




Modeling the Spatial Distribution of Soil Organic Carbon and Carbon Stocks in the Casanare Flooded Savannas of the Colombian Llanos

Javier M. Martín-López¹  · Louis V. Verchot¹ · Christopher Martius² · Mayesse da Silva¹

Received: 8 March 2022 / Accepted: 4 June 2023 / Published online: 12 July 2023
© The Author(s) 2023

Abstract

Flooded savannas are valuable and extensive ecosystems in South America, but not widely studied. In this study, we quantify the spatial distribution of soil organic carbon (SOC) content and stocks in the Casanare flooded savannas. We sampled 80 sites at two soil-depth intervals (0–10 and 10–30 cm), where SOC values ranged from 0.41% in the surface and 0.23% in the sub-surface of drier soils to over 14.50% and 7.51%, in soils that experienced seasonal flooding. Spatial predictions of SOC were done through two digital soil mapping (DSM) approaches: Expert-Knowledge (EK) and Random-Forest (RF). Although both approaches performed well, EK was slightly superior at predicting SOC. Covariates derived from vegetation cover, topography, and soil properties were identified as key drivers in controlling its distribution. Total SOC stocks were 55.07 Mt with a mean density of $83.1 \pm 24.3 \text{ t}\cdot\text{ha}^{-1}$ in the first 30 cm of soil, with 12.3% of this located in areas that experience long periods of flooding (semi-seasonal savannas), which represented only 7.9% of the study area (664,752 ha). Although the study area represents only 15% of the total area of the Casanare department, the intensive pressure of human development could result in the reduction of its SOC stocks and the release of important amounts of greenhouse gases into the atmosphere. At regional level, the impact of a large-scale land use conversions of the flooded Llanos del Orinoco ecosystem area (15 Mha) could transform this area in a future source of important global emissions if correct decisions are not taken regarding the land management of the region.

Keywords Digital soil mapping · Random Forest · Expert knowledge · Carbon sequestration · Wetlands · Driving factors

Resumen

Las sabanas inundadas son ecosistemas extensos y valiosos en América del Sur, pero no han sido ampliamente estudiados. En este estudio, cuantificamos la distribución espacial del contenido y las reservas de carbono orgánico del suelo (COS) en las sabanas inundadas de Casanare. Se muestrearon 80 sitios a dos diferentes profundidades del suelo (0–10 y 10–30 cm), donde los valores de COS variaron de 0,41% en la superficie y 0,23% en la subsuperficie de los suelos más secos, a 14,50% y 7,51% en suelos que experimentan inundaciones estacionales. Las predicciones espaciales de COS se estimaron a través de dos enfoques de mapeo digital de suelos (MDS): Expert-Knowledge (EK) y Random-Forest (RF). Aunque ambos enfoques presentan un buen desempeño, EK fue ligeramente superior en la predicción del COS. Las covariables derivadas a partir de la cobertura vegetal, la topografía y las propiedades del suelo se identificaron como variables clave en explicar su distribución. Las reservas totales de COS fueron de 55,07 Mt, con una densidad media de $83,1 \pm 24,3 \text{ t}\cdot\text{ha}^{-1}$ en los primeros 30 cm del suelo, de los cuales el 12,3% se localizan en zonas que experimentan largos periodos de inundación (sabanos)

✉ Javier M. Martín-López
j.m.martin@cgiar.org

✉ Mayesse da Silva
m.a.dasilva@cgiar.org

¹ International Center for Tropical Agriculture (CIAT),
Multifunctional Landscapes, Km 17 Recta Cali-Palmira,
Valle del Cauca Z.C. 763537 - A.A. 6713 Palmira, Colombia

² Center for International Forestry Research (CIFOR)
Germany gGmbH, Bonn, Germany

semiestacionales), lo que representa solo 7,9% del área de estudio (664.752 ha). Aunque el área de estudio representa solo el 15% del área total del departamento de Casanare, la intensa presión producto del desarrollo humano podría resultar en la reducción de sus reservas de COS y la liberación de importantes cantidades de gases de efecto invernadero a la atmósfera. A nivel regional, el impacto de una conversión del uso del suelo a gran escala en los ecosistemas inundables de los Llanos del Orinoco (15 Mha) podría transformar esta área en una fuente futura de importantes emisiones globales si no se toman decisiones correctas en cuanto al manejo del suelo de la región.

Palabras clave Mapeo digital de suelos · Random forest · Expert Knowledge · Secuestro de carbono · Humedales · Factores de formación

Introduction

Tropical wetlands provide important ecosystem services, including global regulating roles in biogeochemical cycles and biosphere-atmosphere interactions. They store large amounts of carbon, are a sink for atmospheric CO₂, and make up around 60–80% of the natural atmospheric methane (CH₄) source (Chmura et al. 2003; Köchy et al. 2015; Rice et al. 2016; Saunio et al. 2020; Whiting and Chanton 2001; Zhu et al. 2017). In addition to being part of natural hydrology systems and to ensuring water quality and regulating its flow (IPBES 2019), they are also key reservoirs of biodiversity, they support fisheries, they are important sources of food, filter and retain sediments, and are culturally and aesthetically important to many communities and cultures (Mitsch et al. 2015; Shiel et al. 2016). One study estimated that 25% of the global value of ecosystem services is provided by wetlands (Costanza et al. 2014).

Wetlands soils, under regular anaerobic conditions, reduce the decomposition of organic plant materials, resulting in the accumulation of organic matter. This enables wetlands to effectively store substantial amounts of soil organic carbon (SOC) (Nahlik and Fennessy 2016). Despite wetlands only occupying about 5–8% of the earth's surface, (Mitsch and Gosselink 2007), point out that they store more SOC than all types of vegetation on earth. Wetlands are believed to store between 20–30% of the world's soil organic carbon (SOC), and in some cases, the SOC concentration can exceed 40% (Vepraskas and Craft 2016). This characteristic makes wetlands a key ecosystem for regulating water, climate, and biodiversity.

In this order, the development of accurate maps of wetlands and soil organic carbon (SOC) is crucial, as it enables an adequate inventory and monitoring of these carbon-rich ecosystems, leading to better quantifications of SOC stocks and greenhouse gas (GHG) emissions (Page et al. 2011). As stated by Minasny et al. (2020), a better and rapid estimation of peatlands distribution allows stakeholders to make informed decisions regarding the management, conservation, and restoration of these ecosystems. Nonetheless, the extent and state of tropical wetlands and the quantification of soil organic carbon stocks are still uncertain (Xu et al. 2018).

Köchy et al. (2015) estimated that they cover around 9% of the tropical land area and contain around 40 Pg of soil organic carbon (SOC), mostly in floodplains and marshes. In addition to uncertainty about the extent of tropical wetlands, the carbon density or carbon stock per unit area is also uncertain. Assessments of SOC storage vary widely due to both definition and methodological differences, from about 120 Gt to 535 Pg of SOC (Mitra et al. 2005). This is an important constraint in earth systems modelling, where lack of data on the carbon balance and gas fluxes in tropical wetlands limits accurate predictions about future climate and the importance of wetlands in climate change (Melton et al. 2013; Sjögersten et al. 2014).

South America has the largest extent of tropical wetlands, most of them found in floodplains where they are subject to seasonal fluctuations in water levels (Gumbrecht et al. 2017; Junk et al. 2013). The continent has large extents of seasonally flooded savannas, including the Pantanal, Bananal Island, and the Lavrado in Brazil; the Iberá in Argentina, the western Guianan flooded savannas in, Guiana, Venezuela, and Suriname; the Llanos de Moxos in Bolivia; and the Llanos del Orinoco (Barbosa et al. 2012; Hamilton et al. 2002), which is the subject of this study. The Llanos del Orinoco cover about 38.31 Mha (Barreto and Armenteras 2020) in the Orinoco River Basin of Colombia and Venezuela. The Llanos del Orinoco landscape is a mosaic of ecosystems that include forests, extensive areas of Moriche (*Mauritia flexuosa*) palms, upland savannas, and flooded savannas. The flooded area in this region shows significant interannual variation. Hamilton et al. (2004) found that the long-term mean inundation area was about 3.47 Mha, with a median area of 2.53 Mha, which makes it the second largest flooded savanna wetland in South America after the Brazilian Pantanal. For this study we focused on the flooded savannas of the Casanare in Colombia, which drains into the Meta River, one of the major tributaries of the Orinoco River that originates in the Andes. Few data exist for this region, but the area is undergoing rapid change with the introduction of exotic pasture grasses and extensive areas of monoculture cropping. Additionally, fire has been used in this landscape for centuries, although its duration and intensity in the flooded savannas typically is low due to the presence of rocky outcrops, sand dunes, and low biomass on sandy patches, which are spread throughout the landscape (Lasso et al. 2010).

SOC accumulation in flooded savannas appears to be related to length of flooding; short-term measurements in the flooded Colombian Llanos del Orinoco suggest stable to slightly accumulating SOC stocks (Vega et al. 2014). The objectives of the present study were to map the spatial distribution of SOC content and stocks throughout two digital soil mapping (DSM) approaches and determine the factors that influence their variability. To do this, we focused on the eastern Casanare Department in Colombia, where the largest extents of flooded savannas in the Orinoco Basin are located (Gumbrecht et al. 2017; Hamilton et al. 2004).

Material and Methods

Study Area

The study area (Fig. 1), known as the “Sabanas Inundables de Casanare” (Casanare flooded savannas) is part of an extensive floodplain ecosystem located in the “Llanos” ecoregion in the Orinoco River Basin (Olson et al. 2001). With a total area of 664,752 ha, the study area is located between the latitudes 6.30°N and 5.52°N and longitudes 71.07°W and 69.84°W. It has an elevation ranging from 78 to 146 m.a.s.l. and an average slope of 1.55 degrees. The annual average temperature in the region is approximately 26°C, with an average maximum temperature of 33°C during the dry season and an average minimum of 21°C during the

wet season. Rainfall distribution follows a unimodal regime with total annual average precipitation of 2,684 mm, characterized by a rainy season from April to November, during which most of the area is flooded, and a short dry season from December to March when the water drains and many of the flooded wetlands dry out (Buriticá Mejía 2016; Castillo-Figueroa et al. 2019; Lasso et al. 2010).

The land cover is mainly composed of native savanna, native forest, and permanent water bodies. Native savanna is sub-classified in this study as seasonal savanna, hyperseasonal savanna, and semiseasonal savanna, based on its vegetation and soil-water conditions (Cabrera-Amaya et al. 2020; Mora-Fernández and Peñuela-Recio 2013; Peñuela et al. 2014; Romero Duque et al. 2018; Sarmiento 1983). Seasonal savanna is a tree savanna, characterized by discontinuous grasses, dicotyledons and herbaceous vegetation, with the presence of scattered trees and shrubs. It develops under a bi-seasonal regime marked by a dry season (4 to 6 months) and a wet season, in which it remains flood free. In the dry season, the vegetation acquires a wilted appearance, and large numbers of plants die off, increasing fire frequency. This ecosystem is located mainly on plain-convex banks, continental dunes and sandy deposits, where coarse-textured soils, compaction, and low water availability create difficult conditions for vegetation growth. Hyperseasonal savanna is the most widely spread ecosystem in the region. It is mostly a treeless

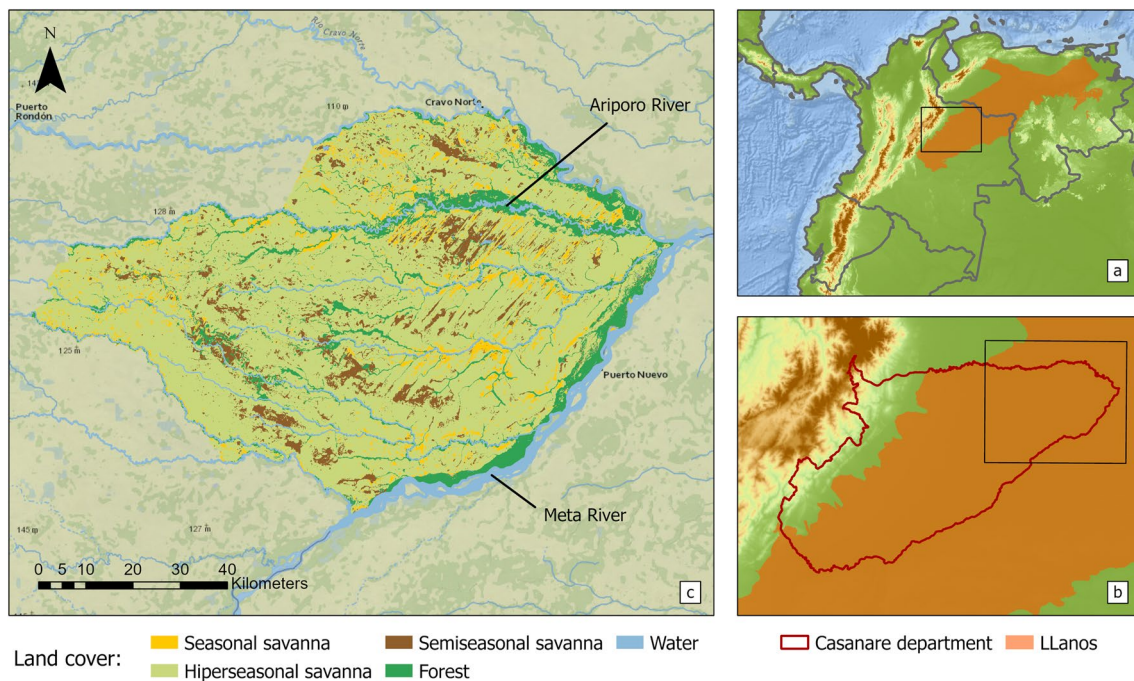


Fig. 1 Study location in the Casanare Flooded Savannas, Colombia. a) location within the tropical savannas ecoregion “Llanos” (Olson et al. 2001) b) position within the Casanare department (dark red line) and c) local land cover map of the study area

savanna, where perennial grasses and annual dicotyledons predominate and sporadic woody plants, shrubs, and palms are found. This ecosystem is temporarily flooded during the rainy season, where water table level can reach up to 20 cm above the surface, but it dries out quickly after the rain ends. It experiences four seasons: drought, onset of soil saturation, flooding, and decline in soil saturation. Soil texture is mostly represented by sandy-loam and clay-silty loam. Semiseasonal savanna is a treeless savanna, with very rare presence of woody plants or palms. Vegetation (grasses, sedges, march and aquatic plants) varies in function of flooding duration and water table level, which could reach up to 1 m above the surface. This ecosystem is found in topographic depressions with poorly drained conditions, and clay soils. Semiseasonal savannas experience long periods of flooding (between 8 and 11 months) and never present water deficit, even during the dry season, serving as an important food and water source in the region. Similarly, the native forest includes Moriche palm and gallery forests. Moriche palm corresponds to plant communities that are dominated by *Mauritia flexuosa* palms and numerous herbaceous and shrubby species. They are present in flooded areas or in soils with medium-fine textures that do not present water deficits during the dry season. Gallery forest, in turn, is dominated by a great variety of woody trees, palms, and shrubs and are developed close to permanent water bodies, such as riverbank levees, alluvial, and overflow plains of rivers, channels, and ravines.

