fits.



7.4 DISCUSSION

7.4.1 Quantification of SOC and TN in the evaluated areas

Samples collected from the R1 region showed slightly higher average % SOC levels compared to those from R2, with mean values of 1.102% and 0.889%, respectively (Table 7). Based on granulometric analysis conducted prior to this study, R1 samples were classified as sand texture, while R2 samples were sandy loam, both with high sand content (Appendix D).

In R1 sampling areas, animal manure residues, grass and legume straw, and charcoal from the burning of previous crops predominated, whereas R2 samples mainly contained wood chips, residual shells from nut byproducts, and fumigation residues.

According to Wang et al. (2021) and Schjønning (2023), the residues found in R1 significantly contribute to SOC availability, particularly residues of straw and burning, due to their rapid decomposition under high-temperature irrigated conditions. This aligns with observations from this research, as temperatures during the sampling period ranged between 26°C and 33.5°C.

On the other hand, the wood chips and nut byproducts found in R2, which contain high levels of structural carbon decompose and mineralize more slowly due to their high lignin and cellulose content compared to plant residues and manure (Mays et al., 2014; Bolinder et al., 2020). The opposite trend is observed in TN levels between the two regions (Table 7), R2 samples presented higher average TN levels than R1 samples. Generally, R1 areas showed residues rich in organic carbon at the expense of nitrogen, despite the presence of legume residues and manure in some plots.

The high TN content in R2 can be attributed to the fact that the areas evaluated in the Central Valley region were subjected to experiments with different nitrogen doses in the orchards, as well as the presence of wood chips and fumigation residues at the site during the sample collection. The fumigation process can temporarily increase the availability of mineral nitrogen (NH_4^+) and reduce organic matter decomposition, lowering organic carbon release and significantly altering the carbon and nitrogen contents and dynamics at the surface (Fang et al., 2019; Sennett et al., 2022; Yan et al., 2022).

In terms of variability, the coefficient of variation was high and quite similar across all datasets, ranging from a minimum of 39.45% for TN in R1 samples to a maximum of 50.79%, also for TN in the generalized R1&R2 dataset. The results therefore highlight the absence of influence of the quantification method on the variability observed in SOC and TN across all datasets. Thus, the variation in element content across sets results from the specific applications in each evaluated area.

The moderate to high variability found in this study aligns with previous findings for SOC and TN in agricultural surface soils (Seema et al., 2020; Chen et al., 2022; Zhang et al., 2022). This reflects the influence of factors such as agricultural management, soil type, climate, and fertilization practices (Bradari et al., 2024). These results are supported by the

similar maximum and minimum SOC contents in the two regions studied (Table 7). In R1 samples, SOC levels ranged from 0.267% to 2.504%, while in R2 the range was 0.320% to 2.320%. For TN, the minimum and maximum variations were slightly different: from 0.197 g/kg to 1.142 g/kg in R1 and from 0.230 g/kg to 2.13 g/kg in R2.

7.4.2 Spectral Analysis and Reflectance Transformations

The high reflectance observed in the average spectrum of R1 samples (Figure 28), is mainly due to the sandy texture of the samples. According to Jensen (2011), quartz minerals do not exhibit energy absorption properties, and thus sandy soils tend to have higher albedo compared to soils with lower sand content.

The presence of kaolinite is significantly evidenced by the prominent “step” features at 1400 nm and 2200 nm, with its concentration indicated by the depth of absorption at these two wavelengths (Stenberg et al., 2010; Coblinski et al., 2021). In addition to these features, kaolinite may also present a spectral signature around 2300 nm (Kokaly et al., 2017). 2:1 clay mineral, such as smectite, montmorillonite, and illite, formed through a relatively recent weathering process, are also indicated by absorption at 2200 nm but without the step-like feature and with a prominent peak at 1900 nm (Fang et al., 2018), making these wavelengths reliable markers for mineralogical identification.

Previous studies with various soil samples have also highlighted the 2350 nm band as indicative of illite presence (Poppiel et al., 2019; Coblinski et al., 2021). This wavelength showed strong absorption peaks in the spectral behavior of R1 samples, suggesting that this clay mineral may be present on the surface of the evaluated soils.

Samples from R2, in addition to showing a pronounced feature at 1900 nm characteristic of 2:1 minerals, exhibited weak absorption at 1400 nm and 2200 nm, as well as reduced albedo across the spectrum, indicating higher energy absorption, possibly attributable to the less sandy texture of the soils in the region (Jensen, 2011).

The higher spectral variability observed in the shaded area of Figure 28 was more pronounced in R2 samples than in R1, which was also reflected in the variability of SOC and TN contents. This can be attributed to the edaphic heterogeneity of the Central Valley, captured by the diverse samples and the greater sampling depth (0-10 cm). In contrast, R1 samples, collected from the topsoil (0-1 cm), showed less variability, likely due to the shallower sampling depth.

Soil properties such as texture, clay content, organic matter, and moisture vary with depth, influencing reflectance spectra (Coblinski et al., 2020; Zeng et al., 2021). Therefore, sampling depth is a key factor in spectral variability of samples and is essential for accurate analyses and proper interpretation of results.

These findings suggest that the soil in the R1 region is spectrally more homogeneous, with lower intraregional variability, as expected given that the sampling area is smaller than R2's. In this context, spectral heterogeneity combined with chemical variation becomes a valuable source of information for detailed analyses, provided appropriate data processing and interpretation techniques are applied (Jin et al., 2011; Sano et al., 2024).

Differences in spectral behavior between samples become even more apparent after continuum removal of reflectance (Figure 29a) and when using the first derivative (Figure 29b). Continuum removal normalizes the curve around absorption features, setting the maximum reflectance to 1.0 at each feature (Clark et al., 1987). The continuum-removed reflectance value reflects the relative depth of absorption at specific features (Poppiel et al., 2019), allowing observation of absorption nuances and subtle spectral differences between R2 and R1 samples.

The prolonged absorption feature between 430 and 850 nm in R1 samples (Figure 29a) can be attributed to the accumulation of organic structures, especially the C-H group stretches (Stenberg et al., 2010). The complex structure of organic matter, rich in aromatic molecules and carboxylic groups, imparts a strong capacity to absorb energy, masking important characteristics such as the response of iron oxides in the visible spectrum and amplifying absorption peaks (Pearlshtien; Ben-Dor, 2020; Pudelko; Chodak, 2020; Ribeiro et al., 2021).

On the other hand, R2 samples show clear features of iron oxides, highlighted by a peak between 415 and 482 nm, important for detecting goethite in various soils with continuum-removed spectra, as well as characteristic hematite features between 600 and 870 nm (Fang et al., 2018; Seema et al., 2020; Coblinski et al., 2021).

The prominent absorption at 1900 nm, evidenced by reflectance transformed to the first derivative (Figure 29b), indicates the presence of 2:1 type clay mineral in both datasets, with a more pronounced illite behavior in R2 samples, showing strong absorption at 1900 nm, features at 2350 nm, and less expression at 1400 nm and 2200 nm.

In the first derivative, the value is zero at the wavelength where reflectance reaches its maximum and increases according to the rate of change of reflectance with respect to wavelength (Bou-Orm et al., 2020). Sharp variation peaks captured by the first derivative highlight the strong influence of clay mineral absorption, even in sandier soils. The colloids present are well indicated by the spectral behavior in the regions of 1400 nm, 1900 nm, 2200 nm, and 2350 nm.

7.4.3 Predictive Models

7.4.3.1 Multiple Linear Regression (MLR)

For the R1 sample set, % SOC prediction was most robust when using first derivative transformed reflectance data, with 25 wavelengths selected as predictor variables. While untransformed reflectance and continuum-removed spectra produced reasonably accurate models (Table 8), their RPD values of 1.67 and 1.87 indicate model refinement is necessary according to Chang et al. (2001).

For TN (g/kg) prediction, untransformed reflectance and continuum-removed data also performed well with MLR, but the most accurate model again used first derivative reflectance as predictors (Figure 30b). These results are consistent with Gholizadeh et al. (2023), who predicted SOC and TN in litter samples with Adj. R^2 ranging from 0.64 to 0.77 using reflectance spectra collected in darkroom conditions with a contact probe.

With the R2 samples, the best predictive performance for both attributes was also achieved with first derivative reflectance (Figures 30c and 30d). Although accuracy parameters were similar to those of continuum-removed spectra (Table 8), the first derivative offered computational cost advantages, using 21 spectral bands as predictors.

Spectral models often deal with high-dimensional data, making processing computationally intensive and costly (Deneer et al., 2024). Reducing computational cost enables the application of these models at large scale and in resource-limited settings, providing fast, efficient responses without sacrificing precision (Sun; Scanlon, 2019; Akulich et al., 2022; Matsui; Sasaki, 2023).

Specific models for R2 samples showed the best approximation with measured SOC and TN values ($0.85 \leq$ Adj. $R^2 \leq 0.91$), with RPD = 3.34 for SOC and 3.49 for TN. The high accuracy of the local Central Valley model may be attributed to the use of the automated

LECO® Elemental Analyzer, which reduces noise-related variation (Wang; Anderson, 1998; Gazulla et al., 2012; Even et al., 2025).

When combining R1 and R2 samples into a generalized dataset, models showed intermediate accuracy for % SOC using raw reflectance and continuum-removed data (RPD < 2.00, Adj. R^2 = 0.60 and 0.73, minimum RMSE of 0.211 and MAE of 0.172), similar to the local R1 model. SOC and TN prediction stood out with first derivative reflectance data, achieving Adj. R^2 = 0.82 and RPD = 3.12 for SOC (Figure 30e), and Adj. R^2 = 0.90 and RPD = 3.44 for TN (Figure 30f).

The predictive performances found here are more accurate than those reported by Reda et al. (2019), who used neural network models to estimate TN in diverse soils, obtaining R^2 = 0.75, RMSE = 1.10 g/kg, and RPD = 2.55 with reflectance data.

An increase in the number of selected spectral bands was noted in the generalized R1&R2 model (Table 8), due to the higher spectral variability resulting from combining soils with different physicochemical and mineralogical compositions (Sano et al., 2024; Mishra et al., 2020). This required capturing a broader range of absorption features significant to diverse edaphoclimatic contexts. Xu et al. (2017) and Zeng et al. (2021) suggest that effective predictive models should consider multiple spectral variables to isolate influences of different soil components, as spectral response is often affected by cumulative properties.

Although the best-performing models are highlighted, all developed models—even those with fewer spectral variables—proved reliable and computationally efficient. Complete equations for all multiple linear regression models are provided in Appendix E.

7.4.3.2 Partial Least Squares Regression (PLSR)

Table 9 shows that the use of the first derivative generally improves model accuracy, as evidenced by higher Adj. R^2 and RPD values, particularly in the R1 and R2 datasets, highlighting the importance of this spectral transformation to enhance accuracy by emphasizing relevant spectral variations for soil attribute quantification (Ribeiro et al., 2021).

For the R1 region, the use of the first derivative yielded accurate models, with % SOC prediction achieving an Adj. R^2 of 0.81, RPD of 2.30, and low errors (Figure 31a). TN prediction was even more accurate, with Adj. R^2 = 0.87, RPD = 2.81, RMSE = 0.065, and MAE = 0.054, reflecting the model's superior ability to predict total nitrogen levels (Figure 31b).

These results align with findings by Pudelko and Chodak (2019), who developed PLSR models for SOC and TN prediction based on reflectance spectra, obtaining an RPD of approximately 2.19. Chen et al. (2022) also reported similar results modeling various nitrogen types in soils via reflectance spectroscopy, with Adj. R^2 ranging from 0.88 to 0.96 and RPD varying between 2.92 and 5.65.

For R2 samples, the % SOC model reached Adj. $R^2 = 0.92$, RPD = 3.34, RMSE = 0.117, and MAE = 0.087 using 5 factors (Figure 31c), while TN prediction attained Adj. $R^2 = 0.92$, RPD = 3.49, RMSE = 0.106, and MAE = 0.080 with 7 factors (Figure 31d), confirming the efficacy of the first derivative in modeling these attributes, with reduced dispersion between predicted and observed values.

When the regional datasets were combined into R1&R2, % SOC prediction exhibited intermediate performance (Figure 31e), falling between the local models of R1 and R2. TN prediction also showed intermediate accuracy metrics (Figure 31f), with Adj. $R^2 = 0.90$, RPD = 2.86, RMSE = 0.104, and MAE = 0.091.

Findings indicate that combining samples from distinct regions may slightly reduce accuracy due to data heterogeneity, yet still maintain strong predictive capability (Xu et al., 2017). Results align with Tahmasbian et al. (2018), who developed PLSR models for TN using hyperspectral data, achieving $R^2 = 0.86$ and RPD = 2.08.

The predictive behavior of the combined R1&R2 model, observed in both MLR and PLSR, suggests that the global model is more robust than the local R1 model, generalizing better and providing useful predictions for heterogeneous soils. However, its precision does not match the isolated R2 model, possibly due to combining samples quantified by different methodologies, with R1 being more affected by external noise than R2. Global predictive models, nonetheless, offer operational and scalability advantages, enabling application across multiple regions (Alaminos et al., 2016; Ye et al., 2021; Zhou et al., 2021).

Despite methodological differences between R1 and R2, both models demonstrated high efficiency in MLR and PLSR across all trials (Table 8 and Table 9), reinforcing the robustness of the applied methods and their capacity to capture SOC and TN variation irrespective of measurement approaches.

The similarity in accuracy between MLR and PLSR reflects the parsimony principle, which favors simpler solutions that adequately explain the data. Although complex

methods are employed, simpler approaches should be considered when they do not yield significant improvements in prediction (Kargah-Ostadi, 2014; Ren et al., 2021).

7.5 Influential Wavelengths for Predictive Models

For all regression algorithms and datasets, spectral information using the first derivative of reflectance demonstrated the best predictive performance for soil organic carbon and total nitrogen. The main difference between MLR and PLSR lies in the interpretation of the coefficients (slopes). In MLR, the slopes are directly associated with predictor variables (spectral bands) and represent the influence of them on the variation of the predicted variable (SOC or TN). In PLSR, the slopes reflect the influence of each predictor variable within latent factors, which are combinations of those variables in predict model. This characteristic enables PLSR to capture complex spectral variations in a more global manner (Wang et al., 2016; Ribeiro et al., 2021; Gholizadeh et al., 2023).

Figure 32a shows the overlap in the influence of certain spectral bands on % SOC prediction in R1 samples by both MLR and PLSR. Spectral bands at 772 nm, 818 nm, 845 nm, 1707 nm, 2192 nm, and 2313 nm demonstrate significant and consistent influence across both methods based on the first derivative of reflectance.

Wavelengths near 800 nm, according to Alsaleh et al. (2025), are linked to absorption by humic substances rich in carboxylic groups, particularly found in cellulose and lignin from plant residues on the soil surface. Rossel and Behrens (2010) indicate that wavelengths near 850 nm correspond to spectral signatures of compounds containing C-H groups, while bands around 1700 nm relate to absorption by methyl and alkyl groups. Ribeiro et al. (2023) also associated regions near 1700 nm with aromatic groups in humic substances, observing negative slopes in the model, confirmed for the 1707 nm band in this study.

Spectral bands exhibiting similar slopes in MLR and PLSR usually indicate a robust relationship with the target attribute, especially when associated with high predictive performance (Tang et al., 2022). In soil spectroscopy studies from China, Wang et al. (2016) demonstrated that both methods identified similar bands related to soil organic matter, reinforcing the statistical and physical relevance of the convergent bands.

In the R1 dataset, PLSR smoothed the slopes compared to MLR, interpreted as correction for multicollinearity and noise in the spectral and SOC data (Vašát et al., 2014; Wang et al., 2023). By decomposing multicollinearity between adjacent bands, PLSR reduced extreme

slopes by up to 40% in the local model R1, enhancing stability (Figure 32a). This effect is evident in the 952/953 nm pair, where MLR shows slopes with opposite signs due to high first derivative variability, possibly caused by noise. In PLSR, the regularization aligned signs and reduced magnitudes, indicating the discrepancy arose from correlation among neighboring bands rather than actual physicochemical differences. Even after attenuation by PLSR, the bands at 802 nm and 818 nm remained relevant for SOC prediction, indicating that the spectral features were preserved and recognized by both models.

According to Vašát et al. (2014) and Wang et al. (2016), PLSR regression coefficients tend to be smooth for spectral variables, reflecting the continuous nature of informative data, while non-regularized methods like MLR may produce artificial peaks or sign inversions.

For the R2 samples, PLSR maintained the coefficient structure quite similarly to MLR, with minor differences and maximum attenuations of up to 10% in coefficient values (Figure 32b). This suggests lower redundancy among bands in this predictive model, allowing PLSR to preserve nearly the entire amplitude of the direct linear model. In the R2 dataset, the strongest negative correlations between the first derivative of reflectance and SOC prediction occurred at 599 nm, 1165 nm, 1628 nm, 1730 nm, 1906 nm, 1983 nm, and 2016 nm. These regions have previously been identified as influential for organic carbon detection in various soil samples (Seema et al., 2020; Ribeiro et al., 2021; Alsaleh et al., 2025). Additionally, Stenberg et al. (2010) highlighted that regions around 1100 nm and between 1600 and 1800 nm are strongly influenced by organic structures containing phenolic compounds and aliphatic carboxyl groups.

