

# Near-Infrared Spectroscopy: Assessment of Soil Organic Carbon Stock in a Colombian Oxisol

## Espectroscopia de infrarrojo cercano: evaluación del almacenamiento de carbono orgánico del suelo en un oxisol colombiano

Felipe Fernández-Martínez<sup>1</sup>, Jesús Hernán Camacho-Tamayo<sup>2</sup>, and Yolanda Rubiano-Sanabria<sup>3</sup>

### ABSTRACT

Soil organic carbon (SOC) is a property known for its influence on the physical, chemical, and biological characteristics of soils, which are essential when assessing their quality. SOC stock (SOCS) monitoring is a key task in climate change mitigation studies. However, the resources necessary to obtain the information required by these studies tend to be high. The objective of this study was to develop a model for estimating the SOCS of a Colombian oxisol using near-infrared (NIR) diffuse reflectance spectroscopy. In a sampling scheme of 70 points distributed over 248 ha, 313 soil samples were collected in five defined depth intervals of 10 cm each, from 0 to 50 cm. SOC was determined through an elemental analyzer, and bulk density (BD) by means of sampling cylinders. A NIRFlex spectrometer was used to acquire spectral signatures in the NIR range from the processed soil samples, and, together with the data measured in the laboratory, a statistical analysis was performed using partial least squares regression (PLSR) in order to calibrate the spectral models. Based on the residual prediction deviation (RPD), the root mean square error (RMSE), and the coefficient of determination ( $R^2$ ) of the validation groups, a highly representative model was achieved for the estimation of SOCS ( $R^2 = 0,93$ ; RMSE = 2,12 tC ha<sup>-1</sup>; RPD = 3,69), which was also corroborated with geostatistical interpolation surfaces and depth splines. This research showed NIR diffuse reflectance spectroscopy to be a viable technique for SOCS estimation in the study area.

**Keywords:** soil organic carbon, bulk density, soil spectroscopy, spline, geostatistics

### RESUMEN

El carbono orgánico del suelo (COS) es una propiedad conocida por su influencia en las propiedades físicas, químicas y biológicas de los suelos, que son fundamentales para evaluar su calidad. El monitoreo del stock de COS (SCOS) es una labor clave en los estudios de mitigación del cambio climático. Sin embargo, los recursos necesarios para obtener la información requerida en estos estudios suelen ser elevados. El objetivo de este estudio fue desarrollar un modelo para estimar el SCOS de un oxisol colombiano utilizando espectroscopia de reflectancia difusa de infrarrojo cercano (NIR). En un esquema de muestreo de 70 puntos distribuidos en 248 ha, se recolectaron 313 muestras de suelo en cinco intervalos de profundidad definidos de 10 cm cada uno, de 0 a 50 cm. El COS se determinó mediante un analizador elemental, y la densidad aparente (DA) mediante cilindros de muestreo. Se utilizó un espectrómetro NIRFlex para adquirir firmas espectrales en el rango NIR de las muestras de suelo procesadas, y, junto con datos medidos en laboratorio, se realizó un análisis estadístico usando regresión de mínimos cuadrados parciales (RMCP) para calibrar los modelos espectrales. Con base en la desviación de predicción residual (DRP), raíz del error cuadrático medio (RECM) y el coeficiente de determinación ( $R^2$ ) de los grupos de validación, se logró un modelo de alta representatividad para la estimación de SCOS ( $R^2 = 0,93$ ; RECM = 2,12 tC ha<sup>-1</sup>; DRP = 3,69), lo cual también se corroboró con superficies de interpolación geoestadística y splines de profundidad. Esta investigación mostró que la espectroscopia de reflectancia difusa NIR es una técnica viable para la estimación de SOCS en el área de estudio.

**Palabras clave:** carbono orgánico del suelo, densidad aparente, espectroscopia de suelos, spline, geoestadística

**Received:** October 19<sup>th</sup>, 2021

**Accepted:** March 21<sup>th</sup>, 2022

### Introduction

In recent decades, scientists around the world have been increasingly interested in studying soil organic carbon (SOC) and its relationship with climate change. A potential impact of global warming is the accelerated decomposition of SOC and an increase in the carbon released into the atmosphere (Jia *et al.*, 2017). Rising levels of CO<sub>2</sub> concentration in the atmosphere have led a considerable part of the scientific community to contemplate SOC sequestration as an alternative to mitigate climate change, which can be achieved through various management practices based on increasing the SOC stock (SOCS) (Huang *et al.*, 2019).

<sup>1</sup> Agricultural engineer, MSc in Agricultural Sciences, Universidad Nacional de Colombia, Colombia. Affiliation: Graduate research assistant and PhD candidate in Civil Engineering – Water Resources Engineering, Mississippi State University, United States. Email: ff181@msstate.edu

<sup>2</sup> Agricultural engineer, Universidad Nacional de Colombia, Colombia. MSc in Agricultural Engineering, Universidad de Campinas, Brazil. PhD in Agricultural Sciences, Universidad Nacional de Colombia, Colombia. Affiliation: Associate professor, Universidad Nacional de Colombia, Colombia. Email: jhcamachot@unal.edu.co

<sup>3</sup> Agrologist, Universidad Jorge Tadeo Lozano, Colombia. PhD in Agricultural Sciences, Universidad Nacional de Colombia, Colombia. Affiliation: Associate professor, Universidad Nacional de Colombia, Colombia. Email: yrubianos@unal.edu.co



Through photosynthesis, plants are able to capture carbon from the atmosphere and integrate it into their own structure (biomass), as well as into the soil through their radicular systems. This phenomenon can also be called *carbon sequestration* (Stockmann *et al.*, 2013) and has a special focus on agricultural land due to the fact that around 37% of the planet's surface is destined for agricultural use (Sommer and Bossio, 2014).

In order to monitor the SOCS, it is necessary to acquire information on the soil bulk density (BD) and the SOC and subsequently analyze the changes in these variables over different periods of time according to the established needs. However, the resources needed to conduct these kinds of studies are expensive and time-consuming because of the standard laboratory analyses and procedures involved (Camacho-Tamayo *et al.*, 2014; Liu *et al.*, 2019). Soils are also characterized by being a heterogeneous material in both spatial and temporal scales, often causing SOCS studies to require a high sampling density (Cambou *et al.*, 2016). Therefore, it is necessary to explore SOCS analysis techniques that represent a reduction in the resources required.

Chemometric approaches constitute a potential tool to evaluate soil attributes rapidly and accurately in the laboratory (Ben Dor *et al.*, 2015; Davari *et al.*, 2021). Diffuse reflectance spectroscopy is a technique developed to accelerate data acquisition regarding materials' properties. Soil spectroscopy is based on the idea that the characteristics of the radiation reflected by a material when exposed to an electromagnetic spectrum depend on the composition of the material itself. This means that studying soil reflectance can provide information about the properties and condition of the soil under study (Davari *et al.*, 2021; Viscarra Rossel and Webster, 2012). The features of the reflected spectra of a soil sample in the NIR spectrum have allowed for optimal results in terms of predicting soil properties, especially when dealing with SOC assessments (Liu *et al.*, 2019; Nawar and Mouazen, 2019; Nocita *et al.*, 2014).

