

Prediction of organic carbon and calcium carbonates in agricultural soils with Vis-NIR spectroscopy

Procjena organskog ugljika i kalcijevih karbonata u poljoprivrednim tlima Vis-NIR spektroskopijom

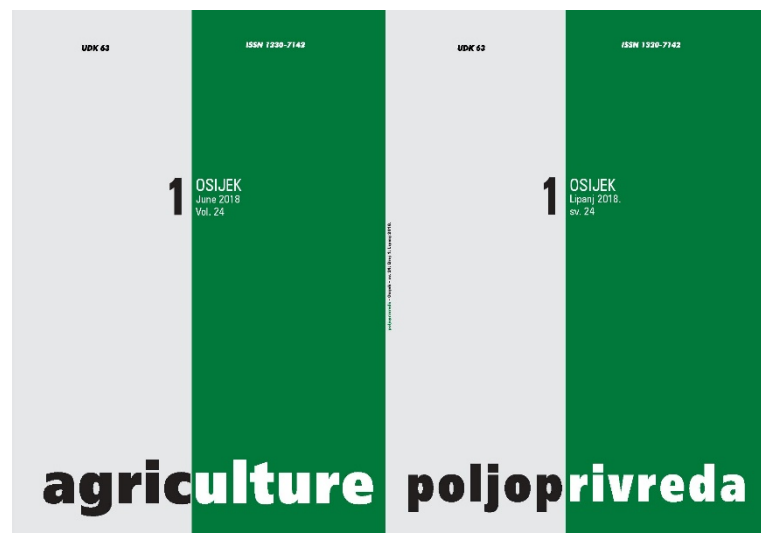
Miloš, B., Bensa, A.

Poljoprivreda/Agriculture

ISSN: 1848-8080 (Online)

ISSN: 1330-7142 (Print)

<http://dx.doi.org/10.18047/poljo.24.1.6>



Poljoprivredni fakultet u Osijeku, Poljoprivredni institut Osijek

Faculty of Agriculture in Osijek, Agricultural Institute Osijek

PREDICTION OF ORGANIC CARBON AND CALCIUM CARBONATES IN AGRICULTURAL SOILS WITH VIS-NIR SPECTROSCOPY

Miloš B.⁽¹⁾, Bensa A.⁽²⁾

Original scientific paper

Izvorni znanstveni članak

SUMMARY

The objectives of this study were to (i) evaluate the ability of Vis-NIR spectroscopy to predict soil organic carbon (SOC) and CaCO₃ content in the heterogeneous agricultural soils from Dalmatia, Croatia and (ii) to compare the performance of two multivariate calibration techniques: partial least square regression (PLSR) and support vector machine regression (SVMR). The reflectance spectra of a total of 250 top-soils (0-25 cm) samples were collected in the laboratory using a portable Terra Spec 4 Hi-Res Mineral Spectrometer with a wavelength range 350-2500 nm. The coefficient of determination (R²), the residual prediction deviation (RPD) and the root mean square error (RMSE) were used for the model evaluation. The CaCO₃ prediction derived by PLSR and SVMR with R² (0.86 and 0.88) and RPD (2.67 and 2.82), respectively are considered good prediction models. The SOC prediction with SVMR (R² 0.84 and RPD 2.43) indicates good prediction and approximate quantitative prediction with PLSR with R² of 0.78 and RPD of 1.94. Our results showed that (i) CaCO₃ and SOC estimations were obtained with acceptable accuracy using Vis-NIR spectroscopy, (ii) the SVMR method produced more accurate estimations of selected soil properties compared to PLSR, and (iii) Vis-NIR spectroscopy, in combination with SVMR can be recommended as a rapid and inexpensive method for screening of the CaCO₃ and SOC content.

Key-words: diffuse reflectance spectroscopy, PLSR, regression, SVMR

INTRODUCTION

Organic carbon and carbonates play a key role in many of the soil chemical and physical processes. These are basic soil properties widely used in soil research, precision agriculture, land and soil suitability assessment, environmental monitoring and modelling that require rapid and accurate detection. Diffuse reflectance spectroscopy (DRS) is a rapid, non-destructive and eco-friendly method, which offers a cheap solution in the estimation of different soil properties. The Vis-NIR spectra are difficult to interpret due to the (i) overlap of weak overtones and combinations of fundamental vibrational bands and (ii) very complex relation between the soil spectra reflectance and characteristic parameters of the soil samples such as SOC and CaCO₃ content. As a result, for quantitative analysis of soil properties,

multivariate calibration is required. Various methods have been used to relate soil spectra and soil properties. Among them, the partial least-squares regression (PLSR) developed by Wold et al. (2001) is a standard multivariate statistical technique and the most common method for spectral calibration and prediction of soil properties.

