PREDICTION STATISTICAL MODEL FOR SOIL ORGANIC CARBON MAPPING IN CROP AREAS USING THE LANDSAT/OLI SENSOR

Manuela Corrêa de Castro Padilha[1], Luiz Eduardo Vicente[2], José Alexandre Melo Demattê[1], Daniel Gomes[2], Diego Urbina Salazar[1], Andrea Koga Vicente[2], Luciana Spinelli-Araujo[2], Celso Vainer Manzatto[2]

[1] University of São Paulo/ "Luiz de Queiroz" College of Agriculture, Av. Pádua Dias, 11 - Cx. Postal 9 - Piracicaba – SP CEP 13418-900 - PABX: +55 (19) 3429-4100, manuela.padilha2111@gmail.com, jamdemat@usp.br, diego_urbina_salazar@usp.br; [2]Low Carbon Agriculture Platform/ Embrapa Environment, Cx. Postal 69 – 13820 -000 – Jaguariúna – SP, Brasil, luiz.vicente@embrapa.br, daniel.gomes@embrapa.br, andrea.kvicente@gmail.com, luciana.spinelli@embrapa.br, celso.manzatto@embrapa.br.

ABSTRACT

The quantification of soil organic carbon (SOC) is essential to agriculture and sustainable use of the land. However, there are difficulties to estimate it in large areas due to high cost of soil sample extraction, and laboratory preparations. There are approaches that may facilitate the estimation of SOC, such as the use of satellite imagery and the application of statistical models based on the spectral bands of the satellite under study. In July of 2017, this study proposed a prediction statistical model from optical-orbital data of the series Landsat, OLI sensor for estimating SOC content.

Key words **—** linear model, linear regression, soil organic carbon, Landsat OLI.

1. INTRODUCTION

The land use/land cover may change (Land use land cover change – LULUC) the physical and chemical properties of soil natural ecosystems as a result of the production systems adopted by farmers. The production systems result in an increase or decrease in Soil Organic Carbon (SOC), and in tropical soils its impact affect crop productivity.

Thus, information about soil physical properties is extremely important to evaluate the land management practices that promote the increase of carbon in the terrestrial environment [1]. Therefore, knowledge on SOC spatial distribution pattern provides valuable information to agricultural systems, such as the sustainable land use, and for identifying soil uses and management that promote carbon sequestration [2].

In this context, the use of Geographic Information Systems (GIS) may be used as a monitoring tool to identify soil use sweeps, slope, hydrology, geology, among others, in the landscape. By using this tool, farmers may plan a productive system. The variables may be related to hydrological, geomorphological, and biological processes [3], aiming productivity gains and optimization of agricultural production costs.

The objective of the study is to evaluate the proposition of a predictive model by spatial interpolation/pedometric techniques considering several factors that affect the distribution of SOC for its estimation, aiming to optimize agricultural planning and management.

2. MATERIAL AND METHODS

The area of study is located in the Piracicaba region, state of São Paulo, Brazil, and comprises the municipalities of Charqueada, Piracicaba, Rio das Pedras, Saltinho, Mombuca, Tietê and Rafard, all belonging to the Tietê basin. According to [4], the climatic classification is tropical Cwa type, with dry winters and rainy summers. The region is recognized by the growing economic polo of agribusiness in soybean crops and mainly sugar cane.

For sampling, it was necessary to evaluate the quality of the scene, and to select the image pixels corresponding to the exposed soil visible from the creation of an exposed soil mask (ESM), according to the methodology adapted from [5].

The sampling points were marked in the same geographical position as each ESM. They were selected according to the variation of soil type, which is the largest spectral variation of the study area, and also the amount of exposed soil present in the images of interest.

The contents of the chemical and structural elements of the soil, such as organic matter, clay, sand, silt, and other elements were obtained in laboratory analysis.

We followed the studies of [6] to obtain a statistical model of SOC prediction through the spectral bands of the Landsat, due to the fact that the estimation of organic carbon uses the spectral range VIS-NIR-SWIR (350 to 2500nm).

The carbon estimation model proposed in this study was applied in the Landsat scene on July 17th, 2017. The Linear Regression (LR) was used with samples from 227 boreholes. The reflectance values of bands 5 and 7 were extracted from the pixels of the sampled points. [7], [8] and [9] were used to perform the underlying assumptions of linear models, such as normality of residuals, independency of residuals, and constant variance. The p-value indicates the acceptance of the null hypothesis for each test (residuals normally distributed, independent and with constant variance, respectively), and adopted significance level threshold was 0.05 [10]. The predictive performance of the model was evaluated based on the adjusted coefficient of determination (R^2), the Root Mean-Squared Error (RMSE), and Ratio of Performance to Interquartile Range (RPIQ) obtained in data validation.

Figure 1 shows the methodology used in this study in an illustrated way.