It is important to mention that the acquisition of spectral signatures can be carried out *in situ* or using laboratory equipment. This can affect the information obtained from the reflectance spectra, causing variations in the model performance with regard to the estimated properties (Ahmadi *et al.*, 2021; Ben Dor *et al.*, 2015). When spectral signatures are taken *in situ*, factors such as the contact probe device of the equipment and its operator, the moisture of the soil sample, and environmental conditions can alter the spectral reflectance of the samples in comparison with those acquired in the laboratory. The latter would then be assumed to have been acquired under controlled environmental conditions, as well as after being dried and sieved for the sake of homogeneity.

The objective of this research was to evaluate the potential of NIR spectroscopy in order to estimate the SOCS of oxisols in the study area. Several models were developed by varying

the number of samples used for calibration, followed by the evaluation of the spatial variability of the measured and estimated data by means of depth splines and geostatistical interpolated surfaces.

## Materials and methods

### Study area and field sampling

This study was carried out at the Carimagua Experimental Station of Agrosavia, located in the municipality of Puerto Gaitán (Meta, Colombia), with geographic coordinates 4° 34' 48" N, 71° 21' 00" W, an average altitude of 175 masl, an average annual temperature of 28 °C, an average rainfall of 2 240 mm, and a warm-humid climate. The relief of the study area is flat to slightly undulating, with slopes ranging between 0 and 7%. Soils are mostly oxisols whose dominant taxonomic component is typic hapludox. These soils have a low pH (<5), are well drained, and are covered with native savanna. Their predominant use is extensive livestock farming.

The study area has an extension of 248 ha, where a directed sampling scheme was established. Sampling took place in 2018 and consisted of 70 points spaced perpendicularly by 200 m, from which soil samples were extracted in depth intervals of 0-10, 10-20, 20-30, 30-40, and 40-50 cm, for a total of five samples per trunk (Figure 1). The total number of samples collected was 313, out of which 70 corresponded to the first and second depths, 68 to the third, 59 to the fourth, and 46 to the last. The difference in samples collected per depth is attributed to the presence of the water table in the study area. Each of the 313 soil samples consisted of a disturbed sample and an undisturbed cylinder sample.

### Laboratory analysis

The disturbed samples were dried at 35 °C and then sifted through a 2 mm mesh. Total soil carbon was determined using an elemental analyzer (TruSpec CN Carbon Nitrogen Determinator, LECO Corp., St. Joseph, MI, USA). Since the soil samples did not show the presence of carbonates in the hydrochloric acid test, the total carbon was considered to be the SOC. The undisturbed soil sample cylinders were oven-dried at 105 °C for 48 h, and then the BD was calculated by dividing the mass of the soil by the volume of the cylinder. The SOCS was determined according to Equation (1):

$$SCOS \left[ \frac{t}{ha} \right] = DT [dm] * SOC \left[ \frac{gC}{kg} \right] * (1 - Pd * 0,01) \quad (1)$$

where DT stands for the depth and thickness of the soil layer, and Pd stands for the stoniness factor. In this case, Pd was assumed to be 0 because no significant stoniness was evidenced in the sampling process. Spectral signatures were recorded using a NIRFlex N-500 spectrometer (BÜCHI Labortechnik AG), which recorded measurements in

reflectance units in the NIR range, from 1 000 nm to 2 500 nm, averaging 32 scans for each wavelength. The spectral signatures were taken from the dried and sieved soil samples.

### Spectral modeling

To perform spectral modeling, it was first necessary to define the calibration and validation groups of the spectral models. The Kennard-Stone algorithm (Kennard and Stone, 1969) made it possible to establish uniformly distributed groups according to the variance of the spectral signatures. The calibration groups for BD and SOC were defined with 70% of the total signatures, while the validation group was assigned the remaining 30%. For SCOS spectral modeling, calibration groups consisting of 10, 20, 30, 60, 90, 120, 150, 180, 210, 240, and 270 samples were built with the Kennard-Stone algorithm. A subsequent validation of each group was carried out with the samples that did not take part in the calibration model.

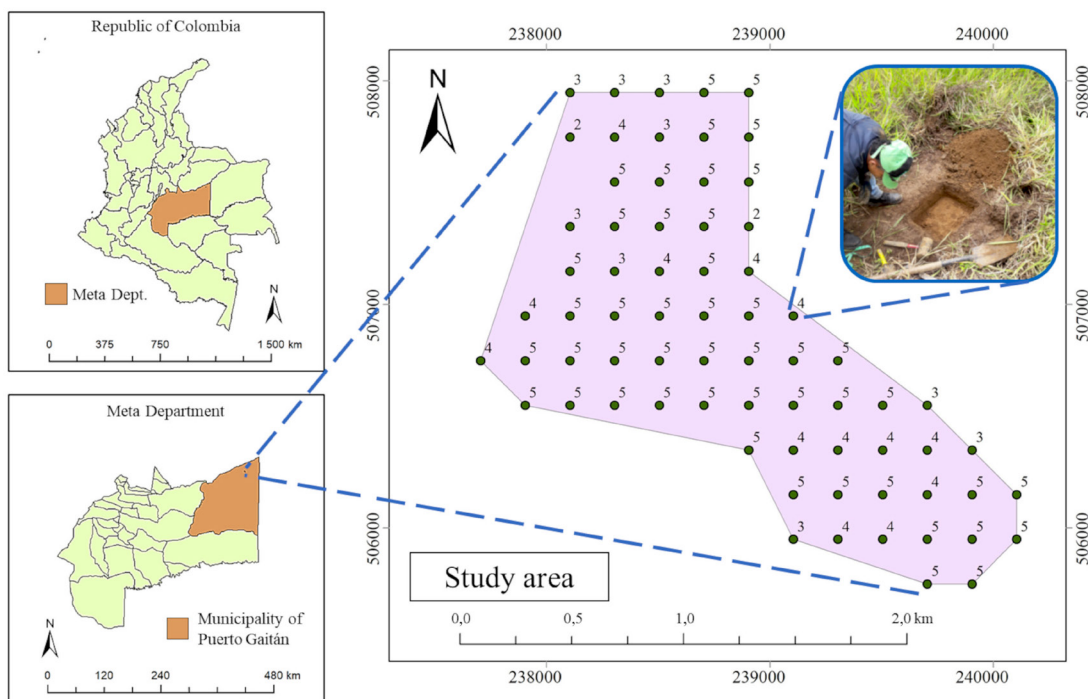
In the model calibrations, spectral data were submitted to a series of pretreatments to normalize the responses and smooth out any possible noise that could be present in the signatures. The pretreatments considered were the transformation of raw spectra to absorbance, the Savitsky-Golay derivate, and standard normal variation. Once the combinations of pretreatments were applied, the model was built by means of PLSR, which is a commonly used technique when dealing with high-dimensional data (such as the wavelengths of the NIR spectral range) with highly collinear predictor variables. This technique summarizes the data into

a few orthogonal factors, which are linear combinations of the predictor variables, so that the covariance of both the predictor and dependent variables is maximized. These orthogonal factors are then employed to develop a linear model to estimate the soil's attribute of interest (Katuwal *et al.*, 2020).

