

Using Visible-Near Infrared Spectroscopy to Predict Soil Properties of Mugan Plain, Azerbaijan

Fikrat Feyziyev¹, Maharram Babayev¹, Simone Priori², Giovanni L'Abate²

¹Institute of Soil Science and Agrochemistry of National Academy of Science of Azerbaijan, Baku, Azerbaijan

²CREA-Council for Agricultural Research and Economics, Florence, Italy

Email: fikrat.fm@gmail.com

Received 16 February 2016; accepted 28 March 2016; published 31 March 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The potential ability of the Vis-NIR (350 - 2500 nm) laboratory spectroscopy for prediction of soil properties has been demonstrated in the literature. The aim of this work was to predict different soil properties of soils collected in Mugan plain (Azerbaijan) using PLSR models and cross-validation. Carbonatation and salinisation are the main pedogenetic processes in Muganplain, therefore there is a need to monitor total carbonates and soil electrical conductivity (EC). The result of work was positive and both total carbonates and EC showed the best result (SE = 2.90, R²-0.90 and SE-0.09, R²-0.82, respectively). This parameter using PLSR, SOM (SE-0.54, R²-0.83), CEC (SE-0.43, R²-0.62), and Total P (SE-0.50, R²-0.73) is predicted and result is over optimist. Total N (SE-0.04, R²-0.44), and pH (SE-0.02, R²-0.51) demonstrated low prediction quality.

Keywords

PLSR, Prediction, VIS-NIR, Soil Reflectance, Soil Properties

1. Introduction

There is widespread interest for using Vis-NIR spectroscopy to predict soil properties due to cost-effective, rapid and minimal soil preparation process and no hazardous to environment [1]. The demographic increasing of world population requires that soil resources should be carefully used. It is important that, soil resources should be investigated using modern technology. Conventional soil analysis is not efficiently because they are very slow and expensive [2] with their ecological hazard. Visible (Vis, 400 - 700 nm) and Near-Infrared (NIR, 700 - 2500 nm) diffuse reflectance spectroscopy (DRS) has shown to be an efficient tool for the rapid and cheap pre-

How to cite this paper: Feyziyev, F., Babayev, M., Priori, S. and L'Abate, G. (2016) Using Visible-Near Infrared Spectroscopy to Predict Soil Properties of Mugan Plain, Azerbaijan. *Open Journal of Soil Science*, 6, 52-58.

<http://dx.doi.org/10.4236/ojss.2016.63006>

diction of soil properties [3]. The prediction of soil properties requires the creation of a spectral library relating spectra with reference data [4]. Such a library should be designed to represent the variation in soil properties of the soil types of interest. Multivariate regressions are then used to infer properties [5].

The most important thing to enhance the accuracy of Vis-NIR measurement of soil properties is the optimal selection of calibration model. Five multivariate techniques, namely, stepwise multiple linear regression, PCR, PLSR, regression tree and committee trees were compared by Vasques *et al.* 2008 [6] with the aim of identifying the best combination of multivariate statistics and spectra pre-processing to predict soil carbon. They concluded that PLSR performed the best compared to other techniques tested. Moreover, linear PCR and PLSR analyses are the most common techniques for spectral modelling [1] [7].

There are several works predicted soil properties using Vis-NIR spectroscopy by PLSR model. Soil organic carbon and soil organic matter prediction is highly variable in Vis-NIR region but often reported that OM signals are weak [2]. Different authors explain it with relating soil spatial variability and soil condition and according to Ben-Dor and Banin [8] suggested that the organic matter itself changes due to decomposition. There is support that soil spatial variation would affect to Vis-NIR prediction but in small range so field or farm-scale calibration is better than regional or coarse [9]. As an alternative Vis-NIR widely used and there are a lot of powerful results. NIR reflectance spectroscopy is sensitive to organic carbon and mineral soil composition which it possible to predict of various soil properties form single scan [15].

2. Materials and Methods

2.1. Study Area

Soil samples were collected on the Mugan plain of Azerbaijan. The study are has a arid climate with a mean annual precipitation, evaporation, temperature of 30 mm, 1000 mm, 17°C, respectively. Study area (**Figure 1**) is mainly located under sea level and the dominant soilsare Calcisols, Solonchaks and Calcacic Fluvisols. Salinization, gleyfication and carbonatation processes are very appreciable and plays significant role in these soils genesis and morphology [10].

