International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319-6505, Impact Factor: 6.614
Available Online at www.journalijcar.org
Volume 9; Issue 04 (C); April 2020; Page No.21972-21976
DOI: http://dx.doi.org/10.24327/ijcar.2020.21976.4327

Research Article

SOIL PROPERTIES PREDICTION ON THE BASIS REFLECTANCE SPECTROSCOPY USING PLSR MODELING

A.D.Abdellatif[1]*, Mohamed E. Abou-Kota[1] and Ali R. A. Moursy[2]

[1]Soil, Water and Environment Research Institute, Agricultural Research Center, Egypt
[2]Division of Soil and Water Sciences, Faculty of Agriculture, Sohag University, Egypt

A R T I C L E I N F O

Article History:
Received 6th January, 2020
Received in revised form 15th February, 2020
Accepted 12th March, 2020
Published online 28th April, 2020

Key words:
Spectroradiometer, Vis-NIR, PLSR, hyperspectral

A B S T R A C T

This study aimed to apply the partial least square regression algorithm to estimate different soil parameters. Surface soil samples were collected and analyzed for some parameters (i.e. pH, sand, silt, clay, and CEC) using the conventional methods of soil analysis. Hyperspectral signatures of soil samples were collected in the range of Vis-NIR spectra (350-2500nm) using the analytical spectroradiometer device (ASD) in the laboratory conditions. The PLSR model was applied to soil spectra and soil parameters' data to develop the calibration and validation models. The obtained results showed that sand and CEC soil parameters recorded an excellent predictability with R^2 values 0.91 and 0.86, and RPD values 3.00 and 2.41, respectively. The rest soil parameters such as pH, silt and clay were having moderate predictability whereas R^2 values were 0.68, 0.50, and 0.62, and RPD values were 1.84, 1.41 and 1.70, respectively. The diffuse reflectance spectroscopy integrated with multivariate regression models such as PLSR could successfully estimate soil parameters with good prediction. This technique was found to be promising for soil parameters' prediction. It saves time, effort, chemicals, and many soil parameters that can be estimated simultaneously.

INTRODUCTION

Soil is one of the basic natural resources, and its description and evaluation are critical (Soil Survey Staff, 2011). For a fact, the most reliable method for soil analysis is the routine conventional or traditional laboratory method. Unfortunately, the conventional methods of soil sampling and laboratory analysis are expensive, time and chemicals consuming, laborious, and require a lot of preparation stages (Disla*et al.,* 2014). The global need is for faster, cheaper, more cost-efficient methods and also having enough soil spatial data which will be used in environmental monitoring, modeling, mapping, and precision agriculture (Dammate*et al.,* 2015).

Over the past 35 years, Diffuse Reflectance Spectroscopy (DRS) has proven high efficiency for estimating soil properties. This technique can be applied in both laboratory and field conditions by using spectroradiometers. It can estimate many soil properties at the same time with minimum or without samples preparation (Kadupitiya*et al.,* 2010). The Vis-NIR-MIR spectral range (0.35 to 25 μm) is suitable for estimating the majority of soil properties (Ogen*et al,* 2019). Nowadays, multivariate statistics and chemometrics are used in the prediction of soil parameters by quantitative soil spectroscopy, and these techniques still growing (Chabrillat*et al.,* 2013).

Corresponding author:* **A.D.Abdellatif
Soil, Water and Environment Research Institute, Agricultural Research Center, Egypt

Data analysis techniques are dependent on the number of spectral variables of the soil spectral data. The spectral data obtained from field or laboratory conditions by ground spectrometers are noisy and hard to be evaluated. Here the role of spectral transformation appears to clean noises, correct non-linearity measurement, sample variations and develop fit soil spectral curves (Stenberg, 2010). Partial Least Square Regression (PLSR) is the most popular and widely used technique in chemometrics for quantitative analysis of reflectance spectra (Wold*et al.,* 2001).The ability of the hyperspectral RS technique to predict a soil property could be evaluated using statistical parameters such as the correlation coefficient (R^2), the Root Mean Square Error (RMSE) and Ratio of Performance Deviation (RPD) which are commonly used for the DRS technique (Woodcock, 2006).

