

Prediction of soil organic carbon using VIS-NIR spectroscopy: Application to Red Mediterranean soils from Croatia

Boško Miloš^a, Aleksandra Bensa^{b,*}

^a Institute for Adriatic Crops and Karst Reclamation, Put Duilova 11, 21 000 Split, Croatia

^b University of Zagreb, Faculty of Agriculture, Soil Science Department, Svetošimunska 25, 10 000 Zagreb, Croatia

Abstract

The objectives of this research were: (i) to assess the accuracy of diffuse reflectance spectroscopy (DRS) in predicting the soil organic carbon (SOC) content, and (ii) determine the importance of wavelength ranges and specific wavelengths in the SOC prediction model. The reflectance spectra of a total of 424 topsoils (0-25 cm) samples were measured in a laboratory using a portable Terra Spec 4 Hi-Res Mineral Spectrometer with a wavelength range 350-2500 nm. Partial least squares regression (PLSR) with leave-one-out cross validation was used to develop calibration models for SOC prediction. The accuracy of the estimate determined by the coefficient of determination (R^2), the concordance correlation coefficient (ρ_c), the ratio of performance to deviation (RPD), the range error ratio (RER) and the root mean square error (RMSE) values of 0.83, 0.90, 2.22, 14.2 and 2.47 g C kg⁻¹ respectively, indicated good model for SOC prediction. The near infrared (NIR) and the short-wave infrared (SWIR) spectrums were more accurate than those in the visible (VIS) and short-wave near-infrared (SWNIR) spectral regions. The wavelengths contributing most to the prediction of SOC were at: 1925, 1915, 2170, 2315, 1875, 2260, 1910, 2380, 435, 1960, 2200, 1050, 1420, 1425 and 500 nm. This study has shown that VIS-NIR reflectance spectroscopy can be used as a rapid method for determining organic carbon content in the Red Mediterranean soils that can be sufficient for a rough screening.

Keywords: Chemometrics, PLSR, Spectral regions, wavelengths.

© 2017 Federation of Eurasian Soil Science Societies. All rights reserved

Article Info

Received : 09.03.2017

Accepted : 24.05.2017

Introduction

Soil organic carbon content (SOC) is a fundamental constituent of the soil. It governs physical, chemical and biological processes in soils. The decline of SOC content in soils is one of the main threats for soil degradation. There is an increasing interest for large amounts of the accurate SOC data to be used in protection and enhancement of the global environment, soil quality assessment and precision agriculture.

In last few decades diffuse reflectance spectroscopy (DRS) due to its cost-efficiency, ease of handling, rapidity, minimal sample preparation and the development of chemometrics is being considered as a possible alternative to the conventional soil laboratory analyses. The visible-near-infrared (VIS-NIR) spectroscopy provides a large number of information on the organic and inorganic soil components. Absorption in the visible range provides a measure of soil colour, organic matter (Ben-Dor et al., 1999) and Fe minerals, mainly haematite and goethite (Sherman and Waite, 1985). The near-infrared (NIR) portions of the electromagnetic spectrum are associated with the stretching and bending of NH, OH and CH groups (Dalal and Henry, 1986; Clark, 1999; Viscarra Rossel and Behrens, 2010).

* Corresponding author.

University of Zagreb, Faculty of Agriculture, Soil Science Department, Svetošimunska 25, 10 000 Zagreb, Croatia

Tel.: +385 1 239-3802

e-ISSN: 2147-4249

E-mail address: abensa@agr.hr

DOI: [10.18393/ejss.319208](https://doi.org/10.18393/ejss.319208)

Published data pointed out that different spectral ranges and wavelengths can be responsible for the prediction of the SOC content (Brown et al., 2006; Viscarra Rossel et al., 2006a; Viscarra Rossel and Behrens, 2010). Sudduth and Hummel (1991) reported that near-infrared (NIR) reflectance data provided considerably better predictions of soil organic matter content than visible (VIS). Islam et al. (2003) evaluated the ability of reflectance spectroscopy in the UV, VIS, and NIR ranges to predict several soil properties. They demonstrated that the overall prediction of SOC was better with the whole spectral range (VIS-NIR) than NIR or VIS. Several studies have examined the question of the spectral ranges most suitable for predicting SOC concentration, considering the VIS, VIS-NIR, NIR and mid-infrared (MIR) (Viscarra Rossel et al., 2006b). They indicated that the MIR was more suitable than the NIR or VIS, but also noted that the MIR spectroscopy is more expensive and complex than VIS-NIR measurement.

