

# Estimation of SOC Content in Anthropogenic Soils from Flysch Deposits Using Vis-NIR Spectroscopy

---

Boško MILOŠ<sup>1</sup>

Aleksandra BENSA<sup>2</sup> (✉)

## Summary

---

The objectives of this work were to estimate the ability of Vis-NIR diffuse reflectance spectroscopy for the prediction of soil organic carbon (SOC) content in terraced soils from Flysch deposits in Dalmatia, Croatia and to determine the significance of spectral wavelengths and regions. In a total of 159 top-soil samples (0-25 cm) SOC content was determined in the laboratory and reflectance spectra were collected using a portable Terra Spec 4 Hi-Res Mineral Spectrometer with a wavelength range 350-2500 nm. The partial last square regression (PLSR) with leave-one-out cross-validation method was used for calibrating the Vis-NIR spectra and SOC content. The SOC content varied from 2.79 to 28.66 g kg<sup>-1</sup> with an average value of 13.47 g kg<sup>-1</sup>. The SOC model prediction parameters, the coefficient of determination (R<sup>2</sup>), the ratio of performance to deviation (RPD) and the range error ratio (RER) were 0.73, 1.76 and 8.19, respectively, indicating moderately useful calibration model which is acceptable for a rapid sample screening for SOC content determination. The wavelengths located near 1400, 1900, 2000, 2025, 2200, 2225, 2275, 2325 and 2355 nm in NIR, near 560 nm in visible and near 825 nm in the short-wave NIR (700-1100 nm) were identified as important wavelengths for PLSR SOC modelling.

## Key words

---

calibration, diffuse reflectance spectroscopy, PLSR, soil organic carbon

<sup>1</sup> Institute for Adriatic Crops and Karst Reclamation, Put Duilova 11, 21000 Split, Croatia

<sup>2</sup> University of Zagreb, Faculty of Agriculture, Svetošimunska 25, 10000 Zagreb, Croatia

✉ e-mail: [abensa@agr.hr](mailto:abensa@agr.hr)

Received: October 16, 2017 · Accepted: November 29, 2017

## Introduction

Soil organic carbon (SOC) is crucial for sustaining soil fertility since it affects physical, chemical and biological soil properties. The need of rapid and accurate information on SOC content has long been recognized. During the last two decades many researchers have demonstrated that Vis-NIR diffuse reflectance spectroscopy (DRS) is capable of providing low cost, fast and reliable tool for prediction of chemical and physical soil properties compared to laboratory analyzes (Ben-Dor and Banin, 1995; Chang and Laird., 2002; Sankey et al., 2008; Ben-Dor et al., 2009; Viscarra Rossel et al. 2006; 2010; Stevens et al. 2013). Soil organic carbon (SOC) is a soil property that has been the most often quantified by Vis-NIR DRS with substantial differences in prediction accuracy as summarized in the review articles by Viscarra Rossel et al. (2006), and Stenberg et al. (2010). There are several possible explanations for highly variable performance of SOC prediction such as: quality (Ben Dor 1995; Brown et al. 2005) and quantity of organic matter (Montgomery and Baumgardner, 1974), heterogeneity of soil types, and parent material (Vasquez et al. 2010; Wijewardane et al. 2016).

The SOC significantly influences the reflectance characteristics of a soil and has spectral activity over a whole Vis-NIR spectral range (Ben-Dor et al., 2009). The soil spectrum characterizes complex absorption patterns with a large number of predictor variables that are highly collinear, and therefore analyses of diffuse reflectance spectra require the use of multivariate calibrations (Martens and Naes, 1989). The most common calibration method for analyses of SOC (and soil) spectra is partial last square regression (PLSR), developed by Wold et al. (2001). It raises a question whether PLS models are based directly upon the absorptions of organic matter (OM), or they are built indirectly on OM correlations with other soil constituents i.e. carbonates, Fe oxides and/or clay minerals. Anthropogenic soils from Flysch deposits have a low SOC content, very high carbonate contents and variable mineral matrix. Given that, the question remains of usability of Vis-NIR with PLSR for SOC prediction.

The objectives of this work were to (i) estimate the ability of Vis-NIR diffuse reflectance spectroscopy in the combination with PLSR for the prediction of SOC content in surface horizon of terraced soils derived from Flysch deposits; (ii) to determine the significance of spectral wavelengths and regions for SOC predicting, and (iii) whether this technique is reliable enough as an alternative for traditional laboratory analysis.