Differences in quantification methodologies may introduce noise into the results, affecting the consistency of the calculated coefficients. The titration method used for the R1 samples is less sensitive to soils with low carbon concentration and sandy texture, such as those in the studied region (Avramidis; Bekiari, 2021). In contrast, the elemental analyzer applied to the R2 samples is more sensitive, providing more accurate results, especially for soils with low SOC content (Gazulla et al., 2012; Even et al., 2025). The higher sensitivity and lower interference of the elemental analyzer make it a more reliable methodology for sandy loam soils like those of R2.

In the combined R1&R2 dataset, the model based on the first derivative of reflectance highlighted several particularly influential wavelengths (Figure 32c). The 818 nm

band exhibited the highest coefficient, remaining a dominant carbon marker even after smoothing by PLSR. Other wavelengths such as 558 nm, 1077 nm, 1252 nm, 1340 nm, between 1591 and 1664 nm, 1864 nm, and 1996 nm also proved influential for SOC prediction in both MLR and PLSR.

Influences observed near 550 nm can be attributed to the presence of organic matter associated with iron oxides, responsible for peaks and absorption features in the visible region (Pearlshtien; Ben-Dor, 2020; Cambou et al., 2022). Ribeiro et al. (2023) identified bands at 1608 nm and near 1840 nm as influential in predicting organic carbon associated with humic substances in sandy soils. Additionally, Stuart (2004) reported that bands between 1650 and 1800 nm correspond to the first overtone of C-H stretching absorption in aromatic compounds.

The convergence of slopes in these regions demonstrates that, despite the diversity in the clay contents and SOC quantification methodologies, the generalized model was able to identify spectrally coherent features physically consistent with organic carbon distribution, maintaining robust performance due to the selective attenuation of extreme coefficients applied by PLSR.

For TN prediction in R1 samples, regression methods showed significant differences among spectral band slopes. In the MLR model, the bands at 868 nm, 879 nm, 978 nm, 1179 nm, 1244 nm, 1302 nm, 1445 nm, 1502 nm, 1745 nm, and 2019 nm were the most influential, reflecting the spectral behavior of the samples (Figure 33a). Upon applying PLSR, extreme variations in slopes were notably smoothed. Bands near 868 nm and 879 nm experienced coefficient reductions of up to 70%, demonstrating PLSR's capability to stabilize slopes. This adjustment contributed to more stable and robust models with greater potential for generalization to new datasets by reducing noise impact and regularizing the model (Wang et al., 2016).

Considering the spectral behavior of R1 samples, bands at 1179 nm, 1244 nm, 1502 nm, 1745 nm, and 2019 nm were the most influential in total nitrogen variation, particularly in the fifth latent variable of the PLSR, highlighting key spectral interaction points for this element in the soil samples. These findings align with studies such as Wei et al. (2017), which identified bands near 1239 nm and 1770 nm as relevant for nitrate detection in soil.

Fidêncio et al. (2002) noted that amine-containing structures (N-H) tend to absorb energy via vibrations in the 2000–2100 nm range. Gholizadeh et al. (2023) also observed

spectral bands near 1100 nm influenced by amine molecule absorption. Seema et al. (2020) highlighted this region as strongly correlated with SOC detection across various soils.

Given the strong influence of amine and carboxyl groups in organic matter structures, the nitrogen-carbon association likely contributed to energy absorption and reflectance variation near 1170 nm (Stenberg et al., 2010). This pattern was also observed near 1700 nm.

Although both methods are efficient, differences between MLR and PLSR reflect the regularization effect of PLSR, especially in datasets with high noise or multicollinearity. In the R1 set, dominated by sandy soils, PLSR helped reduce the impact of local spectral anomalies, yielding more reliable TN predictions by maximizing the relationship between the response variable and predictors. Studies show that PLSR filters spectral noise from collinear bands, resulting in models with more interpretable slopes aligned with actual soil chemical properties (Vašát et al., 2014).

Analyzing the regression coefficients for R2 samples (Figure 33b) revealed higher consistency in slopes magnitudes for both MLR and PLSR, similar to SOC predictive model behavior. This stability may be attributed to the higher precision of the elemental analyzer used for TN quantification. According to Wang and Anderson (1998), this technique carries a lower risk of external contamination, likely leading to more consistent TN values, with less variable coefficients in MLR compared to the maximum 30% smoothing observed in PLSR, notably at 1436 nm.

For the R2 dataset, the first derivative wavelengths exhibiting the strongest inverse influence on total nitrogen (TN) prediction were 465 nm, 529 nm, 605 nm, 1949 nm, 1943 nm, and 1999 nm. This indicates that these spectral bands are negatively correlated with TN levels; as total nitrogen content increases, energy reflected around these bands reduces, resulting in detectable spectral variations in the first derivative.

Conversely, bands at 953 nm and 1436 nm showed positive slopes for TN prediction in R2 samples. The 953 nm band was characterized by Stuart (2004) and Stenberg et al. (2010) as influential in detecting molecular vibrations of N-H groups, especially at wavelengths below 1000 nm. Vibrations in the visible range (465–605 nm) are likely associated with partially decomposed plant residues present in the soil samples (Tahmasbian et al., 2018; Sukhova et al., 2019). These findings align with Misbah et al. (2024), who identified strong influences of spectral bands between 500 and 600 nm for TN prediction in diverse soils.

The spectral bands between 1949 nm and 1999 nm, highlighted as influential, were previously linked by Stuart (2004) to the detection of amide molecules, especially with vibrations near 1450 nm.

In the global model (Figure 33c), the first derivative slopes demonstrated higher consistency compared to isolated R1 models. The R1&R2 model showed intermediate performance, effectively representing data variability. The smoothing effect of PLSR on the global model was less pronounced than in the isolated R1 model, reducing extreme MLR coefficients by up to 51% (at the 617 nm band).

The most influential and consistent wavelengths for both methods were 1440 nm, 1592 nm, 1604 nm, and 1632 nm, exerting direct influence on the first derivative spectrum, while 1479 nm, 1841 nm, and 2254 nm showed inverse effects. These spectral regions are recognized as important for detecting nitrogen-containing organic compounds in soil (Wei et al., 2017; Liang et al., 2025; Mochen et al., 2025). Chen et al. (2022) also emphasize that ammonium, nitrate, and urea affect energy absorption around 1000–1100 nm, 1400 nm, and 1600 nm, with spectral variations dependent on their concentrations, even at low levels.

Bands near 1440 nm and 2250 nm have been associated by Ruma et al. (2024) with indirect prediction of mineralized nitrogen, due to interactions between amine and amide molecules and hydroxyl groups of soil colloids such as kaolinite, illite, smectite, and carboxylic acids. Although bands near 1100 nm did not show strong influence in the model, the first derivative in this region was effective for TN prediction in the R1&R2 dataset. Absorption peaks associated with urea molecules were observed near 1160 nm (Chen et al., 2022), corroborating results for R1 and R1&R2. The equations for the first derivative-based models are provided in Appendix F.

These results demonstrate that simple models like Multiple Linear Regression effectively predict SOC and TN using reflectance spectroscopy, although PLSR better smooths and balances regression coefficients. Combining simplified regression algorithms with generalized predictive models appears promising, particularly for sandy soils, enhancing the application of remote sensing for soil monitoring under diverse environmental conditions. These findings contribute significantly to the scientific community, broadening the applicability of soil monitoring techniques across varied contexts.

7.6 CONCLUSION

This research effectively assessed the application of reflectance spectroscopy in predicting soil organic carbon (SOC) and total nitrogen (TN) across two distinct regions (R1 and R2), with different soil properties and agricultural management approaches. Spectral transformations, particularly the first derivative, significantly improved the models' capability to discriminate subtle variations in SOC and TN between the regions.

Multiple Linear Regression (MLR) and Partial Least Squares Regression (PLSR) models demonstrated high predictive accuracy for SOC and TN, particularly when using first derivative transformed reflectance data. The models for R2 exhibited the best performance, with robust predictions. The generalized models also showed highly reliable performance, proving to be efficient and computationally feasible, highlighting their potential for large-scale applications.

Finally, the spectral bands identified as most influential for SOC and TN prediction were consistent across both MLR and PLSR models, reinforcing their importance in soil characterization. Bands located at 818 nm, 1707 nm, and 2000 nm further emphasized the relevance of spectral data in soil monitoring.

This study therefore demonstrates the feasibility of using reflectance spectroscopy to predict SOC and TN in soils with distinct characteristics and management histories. The high predictive performance of the MLR and PLSR models, along with the identification of key spectral bands, provides valuable insights for the field of soil science, offering practical applications for large-scale monitoring, particularly in agricultural contexts with limited resources.

8 FINAL CONSIDERATIONS

Methodologies for monitoring and quantifying soil attributes have evolved in response to the growing demand for sustainable, cost-effective, and low-maintenance analyses. In this context, reflectance spectroscopy and hyperspectral imaging have become essential complements to traditional chemical techniques, especially for analyzing surface soil samples.

The study on EnMAP confirmed its effectiveness as a valuable data source for environmental monitoring, validating its spectral responses against established multispectral sensors with higher spatial and temporal resolutions. EnMAP's high-resolution spectral capabilities surpass traditional multispectral sensors, making it particularly useful in NDVI studies, and offering greater precision in distinguishing soil from vegetation. The combination of EnMAP with Sentinel-2 and PlanetScope has proven to be deeper insights for environmental research and monitoring. The research emphasized that EnMAP, with its multispectral and hyperspectral data integration, holds significant promise for advancing environmental monitoring in a variety of applications, from soil analysis to broader ecological research.

The research evaluating the performance of Multiple Linear Regression (MLR) models for predicting SOC and TN using spectral data from EnMAP and the ASD FieldSpec® sensor further demonstrated the reliability and accuracy of satellite-based predictive models. With ideal study conditions, such as low cloud cover, proximity of sampling and imagery, and enough samples, EnMAP proved to be a practical and efficient alternative for large-scale monitoring. This approach eliminates the need for transporting expensive and cumbersome equipment to remote locations, making it particularly advantageous in areas with limited access and resources. The study also revealed that model calibration for specific SOC and TN ranges could improve predictive accuracy, particularly for soils with lower concentrations of these attributes.

Additionally, laboratory spectral studies conducted on oven-dried soil samples collected from different depths and land uses confirm the efficacy of reflectance spectroscopy for predicting organic carbon and total nitrogen. The development of global predictive models capable of generalizing results to soils with textures ranging from sand to sandy loam further expands the potential applicability of these methodologies across diverse agricultural contexts.

Therefore, this research demonstrates the feasibility and effectiveness of using advanced remote sensing and spectral analysis techniques to predict soil properties such as SOC and TN across diverse regions with different soil characteristics and management histories. The

high predictive performance of the models, alongside the identification of influential spectral bands, provides a solid foundation for future applications in large-scale soil monitoring. These methods offer valuable practical applications in agricultural contexts, especially in areas with limited resources. By reducing operational costs and fieldwork challenges, these technologies open new pathways for more sustainable, efficient, and large-scale environmental monitoring. The results contribute to advancing the field of soil science, encouraging further exploration of these techniques in other environmental contexts and supporting the development of more informed and effective land management strategies.

BIBLIOGRAPHY

ADECE (Agência de Desenvolvimento do Estado do Ceará). **Perímetros públicos irrigados do Ceará**. Online, 2011. Available at: https://www.adece.ce.gov.br/wp-content/uploads/sites/98/2012/10/perimetros_publicos_do_ceara_sb-7.pdf. Accessed on 18 sep. 2024.

AKULICH, F.; ANAHIDEH, H.; SHEYYAB, M.; AMBRE, D. Explainable predictive modeling for limited spectral data. **Chemometrics and intelligent laboratory systems (Print)**, [*s. l.*], v. 225, p. 104572, 2022.

ALAMINOS, D.; DEL CASTILLO, A.; FERNÁNDEZ, M. Á. A Global Model for Bankruptcy Prediction. **PLOS ONE**, [*s. l.*], v. 11, n. 11, p. e0166693, 2016.

ALSALEH, A. R. S.; ALCIBAHY, M.; GAFOOR, F. A.; HASHEMI, H. A.; ATHAMNEH, B.; AL HAMMADI, A. A.; SENEVIRATNE, L.; AL SHEHHI, M. R. Estimation of soil organic carbon in arid agricultural fields based on hyperspectral satellite images. **Geoderma**, [*s. l.*], v. 453, p. 117151, 2025.

AMELSE, J. Achieving Net Zero Carbon Dioxide by Sequestering Biomass Carbon. **Preprints**, [*s. l.*], 2020.

AMOLI, M. G.; HASANLOU, M.; MEHRJARDI, R. T.; SAMADZADEGAN, F. Exploring the Potential of PRISMA Satellite Hyperspectral Image for Estimating Soil Organic Carbon in Marvdasht Region, Southern Iran. **Remote Sensing**, [*s. l.*], v. 16, n. 12, p. 2149, 2024.

ANDRADE, R.; MANCINI, M.; TEIXEIRA, A. F. dos S.; SILVA, S. H. G.; WEINDORF, D. C.; CHAKRABORTY, S.; GUILHERME, L. R. G.; CURI, N. Proximal sensor data fusion and auxiliary information for tropical soil property prediction: Soil texture. **Geoderma**, [*s. l.*], v. 422, p. 115936, 2022.

ANGELOPOULOU, T.; BALAFOUTIS, A.; ZALIDIS, G. C.; BOCHTIS, D. From Laboratory to Proximal Sensing Spectroscopy for Soil Organic Carbon Estimation—A Review. **Sustainability**, [*s. l.*], v. 12, n. 2, p. 443–443, 2020.

ANGELOPOULOU, T.; TZIOLAS, N.; BALAFOUTIS, A.; ZALIDIS, G.; BOCHTIS, D. Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review. **Remote Sensing**, [*s. l.*], v. 11, n. 6, p. 676, 2019.

ASSMANN, J. J.; KERBY, J. T.; CUNLIFFE, A. M.; MYERS-SMITH, I. H. Vegetation monitoring using multispectral sensors — best practices and lessons learned from high latitudes. **Journal of Unmanned Vehicle Systems**, [*s. l.*], v. 7, n. 1, p. 54–75, 2019.

AVDAN, U.; KAPLAN, G.; AVDAN, Z. Y.; MATCI, D. K.; ERDEM, F.; MIZIK, E. T.; DEMIRTAS, I. Comparison of Remote Sensing Soil Electrical Conductivity from PlanetScope and Ground Measured Data in Wheat and Beet Yields. **Biology and Life Sciences Forum**, [*s. l.*], p. 48, 2021.

AVRAMIDIS, P.; BEKIARI, V. Application of a catalytic oxidation method for the simultaneous determination of total organic carbon and total nitrogen in marine sediments and soils. **PLOS ONE**, [*s. l.*], v. 16, n. 6, p. e0252308, 2021.

BADRARI, H.; GHIMIRE, R.; ARYAL, D. R.; MESBAH, A. Soil profile distribution of organic, inorganic, and labile carbon and nitrogen fractions vary in semi-arid drylands under long-term conservation tillage systems. **Soil Use and Management**, [*s. l.*], v. 40, n. 4, 2024.

BEAMISH, A.; BERGER, K.; BUDDENBAUM, H.; CHABRILLAT, S.; COOPER, S.; DANNER, M.; DOERFFER, R.; DOTZLER, S.; FÖRSTER, S.; GUANTER, L.; HANK, T. **Environmental Mapping and Analysis Program: Mission overview and applications.** Bonn: EnMAP Consortium, 2022.

BELLON-MAUREL, V.; FERNANDEZ-AHUMADA, E.; PALAGOS, B.; ROGER, J.-M.; MCBRATNEY, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. **TrAC Trends in Analytical Chemistry**, [*s. l.*], v. 29, n. 9, p. 1073–1081, 2010.

BEN-DOR, E.; BANIN, A. Determination of organic matter content in arid-zone soils using a simple "loss-on-ignition" method. **Communications in Soil Science and Plant Analysis**, [*s. l.*], v. 20, n. 15-16, p. 1675–1695, 1989.

BEZERRA, U. A.; OLIVEIRA, L. M. M.; CANDEIAS, A. L. B.; SILVA, B. B.; LEITE, A. C. L. S.; SILVA, L. T. M. S. Comparison of the Normalized Difference Vegetation Index (NDVI) Between the Sensors OLI-Landsat Satellite-8 and MSI-Sentinel-2 Satellite in Semi-Arid Region. **Anuário do Instituto de Geociências - UFRJ**, [*s. l.*], v. 41, n. 3, p. 167–177, 2018.

BISHOP, J. L.; M. DANIEL LANE; DYAR, M. D.; BROWN, A. S. Reflectance and emission spectroscopy study of four groups of phyllosilicates: smectites, kaolinite-serpentines, chlorites and micas. **Clay Minerals**, [*s. l.*], v. 43, n. 1, p. 35–54, 2008.

