CHAPTER 3

SOIL SPECTRAL LIBRARY AND SOIL CARBON CONTENT ASSESSMENT FOR THE STATE OF RIO DE JANEIRO

Gustavo M. Vasques, Levi B. Luz, Fabiano C. Balieiro, Monise A. F. Magalhães, Telmo B. Silveira Filho, Tatiane Mesquita Araújo, Marcelo T. Andrade

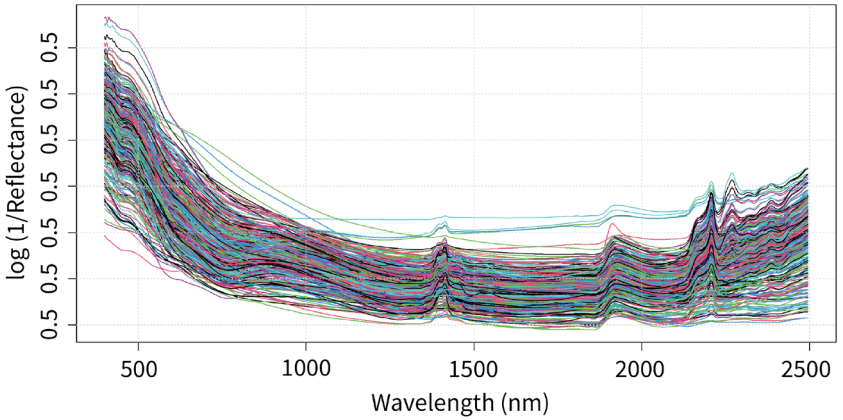
The Brazilian National Forest Inventory (NFI) is a cornerstone program for monitoring the country's forest resources, providing data to support public policies and meet international climate commitments. A key component of the forest functioning and dynamics is soil carbon. To support the NFI, soil spectral libraries (SSL) along with spectroscopy-based models can be developed to rapidly and cost-effectively estimate soil carbon content, allowing to calculate carbon stocks and changes over time.

For this aim, a visible-near infrared (Vis-NIR) SSL was built for the state of Rio de Janeiro, and Vis-NIR spectroscopy model was derived to predict the carbon contents in NFI soil samples across the state. Soil Vis-NIR spectroscopy is non-destructive, rapid, inexpensive, and precise, and has been used to estimate various soil properties, including carbon (Viscarra Rossel *et al.*, 2006; Demattê *et al.*, 2019).

To produce the SSL, soil samples at 0-20 and 30-50 cm depths were used, obtained from the Rio de Janeiro state NFI (SFB, 2018), which took place from 2013 to 2016 and employed a systematic grid sampling of 20 x 20 km, with a total of 251 sampling sites. A total of 355 samples (174 at 0-20 cm, and 181 at 30-50 cm) were used in the study. The NFI soil samples were ground, sieved (2 mm), and dried at room temperature. Then, the soil carbon content was measured by dry combustion in a CHNS 2400 elemental analyzer (Perkin Elmer, Waltham, EUA).

Soil spectral curves at the Vis-NIR range (350-2500 nm) were obtained in the laboratory using an ASD FieldSpec 4 spectroradiometer (Malvern Panalytical, Malvern, United Kingdom), averaging 100 repetitions per sample, and Spectralon® (Labsphere, North Sutton, USA) as a white reference. Before spectral reading, the samples were dried at 45 °C overnight to harmonize the water content. Once collected, the soil spectral curves were transformed to log(1/reflectance) (Figure 1).

FIGURE 1. Soil visible-near infrared spectral curves transformed to log(1/reflectance)



Source: Authors.