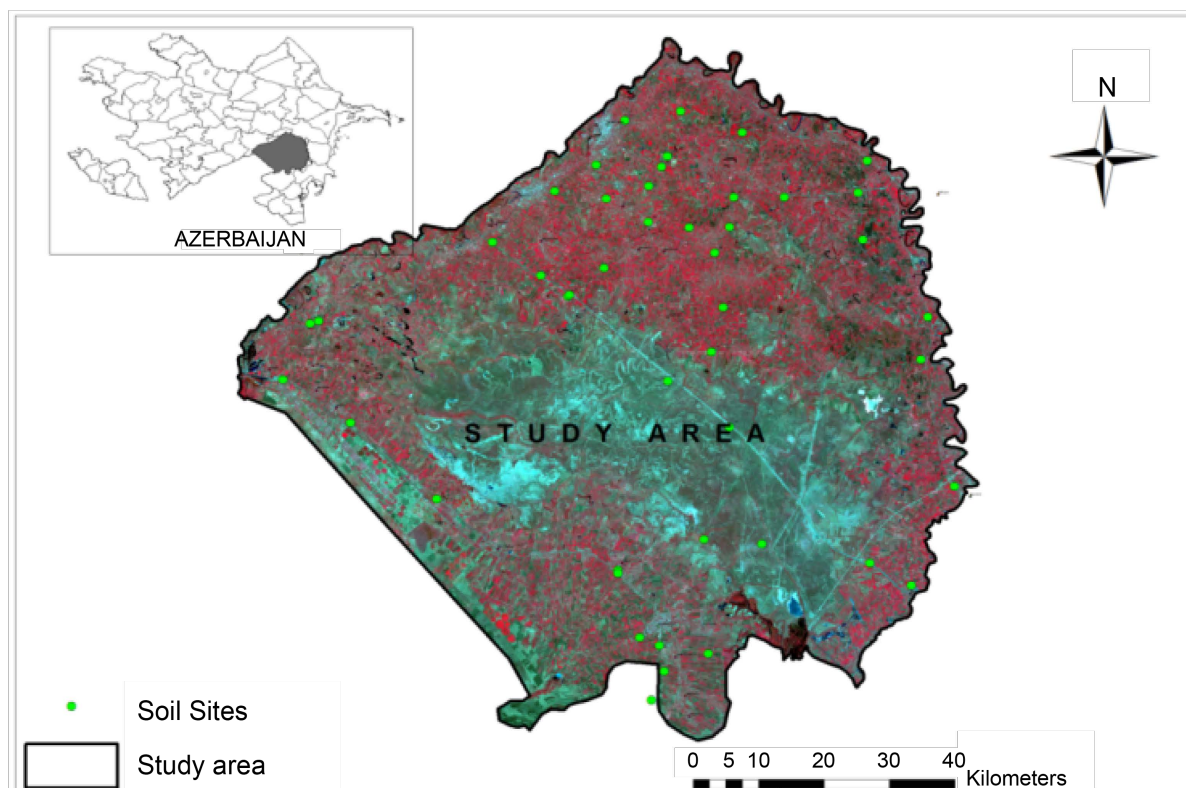


Figure 1. The false colour composite of the study area.

2.2. Soil Samples and Laboratory Analysis

For representative result of soil properties have been selected 46 sites and 194 samples were collected over genetic horizon of soil sites. Soil samples were air dried and 2 mm sieved. Particle size distribution were determined by Pipette method, Organic matter content (OM) and N contents were determined using Walkey-Black and Kjeldahl methods, respectively; CaCO₃ content was determined with the gas-volumetric method; CEC was determined through the extraction in ammonium acetate. Soil pH and electric conductivity (EC) were determined using a 1:5 soil-water suspension

2.3. Soil Vis-NIR Spectral Measurement and Pre-Processing

Air dried and 2 mm sieved soil samples were put in a glass Petri dish for spectral scanning. The samples were scanned using ASD FielSpec 3 radiometer (Analytical Spectral Devices, Boulder, Colo.) at the range of 350 - 2500 nm wavelength, using a contact probe with artificial light. Each soil spectrum was obtained as the mean of 10 scans. Spectralon® white reflectance standard was scanned after every 10 soils samples for white reference.

