Natural Resources Conservation and Research 2024, 7(1), 4913.
https://doi.org/10.24294/nrcr.v7i1.4913

Article

Modeling soil organic carbon based on field soil-texture measurements

Talita Nogueira Terra[1,*], Rozely Ferreira dos Santos[2], Rubens Augusto Camargo Lamparelli[3]

[1] Faculty of Agricultural Engineering, University of Campinas, Campinas-SP 13083-896, Brazil
[2] Biosciences Institute, University of São Paulo, São Paulo 05508-090, Brazil
[3] Interdisciplinary Center on Energy Planning, University of Campinas, Campinas-SP 13083-896, Brazil
*** Corresponding author:** Talita Nogueira Terra, talitanogueiraterra@gmail.com

CITATION

Terra TN, dos Santos RF, Lamparelli RAC. Modeling soil organic carbon based on field soil-texture measurements. Natural Resources Conservation and Research. 2024; 7(1): 4913. https://doi.org/10.24294/nrcr.v7i1.4913

ARTICLE INFO

Received: 29 February 2024
Accepted: 29 March 2024
Available online: 11 May 2024

Abstract: There are many studies about soil organic carbon (SOC) around the world but, in extensive territories, it is more difficult to obtain data due to the number of variables involved in the models and their high cost. In large regions with poor infrastructure, low-cost SOC models are needed. With this in mind, our objective was to estimate the SOC using a simple model based on soil textural data. The work was focused on savanna soil and validated the model in the Brazilian Savanna. Two models were constructed, one for topsoil (0–0.3 m) and other for subsoil (0.3–1.0 m). The SOC models can be carried out in a textural triangle together with SOC values. The results showed that subsoil models were more accurate than topsoil models, but both had good performance. The models give support to SOC-related preliminary research in gross and fast estimates, requiring only reduced financial contribution to calculate SOC in a large region of interest.

Keywords: soil organic carbon; soil texture; model; savanna

1. Introduction

Many authors have been estimating the soil organic carbon (SOC) content for years because the data allow interpretation of soil conservation conditions in both natural [1] and anthropic areas [2]. One good advantage is that results help to understand the damage caused by human actions in nature [3].

SOC modelling often employs machine learning techniques [4] based on several soil parameters and properties (textural data, N, pH, Ca^{2+}, Mg^{2+}, Al, Fe), and their covariates related to climate, organisms, topography, parent material, time and site [5,6]. However, the availability of these extensive datasets can often be limited or lacking in certain study sites, particularly in regions with limited financial resources and skilled labor for database construction [7], as commonly observed in tropical savannas of countries like Brazil. As highlighted by Mutuku et al. [8], smallholder farmers usually cannot afford soil laboratory tests, and methods for visual soil evaluations need to be developed.

Garsia et al. [9] evaluated the effectiveness of 221 soil organic carbon models and found that most of them are not validated (71%), only just four included Brazil as a study area. Despite being one of the world's biodiversity hotspots [10], the Brazilian Savanna has experienced significant deforestation, losing over 50% of its native vegetation cover due to agricultural expansion [11]. It is expected that changes in land use-cover affect the SOC in the whole national territory, as we found in several other regions of the world [3,12–15]. However, as in other countries, it is difficult to study SOC in the whole Brazilian territory due to its large extension and the soil data are highly discontinuous in the territory [16].

In this context, it is usual to find studies that rely on a limited number of samples, highlighting the challenge of scarce soil data and without methodological validation. Zinn et al. [17], for example, examined the relationship between soil texture and SOC in Brazilian Savanna soils, analyzing 17 soil samples across seven different depths. Ruggiero et al. [18] investigated the correlation between vegetation and soil properties in the Brazilian Savanna by randomly sampling 10 quadrants of 10 × 10 m. They discovered that soil properties were more closely associated with variations in vegetation physiognomy in the superficial soil layers than at deeper depths. Neufeldt et al. [19] also studied Brazilian Savanna Oxisols, focusing on the effects of texture and land use on soil organic matter (SOM). Their study, based on five undisturbed topsoil samples, concluded that SOM content was correlated with the clay fraction. These studies, while informative, based on a small number of samples, are often related to low financial resources and skilled labor to construct strong systematic datasets [7,20].

We know that it is vital to understand SOC behavior as well as the consequences of agricultural activities. Thus, it is crucial to create ways to allow SOC estimates—with efficiency, effectiveness, and low cost—in large territories and from information easily obtained. Bearing this in mind, we aim to elaborate a model to obtain an exploratory estimate of the SOC based on soil textural data in Brazilian Savannas—the simpler parameter to be evaluated in the soil. This approach should prove valuable for areas where intensive soil surveys and analytical efforts are not feasible.

2. Materials and methods

We developed SOC models based on three steps: dataset organization; mathematical models' construction; and validation under different land use-cover (**Figure 1**).

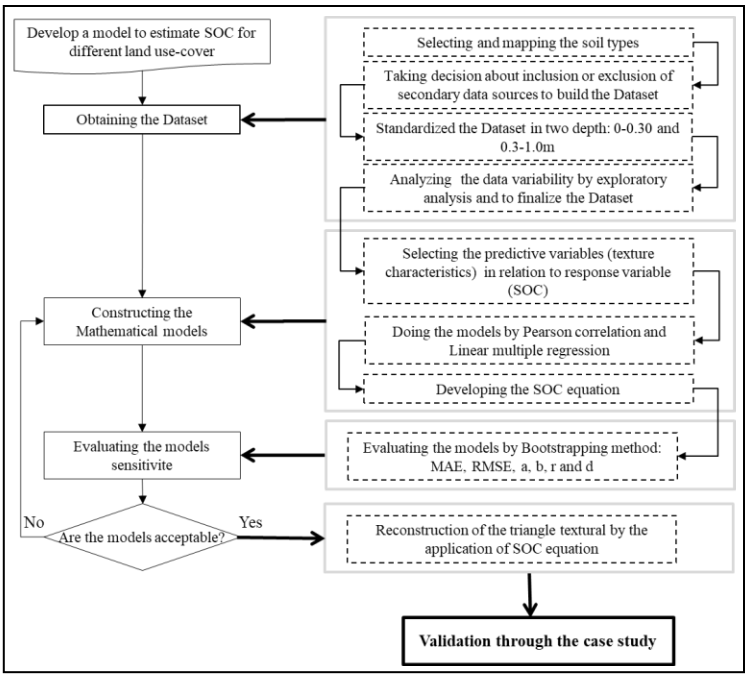

Figure 1. Framework of procedure used.