The criteria to establish the performance of each model were set by the values obtained for the coefficient of determination ( $R^2$ ), root mean square error (RMSE), and residual prediction deviation (RPD), the latter being the ratio of the standard deviation of the laboratory-measured data to the validation RMSE. The lower the RMSE and the higher the  $R^2$ , the better the estimations reached. According to Viscarra Rossel *et al.* (2006), the RPD can rate a model as unrepresentative if it reaches a value lower than 1,4; as regular with values between 1,4 and 1,8; as well-performing with 1,8-2,0; as a very good model with 2,0-2,5; and as an excellent model with a RPD value greater than 2,5. The construction and evaluation of spectral models was carried out with the R software (R Core Team) and the libraries *prospectr* (Stevens and Ramírez-Lopez, 2020) and *pls* (Liland *et al.*, 2021).

### Spatial variation analysis

The vertical behavior of the estimated and measured data was verified by means of depth splines. Splines are a set of quadratic functions joined by nodes (depths measured) that represent a smoothed curve (Ponce-Hernandez *et al.*, 1986). Regarding the horizontal plane, by means of geostatistical interpolation surfaces, the likeness of the



**Figure 1.** Location of the study area detailing the number of samples obtained for each trunk

**Source:** Authors

measured and estimated SOCS was observed. To this effect, experimental semivariograms were built and then adjusted to fit the theoretical models. These models have three common parameters: the nugget effect ( $C_0$ ), the sill ( $C_0 + C_1$ ), and the range. The adjustment of the theoretical model was based on the lowest value of the sum of the squared residuals, the highest values of  $R^2$  of the fitted model, and the cross-validation coefficient (CVC). Afterwards, spatial interpolation was carried out by ordinary kriging, which is regarded as an optimal estimation method that provides unbiased predictions while minimizing the variance (Oliver and Webster, 2015). In addition, the degree of spatial dependence (DSP) was assessed, which is the relationship between the nugget effect and the sill of the semivariograms ( $C_1 * (C_0 + C_1)^{-1}$ ). According to Cambardella *et al.* (1994), the DSP is classified as weak if it is lower than 25%, as moderate if it is between 25 and 75%, and as strong above 75%. Geostatistical processing was performed with the GS+ software v.9 (Gamma Design Software, LLC, Plainwell, MI, USA), in combination with ArcGIS v.10.8 (ESRI), while depth splines were built with R (R Core Team) and the *ithir* library (Malone, 2016).

## Results and discussion

### Exploratory analysis and spectral signatures

For the statistical analysis of each variable, an exploratory data study was initially carried out with R, calculating measures of location (mean, median, minimum, and maximum), variability (coefficient of variation, CV), and central tendency (skewness and kurtosis). Values close to zero for skewness and kurtosis indicated that the precept of normality was met. The descriptive statistics for the behavior of BD and SOC across the studied profile are shown in Table 1. It was evidenced that, across the soil profile, the behaviors of SOC and BD are inversely proportional. The BD reached a mean

value of 1,33 g cm<sup>-3</sup> at the first depth, which kept increasing until 1,49 g cm<sup>-3</sup> for the last interval. For the SOC, the average content in the range of 0-10 cm was 2,61%, which kept decreasing until reaching a value of 0,95% at 40-50 cm. The higher values of SOC at lower depths can be attributed to the presence of vegetation and the natural contribution of surface residues. A similar behavior of these two properties was also reported by other investigations in the oxisols of the Eastern Plains in Colombia (Camacho-Tamayo *et al.*, 2014; Ramirez-Lopez *et al.*, 2008)

The average spectral signature for all soil samples grouped by soil depth is shown in Figure 2. The behavior shown by the spectral signatures at each sampling depth can also be qualitatively related to some of the soil properties. In the NIR, at wavelength ranges of 1 000-1 350 nm and 1 400-2 200 nm, spectral responses tend to be smoother and have a lower magnitude of reflectance in soils with a higher content of organic material (Poppiel *et al.*, 2018). This is because organic matter affects the albedo of soils, which makes them reflect less light at higher levels of organic content (organic carbon) (Camacho-Tamayo *et al.*, 2014). The spectral signatures obtained in this study showed a spectral reflectance according to these characteristics, as higher SOC contents were observed at lower sampling depths, which is consistent with the lower magnitudes of reflectance at a greater proximity to the soil surface.

### BD and SOC modeling performance

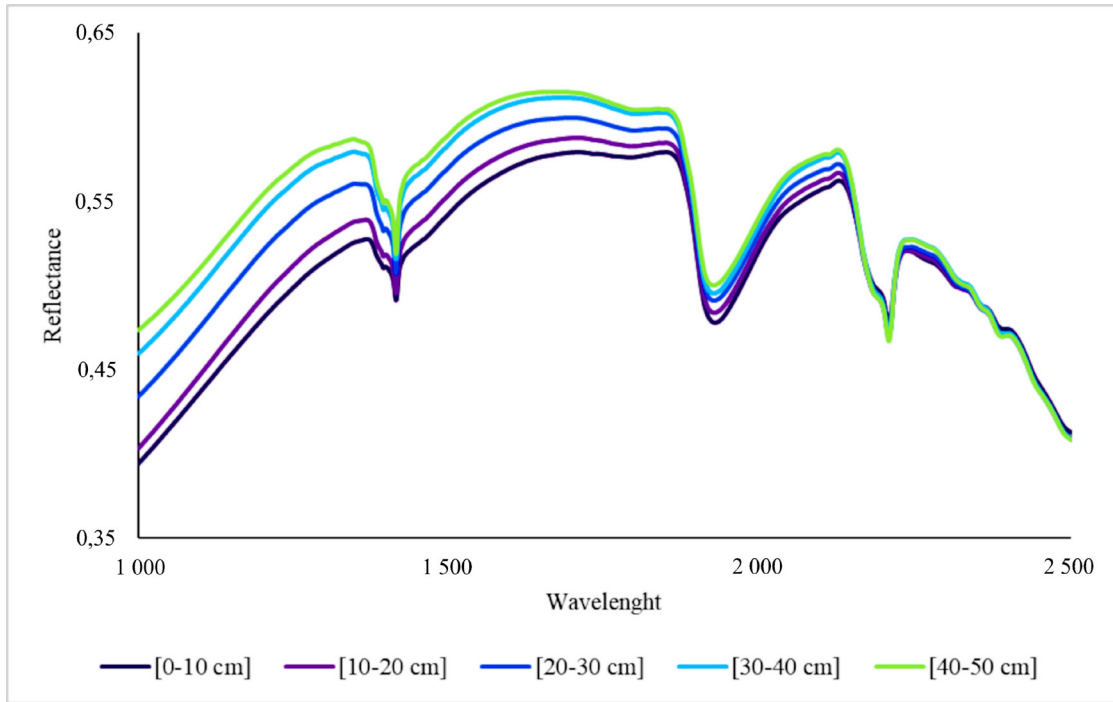
The performance of the most accurate spectral models for BD and SOC is shown in Table 2. These were calibrated with 70% of the total signatures and validated with the remaining 30%. The BD spectral model turned out to be unrepresentative, reaching a RPD of 1,35 with an  $R^2$  equal to 0,41 during its validation. On the other hand, SOC showed excellent performance, reaching RPD and  $R^2$  values of 5,63 and 0,97, respectively.