The contribution of each wavelength of the VIS-NIR reflectance spectra for the SOC prediction was the subject of numerous studies (Ben-Dor and Banin, 1995; Dalal and Henry, 1986; Chang and Laird, 2002; Lee et al., 2009; Stenberg et al., 2010; Sarkhot et al., 2011; Xu et al. 2016). They have identified different wavelengths as important for estimation of SOC.

Numerous studies used VIS-NIR spectroscopy to analyse SOC content on the global (Brown et al., 2006; Viscarra Rossel et al., 2016), continental (Stevens et al., 2013), national (Vasquez et al., 2010; Knadel et al., 2012; Shi et al., 2015; Wijewardane et al., 2016), regional (Ben-Dor and Banin, 1995; Islam et al. 2003; Lee et al., 2009; Summers et al., 2011; Sarkhot et al., 2011; Leone et al., 2012; Shi et al., 2015) and local/field scales (Kuang and Mouazen, 2012; Wetterlind et al., 2010; Fontán et al., 2011). The most studies were conducted on the heterogeneous sets of soil samples with respect its geographical origin and varying of soil forming factors (Climate, Organisms, Topography, Parent material and Time). The exposed literature sources showed that the estimates of the SOC have been validated with a substantial range of the accuracy. The possible factors of the relatively large differences in the accuracy of the SOC estimation related with a number of factors (high SOC variability, the heterogeneity of soil types and parent material, land uses and size of the sampling area). Numerous studies have shown that geographically closer and more homogeneous sample sets with similar characteristics resulted in the improved SOC prediction models (Vasques et al., 2010, Wijewardane et al., 2016).

In this context, we decided to explore the accuracy of VIS-NIR spectroscopy for prediction of SOC content using soil samples that are similar with respect to soil types, parent materials and land use. We have chosen Red Mediterranean soils from Dalmatia, Croatia on limestones and dolomites used in agricultural production. The objectives of this work were (i) to assess the accuracy of the diffuse reflectance spectroscopy and PLSR for the SOC measurement and (ii) to determine the importance of different VIS-NIR spectral ranges and the wavelengths for the estimating SOC content.

Material and Methods

Study area and soil data

In our analyses, we used the SOC content data and soil spectra for a total of 424 topsoil samples selected from a Spectral library of soils from Dalmatia (Miloš, 2013). Dalmatia occupies the middle part of the Adriatic coastal zone of Croatia. The SOC content was determined using the Kottman method (JDPZ, 1966). Selected data belong to soils with a characteristic reddish-brown to red colour what is commonly referred to as a Red Mediterranean soil or Terra Rossa on hard limestone and dolomites. According to the Croatian soil classification system, these soils belong to the class of Cambic soils. The World Reference Base for soil resources (FAO, 2014) equivalent is Chromic Cambisols and Rhodic Cambisols. The climate of the study area is Mediterranean, with dry, hot summers and mild, moderately rainy winters. Agricultural areas of this typical karst environment consists a large number of dislocated smaller areas and parcels of olive groves, vineyards, fruit orchards and abandoned agricultural land.

Spectra measurements and data pre-processing

The reflectance spectra of soil samples were measured in a laboratory using a portable Terra Spec 4 Hi-Res Mineral Spectrometer with a wavelength range 350-2500 nm. Soil samples were air-dried and sieved through 2 mm sieve. The correction with a standardised white Spectralon® panel (Analytical Spectral Devices, Boulder, CO, USA) with 100% reflectance was made prior to the first scan and after every ten samples. All spectra were recorded at 1 nm data spacing interval over the wavelength range of 400-2500 nm. Pre-processing method: first derivative with Savitzky-Golay smoothing (Savitzky and Golay, 1964), with a second order polynomial fit. The spectral range of the soil spectra was first reduced to 400 - 2490 nm to

eliminate the noise at the edges of each spectrum. Spectral shortening was followed by averaging of the adjacent 5 nm wavelength. Thus, the total number of wavelengths for VIS-NIR modelling were 419 wavelengths.

Model Calibration and Validation

We used the partial least squares regressions (PLSR) with leave-one-out cross-validation method for calibrating the spectral data with the reference data and SOC model construction. For more details on the PLSR see e.g. [Martens and Naes \(1989\)](#) and [Wold et al. \(2001\)](#). Leave-one-out cross-validation method ([Efron and Tibshirani, 1994](#)) we used to determine the optimal number of the factors to retain in the calibration model. The regression coefficients for each of the 419 wavelengths in the range of 400 to 2490 nm were calculated. The best SOC model was generated when the PLS analyses were run using only significant wavelengths ($p < 0.05$) identified with Marten's uncertainty test.