Pre-processing is important for remove undesirable affects from spectral data. A spectral pre-treatment was performed to reduce the influence of particle size and optical path length variation in the reflectance spectra [11]. The head and the tail (350 - 400 nm and 2450 - 2500 nm) of the spectra showed very high noise and were removed manually from data matrix. The remaining spectra were subjected to Multiplicative Scatter Correction (MSC) for removing scatter effect and first derivative with Savinski Golay algorithms to highlight the peaks of the spectra [12]. All pre-processing data were carried out using Unscrambler 9.7 (CAMO) (See **Figure 2**).

2.4. Spectral Modeling, Calibration and Cross-Validation

The spectral modeling process was carried out by Partial Least Squareregression (PLSR), which is one of the most common methodin Vis-NIR chemometrics analysis. PLSR is a method for relating to data matrix X and Y through a linear multivariate model [13]. PLSR is closely related to Principal component regression (PCR), but PLSR algorithm integrates the compression and regression steps and it selects successive orthogonal factors that maximize the covariance between predictor and response variables [14]. Since the number of soil samples set was not high, but very representative of the soil typologies in Mugan plains, we decided to use a leave-one-out cross validation. Then, PLSR model was calculated on all the 194 samples, whereas the validation of the model was carried out on “total set-1 sample”, iteratively.

3. Result and Discussion

3.1. Soil Properties

Summary statistic of measured laboratory data is provided in **Table 1**. The range of measured data have a variable respectively CaCO₃ (215 - 3.60 g/kg), P (49 - 1 mg/kg), SOM (67.10 - 3.80 g/kg) and for the variable pH (9.50 - 8), K (0.11 - 4.40), have narrow range in the data set opposite to other properties. Standard deviation (SD) for total N, CaCO₃, P, pH and SOM is respectively 0.59, 39.4, 0.97, 7.85, 0.25, 9.37.

3.2. Prediction of Soil Properties

The cross validation approach was used with PLSR model to calibrate soil variables based on entire soil data set. Method was estimated using cross validation R² and RMSE for the soil properties. The result of PLSR model applied to soil properties statistics are shown in **Table 2** and **Figure 3** for variable soil measured and validated properties. PLSR applied to total N, CaCO₃, total P, CEC, EC, SOM and pH. The results showed that there was strong correlation between Vis-NIR spectra and most of the measured soil properties. Only pH showed low correlation with soil spectra. The best calibration models were obtained for EC (R²-0.96), CaCO₃ (R²-0.94), Total N (R²-0.86), Total P (R²-0.86), CEC (R²-0.83), SOM (R²-0.78). The poor calibration result for pH (R²-0.65) but these properties inherently poorly related with Vis-NIR spectroscopy [1].

The statistics of cross-validation and result of PLSR models for CaCO₃, CEC, SOM, Total N, Total P, and pH in soils of Mugan plain are presented in **Table 2**. Standard error for measured vs predicted CaCO₃ (4.22 - 4.11), CEC (0.88 - 0.79), SOM (0.99 - 0.70), Total N (0.06 - 0.05), Total P (0.83 - 0.66) and pH (0.03 - 0.02) respectively.