The PLSR can successfully model the linear relationship between spectral data and soil properties especially when multicollinearity exists in raw spectra data. However, nonlinearity between the spectra and soil properties often exists due to soil heterogeneity. Thus, a nonlinear calibration method as support vector machine

(1) Assist. Prof. Boško Miloš, Institute for Adriatic Crops and Karst Reclamation, Put Duilova 11, 21000 Split, Croatia, (2) Assoc. Prof. Aleksandra Bensa (abensa@agr.hr), University of Zagreb, Faculty of Agriculture, Soil Science Department, Svetošimunska 25, 10000 Zagreb, Croatia

regression (SVMR) can provide a more reasonable solution than linear methods. The SVMR is a kernel-based learning method from statistical learning theory (Vapnik, 1995). The SOC and CaCO_3 have a strong influence on soil reflectance characteristics. Soil organic matter has spectral activity over a wide spectral range (Ben-Dor et al. 1999). The carbonates have several diagnostic vibrational absorptions in the NIR spectral region. The strongest is at 2300-2350 nm and other three weaker bands occur near 2120-2160 nm, 1997-2000 nm and 1850-1870 nm (Clark 1999). Ben-Dor and Banin (1990) found the strong and sharp absorption features of the carbonate in soil at 1800 nm, 2350 nm, and 2360 nm and concluded that DRS provide a reliable tool for quantifying soil carbonate content. The SOC has been often related to reflectance data in Vis-NIR region acquired by spectrometers in laboratory conditions and its content was determined, generally, with a good accuracy, i.e. Viscarra Rossel et al. (2016) reported coefficient of determination (R^2) of 0.89 and residual prediction deviation (RPD) of 3.3 for global data set. Wijewardane et al. (2016) found the R^2 value of 0.83 and RPD of 2.4 for the SOC model calibrated with PLSR using nearly 20,000 soil samples collected across the United States. Leone et al. (2012) achieved R^2 in range 0.84 - 0.93 and RPD between 2.36 and 3.03 for regional and local predictions of SOC in Mediterranean soils from southern Italy. Carmon and Ben-Dor (2017) achieved excellent SOC prediction with R^2 values of 0.95 and RPD of 4.36. However, some studies reported the lower accuracy of SOC content estimation, i.e. Summers et al. (2011) obtained R^2 values of 0.57 and RPD 1.8. A wide ranges of R^2 values (0.66-0.96) and RPD between 1.55 and 4.20 were reported by Nduwamungu et al. (2009) based on 15 literature findings, indicating approximate quantitative to excellent SOC prediction using Vis-NIR spectroscopy.

Many authors (Volkan Bilgili et al., 2010; Summers et al., 2011; Canasveras et al., 2012; Gomez et al., 2012 and 2013; Leone et al., 2012; Gras et al., 2014; Carmon and Ben-Dor, 2017) have estimated CaCO_3 content using Vis-NIR with substantial differences in prediction accuracy. The predictive parameters, R^2 and RPD, varied in the mentioned studies in a wide range of 0.64-0.99 and 1.63-8.60, respectively. The wide range in SOC and CaCO_3 estimation accuracy is possibly related to the variability of analysed soil properties, heterogeneity of soils and parent material, land use, the number of samples, as well as differences in multivariate calibration techniques and data pre-processing. A comparison between SVMR and PLSR method for prediction of soil properties has been the subject of numerous researches. Most of them (Viscarra Rossel and Behrens, 2010; Gao et al., 2014; Peng et al., 2014; Luca et al., 2017) showed that SVMR produced better results than PLSR.

The objectives of this study were to (i) evaluate the ability of Vis-NIR spectroscopy to predict SOC and CaCO_3 content in the heterogeneous agricultural soils from Dalmatia, Croatia and (ii) to compare the performance of two multivariate calibration techniques: partial least

square regression (PLSR) and support vector machine regression (SVMR) for estimation soil properties.