Model Performance Evaluation

Predictive performances of the SOC model were evaluated by calculating the root mean square error (RMSEP) of prediction, the ratio of performance to deviation (RPD), range error ratio (RER) the coefficient of determination (R^2) and the concordance correlation coefficient (ρ_c). RMSEP is defined as the square root of the average of squared differences between predicted and measured Y values of the validation objects and evaluated by Equation (1):

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

where y_i and \hat{y}_i are the measured and predicted values of sample i , respectively, and N is the number of samples.

The RPD was initially used by [Williams \(1987\)](#) for assessing the goodness of fit for NIR calibrations. The RPD represents the division between the reference data standard deviation and the standard error of the prediction and deduced with Equation (2):

$$\text{RPD} = \text{SD} / \text{SEP} \quad \text{Eq. (2)}$$

where SD is the standard deviation of the validation dataset and SEP (Equation 3), which is standard error of prediction. SEP is the RMSEP corrected for bias (Equation 3). Bias is the average value of the difference between predicted and measured values (Equation 4).

$$\text{SEP} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i - \text{Bias})^2} \quad \text{Eq. (3)}$$

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i) \quad \text{Eq. (4)}$$

To interpret the RPD values we adopted the six level interpretations given by [Viscarra Rossel et al. \(2006b\)](#) as follows: RPD < 1.0 indicate very poor model/predictions and their use is not recommended; RPD between 1.0 and 1.4 indicate poor model/predictions where only high and low values are distinguishable; RPD between 1.4 and 1.8 indicate fair model/predictions which may be used for assessment and correlation; RPD values between 1.8 and 2.0 indicate good model/predictions where quantitative predictions are possible; RPD between 2.0 and 2.5 indicate very good, quantitative model/predictions and RPD > 2.5 indicate excellent model/predictions.

Range Error Ratio (RER, Equation 5) is the ratio of the range of the reference data to the SE in the validation set ([Starr et al., 1981](#)).

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

In Equation 5, Max and Min are the maximum and the minimum values observed in the reference data. The RPD and RER are dimensionless statistics, meaning they can be compared on the same basis between different models. The RER values we interpreted according to the thresholds given by Malley et al. (2004) as follows: $RER > 20$ indicate excellent prediction model; $15 \leq RER \leq 20$ successful; $10 \leq RER < 15$ moderately successful and $8 \leq RER < 10$ indicate moderately useful prediction model.

To interpret the predictive performances of the SOC models calculated by the coefficient of determination (R^2) we adopted threshold values given by Saeys et al. (2005) as follows: a value for R^2 between 0.66 and 0.80 indicates approximate quantitative predictions, R^2 between 0.81 and 0.90 reveals good prediction and $R^2 > 0.90$ is considered to be an excellent prediction. The concordance correlation coefficient (ρ_c) was proposed by Lin (1989) for assessment of concordance (agreement) between two measures of the continuous data. Like a correlation, Lin's concordance correlation coefficient (ρ_c) ranges from -1 to 1, where a value of 1 denotes perfect agreement; values > 0.90 suggest excellent agreement; values between 0.80 and 0.90 substantial agreement; between 0.65 and 0.80 moderate agreement and values < 0.65 poor agreement.

Results and Discussion

The soil organic carbon statistics and soil spectral properties

The descriptive statistics of the soil organic carbon content analysed using conventional laboratory method analysis (reference dataset), and their calibration and cross-validated PLSR predictions for the four spectral regions: VIS (400-700 nm), NIR (700-2490 nm), SWNIR (700-1100 nm), SWIR (1100-2490 nm) and combined VIS-NIR (400-2490 nm) is given in Table 1. The SOC content for the reference dataset varies from 0.81 to 37.87 g C kg⁻¹ with an average value of 21.35 g C kg⁻¹ (Table 1).

Table 1. Statistical description of the soil organic carbon content (g kg⁻¹) for the reference, calibration and validation datasets for different spectral ranges

Statistics	Reference	Calibration		Validation			
		VIS-NIR	VIS-NIR	NIR	VIS	SWIR	SWNIR
Mean	21.35	21.35	21.36	21.37	21.36	21.37	21.36
Minimum	0.81	0.36	0.25	0.32	0.05	0.25	0.02
Maximum	37.87	35.35	35.28	33.84	31.11	33.51	29.04
Range	37.06	34.99	35.03	33.52	31.06	33.25	28.92
Std.Dev.	5.96	5.52	5.47	5.19	4.37	5.12	4.12
Skewness	-0.55	-0.79	-0.81	-0.84	-1.19	-0.86	-1.51

VIS-NIR – visible and near infrared range (400-2490 nm); NIR – near infrared range (700-2490 nm); VIS – visible range (400-700 nm); SWIR - shortwave infrared range (1100-2490 nm); SWNIR – shortwave near infrared range (700-1100 nm)

The negatively skewed distribution for the reference, calibration and VIS-NIR, NIR and SWIR validation dataset indicate a slightly asymmetrical distribution with a long tail to the left. The skewness for the validation VIS and SWNIR spectral regions are less than -1.0 and indicate substantial skewness and the distribution is far from symmetrical.