BOLINDER, M. A.; CROTTY, F.; ELSEN, A.; FRAC, M.; KISMÁNYOKY, T.; LIPIEC, J.; TITS, M.; TÓTH, Z.; KÄTTERER, T. The effect of crop residues, cover crops, manures and nitrogen fertilization on soil organic carbon changes in agroecosystems: a synthesis of reviews. **Mitigation and Adaptation Strategies for Global Change**, [*s. l.*], v. 25, n. 6, p. 929–952, 2020.

BOREL, C. C.; GERSTL, S. A. W. Nonlinear spectral mixing models for vegetative and soil surfaces. **Remote Sensing of Environment**, [*s. l.*], v. 47, n. 3, p. 403–416, 1994.

BOU-ORM, N.; ALROMAITHI, A. A.; ELRMEITHI, M.; ALI, F. M.; NAZZAL, Y.; HOWARI, F. M.; AL AYDAROOS, F. Advantages of first-derivative reflectance spectroscopy in the VNIR-SWIR for the quantification of olivine and hematite. **Planetary and Space Science**, [*s. l.*], v. 188, p. 104957, 2020.

BRUNET, D.; BARTHÈS, B. G.; CHOTTE, J.-L.; FELLER, C. Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. **Geoderma**, [*s. l.*], v. 139, n. 1-2, p. 106–117, 2007.

CAMBOU, A.; BARTHÈS, B. G.; MOULIN, P.; CHAUVIN, L.; FAYE, E. H.; MASSE, D.; CHEVALLIER, T.; CHAPUIS-LARDY, L. Prediction of soil carbon and nitrogen contents using visible and near infrared diffuse reflectance spectroscopy in varying salt-affected soils in Sine Saloum (Senegal). **CATENA**, [*s. l.*], v. 212, p. 106075, 2022.

CARVALHO, T.; BROSINSKY, A.; FOERSTER, S.; TEIXEIRA, A.; MEDEIROS, P. Reservoir sediment characterisation by diffuse reflectance spectroscopy in a semiarid region to support sediment reuse for soil fertilization. **Journal of Soils and Sediments**, [*s. l.*], v. 22, n. 9, p. 2557–2577, 2022.

CASTALDI, F.; PALOMBO, A.; PASCUCCI, S.; PIGNATTI, S.; SANTINI, F.; CASA, R. Reducing the Influence of Soil Moisture on the Estimation of Clay from Hyperspectral Data:

A Case Study Using Simulated PRISMA Data. **Remote Sensing**, [*s. l.*], v. 7, n. 11, p. 15561–15582, 2015.

CASTALDI, F.; PALOMBO, A.; SANTINI, F.; PASCUCCI, S.; PIGNATTI, S.; CASA, R. Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. **Remote Sensing of Environment**, [*s. l.*], v. 179, p. 54–65, 2016.

CHABRILLAT, S.; BEN-DOR, E.; CIERNIEWSKI, J.; GOMEZ, C.; SCHMID, T.; VAN WESEMAEL, B. Imaging Spectroscopy for Soil Mapping and Monitoring. **Surveys in Geophysics**, [*s. l.*], v. 40, n. 3, p. 361–399, 2019.

CHABRILLAT, S. et al. The EnMAP spaceborne imaging spectroscopy mission: Initial scientific results two years after launch. **Remote Sensing of Environment**, [*s. l.*], v. 315, p. 114379, 2024.

CHABRILLAT, S. et al. EnMAP Science Plan. **GFZ Data Services**, Postdam 2022.

CHANG, C.-W.; LAIRD, D. A.; MAUSBACH, M. J.; HURBURGH, C. R. Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties. **Soil Science Society of America Journal**, [*s. l.*], v. 65, n. 2, p. 480–490, 2001.

CHEN, Z.; REN, S.; QIN, R.; NIE, P. Rapid Detection of Different Types of Soil Nitrogen Using Near-Infrared Hyperspectral Imaging. **Molecules**, [*s. l.*], v. 27, n. 6, p. 2017, 2022.

CHEN, S.; ZHANG, X.; SHAO, L.; SUN, H.; NIU, J.; LIU, X. Effects of straw and manure management on soil and crop performance in North China Plain. **Catena**, [*s. l.*], v. 187, p. 104359–104359, 2020.

CHEN, Y.; ZHANG, S.; WANG, Y. Distribution Characteristics and Drivers of Soil Carbon and Nitrogen in the Drylands of Central Asia. **Land**, [*s. l.*], v. 11, n. 10, p. 1723, 2022.

CHI, Y.; ZHAO, M.; SUN, J.; XIE, Z.; WANG, E. Mapping soil total nitrogen in an estuarine area with high landscape fragmentation using a multiple-scale approach. **Geoderma**, [*s. l.*], v. 339, p. 70–84, 2019.

CLARK, R. N.; KING, T. V. V. **Automatic continuum analysis of reflectance spectra**. Denver, CO: JPL Proceedings of the 3rd Airborne Imaging Spectrometer Data Analysis Workshop, 1987.

COBLINSKI, J. A.; GIASSON, E.; ALEXANDRE, J.; DOTTO, A. C.; COSTA, J. C.; VAŠÁT, R. Prediction of soil texture classes through different wavelength regions of reflectance spectroscopy at various soil depths. **Catena**, [*s. l.*], v. 189, p. 104485, 2020.

COBLINSKI, J. A.; INDA, A. V.; DEMATTÊ, J. A. M.; DOTTO, A. C.; GHOLIZADEH, A.; GIASSON, É. Identification of minerals in subtropical soils with different textural classes by VIS–NIR–SWIR reflectance spectroscopy. **CATENA**, [*s. l.*], v. 203, p. 105334, 2021.

COLLIN, A. M.; ANDEL, M.; JAMES, D.; CLAUDET, J. The superspectral/hyperspatial worldview-3 as the link between spaceborne hyperspectral and airborne hyperspatial sensors: the case study of the complex tropical coast. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, [*s. l.*], v. XLII-2/W13, p. 1849–1854, 2019.

CURCIO, D.; CIRAOLO, G.; D'ASARO, F.; MINACAPILLI, M. Prediction of Soil Texture Distributions Using VNIR-SWIR Reflectance Spectroscopy. **Procedia Environmental Sciences**, [*s. l.*], v. 19, p. 494–503, 2013.

DALAL, R. C.; HENRY, R. J. Simultaneous Determination of Moisture, Organic Carbon, and Total Nitrogen by Near Infrared Reflectance Spectrophotometry. **Soil Science Society of America Journal**, [*s. l.*], v. 50, n. 1, p. 120–123, 1986.

DALMOLIN, R. S. D.; GONÇALVES, C. N.; KLAMT, E.; DICK, D. P. Relação entre os constituintes do solo e seu comportamento espectral. **Ciência Rural**, [*s. l.*], v. 35, n. 2, p. 481–489, 2005.

DEMATTÊ, J. A. M. et al. The Brazilian Soil Spectral Library (BSSL): A general view, application and challenges. **Geoderma**, [*s. l.*], v. 354, p. 113793–113793, 2019.

DEMATTÊ, J. A. M.; FONGARO, C. T.; RIZZO, R.; SAFANELLI, J. L. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. **Remote Sensing of Environment**, [*s. l.*], v. 212, p. 161–175, 2018.

DEMATTÊ, J. A. M.; TERRA, F. da S. Spectral pedology: A new perspective on evaluation of soils along pedogenetic alterations. **Geoderma**, [*s. l.*], v. 217-218, p. 190–200, 2014.

DENEER, A.; MOLENAAR, J.; FLECK, C. Spectral expansion methods for prediction uncertainty quantification in systems biology. **Frontiers in Systems Biology**, [*s. l.*], v. 4, 2024.

DENG, X.; MA, W.; REN, Z.; ZHANG, M.; GRIENEISEN, M. L.; CHEN, X.; FEI, X.; QIN, F.; ZHAN, Y.; LV, X. Spatial and temporal trends of soil total nitrogen and C/N ratio for croplands of East China. **Geoderma**, [*s. l.*], v. 361, p. 114035, 2020.

DEO, R. N.; SOUNTHARARAJAH, A.; KODIKARA, J. Reflectance spectroscopy of asphalt pavement under moisture and diesel perturbations: A laboratory study. **Construction and Building Materials**, [*s. l.*], v. 415, p. 135076, 2024.

DEODORO, S. C.; ANTÔNIO, M.; GARCIA, R. A.; TEMBA, P. da C.; BERTOLINI, W. Z. Classificação e mapeamento da textura superficial do solo a partir de dados de sensoriamento remoto e análise discriminante, na região de Volta Grande do Rio Uruguai – Brasil. **Revista Brasileira de Geografia Física**, [*s. l.*], v. 14, n. 1, p. 340–356, 2021.

DINAKARAN, J.; BIDALIA, A.; KUMAR, A.; HANIEF, M.; MEENA, A.; RAO, K. S. Near-Infrared-Spectroscopy for Determination of Carbon and Nitrogen in Indian Soils. **Communications in Soil Science and Plant Analysis**, [*s. l.*], v. 47, n. 12, p. 1503–1516, 2016.

DOTZLER, S.; HILL, J.; BUDDENBAUM, H.; STOFFELS, J. The Potential of EnMAP and Sentinel-2 Data for Detecting Drought Stress Phenomena in Deciduous Forest Communities. **Remote Sensing**, [*s. l.*], v. 7, n. 10, p. 14227–14258, 2015.

EUROPEAN SPACE AGENCY (ESA). **S2 Mission**. [*S. l.: s. n.*], 2025. Available at: https://sentiwiki.copernicus.eu/web/s2-mission. Accessed on: 29 Apr. 2025.

EVEN, R. J.; MACHMULLER, M. B.; LAVALLEE, J. M.; ZELIKOVA, T. J.; M. FRANCESCA COTRUFO. Large errors in soil carbon measurements attributed to inconsistent sample processing. **SOIL**, [*s. l.*], v. 11, n. 1, p. 17–34, 2025.

FANG, Q.; HONG, H.; ZHAO, L.; KUKOLICH, S.; YIN, K.; WANG, C. Visible and Near-Infrared Reflectance Spectroscopy for Investigating Soil Mineralogy: A Review. **Journal of Spectroscopy**, [*s. l.*], v. 2018, p. e3168974, 2018.

FANG, W.; YAN, D.; WANG, Q.; HUANG, B.; REN, Z.; WANG, X.; WANG, X.; LI, Y.; OUYANG, C.; QUIRICO MIGHELI; CAO, A. Changes in the abundance and community composition of different nitrogen cycling groups in response to fumigation with 1,3-dichloropropene. **Science of The Total Environment**, [*s. l.*], v. 650, p. 44–55, 2019.

FARINA, A.; PIERGALLINI, R.; DOLDO, A.; SALSANO, E. P.; ABBALLE, F. The determination of C-H-N by an automated elemental analyzer. **Microchemical Journal**, [*s. l.*], v. 43, n. 3, p. 181–190, 1991.

FIDÊNCIO, P. H.; POPPI, R. J.; CARLOS; CANTARELLA, H. Determination of organic matter in soil using near-infrared spectroscopy and partial least squares regression. **Communications in Soil Science and Plant Analysis** , [*s. l.*], v. 33, n. 9-10, p. 1607–1615, 2002.

FOWLER, D. et al. The Global Nitrogen Cycle in the twenty-first Century. **Philosophical Transactions of the Royal Society B: Biological Sciences**, [*s. l.*], v. 368, n. 1621, p. 20130164, 2013.

FRUTH, T.; LENZEN, C.; GROSS, E.; MROWKA, F. The EnMAP Mission Planning System | SpaceOps Conferences. **SpaceOps Conferences**, [*s. l.*], 2018.

GAZULLA, M. F.; RODRIGO, M.; ORDUÑA, M.; GÓMEZ, C. M. Determination of Carbon, Hydrogen, Nitrogen and Sulfur in Geological Materials Using Elemental Analysers. **Geostandards and Geoanalytical Research**, [*s. l.*], v. 36, n. 2, p. 201–217, 2012.

GERENFES, D.; GIORGIS, A.; NEGASA, G. Comparison of organic matter determination methods in soil by loss on ignition and potassium dichromate method. **International journal of horticulture and food science**, [*s. l.*], v. 4, n. 1, p. 49–53, 2022.

GHOLIZADEH, A.; SABERIOON, M.; POULADI, N.; BEN-DOR, E. Quantification and depth distribution analysis of carbon to nitrogen ratio in forest soils using reflectance spectroscopy. **International Soil and Water Conservation Research**, [*s. l.*], v. 11, n. 1, p. 112–124, 2023.

GIRÃO, R. de O.; MOREIRA, L. J. da S.; GIRÃO, A. L. de A.; ROMERO, R. E.; FERREIRA, T. O. Soil genesis and iron nodules in a karst environment of the Apodi Plateau. **Revista Ciência Agronômica**, [*s. l.*], v. 45, n. 4, p. 683–695, 2014.

GOCKE, M. I. et al. Interactive effects of agricultural management on soil organic carbon accrual: A synthesis of long-term field experiments in Germany. **Geoderma**, [*s. l.*], v. 438, p. 116616–116616, 2023.

GOMES, A. F.; QUEIROZ, D. M.; VALENTE, D. S. M.; PINTO, F. A. C.; ROSAS, J. T. F. Comparing a single-sensor camera with a multisensor camera for monitoring coffee crop using Unmanned Aerial Vehicles. **Engenharia Agrícola**, [*s. l.*], v. 41, n. 1, p. 87–97, 2021.

GUANTER, L. et al. The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. **Remote Sensing**, [*s. l.*], v. 7, n. 7, p. 8830–8857, 2015.

GUO, L.; FU, P.; SHI, T.; CHEN, Y.; ZHANG, H.-T.; MENG, R.; WANG, S.-Q. Mapping field-scale soil organic carbon with unmanned aircraft system-acquired time series multispectral images. **Soil and Tillage Research**, [*s. l.*], v. 196, p. 104477, 2020.

GUO, L.; SUN, X.; FU, P.; SHI, T.; DANG, L.; CHEN, Y.; LINDERMAN, M.; ZHANG, G.; ZHANG, Y.; JIANG, Q.; ZHANG, H.; ZENG, C. Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas. **Geoderma**, [*s. l.*], v. 398, p. 115118, 2021.

GUOJU, X.; YANBIN, H.; QIANG, Z.; JING, W.; MING, L. Impact of cultivation on soil organic carbon and carbon sequestration potential in semiarid regions of China. **Soil Use and Management**, [*s. l.*], v. 36, n. 1, p. 83–92, 2019.

HERRANZ-LUQUE, J. E. et al. Monitoring Land Management Practices Using Vis–NIR Spectroscopy Provides Insights into Predicting Soil Organic Carbon and Limestone Levels in Agricultural Plots. **Agronomy**, [*s. l.*], v. 14, n. 6, p. 1150, 2024.

HOGBERG, P.; READ, D. J. Towards a more plant physiological perspective on soil ecology. **Trends in Ecology & Evolution**, [*s. l.*], v. 21, n. 10, p. 548–554, 2006.

HONG, Y.; LIU, Y.; CHEN, Y.; YU, L.; LIU, Y.; CHENG, H. Application of fractional-order derivative in the quantitative estimation of soil organic matter content through visible and near-infrared spectroscopy. **Geoderma**, [*s. l.*], v. 337, p. 758–769, 2019.

HU, J.; WU, J.; QU, X. Effects of organic wastes on labile organic carbon in semiarid soil under plastic mulched drip irrigation. **Archives of Agronomy and Soil Science**, [*s. l.*], v. 65, n. 13, p. 1873–1884, 2019.

HUANG, Z.; CAO, C.; CHEN, W.; XU, M.; DANG, Y.; SINGH, R. P.; BASHIR, B.; XIE, B.; LIN, X. Remote Sensing Monitoring of Vegetation Dynamic Changes after Fire in the Greater Hinggan Mountain Area: The Algorithm and Application for Eliminating Phenological Impacts. **Remote Sensing**, [*s. l.*], v. 12, n. 1, p. 156, 2020.

IBM CORP. **IBM SPSS Statistics for Windows, Version 21.0**. Armonk, NY: IBM Corp, 2012.

ISSAOUI, W.; ALEXAKIS, D. D.; NASR, I. H.; ARGYRIOU, A. V.; ALEVIZOS, E.; PAPADOPOULOS, N.; INOUBLI, M. H. Monitoring Olive Oil Mill Wastewater Disposal Sites Using Sentinel-2 and PlanetScope Satellite Images: Case Studies in Tunisia and Greece. **Agronomy**, [*s. l.*], v. 12, n. 1, p. 90–90, 2021.

JAKIMOW, B.; JANZ, A.; THIEL, F.; OKUJENI, A.; HOSTERT, P.; SEBASTIAN. EnMAP-Box: Imaging spectroscopy in QGIS. **SoftwareX**, [*s. l.*], v. 23, p. 101507, 2023.