Many researchers reported good results with regression analysis for soil properties characterization. Li *et al.* (2015) used the PLSR method to estimate the soil nitrogen; R^2 was 0.57. Sh. *et al.* (2015) used Vis-NIR spectral libraries to predict SOM (R^2=0.74) from newly acquired spectral by applying the PLSR method. The same technique was used in Alaska, USA, where soil OC was predicted in the NIR range of spectra with R^2=0.41 (Zabowski*et al.,* 2011). Curcio*et al.* (2013) applied the PLSR technique on Vis-NIR and SWIR reflectance data to predict soil sand, silt and clay (R^2=0.87, 0.60 and 0.80, respectively). Rossel*et al.* (2006) applied DRS in across the Vis-NIR-MIR spectra to analyze the soil properties using the PLSR algorithm for prediction. They found that R^2 values for pH, EC, OC, clay, silt and sand were 0.73, 0.29, 0.72, 0.67, 52 and 0.75, respectively. Antonio *et al.* (2012) quantitatively predicted soil parameters using PLSR in Italy soils (R^2=0.57, 0.50, 0.82 and 0.90 for sand, silt, clay, and OC) in vis-NIR-MIR regions. The integration between Vis-NIR and PLSR model has an advantage for determining soil EC with R^2=0.90 (Fikrat*et al.*, 2016). Mouazen*et al.* (2010) found that the PLSR model was able to predict SOC with R^2=0.80. Srivastava *et al.,* (2004) applied linear regression-NIR modeling to predict some soil parameters in a part of central India, R^2 values of pH, OC and clay were 0.77, 0.78 and 51, respectively. By using the PCR model, Kadupitiya *et al.* (2010) were able to predict soil properties (pH, EC, and OC) in Punjab soils with R^2=0.82, 0.85 and 0.79, respectively. The previous introduction reviewed the importance of diffuse reflectance spectroscopy with the application of the multivariate regression model for the prediction of soil parameters. Thus, the current study aimed to use the hyperspectral remote sensing technique integrated with PLSR for characterizing and predicting soil parameters, and also to assess the performance of the applied prediction model.

MATERIALS AND METHODS

The study area

The study area is a part of the Cairo-Ismailia agriculture road and Ismailia canal with a total area of 760km^2. It locates in the northeastern part of Nile delta with latitudes 30^{o}27$^{"}$6.4$^{"}$to 30^{o}39'10.7" and longitudes 31^{o}41'54.16" to 32^{o}05'52.3".The climatic data of Ismailia Governorate show a very low annual precipitation (≈22 mm/y). The minimum and maximum average of the annual temperature is 16.2° Cand 28.9° C (CLAC, 2010). The main water source is Nile water through the Ismailia Canal. The studied area has a hyperthermic temperature regime with a torric soil moisture regime.

Soil sampling

In the year 2019, thirty-one representative surface soil samples (0–25 cm) were collected according to the corresponding physiographic units in the study area. The soil samples were air-dried, ground and 2mm sieved to be scanned using the spectroradiometer and also analyzed for their properties. Figure (1) showed the location map of the study area and also the soil sampling locations.

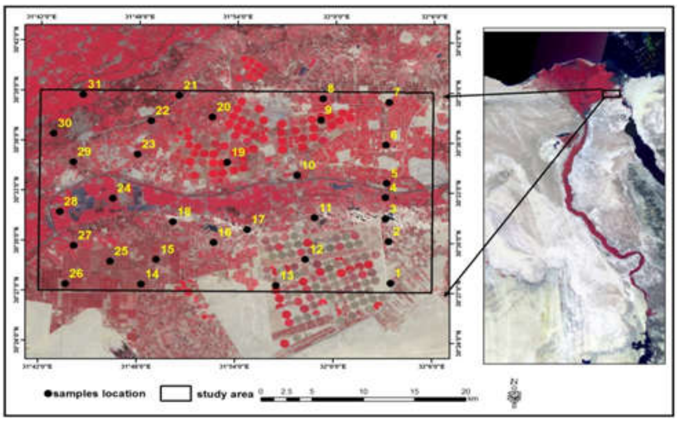

Fig 1 Location map of the study area and soil sampling locations.

Hyperspectral data collection

The hyperspectral reflectance data of each soil sample was obtained in the laboratory conditions using the Analytical Spectroradiometer Device (ASD) Model: PSR-3500 Serial: 1166005 produced by Spectral Evolution, USA at the Vis-NIR spectral range (0.35-2.5μm). Soil samples were placed at a 5-cm diameter Petri dish and the spectra were collected using a high-intensity light source probe. The instrument was optimized and calibrated using white spectral on for obtaining absolute reflectance readings before the samples were recorded. The average of 3 spectra was recorded at each soil sample to minimize noise produced by the instrument for obtaining the final spectra. The RS3 (version 6.3) inbuilt software was used to record the reflected spectra.