MATERIAL AND METHODS

Study area and soil data

The research included 250 top-soil samples and laboratory data (SOC and CaCO_3 content) obtained for various scientific and technical projects in the area of Middle Dalmatia, centred around 43° 32' N; 16° 29' E. This area has a Mediterranean climate characterised by hot summer and mild, moderately rainy winters. The area is built of the Jurassic and Cretaceous limestones and dolomites, Tertiary marls, breccias, conglomerates, Flysch marls, sandstones and siltstones and Quaternary deposits. According to the World Reference Base for Soil Resources (IUSS Working Group WRB, 2014), we classified the investigated soils as Leptic, Rhodic, Chromic, Eutric Cambisols, Colluvic, Skeletic, Calcic Regosols and Terric Anthrosols. Current agriculture is adjusted to typical karst topography by the small, mixed and dislocated parcels of Mediterranean species (olive groves, vineyards, orchards), arable and abandoned agricultural land. The SOC content was determined by oxidation with solution of KMnO_4 using the Ketzman method (JDPZ, 1966) and the CaCO_3 content was analysed with the modified volumetric method (JDPZ, 1966).

Spectra measurements and pre-processing

The spectra measurements of air-dried and sieved (2mm) soil samples were obtained in the laboratory using a portable TerraSpec 4 Hi-Res Mineral Spectrometer with a wavelength range of 350-2500 nm. 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. Pre-processing treatment included (i) Savitzky-Golay smoothing algorithm (5 nm for the whole region 400–2500 nm) and (ii) first derivative with a second order polynomial fit (Savitzky and Golay, 1964). Furthermore, to eliminate the noise at the edges of each spectrum the spectral range of the soil spectra was reduced to 400-2490 nm range.

Selection of calibration model

The partial least squares regression (PLSR) and the support vector machine regression (SVMR) were used for calibrating the reflectance spectra with analysed soil properties. For more details on the PLSR see e.g. Martens and Næs (1989) and Wold et al. (2001). The validation for SOC and CaCO_3 calibration and validation model was performed using leave-one-out cross-validation method (Efron and Tibshirani, 1994), also known as full cross validation. This method used same samples both for model of calibration and validation. In the each iteration of this method, only one sample was left out and the model was constructed on the remaining data points. The process was repeated with another sample and so on until every sample was left out once. Leave-

one-out cross-validation was also used to determine the optimum number of factors to retain in the PLSR model of calibration and validation (prediction). The support vector machine regression (SVMR) is a supervised, nonparametric and statistical machine learning method for the calibration of nonlinear relationships, introduced by Vapnik (1995).

Model Performance Evaluation

The performance of the SOC and CaCO₃ prediction models established with the PLSR and the SVMR method were evaluated by parameters such as the root mean square error of prediction (RMSEP), the ratio of performance to deviation (RPD) and the correlation coefficient (R²). The RMSEP is a measure of the dispersion of the validation samples around the regression line and measures the average accuracy of prediction. The RMSEP is defined as the square root of the average of squared differences between predicted and measured Y values of the validation objects and computed by Equation (1),

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

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

The coefficient of determination (R²) was used as the measure of how close the data are to the fitted regression line. The R² value between 0.50 and 0.66 indicates only the ability to discriminate between high and low values; R² between 0.66 and 0.80 indicates approximate quantitative predictions; R² between 0.81 and 0.90 reveals good prediction and R² > 0.90 is considered to be an excellent prediction (Saeys et al., 2005). The RPD was initially used by Williams (1987). It represents the division between the standard deviation (SD) of the validation dataset and the standard error of the prediction (SEP) and deduced with Equation (2):

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

The SEP is the RMSEP corrected for bias (Eq. 3).

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

Bias is the average value of the difference between predicted and measured values (Equation 4).

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

To interpret RPD values we adopted the interpretations given by Chang et al. (2001) as follows: models with RPD ≥ 2 are considered "good" models, with 1.4 ≤ RPD < 2 they predict „fairly“, while models with RPD < 1.4 predict „poorly“.