Figure 1 and 2 show mean reflectance spectra and mean first-derivative reflectance spectra of the soil samples that were used to develop PLSR calibration model. The shape of the overall reflectance spectra (Figure 1) shows a typical soil spectrum characterised with reflectance increasing with increasing wavelength in the visible range (400–700 nm) and does not contain distinct or sharp peaks that can be directly associated with specific constituents. In the visible range, the mean first-derivative of the SOC reflectance spectra (Figure 2) shows two obvious adsorption bands near 430 and 565 nm which can be attributed to the presence of the chromophorous constituents, mainly iron oxide (Ben-Dor et al., 1999). The mean reflectance spectra (Figure 1), as well as mean first-derivative spectra (Figure 2), show a weak concave shape at the wavelengths around 800-970 nm. This concavity can be attributed mainly to increased iron oxide content since the soil samples contain a low content organic matter which can mask the effect of Fe (Dematté et al., 2004).

Furthermore, soil mean reflectance and first-derivative spectra (Figure 1 and Figure 2) show several important absorptions around 1400, 1900, and 2200 nm. In addition to this, first-derivative spectra shows strong adsorption around 2400 nm (Figure 2). These absorptions can be attributed to water molecule vibrations and OH- groups (1400 nm), water (1900 nm) and mineral influences (2200 nm and 2400 nm).

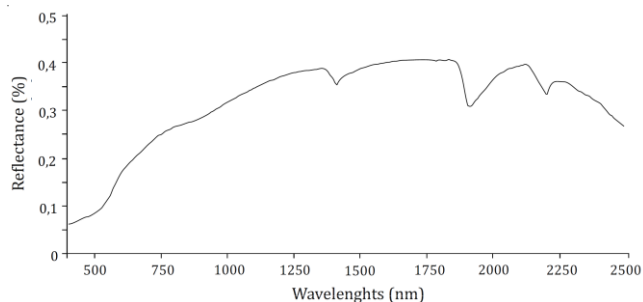


Figure 1. The mean reflectance spectra for 424 soil samples

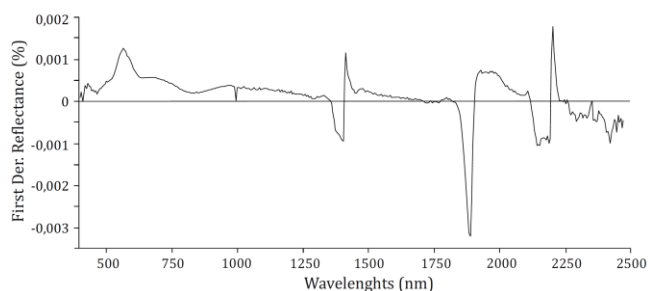


Figure 2. The mean first-derivative spectra for 424 soil samples

Performance assessment of calibration and validation models

Statistical description of the PLSR calibrated and their cross-validated PLSR SOC predictions for different spectral ranges is given in Table 2. The best SOC prediction model for the each spectral range was generated when the PLSR analyses included five factors. The PLSR SOC model with the lowest RMSEP and highest R^2 , ρ_c , RPD and RER is considered as the best model. The best prediction for SOC was obtained from the VIS-NIR spectra with validation R^2 , ρ_c , RPD, RER and RMSEP values of 0.83, 0.90, 2.22, 14.2 and 2.47 g C kg⁻¹ respectively.

Table 2. Calibration and validation results of the SOC models diagnostics for the different spectral ranges and combined VIS-NIR spectral range

Spectral range	Calibration				Validation				
	RMSEC	R^2	RMSEP	Bias	SEP	R^2	ρ_c	RPD	RER
VIS-NIR	2.18	0.88	2.47	-0.01	2.46	0.83	0.90	2.22	14.2
NIR	2.70	0.84	2.99	0.02	3.01	0.75	0.86	1.72	14.1
SWIR	2.96	0.75	3.24	-0.01	3.23	0.70	0.83	1.35	9.6
VIS	4.11	0.52	4.53	-0.01	4.52	0.42	0.61	1.13	7.4
SWNIR	4.23	0.49	4.50	0.01	4.51	0.46	0.61	0.91	6.4

RMSEC - root mean square error of calibration; R^2 - coefficient of determination; RMSEP - root mean square error of prediction; SEP - standard error of prediction, ρ_c - concordance correlation coefficient; RPD - ratio of performance to deviation; RER - range error ratio

According to threshold values for the coefficient of determination (R^2) given by [Saeyns et al. \(2005\)](#) R^2 values of the SOC validation models for the spectral ranges (Table 2) can be interpreted as follows: the SOC model developed using the VIS-NIR spectral range reveals good prediction; NIR and SWIR region indicate approximate quantitative predictions until the calibrations for VIS and SWNIR region are only able to discriminate between high and low SOC values.

