PREDICTION OF SOIL PROPERTIES WITH SPECTRORADIOMETRIC MEASUREMENTS

Yuri Andrei Gelsleichter[1], Mauro Antonio Homem Antunes[2], Robson Altiellys Tosta Marcondes[3], Lúcia Helena Cunha dos Anjos[4], Maria Aparecida Marques[5], Paula Debiasi[6], Helena Saraiva Koenow Pinheiro[7]

[1,5]Postgraduate Program in Science, Technology and Innovation in Agriculture, yuriplanta@gmail.com, mariamarquesadm@gmail.com; [2,6]Institute of Technology, homemantunes@gmail.com, pauladebiasi@yahoo.com.br; [3,4,7]Agronomy Institute, robson.marcondees@gmail.com, lanjos@ufrrj.br, lenask@gmail.com – Federal Rural University of Rio de Janeiro – Postcode 23897-000 – Seropédica – Rio de Janeiro, Brazil

ABSTRACT

Vis-NIR spectroscopy for soil has gained relevance among the development of soil technique analysis. Itatiaia National Park (INP) located at the Southern border of Rio de Janeiro and Minas Gerais states, has a fundamental role for water storage and distribution in the surrounded watersheds. INP is also one of the best exemplars of soil carbon stocks in the tropical regions. The purpose of this study was to test the Random Forest Algorithm for soil properties prediction. The spectral readings were done on 116 soil samples, in a completely dark room laboratory. Results show that the R^2 reached 0.57 for Hydrogen. Random Forest model created a rank of importance for the variables used in the prediction, in this stroke within this soil samples, the most important wavelength band was 2165 nm among others. The homogeneity and the limited number of samples suggest a disadvantage of the prediction models.

Key words – *soil reflectance, proximal soil sensing, soil properties modeling.*

1. INTRODUCTION

Proximal soil sensing specially in the visible and near-infrared reflectance (Vis-NIR) are fast, non-destructive, environmental-friendly and cost-effective [1]. Beginning studies were done by [2], who characterized the spectral signatures of soil minerals. Soil Organic Carbon (SOC) was spatially mapped in bare soils [3]. In south of Brazil, Soil Clay content was predicted and reached a value of R^2 = 0.75 [4].

Besides soil and environmental functions, SOC is an indicator of soil condition. Globally, soil is the second biggest pool of carbon [5], whereas the amount is moving among the pools. These movements can be linked to climate change.

Tropical regions usually are characterized by deeply weathered soils with lower SOC. A different situation is found in the plateau region of Itatiaia National Park (INP), since due to the high altitude and climate conditions, SOC is preserved and incorporated into the soil matrix. Among the environmental functions, the area is an important water catchment. During the raining season, the soils richer in organic matter have a larger capacity of retaining water [6]. The water stored in the soils is released along the year, including the dry season, and that feeds many springs that will supply large areas in Rio de Janeiro and boundary states. The main goal of this study was to test the Random Forest algorithm as a method for prediction of some properties of INP soils.

2. MATERIALS AND METHODS

2.1. Study area

The study area is located at the plateau (above 2.000 m) of the Itatiaia National Park (INP), in the Southwestern part of Rio de Janeiro State, as shown in Figure 1. The INP is located within the coordinates 22º15'E 22º30'S; 44º30'E 44º45'W, with a total area of 28.000 hectares and the climate is classified as Cwb [7]. The dry and cold winter creates favorable conditions to soils covered by grassland vegetation to hold SOC.

A total of 40 points were randomly collected at three soil depths (0-10, 10-20, 20-40 cm), the subseted area of high part is located at 22º22'S 44º42'W coordinates with a range of 42 hectares. In the laboratory, the samples were air dried and passed through a 2 mm mesh sieve. The soil properties, Soil Organic Matter (SOM), Soil Organic Carbon (SOC), C/N: ratio, Hydrogen, Nitrogen, were analyzed by dry combustion method.