Soil analysis

Conventional methods of soil analysis were followed. Soil particle size distribution test was done using an international pipette method (Bashour and Sayegh 2007). The percentages of clay, silt, and sand were calculated. Soil PH was measured in the soil pasteby a potentiometric method using a glass electrode (Jackson, 1967). The cation exchange capacity of soil samples was determined using the method described by Page *et al.* (1982).

Statistical analysis

Simple statistical analysis was applied to the soil parameters' data using MS excel software. Descriptive statistical analysis was done which minimum, maximum, mean, range, standard error, standard deviation; variance, mode, median, kurtosis and skewness parameters were calculated. Pearson correlation was applied to the obtained data, and correlation coefficient was calculated.

Hyperspectral data pre-processing

Soil spectral data collected in the hyperspectral remote sensing laboratory conditions at the range of spectra from 350 to 2500 nm were arranged in text format (.csv files) to be easily processed. The original obtained spectral data were in 1nm interval, so it was converted to be in 5nm intervals using MATLAB (version 2019) software to reduce the spectral variables/bands number and to enhance the quality of the calibration and validation models of soil properties.

Multivariate regression model application

The PLSR is a commonly used technique for quantitative spectral analysis. It is used to develop prediction models when many predictor variables are highly collinear. The PLSR algorithm selects the best orthogonal factors that maximize the covariance between predictor (X spectra) and response variables (y laboratory data). By fitting a PLSR model, a few PLSR factors are selected to explain most of the variation in both predictors and responses. The PLSR decomposes X and y into factor scores (T) and factor loadings (P and q) according to the following equations (1 and 2).

$$\mathrm{X} = \mathrm{TP} + \mathrm{E} \quad (1)$$

$$\mathrm{y} = \mathrm{Tq} + \mathrm{f} \quad (2)$$

whereas, X and y are mean-centered before decomposition. The decomposition is performed simultaneously and in such a way that the first few factors explain most of the variation in X and y. The remaining factors relate to noise and can be ignored, hence the addition of residuals E and f. Generally, the resulting matrices and vectors have a much lower dimension than X and y. Therefore, given a new spectrum x, the soil property y can be estimated as a bilinear combination of the factor scores and factor loadings of x (Martens and Næs 1989). The PLS package in R studio software was used for developing the calibration and validation models of the different studied soil parameters. Soil spectral data and the laboratory soil data were combined in (.csv) files to be used in R software. Moreover, the spectral data, as well as the soil parameters' data, were processed through different stages.

For enhancing the modeling performance and the predictability of different soil parameters, data processing was done through the following stages: (i) data normalization (giving values between 0 and 1 for the soil spectral data); (ii) data dividing (the whole spectral and soil data were divided into two data sets; 80% of the data (n=24) was separated to be as a calibration data set and 20% of the data (n=7) for a validation); (iii) data sorting (data arrangement for the randomized distributing the values depending on their weights among the calibration and validation data sets); and (iv) removing the outliers (the much higher or lower values in the whole soil parameters' data set were removed as outliers).

Models quality evaluation (Validation of the developed prediction models)

Two statistical indices were used for validation of developed prediction models and were the R^2, Randomized Mean Square Error (RMSE), Ratio Prediction Deviation (RPD) as described by (Islam *et al.,* 2003) and shown in equations (3, 4 and 5).

The correlation coefficient (R^2)

$$R^2 = 1 - \left(\frac{\sum(Y_{pred} - Ymeas)^2}{\sum(Y_i - Ymeas)^2}\right) \tag{3}$$

Where, Ypred = predicted values; Ymean = mean of measured values; Ymeas = measured values; n= number of predicted or measured values with I = 1, 2, ...n.

Room Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n\Sigma(y-x)2}} \tag{4}$$

Where y is a predicted value of soil parameter and x is a measured value.

Ratio of Performance Deviation (RPD)

$$RPD = \frac{SD}{RMSE} \tag{5}$$

where SD is the standard deviation of measured values in the validation dataset; and RMSE= root mean square error of prediction in the validation dataset.

Chang *et al.* (2001) categorized the ability of NIR spectra to predict soil properties into three categories based on the ratio of performance deviation (RPD) and the Correlation coefficient (R^2) values as shown in table (1).