For mapping purposes, a local land cover map derived from remote sensing images (Fig. 1) was created through a supervised image classification algorithm (further information in “Environmental Covariates” section). After classification, five categories were defined: seasonal savanna, hyperseasonal savanna, semiseasonal savanna, forest, and water. Gallery forest and Moriche palm were combined within a unique forest category since their spectral signatures were not clearly differentiated.

Soils of Casanare flooded savannas are predominantly Inceptisols and Entisols formed from alluvial, fluvio-torrential, and fluvio-glacial sediments deposited from the upper Pleistocene to the lower Holocene, upon which Aeolian materials subsequently accumulated (IGAC 2014). Relief is mainly composed of large extensions of aeolian and alluvial terraces, dissected by depressions and channels covered by fine materials (floodplains), which remain flooded for a great part of the year, and some elevated sandy areas called “Medanos” or continental dunes, which have been formed longitudinally by the northeasterly trade winds (IGAC 2014). The combination of the regional geomorphology and the high annual precipitation creates the hydrological connectivity between rivers and

wetlands, resulting in the maintenance of water throughout the region for most of the year (Sarmiento 1984).

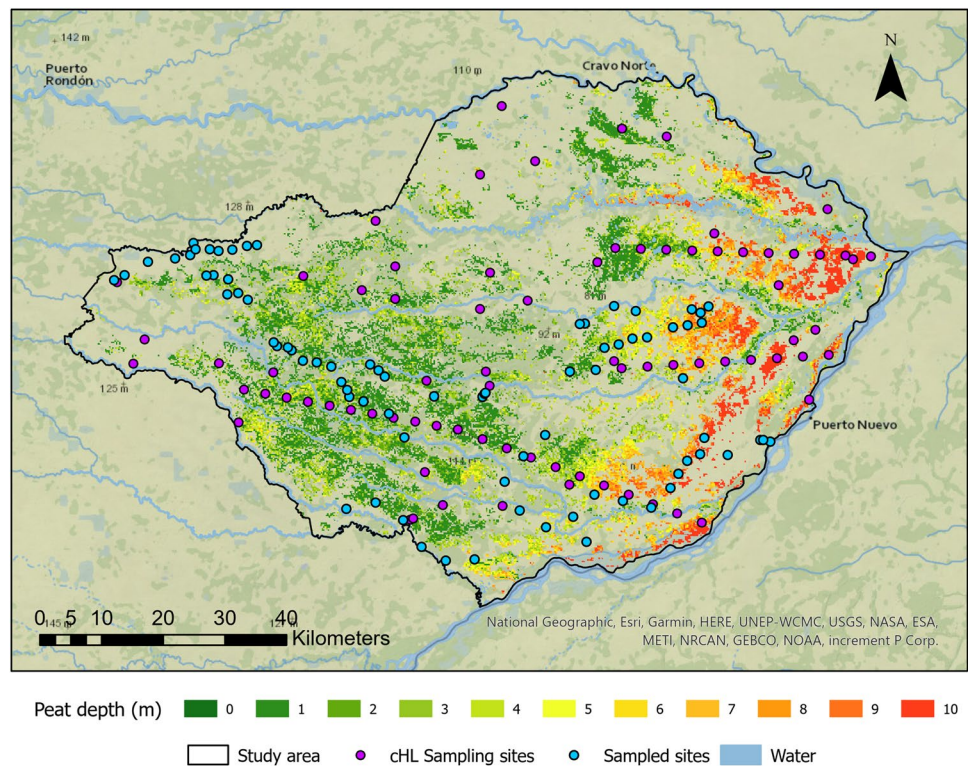
Data Collection and Soil Characterization

A total of 80 sites were sampled in February 2019 at two soil depth intervals (0–10 cm and 10–30 cm), resulting in 160 soil samples (disturbed and non-disturbed). The sites were sampled following two strategies: 43 sites located in four transects along the peat depth distribution predicted by the Global Wetlands Map (Gumbricht et al. 2017) and 37 sites distributed across the study area based on the conditioned Latin hypercube sampling (cLHS) strategy (Fig. 2). The cLHS method is a stratified random procedure for sampling variables from a multidimensional distribution of covariates, which defines sampling locations that capture the greatest variability among covariates (Minasny and McBratney 2006). Maps of covariates used in the cLHS to define the sampling site locations included: the Orinoquía land cover map at 30 m resolution; geology and lithology at a scale of 1:100,000 (IGAC 2014); peat depth and wetland predictions with a spatial resolution of 123 m (Gumbricht et al. 2017); landforms (Jasiewicz and Stepinski 2013) with a spatial resolution of 12 m (derived from the DEM); and the cost distance, which represents the cost of reaching each point in the landscape from the road network based on topography (slope at 12 m resolution) and distance from roads (Aitken 1977; Roudier et al. 2012). The Orinoquía land cover map of Colombia was developed from Landsat 5–8 images at a resolution of 30 m, the landforms were mapped using the DEM ALOS/PALSAR scaled up at a resolution of 12 m based on the geomorphons approach (Jasiewicz and Stepinski 2013), and cost distance was mapped according to Douglas, (1994). Due to the adverse field conditions, some sampling sites needed to be readjusted to locations with similar conditions, keeping original sampling characteristics (Fig. 2).

The soil samples were air-dried, sieved, and analyzed at the International Center for Tropical Agriculture (CIAT) laboratory for determining bulk density (BD) and soil texture (clay, sand, silt). BD and soil texture were performed by the volumetric ring (Blake 1965) and hydrometer (Bouyoucos 1936) methods, respectively. SOC, stable ^{13}C isotope ($\delta^{13}\text{C}$), total nitrogen (Nt), and stable ^{15}N isotope ($\delta^{15}\text{N}$) analyses were performed using the dry combustion method and mass spectrometry at the UC Davis Laboratory. Carbon to nitrogen ratios (C:N ratios) were calculated from SOC and Nt.

Other characteristics such as land cover (seasonal savanna, hyperseasonal savanna, semiseasonal savanna, gallery forest and Moriche palm); geomorphic classification (rise: slightly elevated areas (1–3% slope), talf: very low slope gradients (0–1%), dip: depressions, and continental dunes: aeolian mounds of sand (> 3% slope)) defined according to Haskins et al. (1999) and Soil Science Division

Fig. 2 Spatial distribution of the original cLHS sampling sites (purple points), field sampled sites (blue points) and peat depth (m) distribution predicted by Gumbrecht, et al. (2017)



Staff (2017); and generalized soil texture (sandy – coarser textures: sand, loamy sand, sandy loam; loamy – medium textures: loam, silt loam, silt, sandy clay loam, clay loam, silty clay loam; and clayey – finer textures: sandy clay, silty clay, clay) simplified from the soil texture taxonomy classification (Soil Science Division Staff 2017) at each sampling site were reported.

Statistical Analysis

Descriptive statistics of chemical and physical soil properties (0-10 cm and 10-30 cm) were done using SAS Studio. An ANOVA on log-transformed data, followed by the Student-Newman-Keuls test was carried out for normally distributed data, while the Kruskal-Wallis non-parametric test followed by a pair-wise separation using the Dwass, Steel, Critchlow-Flinger method was applied for data that could not be normalized.

SOC content at each depth was analyzed according to different environmental conditions: land cover, geomorphic feature, and generalized soil texture. Statistical analyses considered were data normality, homogeneity of variances, and design balance analysis tested using Shapiro-Wilk (“stats” R package), Levene (“car” R package), and the balance test (“nlme” R-package) functions. A Kruskal-Wallis non-parametric test followed by the Dunn post-hoc test was performed to identify significant differences among SOC

comparisons. In both cases, significant tests were carried out using the “PMCMRplus” R-package including a Bonferroni p-adjustment method.

Environmental Covariates

To map the distribution of SOC content at 0-10 cm and 10-30 cm using digital soil mapping (DSM) (McBratney et al. 2003) a set of 32 environmental covariates was considered (see Table S1 in the supplementary material). Terrain attributes were derived from the DEM – ALOS/PALSAR scaled up at 12 m spatial resolution (Shimada et al. 2014). Geomorphology maps at a scale of 1:100,000 were obtained from the geoportal of the Instituto Geográfico Agustín Codazzi (IGAC). Because of lack of soil texture differences at the two depths, soil texture was generated by using the DSM approach random forest at depth of 0-30 cm. Vegetation indices (VI) were calculated in Google Earth Engine from SENTINEL-2 spectral images for 3 distinct periods (dry: December (2018) – February (2019), wet: May (2018) – October (2018), transition dry-wet: November (2018) – December (2018)). Climate data (temperature and precipitation from 2010 to 2018) was obtained from CHELSA and MODIS spectral images. The local land cover (Fig. 1) map was created through a supervised classification method combining field data and SWIR-1 (B11), near-infrared (B8), and blue (B2) bands from Sentinel-2 images.

Digital Soil Mapping

DSM combines different environmental covariates according to the soil-forming factors defined by Jenny (1941) with field soil point data in a mathematical model to produce high-resolution soil maps (raster/pixel-based) that are continuum and variable in the space (McBratney et al. 2003; Moore et al. 1993; Zhu et al. 2001). Two DSM approaches, Expert Knowledge (EK) and Random Forest (RF), were used to predict and map SOC content distribution at 0–10 cm and 10–30 cm of soil depth. Both models were calibrated using a random set of 70% of the field data (training dataset; same points for each model) and the other 30% of the field data (validation dataset) were used for model assessment.

The EK model was based on the Soil Inference Engine (SIE) system described by Shi et al. (2004). This model combines expert knowledge of soil scientists with fuzzy logic by creating instance rules based on the spatial correlation between soil properties and environmental covariates (Ashtekar and Owens 2013; Heitkamp et al. 2020; Zhu 1997; Zhu et al. 2001). To perform the model, four units were defined according to the most dominant local land covers (Fig. 1) and their respective instance rules were set to create a Membership Function (MF) map (Zhu 1999, 1997) using the training dataset. The MFs were created using the FRBS and sets R-packages in which two-piece normal and sigmoid functions were identified as more frequent descriptors of the soil-landscape relationship within every unit. After individual environmental covariate's fuzzification was done, a weighted average was computed between the MF within every instance in order to get the overall optimality value for the unit.

The covariates used for mapping SOC with EK were chosen according to the SOC-landscape relationship observed during the sampling campaign, prior knowledge from the literature, information collected during the field trip, and the distribution of soil samples data across the environmental covariates. Scatter plots were created for each covariate using SOC content values as dependent variable to understand and create SOC-landscape relationships. Categorical covariates which did not provide clear differentiation by grouping many sites into each category were discarded.

The MF maps were used to map the continuous SOC content from the weighted average of the typical property values sampled in the field for each soil unit according to the following equation (Zhu 1997):

$$V_{ij} = \frac{\sum_{k=1}^n S_{ij}^k v^k}{\sum_{k=1}^n S_{ij}^k} \quad (1)$$

Where: V_{ij} is the estimated potential value at location i, j ; S_{ij}^k are the fuzzy membership values at each location (similarity values); v^k is a typical value of soil type k .

Random Forest (RF) is a machine learning method used to predict continuous and categorical variables from a large collection of non-correlated decision trees (Breiman 2001). The construction of each tree depends on the classification of the environmental covariates from the training dataset. To select the most appropriate covariates to enter the model, a Variance Inflation Factor (VIF) diagnosis was carried out and covariates with high multi-collinearity were identified (Alin 2010). VIF values close to zero represent null or scarce collinearity, while values above 5 indicate high collinearity between variables (Allam et al. 2020; Vu et al. 2015). In this study, the VIF assessment was done independently for each depth, removing variables with VIF greater than 5 in order to reduce overfitting and improve models' performances (Valbuena et al. 2017).

The RF was performed with the randomForest R-package (Liaw and Wiener 2003), using the training dataset and ntree: 1,000. In order to identify which covariates mainly influence the prediction of the SOC, we computed the variable importance assessment mean decrease in accuracy (%IncMSE), which is the percentage increase in mean squared errors (Zhou et al. 2019). The most important variables were defined as those that resulted in a significant reduction of predictive power when removed from the model (Behrens et al. 2014). The most relevant covariates for predicting SOC content based on the RF approach in the Casanare flooded savannas were established from the top-five most important covariates used at 0–10 cm and 10–30 cm soil depth models.

DSM Assessment

Model performance for EK and RF approaches was assessed based on three statistical indices applied to the validation dataset: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). Lower values of RMSE and MAE indicated better model performance, while higher values of R^2 indicated that the model explained a higher proportion of the variation.

SOC Stocks Spatial Distribution

SOC stocks (t ha^{-1}) at 0–30 cm soil depth were calculated for each sampling point by multiplying the SOC content and bulk density using equation (2) and subsequently compared under different geomorphic features and land cover conditions through a mean comparison test ("Statistical Analysis" section).

$$SOC \text{ Stock (ton/ha)} = SOC \cdot BD \cdot Depth \cdot (1 - G/100) \cdot 10 \quad (2)$$

Where: *SOC* is the soil organic carbon content (%); *BD* is the soil bulk density (Mg/m^3); *Depth* is the depth of the soil layer (*m*); *G* is the volume (%) occupied with coarse fragments (gravel or stones) and 10 is the conversion factor to t/ha. Note that sampled soils in the area did not contain coarse fragments.