## Material and methods

### Study area and soil data

The study area is located in the middle part of Adriatic coastal area of Croatia near the city of Split, centred on 43°32' N; 16°29' E. This area has a Mediterranean climate characterized by hot summers and mild, moderately rainy winters classified as Csa. The mean annual air temperature of Split for the period between 1981 and 2010 was 15.9 °C, and the mean annual precipitation for the same period was 1052 mm. This area was built of Eocene Flysch marls, sandstones and siltstones and it is characterized with a high proportion and wide range of carbonate component (Marinčić et al. 1971; Marinčić et al. 1976). Soils of the area have very low to medium humus content, silty loam texture and alkaline reaction (Miloš and Maleš, 1998). According to the World Reference Base for Soil Resources (IUSS Working Group WRB, 2014), soils used for the analysis were classified as Terric Anthrosols (Calcaric, Siltic,

Escalic). Agricultural production is characterized by small, mixed and dislocated parcels of the olive groves, vineyards, Mediterranean orchards and abandoned terraced soil. For PLSR predictions, laboratory and spectral measurement of organic carbon (OC) were used in total of 159 top-soil samples selected from a Soil spectral library of Dalmatia, Croatia described by Miloš (2013). Soil sampling was carried out using the random point sampling method throughout an entire area. The SOC content was determined using the Kottzman method (JDPZ 1966).

### Spectra measurements, data pre-processing and selection of the optimal PLSR models

The spectra measurements of air-dried and sieved (2 mm) soil samples were obtained in a laboratory using a portable TerraSpec 4 Hi-Res Mineral Spectrometer (Analytical Spectral Devices Inc., Boulder, Colorado, USA) with a wavelength range of 350-2500 nm and recorded output on a 1 nm interval. The PLSR model was optimized by spectral data pre-processing treatments that included (i) a wavelength reduction to 5 nm for the whole region of 400–2500 nm using Savitzky-Golay smoothing algorithm and (ii) first-order derivative algorithm with a second order polynomial fit (Savitzky and Golay, 1964). The reflectance spectra were reduced to 400–2450 nm to eliminate the noise at spectral edges. The PLS regression with leave-one-out cross-validation method by Martens and Naes (1989) and Wold et al. (2001) was used for calibrating the spectra and soil organic carbon.

### Model performance evaluation

The performance of the PLSR model was evaluated using next parameters: the root mean square error of prediction (RMSEP), the standard error of prediction (SEP), the Bias, the ratio of performance to deviation (RPD), the range error ratio (RER) and the coefficient of determination ( $R^2$ ). The RMSEP, SEP, RPD and RER were calculated based on the following equations:

$$\text{RMSEP} = \sqrt{\sum_{i=1}^N \frac{(\hat{y}_i - y_i)^2}{N}} \quad (1)$$

$$\text{SEP} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i - \text{Bias})^2} \quad (2)$$

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i) \quad (3)$$

$$\text{RPD} = \frac{\text{SDv}}{\text{SEP}} \quad (4)$$

$$\text{RER} = \frac{\text{Max-Min}}{\text{SEP}} \quad (5)$$

where:

- $\hat{y}_i$  and  $y_i$  are the observed and predicted values of sample  $i$ ,
- $N$ , number of samples,
- SDv, standard deviation of the validation dataset
- Min and Max, minimum and maximum values of the validation set

The PLSR model parameters ( $R^2$ , RPD and RER) were used to classify the success of the predictions according to the guidelines proposed by Malley et al. (2004). Excellent calibrations were those

with  $R^2 > 0.95$ ,  $RPD > 4$ , and  $RER > 20$ . Successful calibrations had  $0.90 \leq R^2 \leq 0.95$ ,  $3 \leq RPD \leq 4$ , and  $15 \leq RER \leq 20$ . Moderately successful calibration had  $0.8 \leq R^2 < 0.90$ ,  $2.25 \leq RPD < 3$ , and  $10 \leq RER < 15$ ; and moderately useful ones had  $0.7 \leq R^2 < 0.8$ ,  $1.75 \leq RPD < 2.25$ , and  $8 \leq RER < 10$ . Hence, calibration models with  $R^2 < 0.7$ ,  $RPD < 1.75$  and  $RER < 8$  were considered less reliable.

## Results and discussion

### Soil and spectral properties

Table 1 shows the descriptive statistics of organic carbon (SOC) content analysed using conventional laboratory method analysis (reference dataset) and their calibrated and cross-validated PLSR predictions of the 159 soil samples. The SOC content varies from 2.79 to 28.66 g kg<sup>-1</sup> with an average value of 13.47 g C kg<sup>-1</sup>, indicating soils poor in organic carbon content. The positively skewed distribution of SOC dataset (skew=0.47) indicates a slightly asymmetrical distribution with a long tail to the right (Table 1).