Lin's ρ_c values of the SOC validation models for the VIS-NIR range (Table 2) suggests good agreement; the NIR and SWIR region denote substantial agreement and VIS and SWNIR region suggest a poor agreement between reference data and the cross-validated SOC values.

According to the stated interpretation given by [Viscarra Rossel et al. \(2006b\)](#), follows that the VIS-NIR spectral range with the RPD value of 2.22 (Table 2) indicates very good quantitative model for the SOC prediction; NIR indicates fair model; SWIR and VIS regions with the RPD values between 1.0 and 1.4 indicate poor prediction model and SWNIR regions with values of the RPD <1.0 and indicates very poor SOC prediction model.

The RER values for different spectral ranges (Table 2) we interpreted according to the thresholds given by [Malley et al. \(2004\)](#) as follows: the SOC model created using VIS-NIR and NIR range ($10 \leq RER \leq 15$) indicate moderately successful prediction; SWIR spectral range with $8 \leq RER < 10$ indicates moderately useful prediction model and model created using the VIS and SWNIR region $RER < 8$ indicate poor prediction.

The RMSEP is the most efficient measure of the average uncertainty in spectral predictions that can be expected for future predicting new samples. The future predictions of the samples in the test set can be considered that 2 times RMSEP represents a 95% confidence interval for the real values. Since RMSEP values for the SOC predicted models (Table 2) ranged from 2.47 g C kg⁻¹ (VIS-NIR) to 4.53 g C kg⁻¹ (VIS) with its

mean value of 21.36 g C kg⁻¹ (Table 1), then there is a 95% chance that the mean value of the SOC content, as measured in laboratory, lies between 16.42 to 26.30 g C kg⁻¹ and 12.30 to 30.42 g C kg⁻¹ respectively.

The predictive performance of the SOC prediction model in this study (Table 2) is a similar accuracy as those obtained by [Gras et al. \(2014\)](#), [Wijewardane et al. \(2016\)](#) and [Knadel et al. \(2012\)](#) who achieved similar R² values of 0.82, 0.83 and 0.81 and RPD of 2.40, 2.41 and 2.40, respectively. Some authors ([Brown et al., 2006](#); [Chang and Laird, 2002](#) and [Viscarra Rossel et al., 2016](#)) reported the higher accuracy of SOC prediction models than our with R² values of 0.87, 0.89 and 0.89, respectively. [Leone et al. \(2012\)](#) obtained a high accuracy (R² 0.84-0.93 and 2.36-2.53 RPD) for local SOC predictive models in the soils of southern Italy. Study of [Kuang and Mouazen \(2012\)](#) at the farm-scale in four different European countries showed a very wide range of R² and RPD, 0.12-0.96 and 1.07-4.95, respectively. [Wetterlind et al. \(2010\)](#) reported a very good prediction of SOM (R² = 0.89, RMSE = 4.70 g kg⁻¹ and RPD= 3.0) achieved for a farm-scale calibration model using only 25 soil samples. However, numerous studies achieved a less predictive capacity of SOC models than our e.g. [Gao et al. \(2014\)](#), [Shi et al. \(2015\)](#), [Summers et al. \(2011\)](#) and [Fontán et al. \(2011\)](#). In the mentioned studies R² ranged from 0.55 to 0.79 and RPD from 1.80 to 2.01. The exposed large differences in the accuracy of the SOC estimates related with a number of factors (high SOC variability, the heterogeneity of soil types and parent material, land uses and size of the sampling area) and it cannot be easily identified.

In general, all parameters of cross-validated predictions diagnostic (R², Lin's ρ_c , RPD, RER and RMSP) given in Table 2, showed that the SOC prediction model created using combined VIS-NIR spectra provides the most accurate predictions. So, the NIR and the short-wave infrared (SWIR) spectrum predictions were more accurate than those in the VIS and short-wave near-infrared (SWNIR) spectral region. Exposed is in accordance with previous findings, e.g. [Sudduth and Hummel \(1991\)](#) who reported that NIR reflectance data provided considerably better predictions of soil organic matter content than those in the visible range. [Islam et al. \(2003\)](#) demonstrated that the overall prediction of SOC was better with the whole VIS-NIR spectral range than in the NIR.