The spatial distribution of SOC stocks was modeled using the best-performing DSM approach for SOC content at 0–10 cm and 10–30 cm and the respective BD maps developed previously for the study area with a random forest model. Basic statistics of predicted SOC stocks per land cover were calculated.

Results

Soil Properties Characterization

Descriptive statistics of the measured physicochemical soil properties are presented in Table 1. The highest mean values of SOC and Nt contents were found in the semiseasonal savannas and the lowest values were found in the seasonal savanna. Isotopic signatures of $\delta^{13}C$ indicate the relative proportion of carbon originating from trees and shrubs with a C_3 photosynthetic pathway or grasses with a C_4 photosynthetic pathway. As expected, depleted values in the forest stands showed carbon from predominantly C_3 sources, while both seasonal savanna and hyperseasonal savanna showed predominantly C_4 sources. Depletion was intermediate in the moriche palm and semiseasonal savanna showing mixed sources of SOC in these systems. The measurement of $\delta^{15}N$ provides an integrated index of the N cycle and all sites

Table 1 Descriptive statistics of chemical and physical soil properties for the different land cover in the flooded savanna landscape of Casanare. Values are presented for two soil layers 0–10 cm and 10–30

cm depths. Values are mean \pm SD and values for each parameter followed by the same letter are not significantly different from each other ($P < 0.05$)

	SOC (%)	$\delta^{13}C$ (‰)	Nt (%)	$\delta^{15}N$ (‰)	C:N
0–10 cm					
Semiseasonal savanna (<i>n</i> = 13)	8.71 \pm 0.94 ^a	-18.36 \pm 1.05 ^b	0.86 \pm 0.11 ^a	2.42 \pm 0.27 ^{bc}	10.79 \pm 0.55 ^{ab}
Forest (<i>n</i> = 10)	2.09 \pm 0.49 ^c	-26.18 \pm 0.97 ^c	0.17 \pm 0.04 ^c	3.68 \pm 0.41 ^{ab}	11.93 \pm 0.77 ^{ab}
Moriche palm (<i>n</i> = 5)	4.51 \pm 0.93 ^b	-21.01 \pm 0.60 ^{bc}	0.45 \pm 0.10 ^b	2.05 \pm 0.17 ^c	10.27 \pm 0.24 ^b
Hyperseasonal savanna (<i>n</i> = 27)	3.73 \pm 0.38 ^b	-15.62 \pm 0.31 ^a	0.31 \pm 0.04 ^b	3.37 \pm 0.27 ^b	12.70 \pm 0.29 ^a
Seasonal savanna (<i>n</i> = 25)	1.51 \pm 0.20 ^c	-15.00 \pm 0.33 ^a	0.12 \pm 0.01 ^c	4.74 \pm 0.26 ^a	12.39 \pm 0.36 ^a
10–30 cm					
Semiseasonal savanna	3.03 \pm 0.52 ^a	-17.47 \pm 1.04 ^b	0.29 \pm 0.06 ^a	4.42 \pm 0.39 ^b	11.08 \pm 0.63
Forest	1.50 \pm 0.56 ^{bc}	-23.35 \pm 1.37 ^c	0.14 \pm 0.06 ^{bc}	5.16 \pm 0.48 ^{ab}	11.21 \pm 0.83
Moriche palm	1.92 \pm 0.29 ^b	-20.32 \pm 0.61 ^{bc}	0.19 \pm 0.03 ^{ab}	3.33 \pm 0.43 ^b	10.69 \pm 0.51
Hyperseasonal savanna	1.84 \pm 0.20 ^{ab}	-14.67 \pm 0.36 ^{ab}	0.15 \pm 0.02 ^b	5.10 \pm 0.30 ^b	12.66 \pm 0.31
Seasonal savanna	0.81 \pm 0.10 ^c	-13.66 \pm 0.30 ^a	0.06 \pm 0.01 ^c	6.98 \pm 0.38 ^a	12.50 \pm 0.45
	BD ($g\ cm^{-3}$)	Sand (%)	Silt (%)	Clay (%)	SOC Stock ($t\ ha^{-1}$)
0–10 cm					
Semiseasonal savanna	0.74 \pm 0.06 ^c	30.8 \pm 2.9	38.5 \pm 3.1	30.7 \pm 2.3	59.06 \pm 3.62 ^a
Forest	1.10 \pm 0.09 ^b	34.4 \pm 5.2	42.1 \pm 4.6	23.5 \pm 2.3	19.42 \pm 2.74 ^c
Moriche palm	1.06 \pm 0.12 ^b	24.4 \pm 1.8	43.9 \pm 4.0	31.6 \pm 4.7	44.08 \pm 7.21 ^b
Hyperseasonal savanna	1.16 \pm 0.05 ^b	28.2 \pm 2.2	43.2 \pm 1.9	28.6 \pm 2.0	39.08 \pm 2.78 ^b
Seasonal savanna	1.42 \pm 0.02 ^a	42.8 \pm 5.1	34.9 \pm 4.0	22.3 \pm 1.9	20.59 \pm 2.47 ^c
10–30 cm					
Semiseasonal savanna	1.02 \pm 0.04 ^c	23.3 \pm 3.0	40.1 \pm 2.8	36.5 \pm 3.9	58.54 \pm 8.02 ^a
Forest	1.25 \pm 0.08 ^b	33.2 \pm 6.1	41.8 \pm 4.2	24.9 \pm 3.7	29.86 \pm 6.70 ^b
Moriche palm	1.19 \pm 0.11 ^b	28.6 \pm 4.7	42.5 \pm 3.7	28.9 \pm 3.9	43.81 \pm 5.38 ^a
Hyperseasonal savanna	1.27 \pm 0.03 ^b	27.1 \pm 2.4	45.0 \pm 2.2	27.8 \pm 1.7	44.13 \pm 3.73 ^a
Seasonal savanna	1.46 \pm 0.02 ^a	40.6 \pm 5.6	35.3 \pm 4.2	24.1 \pm 2.7	23.23 \pm 2.83 ^b

SOC Soil organic carbon; *Nt* Total nitrogen; $\delta^{13}C$ Stable 13C isotope; $\delta^{15}N$ Stable 15N isotope; *C:N ratio* Carbon to nitrogen ratio; *BD* Bulk density

showed $\delta^{15}\text{N}$ enrichment indicating an open or “leaky” N cycle (Craine et al. 2015). Enrichment was lower in the Moriche palm and semiseasonal savannas compared to the other land covers, suggesting that these ecosystems were less leaky. Carbon to nitrogen ratios were similar across land covers.

Soil bulk density ranged from 0.74 to 1.42 g cm⁻³ (0-10 cm) and from 1.02 to 1.46 g cm⁻³ (10-30 cm). Higher values of BD were measured in the seasonal savanna, which are characterized by a high sand content, while lower BD values were found in the semiseasonal savannas as a result of higher soil organic matter and clay contents. SOC stocks were significantly higher in clay soils and semiseasonal savannas compared to seasonal savanna and forest.

SOC Distribution Across the Different Environmental Conditions

Measured SOC content at soil depths of 0-10 cm and 10-30 cm and its distribution across the environmental conditions of geomorphic features, soil textures, and land cover are presented in Fig. 3. In general, SOC content was around 1.64% higher on average at soil depth 0-10 cm than at 10-30 cm in all environmental conditions.

Among the different geomorphic features, dips had significantly higher SOC content than talfs, rises, and continental dunes. Rise features did not show differences

with talfs and continental dunes, while these last two differed significantly at both soil depths. The highest SOC content of the study area (14.5%) was found at soil depth 0-10 cm in the dip feature, while the highest value in talf, rise, and dune for the same depth were 6.3%, 5.5%, 1.3%, respectively. The largest mean reduction of SOC content between soil depths occurred in the dip feature, where mean SOC content decreased from 7.2% at 0-10 cm to 2.82% at 10-30 cm, followed by talf (2.48% » 1.28%), rise (2.01% » 1.02%) and continental dune (0.61% » 0.38%).

Differences in SOC content were also observed among the different land covers. For soil depth 0-10 cm, SOC content was higher in semiseasonal savannas, with values ranging from 4.2% to 14.5%, and lower in seasonal savanna, with values ranging from 0.4% to 3.6%. For soil depth 10-30 cm the same pattern was observed, but the differences in SOC content among land covers were lower than those observed in the surface layer.

Soil texture also played an important role in SOC accumulation for this region. Although loamy soils presented SOC content at the surface up to 14.5%, the average was 3.8% and this was much lower than the minimum SOC content (4.2%) found in clayey soils. Coarse textured soils had the lowest SOC content, with values inferior to 3.6% (0-10 cm). At 10-30 cm soil depth, higher values of SOC were found on clayey and loamy soils and lower content in coarse soils.

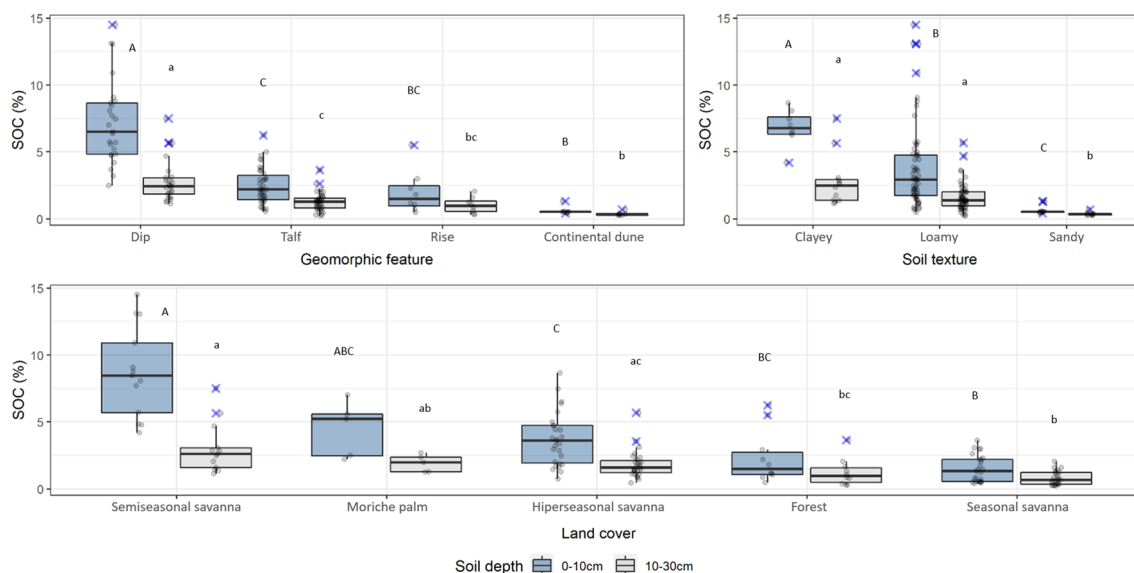


Fig. 3 Boxplot diagrams showing soil organic carbon (SOC) content (%) distribution at soil depth 0-10 cm and 10-30 cm under different environmental conditions in the Casanare flooded savannas, Colombia. Grey dots represent single values of sampled soils, the black horizontal line in the middle of the box represents the median value, lower and upper boundaries indicate the 25th and 75th percentiles,

blue crosses represent outliers, vertical black lines (whiskers) denote the maximum and minimum values (excluding outliers). Uppercase and lowercase letters denote statistically significant differences at $P < 0.05$ using the Dunn post-hoc test at 0-10 cm and 10-30 cm, respectively

Table 2 Model performance of soil organic carbon (SOC) content (%)

Soil depth	Method	RMSE	MAE	R ²
0–10 cm	RF	1.76	1.48	0.49
	EK	1.81	1.45	0.42
10–30 cm	RF	0.74	0.59	0.32
	EK	0.70	0.53	0.36

EK Expert knowledge; RF Random forest; RMSE Root mean squared error; MAE Mean absolute error; R² Correlation coefficient

Assessment of DSM Approaches

The performance assessment using the validation dataset showed small differences of RMSE, MAE, and R² between the EK and RF approaches at both soil depths, with smaller prediction errors (RMSE and MAE) at the 10–30 cm depth (Table 2). The RF presented lower values of RMSE and higher values of MAE and R² compared to the EK at 0–10 cm, while at 10–30 cm EK presented lower values of RMSE and MAE and higher R² values. Figure 4 shows the scatter plots of the relationship between the dataset used for SOC (%) model validation using RF and EK approaches and the measured data.

Digital Soil Mapping of SOC Content

Prior to implementing the EK and RF models, environmental covariates for each approach were selected. Covariates selected for implementing the EK at 0–10 cm and 10–30 cm were land cover, moisture stress index for the dry period – MSI (dry), green-red vegetation index for the dry and transition: wet-dry period – GRVI (dry and trans), relative slope position – RSP, clay, silt, sand, and the SAGA wetness Index, which is a 'Topographic Wetness Index' (TWI), based on a modified catchment area

calculation. For the RF model, an independent set of covariates was selected for each depth after the elimination of highly collinear covariates through the VIF assessment (see Table S1 in supplementary material). Figure 5 shows the top-five most relevant covariates for predicting SOC ranked by the %IncMSE for the RF model at each depth. The variable importance ranking identified local land cover, normalized difference vegetation index for the wet period – NDVI (wet), sand, and SAGA wetness index as common covariates for predicting SOC at both depths. Local land cover was found to be the most important covariate at 0–10 cm, contrary to the 10–30 cm depth in which it was placed at the bottom of the ranking. Normalized difference moisture index for the transition: wet-dry period – NDMI (trans), in turn, was the most important variable at 10–30 cm. NDVI (wet) and sand remained at the same position of the ranking at both depths, while the SAGA wetness index was more important for the 10–30 cm depth. Several covariates, such as local land cover, SAGA wetness index and sand, were utilized in both EK and RF approaches. Additionally, covariates regarding the vegetation properties were the most frequently used in both DSM approaches.