**Table 1.** Descriptive statistics for BD and SOC per depth with measured data

Property	Depth [cm]	Mean	Median	Max,	Min,	SD	CV [%]	Skewness	Kurtosis
BD [g cm <sup>-3</sup> ]	0-10	1,33	1,33	1,58	1,14	0,10	7,78	0,18	-0,38
	10-20	1,41	1,42	1,55	1,28	0,06	4,43	-0,01	-0,62
	20-30	1,41	1,41	1,58	1,25	0,09	6,14	-0,06	-0,74
	30-40	1,43	1,46	1,63	1,24	0,10	7,07	-0,11	-0,89
	40-50	1,49	1,50	1,73	1,29	0,10	6,70	-0,09	-0,28
SOC [%]	0-10	2,61	2,63	3,24	1,98	0,26	10,08	-0,01	-0,12
	10-20	1,75	1,72	2,16	1,33	0,18	10,51	0,07	-0,51
	20-30	1,39	1,37	1,75	1,08	0,16	11,33	0,38	-0,42
	30-40	1,12	1,08	1,50	0,85	0,17	14,78	0,57	-0,59
	40-50	0,95	0,98	1,31	0,58	0,16	16,50	-0,06	-0,30

SD: standard deviation; CV: coefficient of variation; Max: maximum; Min: minimum

Source: Authors



**Figure 2.** Spectral signatures averaged for each sampling depth  
**Source:** Authors

**Table 2.** Calibration and validation results for the BD and SOC spectral models

Property	Model	R <sup>2</sup>	RMSE	RPD	Range
BD [g cm <sup>-3</sup> ]	Calibration	0,45	0,08	1,36	1,14 - 1,73
	Validation	0,41	0,08	1,35	1,15 - 1,63
SOC [%]	Calibration	0,97	0,11	5,46	0,58 - 3,24
	Validation	0,97	0,11	5,63	0,75 - 3,24

**Source:** Authors

The underperformance obtained for the spectral model of BD has also been reported in other studies (Katuwal *et al.*, 2020; Moreira *et al.*, 2009). Particularly, the work by Moreira *et al.* (2009) consisted of 1 184 samples with BD values in the range of 0,45-1,95 g cm<sup>-3</sup>. They managed to validate models with R<sup>2</sup> between 0,10 and 0,34 by applying different pretreatments to the spectral signatures. Meanwhile, Katuwal *et al.* (2020) achieved slightly higher performances, with R<sup>2</sup> between 0,10 and 0,52 while working with a BD range of 1,02-2,01 g cm<sup>-3</sup>. BD is a soil property whose estimation through spectral signatures depends on the correlations that it may have with other attributes with recognized associations in the spectral range (Askari *et al.*, 2015; Viscarra Rossel and Webster, 2012). For example, Al-Asadi and Mouazen (2014) coupled auxiliary variables such as water content, texture fractions, and organic matter with spectral models of BD in order to improve estimations, going from R<sup>2</sup> values between 0,23 and 0,69 to a 0,70-0,81 range.

The level of performance obtained for the SOC estimations has also been reported in other works, corroborating the potential use of NIR spectroscopy for the prediction of this property (Jia *et al.*, 2017; Nawar and Mouazen, 2019; Nocita *et al.*, 2014). An example of this can be found in the large-scale study conducted by Liu *et al.* (2019), in which a model composed of 11 213 soil samples with a mean SOC content of 1,84% and magnitudes ranging from 0,02% to 9,96% achieved a validation R<sup>2</sup> equal to 0,96 and a RMSE of 0,29%.

### SOCs spectral estimation

The result of the product between BD [kg dm<sup>-3</sup>] and SOC [gC kg<sup>-1</sup>] is the volumetric SCOS [gC dm<sup>-3</sup>] (SCOSv), which was the value estimated from spectral models. The performance of the models calibrated via the established groups ranging from 10 to 300 samples is shown in Table 3. It was evidenced that, by increasing the number of samples used to calibrate the models, the RMSE decreased while the RPD and R<sup>2</sup> increased. This bears out the fact that more robust estimation models can be achieved with more samples. The stabilization of these parameters occurs at around 90 samples for the calibration model, so this number was chosen as optimal for SCOSv spectral modeling. However, it is worth highlighting that each of the models defined in Table 3 showed RPD and R<sup>2</sup> values greater than 2,0 and 0,80 in the validation parameters.

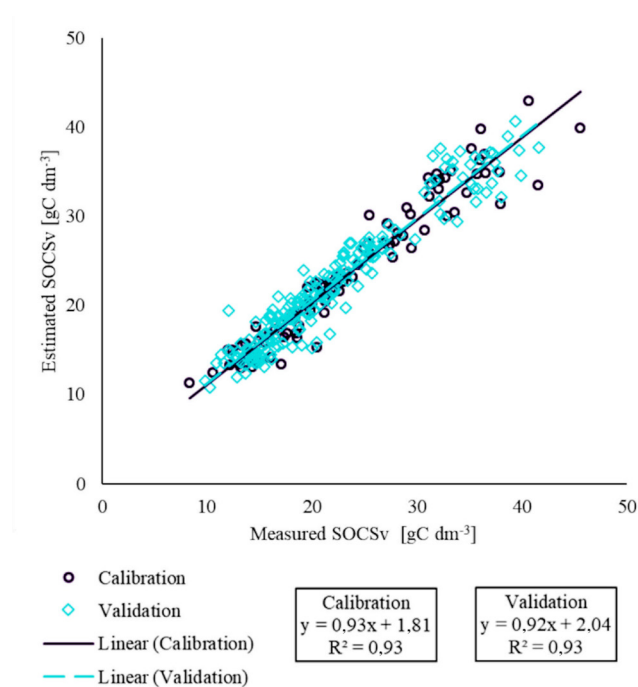
According to the RPD, the model made with the calibration group of 90 soil samples gave excellent estimates for SCOSv, with a RMSE of 2,12 gC dm<sup>-3</sup> and a R<sup>2</sup> of 0,93 (Figure 3). The

calibration range for this model was 8,39-45,55 gC dm<sup>-3</sup>, while the range was 9,81-41,61 gC dm<sup>-3</sup> for the validation group. The behavior of the measured and estimated SCOS in the soil profile is shown in Table 4. No large differences between the measured and estimated SOCS were evidenced. The datasets showed similar magnitudes regarding the mean and median, as well as for skewness, kurtosis, and the CV, which was in fact lower at each one of the sampling ranges (estimated vs. measured SOCS). The average content regarding the SOCS measured in the whole depth range of 0-50 cm was 109,28 tC ha<sup>-1</sup>. This value was 110,56 tC ha<sup>-1</sup> for the estimated data.