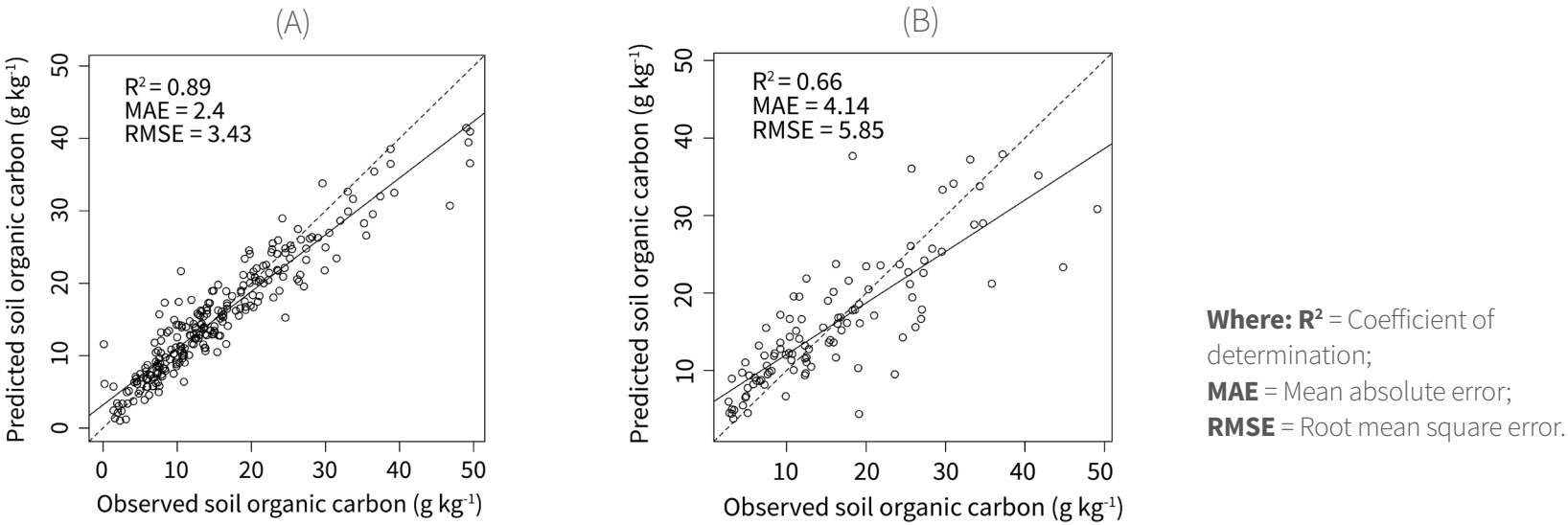
The Vis-NIR-based soil carbon content prediction model was derived using cubist (Quinlan, 1993) via the Cubist package (Kuhn and Quinlan, 2024) in R (The Comprehensive R Archive Network, 2024). Cubist is a multivariate method that combines decision tree-based modeling to split the dataset into homogeneous subsets with multiple linear regression to predict the target variable in each subset. This method has been successfully used to predict soil carbon in other regions in Brazil (Demattê *et al.*, 2019; Moura-Bueno *et al.*, 2021).

Prior to modeling, the samples were randomly split into a training (248 samples ~70%) and a testing set (107 samples ~30%). The training samples

were exclusively used to calibrate the cubist model, whereas the testing set served to validate the model soil carbon predictions. A committee of five cubist models was used to improve the accuracy of the predictions.

Soil carbon contents in the Rio de Janeiro NFI vary from 0.1 to 49.5 g kg^{-1}, with a mean, median and standard deviation of 15.5, 13.2 and 9.9 g kg^{-1}, respectively. Soil carbon content predictions vary from 1.0 to 41.4 g kg^{-1}, with a mean, median and standard deviation of 15.5, 14.1 and 8.2 g kg^{-1}, respectively. Observed and predicted soil carbon contents were highly correlated in both model training, with a R2 of 0.89 and root mean square error (RMSE) of 3.4 g kg^{-1} (Figure 2A), and validation, with a R2 of 0.66 and RMSE of 4.1 g kg^{-1} (Figure 2B).

FIGURE 2. Predicted versus observed soil carbon content in the of: (A) model training; and (B) external validation



Where: R^2 = Coefficient of determination;
MAE = Mean absolute error;
RMSE = Root mean square error.

Source: Authors.

The cubist models obtained soil carbon content predictions with similar or better accuracy than previous studies in other regions (Moura-Bueno *et al.*, 2021) or nationwide (Demattê *et al.*, 2019). Regionally, Moura-Bueno *et al.* (2021) achieved RMSE of 5.3 to 6.6 g kg^{-1} for soil carbon prediction from Vis-NIR data alone or combining Vis-NIR, spectral classes and environmental covariates, whereas Demattê *et al.* (2019) found a a slightly higher RMSE of 6.9 g kg^{-1}.

The Vis-NIR spectroscopy models for soil carbon content prediction at 0-20 and 30-50 cm for Rio de Janeiro state allow assessing soil carbon contents, stocks and changes, supporting future NFI campaigns across the state. Other statewide soil-related projects may benefit from the model produced here to expedite soil analysis. Accordingly, SSL and Vis-NIR spectroscopy-based prediction models may be extended to other soil properties as well as other states and regions in Brazil.

REFERENCES

DEMATTÊ, J. A. M. *et al.* The Brazilian Soil Spectral Library (BSSL): a general view, application and challenges. **Geoderma**, v. 354, p. 113793, 2019. DOI: https://doi.org/10.1016/j.geoderma.2019.05.043. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0016706118318548. Acessed: 01 set. 2025.

KUHN, M.; WESTON, S.; KEEFER, C.; COULTER, N.; QUINLAN, R. Cubist: rule- and instance-based regression modeling. Versão 0.5.0. *In*: THE COMPREHENSIVE R ARCHIVE NETWORK. [Viena, Áustria]: CRAN, 2024. Available at: https://CRAN.R-project.org/package=quantregForest. Acessed: 01 set. 2025.