Data source and analysis methods

The study area was the Brazilian Savanna (Cerrado Biome), covered with natural vegetation or pasture and sugarcane, totaling more than 2 million km^2, equivalent to 22% of the national territory [21]. Due to its large extension, three climate zones are found: (a) the largest one is tropical climate with 88% of the area, (b) 11% of the area is humid subtropical climate, and less than 1% is dry climate [22]. The average annual temperature varies from 20 ℃ to 26 ℃, and the average annual precipitation from 1000 to 1800 mm. Soil fertility is the key factor in determining the Savanna biome vegetation [23]. To represent this diversity, we collected data from the predominant soil type, Oxisols, occupying 41% of the biome [24] in the three climatic zones (**Figure 2**), with natural vegetation and farming production. The remnants of Savanna are established on ancient soils, which are acidic, depleted of nutrients, and rich in aluminum [25,26]. The vegetation upon these soils displays a mosaic of the structures from savanna-like formations to forests.

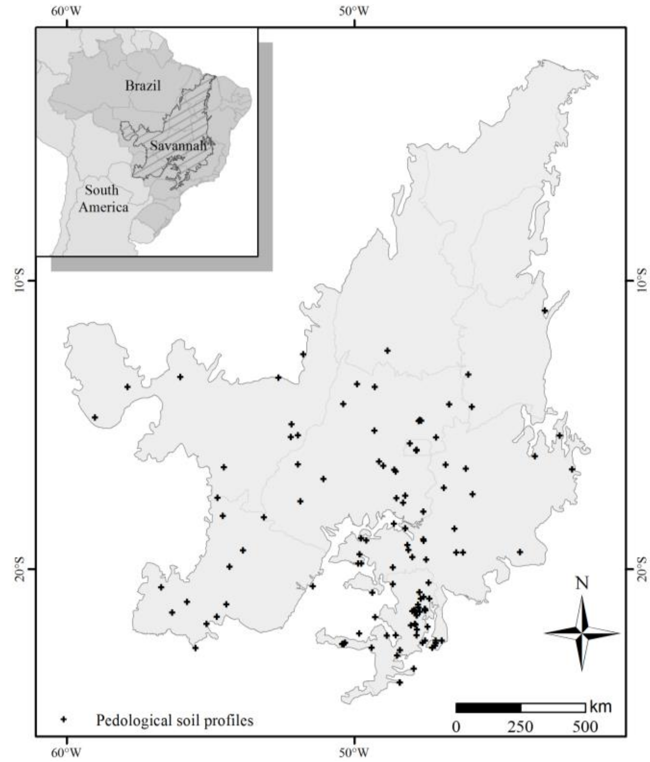

Figure 2. Cerrado soil data distribution obtained from literature review on pedological profiles and ground surveys.

The 10,796 data points were obtained from secondary sources of pedological profiles and ground surveys. They were stored in an information system operated as an Open Access library of the georeferenced data PANGAEA [27]. The dataset considered all pedological profiles to have at least four depths sampled up to a minimum of one meter and located in natural vegetation, pasture, or sugarcane according to Landsat satellite images from soil analysis dates. The available dataset is limited to the savanna center-south region.

The textural data were submitted to a quadratic spline function of the same area

(λ = 0.1) [28,29] using a MATLAB routine [30] to obtain the average values of sand, clay, and silt for two depths, 0–0.3 m (topsoil layer) and 0.3–1.0 m (subsoil layer) as it was done for the SGDBE (Soil Geographical Database of Europe). We checked the dataset variability by mean, standard deviation, and coefficient of variation in Minitab program excluding the outliers. The Pearson correlation method was applied to identify the textural factors (predictable variable) that could be correlated with SOC (response variable), with $0.01 < p\text{-value} > 0.5$. The correlated data in the two depths were included in the linear multiple regression and stepwise (forward and backward) to adjust the SOC estimation models in the statistical program Minitab [17,31,32].

The acceptability of the models was evaluated by mean absolute error (MAE), root mean square error (RMSE), concordance index (d) and linear coefficient (a), angular coefficient (b) and correlation coefficient (r) applying the bootstrap method [33] coded in MATLAB [34] for 10.000 iterations. The accepted models were validated by 350 soil points (139 points for 0–0.3 m and 211 points for 0.3–1.0 m) located in the Brazilian Savanna with different depths by MAE, RMSE and Adjusted R-squared (R^2).

To explore the potential of elaborating SOC maps from the models proposed in this work, we elaborated a SOC map for São Paulo State—Brazil. We obtained the soil textures from the Pedological Map of São Paulo State [35] and estimated the SOC values from a textural triangle. For this area we have 14 SOC values for topsoil and 40 for subsoil obtained in field points and with lab analysis. Then, we calculated the percentual concordance between these values and the estimated values by the triangle to the same points as a way of evaluating the effectiveness.

3. Results

3.1. Application of the mathematical models in study case

The 10,796 Oxisols soil data points were submitted to a quadratic Spline function obtaining 164 pedological profiles for 0–0.3 m depth and 164 profiles for 0.3–1.0 m, composed of 656 data points of sand, clay, silt, and SOC. These soils displayed textural data values with large-scale variability, from 92 to 515 $g \cdot kg^{-1}$, and high standard deviations (**Table 1**). The dataset covers all the textural variation characteristics of Oxisols [4]. The layer 0–0.3 m had higher SOC values ($1.2\ g \cdot kg^{-1} < SOC < 47.1\ g \cdot kg^{-1}$) than the layer 0.3–1.0 m ($1.0\ g \cdot kg^{-1} < SOC < 14.7\ g \cdot kg^{-1}$). The SOC data present normality for both layers by the Anderson-Darling Test (Appendix **Figure A1**).

Table 1. Statistical variation for four soil attributes in two depths.