**Table 3.** Results for the calibration and validation of SOCSv spectral modeling

Number of samples	Calibration			Validation		
	R <sup>2</sup>	RMSE [gC dm <sup>-3</sup> ]	RPD	R <sup>2</sup>	RMSE [gC dm <sup>-3</sup> ]	RPD
10	0,98	1,04	2,87	0,82	3,35	2,34
20	0,92	2,36	1,97	0,84	3,13	2,50
30	0,88	2,74	2,11	0,86	2,95	2,66
60	0,93	2,14	2,95	0,92	2,19	3,58
90	0,93	2,16	3,01	0,93	2,12	3,69
120	0,93	2,17	3,08	0,92	2,15	3,64
150	0,92	2,16	2,97	0,93	2,13	3,68
180	0,93	2,15	3,20	0,93	2,07	3,78
210	0,93	2,09	3,34	0,93	2,04	3,84
240	0,93	2,08	3,42	0,93	2,08	3,77
270	0,93	2,09	3,44	0,93	2,08	3,77

Source: Authors



**Figure 3.** Spectrally estimated vs. measured values of SOCSv  
Source: Authors

A study carried out by Cambou *et al.* (2016) reported acceptable models regarding the estimation of SCOS from spectral signatures acquired in the field for a haplic luvisol, reaching R<sup>2</sup> and RPD values of 0,70 and 1,80, respectively, for SCOSv values ranging between 7,90 and 27,70 gC dm<sup>-3</sup>. Part of the flaws in the estimation model of said research were attributed to the fact that the spectral signatures were collected by manually extracted auger soil cores that altered the natural structure of the samples. This, in addition to the fact that conventional laboratory analyses for BD and SOC were not performed on the soil samples from which the spectral signatures were collected (Cambou *et al.*, 2016). On the other hand, Allory *et al.* (2019) compared the performance of models calibrated for urban soils using spectral signatures acquired in the field and in the laboratory. With a SCOSv range of 0,30-54,60 gC dm<sup>-3</sup>, a validation model was obtained with a RPD equal to 2,20 and an R<sup>2</sup> of 0,78 for spectral signatures acquired *in situ*, while, for the laboratory-acquired spectra, the model parameters reported were 3,1 for the RPD and 0,89 for the R<sup>2</sup>. The performance improvement of the latter was attributed to the more stable measurement conditions in the laboratory, in comparison with those in the field.

For tropical volcanic soils that included andosols, cambisols, and ferralsols, Allo *et al.* (2020) calibrated models with spectral signatures obtained in the field and in the laboratory, obtaining better results in the validation of the model elaborated from signatures taken *in situ* (R<sup>2</sup> = 0,91 and RPD = 3,29) when compared to the validation of the laboratory spectral model (R<sup>2</sup> = 0,86 and RPD = 2,63). Allo *et al.* (2020) used a mechanical column auger to collect spectra in the field, allowing them to collect 30 signatures uniformly every 10 cm of the sampled soil column to establish an average signature for a single sample, achieving a closer representation of natural conditions. Doing this could have considerably improved the predictive ability of the model with spectra taken *in situ* vs. those acquired in the laboratory.

Given that the magnitudes of SOC and SCOS tend to increase at depths closer to the surface and decrease at lower depths in the soil profile (FAO, 2019; Stockmann *et al.*, 2013), the Intergovernmental Panel on Climate Change recommends monitoring these organic C concentrations at least in the depth range of 0-30 cm (IPCC, 2019). This recommendation is also motivated by the fact that the first 30 cm from the soil surface usually experience substantial changes in SOC due to management practices. Hence, the detection of variations in SOC reserves at this depth range could be noticed in shorter periods of time, depending on the activities carried out on them (Huang *et al.*, 2019). In this regard, the oxisols of the Eastern Plains in the study area show a cumulative mean value of 78,85 tC ha<sup>-1</sup> in the first 30 cm of soil depth, which, for estimated spectral data, equals 79,86 tC ha<sup>-1</sup> and corresponds to approximately 72% of the total SOCS present in the profile considered.

**Table 4.** Results for the calibration and validation of SOCSv spectral modeling

SCOS [tC ha <sup>-1</sup> ]	Depth [cm]	Mean	Median	Max,	Min,	SD	CV [%]	Skewness	Kurtosis
Measured	0-10	34,64	34,57	45,55	25,44	3,69	10,64	0,18	0,82
	10-20	24,74	25,01	29,85	19,18	2,43	9,83	-0,05	-0,56
	20-30	19,47	19,55	25,81	14,68	2,35	12,05	0,34	-0,07
	30-40	16,11	15,91	21,69	12,05	2,67	16,55	0,42	-0,79
	40-50	14,32	14,64	20,41	8,39	2,43	16,99	-0,10	0,35
Estimated	0-10	34,38	34,31	42,90	27,64	3,08	8,95	0,16	0,12
	10-20	25,32	25,58	29,11	21,56	1,91	7,55	-0,17	-0,85
	20-30	20,16	20,21	25,44	15,76	1,95	9,65	0,22	0,38
	30-40	16,21	15,69	20,89	12,41	2,05	12,65	0,59	-0,38
	40-50	14,49	14,44	18,34	10,84	1,75	12,07	0,28	-0,04

SD: standard deviation; CV: coefficient of variation; Max: maximum; Min: minimum

Source: Authors

### SOCS spatial variation

The vertical variation of the soil properties is shown in Figure 4. The splines allowed to estimate the value of soil attributes within each measured depth. Furthermore, the addition of boxplots to these diagrams aided in examining the similarity between measured and estimated data. At first glance, it can be noted that the resulting splines for BD have clear differences, which are attributed to the low representativeness of the spectral model. The interquartile range, as well as the whiskers of the boxplots, showed a considerable reduction for the estimated data in comparison with the measured values. The unrepresentativeness obtained in the spectral modeling for BD can be associated with the alteration of the structure of the soil samples prior to the signature acquisitions, as the soil samples were dried and sieved.