Maps of SOC content with a spatial resolution of 12 m developed using both the EK and RF methods at soil depths of 0–10 cm and 10–30 cm are shown in Fig. 6. In general, RF overestimated the minimum and underestimated the maximum SOC contents while EK underestimated both minimum and maximum (Table 3). Comparing between minimum values of the predicted SOC content maps (EK and RF) and the measured data, we observed smaller differences for EK than RF at both depths, while RF had smaller differences for predicting the maximum values. At 0–10 cm, both approaches presented similar mean_w SOC content of 4.20% (RF) and 4.22% (EK), representing a higher mean_w prediction in comparison to the measured mean_w value (3.63%). At soil depth 10–30 cm,

Fig. 4 Scatter plots showing the relationship between the dataset used for SOC (%) model validation using RF (blue dots) and EK (red dots) approaches and data measured in the Casanare flooded savannas, Colombia for 0–10 cm and 10–30. Dots represent the values for each variable while the lines represent the best fit for the correlation between them

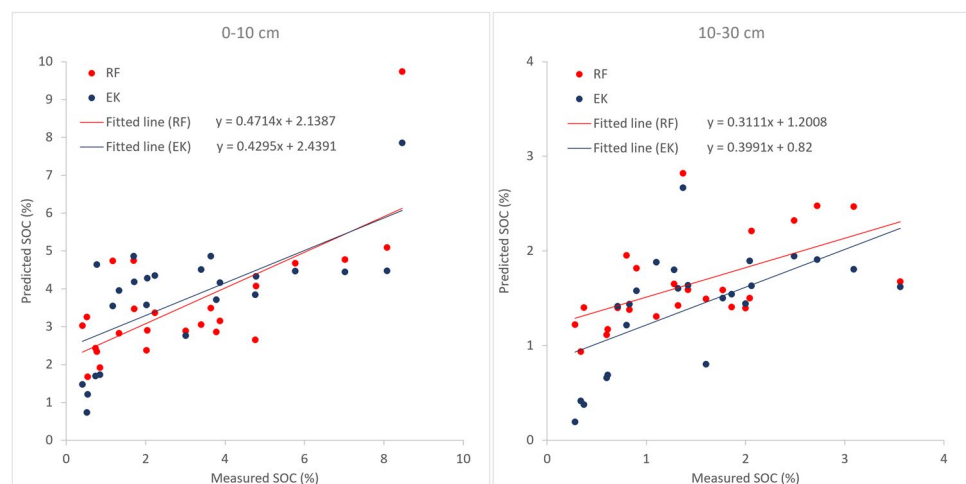
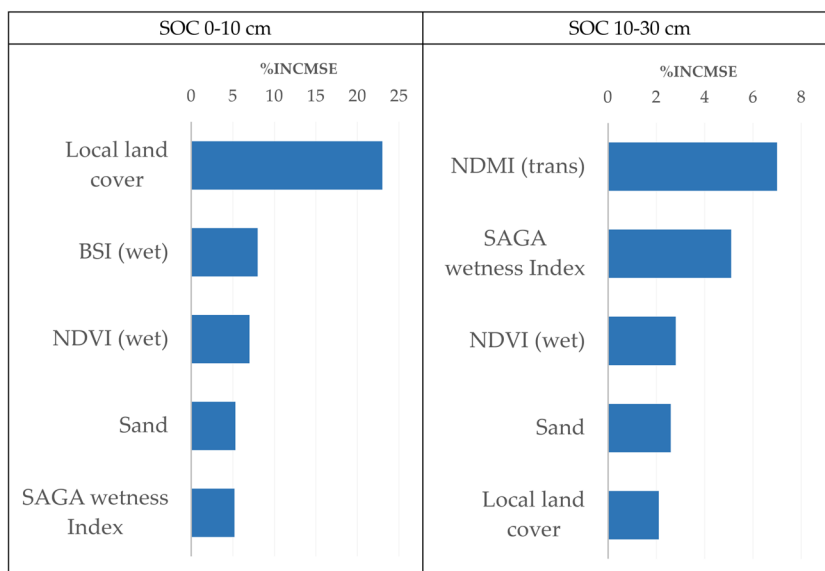


Fig. 5 Importance of environmental covariates used in the RF models for the study area according to the mean %IncMSE (percentage increase in mean square error). High values represent a high importance of the covariates in predicting soil organic carbon (SOC) content (%). BSI: Bare Soil Index, NDVI: Normalized Difference Vegetation Index, NDMI: Normalized Difference Moisture Index, wet: wet period -> May (2018) - October (2018), trans: transition dry-wet: November (2018) - December (2018)



RF overestimated the mean_w value compared to the measured value (difference of 0.28%), while EK slightly underestimated the mean_w values (difference of 0.06%).

In general, the two approaches presented a similar pattern regarding the spatial distribution of SOC content (Fig. 6). The highest values of predicted SOC content were found at locations

that experienced longer periods of seasonal flooding (darker areas) and the lowest values were predicted in the continental dunes with sandy soils and no seasonal flooding. However, there are differences in the predicted SOC values in some areas such as the large gallery forests located in the floodplains of the Ariporo and Meta Rivers where RF estimated higher values of SOC

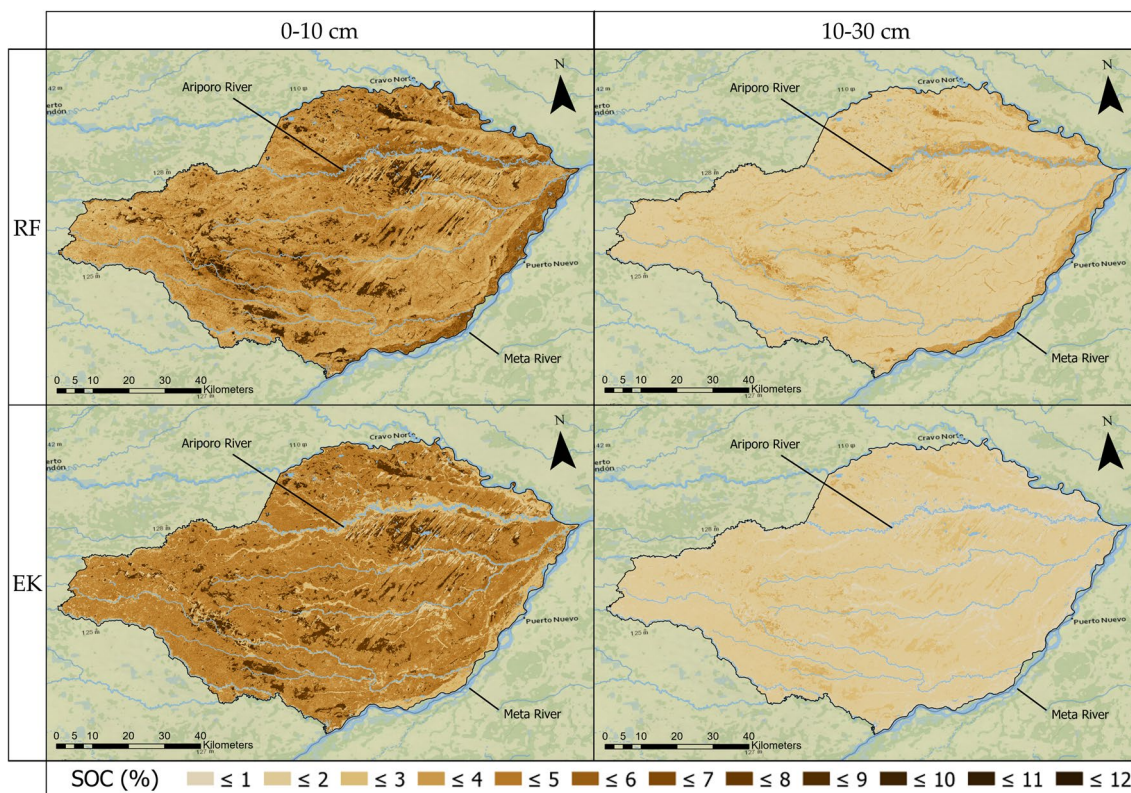


Fig. 6 Spatial distribution of SOC (%) using two different DSM approaches (EK and RF) in the Casanare flooded savannas, Colombia at depth intervals: 0-10 cm and 10-30 cm

Table 3 Descriptive statistics of SOC (%) predicted using EK and RF approaches and measured data collected in the Casanare flooded savannas, Colombia for 0-10 cm and 10-30 cm

Depth	Method	Min	Max	Mean _w	sd _w
SOC (%) (0-10 cm)	Field Data	0.41	14.52	3.63	1.90
	Map (RF)	0.85	11.92	4.20	0.85
	Map (EK)	0.30	11.30	4.22	0.51
SOC (%) (10-30 cm)	Field Data	0.23	7.51	1.65	1.08
	Map (RF)	0.48	5.09	1.93	0.45
	Map (EK)	0.10	4.81	1.59	0.19

Mean_w: weighted arithmetic mean and sd_w: weighted arithmetic standard deviation based on the proportion of data per land cover

for both soil depths (0-10 cm and 10-30 cm). Predicted values of SOC in the hyperseasonal savannas tend to be lower for the RF than for the EK model at 0-10 cm; while at 10-30 cm, predicted SOC values across the hyperseasonal savannas are similar.

SOC Stock Distribution

SOC stocks over the entire 30 cm profile determined using field data and scaling by geomorphic feature and land cover type are summarized in Fig. 7. Comparisons between measured SOC stocks across the different geomorphic features and land covers showed significant differences ($P < 0.05$) between dip ($111 \pm 30 \text{ t ha}^{-1}$) and the other features, while no significant differences were identified between talf ($63 \pm 24 \text{ t ha}^{-1}$), rise ($49 \pm 28 \text{ t ha}^{-1}$), and continental dune ($20 \pm 8 \text{ t ha}^{-1}$). On an ecosystem basis, semiseasonal savannas ($118 \pm 34 \text{ t ha}^{-1}$) differed significantly from forests ($48 \pm 26 \text{ t ha}^{-1}$) and seasonal savannas ($44 \pm 25 \text{ t ha}^{-1}$) and did not show significant differences with Moriche palm ($88 \pm 28 \text{ t ha}^{-1}$) or hyperseasonal savanna ($83 \pm 29 \text{ t ha}^{-1}$). Overall, higher mean measured values were attributed

to semiseasonal savanna land cover and dip geomorphic features, where lower values of BD and higher values of SOC contents were identified (Table 1). Lower SOC stock values were attributed to seasonal savanna land cover and continental dunes, which were characterized by lower SOC contents and high BD.

Due to the slightly better capacity of EK to describe and map the spatial distribution of SOC content for the study area, the map of SOC stocks at 0-30 cm soil depth was developed using this approach and selected covariates described in “Digital Soil Mapping” section. The corresponding model performance parameters were RMSE: 26 -, MAE: 23, and R^2 : 0.69.

(Circles indicate the larger areas of higher SOC stock accumulation)

Predicted SOC stocks ranged from 6 to 210 t ha^{-1} , with average stocks of $83.13 \pm 24.32 \text{ t C ha}^{-1}$ (Fig. 8). Similar to SOC content, high ($150\text{-}180 \text{ t ha}^{-1}$) and very high ($> 180 \text{ t ha}^{-1}$) SOC stocks were found in areas characterized by their proximity to water bodies, their location in floodplains with clay-loam soils, and the presence of long-term flooded conditions (circles indicate the larger areas of higher SOC stock accumulation). Medium-high values ($120\text{-}150 \text{ t ha}^{-1}$) were found adjacent to areas with higher SOC stocks. The hyperseasonal savanna land cover, which occupies the larger proportion of the study area, is characterized as a mix mainly between medium-high and medium SOC stocks ($90\text{-}120 \text{ t ha}^{-1}$) in dip features and the medium and medium-low ($60\text{-}90 \text{ t ha}^{-1}$) stocks in well-drained talf features. Low values ($30\text{-}60 \text{ t ha}^{-1}$) of SOC stocks were largely predicted in riversides along the Ariporo River and the Meta River as well as along other rivers, where coarser soil textures predominate. Very low values ($< 30 \text{ t ha}^{-1}$), in turn, were characteristic of areas where continental dunes (northeasterly elongated geomorphic features) and sandy patches predominate.