Depth (m)		Sand ($g \cdot kg^{-1}$)	Silt ($g \cdot kg^{-1}$)	Clay ($g \cdot kg^{-1}$)	SOC ($g \cdot kg^{-1}$)
0–0.3	(a)	514.3 ± 268.2	101.7 ± 70.0	383.9 ± 223.3	13.3 ± 7.8
	(b)	52.2	68.8	58.2	58.6
0.3–1.0	(a)	473.2±265.2	92.2 ± 64.2	434.8 ± 230.4	6.7 ± 3.0
	(b)	56.1	69.6	53.0	45.5

(a) Mean standard deviation; (b) coefficient of variation (%).

The clay and sand data were correlated with SOC by Pearson correlation (**Table 2**), presenting a direct positive correlation between clay and carbon, and an inversed negative correlation with sand.

Table 2. Pearson correlation for SOC response variable.

Depth (m)	Response variable	Sand	Silt	Clay + Silt	Clay
0–0.3	SOC	−0.63	0.38	0.63	0.64
0.3–1.0		−0.72	0.35	0.72	0.74

The best models obtained by multiple linear regression were two for each depth (**Table 3**).

Table 3. Predictive models of SOC ($g \cdot kg^{-1}$) occurring at two depths in soils.

Depth (m)	Models	R^2
0–0.3	SOC = 580 − 0.575 × Sand − 0.557 × (Clay + Silt)	0.44
0–0.3	SOC = 4.771 + 0.02235 × Clay	0.41
0.3–1.0	SOC = 2.452 + 0.009743 × Clay	0.54
0.3–1.0	SOC = 3 − 0.00057 × Sand + 0.0091 × Clay	0.54

3.2. Evaluation of the models sensitive

All the models presented a significant regression ((*b*) ≠ 0) and were more adjusted in the subsoil layer ((*a*) closer to zero). The statistical parameters ((MAE); (RMSE)) of this study also indicated the topsoil layer with low performance, which the SOC was overestimated in 5 $g \cdot kg^{-1}$. The concordance index (*d*) agreed with the MAE/RMSE, ranging from 0.33 to 0.45 (**Figure 3**; Appendix **Figure A2**). For the subsoil layer (0.3–1.0 m) the SOC was overestimated in approximately 1.8 $g \cdot kg^{-1}$ and (*d*) was near to 0.52. The correlation coefficient (*r*) indicated the subsoil layer (0.3–1.0 m) as a better combination between the real SOC and the estimated SOC than the topsoil layer models.

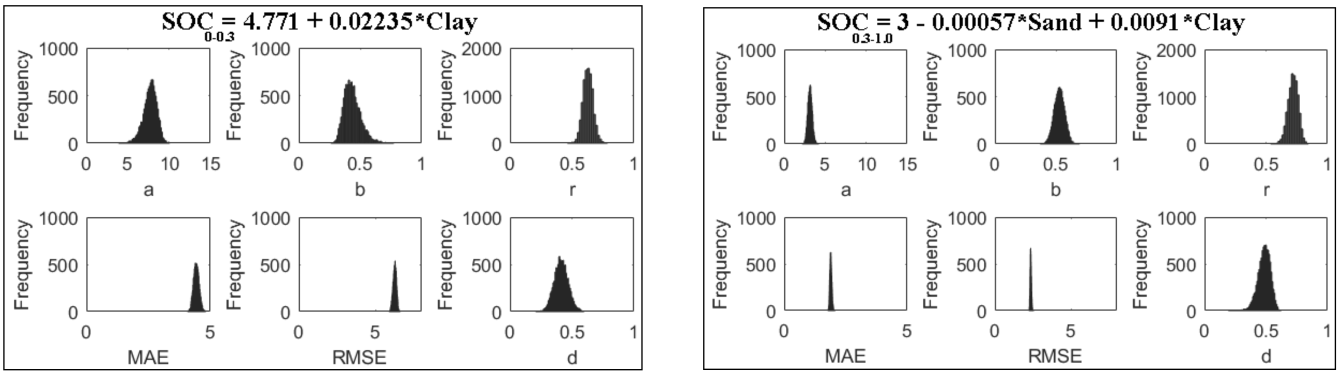

Figure 3. The values estimated by the bootstrap for the best models in the two depths were representing by the histograms above, where (*a*) linear coefficient, (*b*) angular coefficient, (*r*) correlation coefficient, (MAE), (RMSE) and (*d*) statistical parameters.

These equations resulting from the SOC models provided the basis to estimate SOC based on a textural triangle [36] and making it possible to create a new triangle by placing the estimated values of organic carbon on the points of junction of textural values. The best equations are represented in the **Figure 4** and the others can be seen in the Appendix **Figure A3**.

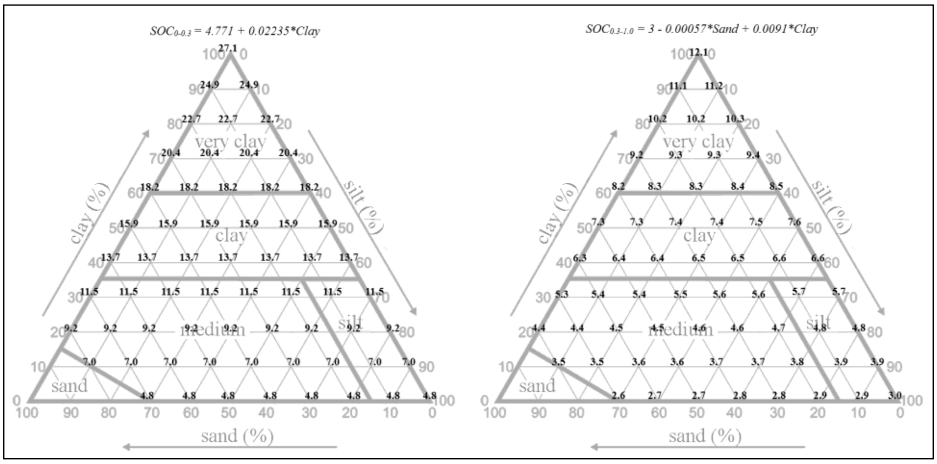

Figure 4. Estimated SOC (g·kg^{-1}) in Savanna Oxisols based in the textural triangle.

3.3. Models validation

The clay values of 139 data points were inserted in the textural triangle which represents the topsoil (**Figure 4**) to estimate the SOC. Then, these values were compared with the measured SOC presented in their respective studies (**Figure 5**). The same approach was utilized for the subsoil (0.3–1.0 m) but with clay and sand data from 211 data points.

