Soil carbon contents derived from VisNIR in the Rio de Janeiro State

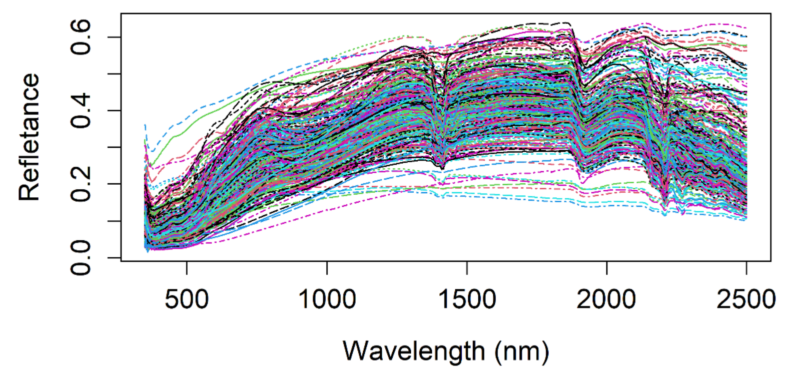
Introduction

Measuring soil carbon content dry combustion or wet chemistry requires extensive sample preparation and labor, making these methods too costly for monitoring soil carbon. Visible-near infrared (VisNIR) spectroscopy expedites soil carbon assessment with minimum sample preparation and reduced labor. The objective is to produce soil carbon content prediction models from soil VisNIR spectral curves for the Rio de Janeiro state, comparing different combinations of spectral pretreatments and prediction methods.

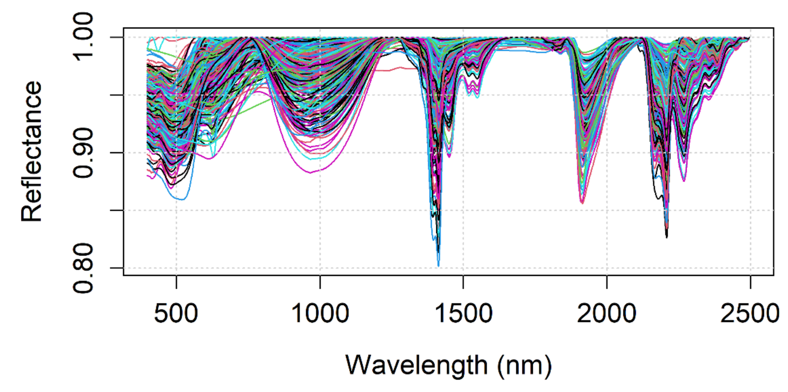
Methods

The soil samples were obtained from the 2013-2016 National Forest Inventory of Rio de Janeiro state (SFB, 2018) collected at 0-20 and 30-50 cm at 188 sites, comprising 376 samples. Training and testing samples were split by sampling site using a 70/30 percent ratio. Five pretreatments and five multivariate prediction methods were combined, and ten-fold cross-validation (CV) was used to optimize the method hyperparameters and find the best method-pretreatment combinations for soil carbon content prediction.