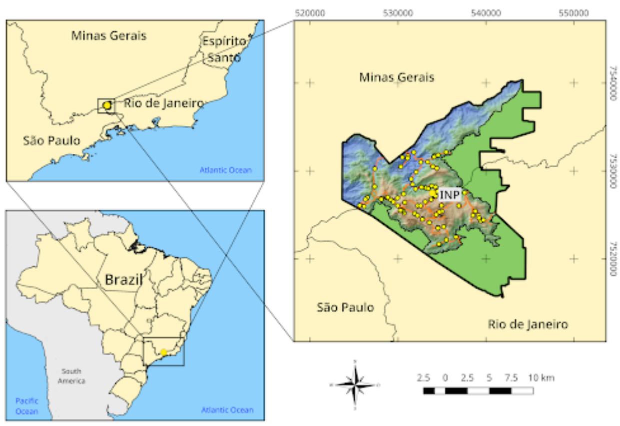

Figure 1: Location of Itatiaia National Park. Adapted from [8].

2.2. Spectral readings and processing

To ensure the moisture would not influence the spectral readings, soil samples were placed in the oven for 48 h at the temperature of 45 Celsius. All spectral readings were proceeded in a dark room using an ASD Fieldspec 4® with 30 scans per sample, and the samples placed in 9 cm petri dishes. Every 24 readings (approximately) the sensor was readjusted for light conditions using the *optimize* and *white reference* commands with a spectralon® plaque. The light source (70 W halogen bulb lamp) was placed in a distance of 40 cm and 15° from nadir. Fiber glass was positioned with an 8° lens at zero degree to the nadir in relation to the soil sample surface.

The spectral data were interpreted and processed R [9] software. Random Forest was performed using the randomForest [10, 11] R Package.

3. RESULTS AND DISCUSSION

With the application of Random Forest for the soil properties, Hydrogen estimated reached the highest R^2 with 0.57, (Figure 2), followed by SOM (0.43), and SOC (0.41). Cross validation was also used, and by applying pretreatments in the row spectral data [1] it was obtained the R^2 value of 0.85 for SOM. The stats coefficients for the other variables are shown in the Table 1.

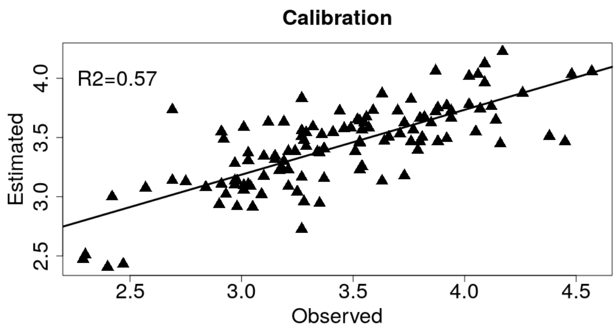

Figure 2: Estimated vs Observed values for hydrogen in the samples from INP.

Table 1: Soil properties coefficients for the INP soil samples.

Stats coef	Hydrogen	SOM	SOC	Nitrogen	C/N: ratio
R^2	0.57	0.43	0.41	0.28	0.05
RMSE	0.19	3.12	1.85	3.05	1.29
Min	2.29	16.93	9.82	0.49	14.00
Median	3.37	32.18	18.65	10.28	17.77
Max	4.57	49.72	28.80	17.25	25.98

Random Forest algorithm build a rank of importance for predictor variable. Algorithm stroke for Hydrogen, the most representative region in the spectrum wavelengths for prediction of Hydrogen were around 2000, 1000 and 900 nm; for the bands 2165, 1394, 1147 and 890 nm (Figure 3). It is noteworthy that these are not regions of water absorbance. A pretreatment of the specters could improve the results.