JENSEN, J. R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. São José Dos Campos, Sp: Parêntese, 2011.

JIANG, C. Y. H. Digital Elevation Model and Satellite Imagery Based Bushfire Simulation. **American Journal of Geographic Information System**, [*s. l.*], v. 2, n. 3, p. 47–65, 2013.

JIANG, Y.; WU, Y.; DI, H.; TIAN, X.; WANG, T.; WANG, R.; ZHAO, J.; HU, R.; SHAABAN, M. Impact of biocrust on soil nitrogen maintenance and microbial composition in citrus orchards under varying urea application rates. **Applied Soil Ecology**, [*s. l.*], v. 201, p. 105497–105497, 2024.

JIN, J.-W.; CHEN, Z.-P.; LI, L.-M.; STEPONAVICIUS, R.; THENNADIL, S. N.; YANG, J.; YU, R.-Q. Quantitative Spectroscopic Analysis of Heterogeneous Mixtures: The Correction of Multiplicative Effects Caused by Variations in Physical Properties of Samples. **Analytical Chemistry**, [*s. l.*], v. 84, n. 1, p. 320–326, 2011.

JIN, Y.; HE, R.; MARINO, G.; WHITING, M.; KENT, E.; SANDEN, B. L.; CULUMBER, M.; FERGUSON, L.; LITTLE, C.; GRATTAN, S.; KYAW; LAGOS, L. O.; SNYDER, R. L.; ZACCARIA, D. Spatially variable evapotranspiration over salt affected pistachio orchards analyzed with satellite remote sensing estimates. **Agricultural and forest meteorology**, [*s. l.*], v. 262, p. 178–191, 2018.

KARGAH-OSTADI, N. Comparison of Machine Learning Techniques for Developing Performance Prediction Models. **Computing in Civil and Building Engineering**, [*s. l.*], 2014.

KAUFFMAN, E. Climate and Topography. **Atlas of the Biodiversity of California**, [*s. l.*], v. 12, p. 15, 2003. Online. Available at: https://www.coastal.ca.gov/coastalvoices/resources/Biodiversity_Atlas_Climate_and_Topography.pdf. Access on jan. 2025

KAUFMAN, Y. J.; SENDRA, C. Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. **International Journal of Remote Sensing**, [*s. l.*], v. 9, n. 8, p. 1357–1381, 1988. Disponível em: Acesso em: 13 fev. 2022.

KJELDAHL, J. Neue Methode zur Bestimmung des Stickstoffs in organischen Körpern. **Fresenius' Zeitschrift für analytische Chemie**, [*s. l.*], v. 22, n. 1, p. 366–382, 1883.

KOKALY, R. F.; CLARK, R. N.; SWAYZE, G. A.; LIVO, K. E.; HOEFEN, T. M.; PEARSON, N. C.; WISE, R. A.; BENZEL, W. M.; LOWERS, H. A.; DRISCOLL, R. L.; KLEIN, A. J. USGS Spectral Library Version 7. **Data Series**, [*s. l.*], 2017.

KUKUĻS, I.; KĻAVIŅŠ, M.; NIKODEMUS, O.; KASPARINSKIS, R.; BRŪMELIS, G. Changes in soil organic matter and soil humic substances following the afforestation of former agricultural lands in the boreal-nemoral ecotone (Latvia). **Geoderma Regional**, [*s. l.*], v. 16, p. e00213–e00213, 2019.

LAAMRANI, A.; BERG, A.; VORONEY, P.; HANNES FEILHAUER; BLACKBURN, L.; MARCH, M.; DAO, P. D.; HE, Y.; MARTIN, R. C. Ensemble Identification of Spectral Bands Related to Soil Organic Carbon Levels over an Agricultural Field in Southern Ontario, Canada. **Remote Sensing**, [*s. l.*], v. 11, n. 11, p. 1298, 2019.

LAL, R. Soil carbon sequestration to mitigate climate change. **Geoderma**, [*s. l.*], v. 123, n. 1-2, p. 1–22, 2004.

LANFRANCHI, R. A.; CRUZ, S. C. P.; ROCHA, W. F. Application of remote sensing and reflectance spectroscopy to explore iron-enriched domains in the north region of the intracontinental sector of the Araçuaí West Congo Orogen. **Ore Geology Reviews**, [*s. l.*], v. 128, p. 103916, 2021.

LANGE, M.; DECHANT, B.; REBMANN, C.; VOHLAND, M.; CUNTZ, M.; DOKTOR, D. Validating MODIS and Sentinel-2 NDVI Products at a Temperate Deciduous Forest Site Using Two Independent Ground-Based Sensors. **Sensors**, [*s. l.*], v. 17, n. 8, p. 1855, 2017.

LÉGARÉ, B.; BÉLANGER, S.; SINGH, R. K.; BERNATCHEZ, P.; CUSSON, M. Remote Sensing of Coastal Vegetation Phenology in a Cold Temperate Intertidal System: Implications for Classification of Coastal Habitats. **Remote sensing**, [*s. l.*], v. 14, n. 13, p. 3000, 2022.

LESAIGNOUX, A.; FABRE, S.; BRIOTTET, X. Influence of soil moisture content on spectral reflectance of bare soils in the 0.4–14 μm domain. **International Journal of Remote Sensing**, [*s. l.*], v. 34, n. 7, p. 2268–2285, 2012.

LI, H.; WANG, J.; ZHANG, J.; LIU, T.; ACQUAH, G. E.; YUAN, H. Combining Variable Selection and Multiple Linear Regression for Soil Organic Matter and Total Nitrogen Estimation by DRIFT-MIR Spectroscopy. **Agronomy**, [*s. l.*], v. 12, n. 3, p. 638, 2022.

LI, G.; ZHANG, M.; WU, C. Short-term fallow practices drive soil bacterial community changes: A case study from China. **Applied Soil Ecology**, [*s. l.*], v. 165, p. 103988, 2021.

LIANG, H.; SONG, Y.; DAI, Z.; LIU, H.; ZHONG, K.; FENG, H.; XU, L. Soil total nitrogen content and pH value estimation method considering spatial heterogeneity: Based on GNNW-XGBoost model. **Spectrochimica Acta Part A Molecular and Biomolecular Spectroscopy**, [*s. l.*], v. 330, p. 125716, 2025.

LO, C. P.; YEUNG, A. K. W. **Concepts and Techniques of Geographic Information Systems**. New Jersey: Prentice Hall, 2002.p. 265–304.

MADHAVAN, D. B. et al. Rapid prediction of particulate, humus and resistant fractions of soil organic carbon in reforested lands using infrared spectroscopy. **Journal of Environmental Management**, [*s. l.*], v. 193, p. 290–299, 2017.

MARTY, C.; HOULE, D.; GAGNON, C.; COURCHESNE, F. The relationships of soil total nitrogen concentrations, pools and C:N ratios with climate, vegetation types and nitrate deposition in temperate and boreal forests of eastern Canada. **Catena**, [*s. l.*], v. 152, p. 163–172, 2017.

MATSUI, K.; SASAKI, Y. Computational reduction of the spectral division method for synthesizing moving sources by source trajectory approximation. **The Journal of the Acoustical Society of America**, [*s. l.*], v. 153, n. 1, p. 159–167, 2023.

MAYS, N.; BRYE, K. R.; ROM, C. R.; SAVIN, M. C.; GARCÍA, M. Groundcover Management and Nutrient Source Effects on Soil Carbon and Nitrogen Sequestration in an Organically Managed Apple Orchard in the Ozark Highlands. **Hortscience**, [*s. l.*], v. 49, n. 5, p. 637–644, 2014.

MENDIBURU, F. de. **Agricolae: Statistical Procedures for Agricultural Research**. [*s.l.*], 2023. Available at: https://cran.r-project.org/package=agricolae. Accessed on 15 Sep 2024

MENESES, P. R.; ALMEIDA, T. de ; MACEDO, G. **Reflectância dos materiais terrestres**. São Paulo: Oficina de Textos, 2019.

MINHONI, R. T. de A.; SCUDIERO, E.; ZACCARIA, D.; SAAD, J. C. C. Multitemporal satellite imagery analysis for soil organic carbon assessment in an agricultural farm in southeastern Brazil. **The Science of The Total Environment**, [*s. l.*], v. 784, p. 147216, 2021.

MISBAH, K.; LAAMRANI, A.; VORONEY, P.; KHECHBA, K.; CASA, R.; ABDELGHANI CHEHBOUNI. Ensemble Band Selection for Quantification of Soil Total Nitrogen Levels from Hyperspectral Imagery. **Remote Sensing**, [*s. l.*], v. 16, n. 14, p. 2549–2549, 2024.

MISHRA, U.; GAUTAM, S.; RILEY, W. J.; HOFFMAN, F. M. Ensemble Machine Learning Approach Improves Predicted Spatial Variation of Surface Soil Organic Carbon Stocks in Data-Limited Northern Circumpolar Region. **Frontiers in Big Data**, [*s. l.*], v. 3, 2020.

MOCHEN, L.; KUANKUAN, Y.; YINFA, Y.; ZHANHUA, S.; FUYANG, T.; FADE, L.; ZHENWEI, Y.; ZHANG, R.; QINGLU, Y.; YAO, L. Multi-spectral evaluation of total nitrogen, phosphorus and potassium content in soil using Vis-NIR spectroscopy based on a modified support vector machine with whale optimization algorithm. **Soil and Tillage Research**, [*s. l.*], v. 252, p. 106567, 2025.

MONTEIRO, D. N. C.; LIMA, R. H. C.; COSTA, S. dos S.; SANTOS, E. T. V. dos. Análise dos parâmetros texturais e mineralógicos do solo por espectroscopia na região do Parque Nacional de Anavilhanas (AM). **Revista Geonorte**, [*s. l.*], v. 15, n. 48, 2024.

MUTANGA, O.; VAN AARDT, J.; KUMAR, L. Imaging spectroscopy (hyperspectral remote sensing) in southern Africa: an overview. **South African Journal of Science**, [*s. l.*], v. 105, n. 5/6, 2010.

MZID, N.; CASTALDI, F.; TOLOMIO, M.; PASCUCCI, S.; CASA, R.; PIGNATTI, S. Evaluation of Agricultural Bare Soil Properties Retrieval from Landsat 8, Sentinel-2 and PRISMA Satellite Data. **Remote Sensing**, [*s. l.*], v. 14, n. 3, p. 714, 2022.

NAUE, C. R.; MARQUES, M. W.; LIMA, N. B.; GALVÍNCIO, J. D. Sensoriamento remoto como ferramenta aos estudos de doenças de plantas agrícolas: uma revisão (Remote Sensing as a Toll for the Study of Plant Diseases on Agriculture: a Revision). **Revista Brasileira de Geografia Física**, [*s. l.*], v. 3, n. 3, p. 190–195, 2011.

NOWKANDEH, S. M.; NOROOZI, A. A.; HOMAEE, M. Estimating soil organic matter content from Hyperion reflectance images using PLSR, PCR, MinR and SWR models in semi-arid regions of Iran. **Environmental Development**, [*s. l.*], v. 25, p. 23–32, 2018.

O'HAVER, T. C. Derivative and wavelength modulation spectrometry. **Analytical Chemistry**, [*s. l.*], v. 51, n. 1, p. 91A99A, 1979.

OLIVEIRA, M. R. R. de; RIBEIRO, S. G.; MAS, J.-F.; TEIXEIRA, A. dos S. Advances in hyperspectral sensing in agriculture: a review. **Revista Ciência Agronômica**, [*s. l.*], v. 51, n. 5, 2020.

PEARLSHTIEN, D. H.; BEN-DOR, E. Effect of Organic Matter Content on the Spectral Signature of Iron Oxides across the VIS–NIR Spectral Region in Artificial Mixtures: An Example from a Red Soil from Israel. **Remote Sensing**, [*s. l.*], v. 12, n. 12, p. 1960, 2020.

PEREIRA, A. A.; TEIXEIRA, F. B.; LIBONATI, R.; MELCHIORI, E. A.; CARVALHO, L. Avaliação de índices espectrais para identificação de áreas queimadas no Cerrado utilizando dados Landsat TM. **Revista Brasileira de Cartografia.**, [*s. l.*], v. 68, n. 8, p. 1665–1680, 2016.

PLANET LABS INC. **Understanding PlanetScope Instruments**. Online. 2025. Available at: https://developers.planet.com/docs/apis/data/sensors/. Accessed on: 29 Apr. 2025.

POPPIEL, R. R.; PINTO, M.; ALEXANDRE, J.; OLIVEIRA, M. M.; GALLO, B. C.; SAFANELLI, J. L. Pedology and soil class mapping from proximal and remote sensed data. **Geoderma**, [*s. l.*], v. 348, p. 189–206, 2019.

PUDEŁKO, A.; CHODAK, M. Estimation of total nitrogen and organic carbon contents in mine soils with NIR reflectance spectroscopy and various chemometric methods. **Geoderma**, [*s. l.*], v. 368, p. 114306, 2020.

PURNAMASARI, E.; KAMAL, M.; PRAMADITYA WICAKSONO. Comparison of vegetation indices for estimating above-ground mangrove carbon stocks using PlanetScope image. **Regional Studies in Marine Science**, [*s. l.*], v. 44, p. 101730, 2021.

R CORE TEAM. **R: The R Project for Statistical Computing**. Online. 2024. Available at: https://www.r-project.org. Accessed on: 10 Dec. 2024.

RAHEB, A.; HEIDARI, A.; MAHMOODI, S. Organic and inorganic carbon storage in soils along an arid to dry sub-humid climosequence in northwest of Iran. **CATENA**, [*s. l.*], v. 153, p. 66–74, 2017.

REDA, R.; SAFFAJ, T.; ILHAM, B.; SAIDI, O.; ISSAM, K.; BRAHIM, L.; EL HADRAMI, E. M. A comparative study between a new method and other machine learning algorithms for soil organic carbon and total nitrogen prediction using near infrared spectroscopy. **Chemometrics and Intelligent Laboratory Systems**, [*s. l.*], v. 195, p. 103873, 2019.

REN, S.; GUO, B.; WANG, Z.; WANG, J.; FANG, Q.; WANG, J. Optimized spectral index models for accurately retrieving soil moisture (SM) of winter wheat under water stress. **Agricultural Water Management**, [*s. l.*], v. 261, p. 107333, 2022.

REN, X.-Y.; HAN, R.-S.; CHEN, L. Learning impurity spectral functions from density of states. **Journal of Physics: Condensed Matter**, [*s. l.*], v. 33, n. 49, p. 495601, 2021.

REYES-ROJAS, J. et al. Vertical distribution and variability of soil organic carbon and CaCO3 in deep Colluvisols modeled by hyperspectral imaging. **Geoderma**, [*s. l.*], v. 453, p. 117146, 2025.

RIBEIRO, J. M.; FRAZÃO, L. A.; FERNANDES, L. A.; SAMPAIO, R. A.; CARDOSO, P. H. S.; OLIVEIRA, A. L. G. Fertilidade do solo e estoques de carbono e nitrogênio sob sistemas agroflorestais no Cerrado mineiro. **Ciência Florestal**, [*s. l.*], v. 29, n. 2, p. 913–923, 2019.

RIBEIRO, S. G.; OLIVEIRA, M. R. R. de; LOPES, L. M.; COSTA, M. C. G.; TOMA, R. S.; ARAÚJO, I. C. da S.; MOREIRA, L. C. J.; TEIXEIRA, A. dos S. Reflectance spectroscopy in the prediction of soil organic carbon associated with humic substances. **Rev. Bras. Ciênc. Solo**, [*s. l.*], v. 47, 2023.

RIBEIRO, S. G.; TEIXEIRA, A. dos S.; DE OLIVEIRA, M. R. R.; COSTA, M. C. G.; ARAÚJO, I. C. da S.; MOREIRA, L. C. J.; LOPES, F. B. Soil Organic Carbon Content Prediction Using Soil-Reflected Spectra: A Comparison of Two Regression Methods. **Remote Sensing**, [*s. l.*], v. 13, n. 23, p. 4752, 2021.

ROSSEL, R. A. V.; BEHRENS, T. Using data mining to model and interpret soil diffuse reflectance spectra. **Geoderma**, [*s. l.*], v. 158, n. 1-2, p. 46–54, 2010.

ROSSEL, R. A. V.; WALVOORT, D. J. J.; MCBRATNEY, A. B.; JANIK, L. J.; SKJEMSTAD, J. O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. **Geoderma**, [*s. l.*], v. 131, n. 1-2, p. 59–75, 2006.

RUDORFF, C. M.; DE, L.; LÊNIO SOARES GALVÃO; WATERLOO PEREIRA FILHO. Análise derivativa de dados hiperespectrais medidos em nível de campo e orbital para caracterizar a composição de águas opticamente complexas na Amazônia. **Acta Amazonica**, [*s. l.*], v. 37, n. 2, p. 269–280, 2007.