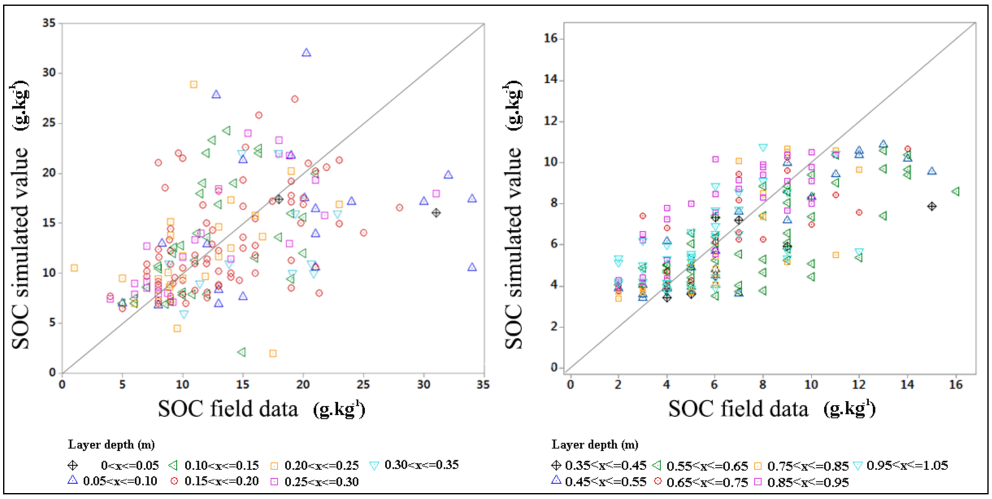

Figure 5. Field measuring versus simulated value of the soil SOC by the textural triangle.

The statistical parameters ((MAE); (RMSE); (R^2)) of the study case indicated the topsoil layer with lower performance than the subsoil (**Table 4**).

Table 4. Statistical tests applied to check the agreement between simulated and observed values, where n is the number of data used in each depth.

Statistical parameters	Depth (m)						
	0 < x ≤ 0.05	**0.05 < x ≤ 0.1**	**0.1 < x ≤ 0.15**	**0.15 < x ≤ 0.2**	**0.2 < x ≤ 0.25**	**0.25 < x ≤ 0.3**	**0.3 < x ≤ 0.35**
RMSE	10.6	9.7	3.9	4.7	3.7	5.4	6.5
MAE	8	7.8	3.1	3.5	2.6	4	5.6
R^2	-	0.2	0.5	0.3	0.5	0.5	0.1
n	2	16	18	57	18	17	11
	0.35 < x ≤ 0.45	**0.45 < x ≤ 0.55**	**0.55 < x ≤ 0.65**	**0.65 < x ≤ 0.75**	**0.75 < x ≤ 0.85**	**0.85 < x ≤ 0.95**	**0.95 < x ≤ 1.05**
RMSE	3	2	2.7	1.9	1.9	2	2.1
MAE	2	1.6	2.1	1.5	1.5	1.7	1.6
R^2	0.4	0.8	0.5	0.6	0.6	0.7	0.3
n	7	24	56	28	23	33	40

We also applied the estimated SOC values in Pedological Map of São Paulo State [35], with the aim to assert the effectiveness of the models. We investigated part of the Brazilian Savanna (Sao Paulo State Savanna), which in turn had less environmental variability in terms of climate, relief, and land use-cover, then with expectations of good results for topsoil [13]. The soil organic carbon mapping represented in **Figure 6** is an expression of the indirect measure obtained by the textural triangle. Our results presented an accuracy of 86% for topsoil and 70% for subsoil. However, if the analyzes were made between 0.6 m and 1 m, the accuracy goes up to 80%.

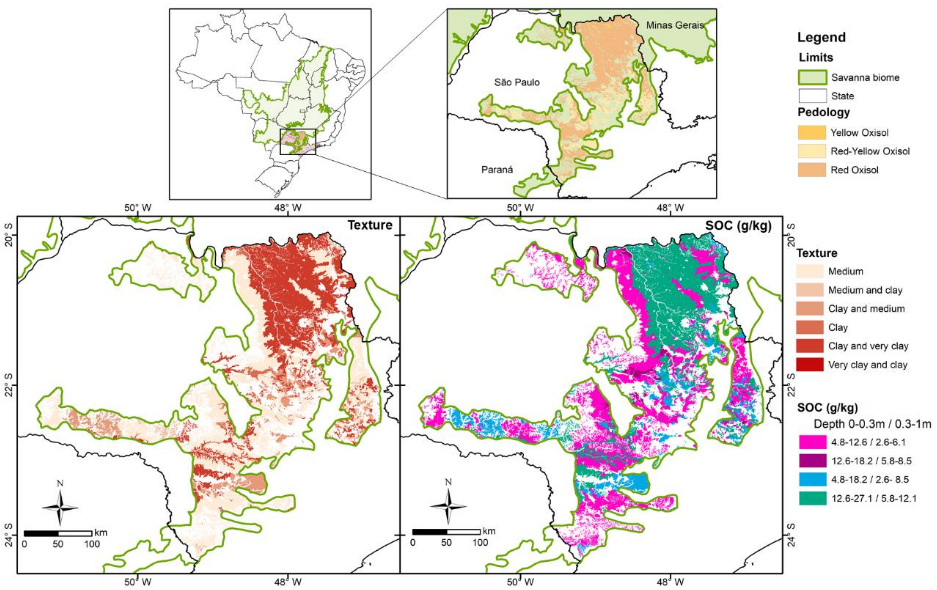

Figure 6. Soil organic carbon distribution in topsoil and subsoil according to the texture.

All results presented so far were restricted to Oxisol order. However, it is an important issue that these models have applicability to other soil orders. Thus, we evaluated eleven different soil orders that occur in São Paulo state, using the same procedure adopted for the Oxisol (Appendix **Figure A4**). The soils of the orders Mollisol and Psamment had an accuracy above 75% for the subsoil and Ultisol an accuracy of 67%. In the case of topsoil, the highest accuracy was 67% for Psamment. These results are only indicative that our models can assist in other soil SOC estimations However, we cannot confirm the SOC models effectiveness for other soils since there is no statistically significant and comparable sampling of soil profiles between soil orders.