4. CONCLUSION

The Random Forest algorithm has good potential for prediction of soil properties such as in this study. However, since the soil samples were from relatively homogeneous soils (high SOC), the data set limited the extrapolation for

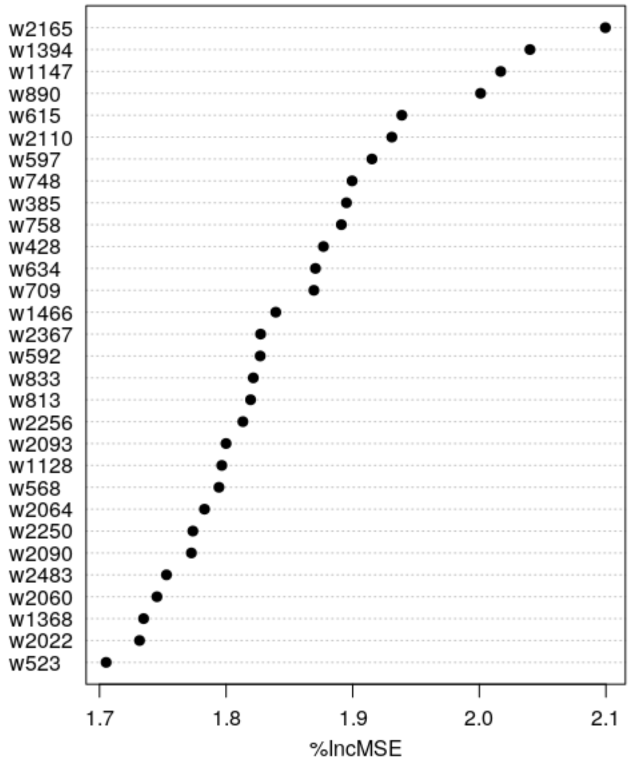

Figure 3: Importance of predictor variables for hydrogen in the samples from INP.

other classes with a wider variation of properties; also the small number of samples lead to some disadvantage for the prediction models.

ACKNOWLEDGMENTS

To the Federal Rural University of Rio de Janeiro, Graduate Program in Science Technology and Innovation in Agriculture (PPGCTIA), FAPERJ for the financial support of the research, FINEP for the financial support for the spectroradiometer, Itatiaia National Park for the cooperation, CNPq (IC/PIBIC) for the undergraduate scholarship. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

5. REFERENCES

[1] KOPAČKOVÁ, V. et al. Modelling Diverse Soil Attributes with Visible to Longwave Infrared Spectroscopy Using PLSR Employed by an Automatic Modelling Engine. *Remote Sensing*, v. 9, n. 2, p. 134, feb 2017. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/9/2/134>.

[2] HUNT, G. R. Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics*, v. 42, n. 3, 1977. Disponível em: <https://courses.eas.ualberta.ca/eas451/hunt1977.pdf>.

[3] GOMEZ, C.; Viscarra Rossel, R. A.; MCBRATNEY, A. B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*, v. 146, n. 3-4, p. 403–411, 2008. ISSN 00167061.

[4] DOTTO, A. C. et al. Potential of spectroradiometry to classify soil clay content. *Revista Brasileira de Ciencia do Solo*, v. 40, p. 1–8, 2016. ISSN 18069657.

[5] STOCKMANN, U. et al. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems and Environment*, v. 164, p. 80–99, jan 2013. ISSN 01678809. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0167880912003635>.

[6] AXIMOFF, I. A.; ALVES, R. G.; RODRIGUES, R. d. C. Campos de Altitude do Itatiaia: Aspectos Ambientais, Biológico e Ecológicos. *Boletim do Parque Nacional do Itatiaia Nº 18*, p. 74, 2014.

[7] KÖPPEN, W. P. *Climatologia: con un estudio de los climas de la tierra.* Primera ed. México: [s.n.], 1948. 496 p. Disponível em: <https://issuu.com/lucaspestana/docs/koeppen{_}climatolo>.

[8] IBGE. Metadados do produto Modelo Digital de Elevação 1:25.000 do Projeto RJ-25. *Instituto Brasileiro De Geografia E Estatística.*, v. 55, n. 21, p. 2, 2008.

[9] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2018. Disponível em: <https://www.r-project.org/>.

[10] BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001. ISSN 08856125.

[11] BREIMAN, L. et al. *randomForest: Breiman and Cutler's Random Forests for Classification and Regression*. [S.l.], 2018. Disponível em: <https://cran.r-project.org/package=randomForest>.