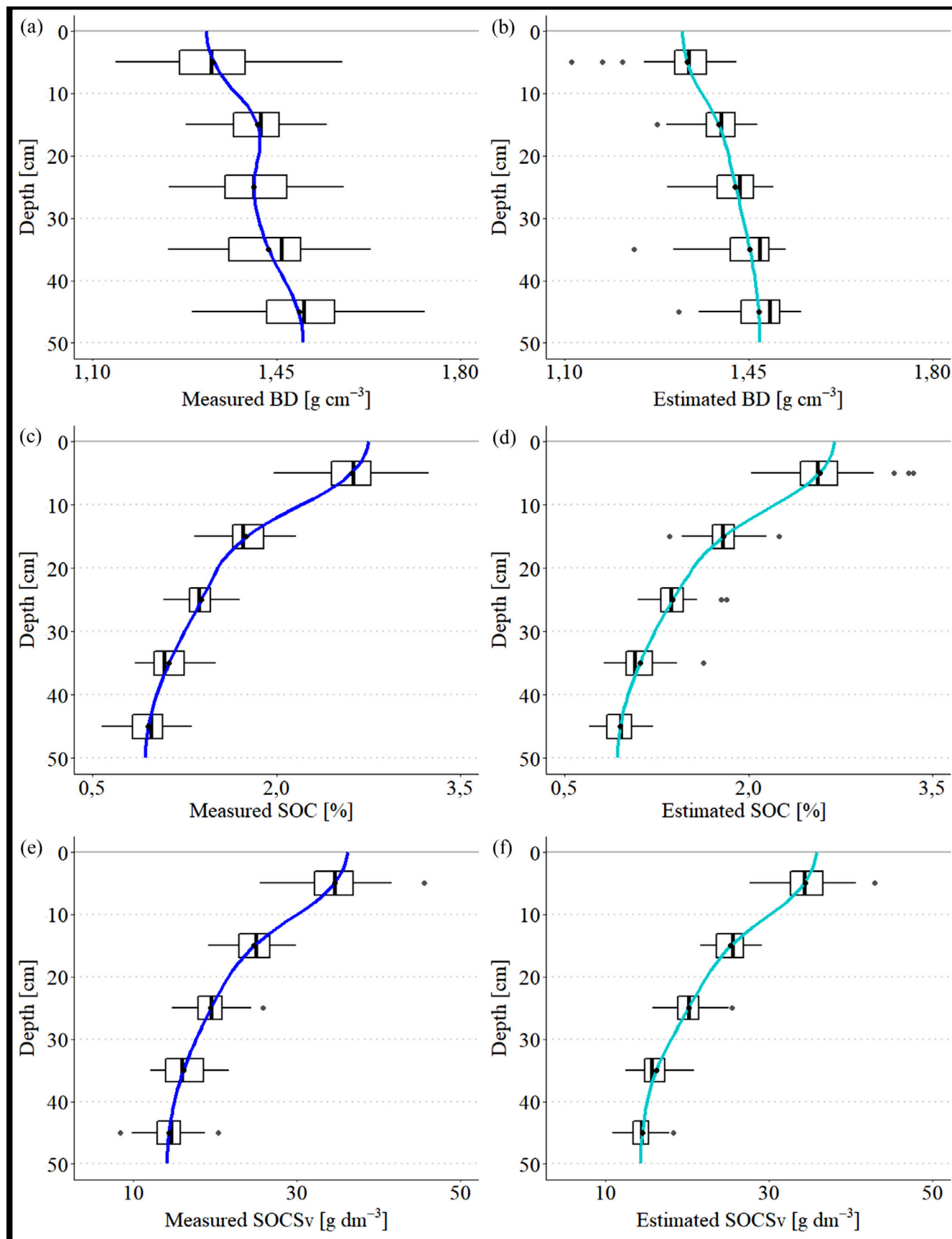
Furthermore, the splines confirmed the excellent performance of spectral models for SOC and SOCSv. Even though a reduction can be perceived in the interquartile ranges of the boxplots of the estimated data when compared to those of the measured data, in general, the ranges of the whiskers of the estimated and measured data are similar. However, it can be pointed out that the loss of performance of the SCOS spectral model in comparison with the SOC one is due to the incorporation of the BD in the estimated variables, since this property could not be estimated directly with the same degree of representativeness.

The results of the geostatistical analysis for the measured and estimated SOCS are shown in Figure 5 and Table 5. The semivariograms were adjusted to spherical

and exponential models, which were the same for the measured and estimated data at every depth range except for 10-20 cm. Moreover, for each of the semivariograms in the five sampling depths, the nugget and sill values were lower for the spectral vs. the measured data. Therefore, it can be stated that the variance is lower in the estimated data. Lower nugget and sill values for the spectral data have also been reported by other authors (Araújo *et al.*, 2015; Bonett *et al.*, 2016; Camacho-Tamayo *et al.*, 2017; Wetterlind *et al.*, 2008). Among the sampling depths, the greatest variability of the SOCS was evidenced in the semivariograms at the closest range to the soil surface, which is consistent with the natural addition processes taking place in the study area.

The use of spectrally estimated data in the construction of semivariograms did not modify the spatial variation tendency of the SOCS, as was verified in the obtained models and the interpolation surfaces illustrated in Figure 5. DSD shows moderate spatial dependence for each of the sampling depths, which is also consistent for the spectral and measured data. In addition, most R<sup>2</sup> and CVC values were above 0,70.