Figure 1 shows the raw soil spectra (Fig. 1a) and first-derivative equivalent (Fig. 1b) for 159 soil samples in this study. The raw soil spectra (Fig. 1a) are characterized by monotone growth with

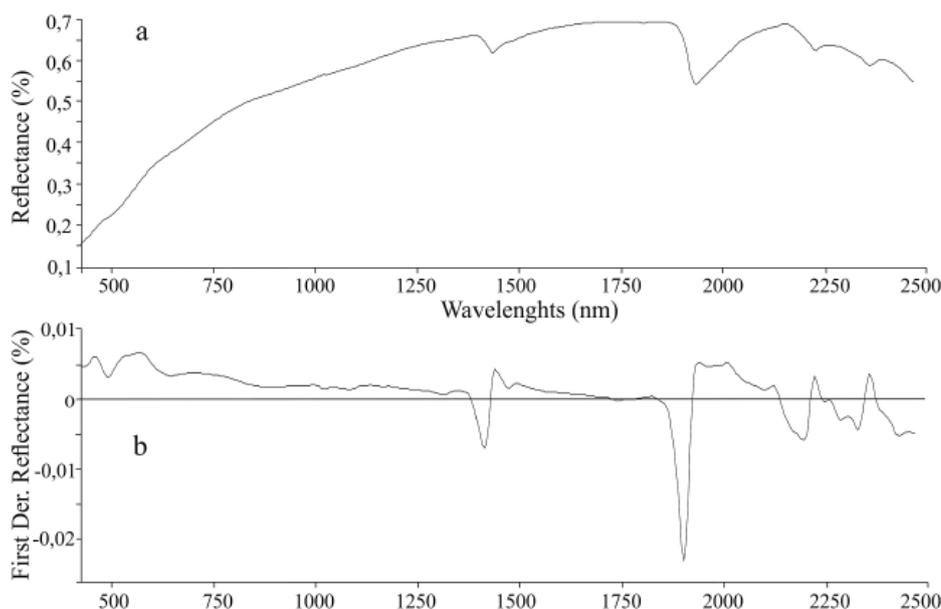


Figure 1. The mean raw spectra (a) and first derivative equivalent (b)

increasing wavelength in the visible range (400–700 nm) and without sharp peaks. The NIR spectrum is characterized by adsorption near 1400 nm due to the first overtone of the O-H stretch and the prominent adsorption peak near 1900 nm that may be related to H-O-H bend with the O-H stretches. Two small adsorption peaks between 2000–2400 nm are linked to the characteristics of organic matter and clay minerals.

The first derivative spectra (Fig. 1b) shows two adsorption peaks at 435 and 535 nm in the visible region that can be associated with Fe minerals (e.g. haematite, goethite; Sherman and Waite, 1985). The NIR region shows strong water and OH<sup>-</sup> absorptions in the NIR near 1400 and 1900 nm and a few other prominent absorptions peaks between 2200–2450 nm. These absorptions indicate the presence and the combined effect of secondary clay minerals (Clark 1999; Viscarra Rossel et al. 2006). Fig. 1b also shows characteristic carbonate band with an absorption peak of calcite at 2335 nm.

### Performance of calibration and validation model

Table 2 shows the calibration and cross-validation results of the PLS regression models for the SOC content. The most commonly used parameters for evaluation of prediction accuracy of the model are  $R^2$ , RPD and RER. The SOC prediction model with  $R^2$  0.73, RPD 1.76 and RER 8.19 can be considered as moderately useful according to the thresholds given by Malley et al. (2004).

The standard error of prediction (SEP value) was 2.58 g C kg<sup>-1</sup> (Table 2). The results of future predictions can then be defined by equation:

Reference data = Predicted values  $\pm 2 \cdot \text{SEP}$  (for confidence limit of 95%). Therefore, for this confidence limit, there is a 95% chance that the mean value of the SOC predicted model lies between 8.31 and 18.63 g C kg<sup>-1</sup>.