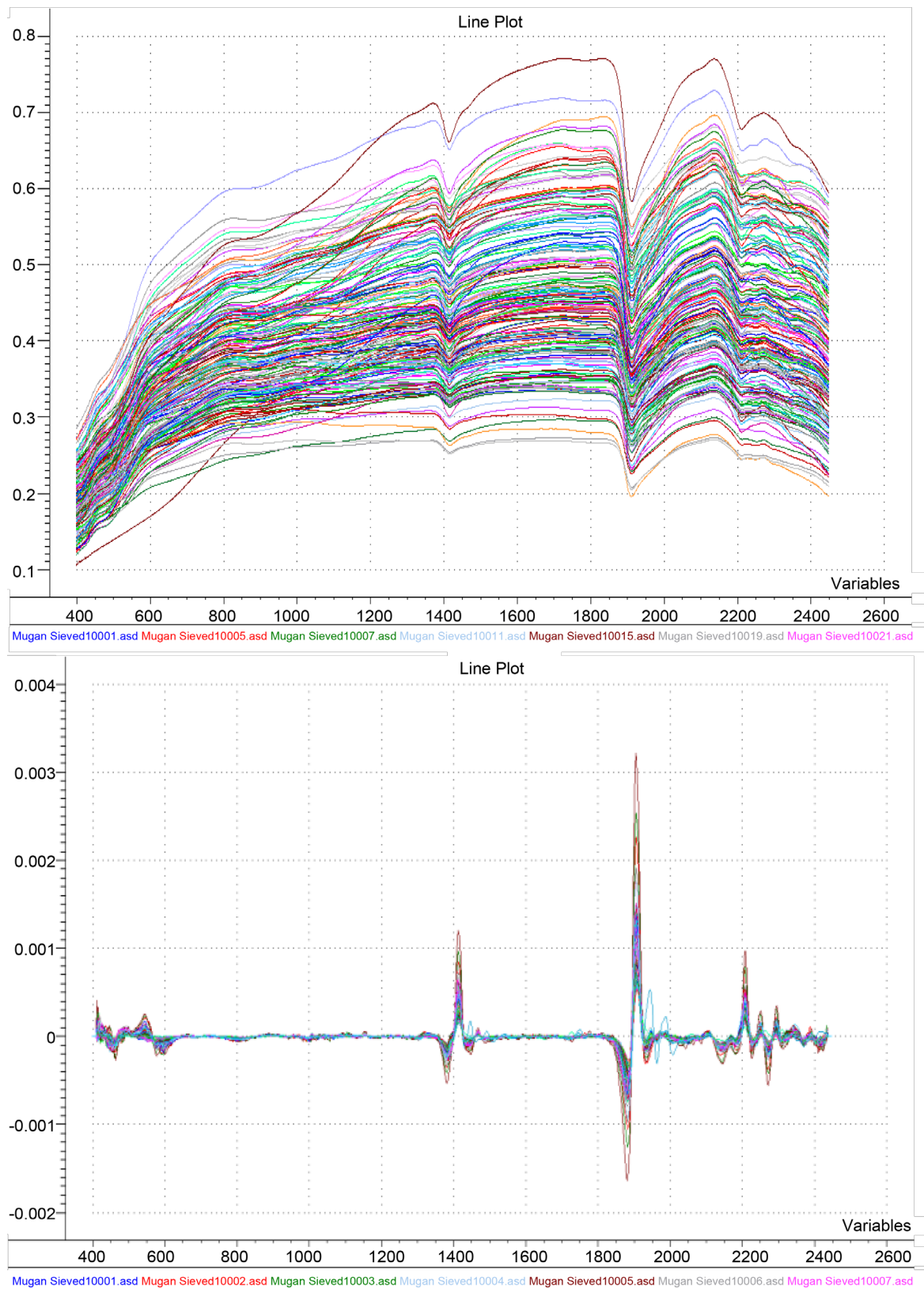


Figure 2. The spectral data plots of soil samples: the raw data (top) and first derivative with Savinski Golay smoothing algorithms (bottom).