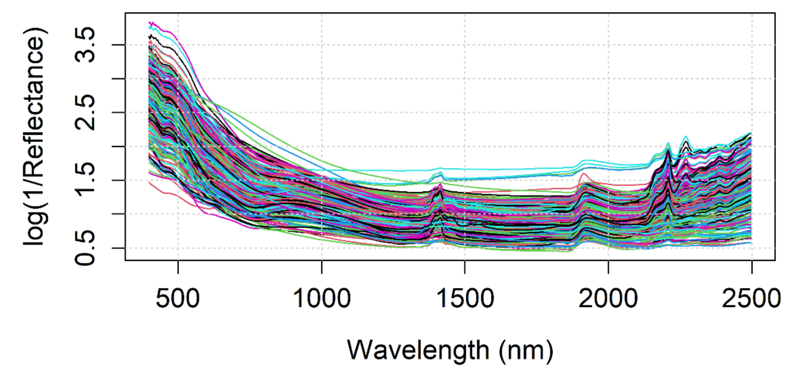
Results and discussion



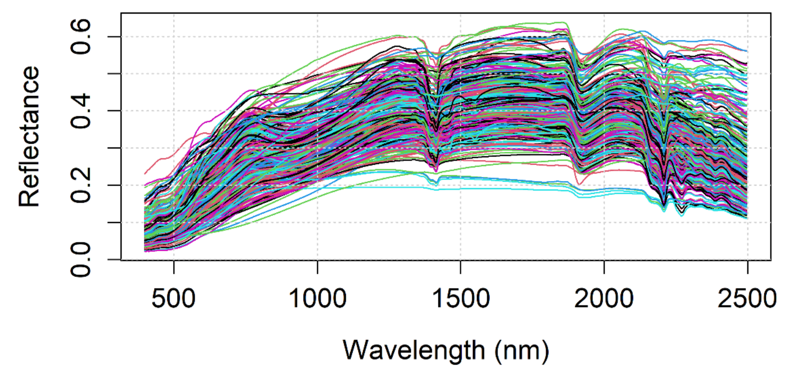
Original curves



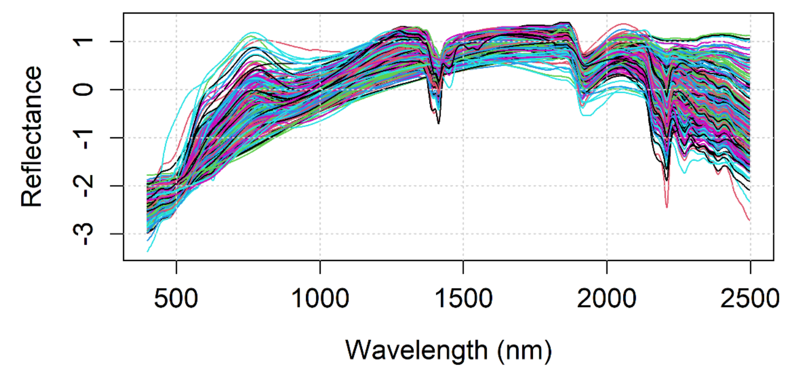
Continuum removal



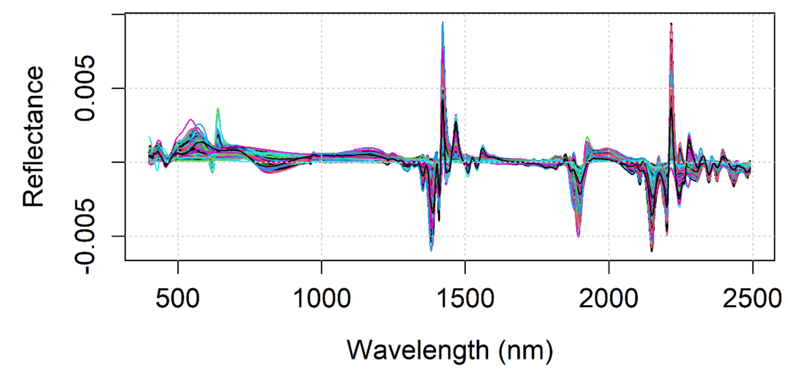
Log(1/Reflectance)



Savitzky-Golay smoothing



Standard normal variate



Savitzky-Golay 1st derivative

Descriptive statistics

	Min	Max	Median	Mean	SD
All	0.1	49.5	13.2	15.5	9.8
Training	0.1	49.4	13.3	15.3	9.7
Validation	2.7	49.1	12.8	15.9	10

Optimized methods and error metrics

Method	Pretreatment	R^2_{CV}	MAE_{CV}	$RMSE_{CV}$	R^2_T	MAE_T	$RMSE_T$	R^2_V	RPD_V	MAE_V	$RMSE_V$
Cubist	LOG	0.73	3.5	5.1	0.89	2.4	3.4	0.66	1.72	4.1	5.8
SVM	SNV	0.71	3.9	5.4	0.69	3.6	5.6	0.43	1.32	5.9	7.6
PLSR	LOG	0.71	4	5.4	0.76	3.4	4.7	0.66	1.72	4.1	5.8
Elastic net	SGD	0.67	4.3	5.6	0.8	3.2	4.3	0.57	1.53	4.8	6.5
RF	SGD	0.65	4.6	6.1	0.97	1.7	2.4	0.55	1.44	5.4	6.9

R^2: Coefficient of determination; CV: Cross-validation; MAE: Mean absolute error; RMSE: Root mean square error; T: Training; V: Validation; RDP: Residual prediction deviation; LOG: Log(1/reflectance); SVM: support vector machine; SNV: Standard normal variate; PLSR: Partial least squares regression; SGD: Savitzky-Golay 1st derivative; RF: Random forest.

Conclusion

The best soil carbon predictions were obtained from Cubist-Log(1/R), followed by SVM-SNV, and PLSR-Log(1/R). These models can be used to predict soil carbon content with reasonable accuracy. They offer a rapid alternative to assess soil carbon in different types of soils and landscapes across the Rio de Janeiro state, supporting soil carbon monitoring and credit accounting projects.

AUTHORS

Levi Luz[1], Gustavo Vasques[1], Fabiano Baliero[1], Telmo Silveira Filho[2], Monise Magalhães[2], Tatiane Araújo[3], Bárbara Andrade[4], Lygia Roque[4]

AFFILIATION OF AUTHORS

[1]Embrapa, [2]Secretaria de Estado do Ambiente e Sustentabilidade do Rio de Janeiro, [3]Universidade Federal do Rio de Janeiro, [4]Universidade Federal Rural do Rio de Janeiro

REFERENCES

SFB (Serviço Florestal Brasileiro) (2018). Inventário Florestal Nacional. Rio de Janeiro. Principais Resultados. Ministério do Meio Ambiente, Brasilia, Brazil.

ACKNOWLEDGEMENTS

Data provider: Serviço Florestal Brasileiro. Institutional support: SEAS/RJ and Embrapa Cooperation Agreement 25100.23/0109-8. Financial support: Climate Group Under2 Coalition Future Fund and Mata Atlântica Fund.

INDICATION OF THE CORRESPONDING AUTHOR

Gustavo Vasques
gustavo.Vasques@embrapa.br
Embrapa Solos

LATIN AMERICAN & CARIBBEAN
Soil Carbon Research Symposium

Rio de Janeiro, RJ, Brazil
June 25-28, 2025

CO-ORGANISED AND PROMOTED BY

ORGANISED BY