Table 1 Categories of NIR predictability of soil parameters

Category	RPD	R^2	Parameters
A	>2	1-0.8	TC, TN, moisture, sand, silt, exch.Ca and CEC.
B	2-1.4	0.8-0.5	Clay, pH, mineralizable N, extractable K, Ca, Mg, Fe and Mn
C	<1.4	<0.5	Extractable Cu, P, Zn and Na.

RESULTS AND DISCUSSION

Soil characterization

Descriptive statistical analysis of the soil properties data was given in table (2). From the obtained data, the soil of the studied area was ranged from non-alkaline to strongly alkaline whereas soil pH values were ranged from 7.14 to 8.29 with an average of 7.85. The dominant textural classes of these soils were sand and sandy loam. Clay loam and sandy clay loam classes were recorded for few samples. Sand fraction ranged from 28.50 to 84.60% with an average of 63.85%. The minimum and maximum values of silt were 10.90 and 36.3% with an average of 21.13%. Clay content ranged between 3.20and35.20%. The CEC was low in general for these soils and ranged from 8.10to23.50Cmole $(p^+).kg^{-1}$. Table (3) showed the correlation among soil parameters.

Table 2 Descriptive statistics of soil parameters.

	pH	Sand	Silt	Clay	CEC
		(%)			Cmole $(p^+).kg^{-1}$
Mean	7.85	63.85	21.13	14.81	13.64
Standard Error (S.E)	0.04	3.70	1.54	2.20	1.07
Median	7.77	74.60	17.80	6.90	9.90
Mode	7.75	83.20	12.20	5.70	8.20
Standard Deviation (S.D)	0.23	20.60	8.56	12.23	5.96
Sample Variance	0.05	424.41	73.28	149.63	35.54
Kurtosis	1.99	-1.35	-1.30	-1.42	-1.52
Skewness	-0.08	-0.65	0.57	0.66	0.57
Range	1.15	56.10	25.40	32.00	15.40
Minimum	7.14	28.50	10.90	3.20	8.10
Maximum	8.29	84.60	36.30	35.20	23.50

Table 3 Pearson correlation between studied soil parameters.

	pH	sand	silt	clay	CEC
pH	1				
sand	-0.44	1			
silt	0.41	-0.99	1		
clay	0.45	-0.99	0.98	1	
CEC	0.39	-0.88	0.87	0.89	1

High negative correlation was recorded between sand and silt, clay and CEC with correlation coefficient (*r*= -0.99, -0.99 and -0.88), respectively. Low correlation coefficient value was observed for pH (*r*=-0.44). High positive correlation was for silt-CEC (*r*= 0.87), silt-clay (*r*= 0.98) and clay-CEC (*r*= 0.89). Soil pH was positively correlated with silt, clay and CEC with correlation coefficient values (*r*=0.41, 0.45 and 0.39), respectively.

Soil hyperspectral signature

From the figure (2), it was shown that reflectance spectra of soil samples followed the same basic shape as observed by many researchers, with prominent absorption bands around 1400, 1900, and 2200 nm (Shepherd and Walsh 2002). These bands are associated with clay minerals, for example, OH features of free water at 1400 and 1900 nm, and lattice OH features at 1400 and 2200 nm (Hunt 1980).

PLSR-ASD modeling

Table (4) showed the obtained results of PLSR-ASD modeling for the calibration and validation model using the data set of the study area. The obtained data of the validation model were plotted against the measured data of the different best-predicted soil parameters and shown in the figure (3).

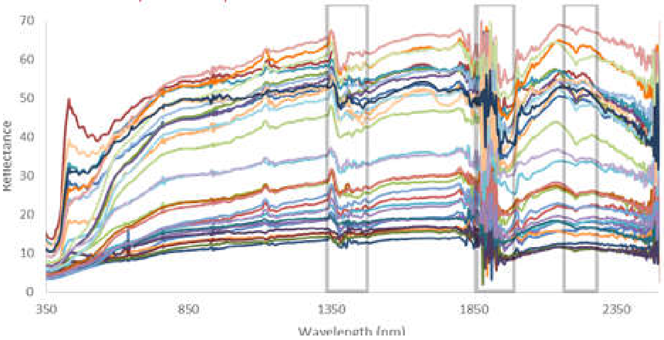

Fig (2) Soil Spectral signature of soil samples.

Table 4 The performance assessment of the PLSR calibration and validation model of ASD data of the study area.