RUMA, F. Y.; MUHAMMAD ABDUL MUNNAF; NEVE, S. D.; MOUAZEN, A. M. Visible-to-near-infrared spectroscopy for prediction of soil nitrogen mineralization after sample stratification by textural homogeneity criteria. **Soil and Tillage Research**, [*s. l.*], v. 244, p. 106250, 2024.

SALAZAR, S. E.; COFFMAN, R. A. Validation of a ground-based telescope-assisted hyperspectral remote sensor for soil measurements. **Journal of Applied Remote Sensing**, [*s. l.*], v. 14, n. 02, p. 1, 2020.

SALEHI, M. H.; BENI, O. H.; HARCHEGANI, H. B.; BORUJENI, I. E.; MOTAGHIAN, H. R. Refining Soil Organic Matter Determination by Loss-on-Ignition. **Pedosphere**, [*s. l.*], v. 21, n. 4, p. 473–482, 2011.

SANO, M.; YAMASHITA, T.; KITAMURA, Y.; KOKAWA, M. Use of microscale heterogeneity in samples for spectral factorization—A strategy to build robust prediction models for nondestructive analyses. **Food Chemistry**, [*s. l.*], v. 460, p. 140591, 2024.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. **Analytical Chemistry**, [*s. l.*], v. 36, n. 8, p. 1627–1639, 1964.

SCHJØNNING, P. Straw management in small grain cereal crop production – The long-term effects on soil carbon and soil pore characteristics. **Geoderma**, [*s. l.*], v. 435, p. 116499–116499, 2023.

SEEMA; GHOSH, A. K.; DAS, B. S.; REDDY, N. Application of VIS-NIR spectroscopy for estimation of soil organic carbon using different spectral preprocessing techniques and multivariate methods in the middle Indo-Gangetic plains of India. **Geoderma Regional**, [*s. l.*], v. 23, p. e00349, 2020.

SEGL, K.; GUANTER, L.; ROGASS, C.; KUESTER, T.; ROESSNER, S.; KAUFMANN, H.; SANG, B.; V. MOGULSKY; HOFER, S. EeteS—The EnMAP End-to-End Simulation Tool. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, [*s. l.*], v. 5, n. 2, p. 522–530, 2012.

SENNETT, L. B.; BURTON, D. L.; GOYER, C.; ZEBARTH, B. J. Chemical fumigation alters soil carbon and nitrogen dynamics in soils amended with substrates of contrasting carbon availability. **Geoderma**, [*s. l.*], v. 419, p. 115878, 2022.

SENNETT, L.; BURTON, D. L.; GOYER, C.; ZEBARTH, B. J. Influence of chemical fumigation and biofumigation on soil nitrogen cycling processes and nitrifier and denitrifier abundance. **Soil Biology and Biochemistry**, [*s. l.*], v. 162, p. 108421, 2021.

SHAKOOR, A.; BOSCH-SERRA, À. D.; LIDON, A.; GINESTAR, D.; BOIXADERA, J. Soil mineral nitrogen dynamics in fallow periods in a rainfed semiarid Mediterranean agricultural system. **Pedosphere**, [*s. l.*], v. 33, n. 4, p. 622–637, 2023.

SILVA, S. M. P.; CRÓSTA, A. P.; ANGÉLICA, R. S.; BEURLEN, H. Dados EO-1 Hyperion no mapeamento mineralógico de pegmatitos na porção sul da Província Pegmatítica da Borborema (PPB), Nordeste do Brasil. *In*: XIV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2009, Natal, Brasil. **Anais** [...]. Natal, Brasil: INPE, 2009. p. 1207–1214.

SKJEMSTAD, J. O.; BALDOCK, J. A. Total Organic Carbon. *In*: SOIL CHEMICAL METHODS : AUSTRALASIA. Collingwood, Vic. 3066: Csiro Publishing, 2011. p. 225–227.

STEIN, L. Y.; KLOTZ, M. G. The nitrogen cycle. **Current Biology**, [*s. l.*], v. 26, n. 3, p. R94–R98, 2016.

STEINBERG, A.; CHABRILLAT, S.; STEVENS, A.; SEGL, K.; FOERSTER, S. Prediction of Common Surface Soil Properties Based on Vis-NIR Airborne and Simulated EnMAP Imaging Spectroscopy Data: Prediction Accuracy and Influence of Spatial Resolution. **Remote Sensing**, [*s. l.*], v. 8, n. 7, p. 613, 2016.

STENBERG, B.; ROSSEL, R. A. V.; MOUAZEN, A. M.; WETTERLIND, J. Visible and Near Infrared Spectroscopy in Soil Science. **Advances in Agronomy**, [*s. l.*], p. 163–215, 2010.

STEVENS, A.; RAMIREZ-LOPEZ, L. **An introduction to the prospectr package**. [*S. l.*]: R package version 0.2.7, 2024.

STORCH, T. et al. The EnMAP imaging spectroscopy mission towards operations. **Remote Sensing of Environment**, [*s. l.*], v. 294, p. 113632, 2023.

STUART, B. H. **Infrared Spectroscopy: Fundamentals and Applications Stuart/Infrared Spectroscopy: Fundamentals and Applications**. [*S. l.*]: Chichester, Uk John Wiley & Sons, Ltd, 2004.

SUKHOVA, E.; YUDINA, L.; AKINCHITS, E.; VODENEEV, V.; SUKHOV, V. Influence of electrical signals on pea leaf reflectance in the 400–800-nm range. **Plant Signaling & Behavior**, [*s. l.*], v. 14, n. 7, p. 1610301, 2019.

SUN, A. Y.; SCANLON, B. R. How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. **Environmental Research Letters**, [*s. l.*], v. 14, n. 7, p. 073001, 2019.

TAHMASBIAN, I.; XU, Z.; BOYD, S.; ZHOU, J.; ESMAEILANI, R.; CHE, R.; BAI, S. H. Laboratory-based hyperspectral image analysis for predicting soil carbon, nitrogen and their isotopic compositions. **Geoderma**, [*s. l.*], v. 330, p. 254–263, 2018.

TANG, S.; DU, C.; NIE, T. Inversion Estimation of Soil Organic Matter in Songnen Plain Based on Multispectral Analysis. **Land**, [*s. l.*], v. 11, n. 5, p. 608, 2022.

TAO, Y.; XIONG, S.; SONG, R.; MULLER, J.-P. Towards Streamlined Single-Image Super-Resolution: Demonstration with 10 m Sentinel-2 Colour and 10–60 m Multi-Spectral VNIR and SWIR Bands. **Remote Sensing**, [*s. l.*], v. 13, n. 13, p. 2614, 2021.

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

TIWARI, S. K.; SAHA, S. K.; KUMAR, S. Prediction Modeling and Mapping of Soil Carbon Content Using Artificial Neural Network, Hyperspectral Satellite Data and Field Spectroscopy. **Advances in Remote Sensing**, [*s. l.*], v. 04, n. 01, p. 63–72, 2015.

VANGURI, R.; LANEVE, G.; HOŚCIŁO, A. Mapping forest tree species and its biodiversity using EnMAP hyperspectral data along with Sentinel-2 temporal data: An approach of tree species classification and diversity indices. **Ecological Indicators**, [*s. l.*], v. 167, p. 112671, 2024.

VAŠÁT, R.; KODEŠOVÁ, R.; BORŮVKA, L.; KLEMENT, A.; JAKŠÍK, O.; GHOLIZADEH, A. Consideration of peak parameters derived from continuum-removed spectra to predict extractable nutrients in soils with visible and near-infrared diffuse reflectance spectroscopy (VNIR-DRS). **Geoderma**, [*s. l.*], v. 232-234, p. 208–218, 2014.

VASAVA, H. B.; GUPTA, A.; ARORA, R. K.; DAS, B. S. Assessment of soil texture from spectral reflectance data of bulk soil samples and their dry-sieved aggregate size fractions. **Geoderma**, [*s. l.*], v. 337, p. 914–926, 2019.

VENABLES, W. N.; RIPLEY, B. D. **Modern applied statistics with S.** 4th. ed. New York: Springer, 2002.

VERMOTE, E.; JUSTICE, C.; CLAVERIE, M.; FRANCH, B. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. **Remote Sensing of Environment**, [*s. l.*], v. 185, p. 46–56, 2016.

VILLAMIL, M. B.; LITTLE, J.; NAFZIGER, E. D. Corn residue, tillage, and nitrogen rate effects on soil properties. **Soil and Tillage Research**, [*s. l.*], v. 151, p. 61–66, 2015.

WALKLEY, A.; BLACK, I. A. An examination of the Degtjareff method for determining Soil Organic Matter, and a proposed modification of the Chromic Acid Titration method. **Soil Science**, [*s. l.*], v. 37, n. 1, p. 29–38, 1934.

WANG, D.; ANDERSON, D. W. Direct measurement of organic carbon content in soils by the Leco CR-12 carbon analyzer. **Communications in Soil Science and Plant Analysis**, [*s. l.*], v. 29, n. 1-2, p. 15–21, 1998.

WANG, C.; FENG, M.; YANG, W.; DING, G.; WANG, H.; LI, Z.; SUN, H.; SHI, C. Use of Spectral Character to Evaluate Soil Organic Matter. **Soil Science Society of America Journal**, [*s. l.*], v. 80, n. 4, p. 1078–1088, 2016.

WANG, Z.; MIAO, Z.; YU, X.; HE, F. Vis-NIR spectroscopy coupled with PLSR and multivariate regression models to predict soil salinity under different types of land use. **Infrared Physics & Technology**, [*s. l.*], v. 133, p. 104826, 2023.

WANG, L.; QU, J. J.; HAO, X.; ZHU, Q. Sensitivity studies of the moisture effects on MODIS SWIR reflectance and vegetation water indices. **International journal of remote sensing**, [*s. l.*], v. 29, n. 24, p. 7065–7075, 2008.

WANG, Q.; WANG, W.; HE, X.; ZHENG, Q.; WANG, H.; WU, Y.; ZHONG, Z. Changes in soil properties, X-ray-mineral diffractions and infrared-functional groups in bulk soil and fractions following afforestation of farmland, Northeast China. **Scientific Reports**, [*s. l.*], v. 7, n. 1, 2017.

WANG, Y.; WU, P.; MEI, F.; LING, Y.; QIAO, Y.; LIU, C.; LEGHARI, S. J.; GUAN, X.; WANG, T. Does continuous straw returning keep China farmland soil organic carbon continued increase? A meta-analysis. **Journal of Environmental Management**, [*s. l.*], v. 288, p. 112391, 2021.

WANG, J.; YANG, R.; BAI, Z. Spatial variability and sampling optimization of soil organic carbon and total nitrogen for Minesoils of the Loess Plateau using geostatistics. **Ecological Engineering**, [*s. l.*], v. 82, p. 159–164, 2015.

WANG, X.; ZHANG, Y.; ATKINSON, P. M.; YAO, H. Predicting soil organic carbon content in Spain by combining Landsat TM and ALOS PALSAR images. **International Journal of Applied Earth Observation and Geoinformation**, [*s. l.*], v. 92, p. 102182, 2020.

WEI, Y.; ZHU, X.; LI, C.; CHENG, L.; WANG, L.; ZHAO, G.; JIANG, Y. Monitoring Soil Nitrate Nitrogen Based on Hyperspectral Data in the Apple Orchards. **Agricultural Sciences**, [*s. l.*], v. 08, n. 01, p. 21–32, 2017.

WEST, H.; QUINN, N.; HORSWELL, M.; WHITE, P. Assessing Vegetation Response to Soil Moisture Fluctuation under Extreme Drought Using Sentinel-2. **Water**, [*s. l.*], v. 10, n. 7, p. 838, 2018.

XIAOJU, N.; TONGQIAN, Z.; YANYAN, S. Fossil fuel carbon contamination impacts soil organic carbon estimation in cropland. **CATENA**, [*s. l.*], v. 196, p. 104889, 2021.

XU, D.; LIU, Y.; XU, W.; GUO, X. The Impact of NPV on the Spectral Parameters in the Yellow-Edge, Red-Edge and NIR Shoulder Wavelength Regions in Grasslands. **Remote Sensing**, [*s. l.*], v. 14, n. 13, p. 3031, 2022.

XU, C.; XU, X.; LIU, M.; LIU, W.; YANG, J.; LUO, W.; ZHANG, R.; KIELY, G. Enhancing pedotransfer functions (PTFs) using soil spectral reflectance data for estimating saturated hydraulic conductivity in southwestern China. **CATENA**, [*s. l.*], v. 158, p. 350–356, 2017.

YAN, D.; WANG, Q.; SONG, Z.; FANG, W.; WANG, Q.; LI, Y.; CAO, A. Activation effect of soil available nitrogen, manganese and cobalt after addition of different fumigants. **Environmental Research Communications**, [*s. l.*], v. 4, n. 4, p. 041002, 2022.

YE, Y.; JIANG, Y.; KUANG, L.; HAN, Y.; XU, Z.; GUO, X. Predicting spatial distribution of soil organic carbon and total nitrogen in a typical human impacted area. **Geocarto International**, [*s. l.*], v. 37, n. 15, p. 4465–4482, 2021.

YEOMANS, J. C.; BREMNER, J. M. A rapid and precise method for routine determination of organic carbon in soil. **Communications in Soil Science and Plant Analysis**, [*s. l.*], v. 19, n. 13, p. 1467–1476, 1988.

ZAGAJEWSKI, B.; KLUCZEK, M.; ZDUNEK, K. B.; HOLLAND, D. Sentinel-2 versus PlanetScope Images for Goldenrod Invasive Plant Species Mapping. **Remote Sensing**, [*s. l.*], v. 16, n. 4, p. 636, 2024.

ZENG, R.; ZHANG, J. P.; CAI, K.; GAO, W. C.; PAN, W. J.; JIANG, C. Y.; ZHANG, P. Y.; WU, B. W.; WANG, C. H.; JIN, X. Y.; LI, D. C. How similar is “similar,” or what is the best measure of soil spectral and physiochemical similarity? **PLOS ONE**, [*s. l.*], v. 16, n. 3, p. e0247028, 2021.

ZHANG, L.; GUI-CHENG, Z. Genetic algorithm-based parameter optimization for EO-1 Hyperion remote sensing image classification. **European Journal of Remote Sensing**, [*s. l.*], v. 53, n. 1, p. 124–131, 2020.

ZHANG, K.; SU, Y.; YANG, R. Variation of soil organic carbon, nitrogen, and phosphorus stoichiometry and biogeographic factors across the desert ecosystem of Hexi Corridor, northwestern China. **Journal of Soils and Sediments**, [*s. l.*], v. 19, n. 1, p. 49–57, 2018.

ZHANG, P.; WEI, T.; LI, Y.; WANG, K.; JIA, Z.; HAN, Q.; REN, X. Effects of straw incorporation on the stratification of the soil organic C, total N and C:N ratio in a semiarid region of China. **Soil and Tillage Research**, [*s. l.*], v. 153, p. 28–35, 2015.

ZHOU, T.; GENG, Y.; JI, C.; XU, X.; WANG, H.; PAN, J.; BUMBERGER, J.; HAASE, D.; LAUSCH, A. Prediction of soil organic carbon and the C:N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images. **Science of the Total Environment**, [*s. l.*], v. 755, p. 142661–142661, 2021.

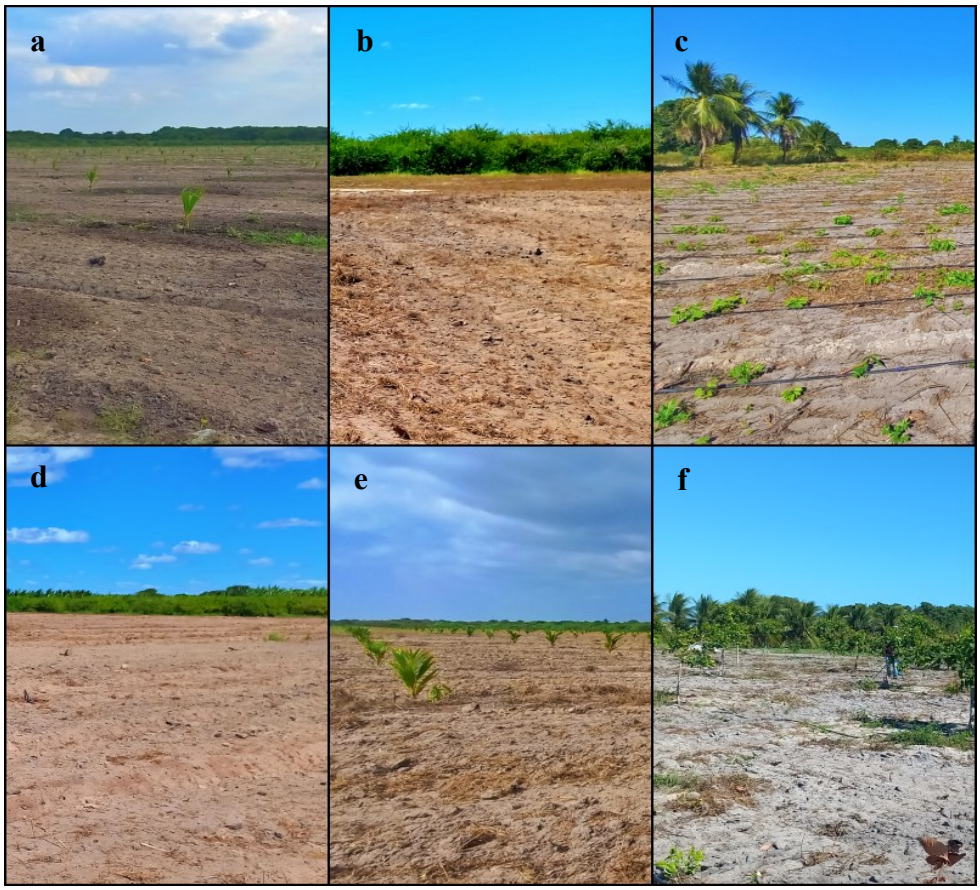
ZHOU, K.-F.; WANG, S.-S. Spectral properties of weathered and fresh rock surfaces in the Xiemisitai metallogenic belt, NW Xinjiang, China. **Open Geosciences**, [*s. l.*], v. 9, n. 1, 2017.

