

Carbon and nitrogen assessment in soils under unburned sugarcane using NIRS-LS-SVM models

Sandra Maria Oliveira SÁ¹, Christian FELLER², Marco Flores FERRÃO³, Marcelo Valadares GALDOS¹, Carla Maris Machado BITTAR⁴, Carlos Clemente CERRI¹, Roney Jesus POPPI⁵

¹ Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, 13400-970, Piracicaba SP, Brazil.

² Institut de Recherche pour le Développement, BP 434, 101 Antananarivo, Madagascar.

³ Departamento de Química e Física, Universidade de Santa Cruz do Sul, 96815-900, Santa Cruz do Sul RS, Brazil.

⁴ Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 13418-900, Piracicaba SP, Brazil.

⁵ Instituto de Química, Universidade Estadual de Campinas, 13083-970, Campinas, SP, Brazil.

Corresponding authors: smolivei@cena.usp.br and christian.feller@ird.fr

1. Introduction

Near infrared reflectance spectroscopy (NIRS) is a rapid, non-destructive, and non-polluting technology. This technique mainly measures overtones and combination bands of fundamental vibrations of O-H, N-H and C-H bonds relatives to several functional groups present in the organic compounds [1]. The use of NIRS for assessment of carbon and nitrogen, as well as, other soil properties and for evaluating changes of soil C and N stocks through time due to adoption of different management practices has been gaining attention in the scientific community. The aim of this work was to propose the use of NIRS associated with least square support vector machine (LS-SVM) as a methodology to quantify total carbon and nitrogen concentrations in soil samples under unburned sugarcane. Also, a comparative study between LS-SVM and PLS regression is presented.

2. Experimental

A total of 250 soil samples in 10 cm increments up to 100 cm depth were collected in an area with a long term history of unburned sugarcane in Brazil, at the São Martinho mill, in Pradópolis, São Paulo State. The soil in the area is classified as a clayey Oxisol. A chronosequence of plots where sugarcane had been cropped for 2, 4, 6 and 8 years without pre-harvest burning (235 samples) was studied. An adjacent area of native forest was sampled as a reference (15 samples). Reference analyses were performed by dry combustion on a LECO CN 2000 elemental analyzer. NIRS spectra were recorded on a NIRS 5000 scanning monochromator (Foss NIRSystems, MD). Samples were scanned in a spinning micro sample cup and the spectra were recorded at 2-nm intervals in the range of 1100 - 2498 nm by using WinISI II version 1.05 software (Infrasoft International, Silver Spring, MD) for data acquisition. A ceramic standard was used for the background spectra and the spectra was collected as $\log(1/R)$,

where R is the reflectance. For NIR calibration, two multivariate regression methods were used: (i) PLS using the PLS program from PLS-Toolbox version 3.5 with Matlab from Eigenvector Research Inc.[2]; and (ii) LS-SVM using the LS-SVMlab (Matlab/C Toolbox for Least Squares Support Vector Machines)[3].

3. Results and discussion

Multiplicative scatter correction (MSC) was performed to correct the baseline drift between the spectra. The three most important principal components PC₁, PC₂ and PC₃ explained 99.90% of the data variance. Table 1 presents the results for LS-SVM and PLS models for total-C and total N determination, respectively. In the range of 0.401-3.101 %, for total-C, and 0.030-0.252 %, for total-N suitable models were developed using LS-SVM compared with PLS regression, with low prediction errors, indicated by root mean square errors of cross-validation (RMSECV) and prediction (RMSEP). All two methods can give satisfactory prediction results both for training and prediction sets, and the most accurate model was obtained by the LS-SVM approach through the comparison of performance.

Table 1. Performance results comparison between PLS and LS-SVM for total-C and total-N quantification.

| | total-C | | | total-N | |
|-----------------|---------|---------|--------|---------|--------|
| | PLS(6) | PLS(11) | LS-SVM | PLS | LS-SVM |
| R^2_{cal} | 0.7842 | 0.9379 | 0.9772 | 0.8775 | 0.9748 |
| RMSEV (%) | 0.1498 | 0.1388 | 0.1139 | 0.0099 | 0.0078 |
| RMSEP (%) | 0.1459 | 0.1438 | 0.1036 | 0.0098 | 0.0082 |
| LV ^a | 6 | 11 | - | 9 | - |
| γ | - | - | 2350 | - | 1300 |
| σ^2 | - | - | 1200 | - | 1600 |

4. Conclusions

The diffuse reflectance spectroscopy in the near infrared region with data treatment by LS-SVM can be used with advantages as an analytical method for rapid, accurate, reliable and cost-effective routine analysis of soil total carbon and nitrogen. The total carbon stocks, after correction for density differences, were also higher in the areas with eight years of green cane (unburned) management.

5. References

- Murray, In: A. Davies, R.D. Baker, S.A. Grant, A.S. Laidlaw (Eds), Forage analysis by near infrared spectroscopy, Sward Herbage Measurement Handbook, Vol.14, British Grassland Society, Reading, UK, 1993, pp.285-312.
- B. M. Wise, N. B. Gallagher, R. Bro, J. M. Shaver, W. Windig, R. S. Koch, PLS Toolbox 3.5 for use with MATLAB, Eigenvector Research Inc., Manson, WA, 2005.
- J.A.K. Suykens, T. van Gestel, J. de Brabanter, B. de Moor and J. Vandewalle, Least-Squares Support Vector Machines. World Scientific, Singapore, 2002.

26

Avril 2008

**TERRE
MALGACHE**



**SPECIAL
SEMIS DIRECT**

**TANY
MALAGASY**



MACARTHUR

The John D. and Catherine T. MacArthur Foundation

**UNIVERSITE D'ANTANANARIVO
ECOLE SUPERIEURE DES SCIENCES AGRONOMIQUES**