Identification of the important wavelengths for the SOC prediction

The contribution of each wavelength of the VIS-NIR spectral range for the SOC prediction model, marked with PLSR regression coefficient, is shown in Figure 3. The best SOC prediction model included only those wavelengths for which Marten's test of uncertainty showed that significantly contribute ($p < 0.01$) to the prediction. Figure 3 illustrates that the final SOC model retained a very large number of wavelengths (total number of selected wavelengths were 176) and that the significant wavelengths with large regression coefficients were observed throughout the spectrum.

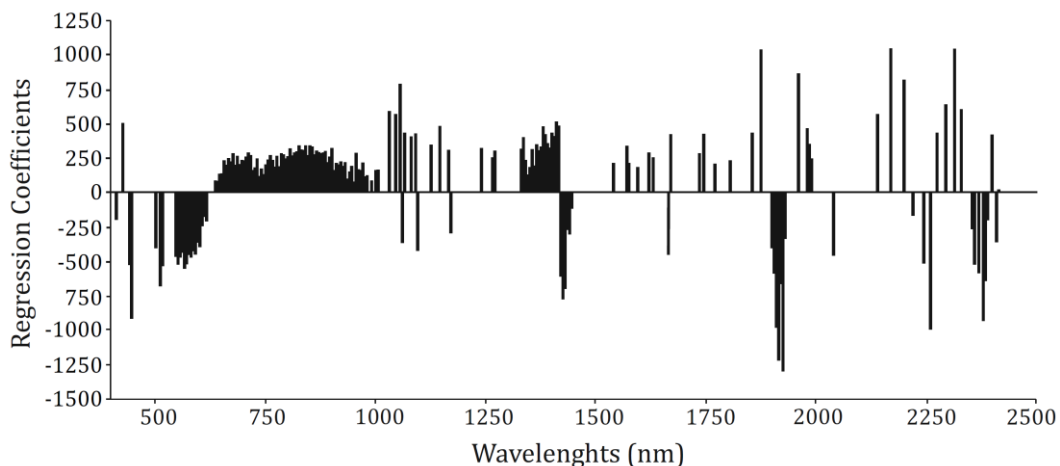


Figure 3. Regression coefficients of the wavelengths retained as a significant ($p < 0.01$) in the final a fifth-component SOC prediction model for VIS-NIR range

In the visible range the significant wavelengths with the maximum regression coefficients of the SOC prediction model were observed as individual peaks at 430 and 435 nm and a broad spectral range from 500 to 590 nm, with maximum peaks at 500 and 555 nm (Figure 3). In the SWNIR (700-1100 nm) spectral region there is noticeable a broad region around 800 nm with a maximum contribution at 815, 830 and 840 nm. Wavelengths in these spectral regions associated with the chromophorous constituents - iron oxides - mainly haematite and goethite ([Sherman and Waite, 1985](#)) and organic matter ([Ben-Dor et al., 1999](#)). The

highest correlation coefficients were observed in the short-wave infrared (SWIR, 1100-2500 nm) region (Figure 3). This spectral region is characterised with featured absorption bands around 1400 (max. peak at 1420 nm), 1900 (max. peak at 1925 nm), 2200 and 2300 nm due to OH⁻ and water (1400 nm and 1900 nm) and mineral influences (2200 and 2300 nm). These absorption bands associated also with the overtones and combination absorptions of C-H, O-H, S-H, C=O, and N-H (Dalal and Henry, 1986; Clark, 1999; Viscarra Rossel and Behrens, 2010) and consequently contain information about the concentration of the organic composition of the scanned soil sample. The high contributing wavelengths of the water adsorption features are consistent with research findings by Viscarra Rossel et al. (2006a) and Sarkhot et al. (2011). The wavelengths contributing most to the SOC prediction model are found around 2200 nm (max. peak at 2170 nm) and around 2300 nm (max. peaks at 2315 and 2380) due to Al-OH bend plus O-H stretch combinations, that are diagnostic absorption features in clay mineral identification (Clark et al., 1990). Many other authors (Ben-Dor and Banin, 1995; Dalal and Henry, 1986; Stenberg et al. (2010) and Viscarra Rossel and Behrens (2010) have also identified the bands around 2200 and 2300 nm as being important for SOC calibration.