ZHU, Q.; LIAO, K.; LAI, X.; LV, L. Scale-dependent effects of environmental factors on soil organic carbon, soil nutrients and stoichiometry under two contrasting land-use types. **Soil Use and Management**, [*s. l.*], 2021.

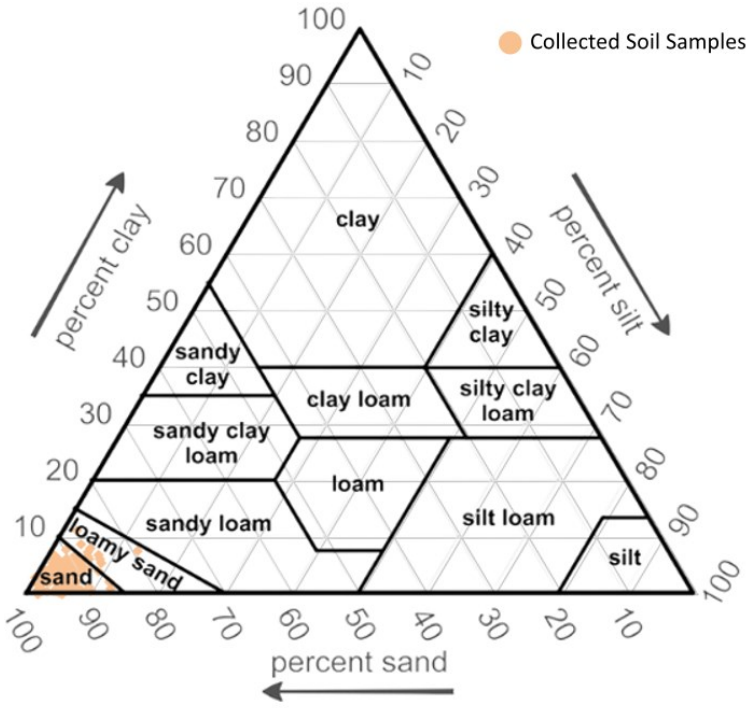
ZINN, Y. L.; MARRENJO, G. J.; SILVA, C. A. Soil C:N ratios are unresponsive to land use change in Brazil: A comparative analysis. **Agriculture, Ecosystems & Environment**, [*s. l.*], v. 255, p. 62–72, 2018.

APPENDIX A – RECORDS OF THE STUDY AREAS

Figure A1. Representative examples of areas where spectral analyses and soil sampling were conducted in DIBAU. a) coconut cultivation at early growth stage with surface burn residues; b) area undergoing initial soil preparation with leguminous plant residues; c) cassava crop with residual straw; d) area in preparation stage showing presence of charcoal; e) dwarf coconut plantation in early development with inter-row legume residues; f) cashew cultivation area with grass straw and burn residues.



APPENDIX B – USDA TEXTURAL TRIANGLE FOR SOIL SAMPLES FROM DIBAU

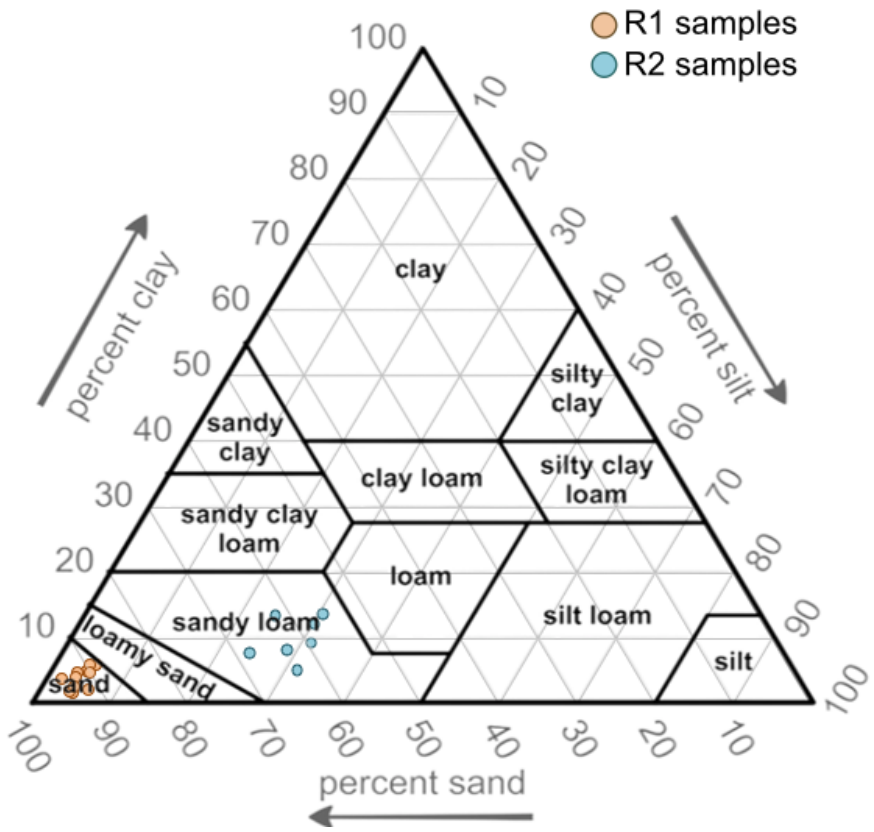


APPENDIX C – LINEAR PREDICTION MODELS

Table C1. Model parameters for Soil Organic Carbon (SOC) and Total Nitrogen (TN) prediction using Approaches I and II, based on reflectance data from the EnMAP sensor and the FieldSpec spectroradiometer *in situ*. The values in parentheses represent the spectral bands selected for each model.

Sensor	Linear Regression Model ($p < 0.05$)	Range (g kg^{-1})	Adj. R^2
EnMAP	SOC (g kg^{-1}) = $-17.35 + 2246.52(\rho_{2191nm}) - 2329.28(\rho_{1175nm}) + 7645.31(\rho_{2207nm}) - 5698.19(\rho_{2182nm}) + 1172.08(\rho_{2174nm}) - 4830.55(\rho_{1128nm}) - 5327.66(\rho_{2282nm}) - 20.65(\rho_{2257nm}) + 720.35(\rho_{2148nm}) - 6657.19(\rho_{2216nm}) + 2827.05(\rho_{2265nm}) + 46.92(\rho_{2306nm}) - 2715.03(\rho_{1234nm}) - 4871.13(\rho_{1211nm}) - 4508.90(\rho_{1258nm}) + 5612.23(\rho_{1139nm}) - 677.33(\rho_{2337nm}) + 2827.05(\rho_{1036nm}) + 8301.07(\rho_{1223nm}) + 273.96(\rho_{2407nm}) - 2348.20(\rho_{1025nm}) + 4956.98(\rho_{1246nm}) + 1385.94(\rho_{2298nm}) + 452.39(\rho_{2353nm})$	(Approach I) 2.67 – 23.6	0.75
FieldSpec *in situ*	SOC (g kg^{-1}) = $21.84 - 2327.67(\rho_{2077nm}) - 1795.10(\rho_{911nm}) + 1700.70(\rho_{895nm}) + 1749.52(\rho_{2095nm}) + 358.64(\rho_{1891nm}) - 3024.26(\rho_{2113nm}) + 2636.88(\rho_{2438nm}) - 2256.84 (\rho_{2119nm}) + 1721.84(\rho_{1534nm}) - 2202.13(\rho_{2439nm}) - 1641.38(\rho_{1557nm}) - 1418.83(\rho_{2349nm}) + 1362.33(\rho_{2406nm}) - 723.22(\rho_{2392nm}) - 637.36(\rho_{1927nm}) + 399.99(\rho_{1949nm})$	(Approach I) 2.67 – 23.6	0.73
EnMAP	SOC (g kg^{-1}) = $-13.13 + 2519.67(\rho_{2191\ nm}) - 264.12\ (\rho_{1175\ nm}) + 697.20(\rho_{2207nm}) - 1571.42(\rho_{2182nm}) - 405.5\ (\rho_{1128nm}) - 6187.92(\rho_{2282nm}) + 597.98(\rho_{2148nm}) - 2125.082(\rho_{2216nm}) + 5253.35(\rho_{2265nm}) + 1820.77(\rho_{1211nm}) - 3730.38\ (\rho_{1258nm}) + 683.54(\rho_{1139nm}) - 99.05(\rho_{2337nm}) - 232.68(\rho_{1036nm}) - 2738.31(\rho_{1223nm}) - 107.530(\rho_{2407nm}) + 353.61(\rho_{1025nm}) + 4557.15(\rho_{1246nm}) + 1000.09(\rho_{2298nm})$	(Approach II) 11.6 – 23.6	0.91
FieldSpec *in situ*	SOC (g kg^{-1}) = $20.05 - 2445.60(\rho_{2077nm}) - 3539.91(\rho_{911nm}) + 3340.0(\rho_{895nm}) + 1195.51(\rho_{2095nm}) + 403.93(\rho_{1891nm}) + 1517.54(\rho_{2113nm}) + 2152.44(\rho_{2438nm}) - 111.57(\rho_{2119nm}) + 3644.82(\rho_{1534nm}) - 1917.32(\rho_{2439nm}) - 3348.71(\rho_{1557nm}) - 1712.0(\rho_{2349nm}) + 1052.79(\rho_{2406nm}) - 269.96(\rho_{1927nm})$	(Approach II) 9.2 – 11.6 & 16.2 – 23.2	0.83
EnMAP	TN (g kg^{-1}) = $0.905 - 18.668(\rho_{936nm}) + 25.445(\rho_{941nm}) - 9.926(\rho_{952nm}) - 13.845(\rho_{961nm}) + 11.087(\rho_{968nm}) + 42.428(\rho_{1128nm}) - 97.091(\rho_{1139nm}) + 58.663(\rho_{1151nm}) - 9.793(\rho_{2361nm}) + 7.519(\rho_{2376nm}) - 19.804(\rho_{2415nm}) + 34.091(\rho_{2422nm}) - 12.841(\rho_{2445nm})$	(Approach I) 0.197 – 0.880	0.61
FieldSpec *in situ*	TN (g kg^{-1}) = $0.579 - 43.434(\rho_{577nm}) + 77.905(\rho_{593nm}) - 47.791(\rho_{1910nm}) + 64.812(\rho_{1953nm}) - 35.688(\rho_{614nm}) - 58.138(\rho_{1944nm}) + 11.398(\rho_{1878nm}) + 46.685(\rho_{1918nm}) - 15.882(\rho_{1988nm}) - 16.112(\rho_{1319nm}) + 14.556(\rho_{1337nm})$	(Approach I) 0.197 – 0.880	0.68
EnMAP	TN (g kg^{-1}) = $1.149 - 8.596(\rho_{936nm}) + 6.240(\rho_{941nm}) + 39.962(\rho_{1128nm}) - 61.805(\rho_{1139nm}) + 23.505(\rho_{1151nm}) - 6.050(\rho_{2361nm}) + 2.103(\rho_{2415nm}) + 4.465(\rho_{2422nm}) - 2.079(\rho_{2445nm})$	(Approach II) 0.550 –0.880	0.82
FieldSpec *in situ*	TN (g kg^{-1}) = $0.751 + 40.269\ (\rho_{577nm}) - 63.620(\rho_{593nm}) + 27.489(\rho_{1910nm}) + 65.692(\rho_{1953nm}) + 27.915(\rho_{614nm}) - 36.397(\rho_{1944nm}) - 7.993(\rho_{1878nm}) - 55.313(\rho_{1918nm}) + 3.809(\rho_{1988nm})$	(Approach II) 0.550 –0.880	0.85

APPENDIX D – USDA TEXTURAL TRIANGLE FOR SOIL SAMPLES FROM DIBAU (R1) AND CCV (R2)



APPENDIX E – MULTIPLE LINEAR REGRESSION (MLR) MODELS USING ASD FIELDSPEC IN DARKROOM

Attribute	Dataset	Spectral data	Equation	Adj. R^2
Soil Organic Carbon (%)	R1	Ref	$3.459 - 61.847(\rho_{2440nm}) - 35.268(\rho_{633nm}) + 15.486(\rho_{1927nm}) - 16.990(\rho_{1899nm}) + 114.868(\rho_{1996nm}) - 73.883(\rho_{435nm}) - 47.425(\rho_{483nm}) + 38.030(\rho_{681nm}) + 125.481(\rho_{2393nm}) + 58.236(\rho_{2433nm}) - 30.170(\rho_{1814nm}) - 42.584(\rho_{433nm}) + 212.209(\rho_{2192nm}) + 62.045(\rho_{2414nm}) - 140.672(\rho_{447nm}) - 68.487(\rho_{2357nm}) - 151.569(\rho_{2179nm}) + 108.128(\rho_{2437nm}) - 106.592(\rho_{2410nm}) + 124.557(\rho_{441nm}) + 177.754(\rho_{445nm}) - 79.615(\rho_{1947nm}) - 184.616(\rho_{2402nm})$	0.64
		CR	$828.082 - 171.433(\rho_{1522nm}) - 291.038(\rho_{1286nm}) + 376.268(\rho_{1626nm}) - 190.789(\rho_{1648nm}) - 245.418(\rho_{1636nm}) - 197.589(\rho_{1347nm}) + 102.656(\rho_{2325nm}) + 108.259(\rho_{1191nm}) - 266.692(\rho_{1216nm}) - 79.788(\rho_{2303nm}) + 178.388(\rho_{1555nm}) - 138.325(\rho_{1544nm}) - 177.975(\rho_{1268nm}) - 290.176(\rho_{1655nm}) + 71.780(\rho_{1194nm}) + 143.551(\rho_{1329nm}) + 157.330(\rho_{2106nm}) - 79.784(\rho_{2100nm}) + 52.109(\rho_{1208nm}) + 66.312(\rho_{1221nm}) + 117.302(\rho_{1463nm}) - 73.055(\rho_{1465nm})$	0.71
		1^{st} derivative	$2.736 - 122.052(\rho_{1915nm}) - 5413.026(\rho_{818nm}) - 1599.414(\rho_{1707nm}) - 1292.692(\rho_{2092nm}) + 1441.557(\rho_{2192nm}) + 3022.852(\rho_{802nm}) - 973.739(\rho_{2408nm}) + 1026.890(\rho_{2313nm}) + 2419.800(\rho_{772nm}) - 2008.228(\rho_{845nm}) - 1527.214(\rho_{2065nm}) - 1652.673(\rho_{865nm}) + 399.943(\rho_{2273nm}) + 2314.095(\rho_{952nm}) - 295.659(\rho_{1026nm}) - 1590.079(\rho_{953nm}) + 1632.657(\rho_{893nm}) - 679.860(\rho_{2252nm}) + 602.625(\rho_{2037nm}) + 290.657(\rho_{1620nm}) + 720.578(\rho_{721nm}) - 436.278(\rho_{581nm}) + 346.403(\rho_{1602nm}) + 733.596(\rho_{1647nm}) - 132.799(\rho_{1688nm})$	0.8
	R2	Ref	$0.615 - 102.545(\rho_{1003nm}) - 14.071(\rho_{534nm}) + 18.581(\rho_{1394nm}) - 9.553(\rho_{1925nm}) + 21.592(\rho_{1801nm}) - 27.810(\rho_{2132nm}) - 6.496(\rho_{782nm}) - 48.767(\rho_{442nm}) + 7.965(\rho_{1901nm}) + 83.557(\rho_{470nm}) - 9.047(\rho_{2206nm}) + 66.865(\rho_{436nm}) + 78.422(\rho_{463nm}) + 12.308(\rho_{2341nm}) + 22.475(\rho_{445nm}) + 91.629(\rho_{993nm}) - 64.412(\rho_{467nm}) - 55.225(\rho_{451nm}) - 51.606(\rho_{494nm})$	0.85
		CR	$-193.184 - 103.074(\rho_{2108nm}) + 41.596(\rho_{1856nm}) + 49.567(\rho_{2322nm}) + 225.150(\rho_{2123nm}) - 42.829(\rho_{760nm}) + 70.169(\rho_{1760nm}) + 3.275(\rho_{442nm}) - 44.439(\rho_{2351nm}) - 58.320(\rho_{1727nm}) - 316.961(\rho_{1261nm}) - 192.147(\rho_{1283nm}) + 286.233(\rho_{1590nm}) - 122.431(\rho_{1558nm}) + 208.129(\rho_{1650nm}) + 243.005(\rho_{1277nm}) + 24.496(\rho_{1514nm}) + 6.788(\rho_{1265nm}) - 8.664(\rho_{852nm}) - 33.189(\rho_{1623nm}) - 75.915(\rho_{1551nm}) - 113.539(\rho_{1649nm}) + 77.291(\rho_{826nm}) + 70.797(\rho_{1681nm})$	0.9
		1^{st} derivative	$1.213 - 650.002(\rho_{599nm}) - 1088.239(\rho_{1906nm}) - 342.145(\rho_{570nm}) + 830.420(\rho_{1056nm}) - 1427.657(\rho_{2016nm}) + 1077.287(\rho_{1962nm}) - 838.816(\rho_{1730nm}) - 1167.123(\rho_{1975nm}) + 1133.844(\rho_{1031nm}) - 865.178(\rho_{1628nm}) + 873.966(\rho_{1739nm}) - 864.282(\rho_{2128nm}) - 804.893(\rho_{1599nm}) + 971.116(\rho_{967nm}) - 1140.037(\rho_{1983nm}) - 962.676(\rho_{1165nm}) + 532.710(\rho_{1653nm}) + 561.395(\rho_{1270nm}) - 631.504(\rho_{584nm}) - 566.553(\rho_{1085nm}) + 617.064(\rho_{1131nm})$	0.9
	R1&R2	Ref	$1.329 - 45.317(\rho_{2440nm}) + 43.848(\rho_{1269nm}) + 39.075(\rho_{2207nm}) - 15.160(\rho_{1840nm}) - 13.565(\rho_{1413nm}) - 39.576(\rho_{2147nm}) + 14.044(\rho_{1903nm}) - 72.191(\rho_{679nm}) + 13.496(\rho_{2227nm}) - 63.077(\rho_{435nm}) + 84.516(\rho_{737nm}) + 41.805(\rho_{445nm}) - 46.922(\rho_{1001nm}) + 50.411(\rho_{497nm}) + 30.901(\rho_{2297nm}) - 23.155(\rho_{577nm})$	0.6
		CR	$280.504 - 137.252(\rho_{1929nm}) - 1134.819(\rho_{1280nm}) + 129.928(\rho_{1934nm}) + 45.584(\rho_{1875nm}) - 51.316(\rho_{2440nm}) + 148.199(\rho_{1620nm}) - 99.936(\rho_{2098nm}) - 89.876(\rho_{2395nm}) + 220.593(\rho_{2113nm}) - 134.254(\rho_{2092nm}) - 181.212(\rho_{1703nm}) + 94.526(\rho_{1760nm}) + 15.137(\rho_{2316nm}) + 1208.829(\rho_{1279nm}) - 95.128(\rho_{1740nm}) - 477.684(\rho_{1278nm}) + 31.827(\rho_{2430nm}) + 191.946(\rho_{1571nm}) + 56.514(\rho_{2399nm}) + 5.716(\rho_{783nm}) - 110.207(\rho_{1582nm}) + 96.643(\rho_{1690nm}) - 12.569(\rho_{999nm}) - 2.131(\rho_{663nm})$	0.73