RESULTS AND DISCUSSION

Descriptive Statistics

Table 1 shows the descriptive statistics of the carbonates (CaCO₃) and soil organic carbon (SOC) content analysed using conventional laboratory methods (reference dataset) and their calibrated and cross-validated PLSR and SVMR predictions of the 250 soil samples. The average CaCO₃ content in the reference dataset is 256.4 g kg⁻¹ and varies within a very wide range from 0.0 to 864.0 g kg⁻¹. It has a positively skewed distribution indicating a slightly asymmetrical distribution with a long tail to the right. The statistical parameters for the CaCO₃ content produced with the PLSR and SVMR calibration and validation methods (Table 1) are very similar to those described for the reference set. The SOC content of the reference dataset varies from 0.11 to 36.92 g kg⁻¹ with an average value of 17.53 g C kg⁻¹ indicating that soils with a low content of the organic carbon dominate. The skewness values for the SOC content showed approximately symmetrical distributions with a long tail to the left. All statistical parameters for the SOC content produced with the PLSR and SVMR calibration and validation methods are very similar to those described for the reference set.

Table 1. Statistical description of the CaCO₃ and soil organic carbon content for the reference, calibration and validation datasets for 250 soil samples

Tablica 1. Deskriptivna statistika za CaCO₃ i organski ugljik tla za referentni, kalibracijski i validacijski skup za 250 uzoraka tla

Statistics Statistika	CaCO ₃ g kg ⁻¹					SOC g kg ⁻¹				
	Reference Referentni	Calibration Kalibracija		Validation Validacija		Reference Referentni	Calibration Kalibracija		Validation Validacija	
		PLSR	SVMR	PLSR	SVMR		PLSR	SVMR	PLSR	SVMR
Mean	256.4	256.4	256.3	256.2	255.2	17.53	17.54	17.53	17.53	17.52
Max.	864.0	871.2	865.2	858.4	872.1	36.92	34.33	36.87	33.53	35.13
Min.	0.0	0.0	0.0	0.0	0.0	0.11	0.28	0.09	0.37	0.05
Range	864.0	871.2	865.2	858.4	872.1	36.81	34.05	35.21	33.16	35.12
St. Dev.	248.0	234.0	243.8	232.3	239.2	7.73	7.30	7.48	7.15	7.38
Skewness	0.69	0.80	0.80	0.81	0.80	-0.15	-0.26	-0.14	-0.39	-0.08

Interpretation of Vis-NIR spectra

Figure 1 shows mean soil reflectance spectra (Figure 1a) and first-derivative of 5 nm reflectance spectra (Figure 1b) of the 250 soil samples. The overall soil spectra (Figure 1a) is a typical soil spectrum characterised with an increasing pattern between 415 and 1350 nm with no distinct or sharp peaks and then oscillates between 1350 and 2500 nm. The mean reflectance curve in the NIR region shows strong water and OH⁻ absorptions near 1400 nm and 1900 nm (Dalal and Henry, 1986; Ben-Dor and Banin, 1995; Clark, 1999).

The first-derivative spectra in the Vis range (Figure 1b) shows adsorption peak around 465 nm which can be due to the presence of the chromophorous constituents, mainly Fe oxides and darkness of the organic constituents (Ben-Dor et al., 1999). Furthermore, in addition to strong water and OH⁻ absorptions around 1400 and 1900 nm NIR range shows prominent absorptions peaks around 2200 nm and 2300 nm indicating mineral influences (Dalal and Henry, 1986; Clark, 1999; Viscarra Rossel and Behrens, 2010). This region shows strong adsorption at 2335 nm (Figure 1b). It can be attributed to the calcite absorption (Clark 1999).

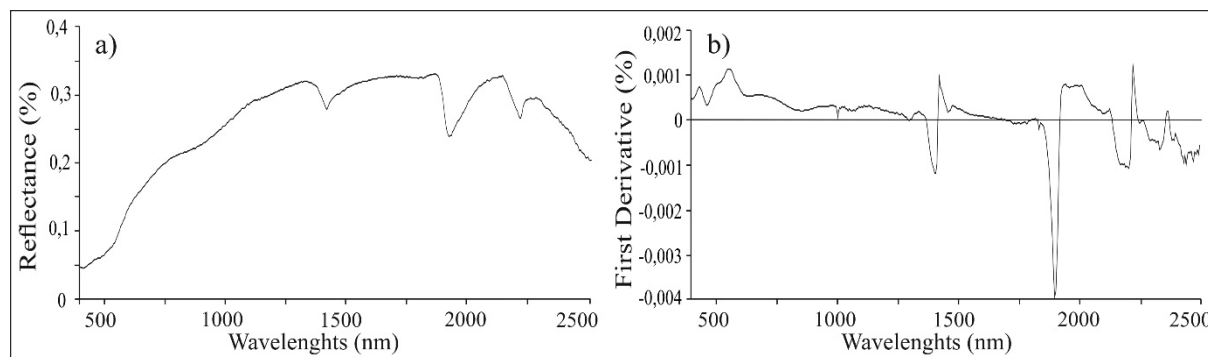


Figure 1. Mean original soil reflectance spectra (a) and first-derivative of 5 nm reflectance spectra (b)