Fig. 7 Boxplot diagrams showing SOC stocks (t ha^{-1}) distribution at soil depth 0-30 cm under different environmental conditions. Grey dots represent single values of sampled soils, the black horizontal line in the middle of the box represents the median value, lower and upper boundaries indicate the 25th and 75th percentiles, blue crosses represent outliers, vertical black lines (whiskers) denote the maximum and minimum values (excluding outliers). Lowercase letters denote statistically significant differences at $P < 0.05$ using the Dunn post-hoc test

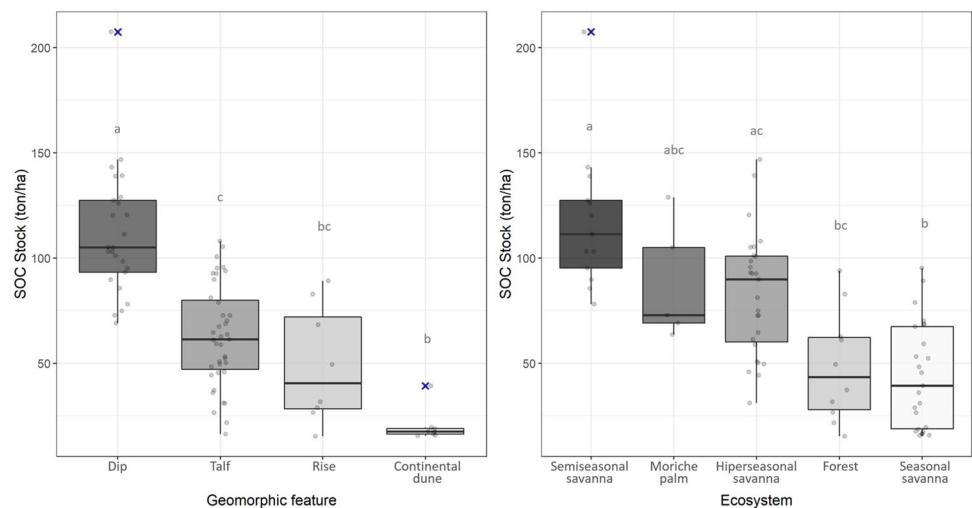
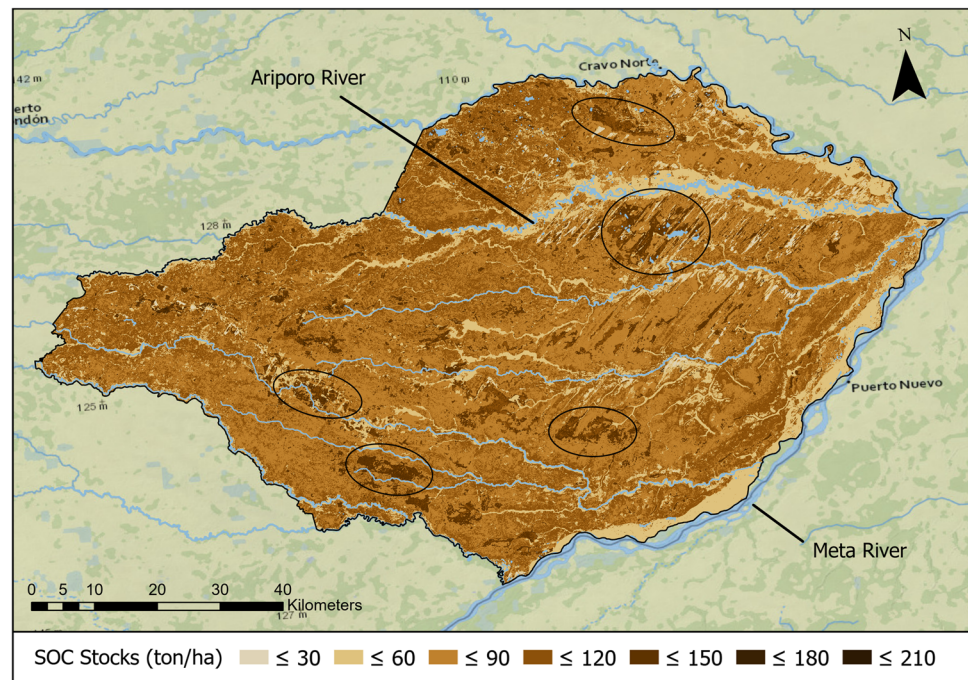


Fig. 8 Map of SOC stocks ($t\ ha^{-1}$) predicted from EK model at 0–30 cm soil depth in the Casanare flooded savannas, Colombia.



Using the land cover map (Fig. 1) as a mask, descriptive statistics of the predicted SOC stock map per land cover were calculated (Table 4). We found that the highest mean predicted SOC stocks values were in the semiseasonal savannas, followed by the hyperseasonal savanna, seasonal savanna, and forest, which combines riparian forest and Moriche palm. Total soil organic carbon stock predicted for the first 30 cm of soil depth in the entire study area (Fig. 8) was equivalent to 55.07 Mt of C. Hyperseasonal savannas contributed 65% (42.73 Mt) of the total stocks, while forest contributed with 22% (3.55 Mt), semiseasonal savanna with 7% (6.79 Mt), and seasonal savanna with 6% (2.00 Mt).

Discussion

SOC Characterization

Carbon accumulation in tropical wetlands is typically the result of small differences between input and outputs, and

is often associated with topography that concentrates water and creates anaerobic conditions that slow decomposition (Gumbrecht et al. 2017; Paul 2016; Sjögersten et al. 2014). Carbon accumulation is also regulated by the types and quality of vegetation, particularly the proportion of plant-derived C and N that is incorporated into soil organic matter, and by soil matrix interactions that control its stabilization (Cotrufo et al. 2013; Paul 2016). Clay and silt fractions play important roles in stabilizing soil organic matter through both physical and chemical protection. We found strong associations between geomorphology, plant communities, and soil texture that affected the SOC concentrations in the flooded Llanos del Orinoco and in the surrounding ecosystems. There are few data from South American floodplain wetlands for comparison and available data do not represent systematic sampling, but we found that mean SOC contents are higher in the Casanare flooded savannas compared to the levels reported for flooded savannas in the Brazilian Pantanal (Vega et al. 2014), the Brazilian Cerrado (Wantzen et al.

Table 4 Basic statistics of predicted SOC stocks ($t\ ha^{-1}$) per land cover within the study area

Land cover	Range SOC stock ($t\ ha^{-1}$)	Mean \pm SD SOC stock ($t\ ha^{-1}$)	Total SOC stock (Mt)	Landscape area occupied (%)
Water	-	-	-	0.5
Seasonal savanna	6.27 – 119.83	47.13 \pm 14.16	2.00	6.39
Semiseasonal savanna	25.68 – 210.52	129.59 \pm 11.43	6.79	7.89
Forest	22.88 – 108.38	42.64 \pm 8.19	3.55	12.51
Hyperseasonal savanna	36.61 – 138.02	88.39 \pm 10.33	42.73	72.71
Total			55.07	100

2012) or the Bolivian Llanos do Moxos (Boixadera 2003), and they were similar to flooded savannas in Roraima, Brazil (Barbosa et al. 2012).

We found high SOC contents in locations characterized by clayey soils with high grass productivity (semiseasonal savannas), and depressions (dip features). SOC accumulation in these areas have been favored by frequent and excessive precipitation, high concentrations of sediments, and long periods of water accumulation (Barthelmes and Joosten 2018; Sarmiento 1984; Mora-Fernández et al. 2015). Sediment transportation processes occurred during the Tertiary and Quaternary deposited coarser materials, as a result of the erosion of the Andes mountains, along most of the river edges (Goseen 1964), forming banks that do not store much SOC. We found low SOC contents in riparian areas, in continental dunes, and in sandy deposits that are characterized by low water availability, coarser soil textures, and low-density vegetation (seasonal savanna). In riparian areas located specifically in silty-sandy banks, within the first 30 cm prevail dry conditions during a long part of the year, which affects the accumulation of SOC. Aeolian erosion by the Trade winds has acted as an essential factor in shaping the regional topography, particularly in the formation of continental dunes and sandy deposits (FAO 1965). Intermediate SOC contents were associated with Moriche palm and hyperseasonal savanna vegetation and loamy textured soils.

DSM Models Assessment

We compared two different DSM approaches (EK and RF) for predicting SOC content at 0-10 cm and 10-30 cm. The models performed comparably to map SOC variability in the study area and gave a similar spatial prediction pattern. The models both predicted high SOC contents in dip features, where soil textures are finer, flooded conditions last longer, and vegetation has higher biomass production rates, and low SOC values on the continental dunes with coarser soils, low-density vegetation, and no seasonal flooding. However, with validation against field data, we found that the spatial predictions obtained with the EK represented field conditions better. A key major difference between the modelling approaches was associated with the riparian areas along Ariporo and Meta rivers. The RF model predicted much higher values for these areas than the EK model, which showed predictions that were more consistent with field observations. Additionally, EK showed lower prediction errors than RF. Both models were sensitive to the presence of outliers. RF underestimated the highest values and overestimated the lowest, while EK underestimated the highest and lowest values and gave overall predictions around the mean values.

Lower performance in RF was attributed intrinsically to the methodology. The algorithm estimates the values for each pixel, based on an average of the results of all the

independent trees created in the model (Callens et al. 2020), and generates an approximation to the field values based on averages of most similar areas, which limits its ability to describe local extreme values (Szabó et al. 2019). This situation probably caused the misprediction along the rivers, which may have been influenced by other areas with similar environmental conditions. RF has been used effectively in several studies for estimating SOC at landscape scale (Grimm et al. 2008; Wang et al. 2020; Zeraatpisheh et al. 2019; Zhou et al. 2019). Its prediction power has been shown to outperform other DSM approaches, specifically in larger study areas with a diversity of landscape features (Lamichhane et al. 2019), capturing most of the variance of the soil characteristics (Sindayihebura et al. 2017). Other advantages of RF include the modelling of high dimensional data, which would be difficult to afford with EK; the estimation of internal errors and the variable importance assessment; as well as a lower effort for data preparation and modelling time (Breiman 2001; Camera et al. 2017; Grimm et al. 2008). However, Ashtekar and Owens (2013) mentioned that a disadvantage of this method could be that it only works with soil-landscape relationships at specific locations and does not take into consideration extensive differentiation across the landscape. Additionally, Hengl et al. (2015) stated that RF is effective only within the range of values used in the training data, which provides irregular predictions for observations outside of the range (Terra 2017).

EK, in turn, is based mainly on qualitative and quantitative criteria provided by the knowledge of an expert, that takes into account soil-landscape relationships for a specific region based on experience, analysis from sampling data, and knowledge of interactions between covariates that influence the property distribution (Ashtekar and Owens 2013; Menezes et al. 2013; Ngunjiri et al. 2019). The use of EK led to a better control and implementation of inference rules that explain soil-landscape relationships, even with limited field data (Menezes et al. 2018). EK approaches have been satisfactorily applied by other authors (Akumu et al. 2015; Ashtekar et al. 2014; da Silva et al. 2016; Minai et al. 2021; Tsakiridis et al. 2019; Zhu et al. 2010). As evidence of its good performance, this approach has been widely used for the development of soil surveys by the US Department of Agriculture (USDA) (Zhu et al. 2001). On the other hand, EK is time consuming in terms of data selection, analysis, and processing. It requires knowledge about the mapping area and soil-landscape relationships, which may be limited for larger areas.

Drivers of SOC Spatial Variability

Similar to other tropical landscapes studies: (Diwediga et al. 2017; Hamzehpour et al. 2019; Yang et al. 2007), in the Casanare flooded savannas, the vegetation coverage,

topography, and soil texture were found to be the most influential factors of the SOC distribution. The combination of these 3 features leads to different environmental conditions that consequently influence the spatial distribution of SOC, as evidenced in this study.

Vegetation coverage not only controls the inputs of organic material but also protects soils from erosion and conserves soil moisture (Lamichhane et al. 2019; Weil and Brady 2017). Several studies have represented the vegetation coverage in spatial models across the landscape employing categorical land cover maps or continuous vegetation indices derived from remote sensing (Gray et al. 2015; Minai et al. 2021; Mondal et al. 2017; Wan et al. 2019; Wang et al. 2020). Land cover maps and vegetation indices help to classify vegetation in terms of types, biomass production, residues, and soil textures based on changes in leaf colors and leaf density (Motohka et al. 2010). Specific combinations of these characteristics delimit areas from higher carbon fixation such as the semiseasonal savanna to lower carbon fixation such as the seasonal savanna. Vegetation indices also show where to expect more plant water stress and less soil moisture content (Elhag and Bahrawi 2017), being relevant not only for soil moisture variability but also for SOC spatial modeling (Welikhe et al. 2017). Topography plays an imperative role in SOC storage across the landscape (Patton et al. 2019; Wiesmeier et al. 2019). It widely describes water flow paths, water accumulation, soil erosion, and sedimentation, and consequently influences the soil moisture, microbial activity, and SOC accumulation (Sørensen et al. 2006; Yun et al. 2019). Topographic covariates that indicate potential moist/dry areas such as topographical wetness index improves SOC predictions (Adhikari et al. 2014; Lamichhane et al. 2019; Minasny et al. 2013; Siewert 2017; Taghizadeh-Mehrjardi et al. 2017; Wiesmeier et al. 2013).

Additionally, as identified in several studies (Funes et al. 2019; Hounkpatin et al. 2018; Yang et al. 2007; Zinn et al. 2005), the results also showed that texture was also key in explaining the distribution of SOC. Coarser soil textures were correlated negatively with SOC, while finer textures showed a positive relationship. Clay soils have a greater capacity to preserve or protect organic matter from microbial attack as well as to stabilize the organic matter through mineral–organic matter binding. By promoting soil aggregations, silt and clay contents improve SOC protection (Six et al. 2002; Xie et al. 2021), enhance soil moisture (Augustin and Cihacek 2016; Yang et al. 2016), water holding capacity, and plant available water. These conditions reduce or avoid SOC losses by soil erosion, enhance plant biomass production, and, consequently, conserve or accumulate SOC. On the other hand, sandy soils do not physically or chemically protect SOC, they favor nutrient losses by leaching affecting plant density, both of which resulting in low SOC levels (Hamzehpour et al. 2019).

As seen in this section, the identification and understanding of factors that control SOC, as gained from this study, is crucial

to generate better SOC-related models, especially for grassland ecosystems with similar conditions around the world; improve those already developed, generate effective policies for SOC sequestration that contribute to mitigate climate change; as well as to improve soil health, biodiversity conservation and ecosystem services (Gray et al. 2015). Similar landscapes in South America lack carbon studies and their importance in terms of carbon accumulations is often understated. The application of these study's methodologies in other areas would help improve our understanding of their importance and target protection actions in priority areas.

SOC Stocks Distribution

The Casanare flooded savannas landscape holds important SOC stocks (55.07 Mt of C), which may be affected by land use changes and unsustainable management practices. While the pan-tropical wetlands map of Gumbrecht et al. (2017) predicted the presence of peat deposits in this landscape, we found only mineral soils with high SOC contents. Nevertheless, the study area represents only around 4.43% of the flooded Llanos del Orinoco landscape, which accounts approximately 15 Mha (Hamilton et al. 2002). The numbers from our study would suggest that the total carbon stock across this vast ecosystem is on the order of 10^9 tons of C, with about 20 to 30% stored in flooded savanna ecosystems.

The expansion of agriculture and introduction of new land use practices could result in significant wetland loss, contributing to the already high levels of loss in South America and increasing CO₂ emissions into the atmosphere (Guo and Gifford 2002; Laganière et al. 2010; Powers 2005; Zhou et al. 2019). Considered as the last agricultural frontier in Colombia, recent developments in the Colombian Llanos have seen the expansion of industrial crops (African oil palm, sugar cane, rice, and maize) as well as increases in grazing area, which have replaced gallery forests, wetlands, and native vegetation (Usma and Trujillo 2011). This alters both the hydrological regime and the composition and structure of plant communities across the landscape (Lasso et al. 2010). In particular, the wetlands in this landscape have been transformed into minimally productive rice fields, due to the expansion of agriculture in the region. Recurrent burning, a controversial land preparation technique to enhance pasture productivity, has devastated native palm ecosystems and forests, which have been colonized by grasses. In addition, pollution by fertilizers and the demand for water from crops threaten the natural state of flooded ecosystems, Moriches, and other aquatic communities in the savannas, putting their continued existence at risk (Lasso et al. 2014).