Parameter	Calibration			Validation		
	R^2	RMSE	RPD	R^2	RMSE	RPD
pH	0.98	0.01	33.41	0.68	0.12	1.84
Sand (%)	0.98	1.29	15.79	0.91	7.80	3.00
Silt (%)	0.99	0.26	35.00	0.50	7.34	1.41
Clay (%)	0.99	0.58	21.43	0.62	7.37	1.70
CEC (Cmole (p$^+$).kg^{-1})	0.99	0.41	13.52	0.86	6.09	2.41

From the obtained data, the PLSR calibration model was performing well for all soil parameters whereas values of R^2>0.50 and RPD >1.40. The R^2 values of calibration were 0.98, 0.98, 0.99, 0.99 and 0.99 while RPD values of calibration were 33.41, 15.79, 35.00, 21.43 and 13.52 for pH, sand, silt, clay and CEC, respectively.

The same performance was recorded for the prediction model developed for all soil parameters. The R^2 values of validation were 0.68, 0.91, 0.50, 0.62 and 0.86 while RPD values of validation were 1.84, 3.00, 1.41, 1.70 and 2.41, respectively. These results are consistent withmany studies (i.e. the findings of Srivastava *et al.* (2004); Rossel *et al.* (2006); Kadupitiya *et al.* (2010); Curcio *et al.* (2013); etc.).

According to the developed criteria of Chang *et al.* (2001) of the ability of Vis-NIR spectra to predict soil properties, sand and CEC soil parameters were under the 'A' category whereas R^2values between 0.80 and 1.00 and RPD values more than 2.00. Sand and CEC soil parameters can be well predicted using the PLSR prediction model.

The rest soil parameters (pH, silt and clay) could be under the 'B' category whereas R^2 values between 0.50 and 0.80, and RPD values between 1.40 and 2.00. These soil parameters can be moderately predicted using the PLSR prediction model.

CONCLUSION

Soils of the studied area was alkaline and varied between fine and coarse in texture with low CEC. Partial Least Square Regression model was applied for developing the calibration and validation models for predicting the different soil parameters. Soil parameters (i.e. sand and CEC) were predicted well by the PLSR prediction model with R^2 values were between 0.80 and 1.00 and RPD values were more than 2.00. Moderate performance was observed for the PLSR prediction model for estimating the rest soil parameters (pH, silt and clay). Hyperspectral reflectance data in the range of Viz-NIR (350-2500nm) which integrated with the partial least square regression PLSR model as an empirical technique, showed promising performance for soil parameters' prediction. Further studies can be done with an application of several algorithms to enhance the prediction of soil parameters.

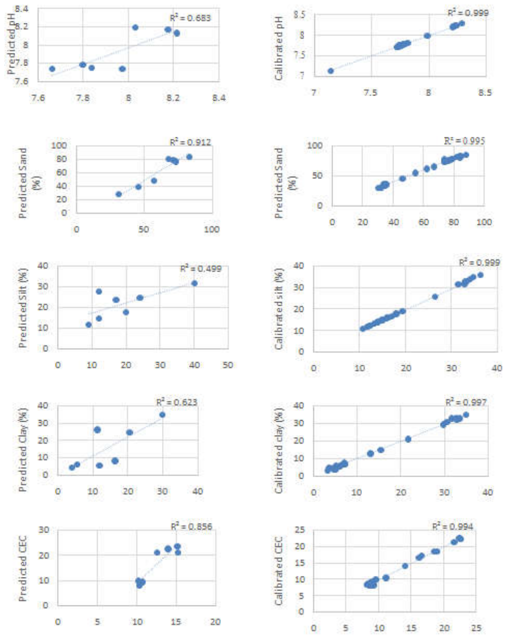

Fig 3 The measured values plotting against the predicted values of soil parameters using the PLSR prediction model; pH, sand, silt, clay, and CEC.

References

Antonio, P.L., Rossel, R.V., Pietro, A., Andrea, B. 2012. Prediction of soil properties with PLSR and Vis-NIR spectroscopy: Application to Mediterranean soils from Southern Italy. Current Analytical Chemistry; 2:283–299.

Bashour, I.I., Sayegh, A.H. 2007. Methods of analysis forsoils ofarid andsemi-arid regions.Food and Agriculture Organization of the United Nations (FAO); pp128.

Chabrillat, S., Ben-Dor, E., Rossel, R.V., Dematte, J.A. 2013. Quantitative soil spectroscopy.Applied and Environmental Soil Science.https://doi.org/10.1155/2013/616578.