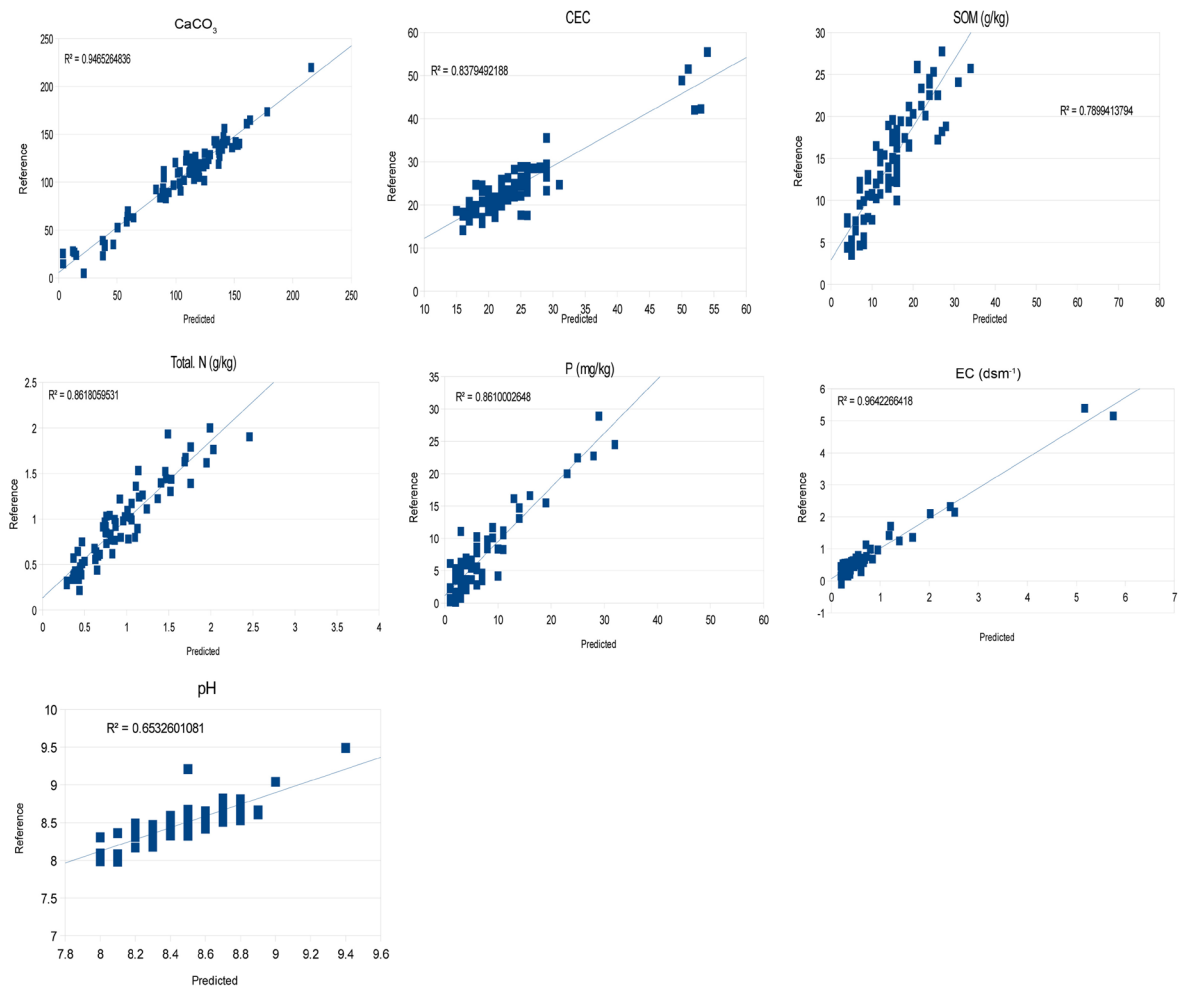


Figure 3. The correlation between measured soil properties and VNIR calibrated soil variables.

Table 1. The statistic of reference data.

	Total N, g/kg	Total CaCO ₃ , g/kg	CEC meq/100g	EC (dS·m ⁻¹)	Fosfor mg/kg P	pH	SOM, g/kg
Min	0.22	3.60	15.60	0.03	1.00	8.00	3.80
Max	3.71	215.80	31.10	5.75	49.00	9.50	67.10
Range	3.49	212.20	15.50	5.72	48.00	1.50	63.30
Mean	1.01	108.02	22.84	0.65	6.49	8.54	15.21
Median	0.93	115.40	22.55	0.33	3.00	8.50	14.80
Sd.error	0.06	4.11	0.38	0.10	0.83	0.03	0.99
SD	0.59	39.04	3.59	0.97	7.85	0.25	9.37
Skew	1.58	-0.78	0.10	3.87	2.99	0.86	2.32

A correlation between soil spectra and carbonates were investigated by using calibration set and measured soil properties data and there is in 3 region (1800, 2350 and 2360 nm) of spectral wavelength to predict soil carbonates [8]. Prediction result demonstrated that in our prediction set is reliable. The predicted soil samples count are 194 and RMSE = 2.89. The best result for quantifying soil carbonate ranged between 10% - 60% and the

average value is 10.8 (**Table 1**) in our samples.