Land use change in the tropics and subtropics is globally the second largest CO₂ source to the atmosphere, after

combustion of fossil fuels (Barrezueta-Unda et al. 2019; Lal 2015; Van der Werf et al. 2009). The use of wetland soils for agriculture, native vegetation conversion, deforestation, and intensification of fires have largely contributed to the GHG emissions (Margono et al. 2014). In Peru, Moriche palm ecosystems are under increasing human pressure (Hergoualc'h et al. 2017; Horn et al. 2012) and we see evidence of this in our study area.

According to Mora-Fernández et al. (2015), the department of Casanare is one of the most transformed and less protected in the country (Romero Duque et al. 2018; Usma and Trujillo 2011). Land use changes, mainly caused by the accelerated expansion of agricultural activities, the increase of wildfires and the mechanical modification of the water regimes, have been the most harmful and stressful human activities for the ecosystems of the floodplain savannas (Mora-Fernández and Peñuela-Recio 2013). These changes cause the loss of its natural vegetation, the variation of the water dynamics, and the alteration of carbon cycle (Ostle et al. 2009). Although the study area represents only 15% of the total area of the department, the intensive pressure of human development could result in the reduction of its SOC stocks and the release of important amounts of GHG into the atmosphere (Ostle et al. 2009; Schreier et al. 1994; Sharma and Sharma 2022). At regional level, the impact of a large-scale land use conversions of the flooded Llanos del Orinoco ecosystem area (15 Mha) could transform this area in a future source of important global emissions if correct decisions are not taken regarding the land management of the region.

Acting on the commitments of the Paris Agreement, several countries have updated their GHG emissions mitigation strategies, aiming to reduce GHG emissions and keeping a rise in global average temperature below 2 °C (Carvajal-Agudelo and Andrade 2020). Given the importance of wetlands for local climate resilience and the significant carbon impacts, the protection and restoration of wetlands should be an integral part of national actions to address the climate change problem and support low emissions, climate resilient development. South America is considered as the wettest continent in the world, counting with 22% of its area covered by wetlands (Junk et al. 2013). Nevertheless, conservation of these important ecosystems is lagging in comparison to other regions due to the slow development of definitions, delineations, and classifications of their wetlands (Pralhad and Kriwoken 2010). The national inventories are not updated, and the ecological importance of these ecosystems is widely undocumented. Thus, the development of accurate SOC and wetland maps plays an essential role, since an adequate inventory and monitoring of these carbon-rich ecosystems will lead to better ecological characterization, quantification of SOC stocks and assessment of GHG emissions (Page et al.

2011), which will contribute to avoiding unwanted impacts on society and maintaining the environmental integrity of South American landscapes. Much attention has been paid recently to swamp forests, but other wetlands in South America are in similar situations. Therefore, it is important to increase the understanding of SOC sequestration and interactions, provide the appropriate assistance to farmers about land management practices and the subsequent creation and application of conservation strategies such as silvopastoral or agroforest systems (Amézquita 1999; Jadán et al. 2015), which have demonstrated a positive balance of SOC storage (Silva-parra 2018), and the definition of global GHG emissions mitigation strategies.

Conclusions

In this study, the status of SOC content and stocks at 0–30 cm soil depth in the Casanare flooded savannas was assessed. Results demonstrated the presence of carbon levels lower than expected, but high relative to other South American flooded savannas. The spatial variability and driving factors of SOC content were evaluated through the implementation of two DSM approaches: Expert Knowledge (EK) and Random Forest (RF). Although both DSM approaches performed very well, EK was considered slightly superior to predict SOC across the study area. Vegetation coverage, topography and soil texture were the most relevant factors in explaining the spatial variability of SOC content, attributing higher SOC contents to dip features, finer soil textures, and high-density vegetation.

Our study showed very high carbon stocks in the study area and a high potential of GHG emissions associated with intensive agricultural development in the region, in addition to loss of large areas of wetlands. SOC stock quantification indicated the importance of maintaining or even improving the carbon sequestration in the area to avoid carbon losses and, consequently, increasing CO₂ emissions to the atmosphere. Inappropriate land and soil management of these ecosystems could hamper the efforts of Colombia and the global community on carbon sequestration and reduction of CO₂ emissions. Additionally, these findings reinforce the view that wetlands, as the Casanare flooded savannas and similar areas elsewhere in South America, are relevant carbon storage ecosystems that must be considered of worldwide interest and protection.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s13157-023-01705-3>.

Acknowledgments We acknowledge funding from the CGIAR Research Program on Forests, Trees, and Agroforestry (CRP-FTA)

with financial support from the CGIAR Fund Donors. We are grateful for the financial support from the Bezos Earth Fund project Using genetic diversity to capture carbon through deep root systems in tropical soils. We also wish to extend our thanks to Cesar Botero who was fundamental during the soil sampling and the farmers who granted us access to their farms.

Author Contributions Javier M. Martín-López: Conceptualization, Field work, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Visualization.

Louis V. Verchot: Conceptualization, Methodology, Investigation, Formal analysis, Writing - Review & Editing, Supervision, Funding acquisition.

Christopher Martius: Writing - Review & Editing, Supervision, Funding acquisition.

Mayesse da Silva: Conceptualization, Methodology, Validation, Investigation, Writing - Review & Editing, Supervision, Project coordination, Funding acquisition.

Funding This work was supported by the CGIAR Research Program on Forests, Trees, and Agroforestry (CRP-FTA) with financial support from the CGIAR Fund Donors and the Bezos Earth Fund project Using genetic diversity to capture carbon through deep root systems in tropical soils.

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Adhikari Kabindra, Alfred E. Hartemink, Minasny Budiman, Bou Kheir Rania, Greve Mette B., Greve Mogens H. (2014) Digital Mapping of Soil Organic Carbon Contents and Stocks in Denmark. *PLoS ONE* 9(8). <https://doi.org/10.1371/journal.pone.0105519>

Aitken R (1977) Wilderness areas in Scotland. Ph. D. thesis, University of Aberdeen

Akumu CE, Johnson JA, Etheridge D, Uhlig P, Woods M, Pitt DG, McMurray S (2015) GIS-fuzzy logic based approach in modeling soil texture: using parts of the Clay Belt and Hornepayne region in Ontario Canada as a case study. *Geoderma* 239–240:13–24. <https://doi.org/10.1016/j.geoderma.2014.09.021>

Alin A (2010) Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2:370–374. <https://doi.org/10.1002/wics.84>

Allam AS, Bassioni HA, Kamel W, Ayoub M (2020) Estimating the standardized regression coefficients of design variables in daylighting and energy performance of buildings in the face of multicollinearity. *Solar Energy* 211:1184–1193. <https://doi.org/10.1016/j.solener.2020.10.043>

Amézquita E (1999) Propiedades físicas de los suelos de los Llanos Orientales y sus requerimientos de labranza. *Palmas* 20:28–30

Ashtekar JM, Owens PR (2013) Remembering knowledge: an expert knowledge based approach to digital soil mapping. *Soil Horizons* 54:0. <https://doi.org/10.2136/sh13-01-0007>

Ashtekar JM, Owens PR, Brown RA, Winzeler HE, Dorantes M, Libohova Z, Dasilva M, Castro A (2014) Digital mapping of soil properties and associated uncertainties in the llanos orientales, south america. *GlobalSoilMap: Basis of the Global Spatial Soil Information System - Proceedings of the 1st GlobalSoilMap Conference*, 367–372. <https://doi.org/10.1201/b16500-67>

Augustin C, Cihacek LJ (2016) Relationships between soil carbon and soil texture in the Northern Great Plains. *Soil Science* 181:386–392. <https://doi.org/10.1097/SS.0000000000000173>

Barbosa RI, Silva dos Santos JR, Souza da Cunha M, Pimentel TP, Fearnside PM (2012) Root biomass, root:shoot ratio and below-ground carbon stocks in the open savannahs of Roraima, Brazilian Amazonia. *Australian Journal of Botany* 60:405. <https://doi.org/10.1071/BT11312>

Barreto JS, Armenteras D (2020) Open data and machine learning to model the occurrence of fire in the ecoregion of “Llanos Colombo–Venezolanos.” *Remote Sensing* 12:3921. <https://doi.org/10.3390/rs12233921>

Barrezueta-Unda S, Velepucha-Cuenca K, Hurtado-Flores L, Jaramillo-Aguilar E (2019) Soil properties and storage of organic carbon in the land use pasture and forest. *Revista de Ciencias Agrícolas* 36:32–45

Barthelmes A, Joosten H (2018) Lineamientos para inventarios de turberas tropicales a fin de facilitar su designación como sitios Ramsar. Nota informativa n°9. Secretaría de La Convención de Ramsar Gland, Sui

Behrens T, Schmidt K, Ramirez-Lopez L, Gallant J, Zhu A-X, Scholten T (2014) Hyper-scale digital soil mapping and soil formation analysis. *Geoderma* 213:578–588. <https://doi.org/10.1016/j.geoderma.2013.07.031>

Blake GR (1965) Bulk Density. *Methods of Soil Analysis C.A. - University of Minnesota St. Paul, Minnesota* 17. <https://doi.org/10.2134/agronmonogr9.1.c30>

Boixadera J (2003) Hydromorphic and clay-related processes in soils from the Llanos de Moxos (Northern Bolivia). *CATENA* 54:403–424. [https://doi.org/10.1016/S0341-8162\(03\)00134-6](https://doi.org/10.1016/S0341-8162(03)00134-6)

Bouyoucos GJ (1936) Directions for making mechanical analyses of soils by the hydrometer method. *Soil Science* 42:225–230. <https://doi.org/10.1097/00010694-193609000-00007>

Breiman L (2001) Random forests. *Machine Learning* 45:5–32. <https://doi.org/10.1023/A:1010933404324>

Buriticá Mejía N (2016) Sabanas inundables de la orinoquía colombiana -documento resumen-. Humboldt: 1–19

Cabrera-Amaya DM, Giraldo-Kalil LJ, Rivera-Díaz O, Castro-Lima F (2020) Riqueza, composición y distribución de las plantas vasculares en sabanas y bosques ribereños de la cuenca baja del río Pauto (Casanare-Colombia). *Revista de La Academia Colombiana de Ciencias Exactas, Físicas y Naturales* 44:1018–1032. <https://doi.org/10.18257/raccefy.1188>

Chmura GL, Anisfeld SC, Cahoon DR, Lynch JC (2003) Global carbon sequestration in tidal, saline wetland soils. *Global Biogeochemical Cycles* 17(4, 1111). <https://doi.org/10.1029/2002GB001917>