4. Discussion

The models were sufficiently validated to prove their merit as a general estimation method or, at least, as a prior trial to the research. We have provided evidence that the models compute the soil organic carbon values close to the real values. This is important mainly to the agricultural regions that do not own a robust soil database or financial resources for large surveys, but they need general mapped data regarding soil quality [37]. The facility to get a notion of the soil carbon in an area where the present status is completely unknown is the greatest strength of these models. Enabling a rough estimation of SOC with only soil texture knowledge means it is possible to work with very low cost in huge areas. According to Vos et al. [38], the soil texture can easily be estimated in the field survey with a relatively high precision using the "finger texturing". However, the models' performance depends on how the variables related to the input data could intervene into the triangle results. We recognize that there are more robust methods [4,39,40], involving other variables that exert influence on SOC and, therefore, resulting in a higher accuracy. Nevertheless, we believe that approach proposed in this work proved to be more than enough for areas where intensive soil survey and analytical effort are not capable of being accomplished.

The precision limits of the models are initially linked to the possible variation of SOC values in short territorial spaces. We found in the analyses a huge textural variation in Oxisols but we expected this result because many other researchers had already reported this fact [4]. However, this variation did not prevent the SOC from being modeled. This outcome is probably due to the strong-positive association between the SOC and clay content [41] and negative association with sand [42] in both studied layers of soil profile [32].

Our data showed more SOC in the topsoil layer than in the subsoil layer, overestimated up to 5 $g{\cdot}kg^{-1}$, which could be attributed both to the vertical distribution of roots [43–45] and for litter—responsible for up 50% of SOC variations in the levels on the soil surface. Around 62%–79% of the roots in the topsoil layer are concentrated in the first 0.2 m of soil [46], showing an exponential decline of the root biomass and its size according to the dept [47]. Besides the root concentration, the higher SOC values in the topsoil layer also reflect the land use [48] and its management, such as crop residues accumulation left on the top of the land for protection [49]. In addition, it must be considered that SOC amount on the soil surface is related to soil compaction, which would be measured by bulk density [50]. However, the higher carbon concentration in topsoil is not a standard, since other authors reported that the subsoil layer contains as much SOC as the topsoil layer [42,51]. On the other hand, the carbon soil variability in the topsoil of the Oxisols that we have studied is less compared to other researchers [18,52]. Ruggiero et al. [18], for example, reported SOC values from 11.7 up to 26.6 g·kg^{-1} in the topsoil and from 6.0 up to 13.7 g·kg^{-1} in the subsoil in Savanna region with both Oxisols and Entisols soils. Carvalho et al. [53] also worked with Oxisol in Savanna, describing SOC values from 19.5 to 24.7 g·kg^{-1} to topsoil (0.3 m). Then, the models' application must be attentive to the possible margin of error due to these ranges.

When we look at the topsoil and subsoil in greater fractionation of the soil column (**Figure 5** and **Table 4**), we can see the part of the soil that has better performance, which is the middle fraction of both topsoil and subsoil. These results can be influenced by land use-cover which provides different carbon inputs to the soil column. In Brazilian Savanna Oxisols 51% of the area is covered by native vegetation [21]. The vegetation upon these soils displays a mosaic of the structures from savanna-like formations to forest, contributing with different soil carbon input. The planted pasture (27%) was the second land use most presented in Brazilian Savanna Oxisols followed by annual agriculture (15%) and perennial agriculture with sugar cane (5%). The other land use-cover totalizing 2% [21]. Zhang et al. [13], for example, studied SOC profile distribution in 120 samples in native vegetation and agriculture areas, finding a small ranging in native vegetation (5.2 g·kg^{-1}—topsoil; 3.5 g·kg^{-1}—subsoil) and a bigger difference in agriculture area (12.0 g·kg^{-1}—topsoil; 4.9 g·kg^{-1}—subsoil). However, when we investigated a land portion of Brazilian Savanna Oxisols (Sao Paulo State), the models had better performance in topsoil, probably because the area was smaller which guarantees less environmental variability in terms of climate, relief, and native vegetation cover. According to TerraClass Cerrado [21], in 2013 this region had 46% of its area covered by perennial agriculture with sugar cane, 17% native vegetation cover, and 6% silviculture, with wide homogeneity in its distribution.

In short, the models' users must be aware of how the variables related to the input data could intervene in the punctual SOC or in SOC final map. However, the validation tests concerning the textural triangle both for individual data and SOC map, point to worthwhile practical applicability to be employed.

5. Conclusion

Our results highlight models grounded on soil textural classes that permit to estimate the value of Soil Organic Carbon in Savanna Oxisols. The data can be obtained from two textural triangles, one for topsoil and other for subsoil, based on the equilateral textural triangle. We warn that the models provide better support to SOC-related preliminary research in gross and fast estimates, giving a general notion of potential Soil Organic Carbon and facilitating the empirical study.

Author contributions: Conceptualization, TNT, RACL and RFdS; methodology, TNT; software, TNT; validation, TNT; formal analysis, TNT, RACL and RFdS; investigation, TNT; resources, RACL; data curation, TNT; writing—original draft preparation, TNT; writing—review and editing, TNT and RFdS; visualization, TNT; supervision, RACL and RFdS; project administration, TNT and RACL; funding acquisition, RACL. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We gratefully acknowledge Alexandre C. Coutinho, Mara A. Marinho and Júlio C. M. D. Esquerdo for their constructive comments and technical advice. We especially thank Felipe F. Bocca and Marcos Adami for statistical advice and for all field trips; Michender W. M. Pereira, Victor Danilo Manabe, Márcio Roberto da S. Melo and Rafael L. Cobra. And not less important, we thank the Embrapa Pecuária Sudeste and the farmers Zé Edílio and Lucia Rigotti for the support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001, CNPq (167618/2014-8) and FUNCAMP.

Conflict of interest: The authors declare no conflict of interest.