Slika 1. Prosječni izvorni reflektantni otisak tla (a) i prva derivacija reflektantnog otiska od 5 nm (b)

Performance of Vis-NIR calibration and validation models

Table 2 shows the calibration and cross-validation results of the PLSR and SVMR methods for the CaCO₃ and SOC prediction. The model with the highest R² and RPD and the smallest RMSEP value is

considered as the best prediction model. The optimum number of factors to retain in the best PLSR CaCO₃ validation model was three. Explained variance in this validation model with the first factor was 69.4, with two factors 85.9 and with first three factors 88.1%. A high percentage of explained variance, as well as a

small number of factors included in the model, indicates a good prediction performance. The CaCO₃ prediction model developed with the PLSR method with validation R² and RPD values of 0.86 and 2.67 respectively, and the SVMR validation with R² and RPD values of 0.88 and 2.82 respectively (Table 2), can be considered good models according to the interpretation of R² values given by Saeys et al. (2005) and RPD values given by Chang et al. (2001).

The RMSEP values for the CaCO₃ estimates with the SVMR and PLSR were 84.65 and 86.94 g CaCO₃ kg⁻¹ respectively and can be considered acceptable.

Table 2. Prediction diagnostics of calibration and validation of the CaCO₃ and soil organic carbon models

Tablica 2. Prognozna dijagnostika za kalibraciju i validaciju modela CaCO₃ i organskog ugljika tla

Soil property <i>Svojstvo tla</i>	Method <i>Metoda</i>	Calibration <i>Kalibracija</i>		Validation <i>Validacija</i>				
		RMSEC	R ²	RMSEP	Bias	SEP	R ²	RPD
CaCO ₃ g kg ⁻¹	PLSR	73.11	0.91	86.94	0.40	86.98	0.86	2.67
	SVMR	68.32	0.94	84.32	0.33	84.65	0.88	2.82
SOC g kg ⁻¹	PLSR	2.87	0.85	3.68	0.01	3.69	0.78	1.94
	SVMR	2.20	0.92	3.05	-0.01	3.04	0.84	2.43

The CaCO₃ prediction model derived with the SVMR method showed slightly better prediction performance compared to the PLSR. Our results showed better prediction accuracy of the CaCO₃ content esti-

mation compared to the study of Volkan Bilgili et al. (2010), Summers et al. (2011) and Gomez et al. (2012 and 2013). The aforementioned authors obtained lower R² and RPD values of 0.64-0.76 and 1.63-2.10, respec-

tively. However, some authors obtained better validation parameters for prediction of the CaCO₃ content compared to ours. For example, Canasveras et al. (2012) achieved R² 0.93 and RPD 3.50, similar to Carmon and Ben-Dor (2017), who obtained R² 0.94 and RPD 3.49. The differences in the prediction accuracy can be attributed to the heterogeneity of sample sets (CaCO₃ content variability and the frequency distribution), a number of samples, as well as differences in data pre-processing and method of calibration.