- Callens A, Morichon D, Abadie S, Delpy M, Lique B (2020) Using Random forest and Gradient boosting trees to improve wave forecast at a specific location. *Applied Ocean Research* 104. <https://doi.org/10.1016/j.apor.2020.102339>
- Camera C, Zomeni Z, Noller JS, Zissimos AM, Christoforou IC, Bruggeman A (2017) A high resolution map of soil types and physical properties for Cyprus: a digital soil mapping optimization. *Geoderma* 285:35–49. <https://doi.org/10.1016/j.geoderma.2016.09.019>
- Carvajal-Agudelo BN, Andrade HJ (2020) Carbon capture regarding biomass from rural land use systems near the municipality of Yopal, Casanare, Colombia. *Captura de Carbono En Biomasa de Sistemas de Uso Del Suelo, Municipio de Yopal, Casanare, Colombia*. 24:13–22
- Castillo-Figueroa D, Martínez-Medina D, Rodríguez-Posada ME, Bernal-Vergara S (2019) Structural differences in mammal assemblages between savanna ecosystems of the Colombian Llanos. *Papéis Avulsos de Zoologia* 59:11. <https://doi.org/10.11606/1807-0205/2019.59.14>
- Costanza R, de Groot R, Sutton P, van der Ploeg S, Anderson SJ, Kubiszewski I, Farber S, Turner RK (2014) Changes in the global value of ecosystem services. *Global Environmental Change* 26:152–158. <https://doi.org/10.1016/j.gloenvcha.2014.04.002>
- Cotrufo MF, Wallenstein MD, Boot CM, Deneff K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Global Change Biology* 19:988–995. <https://doi.org/10.1111/gcb.12113>
- Craine JM, Brookshire ENJ, Cramer MD, Hasselquist NJ, Koba K, Marin-Spiotta E, Wang L (2015) Ecological interpretations of nitrogen isotope ratios of terrestrial plants and soils. *Plant and Soil* 396:1–26. <https://doi.org/10.1007/s11104-015-2542-1>
- da Silva MA, Silva MLN, Owens PR, Curi N, Oliveira AH, Candido BM (2016) Predicting runoff risks by digital soil mapping. *Revista Brasileira de Ciencia Do Solo* 40:1–13. <https://doi.org/10.1590/18069657rbcs20150353>
- de Menezes MD, Silva SHG, de Mello CR, Owens PR, Curi N (2018) Knowledge-based digital soil mapping for predicting soil properties in two representative watersheds. *Scientia Agricola* 75:144–153. <https://doi.org/10.1590/1678-992x-2016-0097>
- de Menezes MD, Silva SHG, Owens PR, Curi N (2013) Digital soil mapping approach based on fuzzy logic and field expert knowledge. *Ciência e Agrotecnologia* 37:287–298. <https://doi.org/10.1590/S1413-70542013000400001>
- Diwediga B, Bao Q, Agodzo S, Wala K (2017) Potential storages and drivers of soil organic carbon and total nitrogen across river basin landscape : the case of Mo river basin (Togo) in West Africa. *Ecological Engineering* 99:298–309. <https://doi.org/10.1016/j.ecoleng.2016.11.055>
- Douglas DH (1994) Least-cost path in GIS using an accumulated cost surface and slopelines. *Cartographica: The International Journal for Geographic Information and Geovisualization* 31:37–51. <https://doi.org/10.3138/D327-0323-2JUT-016M>
- Elhag M, Bahrawi JA (2017) Soil salinity mapping and hydrological drought indices assessment in arid environments based on remote sensing techniques. *Geoscientific Instrumentation, Methods and Data Systems* 6:149–158. <https://doi.org/10.5194/gi-6-149-2017>
- FAO (1965) Soil survey of the Llanos orientales, New York
- Funes I, Savé R, Rovira P, Molowny-Horas R, Alcañiz JM, Ascaso E, Herms I, Herrero C, Boixadera J, Vayreda J (2019) Agricultural soil organic carbon stocks in the north-eastern Iberian Peninsula: drivers and spatial variability. *Science of the Total Environment* 668:283–294. <https://doi.org/10.1016/j.scitotenv.2019.02.317>
- Goseen D (1964) Geomorfología de los Llanos Orientales. *Revista de La Academia de Las Ciencias Exactas, Físicas y Naturales* 12:129–140
- Gray JM, Bishop TFAA, Wilson BR (2015) Factors controlling soil organic carbon stocks with depth in Eastern Australia. *Soil Science Society of America Journal* 79:1741–1751. <https://doi.org/10.2136/sssaj2015.06.0224>
- Grimm R, Behrens T, Märker M, Elsenbeer H (2008) Soil organic carbon concentrations and stocks on Barro Colorado Island - Digital soil mapping using Random Forests analysis. *Geoderma* 146:102–113. <https://doi.org/10.1016/j.geoderma.2008.05.008>
- Gumbrecht T, Roman-Cuesta RM, Verchot L, Herold M, Wittmann F, Householder E, Herold N, Murdiyarso D (2017) An expert system model for mapping tropical wetlands and peatlands reveals South America as the largest contributor. *Global Change Biology* 23:3581–3599. <https://doi.org/10.1111/gcb.13689>
- Guo LB, Gifford RM (2002) Soil carbon stocks and land use change: a meta analysis. *Global Change Biology* 8:345–360. <https://doi.org/10.1046/j.1354-1013.2002.00486.x>
- Hamilton SK, Sippel SJ, Melack JM (2004) Seasonal inundation patterns in two large savanna floodplains of South America: the Llanos de Moxos(Bolivia) and the Llanos del Orinoco(Venezuela and Colombia). *Hydrological Processes* 18:2103–2116. <https://doi.org/10.1002/hyp.5559>
- Hamilton SK, Sippel SJ, Melack JM (2002) Comparison of inundation patterns among major South American floodplains. *Journal of Geophysical Research Atmospheres* 107:LBA 5-1-LBA 5-14. <https://doi.org/10.1029/2000JD000306>
- Hamzehpour N, Shafizadeh-Moghadam H, Valavi R (2019) Exploring the driving forces and digital mapping of soil organic carbon using remote sensing and soil texture. *Catena* 182. <https://doi.org/10.1016/j.catena.2019.104141>
- Haskins DM, Correll CS, Foster RA, Chatoian JM, Fincher JM, Strenger JM, Keys Jr JE, Maxwell JR, King T (1999) A geomorphic classification system. *Abstracts with Programs - Geological Society of America*, 31:254
- Heitkamp F, Ahrends B, Evers J, Steinicke C, Meessenburg H (2020) Inference of forest soil nutrient regimes by integrating soil chemistry with fuzzy-logic: regionwide application for stakeholders of Hesse, Germany. *Geoderma Regional* 23:e00340. <https://doi.org/10.1016/j.geodrs.2020.e00340>
- Hengl T, Heuvelink GBM, Kempen B, Leenaars JGB, Walsh MG, Shepherd KD, Sila A, MacMillan RA, De Jesus JM, Tamene L, Tondoh JE (2015) Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. *PLoS ONE* 10:1–26. <https://doi.org/10.1371/journal.pone.0125814>
- Hergoualc'h K, Gutiérrez-Vélez VH, Menton M, Verchot LV (2017) Characterizing degradation of palm swamp peatlands from space and on the ground: an exploratory study in the Peruvian Amazon. *Forest Ecology and Management* 393:63–73. <https://doi.org/10.1016/j.foreco.2017.03.016>
- Horn CM, Gilmore MP, Endress BA (2012) Ecological and socioeconomic factors influencing aguaje (*Mauritia flexuosa*) resource management in two indigenous communities in the Peruvian Amazon. *Forest Ecology and Management* 267:93–103. <https://doi.org/10.1016/j.foreco.2011.11.040>
- Hounkpatin OKLL, Op de Hipt F, Bossa AY, Welp G, Amelung W (2018) Soil organic carbon stocks and their determining factors in the Dano catchment (Southwest Burkina Faso). *Catena* 166:298–309. <https://doi.org/10.1016/j.catena.2018.04.013>
- IGAC (2014) Estudio general de suelos zonificación de tierras: Departamento de Casanare
- IPBES (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES secretariat, Bonn, Germany. <https://doi.org/10.5281/zenodo.2616458>

- Jadán O, Cifuentes M, Torres B, Selesi D, Veintimilla D, Günter S (2015) Influence of tree cover on diversity, carbon sequestration and productivity of cocoa systems in the Ecuadorian Amazon. *Bois et Forêts Des Tropiques* 325:35–47
- Jasiewicz J, Stepinski TF (2013) Geomorphons—a pattern recognition approach to classification and mapping of landforms. *Geomorphology* 182:147–156. <https://doi.org/10.1016/j.geomorph.2012.11.005>
- Jenny H (1941) *Factors of Soil Formation: A System of Quantitative Pedology*. McGraw-Hill Book Company Inc
- Junk WJ, An S, Finlayson CM, Gopal B, Květ J, Mitchell SA, Mitsch WJ, Robarts RD (2013) Current state of knowledge regarding the world's wetlands and their future under global climate change: a synthesis. *Aquatic Sciences* 75:151–167. <https://doi.org/10.1007/s00027-012-0278-z>
- Köchy M, Hiederer R, Freibauer A (2015) Global distribution of soil organic carbon – Part 1: masses and frequency distributions of SOC stocks for the tropics, permafrost regions, wetlands, and the world. *SOIL* 1:351–365. <https://doi.org/10.5194/soil-1-351-2015>
- Laganière J, Angers DA, Paré D (2010) Carbon accumulation in agricultural soils after afforestation: a meta-analysis. *Global Change Biology* 16:439–453. <https://doi.org/10.1111/j.1365-2486.2009.01930.x>
- Lal R (2015) Restoring soil quality to mitigate soil degradation. 5875–5895. <https://doi.org/10.3390/su7055875>
- Lamichhane S, Kumar L, Wilson B (2019) Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: a review. *Geoderma* 352:395–413. <https://doi.org/10.1016/j.geoderma.2019.05.031>
- Lasso C, Rial A, Colonnello G, Machado-Allison A, Trujillo F (2014) XI. Humedales de la Orinoquia (Colombia- Venezuela). Serie Editorial Recursos Hidrobiológicos y Pesqueros Continentales de Colombia. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (IAvH). Bogotá
- Lasso C, Usma J, Trujillo F, Rial A (2010) Biodiversidad de la cuenca del Orinoco - Bases científicas para la identificación de áreas prioritarias para la conservación y uso sostenible de la biodiversidad, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, WWF Colombia, Fundación Omacha, Fundación La Salle e Instituto de Estudios de la Orinoquia (Universidad Nacional de Colombia). Bogotá, D. C., Colombia. <https://doi.org/10.1017/CBO9781107415324.004>
- Liaw A, Wiener M (2003) Classification and regression by randomForest. *R News* 3:18–22
- Margono BA, Potapov PV, Turubanova S, Stolle F, Hansen MC (2014) Primary forest cover loss in Indonesia over 2000–2012. *Nature Climate Change* 4:730–735. <https://doi.org/10.1038/nclimate2277>
- McBratney AB, Mendonça Santos ML, Minasny B (2003) On digital soil mapping. *Geoderma*. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Melton JR, Wania R, Hodson EL, Poulter B, Ringeval B, Spahni R, Bohn T, Avis CA, Beerling DJ, Chen G, Eliseev AV, Denisov SN, Hopcroft PO, Lettenmaier DP, Riley WJ, Singarayer JS, Subin ZM, Tian H, Zürcher S, Brovkin V, van Bodegom PM, Kleinen T, Yu ZC, Kaplan JO (2013) Present state of global wetland extent and wetland methane modelling: conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences* 10:753–788. <https://doi.org/10.5194/bg-10-753-2013>
- Minai JO, Libohova Z, Schulze DG (2021) Spatial prediction of soil properties for the Busia area, Kenya using legacy soil data. *Geoderma Regional* 25:e00366. <https://doi.org/10.1016/j.geodrs.2021.e00366>
- Minasny B, McBratney AB (2006) A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences* 32:1378–1388. <https://doi.org/10.1016/j.cageo.2005.12.009>
- Minasny B, McBratney AB, Malone BP, Wheeler I (2013) Digital mapping of soil carbon. *Advances in Agronomy* 118:1–47. <https://doi.org/10.1016/B978-0-12-405942-9.00001-3>
- Minasny B, Rudiyanto, Sulaeman Y, Setiawan BI (2020) Open digital mapping for accurate assessment of tropical peatlands. *Tropical Wetlands Innovation in Mapping and Management- Proceedings of the International Workshop on Tropical Wetlands: Innovation in Mapping and Management 2018:3–8*. <https://doi.org/10.1201/9780429264467-1>
- Mitra S, Wassmann R, Vlek PLG (2005) An appraisal of global wetland area and its organic carbon stock. *Current Science* 88:25–35
- Mitsch WJ, Bernal B, Hernandez ME (2015) Ecosystem services of wetlands. *International Journal of Biodiversity Science, Ecosystem Services & Management* 11:1–4. <https://doi.org/10.1080/21513732.2015.1006250>
- Mitsch WJ, Gosselink JG (2007) *Wetlands*, 4th edn. Wiley, Hoboken
- Mondal A, Khare D, Kundu S, Mondal S, Mukherjee S, Mukhopadhyay A (2017) Spatial soil organic carbon (SOC) prediction by regression kriging using remote sensing data. *Egyptian Journal of Remote Sensing and Space Science* 20:61–70. <https://doi.org/10.1016/j.ejrs.2016.06.004>
- Moore ID, Gessler PE, Nielsen GA, Peterson GA (1993) Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal* 57:443–452. <https://doi.org/10.2136/sssaj1993.03615995005700020026x>
- Mora-Fernández C, Peñuela-Recio L (2013) Salud Ecosistémica de las sabanas inundables asociadas a la cuenca del río Pauto, Casanare, Colombia. Yoluka ONG, fundación de investigación en biodiversidad y conservación, Fundación Horizonte Verde y Ecopetrol S.A, Bogotá, Colombia
- Mora-Fernández C, Peñuela-Recio L, Castro-Lima F (2015) Estado del conocimiento de los ecosistemas de las sabanas inundables en la Orinoquia Colombiana. *Orinoquia* 19:253–271
- Motohka T, Nasahara KN, Oguma H, Tsuchida S (2010) Applicability of Green-Red Vegetation Index for remote sensing of vegetation phenology. *Remote Sensing* 2:2369–2387. <https://doi.org/10.3390/rs2102369>
- Nahlik AM, Fennessy MS (2016) Carbon storage in US wetlands. *Nature Communications* 7:1–9. <https://doi.org/10.1038/ncomms13835>
- Ngunjiri MW, Libohova Z, Minai JO, Serrem C, Owens PR, Schulze DG (2019) Predicting soil types and soil properties with limited data in the Uasin Gishu Plateau, Kenya. *Geoderma Regional* 16:e00210. <https://doi.org/10.1016/j.geodrs.2019.e00210>
- Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience* 51:933–938. [https://doi.org/10.1641/0006-3568\(2001\)051\[0933:TEOTWA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2)
- Ostle NJ, Levy PE, Evans CD, Smith P (2009) UK land use and soil carbon sequestration. *Land Use Policy* 26:274–283. <https://doi.org/10.1016/j.landusepol.2009.08.006>
- Page SE, Rieley JO, Banks CJ (2011) Global and regional importance of the tropical peatland carbon pool. *Global Change Biology* 17:798–818. <https://doi.org/10.1111/j.1365-2486.2010.02279.x>
- Patton NR, Lohse KA, Seyfried MS, Godsey SE, Parsons SB (2019) Topographic controls of soil organic carbon on soil-mantled landscapes. *Scientific Reports* 9:1–15. <https://doi.org/10.1038/s41598-019-42556-5>
- Paul EA (2016) The nature and dynamics of soil organic matter: Plant inputs, microbial transformations, and organic matter stabilization. *Soil Biology and Biochemistry* 98:109–126. <https://doi.org/10.1016/j.soilbio.2016.04.001>