References

1. Muscolo A, Panuccio MR, Mallamaci C, et al. Biological indicators to assess short-term soil quality changes in forest ecosystems. Ecological Indicators. 2014; 45: 416-423. doi: 10.1016/j.ecolind.2014.04.047
2. Chenu C, Angers DA, Barré P, et al. Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. Soil and Tillage Research. 2019; 188: 41-52. doi: 10.1016/j.still.2018.04.011
3. Thoumazeau A, Bessou C, Renevier MS, et al. Biofunctool®: a new framework to assess the impact of land management on soil quality. Part A: concept and validation of the set of indicators. Ecological Indicators. 2019; 97: 100-110. doi: 10.1016/j.ecolind.2018.09.023
4. Gomes LC, Faria RM, de Souza E, et al. Modelling and mapping soil organic carbon stocks in Brazil. Geoderma. 2019; 340: 337-350. doi: 10.1016/j.geoderma.2019.01.007
5. Morais VA, Mello JM de, Mello CR de, et al. Spatial distribution of the litter carbon stock in the Cerrado biome in Minas Gerais state, Brazil. Ciência e Agrotecnologia. 2017; 41(5): 580-589. doi: 10.1590/1413-70542017415006917
6. Chen Z, Shuai Q, Shi Z, et al. National-scale mapping of soil organic carbon stock in France: New insights and lessons learned by direct and indirect approaches. Soil & Environmental Health. 2023; 1(4): 100049. doi: 10.1016/j.seh.2023.100049
7. McBratney AB, Santos MLM, Minasny B. On digital soil mapping. Geoderma. 2003; 117: 3-52. doi: 10.1016/S0016-7061(03)00223-4
8. Mutuku EA, Vanlauwe B, Roobroeck D, et al. Visual soil examination and evaluation in the sub-humid and semi-arid regions of Kenya. Soil and Tillage Research. 2021; 213: 105135. doi: 10.1016/j.still.2021.105135
9. Garsia A, Moinet A, Vazquez C, et al. The challenge of selecting an appropriate soil organic carbon simulation model: A comprehensive global review and validation assessment. Global Change Biology. 2023; 29(20): 5760-5774. doi: 10.1111/gcb.16896
10. MMA (Ministério do Meio Ambiente). Available online: http://www.mma.gov.br/biomas/cerrado (accessed on 3 January 2019).
11. Scaramuzza CA de M, Sano EE, Adami M, et al. Land-use and land-cover mapping of the Brazilian Cerrado based mainly on landsat-8 satellite images. Revista Brasileira de Cartografia. 2017; 69(6). doi: 10.14393/rbcv69n6-44309
12. Mello FFC, Cerri CEP, Davies CA, et al. Payback time for soil carbon and sugar-cane ethanol. Nature Climate Change. 2014; 4: 605-609. doi: 10.1016/j.geoderma.2009.05.015
13. Zhang J, Wang X, Wang J. Impact of land use change on profile distributions of soil organic carbon fractions in the Yanqi Basin. CATENA. 2014; 115: 79-84. doi: 10.1016/j.catena.2013.11.019
14. Cherubin MR, Franco ALC, Cerri CEP, et al. Sugarcane expansion in Brazilian tropical soils—Effects of land use change on soil chemical attributes. Agriculture, Ecosystems & Environment. 2015; 211: 173-184. doi: 10.1016/j.agee.2015.06.006
15. Bordonal R de O, Lal R, Ronquim CC, et al. Changes in quantity and quality of soil carbon due to the land-use conversion to sugarcane (Saccharum officinarum) plantation in southern Brazil. Agriculture, Ecosystems & Environment. 2017; 240: 54-65. doi: 10.1016/j.agee.2017.02.016

16. Deng L, Zhu G yu, Tang Z sheng, et al. Global patterns of the effects of land-use changes on soil carbon stocks. Global Ecology and Conservation. 2016; 5: 127-138. doi: 10.1016/j.gecco.2015.12.004
17. Zinn YL, Lal R, Resck DVS. Texture and organic carbon relations described by a profile pedotransfer function for Brazilian Cerrado soils. Geoderma. 2005; 127(1-2): 168-173. doi: 10.1016/j.geoderma.2005.02.010
18. Ruggiero PGC, Batalha MA, Pivello VR, Meirelles ST. Soil-vegetation relationships in Cerrado (Brazilian savanna) and semideciduous forest, Southeastern Brazil. Plant Ecology. 2002; 160: 1-16.
19. Neufeldt H, Resck DVS, Ayarza MA. Texture and land use effects on soil organic matter in Cerrado Oxisols, Central Brazil. Geoderma. 2002; 107: 151-164.
20. Reatto A, Correia JR, Spera ST, Martins ES. Soils of the Cerrado Biome: pedological aspects (Portuguese). In: Sano SM, Almeida SP, Ribeiro JF (editors). Cerrado ambiente e flora. Planaltina: Empresa Brasileira de Pesquisa Agropecuária. 2008. pp. 107-149.
21. Brasil. Ministry of the Environment. Mapping the Use and Cover of the Cerrado: TerraClass Cerrado Project 2013 (Portuguese). Brasília: MMA; 2015. 67p.
22. Alvares CA, Stape JL, Sentelhas PC, et al. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift. 2013; 22(6): 711-728. doi: 10.1127/0941-2948/2013/0507
23. Coutinho LM. Brazilian biomes (Portuguese). São Paulo: Oficina de Textos; 2016.
24. Haridasan M. Aluminium accumulation by some cerrado native species of central Brazil. Plant and Soil. 1982; 65(2): 265-273. doi: 10.1007/bf02374657
25. Embrapa. Brazilian Agricultural Research Corporation—National Soil Research Center (Portuguese). Sistema Brasileiro de Classificação de Solos. Embrapa, Brasília, 4ed. 2014.
26. Lopes AS. Cerrado soils: characteristics, properties and management (Portuguese). 1937—Piracicaba: Instituto da Potassa & Fosfato: Instituto Internacional da Potassa; 1983. 162p.
27. Parizzi TNT, Bento DF. Soil characteristics of savannah soils across Brasil. Published online 2017. doi: 10.1594/PANGAEA.881658
28. Bishop TFA, McBratney AB, Laslett GM. Modelling soil attribute depth functions with equal-area quadratic smoothing Splines. Geoderma, Amsterdam. 1999; 91: 27-45. doi: 10.1016/S0016-7061(99)00003-8
29. Malone BP, McBratney AB, Minasny B, Laslett GM. Mapping continuous depth functions of soil carbon storage and available water capacity. Geoderma. 2009; 154: 138-152. doi: 10.1038/s41597-023-02056-8
30. Pereira MWM. Modeling the continuous variation of organic carbon in soil depth (Portuguese). Dissertação (Mestrado em Engenharia Agrícola)—Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas., Campinas; 2014.
31. Mtama JG, Burras CL, Msanya BM. Pedotransfer Functions for Cation Exchange Capacity, Available Water Holding Capacity and Soil Organic Carbon for Representative Soils of Southern Highland Zone of Tanzania. International Journal of Agriculture and Biological Sciences. 2018; 2522-6584.
32. Mwango SB, Wickama J, Msanya BM, et al. The use of pedo-transfer functions for estimating soil organic carbon contents in maize cropland ecosystem in the Coastal Plains of Tanzania. CATENA. 2019; 172: 163-169. doi: 10.1016/j.catena.2018.08.031
33. Efron B. The Jackknife, the Bootstrap and Other Resampling Plans. doi: 10.1137/1.9781611970319
34. Adami M. Estimating soybean planting dates using MODIS image time series (Portuguese). Tese (Doutorado em Sensoriamento Remoto), INPE, São José dos Campos; 2010. 163p.
35. Rossi M. Pedological map of the state of São Paulo: Revised and expanded (Portuguese). São Paulo: Instituto Florestal; 2017. 118p.
36. Davis ROE, Bennett HH. Grouping of soils on the basis of mechanical analysis. USDA Department Circular. 1927. 419.
37. Laub M, Blagodatsky S, Lang R, et al. A mixed model for landscape soil organic carbon prediction across continuous profile depth in the mountainous subtropics. Geoderma. 2018; 330: 177-192. doi: 10.1016/j.geoderma.2018.05.020
38. Vos C, Don A, Prietz R, et al. Field-based soil-texture estimates could replace laboratory analysis. Geoderma. 2016; 267: 215-219. doi: 10.1016/j.geoderma.2015.12.022
39. Morais VA, Ferreira GWD, de Mello JM, et al. Spatial distribution of soil carbon stocks in the Cerrado biome of Minas Gerais, Brazil. CATENA. 2020; 185: 104285. doi: 10.1016/j.catena.2019.104285
40. Padarian J, Minasny B, McBratney AB. Machine learning and soil sciences: a review aided by machine learning tools. SOIL. 2020; 6(1): 35-52. doi: 10.5194/soil-6-35-2020