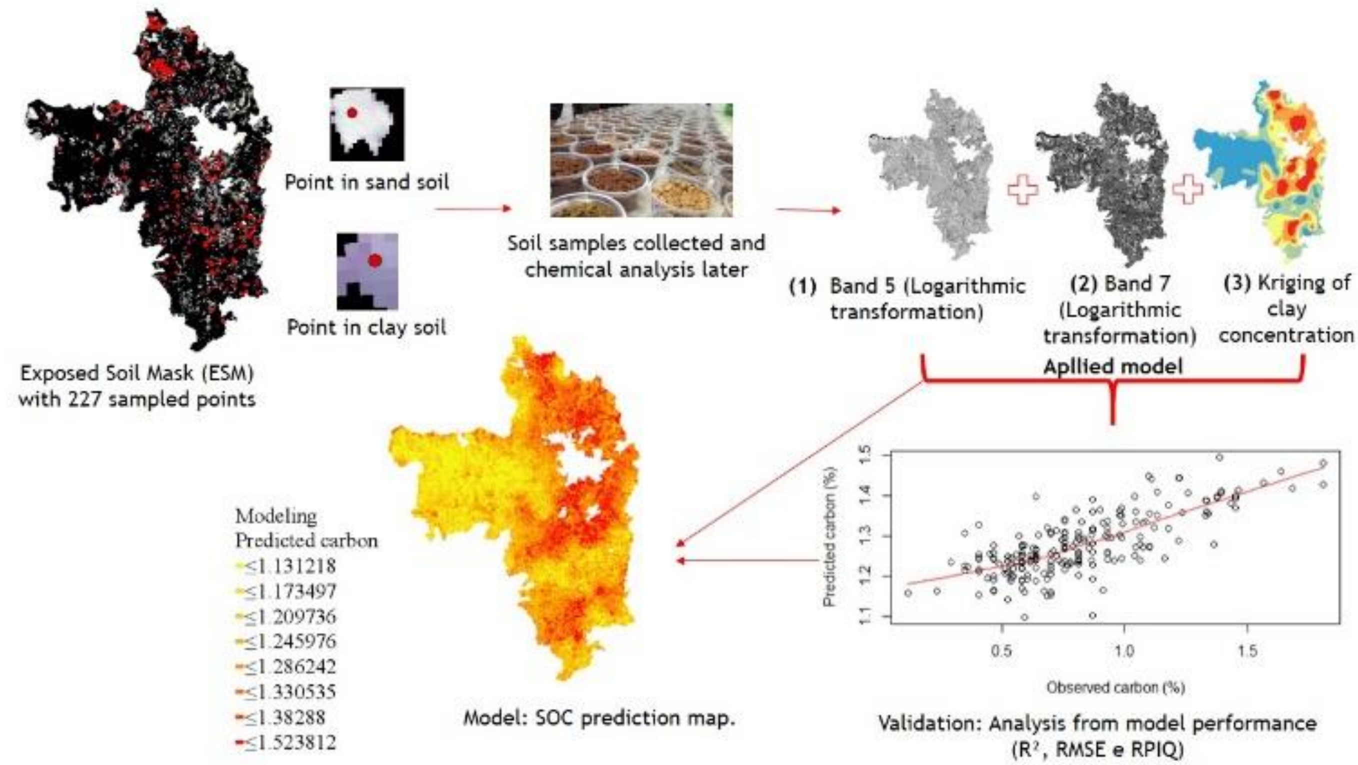

Figure 1. Methodology for applying the SOC prediction model.

3. RESULTS AND DISCUSSION

The linear model was implemented for SOC estimation including different variables. However, these individually assumed variables with carbon were not suitable for making such prediction (Table 1). Using the added variables, the model presented statistical significance (p-value > 0.05).

Table 1. P-values obtained from normality of residuals (Shapiro-Wilk), independency of residuals (Durbin-Watson) and constant variance (Breusch-Pagan) tests in LR model.

Strategy	Normality of residuals	Independency of residuals	Constant variance
		Carbon	
(Band 5_ logarithmic)	0.0467	0	0.5431
(Band 7_ logarithmic)	0.2682	0	0.0402
(Clay)	0.5607	0.058	0.0926
(Band 5_ logarithmic + Band 7_ logarithmic + Clay)	0.1046	0.882	0.0715

For the simple LR, the model presented R^2 values approximately, 0.8 and 0.6 for calibration and validation, respectively. Values of RMSE and RPIQ presented significance, as the error was low according to RMSE. It presented high RPIQ indicating good sampling distribution around the trendline (Table 2 and Figure 2).

Table 2. Calibration and validation results from LR model for prediction of SOC.

Model:	Calibration			Validation		
	R^2	RMSE	RPIQ	R^2	RMSE	RPIQ
(Band 5_ logarithmic + Band 7_ logarithmic + Clay)	0.7563	0.035	8.0377	0.5802	0.5291	0.5316

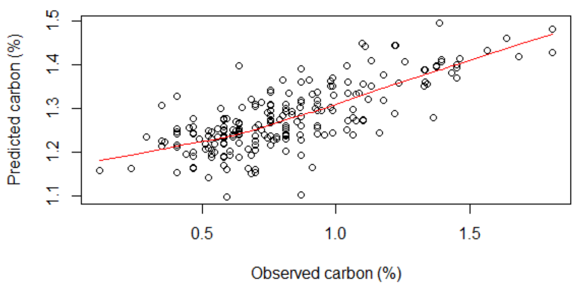

Figure 2. Validation graphic from observed carbon and predicted carbon obteined from LR model.