Ordering 15 wavelengths with the highest regression coefficients in the final a fifth-component SOC model for VIS-NIR range were: 1925, 1915, 2170, 2315, 1875, 2260, 1910, 2380, 435, 1960, 2200, 1050, 1420, 1425 and 500 nm. This is consistent with other research findings. For example, Lee et al. (2009) reported 450–465, 965, 1409, and 1775–2200 nm as important wavelengths for soil organic carbon prediction. Sarkhot et al. (2011) observed the significant wavelengths throughout the spectrum, while the magnitude of PLSR coefficients was higher in the region 400, 1400, 1900 and 2200 or 2300 nm. Chang and Laird (2002) reported that the wavelengths important for organic carbon were in the range between 1770 nm to 2500 nm. This study showed that twelve of the fifteen most significant wavelengths were greater than 1100 nm. The advantage of the wavelengths above 1100 nm is that they are uncorrelated with Fe- oxides that reduced the collinearity problem with organic matter. This is consistent with previous research (Ben-Dor and Banin, 1995; Dalal and Henry, 1986; Stenberg et al., 2010; Xu et al., 2016) which have identified wavelengths around 1100, 1600, 1700 to 1800, 2000, and 2200 to 2400 nm as being particularly important for SOC calibration.

As conclusions, (i) all parameters of cross-validated predictions diagnostic R², Lin's ρ_c , RPD, RER and RMSEP with values of 0.83, 0.90, 2.22, 14.2 and 2.47 g C kg⁻¹ respectively indicated good model for predicting of the organic carbon in Red Mediterranean soil, (ii) the SOC prediction model created using combined VIS-NIR spectra provided the most accurate SOC predictions, (iii) the NIR and the short-wave infrared (SWIR) spectrum predictions were more accurate than those in the VIS and short-wave near-infrared (SWNIR) spectral region and (iv) the wavelengths contributing most to the prediction of SOC content were at 1925, 1915, 2170, 2315, 1875, 2260, 1910, 2380, 435, 1960, 2200, 1050, 1420, 1425 and 500 nm. Taking into account the predictive statistics and accuracy of created model, it can be sufficient for rapid and a rough screening of the organic carbon in Red Mediterranean soils.

References

- 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(2): 364–372.
- Ben-Dor, E., Irons, J., Epema, G., 1999. Soil reflectance. In: Remote sensing for the earth sciences: Manual of remote sensing, 3rd Edition Vol. 3., Rencz, A.N. (Ed.), John Wilen Sons. Inc. New York, USA. pp. 111–188.
- Brown, D.J., Shepherd, K.D., Walsh, M.G., May, M.D., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 132(3-4): 273–290.
- 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: Remote Sensing for the Earth Sciences: Manual of Remote Sensing, 3rd Edition Vol. 3., Rencz, A.N. (Ed.), John Wilen Sons. Inc. New York, USA. pp. 3–58.
- Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research* 95(B8): 12653 – 12680.
- Dalal, R.C., Henry, R.J., 1986. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. *Soil Science Society of America Journal* 50(1): 120–123.
- Demattê, J.A.M., Campos, R.C., Alvesb, M.C., Fiorioa, P.R., Nanni, M.R., 2004. Visible-NIR reflectance: a new approach on soil evaluation. *Geoderma* 121(1-2): 95 – 112.
- Efron, B., Tibshirani, R.J., 1994. An introduction to the bootstrap. Monographs on Statistics and Applied Probability 57. CRC. Press, Boca Raton, Florida, USA. 436p.