41. Sakin E, Sakin ED. Relationships between particle size distribution and organic carbon of soil horizons in the Southeast area of Turkey. Bulgarian Chemical Communications. 2015; 47(2): 526-530.
42. Jobbágy EG, Jackson RB. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecological Applications. 2000; 10(2): 423-436. doi: 10.1890/1051-0761
43. Hilinski TE. Century 5: Implementation of Exponential Depth Distribution of Organic Carbon in the CENTURY Model. Department of Soil and Crop Sciences, Colorado State University; 2001.
44. Lorenz K, Lal R. The depth distribution of soil organic carbon in relation to land use and management and the potential of carbon sequestration in subsoil horizons. Adv. Agron. 2005; 88: 35-66. doi: 10.1016/S0065-2113(05)88002-2
45. Hiederer R. Distribution of Organic Carbon in Soil Profile Data. EUR 23980 EN. Luxembourg: Office for Official Publications of the European Communities; 2009; p. 126.
46. Castro EA, Kauffman JB. Ecosystem structure in the Brazilian Cerrado a vegetation gradient of aboveground biomass, root mass and consumption by fire. Journal of Tropical Ecology. 1998; 14: 263-283. doi: 10.1017/S0266467498000212
47. Smith DM, Inman-Bamber NG, Thorburn PJ. Growth and function of the sugarcane root system. Field Crops Research. 2005; 92(2-3): 169-183. doi: 10.1016/j.fcr.2005.01.017
48. Meersmans J, van Wesemael B, De Ridder F, et al. Modelling the three-dimensional spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma. 2009; 152(1-2): 43-52. doi: 10.1016/j.geoderma.2009.05.015
49. Franzluebbers AJ. Soil organic matter stratification ratio as an indicator of soil quality. Soil & Tillage Research. 2002; 66: 95-106. doi: 10.1016/S0167-1987(02)00018-1
50. Shah AN, Tanveer M, Shahzad B, et al. Soil compaction effects on soil health and cropproductivity: an overview. Environmental Science and Pollution Research. 2017; 24(11): 10056-10067. doi: 10.1007/s11356-017-8421-y
51. Batjes NH, Ribeiro E, van Oostrum A. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). Earth System Science Data. 2020; 12(1): 299-320. doi: 10.5194/essd-12-299-2020
52. Ruggiero PGC, Pivello VR, Sparovek G, et al. Relação entre solo, vegetação e topografia em área de cerrado (Parque Estadual de Vassununga, SP): como se expressa em mapeamentos? Acta Botanica Brasilica. 2006; 20(2): 383-394. doi: 10.1590/s0102-33062006000200013
53. Carvalho JLN, Cerri CEP, Cerri CC, et al. Changes of chemical properties in an oxisol after clearing of native Cerrado vegetation for agricultural use in Vilhena, Rondonia State, Brazil. Soil and Tillage Research. 2007; 96(1-2): 95-102. doi: 10.1016/j.still.2007.04.001