The optimum number of factors to retain in the best PLSR SOC validation model was five. Percentages of explained variance in this validation model were 43.6, 65.8, 71.7, 76.2 and 78.9 for the first five selected factors in the best model. The SOC model developed with the PLSR method showed the R² value of 0.78 and RPD value of 1.94 (Table 2), which indicate an approximate quantitative predictions according to R² values interpretation given by Saeys et al (2005) and prediction of "fairly" according to RPD value interpretation given by Chang et al (2001). The prediction of the SOC content with the SVMR method resulted in higher R² and RPD values of 0.84 and 2.43 respectively (Table 2). These prediction parameters show that the SOC model derived with the SVMR can be considered a good prediction model according to the adopted threshold values for R² given by Saeys et al (2005) and for RPD values given by Chang et al. (2001). The RMSEP values for the SOC model predictions generated with the PLSR and SVMR methods were 3.68 and 3.05 g kg⁻¹, respectively (Table 2). These results confirm that the SVMR method gives more accurate estimates. This is consistent with the study of Gao et al. (2014) that achieved higher R² and RPD values for the SOC predictions generated with the SVMR compared to the PLSR method (0.86 and 2.60; 0.79 and 1.95, respectively). Viscarra Rossel and Behrens (2010) also achieved more accurate SOC estimates with the SVMR method compared to the PLSR. The predictive accuracy of our PLSR SOC model is consistent with numerous studies (Lee et al., 2009; Canasveras et al., 2012; Deng et al., 2013; Gras et al., 2014; Wijewardane et al., 2016) that achieved similar R² values in the range 0.77-0.83 and RPD between 1.9 and 2.4. Some studies (Summers et al., 2012; Gomez et al., 2013) have shown lower prediction performance for the SOC models, with R² values between 0.57 and 0.73 and RPD values of 1.80-1.93. Nevertheless, some studies obtained better accuracy of SOC prediction models than ours, e.g. Leone et al. (2012), with higher R² and RPD values of 0.84-0.93 and 2.36-3.03, respectively.

The results of this study showed that the CaCO₃ models had better predictive performances compared to the SOC models. This is consistent with the research of Canasveras et al. (2010 and 2012), Gomez et al. (2012), Gras et al. (2014) and Summers et al. (2013), who also achieved better accuracy of CaCO₃ content estimation than for SOC.

CONCLUSION

Our results show that: (i) CaCO₃ and SOC content can be estimated with acceptable accuracy using Vis-NIR spectroscopy; (ii) CaCO₃ estimates developed by PLSR and SVMR method achieved similar R² and RPD values (0.86 and 0.88; 2.67 and 2.82 respectively) that can be considered good prediction models; (iii) SOC model derived by SVMR (R² values of 0.84 and RPD of 2.43) indicates good prediction, while PLSR model with R² of 0.78 and RPD of 1.94 indicates only approximate quantitative prediction; (iv) CaCO₃ models have better predictive performances compared to SOC models and (v) the SVMR method produced more accurate estimations of selected soil properties compared to PLSR.

REFERENCES

1. 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.
2. Ben-Dor, E., Irons, J. R., & Epema, J. F. (1999). Soil reflectance: remote sensing for the Earth Science. In: *Manual of Remote Sensing*, Rencz, Andrew N.(Ed.), 3: 111-186.
3. 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. <https://doi.org/10.5424/sjar/2012104-681-11>
4. Carmon, N., & Ben-Dor, E. (2017). An advanced analytical approach for spectral-based modelling of soil properties. *International Journal of Emerging Technology and Advanced Engineering*, 7(3), 90-97. <https://doi.org/10.1002/9781119043553>
5. Chang, C. W., Laird, D. A., Mausbach, M. J., & Hurburgh, C. R. (2001). Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65(2), 480-490.
6. Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. In: *Manual of Remote Sensing*. Rencz, N. (Ed.), John Wiley & Sons, New York, 3-52.
7. 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.
8. Deng, F., Minasny, B., Knadel, M., McBratney, A., Heckrath, G., & Greve, M. H. (2013). Using VIS NIR spectroscopy for monitoring temporal changes in soil organic carbon. *Soil Science*, 178(8), 389-399.
9. Efron, B., & Tibshirani, R.J. (1994). *An introduction to the Bootstrap*. Champan & Hall CRC Press
10. 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*, 66(7), 712- 722. <https://doi.org/10.1366/13-07031>