Continue...

		1st derivative	$2.047 - 857.571(\rho_{2251nm}) - 1721.245(\rho_{1659nm}) - 790.852(\rho_{1705nm}) - 1415.690(\rho_{600nm}) + 688.082(\rho_{567nm}) - 898.081(\rho_{2066nm}) - 392.186(\rho_{999nm}) + 885.485(\rho_{1603nm}) - 739.228(\rho_{1628nm}) - 619.490(\rho_{1542nm}) + 1486.077(\rho_{1340nm}) - 1347.424(\rho_{558nm}) + 354.587(\rho_{436nm}) - 711.001(\rho_{1286nm}) + 1112.103(\rho_{1591nm}) - 785.685(\rho_{2332nm}) + 580.489(\rho_{2184nm}) - 1292.967(\rho_{477nm}) - 801.192(\rho_{1636nm}) + 957.747(\rho_{573nm}) - 1484.535(\rho_{1664nm}) - 369.581(\rho_{2255nm}) + 605.007(\rho_{2098nm}) - 2716.917(\rho_{818nm}) + 1154.400(\rho_{708nm}) + 1249.201(\rho_{1077nm}) + 983.626(\rho_{1864nm}) + 980.400(\rho_{1022nm}) + 1698.578(\rho_{1252nm}) - 533.286(\rho_{1283nm}) - 1114.742(\rho_{699nm}) + 755.629(\rho_{696nm}) + 912.982(\rho_{587nm}) - 363.352(\rho_{2367nm}) - 1202.289(\rho_{1996nm}) + 1017.815(\rho_{1939nm}) - 1069.408(\rho_{604nm}) + 563.804(\rho_{644nm}) - 329.007(\rho_{1688nm}) + 568.070(\rho_{1781nm}) + 884.677(\rho_{811nm}) + 667.289(\rho_{1772nm}) - 805.975(\rho_{963nm})$	0.89
Total Nitrogen (g/kg)	R1	Ref	$0.392 - 31.627(\rho_{1167nm}) - 28.620(\rho_{430nm}) + 31.787(\rho_{633nm}) - 12.430(\rho_{1927nm}) - 19.057(\rho_{1844nm}) - 55.350(\rho_{545nm}) + 22.781(\rho_{873nm}) + 7.509(\rho_{1412nm}) - 3.765(\rho_{2239nm}) - 9.990(\rho_{435nm}) + 30.559(\rho_{1730nm}) - 14.920(\rho_{681nm}) + 56.846(\rho_{2393nm}) + 72.741(\rho_{443nm}) + 33.468(\rho_{2192nm}) - 30.191(\rho_{436nm}) + 10.882(\rho_{2357nm}) - 53.990(\rho_{2410nm}) + 39.770(\rho_{2424nm}) + 30.147(\rho_{432nm}) + 16.855(\rho_{1947nm}) + 2.689(\rho_{2368nm}) - 93.081(\rho_{2402nm})$	0.8
		CR	$124.982 - 106.938(\rho_{2077nm}) + 115.224(\rho_{1242nm}) - 92.534(\rho_{1644nm}) - 19.900(\rho_{2432nm}) + 122.736(\rho_{2124nm}) + 133.007(\rho_{2089nm}) - 72.883(\rho_{1639nm}) - 40.531(\rho_{2413nm}) - 85.247(\rho_{1649nm}) - 57.297(\rho_{1690nm}) - 31.557(\rho_{1275nm}) - 27.475(\rho_{1608nm}) - 57.901(\rho_{1864nm}) - 44.161(\rho_{2426nm}) + 42.844(\rho_{1550nm}) + 13.013(\rho_{2423nm}) + 57.930(\rho_{1633nm}) - 55.455(\rho_{1579nm}) + 73.760(\rho_{1613nm}) - 16.155(\rho_{1612nm}) - 42.186(\rho_{1289nm}) + 1.900(\rho_{1206nm}) + 13.979(\rho_{1303nm}) + 51.042(\rho_{1857nm})$	0.81
		1st derivative	$0.649 + 361.734(\rho_{1302nm}) + 611.685(\rho_{1445nm}) + 530.090(\rho_{1745nm}) - 357.812(\rho_{2240nm}) - 365.730(\rho_{465nm}) + 482.592(\rho_{1502nm}) - 503.008(\rho_{978nm}) + 402.366(\rho_{537nm}) - 474.515(\rho_{534nm}) - 376.660(\rho_{1916nm}) - 894.168(\rho_{868nm}) - 552.215(\rho_{2019nm}) + 541.243(\rho_{879nm}) + 212.257(\rho_{2054nm}) - 188.567(\rho_{1672nm}) + 218.065(\rho_{1279nm}) + 192.203(\rho_{1645nm}) - 377.161(\rho_{1527nm}) - 692.962(\rho_{1244nm}) - 130.439(\rho_{2282nm}) + 93.714(\rho_{1288nm}) + 262.337(\rho_{997nm}) - 220.004(\rho_{2397nm}) - 192.771(\rho_{1680nm}) + 101.082(\rho_{1821nm}) + 458.659(\rho_{1179nm})$	0.88
	R2	Ref	$0.632 - 45.941(\rho_{1003nm}) + 2.324(\rho_{534nm}) + 19.991(\rho_{1394nm}) + 5.140(\rho_{1925nm}) - 10.039(\rho_{782nm}) + 18.079(\rho_{442nm}) + 0.906(\rho_{1901nm}) + 18.808(\rho_{430nm}) + 20.295(\rho_{470nm}) - 12.958(\rho_{2206nm}) + 35.225(\rho 463nm) - 3.355(\rho_{2341nm}) + 42.480(\rho_{993nm}) - 36.311(\rho_{455nm}) - 47.282(\rho_{494nm})$	0.85
		CR	$460.029 - 117.634(\rho_{1933nm}) + 107.985(\rho_{1960nm}) + 91.905(\rho_{1431nm}) - 191.843(\rho_{2240nm}) - 332.070(\rho 1560nm) + 160.241(\rho_{2244nm}) + 70.180(\rho_{1665nm}) - 56.319(\rho_{776nm}) + 48.753(\rho_{1862nm}) + 261.114(\rho_{837nm}) - 407.284(\rho_{1266nm}) - 217.712(\rho_{1282nm}) - 100.650(\rho_{1711nm}) - 169.337(\rho_{848nm}) + 40.179(\rho_{1361nm}) + 191.403(\rho_{2122nm}) - 173.452(\rho_{1634nm}) + 61.104(\rho_{1370nm}) + 126.413(\rho_{1615nm}) - 28.263(\rho_{2097nm}) - 24.641(\rho_{2388nm}) + 60.850(\rho_{1538nm}) - 68.548(\rho_{1643nm}) + 206.215(\rho_{1264nm})$	0.91
		1st derivative	$1.282 - 1088.527(\rho_{465nm}) - 505.927(\rho_{1892nm}) - 1034.352(\rho_{605nm}) - 772.335(\rho_{1927nm}) - 257.019(\rho_{1001nm}) - 1281.967(\rho_{1999nm}) + 1399.386(\rho_{953nm}) + 451.599(\rho_{543nm}) - 620.066(\rho_{1599nm}) - 655.261(\rho_{1874nm}) - 580.388(\rho_{1915nm}) - 905.635(\rho_{1949nm}) - 324.176(\rho_{2409nm}) - 725.779(\rho_{529nm}) - 845.023(\rho_{1943nm}) + 825.097(\rho_{1436nm}) + 283.810(\rho_{432nm}) + 316.019(\rho_{1575nm}) + 498.900(\rho_{1457nm}) + 326.616(\rho_{573nm}) - 287.278(\rho_{1503nm})$	0.91
	R1&R2	Ref	$0.851 - 24.364(\rho_{2440nm}) - 27.019(\rho_{1269nm}) + 15.666(\rho_{2207nm}) - 38.146(\rho_{536nm}) + 5.296(\rho_{1939nm}) + 16.354(\rho_{1413nm}) - 11.504(\rho_{679nm}) - 20.440(\rho_{2227nm}) - 70.702(\rho_{476nm}) - 34.112(\rho_{435nm}) + 10.178(\rho_{1798nm}) - 39.635(\rho_{431nm}) + 18.208(\rho_{450nm}) - 24.513(\rho_{1001nm}) - 74.150(\rho_{442nm}) + 16.662(\rho_{468nm}) + 36.934(\rho_{436nm}) + 40.564(\rho_{997nm}) + 45.755(\rho_{497nm}) + 76.956(\rho_{448nm}) + 13.815(\rho_{2297nm}) + 52.029(\rho_{441nm}) + 15.704(\rho_{577nm})$	0.83

	CR	$265.305 - 111.873(\rho_{1688nm}) - 82.604(\rho_{1815nm}) - 132.306(\rho_{2118nm}) - 33.539(\rho_{1000nm}) + 194.005(\rho_{1538nm}) - 77.238(\rho_{1581nm}) + 96.194(\rho_{2122nm}) - 252.147(\rho_{1282nm}) - 62.583(\rho_{2036nm}) + 4.421(\rho_{803nm}) + 59.891(\rho_{1856nm}) + 70.647(\rho_{1795nm}) + 53.637(\rho_{1304nm}) + 122.896(\rho_{1249nm}) + 133.687(\rho_{1960nm}) - 135.770(\rho_{1964nm}) - 40.816(\rho_{2404nm}) - 127.994(\rho_{1553nm}) + 124.003(\rho_{1500nm}) + 115.533(\rho_{1652nm}) + 83.887(\rho_{1977nm}) - 158.028(\rho_{1567nm}) - 91.408(\rho_{2226nm}) + 63.522(\rho_{1075nm}) - 155.359(\rho_{1261nm}) + 120.194(\rho_{1590nm}) + 76.777(\rho_{2229nm}) - 65.297(\rho_{1932nm}) + 46.290(\rho_{1664nm}) + 22.805(\rho_{1874nm}) - 108.554(\rho_{1541nm}) + 44.845(\rho_{2322nm}) - 39.843(\rho_{2326nm}) - 76.729(\rho_{1783nm}) + 56.131(\rho_{1711nm}) - 45.031(\rho_{1170nm}) + 41.610(\rho_{1787nm})$	0.89
	1^{st} derivative	$1.035 - 742.217(\rho_{1398nm}) - 318.934(\rho_{463nm}) - 553.063(\rho_{2040nm}) - 292.647(\rho_{999nm}) - 547.871(\rho_{2331nm}) - 711.342(\rho_{2120nm}) - 915.834(\rho_{617nm}) - 163.520(\rho_{2396nm}) - 395.541(\rho_{2047nm}) - 371.903(\rho_{2362nm}) + 920.191(\rho_{1632nm}) - 772.867(\rho_{1841nm}) + 936.248(\rho_{1263nm}) - 253.113(\rho_{1580nm}) + 669.168(\rho_{1604nm}) - 497.899(\rho_{2190nm}) - 429.340(\rho_{2045nm}) - 736.093(\rho_{510nm}) + 506.140(\rho_{434nm}) + 337.349(\rho_{1290nm}) + 548.809(\rho_{490nm}) - 525.963(\rho_{1621nm}) + 571.442(\rho_{615nm}) + 882.808(\rho_{1238nm}) - 668.640(\rho_{1479nm}) - 590.048(\rho_{2254nm}) - 390.551(\rho_{1311nm}) + 565.079(\rho_{1592nm}) - 400.830(\rho_{2325nm}) + 616.660(\rho_{1400nm}) - 417.491(\rho_{1853nm}) + 315.264(\rho_{1087nm}) - 540.938(\rho_{1194nm}) - 603.273(\rho_{529nm}) - 84.179(\rho_{1803nm}) + 292.823(\rho_{2066nm}) - 309.769(\rho_{2354nm}) - 575.154(\rho_{1476nm}) + 307.957(\rho_{448nm}) - 445.459(\rho_{1132nm}) + 393.225(\rho_{1119nm}) + 476.134(\rho_{1526nm}) - 454.612(\rho_{1487nm}) + 328.441(\rho_{2118nm})$	0.9

Continue...