Soil organic matter is major constitutes for soil properties and wide spectral range for assess SOM over spectral region. There is strong correlation between soil color and SOM. For this reason there is suggested to use both vis and NIR to get best prediction result [1]. R-square is 0.83 in our sample set and RMSE is 0.54. Bands around 1400, 1890 and 2200 nm are very sensitive in spectra. Therefore, in study area soils have low content of SOM (**Table 1.**) for this reason Vis regions have high reflectance because of light color. Cation exchange capacity (CEC) prediction result is poor than calibration set. R-square is 0.62 but in calibration set R-square was 0.83 and RMSE = 0.43 prediction quality decreased when using more samples in validation set than calibration set [13]. Generally CEC capacity and SOM demonstrate same spectral similarity [16].

Table 3 contains statistic result of predicted properties of soil vis-nir spectroscopy measurement. The low result recorded for Total N, pH and CEC (R^2 0.44, 0.51 and 0.62). Several scientific works demonstrates that pH hasn't direct spectral responses. The pH regulated several properties of soil. One of them is CEC, SOM and clay content. Generally study area soils undergo calcification and salinization process. The positive result is CaCO_3 and EC have a good relation with spectral reflection. As a prediction result, R-square of EC is 0.82 and RMSE 0.09.

3.3. Conclusions

This paper aimed to predict CaCO_3 , CEC, SOM, EC, pH, Total N and Total P content in the arid zone of Azerbaijan using Vis-NIR reflectance Spectroscopy based on PLSR_{CV} models.

Considering of carbonate, CEC, SOM, EC, pH, Total N and Total P importance in soil also the cost of laboratory measurement of them Vis-NIR is a reliable alternative. It is possible to reduce laboratory costs using nis-NIR spectroscopy. The soil of the study area undergoes different ecological problems such as organic matter losses, salinisation and carbonation. To study them and to prevent these negative processes the area soils

Table 2. Statistics of measured and predicted soil data using PLSR calibration model.

	CaCO_3 (g/kg)		CEC meq/100		SOM (g/kg)		Total N g/kg()		P (mg/kg)		pH	
	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted
Min	3.60	4.74	15.00	14.17	3.00	3.54	0.22	0.21	1	0.116	8	7.98
Max	215.80	219.79	54.00	55.43	67.00	27.74	3.71	2	49	28.9	9.4	9.49
Mean	106.14	106.14	24.68	24.12	14.80	14.35	1.01	0.96	6.49	6.43	8.50	8.52
Median	114.95	114.98	23.00	23.24	14.00	13.95	0.92	0.92	3.00	4.76	8.50	8.53
Sd. error	4.22	4.11	0.88	0.79	0.99	0.70	0.06	0.05	0.83	0.66	0.03	0.02
SD	40.07	38.99	8.42	7.26	9.37	6.27	0.60	0.45	7.85	5.96	0.24	0.22

Table 3. Descriptive statistics of predicted soil properties.

	CaCO_3 (g/kg)	CEC (meq/100)	SOM (g/kg)	Total N (mg/kg)	P (mg/kg)	pH	EC $\text{ds}\cdot\text{m}^{-1}$
Min	9.27	8.95	0.88	0.14	0.12	6.99	0.10
Max	168.70	42.26	36.73	3.82	52.95	9.50	5.63
Range	159.44	33.32	35.85	3.68	52.84	2.51	5.53
Mean	101.51	22.92	15.04	1.05	6.66	8.45	0.88
Median	112.27	22.60	15.02	1.03	5.01	8.47	0.45
[*] SE	2.90	0.43	0.54	0.04	0.50	0.02	0.09
[#] SD	39.71	5.93	7.19	0.47	6.61	0.31	1.12
R^2	0.90	0.62	0.83	0.44	0.73	0.51	0.82

SE^{*}: Standart error, SD[#]: Standart deviation.

should be analyzed with powerful and reliable methods. PLSR based prediction model with cross validation algorithms were achieved positive result. The result of this study shows that it is possible to predict soil properties with reasonable accuracy using PLSR methods.