11. Gomez, C., Lagacherie, P., & Coulouma, G. (2012). Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis–NIR data. *Geoderma*, 189, 176-185.
<https://doi.org/10.1016/j.geoderma.2012.05.023>
12. 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, 86-97. <https://doi.org/10.1016/j.geoderma.2013.06.002>
13. 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, 126-134.
<https://doi.org/10.1016/j.geoderma.2013.09.021>
14. IUSS Working Group WRB (2014): World Reference Base for Soil Resources 2014. World Soil Resources Reports No. 106, FAO, Rome.
15. JDPZ (1966): Kemijske metode ispitivanja zemljišta, Beograd.
16. 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.
17. 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. <https://doi.org/10.2174/157341112800392571>
18. Luca, F., Conforti, M., Castrignano, A., & Buttafuoco, G. (2017). Effect of calibration set size on prediction at local scale of soil carbon by Vis-NIR spectroscopy. *Geoderma*, 288, 175-183.
doi: <http://dx.doi.org/10.1016/j.geoderma.2016.11.015>
19. Martens, H., & Næs, T. (1989): Multivariate Calibration. John Wiley & Sons, Inc, New York 504 p.
20. Nduwamungu, C., Ziadi, N., Parent, L. E., Tremblay, G. F., & Thuries, L. (2009). Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review. *Canadian Journal of Soil Science*, 89(5), 531-541. <https://doi.org/10.4141/CJSS08076>
21. Peng, X., Shi, T., Song, A., Chen, Y., & Gao, W. (2014). Estimating Soil Organic Carbon Using VIS/NIR Spectroscopy with SVMR and SPA Methods. *Remote Sensing*, 6(4), 2699-2717.
<https://doi.org/10.3390/rs6042699>
22. Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627-1638. <https://doi.org/10.1021/ac60214a047>
23. 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), 23-131. <https://doi.org/10.1016/j.ecolind.2009.05.001>
24. Vapnik, V. N. (1995): The Nature of Statistical Learning Theory. Springer-Verlag.
25. Volkan Bilgili, A., van Es, H. M., Akbas, F., Durak, A., & Hively, W. D. (2010). Visible near-infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments*, 74(2), 229-238. <https://doi.org/10.1016/j.jaridenv.2009.08.011>
26. Viscarra Rossel, R. A., & Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158(1-2), 46-54. <https://doi.org/10.1016/j.geoderma.2009.12.025>
27. Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D., Dematte, J., Sheperd, K., et al. (2016). A global spectral library to characterise the world's soil. *Earth-Science Reviews*, 155, 198-230.
<https://doi.org/10.1016/j.earscirev.2016.01.012>
28. 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.
<https://doi.org/10.2136/sssaj2016.02.0052>
29. 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 Inc., Saint Paul, MN.143-167.
30. Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109-130.
[https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)

PROCJENA ORGANSKOG UGLJIKA I KALCIJEVIH KARBONATA U POLJOPRIVREDNIM TLIMA VIS-NIR SPEKTROSKOPIJOM

SAŽETAK

Ciljevi ove studije bili su: (i) procijeniti mogućnost Vis-NIR spektroskopije za procjenu sadržaja organskog ugljika tla (OC) i CaCO₃ u heterogenim poljoprivrednim tlima iz Dalmacije, Hrvatska, i (ii) usporediti učinkovitost dviju multivarijantnih kalibracijskih tehnika: partial least square regression (PLSR) i support vector machine regression (SVMR). Spektralni otisci ukupno 250 uzoraka tla (0-25 cm) uzeti su u laboratoriju korištenjem prijenosnoga Terra Spec 4 hi-res spektrometra, s rasponom valne dužine 350-2500 nm. Koeficijent determinacije (R²), rezidualno predikcijsko odstupanje (RPD) i korijen srednje kvadratne greške (RMSE) korišteni su za procjenu modela. Predviđanja CaCO₃ korištenjem PLSR-a uz R² od 0,86 i RPD 2,67 te SVMR-a uz R² od 0,88 i RPD 2,82 indiciraju dobre prognozne modele. Predviđanje OC-a korištenjem SVMR-a (R²=0,84 i RPD=2,43) ukazuje na dobru procjenu, a PLSR metoda (R²=0,78 i RPD=1,94) na približnu kvantitativnu procjenu. Naši su rezultati pokazali (i) prihvatljivu točnost procjene CaCO₃ i OC-a korištenjem Vis-NIR spektroskopije, (ii) da je SVMR metoda dala točnije procjene odabranih svojstava tla od PLSR-a i (iii) da se Vis-NIR spektroskopija, u kombinaciji sa SVMR-om može koristiti kao brza i jeftina metoda za određivanje CaCO₃ i sadržaja organskog ugljika u tlu.

Ključne riječi: difuzna reflektantna spektroskopija, PLSR, regresija, SVMR

(Received on 10 May 2017; accepted on 20 December 2017 - *Primljeno 10. svibnja 2017.; prihvaćeno 20. prosinca 2017.*)