Ref.: Reflectance; CR: Continuum-removed

APPENDIX F – PARTIAL LEAST SQUARE REGRESSION (PLSR) MODELS USING ASD FIELDSPEC IN DARKROOM

Attribute	Dataset	Spectral data	Equation	Adj. R^2
Soil Organic Carbon (%)	R1	Ref	$2.492 - 34.689(\rho_{2440nm}) - 31.474(\rho_{633nm}) - 30.209(\rho_{1927nm}) + 7.815(\rho_{1899nm}) + 90.854(\rho_{1996nm}) - 18.878(\rho_{435nm}) - 90.086(\rho_{483nm}) + 38.054(\rho_{681nm}) + 12.452(\rho_{2393nm}) + 75.494(\rho_{2433nm}) - 38.576(\rho_{1814nm}) - 21.887(\rho_{433nm}) + 177.340(\rho_{2192nm}) + 66.978(\rho_{2414nm}) - 8.833(\rho_{447nm}) - 33.592(\rho_{2357nm}) - 98.010(\rho_{2179nm}) + 23.679(\rho_{2437nm}) - 35.683(\rho_{2410nm}) + 76.622(\rho_{441nm}) + 62.525(\rho_{445nm}) - 34.825(\rho_{1947nm}) - 154.448(\rho_{2402nm})$	0.66
		CR	$659.250 - 172.122(\rho_{1522nm}) - 304.500(\rho_{1286nm}) + 362.296(\rho_{1626nm}) - 193.936(\rho_{1648nm}) - 217.111(\rho_{1636nm}) - 160.133(\rho_{1347nm}) + 73.291(\rho_{2325nm}) + 135.048(\rho_{1191nm}) - 203.827(\rho_{1216nm}) - 50.884(\rho_{2303nm}) + 237.112(\rho_{1555nm}) - 129.505(\rho_{1544nm}) - 143.359(\rho_{1268nm}) - 275.588(\rho_{1655nm}) + 53.753(\rho_{1194nm}) + 100.332(\rho_{1329nm}) + 189.197(\rho_{2106nm}) - 105.593(\rho_{2100nm}) + 39.659(\rho_{1208nm}) + 85.885(\rho_{1221nm}) + 82.665(\rho_{1463nm}) - 61.831(\rho_{1465nm})$	0.77
		1st derivative	$2.410 - 141.122(\rho_{1915nm}) - 4408.478(\rho_{818nm}) - 1619.715(\rho_{1707nm}) - 1723.384(\rho_{2092nm}) + 1285.933(\rho_{2192nm}) + 1748.544(\rho_{802nm}) - 864.174(\rho_{2408nm}) + 903.357(\rho_{2313nm}) + 2553.684(\rho_{772nm}) - 2101.833(\rho_{845nm}) - 1356.814(\rho_{2065nm}) - 1353.990(\rho_{865nm}) + 535.501(\rho_{2273nm}) + 1178.420(\rho_{952nm}) - 102.208(\rho_{1026nm}) - 223.671(\rho_{953nm}) + 1188.576(\rho_{893nm}) - 1006.473(\rho_{2252nm}) + 968.303(\rho_{2037nm}) + 679.970(\rho_{1620nm}) + 933.066(\rho_{721nm}) - 678.082(\rho_{581nm}) + 77.049(\rho_{1602nm}) + 915.384(\rho_{1647nm}) + 326.245(\rho_{1688nm})$	0.81
	R2	Ref	$0.310 + 3.170(\rho_{1003nm}) - 5.442(\rho_{534nm}) + 13.120(\rho_{1394nm}) - 4.492(\rho_{1925nm}) + 7.547(\rho_{1801nm}) - 7.066(\rho_{2132nm}) - 18.791(\rho_{782nm}) + 3.387(\rho_{442nm}) + 2.333(\rho_{1901nm}) + 2.123(\rho_{470nm}) - 0.007(\rho_{2206nm}) + 4.179(\rho_{436nm}) + 2.618(\rho_{2341nm}) - 5.952(\rho_{445nm}) + 3.582(\rho_{993nm}) + 3.162(\rho_{467nm}) + 1.893(\rho_{451nm}) + 2.977(\rho_{494nm})$	0.79
		CR	$-56.641 - 48.213(\rho_{2108nm}) + 35.901(\rho_{1856nm}) + 44.405(\rho_{2322nm}) + 52.944(\rho_{2123nm}) - 47.519(\rho_{760nm}) + 75.264(\rho_{1760nm}) + 3.600(\rho_{442nm}) - 36.715(\rho_{2351nm}) - 120.345(\rho_{1727nm}) - 51.950(\rho_{1261nm}) - 24.435(\rho_{1283nm}) + 55.201(\rho_{1590nm}) - 54.712(\rho_{1558nm}) + 62.100(\rho_{1650nm}) + 23.295(\rho_{1277nm}) - 11.864(\rho_{1514nm}) + 10.529(\rho_{1265nm}) - 32.887(\rho_{852nm}) - 19.589(\rho_{1623nm}) - 28.041(\rho_{1551nm}) + 43.371(\rho_{1649nm}) + 100.307(\rho_{826nm}) + 27.803(\rho_{1681nm})$	0.87
		1st derivative	$1.173 - 551.900(\rho_{599nm}) - 1011.856(\rho_{1906nm}) - 423.091(\rho_{570nm}) + 863.895(\rho_{1056nm}) - 1507.656(\rho_{2016nm}) + 1166.545(\rho_{1962nm}) - 925.538(\rho_{1730nm}) - 1071.087(\rho_{1975nm}) + 1161.392(\rho_{1031nm}) - 826.377(\rho_{1628nm}) + 783.421(\rho_{1739nm}) - 887.592(\rho_{2128nm}) - 750.479(\rho_{1599nm}) + 1002.296(\rho_{967nm}) - 1064.575(\rho_{1983nm}) - 830.232(\rho_{1165nm}) + 502.356(\rho_{1653nm}) + 575.199(\rho_{1270nm}) - 634.634(\rho_{584nm}) - 483.420(\rho_{1085nm}) + 479.362(\rho_{1131nm})$	0.91
	R1&R2	Ref	$1.054 - 65.422(\rho_{2440nm}) + 38.904(\rho_{1269nm}) + 48.799(\rho_{2207nm}) - 26.665(\rho_{1840nm}) - 19.295(\rho_{1413nm}) - 38.222(\rho_{2147nm}) + 20.890(\rho_{1903nm}) - 34.891(\rho_{679nm}) + 38.571(\rho_{2227nm}) - 24.043(\rho_{435nm}) + 37.320(\rho_{737nm}) + 25.347(\rho_{445nm}) - 27.517(\rho_{1001nm}) + 13.470(\rho_{497nm}) + 21.670(\rho_{2297nm}) - 7.465(\rho_{577nm})$	0.57
		CR	$286.584 - 143.332(\rho_{1929nm}) - 186.189(\rho_{1280nm}) + 136.214(\rho_{1934nm}) + 46.627(\rho_{1875nm}) + 44.188(\rho_{2440nm}) - 71.812(\rho_{1620nm}) + 97.059(\rho_{2098nm}) - 104.149(\rho_{2395nm}) - 189.802(\rho_{2113nm}) - 126.214(\rho_{2092nm}) - 153.536(\rho_{1703nm}) + 111.242(\rho_{1760nm}) + 19.357(\rho_{2316nm}) - 77.243(\rho_{1279nm}) - 112.580(\rho_{1740nm}) - 40.679(\rho_{1278nm}) + 36.837(\rho_{2430nm}) + 142.222(\rho_{1571nm}) + 65.477(\rho_{2399nm}) + 6.571(\rho_{783nm}) - 92.636(\rho_{1582nm}) + 78.008(\rho_{1690nm}) - 11.044(\rho_{999nm}) - 2.203(\rho_{663nm})$	0.70

Continue...

		1st derivative	$2.143 - 986.614(\rho_{2251nm}) - 1793.863(\rho_{1659nm}) - 950.781(\rho_{1705nm}) - 1424.268(\rho_{600nm}) + 866.512(\rho_{567nm}) - 794.216(\rho_{2066nm}) - 473.965(\rho_{999nm}) + 1221.544(\rho_{1603nm}) - 993.700(\rho_{1628nm}) - 554.430(\rho_{1542nm}) + 1704.992(\rho_{1340nm}) - 1467.189(\rho_{558nm}) + 387.543(\rho_{436nm}) - 689.202(\rho_{1286nm}) + 1499.068(\rho_{1591nm}) - 921.521(\rho_{2332nm}) + 669.516(\rho_{2184nm}) - 1604.705(\rho_{477nm}) - 714.240(\rho_{1636nm}) + 728.949(\rho_{573nm}) - 1242.819(\rho_{1664nm}) - 200.248(\rho_{2255nm}) + 670.799(\rho_{2098nm}) - 1850.521(\rho_{818nm}) + 704.940(\rho_{708nm}) + 1237.452(\rho_{1077nm}) + 817.340(\rho_{1864nm}) + 830.683(\rho_{1022nm}) + 1156.192(\rho_{1252nm}) - 704.585(\rho_{1283nm}) - 943.364(\rho_{699nm}) + 775.897(\rho_{696nm}) + 891.103(\rho_{587nm}) - 576.586(\rho_{2367nm}) - 1306.040(\rho_{1996nm}) + 800.248(\rho_{1939nm}) - 494.166(\rho_{604nm}) + 638.522(\rho_{644nm}) - 600.979(\rho_{1688nm}) + 546.564(\rho_{1781nm}) + 417.835(\rho_{811nm}) + 826.352(\rho_{1772nm}) - 924.824(\rho_{963nm})$	0.82
Total Nitrogen (g/kg)	R1	Ref	$0.471 - 29.916(\rho_{1167nm}) + 0.182(\rho_{430nm}) + 2.683(\rho_{633nm}) - 2.704(\rho_{1927nm}) - 10.262(\rho_{1844nm}) - 25.259(\rho_{545nm}) + 14.507(\rho_{1412nm}) + 14.396(\rho_{2239nm}) - 1.043(\rho_{435nm}) + 3.735(\rho_{1730nm}) + 23.113(\rho_{681nm}) + 6.049(\rho_{2393nm}) + 7.281(\rho_{443nm}) + 1.964(\rho_{2192nm}) + 17.944(\rho_{436nm}) + 3.581(\rho_{2357nm}) - 4.190(\rho_{2410nm}) - 10.263(\rho_{2424nm}) - 3.569(\rho_{432nm}) + 5.985(\rho_{1947nm}) + 1.264(\rho_{2368nm}) + 2.023(\rho_{2402nm})$	0.74
		CR	$371.861 - 64.282(\rho_{2077nm}) + 93.617(\rho_{1242nm}) - 56.977(\rho_{1644nm}) - 13.346(\rho_{2432nm}) + 16.784(\rho_{2124nm}) + 51.871(\rho_{2089nm}) - 25.871(\rho_{1639nm}) - 48.096(\rho_{2413nm}) - 79.833(\rho_{1649nm}) - 70.915(\rho_{1690nm}) - 23.863(\rho_{1275nm}) - 46.097(\rho_{1608nm}) - 65.263(\rho_{1864nm}) - 30.052(\rho_{2426nm}) + 24.447(\rho_{1550nm}) + 21.733(\rho_{2423nm}) + 26.259(\rho_{1633nm}) - 46.604(\rho_{1579nm}) + 27.492(\rho_{1613nm}) - 8.069(\rho_{1612nm}) - 18.646(\rho_{1289nm}) - 47.675(\rho_{1206nm}) + 12.993(\rho_{1303nm})$	0.79
		1st derivative	$0.615 + 492.178(\rho_{1302nm}) + 404.603(\rho_{1445nm}) + 586.497(\rho_{1745nm}) - 363.862(\rho_{2240nm}) - 371.916(\rho_{465nm}) + 625.714(\rho_{1502nm}) - 311.107(\rho_{978nm}) + 288.626(\rho_{537nm}) - 216.977(\rho_{534nm}) - 395.304(\rho_{1916nm}) - 590.843(\rho_{868nm}) - 540.463(\rho_{2019nm}) + 115.212(\rho_{879nm}) + 125.674(\rho_{2054nm}) - 198.114(\rho_{1672nm}) + 270.072(\rho_{1279nm}) + 55.710(\rho_{1645nm}) - 89.757(\rho_{1527nm}) - 582.611(\rho_{1244nm}) + 0.506(\rho_{2282nm}) - 70.918(\rho_{1288nm}) + 194.018(\rho_{997nm}) - 198.306(\rho_{2397nm}) - 56.353(\rho_{1680nm})$	0.87
	R2	Ref	$0.779 - 2.061(\rho_{1003nm}) + 0.292(\rho_{534nm}) + 17.959(\rho_{1394nm}) + 2.050(\rho_{1925nm}) - 11.007(\rho_{782nm}) + 1.918(\rho_{442nm}) + 4.519(\rho_{1901nm}) + 1.553(\rho_{430nm}) + 1.003(\rho_{470nm}) - 4.333(\rho_{2206nm}) + 1.406(\rho_{463nm}) - 11.226(\rho_{2341nm}) - 1.782(\rho_{455nm}) + 1.703(\rho_{494nm})$	0.78
		CR	$243.411 - 125.593(\rho_{1933nm}) + 108.668(\rho_{1960nm}) + 120.267(\rho_{1431nm}) - 189.954(\rho_{2240nm}) - 281.730(\rho_{1560nm}) + 150.479(\rho_{2244nm}) + 61.722(\rho_{1665nm}) - 49.823(\rho_{776nm}) + 41.382(\rho_{1862nm}) + 171.693(\rho_{837nm}) - 89.173(\rho_{1266nm}) - 200.838(\rho_{1282nm}) - 76.934(\rho_{1711nm}) - 93.156(\rho_{848nm}) + 45.056(\rho_{1361nm}) + 164.105(\rho_{2122nm}) - 81.307(\rho_{1634nm}) + 25.257(\rho_{1370nm}) + 88.950(\rho_{1615nm}) - 28.726(\rho_{2097nm}) - 27.427(\rho_{2388nm}) - 16.771(\rho_{1538nm}) + 7.954(\rho_{1643nm}) + 30.729(\rho_{1264nm})$	0.90
		1st derivative	$1.232 - 1075.175(\rho_{465nm}) - 573.695(\rho_{1892nm}) - 1047.726(\rho_{605nm}) - 756.651(\rho_{1927nm}) - 286.125(\rho_{1001nm}) - 1064.062(\rho_{1999nm}) + 1265.916(\rho_{953nm}) + 492.979(\rho_{543nm}) - 738.600(\rho_{1599nm}) - 509.620(\rho_{1874nm}) - 514.152(\rho_{1915nm}) - 924.649(\rho_{1949nm}) - 303.369(\rho_{2409nm}) - 900.194(\rho_{529nm}) - 677.309(\rho_{1943nm}) + 570.313(\rho_{1436nm}) + 290.187(\rho_{432nm}) + 244.436(\rho_{1575nm}) + 346.233(\rho_{1457nm}) + 481.677(\rho_{573nm}) - 186.704(\rho_{1503nm})$	0.92
	R1&R2	Ref	$0.678 - 39.079(\rho_{2440nm}) - 19.081(\rho_{1269nm}) + 21.450(\rho_{2207nm}) - 4.202(\rho_{536nm}) + 8.391(\rho_{1939nm}) + 14.154(\rho_{1413nm}) - 12.842(\rho_{679nm}) - 18.693(\rho_{2227nm}) - 6.309(\rho_{476nm}) - 3.382(\rho_{435nm}) - 0.065(\rho_{1798nm}) - 16.511(\rho_{431nm}) + 10.121(\rho_{2393nm}) - 23.270(\rho_{445nm}) - 1.948(\rho_{1001nm}) - 1.331(\rho_{468nm}) + 4.588(\rho_{436nm}) + 36.844(\rho_{997nm}) + 3.314(\rho_{497nm}) + 10.636(\rho_{2297nm}) + 5.075(\rho_{441nm}) + 6.013(\rho_{577nm})$	0.82
		CR	$192.724 - 95.079(\rho_{1688nm}) - 79.980(\rho_{1815nm}) - 115.665(\rho_{2118nm}) - 34.478(\rho_{1000nm}) + 150.150(\rho_{1538nm}) - 126.429(\rho_{1581nm}) + 76.205(\rho_{2122nm}) - 210.680(\rho_{1282nm}) - 57.418(\rho_{2036nm}) + 4.702(\rho_{803nm}) + 40.214(\rho_{1856nm}) + 80.752(\rho_{1795nm}) + 37.918(\rho_{1304nm}) + 92.792(\rho_{1249nm}) + 98.335(\rho_{1960nm}) - 102.234(\rho_{1964nm}) - 42.242(\rho_{2404nm}) - 147.268(\rho_{1553nm}) + 118.581(\rho_{1500nm}) + 89.072(\rho_{1652nm}) - 81.872(\rho_{1977nm}) - 51.867(\rho_{1567nm}) + 64.635(\rho_{2226nm}) - 37.237(\rho_{1075nm})$	0.87

	1st derivative	$1.012 - 628.129(\rho_{1398nm}) - 226.778(\rho_{463nm}) - 613.316(\rho_{2040nm}) - 290.693(\rho_{999nm}) - 607.769(\rho_{2331nm}) - 587.196(\rho_{2120nm}) - 447.890(\rho_{617nm}) - 262.155(\rho_{2396nm}) - 406.045(\rho_{2047nm}) - 407.537(\rho_{2362nm}) + 935.901(\rho_{1632nm}) - 720.846(\rho_{1841nm}) + 667.849(\rho_{1263nm}) - 326.632(\rho_{1580nm}) + 726.030(\rho_{1604nm}) - 430.720(\rho_{2190nm}) - 381.524(\rho_{2045nm}) - 787.186(\rho_{510nm}) + 482.449(\rho_{434nm}) + 410.493(\rho_{1290nm}) + 425.210(\rho_{490nm}) - 435.723(\rho_{1621nm}) + 206.930(\rho_{615nm}) + 658.638(\rho_{1238nm}) - 728.577(\rho_{1479nm}) - 585.884(\rho_{2254nm}) - 302.672(\rho_{1311nm}) + 456.733(\rho_{1592nm}) - 423.899(\rho_{2325nm}) + 542.179(\rho_{1400nm}) - 375.486(\rho_{1853nm}) + 266.293(\rho_{1087nm}) - 201.536(\rho_{1194nm}) - 526.395(\rho_{529nm}) - 116.711(\rho_{1803nm}) + 393.493(\rho_{2066nm}) - 344.220(\rho_{2354nm}) + 370.483(\rho_{1476nm})$	0.90

Ref.: Reflectance; CR: Continuum-removed