MOURA-BUENO, J. M.; DALMOLIN, R. S. D.; HORST-HEINEN, T. Z.; GRUNWALD, S.; TEN CATEN, A. Environmental covariates improve the spectral predictions of organic carbon in subtropical soils in southern Brazil. **Geoderma**, v. 393, p. 114981, 2021. DOI: https://doi.org/10.1016/j.geoderma.2021.114981. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0016706121000550. Acessed: 01 set. 2025.

QUINLAN, J. S. Combining instance-based and model-based learning. *In*: INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 10., 1993, Amherst. **Proceedings** [...]. DOI: https://doi.org/10.1016/j.geoderma.2021.114981. Amherst: ACM, 1993. p. 236-243. Available at: https://dl.acm.org/doi/10.5555/3091529.3091560. Acessed: 01 set. 2025.

THE COMPREHENSIVE R ARCHIVE NETWORK. The R Project for Statistical Computing. *In*: THE COMPREHENSIVE R ARCHIVE NETWORK. [Viena, Áustria]: CRAN, [2024]. Available at: https://www.R-project.org. Acessed: 01 set. 2025.

ROSSEL, R. A. V.; WALVOORT, D. J. J.; MCBRATNEY, A. B.; JANIK, L. J.; SKJEMSTAD, J. O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. **Geoderma**, v. 131, n. 1-2, p. 59-75, 2006. DOI: https://doi.org/10.1016/j.geoderma.2005.03.007. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0016706105000728. Acessed: 01 set. 2025.

SERVIÇO FLORESTAL BRASILEIRO. **Inventário Florestal Nacional**: Rio de Janeiro: principais resultados. Brasília, DF: MMA, 2018. Available at: https://www.gov.br/florestal/pt-br/centrais-de-conteudo/publicacoes/relatorios/relatorios-ifn/IFNRJprincipaisresultados.pdf. Acessed: 01 set. 2025.

AUTHORS BIOGRAPHIES

Gustavo de Mattos Vasques
Is a Forestry Engineer from the Federal University of Viçosa (UFV) and holds a Ph.D. in Soil Science from the University of Florida (UF). He has expertise in data science, geophysics, and geotechnologies applied to soil science. Gustavo has been a researcher at Embrapa Soils since 2011. (gustavo.vaques@embrapa.br).

Levi Barreto da Luz
Is an undergraduate student in Industrial Chemistry at the Fluminense Federal University (UFF). He worked as a collaborator at Embrapa Soils between 2023 and 2025. (leviluz2000@gmail.com).

Fabiano de Carvalho Balieiro
Is an Agronomist, with a Master's in Soil and Plant Nutrition from the Federal University of Viçosa (UFV) and a Ph.D. in Soil Science from the Federal Rural University of Rio de Janeiro (UFRRJ). He has been a Researcher at Embrapa Soils since 2007 and a permanent professor in the Graduate Program in Environmental and Forest Sciences at UFRRJ since 2019. His research focuses on nutrient cycling and organic matter dynamics in natural and planted forests, as well as other agroecosystems. (fabiano.balieiro@embrapa.br).

Monise Aguillar Faria Magalhães
Is a Forestry Engineer with a Master's in Environmental and Forest Sciences from the Federal Rural University of Rio de Janeiro (UFRRJ), specializing in conservation units. She has experience in environmental and forest project management and public policies, with over 15 years of professional experience. She currently works for the Superintendent

of Climate Change and Forests at the Subsecretariat of Climate Change and Biodiversity Conservation of the State Secretariat for Environment and Sustainability of Rio de Janeiro (SEAS). (monise.seas@gmail.com).

Telmo Borges Silveira Filho
Is a Forestry Engineer, with a Master's and a Ph.D. in Environmental and Forest Sciences from the Federal Rural University of Rio de Janeiro (UFRRJ). He has experience in forest management and has served as a public servant since 2006. Telmo is the Superintendent of Climate Change and Forests at the Subsecretariat of Climate Change and Biodiversity Conservation of the State Secretariat for Environment and Sustainability of Rio de Janeiro (SEAS). (telmoborges.florestal@gmail.com).

Tatiane Mesquita de Araujo
Is an undergraduate student in Geology at the Federal University of Rio de Janeiro (UFRJ). She has experience in soil geochemistry and pedometrics. From January 2023 to May 2025, she was a collaborator at Embrapa Soils. (tatiane.araujo.7813@gmail.com).

Marcelo Teixeira de Andrade
Is a Chemist with Technological Expertise from the Federal University of Rio de Janeiro (UFRJ) and holds a Master's degree in Chemistry from the same university. He has experience in analytical instrumentation, chemistry of natural products, analytical chemistry, and soil chemistry. Marcelo has been an Analyst at Embrapa since 2006 and at Embrapa Soils since 2008. (marcelo.andrade@embrapa.br).