A map was generated with the SOC prediction model applied to the Landsat satellite image (Figure 3), where it is possible to observe the high and low values (%) of SOC in the study region.

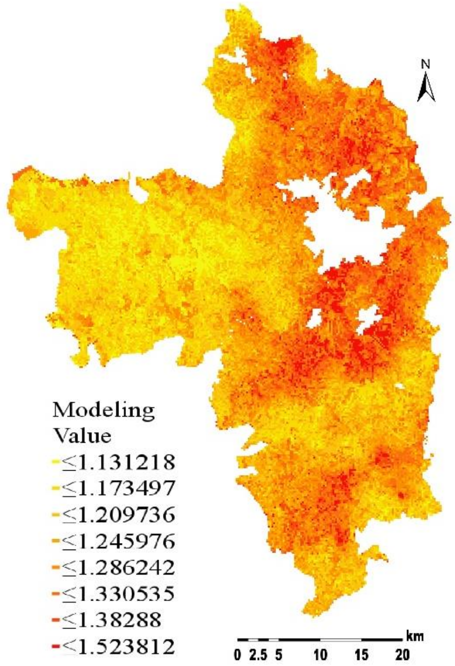

Figure 3. SOC prediction map in the study region with the predominance of sugar cane, but also corn, orange, manioc crops, and pasture too.

5. CONCLUSION

The study evaluated the significance of using spectral bands of Landsat satellite applied for the estimation of SOC in Piracicaba region, state of São Paulo. The application of the model, which contains spectral bands added to the clay concentration of soil revealed significant levels of accuracy. Thus, a more discerning and fast analysis for the producer and for companies that search cost reduction with soil collection, and high productivity in short time spent collecting.

ACKNOWLEDGMENTS

The authors would like to aknowledge CnPq for the scholarship granted to Manuela Corrêa de Castro Padilha.; to the LCA Platform and the Geotechnologies and Quantitative Methods Laboratory teams at Embrapa Environment (site: https://www.embrapa.br/meio-ambiente/plataforma-abc), and to the Research Group on Geotechnology in Soil Science (GeoCis, site: http://esalqgeocis.wixsite.com/geocis) for making available the resources of the laboratories; and the support of the thematic project FAPESP 2014-22262-0.

REFERENCES

[1] Marques, J. D. de O. et al. "Carbono orgânico do solo, estoque de carbono e sua relação com atributos físicos sob solos florestais na amazônia central". *Scielo*, Viçosa, v. 40, n. 2, mar./abr. 2016. Available from: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=s0100-67622016000200197&lng=en&tlng=en>. Accessed on June 4, 2018.

[2] Machado, P. L. O. A. "Carbono do solo e a mitigação da mudança climática global". Rio de Janeiro, vol. 28, n. 2, p. 329-334, 2005.

[3] Moore, I.D. et al. "Digital terrain modeling: a review of hydrological, geomorfological and biological applications Hydrological Processes", v.5, p.3-30, 199.

[4] Rolim, Glauco de Souza et al. "Classificação climática de Köppen e de Thornthwaite e sua aplicabilidade na determinação de zonas agroclimáticas para o estado de são Paulo". Bragantia. Instituto Agronômico de Campinas, v. 66, n. 4, p. 711-720, 2007. Available from: <http://hdl.handle.net/11449/27939>. Accessed from: July 5, 2018.

[5] Demattê J.A.M., Alves, M.R., Terra, F.S., Bosquilia, R.W.D., Fongaro, C.T., Barros, P.P.S. "Is It Possible to Classify Topsoil Texture Using a Sensor Located 800 km Away from the Surface?" *Rev Bras Cienc Solo*. 2016; 40:e0150335.

[6] Knadel M., Y. Peng, K. Schelde, A. Thomsen, F. Deng, and M. H. greve. 2013a. "Optimal sampling and sample preparation for NIR-based prediction of field scale soil properties". Poster presented at European Geosciences Union General Assembly, Vienna

[7] Shapiro, S. S., Wilk, M. B., 1965. "Na analysis of variance test for normality (complete samples)". *Biometrika*. 52, 591-611.

[8] Durbin, J., Watson, G. S., 1950. "Testing for Serial Correlation in Least Squares Regression I". *Biometri*ka. 37, 409-428.

[9] Breusch, T. S., Pagan, A. R., 1979. "A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica." 47, 1287-1294.

[10] Sayão, V. M. "Land surface temperature and reflectance spectra integration obtained from Landsat on the soil atributes quantification." Piracicaba, 2017. *Dissertação (Mestrado) Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz".* 80p.