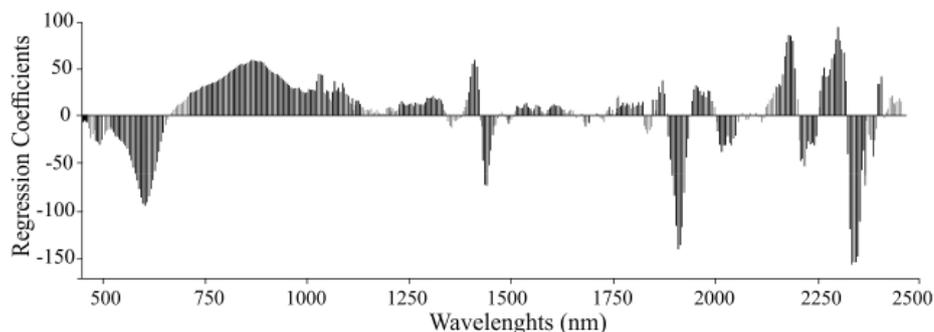
The precision obtained for SOC prediction in our study was similar to those reported in studies by Islam et al. (2003), Canasveras et al. (2012) and Gomez et al. (2013) with  $R^2$  ranging from 0.73 to 0.77 and RPD ranging from 1.7 to 1.9. The less accurate prediction of SOC using PLSR prediction model was established

Table 1. Summary statistics of SOC (g C kg<sup>-1</sup>) for the whole set, calibration and validation sets

Data-set	Mean	Maximum	Minimum	Range	Std. Dev.	Skewness
Whole	13.47	28.66	2.79	25.87	4.84	0.47
Calibration	13.32	25.17	2.59	22.58	4.62	0.20
Validation	13.31	23.25	2.11	21.14	4.54	0.16

Table 2. Calibration and validation results of soil organic content (g C kg<sup>-1</sup>) model diagnostics

Calibration				Validation			
RMSEC	R <sup>2</sup>	RMSEP	Bias	SEP	R <sup>2</sup>	RPD	RER
2.38	0.78	2.57	0.01	2.58	0.73	1.76	8.19



**Figure 2.** Regression coefficients of the wavelengths in the SOC model. The wavelengths retained as a significant ( $p < 0.01$ ) are marked in black

by Kuang and Mouazen (2011), Summers et al. (2011), Stevens et al. (2013) and Vasquez et al. (2010) with lower  $R^2$  and RPD values in ranges of 0.12 - 0.65 and 1.07 - 1.8, respectively. A higher level of accuracy in prediction of SOC was obtained in studies of Chang and Laird (2002), Sankey et al. (2008) and Kuang and Mouazen (2011) with  $R^2$  ranging from 0.89 to 0.96 and RPD ranging from 3.2 to 4.95.

The presented results show that the predictive quality of SOC calibrations is highly variable. It could be caused by the quality and quantity of organic matter, the presence of a chromophore that 'masked' absorptions of organic matter, characteristics of soil types, variable mineral matrix, texture, pre-processing and calibration methods.

#### Importance of the wavelengths and spectral regions

The plot of regression coefficients (Figure 2) illustrates how much each wavelength contributes the SOC prediction model and shows that significant wavelengths were located throughout the spectrum.

The most significant correlations in NIR region were found near 1400, 1900, 2000, 2025, 2200, 2225, 2275, 2325 and 2355 nm indicating the variability mainly in the NIR region. These adsorptions can be attributed to water, clay minerals and organic matter (Viscarra Rossel & Behrens, 2010). The VIS region characterizes a strong adsorption peak around 560 nm. The short-wave NIR (700-1100 nm) spectral region shows a wide range between 765 and 875 nm with a peak near 825 nm. The absorption feature in the visible and short-wave NIR region can be attributed to the Fe oxides in soil (Viscarra Rossel and Behrens, 2010).

#### Conclusions

This study showed that:

- the investigated soils are poor in organic carbon content, with a mean value of 13.47 g C kg<sup>-1</sup>
- the prediction parameters ( $R^2$ , RPD and RER) indicated moderately useful calibration model which is acceptable for a rapid sample screening for SOC content determination
- the wavelengths located throughout the spectrum mainly near 1400, 1900, 2000, 2025, 2200, 2225, 2275, 2325 and 2355 nm in NIR, but also near 560 nm in visible (Vis) and near 825 nm in the short-wave NIR (700-1100 nm) were identified as important wavelengths for PLSR SOC modelling
- this technique is not reliable enough as an alternative for traditional laboratory analysis but it could be used for a rapid and rough sample screening for SOC content.