Appendix

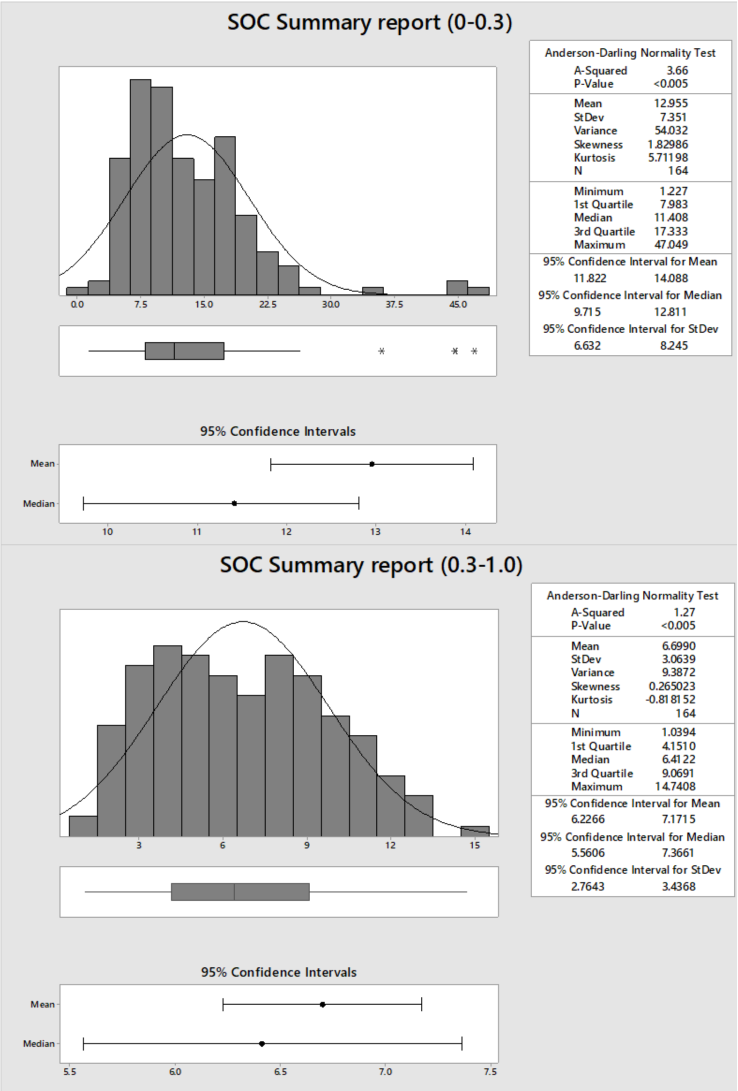

Figure A1. Descriptive analysis of SOC values in the two depth.

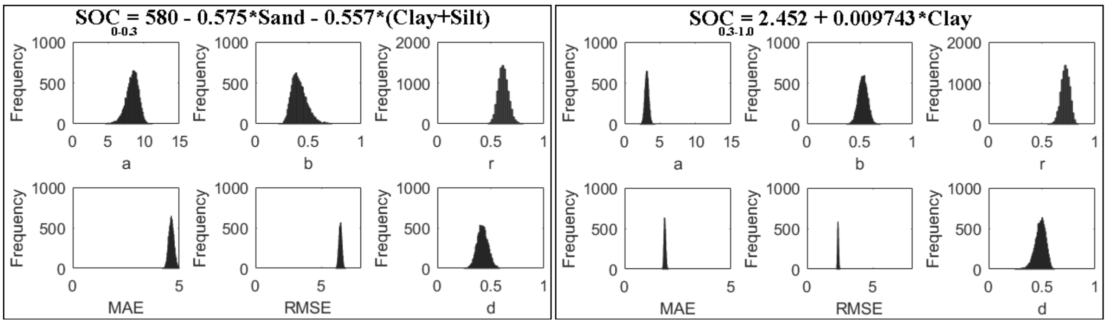

Figure A2. The values estimated by the bootstrap for the best models in the two depths were represented by the histograms above, where (*a*) linear coefficient, (*b*) angular coefficient, (*r*) correlation coefficient, (MAE), (RMSE) and (*d*) statistical parameters.

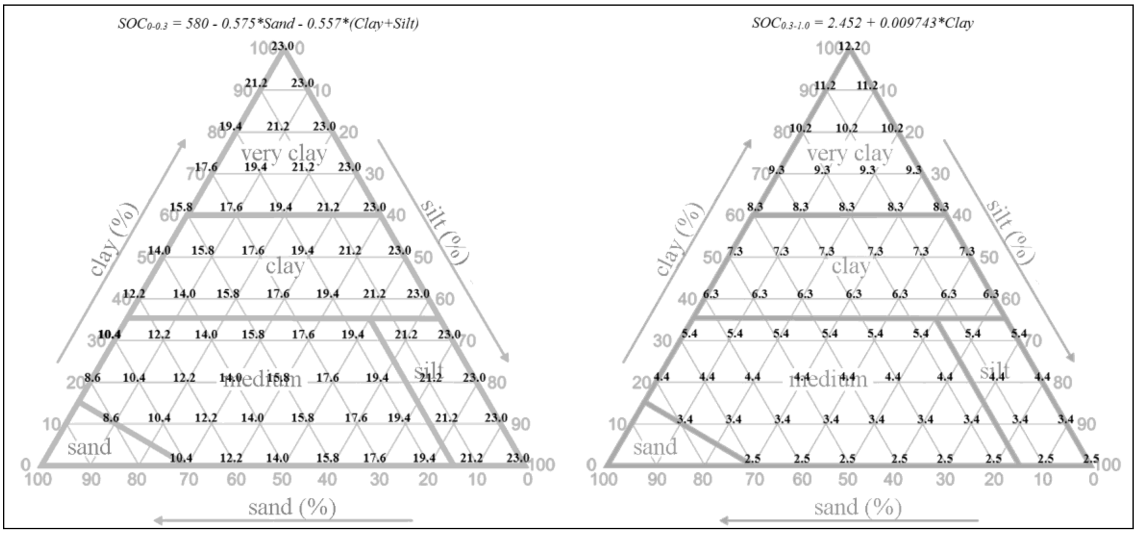

Figure A3. Estimated SOC (g·kg^{-1}) in Savanna Oxisols based on the textural triangle.

Soil order		Alfisol	Aqualf	Histosol	Inceptisol	Mollisol	Nitisol	Oxisol	Psamment	Spodosol	Udalf	Ultisol
0-0.3	N	16	5	3	3	5	30	14	3	2	3	31
(m)	Acurracy	56%	20%	0%	33%	40%	37%	86%	67%	0%	33%	58%
0.3 -1.0	N	20	5	8	2	5	30	40	4	2	5	51
(m)	Acurracy	50%	0%	0%	0%	80%	37%	70%	75%	0%	0%	67%

Figure A4. 287 pedological profile obtained from secondary sources of pedological profiles and ground surveys, between the years 1960 and 2015.

All of them were in São Paulo State. We calculated the percentual concordance between the values presented in pedological profiles and the estimated values by the triangle to the same points as a way of evaluating the models' effectiveness.