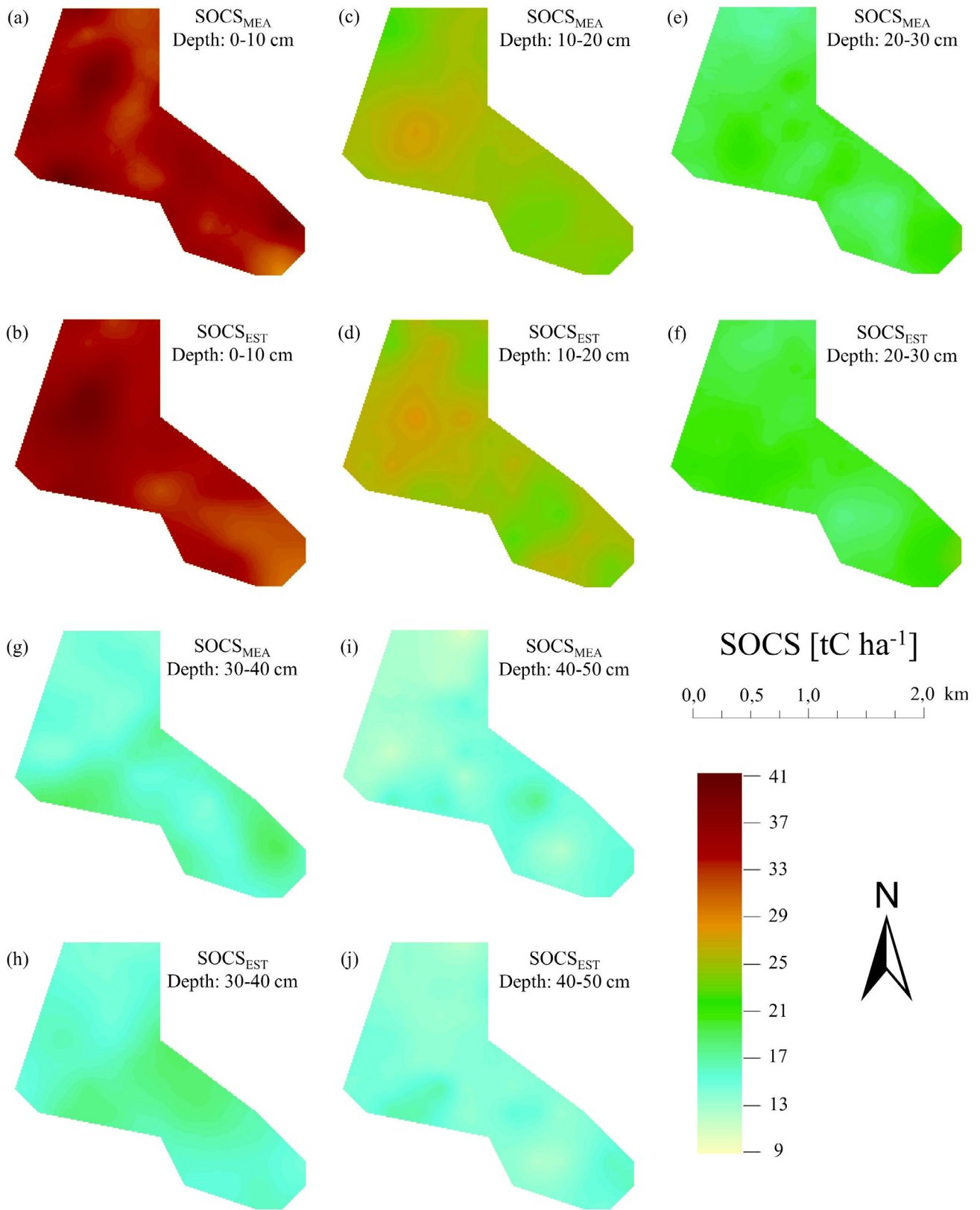
The geostatistical interpolation surfaces obtained from estimated and measured data showed a high degree of correspondence at each sampling depth. This is a positive indicator for approaching SOCS monitoring through spectral models, given that, once a sufficiently robust model is reached with adequate performance regarding the needs, the volume of information obtained could increase without the need to spend resources on conventional laboratory analyses, which is convenient for addressing soil studies in the context of climate change.



**Figure 4.** Depth splines for standard laboratory-measured (a, c, e) and spectral estimated data (b, d, f)

**Source:** Authors





**Figure 5.** Interpolation surfaces of SOCS from measured (a, c, e, g, i) and estimated (b, d, f, h, j) data  
**Source:** Authors

**Table 5.** Parameters of the theoretical semivariogram models for SOCS from standard laboratory-measured and spectrally estimated data

SCOS [tC ha <sup>-1</sup> ]	Depth [cm]	Semivariogram model	C <sub>0</sub>	C <sub>0</sub> + C <sub>1</sub>	Range [m]	DSD [%]	R <sup>2</sup>	CVC
Measured	0-10	Spherical	4,50	11,54	424,82	61	0,99	0,73
	10-20	Spherical	2,80	6,00	1 213,49	53	0,83	0,56
	20-30	Exponential	2,29	4,88	558,00	53	0,86	0,67
	30-40	Spherical	2,64	5,75	531,00	54	0,86	0,81
	40-50	Exponential	2,48	6,15	550,00	60	0,74	0,70
Estimated	0 - 10	Spherical	3,83	8,58	627,00	55	0,81	0,79
	10 - 20	Exponential	0,92	3,12	618,00	71	0,91	0,77
	20 - 30	Exponential	1,64	3,48	910,00	53	0,76	0,82
	30 - 40	Spherical	1,76	4,27	1 006,00	59	0,91	0,85
	40 - 50	Exponential	0,68	2,44	630,00	72	0,78	0,98

C<sub>0</sub>: nugget; C<sub>0</sub> + C<sub>1</sub>: sill; DSD: Degree of Spatial Dependence; CVC: Cross-Validation Coefficient

Source: Authors

## Conclusions

NIR spectroscopy allowed estimating, with a high degree of representativeness, the SOCS of oxisol in the study area. The results suggest that laboratory analyses for SOCS could be largely substituted or complemented by an approach involving spectral techniques in the NIR region. This work also showed that data derived from NIR models could be integrated into geostatistical evaluations for Colombian oxisols.

## Acknowledgements

The authors gratefully acknowledge the funding received from Universidad Nacional de Colombia via the project with HERMES code No. 37685, which allowed conducting this research.

## CRedit author statement

Fernández-Martínez, F., Camacho-Tamayo, J. H., Rubiano-Sanabria, Y.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing (original draft, review, and editing).

## Conflicts of interest

The authors declare no conflict of interest.

## References

Ahmadi, A., Emami, M., Daccache, A., and He, L. (2021). Soil properties prediction for precision agriculture using visible and near-infrared spectroscopy: A systematic review and meta-analysis. *Agronomy*, 11(3), 1-14. <https://doi.org/10.3390/agronomy11030433>

Al-Asadi, R. A., and Mouazen, A. M. (2014). Combining frequency domain reflectometry and visible and near infrared spectroscopy for assessment of soil bulk density. *Soil and Tillage Research*, 135, 60-70. <https://doi.org/10.1016/j.still.2013.09.002>

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

Allory, V., Cambou, A., Moulin, P., Schwartz, C., Cannavo, P., Vidal-Beaudet, L., and Barthès, B. G. (2019). Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions. *Science of the Total Environment*, 686, 764-773. <https://doi.org/10.1016/j.scitotenv.2019.05.192>

Araújo, S. R., Söderström, M., Eriksson, J., Isendahl, C., Stenborg, P., and Demattê, J. A. M. (2015). Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy. *Geoderma*, 237, 308-317. <https://doi.org/10.1016/j.geoderma.2014.09.014>

Askari, M. S., Cui, J., O'Rourke, S. M., and Holden, N. M. (2015). Evaluation of soil structural quality using VIS-NIR spectra. *Soil and Tillage Research*, 146, 108-117. <https://doi.org/10.1016/j.still.2014.03.006>

Ben Dor, E., Ong, C., and Lau, I. C. (2015). Reflectance measurements of soils in the laboratory: Standards and protocols. *Geoderma*, 245-246, 112-124. <https://doi.org/10.1016/j.geoderma.2015.01.002>

Bonett, J., Camacho-Tamayo, J. H., and Vélez-Sánchez, J. (2016). Estimating soil properties with mid-infrared spectroscopy. *Revista U.D.C.A Actualidad & Divulgación Científica*, 19(1), 55-66. [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0123-42262016000100007](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-42262016000100007)