#### References

- Analytical Spectral Devices Inc. (2003). QualitySpec® Pro manual. ASD Document, 600510, Rev. 1. ASD, Boulder, CO.
- Ben-Dor, E., Banin, A. (1995). Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal* 59: 364-372
- Ben-Dor E., Chabrillat S., Demattè J.A.M., Taylor G.R., Hill J., Whiting M.L., Sommer S. (2009). Using Imaging Spectroscopy to study soil properties. *Remote Sensing of Environment* 113: 38-55
- Brown D.J., Bricklemeyer R.S., Miller P.R. (2005). Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma*, 129 (3-4): 251-267
- Cañasveras J.C., Barrón V., del Campillo M. C., Viscarra Rossel R. A. (2012). Reflectance spectroscopy: a tool for predicting soil properties related to the incidence of Fe chlorosis. *Spanish Journal of Agricultural Research* 10 (4): 1133-114.
- Chang, C.W., Laird, D.A. (2002). Near infrared reflectance spectroscopy analysis of soil C and N. *Soil Science* 167(2): 110-116
- Clark, R.N. (1999). Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy. In: N. Rencz (Editor), *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*, 3 ed. John Wiley & Sons, New York, 3-52
- Kuang, B., Mouazen, A.M. (2011). Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *European Journal of Soil Science* 62: 629-636
- Gomez C., Le Bissonnais Y., Annabi M., Bahri H., Raclot D. (2013). Laboratory Vis-NIR spectroscopy as an alternative method for estimating the soil aggregate stability indexes of Mediterranean soils, *Geoderma* 209-210: 86-97
- Islam, K., Singh, B. and McBratney, A. (2003). Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Aust. J. Soil Res.* 41: 1101-1114
- IUSS Working Group WRB, (2014). World Reference Base for Soil Resources 2014. *World Soil Resources Reports No. 106*, FAO, Rome
- JDPZ (1966). *Chemical methods for soil analysis*, Beograd
- Malley, D. F., Martin, P. D. and Ben-Dor, E. (2004). Application in analysis of soils. Pages 729-784 in C. A. Roberts, J. Workman Jr., and J. B. Reeves III, eds. *Near-infrared spectroscopy in agriculture*. *Agronomy* 44. ASA, CSSA, SSSA, Madison, WI.
- Marinčić S., Magaš N., Borović I. (1971). *Osnovna geološka karta 1:100000 List Split K33-21*, Savezni geološki zavod, Beograd
- Marinčić S., Korolija B., Majcen Ž. (1976). *Osnovna geološka karta M 1: 100 000, list Omiš K 33-22*, Savezni geološki zavod, Beograd
- Martens H., Næs T. (1989). *Multivariate Calibration*, John Wiley & Sons, Inc, New York
- Miloš B., Maleš P. (1998). Tla Kaštelanskog zaljeva i problemi njihove zaštite. *Agronomski Glasnik* 4: 185-204
- Miloš B. (2013). *Spectral library of soils from Dalmatia*. Institute for Adriatic Crops and Karst Reclamation. Split, Croatia

- Montgomery O. L., Baumgardner M. F. (1974). The Effects of the Physical and Chemical Properties of Soils on the Spectral Reflectance of Soils. LARS Technical Reports. Paper 134.
- Savitzky, A., Golay M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36:1627–1638
- Sherman, D.M., Waite, T.D. (1985). Electronic spectra of Fe<sup>3+</sup> oxides and oxyhydroxides in the near infrared to ultraviolet. *American Mineralogist* 70: 1262–1269
- Stenberg B., Viscara- Rossel R.A., Mouazen A.M., Wetterlind J. (2010). Visible and Near Infrared Spectroscopy in Soil Science. In Donald L. Sparks, editor: *Advances in Agronomy*, Vol. 107, Burlington: Academic Press, pp. 163-215
- Stevens A, Nocita M, Tóth G, Montanarella L, van Wesemael B. (2013). Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE* 8(6): e66409
- Sankey J.B., Brown D.J., Bernard M.L., Rick L. Lawrence R.L. (2008). Comparing local vs. global visible and near-infrared (Vis NIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma* 148: 149–158
- Summers, D., Lewis, M., Ostendorf, B., Chittleborough, D. (2011). Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecological Indicators*, 11:123-131
- Vasques G.M., Grunwald S., Comerford N.B., Sicman J.O. (2010). Regional modelling of soil carbon at multiple depths within a subtropical watershed. *Geoderma*, 156: 326-336
- Wijewardane N., Ge Y., Wills S., Loecke T. (2016). Prediction of soil carbon in the conterminous united states: visible and near-infrared reflectance spectroscopy analysis of the rapid carbon assessment project, *Soil Science Society America Journal* 80: 973-982
- Viscarra Rossel R.A., Walvoort D.J.J., McBratney A.B., Janik L.J., Skjemstad J.O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131 (1–2): 59-75
- Viscarra Rossel R.A., Behrens T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra, *Geoderma* 158: 46–54
- Wold, S., Sjöström, M., Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory* 58, 109-130

---

acs83\_23