References

- [1] Viscarra, R.A., Cattle, S.R., Ortega, A. and Fouad, Y. (2009) *In Situ* Measurements of Soil Colour, Mineral Composition and Clay Content by Vis-NIR Spectroscopy. *Geoderma*, **150**, 253-266. <http://dx.doi.org/10.1016/j.geoderma.2009.01.025>
- [2] Viscarra, R.A. and McBratney, A.B. (1998) Laboratory Evaluation of a Proximal Sensing Technique for Simultaneous Measurement of Clay and Water Content. *Geoderma*, **85**, 19-39. [http://dx.doi.org/10.1016/S0016-7061\(98\)00023-8](http://dx.doi.org/10.1016/S0016-7061(98)00023-8)
- [3] Islam, K., Singh, B. and McBratney, A. (2003) Simultaneous Estimation of Several Soil Properties by Ultra-Violet, Visible, and Near-Infrared Reflectance Spectroscopy. *Soil Research*, **41**, 1101-1114. <http://dx.doi.org/10.1071/sr02137>
- [4] Nocita, M., Stevens, A., Toth, G., Panagos, P., Wesemael, B. and Montanarella, L. (2014) Prediction of Soil Organic Carbon Content by Diffuse Reflectance Spectroscopy Using a Local Partial Least Square Regression Approach. *Soil Biology & Biochemistry*, **68**, 337-347. <http://dx.doi.org/10.1016/j.soilbio.2013.10.022>
- [5] Shepherd, K.D. and Walsh, M.G. (2002) Development of Reflectance Spectral Libraries for Characterization of Soil Properties. *Soil Science Society of America Journal*, **66**, 988-998. <http://dx.doi.org/10.2136/sssaj2002.9880>
- [6] Vasques, G.M., Grunwald, S. and Sickman, J.O. (2008) Comparison of Multivariate Methods for Inferential Modeling of Soil Carbon Using Visible/Near-Infrared Spectra. *Geoderma*, **146**, 14-25. <http://dx.doi.org/10.1016/j.geoderma.2008.04.007>
- [7] Leone, A.P., Raphael, A., Rossel, V., Amenta, P. and Buondonno, A. (2012) Prediction of Soil Properties with PLSR and Vis-NIR Spectroscopy: Application to Mediterranean Soils from Southern Italy. *Current Analytical Chemistry*, **8**, 283-299.
- [8] Ben-Dor, B.A. (1990) Near Infrared Reflectance Analysis of Carbonate Concentration in Soils. *Applied Spectroscopy*, **44**, 1064-1069. <http://dx.doi.org/10.1366/0003702904086821>
- [9] Stenberg, B., Rossel, R.A.V., Mouazen, A.M. and Wetterlind, J. (2010) Visible and Near Infrared Spectroscopy in Soil Science. In: Sparks, D. L., Ed., *Advances in Agronomy*, Vol. 107, Academic Press, Burlington, 163-215.
- [10] Babayev, M., Hasanov, V., Cafarova, Ch. and Huseynov, S. (2011) The Classification, Nomenclature and Morphogenetic Diagnostics of Azerbaijan Soils. Elm, Baku, 448.
- [11] Williams, P. and Norris, K., Eds. (1987) *Near-Infrared Technology in the Agricultural and Food Industries*. American Association of Cereal Chemists, St. Paul, 330 p.
- [12] Savitzky, A. and Golay, M. (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, **36**, 1627-1639. <http://dx.doi.org/10.1021/ac60214a047>
- [13] Bilgili, V., van Es, H., Akbash, F., Durak, A. and Hively, W. (2010) Visible-Near Infrared Reflectance Spectroscopy for Assessment of Soil Properties in a Semi-Arid Area of Turkey. *Journal of Arid Environments*, **74**, 229-238. <http://dx.doi.org/10.1016/j.jaridenv.2009.08.011>
- [14] Viscarra, R.A. and Behrens, T. (2010) Using Data Mining to Model and Interpret Soil Diffuse Reflectance Spectra. *Geoderma*, **158**, 46-54. <http://dx.doi.org/10.1016/j.geoderma.2009.12.025>
- [15] Minasny, B., McBratney, A.B., et al. (2011) Removing the Effect of Soil Moisture from NIR Diffuse Reflectance Spectra for the Prediction of Soil Organic Carbon. *Geoderma*, **167-168**, 118-124. <http://dx.doi.org/10.1016/j.geoderma.2011.09.008>
- [16] Wei, C.L., Zhao, Y.G., Li, D.C., Zhang, G.L., Wu, D.W. and Chen, J.K. (2014) Prediction of Soil Organic Matter and Cation Exchange Capacity Based on Spectral Similarity Measuring. *Transactions of the Chinese Society of Agricultural Engineering*, **30**, 81-88.