Camacho-Tamayo, J. H., Forero-Cabrera, N. M., Ramírez-López, L., and Rubiano-Sanabria, Y. (2017). Evaluación de textura del suelo con espectroscopía de infrarrojo cercano en un oxisol de Colombia. *Colombia Forestal*, 20(1), 5-18. <http://dx.doi.org/10.14483/udistrital.jour.colomb.for.2017.1.a01>

- Camacho-Tamayo, J. H., Rubiano-Sanabria, Y., and Hurtado, M. del P. (2014). Near-infrared (NIR) diffuse reflectance spectroscopy for the prediction of carbon and nitrogen in an Oxisol. *Agronomía Colombiana*, 32(1), 86-94. <http://www.redalyc.org/9081/html/1803/180330697012/>
- Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., and Konopka, A. E. (1994). Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Science Society of America Journal*, 58(5), 1501-1511. <https://doi.org/10.2136/sssaj1994.03615995005800050033x>
- Cambou, A., Cardinael, R., Kouakoua, E., Villeneuve, M., Durand, C., and Barthès, B. G. (2016). Prediction of soil organic carbon stock using visible and near infrared reflectance spectroscopy (VNIRS) in the field. *Geoderma*, 261, 151-159. <https://doi.org/10.1016/j.geoderma.2015.07.007>
- Davari, M., Karimi, S. A., Bahrami, H. A., Taher Hossaini, S. M., and Fahmideh, S. (2021). Simultaneous prediction of several soil properties related to engineering uses based on laboratory Vis-NIR reflectance spectroscopy. *Catena*, 197, 1-12. <https://doi.org/10.1016/j.catena.2020.104987>
- FAO (2019). *Measuring and modelling soil carbon stocks and stock changes in livestock production systems Guidelines for assessment (Version 1)*. FAO. <http://www.fao.org/3/ca2934en/CA2934EN.pdf>
- Huang, J., Hartemink, A. E., and Zhang, Y. (2019). Climate and land-use change effects on soil carbon stocks over 150 years in Wisconsin, USA. *Remote Sensing*, 11(12), 1504. <https://doi.org/10.3390/rs11121504>
- IPCC (2019). *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Volume 4 - Agriculture, forestry and other land use*. The Intergovernmental Panel on Climate Change (IPCC). <https://www.ipcc-nggip.iges.or.jp/public/2019rf/vol4.html>
- Jia, X., Chen, S., Yang, Y., Zhou, L., Yu, W., and Shi, Z. (2017). Organic carbon prediction in soil cores using VNIR and MIR techniques in an alpine landscape. *Scientific Reports*, 7(1), 2144. <https://doi.org/10.1038/s41598-017-02061-z>
- Katuwal, S., Knadel, M., Norgaard, T., Moldrup, P., Greve, M. H., and de Jonge, L. W. (2020). Predicting the dry bulk density of soils across Denmark: Comparison of single-parameter, multi-parameter, and vis-NIR based models. *Geoderma*, 361, 1-10. <https://doi.org/10.1016/j.geoderma.2019.114080>
- Liland, K. H., Mevik, B.-H., and Wehrens, R. (2021). *pls: Partial least squares and principal component regression* (R package version 2.8-0). <https://cran.r-project.org/package=pls>
- Liu, S., Shen, H., Chen, S., Zhao, X., Biswas, A., Jia, X., Shi, Z., and Fang, J. (2019). Estimating forest soil organic carbon content using vis-NIR spectroscopy: Implications for large-scale soil carbon spectroscopic assessment. *Geoderma*, 348, 37-44. <https://doi.org/10.1016/j.geoderma.2019.04.003>
- Malone, B. (2016). *ithir: Functions and algorithms specific to pedometrics*. R package version 1.0/r126. <https://rdr.io/rforge/ithir/>
- Moreira, C. S., Brunet, D., Verneyre, L., Sá, S. M. O., Galdos, M. V., Cerri, C. C., and Bernoux, M. (2009). Near infrared spectroscopy for soil bulk density assesment. *European Journal of Soil Science*, 60, 785-791. <https://doi.org/10.1111/j.1365-2389.2009.01170.x>
- Nawar, S., and Mouazen, A. M. (2019). On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning. *Soil and Tillage Research*, 190, 120-127. <https://doi.org/10.1016/j.still.2019.03.006>
- Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B., and Montanarella, L. (2014). Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology and Biochemistry*, 68, 337-347. <https://doi.org/10.1016/j.soilbio.2013.10.022>
- Oliver, M. A., and Webster, R. (2015). *Basics steps in geostatistics: The variogram and kriging*. Springer. <https://doi.org/10.1007/978-3-319-15865-5>
- Ponce-Hernandez, R., Marriott, F. H. C., and Beckett, P. H. T. (1986). An improved method for reconstructing a soil profile from analyses of a small number of samples. *Journal of Soil Science*, 37(3), 455-467. <https://doi.org/10.1111/j.1365-2389.1986.tb00377.x>
- Poppiel, R. R., Lacerda P., M., Pereira, M., Almeida Junior, D. O., Demattê, J., Romero, D. J., Sato, M., Almeida J, L. R., and Moreira C, L. F. (2018). Surface spectroscopy of oxisols, entisols and inceptisol and relationships with selected soil properties. *Revista Brasileira de Ciencia Do Solo*, 42, 1-26. <https://doi.org/http://dx.doi.org/10.1590/18069657rbc20160519>
- Ramirez-Lopez, L., Reina-Sánchez, A., and Camacho-Tamayo, J. H. (2008). Variabilidad espacial de atributos físicos de un Typic Haplustox de los Llanos Orientales de Colombia. *Engenharia Agrícola*, 28(1), 55-63. <https://doi.org/10.1590/S0100-69162008000100006>
- Sommer, R., & Bossio, D. (2014). Dynamics and climate change mitigation potential of soil organic carbon sequestration. *Journal of Environmental Management*, 144, 83-87. <https://doi.org/10.1016/j.jenvman.2014.05.017>
- Stevens, A., and Ramírez-López, L. (2020). *An introduction to the prospectr package (R package Vignette R package version 0.2.0)*. <https://github.com/l-ramirez-lopez/prospectr>
- Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., Courcelles, V. de R. de, Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., ... Zimmermann, M. (2013). The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems and Environment*, 164, 80-99. <https://doi.org/10.1016/j.agee.2012.10.001>
- Viscarra Rossel, R. A., McGlynn, R. N., and McBratney, A. B. (2006). Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy. *Geoderma*, 137(1-2), 70-82. <https://doi.org/10.1016/j.geoderma.2006.07.004>
- Viscarra Rossel, R. A., and Webster, R. (2012). Predicting soil properties from the Australian soil visible - near infrared spectroscopic database. *European Journal of Soil Science*, 63, 848-860. <https://doi.org/10.1111/j.1365-2389.2012.01495.x>
- Wetterlind, J., Stenberg, B., and Söderström, M. (2008). The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. *Precision Agriculture*, 9(1-2), 57-69. <https://doi.org/10.1007/s11119-007-9051-z>