- Peñuela L, Solano C, Ardila V, Galán S (2014) Sabana inundable y ganadería, opción productiva de Conservación en la Orinoquía. Proyecto: "Fortalecimiento institucional y de política para incrementar la conservación de la biodiversidad en predios privados en Colombia." Sabana inundable y ganadería, opción productiva de conservación en la Orinoquía. Proyecto: "Fortalecimiento institucional y de política para incrementar la conservación de la biodiversidad en predios privados en Colombia." Bogotá, Colombia
- Powers JS (2005) Regional variation in soil carbon and ^{13}C in forests and pastures of northeastern Costa Rica Regional variation in soil carbon and $\delta^{13}\text{C}$ in forests and pastures of northeastern Costa Rica. <https://doi.org/10.1007/s10533-004-0368-7>
- Prahalad VN, Kriwoken LK (2010) Implementation of the Ramsar convention on wetlands in Tasmania, Australia. *Journal of International Wildlife Law and Policy* 13:205–239. <https://doi.org/10.1080/13880292.2010.486697>
- Rice AL, Butenhoff CL, Teama DG, Röger FH, Khalil MAK, Rasmussen RA (2016) Atmospheric methane isotopic record favors fossil sources flat in 1980s and 1990s with recent increase. *Proceedings of the National Academy of Sciences* 113:10791–10796. <https://doi.org/10.1073/pnas.1522923113>
- Romero Duque LP, Castro Lima F, Rentería Mosquera Á (2018) Contribución al conocimiento de la vegetación de las sabanas de Casanare (Colombia). *Revista U.D.C.A. Actualidad & Divulgación Científica* 21:197–205. <https://doi.org/10.31910/rudca.v21.n1.2018.678>
- Roudier P, Hewitt AE, Beaudette DE (2012) A conditioned Latin hypercube sampling algorithm incorporating operational constraints. *Digital Soil Assessments and Beyond - Proceedings of the Fifth Global Workshop on Digital Soil Mapping*, 227–231. <https://doi.org/10.1201/b12728-46>
- Sarmiento G (1984) *The Ecology of Neotropical Savannas*. Harvard University Press, Cambridge, MA 02138. ISBN 0-674- 22460-4. Harvard University Press. <https://doi.org/10.4159/harvard.9780674418554>
- Sarmiento G (1983) The savannas of tropical America. In: F. Bourliere, F. (Ed): *Ecosystems of the World XIII. Tropical Savannas*, Elsevier, Amsterdam, pp. 245–288., in: *Ecosystems of the World XIII. Tropical Savannas*. pp. 245–288
- Saunois M, Stavert AR, Poulter B, Bousquet P, Canadell JG, Jackson RB, Raymond PA, Dlugokencky EJ, Houweling S, Patra PK, Ciais P, Arora VK, Bastviken D, Bergamaschi P, Blake DR, Brailsford G, Bruhwiler L, Carlson KM, Carrol M, Castaldi S, Chandra N, Crevoisier C, Crill PM, Covey K, Curry CL, Etiope G, Frankenberg C, Gedney N, Hegglin MI, Höglund-Isaksson L, Hugelius G, Ishizawa M, Ito A, Janssens-Maenhout G, Jensen KM, Joos F, Kleinen T, Krummel PB, Langenfelds RL, Laruelle GG, Liu L, Machida T, Maksyutov S, McDonald KC, McNorton J, Müller PA, Melton JR, Morino I, Müller J, Murguía-Flores F, Naik V, Niwa Y, Noce S, O'Doherty S, Parker RJ, Peng C, Peng S, Peters GP, Prigent C, Prinn R, Ramonet M, Regnier P, Riley WJ, Rosentretter JA, Segers A, Simpson IJ, Shi H, Smith SJ, Steele LP, Thornton BF, Tian H, Tohjima Y, Tubiello FN, Tsuruta A, Viovy N, Voulgarakis A, Weber TS, van Weele M, van der Werf GR, Weiss RF, Worthy D, Wunch D, Yin Y, Yoshida Y, Zhang W, Zhang Z, Zhao Y, Zheng B, Zhu Qing, Zhu Qian, Zhuang Q (2020) The global methane budget 2000–2017. *Earth System Science Data* 12:1561–1623. <https://doi.org/10.5194/essd-12-1561-2020>
- Schreier H, Shah PB, Lavkulich LM, Brown S (1994) Maintaining soil fertility under increasing land use pressure in the Middle Mountains of Nepal. *Soil Use and Management* 10:137–142. <https://doi.org/10.1111/j.1475-2743.1994.tb00474.x>
- Sharma G, Sharma LK (2022) Assessment of soil carbon stock and important physicochemical properties in relation to land use patterns in semi-arid region of Rajasthan, India. *Journal of the Indian Society of Soil Science* 70:191–203. <https://doi.org/10.5958/0974-0228.2022.00019.6>
- Shi X, Zhu A-X, Burt JE, Qi F, Simonson D (2004) A case-based reasoning approach to fuzzy soil mapping. *Soil Science Society of America Journal* 68:885–894. <https://doi.org/10.2136/sssaj2004.8850>
- Shiel D, Ladd B, Silva LCR, Laffan SW, van Heist M (2016) How are soil carbon and tropical biodiversity related? *Environmental Conservation* 43:231–241. <https://doi.org/10.1017/S0376892916000011>
- Shimada M, Itoh T, Motooka T, Watanabe M, Shiraiishi T, Thapa R, Lucas R (2014) New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sensing of Environment* 155:13–31. <https://doi.org/10.1016/j.rse.2014.04.014>
- Siewert M (2017) High-resolution digital mapping of soil organic carbon in permafrost terrain using machine-learning: a case study in a sub-arctic peatland environment. *Biogeosciences* 15:1663–1682. <https://doi.org/10.5194/bg-15-1663-2018>
- Silva-parra A (2018) Modeling soil carbon stocks and carbon dioxide emissions (GHG) in production systems of Plain Altillanura. *Orinoquía* 22:1–25. <https://doi.org/10.22579/20112629.525>
- Sindayihebura A, Ottoy S, Dondeyne S, Van Meirvenne M, Van Orshoven J (2017) Comparing digital soil mapping techniques for organic carbon and clay content: Case study in Burundi's central plateaus. *Catena* 156:161–175. <https://doi.org/10.1016/j.catena.2017.04.003>
- Six J, Conant RT, Paul EA, Paustian K (2002) Stabilization mechanisms of soil organic matter: implications for C-saturation of soils. *Plant and Soil* 241:155–176
- Sjögersten S, Black CR, Evers S, Hoyos-Santillan J, Wright EL, Turner BL (2014) Tropical wetlands: a missing link in the global carbon cycle? *Global Biogeochemical Cycles* 28:1371–1386. <https://doi.org/10.1002/2014GB004844>
- Soil Science Division Staff (2017) *Soil Survey Manual*. United States Department of Agriculture Handbook No. 18 18, 1913–5. <https://doi.org/10.1121/1.1903534>
- Sørensen R, Zinko U, Seibert J (2006) On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences* 10:101–112. <https://doi.org/10.5194/hess-10-101-2006>
- Szabó B, Szatmári G, Takács K, Laborczi A, Makó A, Rajkai K, Pásztor L (2019) Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics. *Hydrology and Earth System Sciences* 23:2615–2635. <https://doi.org/10.5194/hess-23-2615-2019>
- Taghizadeh-Mehrjardi R, Neupane R, Sood K, Kumar S (2017) Artificial bee colony feature selection algorithm combined with machine learning algorithms to predict vertical and lateral distribution of soil organic matter in South Dakota, USA. *Carbon Management* 8:277–291. <https://doi.org/10.1080/17583004.2017.1330593>
- Terra DF (2017) Accuracy and uncertainty of digital soil mapping approaches to extract and transfer soil information from reference area. *Dissertação (Mestrado Em Ciência Do Solo)* Universidade Federal de Lavras 114
- Tsakiridis NL, Theocharis JB, Panagos P, Zalidis GC (2019) An evolutionary fuzzy rule-based system applied to the prediction of soil organic carbon from soil spectral libraries. *Applied Soft Computing Journal* 81:105504. <https://doi.org/10.1016/j.asoc.2019.105504>
- Usma JS, Trujillo F (2011) Biodiversidad del departamento Casanare. Identificación de ecosistemas estratégicos. WWF - Colombia, Bogotá, Colombia

- Valbuena R, Hernando A, Manzanera JA, Görgens EB, Almeida DRAA, Mauro F, García-Abril A, Coomes DA (2017) Enhancing of accuracy assessment for forest above-ground biomass estimates obtained from remote sensing via hypothesis testing and overfitting evaluation. *Ecological Modelling* 366:15–26. <https://doi.org/10.1016/j.ecolmodel.2017.10.009>
- Van der Werf GR, Morton DC, Defries RS, Olivier JGJ, Kasibhatla PS, Jackson RB, Collatz GJ, Randerson JT (2009) CO₂ emissions from forest loss. *Nature Geoscience* 2:737–738. <https://doi.org/10.1038/ngeo0671>
- Vega LF, Nunes da Cunha C, Rothaupt K-O, Moreira MZ, Wantzen KM (2014) Does flood pulsing act as a switch to store or release sediment-bound carbon in seasonal floodplain lakes? Case study from the Colombian Orinoco-Llanos and the Brazilian Pantanal. *Wetlands* 34:177–187. <https://doi.org/10.1007/s13157-013-0495-9>
- Vepraskas MJ, Craft CB (2016) *Wetland Soils: Genesis, Hydrology, Landscapes, and Classification*, 2nd Editio. ed, CRC Press. CRC Press, Boca Raton. doi:<https://doi.org/10.1201/b18996>
- Vu DH, Muttaqi KM, Agalgaonkar AP (2015) A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. *Applied Energy* 140:385–394. <https://doi.org/10.1016/j.apenergy.2014.12.011>
- Wan Q, Zhu G, Guo H, Zhang Y, Pan H, Yong L, Ma H (2019) Influence of vegetation coverage and climate environment on soil organic carbon in the Qilian Mountains. *Scientific Reports* 9:1–9. <https://doi.org/10.1038/s41598-019-53837-4>
- Wang X, Zhang Y, Atkinson P, Yao H (2020) Predicting soil organic carbon content in Spain by combining Landsat TM and ALOS PALSAR images. *International Journal of Applied Earth Observation and Geoinformation* 92:102182. <https://doi.org/10.1016/j.jag.2020.102182>
- Wantzen KM, Couto EG, Mund EE, Amorim RSS, Siqueira A, Tielbörger K, Seifan M (2012) Soil carbon stocks in stream-valley-ecosystems in the Brazilian Cerrado agroscape. *Agriculture, Ecosystems and Environment* 151:70–79. <https://doi.org/10.1016/j.agee.2012.01.030>
- Weil R, Brady N (2017) *The Nature and Properties of Soils*, 15th edn. Pearson, United States of America, Columbus
- Welikhe P, Quansah JE, Fall S, Elhenney WMc (2017) Estimation of soil moisture percentage using LANDSAT-based moisture stress index. *Journal of Remote Sensing & GIS* 6(2). <https://doi.org/10.4172/2469-4134.1000200>
- Whiting GJ, Chanton JP (2001) Greenhouse carbon balance of wetlands: methane emission versus carbon sequestration. *Tellus B: Chemical and Physical Meteorology* 53:521–528. <https://doi.org/10.3402/tellusb.v53i5.16628>
- Wiesmeier M, Prietzel J, Barthold F, Spörlein P, Geuß U, Hangen E, Reischl A, Schilling B, von Lütow M, Kögel-Knabner I (2013) Storage and drivers of organic carbon in forest soils of southeast Germany (Bavaria) - implications for carbon sequestration. *Forest Ecology and Management* 295:162–172. <https://doi.org/10.1016/j.foreco.2013.01.025>
- Wiesmeier M, Urbanski L, Hobbey E, Lang B, von Lütow M, Marin-Spiotta E, van Wesemael B, Rabot E, Ließ M, Garcia-Franco N, Wollschläger U, Vogel HJ, Kögel-Knabner I (2019) Soil organic carbon storage as a key function of soils - a review of drivers and indicators at various scales. *Geoderma* 333:149–162. <https://doi.org/10.1016/j.geoderma.2018.07.026>
- Xie E, Zhang Y, Huang B, Zhao Y, Shi X, Hu W, Qu M (2021) Spatiotemporal variations in soil organic carbon and their drivers in southeastern China during 1981–2011. *Soil and Tillage Research* 205:104763. <https://doi.org/10.1016/j.still.2020.104763>
- Xu J, Morris PJ, Liu J, Holden J (2018) PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis. *Catena* 160:134–140. <https://doi.org/10.1016/j.catena.2017.09.010>
- Yang R, Liu F, Zhang G, Zhao Y, Li D, Yang J, Yang Fei, Yang Fan (2016) Mapping soil texture based on field soil moisture observations at a high temporal resolution in an oasis agricultural area. *Pedosphere* 26:699–708. [https://doi.org/10.1016/S1002-0160\(15\)60078-9](https://doi.org/10.1016/S1002-0160(15)60078-9)
- Yang Y, Mohammad A, Feng J, Zhou R, Fang J (2007) Storage, patterns and environmental controls of soil organic carbon in China. *Biogeochemistry* 84:131–141. <https://doi.org/10.1007/s10533-007-9109-z>
- Yun J, Chen X, Liu S, Zhang W (2019) Effects of temperature and moisture on soil organic carbon mineralization. *IOP Conference Series: Materials Science and Engineering* 562. <https://doi.org/10.1088/1757-899X/562/1/012085>
- Zeraatpisheh M, Ayoubi S, Jafari A, Tajik S, Finke P (2019) Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. *Geoderma* 338:445–452. <https://doi.org/10.1016/j.geoderma.2018.09.006>
- Zhou Y, Hartemink AE, Shi Z, Liang Z, Lu Y (2019) Land use and climate change effects on soil organic carbon in North and Northeast China. *Science of the Total Environment* 647:1230–1238. <https://doi.org/10.1016/j.scitotenv.2018.08.016>
- Zhu A-X (1997) A similarity model for representing soil spatial information. *Geoderma* 77:217–242
- Zhu A-X, Hudson B, Burt J, Lubich K, Simonson D (2001) Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal* 65:1463–1472. <https://doi.org/10.2136/sssaj2001.6551463x>
- Zhu A-X, Qi F, Moore A, Burt JE (2010) Prediction of soil properties using fuzzy membership values. *Geoderma* 158:199–206. <https://doi.org/10.1016/j.geoderma.2010.05.001>
- Zhu A-X (1999) A personal construct-based knowledge acquisition process for natural resource mapping. *International Journal of Geographical Information Science* 13(2):119–141. <https://doi.org/10.1080/136588199241382>
- Zhu Q, Peng C, Ciais P, Jiang H, Liu J, Bousquet P, Li S, Chang J, Fang X, Zhou X, Chen H, Liu S, Lin G, Gong P, Wang M, Wang H, Xiang W, Chen J (2017) Interannual variation in methane emissions from tropical wetlands triggered by repeated El Niño Southern Oscillation. *Global Change Biology* 23:4706–4716. <https://doi.org/10.1111/gcb.13726>
- Zinn YL, Lal R, Resck DVSS (2005) Texture and organic carbon relations described by a profile pedotransfer function for Brazilian Cerrado soils. *Geoderma* 127:168–173. <https://doi.org/10.1016/j.geoderma.2005.02.010>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.