- Fontán, J.M., López-Bellido, L., García-Olmo, J., López-Bellido, R.J., 2011. Soil carbon determination in Mediterranean vertisol by visible and near infrared reflectance spectroscopy. *Journal of Near Infrared Spectroscopy* 19(4): 253–263.
- Gao, Y., Cui, L., Lei, B., Zhai, Y., Shi, T., Wang, J., Chen, Y., He, H., Wu, G., 2014. Estimating soil organic carbon content with visible–near infrared (Vis–NIR) spectroscopy. *Applied spectroscopy* 68 (7): 712–722.
- Gras, J.P., Barthès, B.G., Mahaut, B., Trupin, S., 2014. Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma* 214–215: 126–134.
- Islam, K., Singh, B., McBratney, A., 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research* 41(6): 1101–1114.
- FAO, 2014. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106, Food and Agriculture Organization of The United Nations (FAO) Rome, Italy. 192p.
- JDPZ, 1966. Chemical methods for soil analysis, Beograd [in Croatian].
- Knadel, M., Deng, F., Thomsen, A., Greve, M.H., 2012. Development of a Danish national vis-NIR soil spectral library for soil organic carbon determination. Digital Soil Assessments and Beyond. In: Proceedings of the 5th Global Workshop on Digital Soil Mapping. Minasny, B., Malone, B.P., McBratney, A.B., (Eds.). 10-13 April 2012, Sydney, Australia. Pp. 403-408.
- Kuang, B., Mouazen, A.M., 2012. Influence of the number of samples on prediction error of visible and near infrared spectroscopy of selected soil properties at the farm scale. *European Journal of Soil Science* 63(3): 421–429.
- Lee, K.S., Lee, D.H., Sudduth, K.A., Chung, S.O., Kitchen, N.R., Drummond, S.T., 2009. Wavelength identification and diffuse reflectance estimation for surface and profile soil properties. *American Society of Agricultural and Biological Engineers* 52(3): 683–695.
- Leone, A.P., Viscarra-Rossel, R.A., Pietro Amenta, P., Buondonno, A., 2012. Prediction of soil properties with PLSR and vis-NIR Spectroscopy: Application to Mediterranean soils from Southern Italy. *Current Analytical Chemistry* 8(2): 283–299.
- Lin, L.I.K., 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45(1): 255–268.
- Malley, D.F., Martin, P.D., Ben-Dor, E., 2004. Application in analysis of soils. In: Near-infrared spectroscopy in agriculture. Roberts, C.A., Workman J., Reeves, J.B., (Eds.). Agronomy Vol 44. ASA-CSSA-SSSA, Madison, WI, USA. pp. 729–784.
- Martens, H., Naes, T., 1989. Multivariate Calibration. John Wiley & Sons Inc. New York, USA. 504 p.
- Miloš B. 2013. Spectral library of soils from Dalmatia. Institute for Adriatic Crops and Karst Reclamation. Split, Croatia.
- Saeys, W., Mouazen, A.M., Ramon, H., 2005. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosystems Engineering* 91(4): 393–402.
- Sarkhot, D.V., Grunwald, S., Ge, Y., Morgan, C.L.S., 2011. Comparison and detection of total and available soil carbon fractions using visible/near infrared diffuse reflectance spectroscopy. *Geoderma* 164(1-2): 22–32.
- Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36(8): 1627–1638.
- Sherman, D.M., Waite, T.D., 1985. Electronic spectra of Fe³⁺ oxides and oxyhydroxides in the near IR to UV. *American Mineralogist* 70: 1262–1269.
- Shi, Z., Ji, W., Viscarra Rossel, R.A., Chen, S., Zhou, Y., 2015. Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese vis – NIR spectral library. *European Journal of Soil Science* 66(4): 679–687.
- Starr, C., Morgan, A.G., Smith, D.B., 1981. An evaluation of near infra-red reflectance analysis in some plant breeding programmes. *Journal of Agricultural Science* 97(1): 107–118.
- Stenberg, B., Rossel, R.A.V., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. *Advances in Agronomy* 107: 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.
- Sudduth, K.A., Hummel, J.W., 1991. Evaluation of reflectance methods for soil organic matter sensing. *Transactions of the ASAE* 34(4): 1900–1909.
- Summers, D., Lewis, M., Ostendorf, B., Chittleborough, D., 2011. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecological Indicators* 11(1): 123–131.
- Vasques, G.M., Grunwald, S., Harris, W.G., 2010. Spectroscopic models of soil organic carbon in Florida, USA. *Journal Environmental Quality* 39(3): 923–934.
- Viscarra Rossel, R.A., McGlynn, R.N., McBratney, A.B., 2006a. Determining the composition of mineral-organic mixes using UV-VIS-NIR diffuse reflectance spectroscopy. *Geoderma* 137(1-2): 70–82.
- Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006b. 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(1-2): 46–54.

- Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D., Demattê, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aichi, H., Barthès, B.G., Bartholomeus, H.M., 2016. A global spectral library to characterize the world's soil. *Earth-Science Reviews* 155: 198-230.
- Wetterlind, J., Stenberg, B., Söderström, M., 2010. Increased sample point density in farm soil mapping by local calibration of visible and near infrared prediction models. *Geoderma* 156(3-4): 152-160.
- 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(4): 973-982.
- Williams, P.C., 1987. Variables affecting near-infrared reflectance spectroscopic analysis. In: Near-infrared technology in the agricultural and food industries. Williams, P., Norris, K., (Eds.). American Association of Cereal Chemists, Saint Paul, USA. pp. 143 -167.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58(2): 109-130.
- Xu, S., Shi, X., Wang, M., Zhao, Y., 2016. Effects of subsetting by parent materials on prediction of soil organic matter content in a hilly area using Vis-NIR spectroscopy. *PLoS ONE* 11(3): e0151536.