Chang, C.W., Liard, D.A., Mausbach, M.J., Hurburgh, C.R. 2001. Near-infrared reflectance spectroscopy-principal components regression analysis of soil properties.American Journalof Soil ScienceSociety; 65:480-490.

CLAC. 2010. Central Laboratory for Agriculture of Climate. https://www.clad.edu.eg.

Curcio, D., Ciraolo, G., D'Asaro, F., Minacapilli, M. 2013.Prediction of soil texture distributions using VNIR-SWIR reflectance spectroscopy.Procedia Environmental Sciences; 19:494-503.

Dematté, A.M., Marcelo, R.V., Bruna, C.G., Caio, T.F., Marcus, V.S. 2015. Hyperspectral remote sensing as an alternative to estimate soil attributes. RevistaCiênciaAgronômica; 2:223-232.

Disla, S.J., Janik, L.S., Rossel, R.V., MacDonald, L.M., McLaughlin, M.J. 2014.The Performance of Visible, Near, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. Applied Spectroscopy Reviews; 2: 139-186.

Fikrat, F., Maharram, B., Simone, P., Giovanni, L. 2016.Using Visible-Near Infrared Spectroscopy to Predict Soil Properties of Mugan Plain, Azerbaijan. Open Journal of Soil Science; 3:52-58.

Hunt, G.R. 1980. Spectroscopic properties of rocks and minerals.In ‘Handbook of physical properties of rocks’. Ed. CR Stewart, CRC Press: Boca Raton, FL; pp. 295–385.

Islam, K., Singh, B., McBratney, A.B. 2003. Simultaneous estimation of various soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. Australian Journal of Soil Research; 41:1101–1114.

Jackson, M.L. 1967. Soil Chemical Analysis.Prentice-Hall of India Pvt. Ltd., New Delhi; pp498.

Kadupitiya, H.K., Sahoo, R.N., Ray, S.S., Chakraborty, D., Ahmed, N. 2010. Quantitative assessment of soil chemical properties using visible (VIS) and near-infrared (NIR) proximal hyperspectral data. Tropical Agriculture; 158:41–60.

Li, S., Wenjun, J.,Songchao, C., Jie, P., Yin, Y., Zhou, H. 2015.Potential of VIS-NIR-SWIR Spectroscopy from the Chinese Soil Spectral Library for Assessment of Nitrogen Fertilization Rates in the Paddy-Rice Region, China. Remote Sensing;7:7030-7043.

Martens, H., Næs, T. 1989.Multivariate Calibration.John Wiley and Sons, Chichester; pp419.

Mouazen, A.M., Kuang, B., De-Baerdemaeker, J., Ramon, H. 2010. Comparison between principal component, partial least squares and artificial neural network analyses for accuracy of measurement of selected soil properties with visible and near-infrared spectroscopy. Geoderma; 158:23-31.

Ogen, Y., Zaludab, J., Francosb, N., Goldshlegerc, N., Ben-Dor, E. 2019.Cluster-based spectral models for a robust assessment of soil properties.Geoderma; 340:175–184.

Page A.L.,Helmet, P.A., Soltanpour,R.H. 1982. Methods of soil analysis, part 3, American Society of Agronomy, Inc. Madison, Wisconsin. USA.

Rossel, R.V., Walvoort, D.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. Elsevier; 2:59-75.

Sh, Z., Ji, W., Rossel, R.V., 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; 4:679-687.

Shepherd, K.D., Walsh, M.G. 2002.Development of reflectance spectral libraries for characterization of soil properties.American Journal of Soil Science Society; 66:988–998.

Soil Survey Staff. 2011. Soil survey laboratory information manual. Soil Survey Investigations Report No. 45, Version 2.0. USDA–NRCS, Washington, DC.

Srivastava, R., Prasad, J.,Saxena, R. 2004.Spectral reflectance properties of some shrink-swell soils of Central India as influenced by soil properties.Agropedology; 14:45–54.

Stenberg, B. 2010. Effects of soil sample pretreatments and standardized rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. Geoderma; 158:15-22.

Wold, S., Antti, H., Lindgren, F., Ohman, J. 1998.Orthogonal signal correction of near-infrared spectra.Chemom.Intell. Lab. Syst.; 44:175–185.

Woodcock, C.E. 2006.Uncertainty in Remote Sensing.Uncertainty in Remote Sensing and GIS.Chichester, UK: John Wiley and Sons; pp19−24.

Zabowski, D., Whitney, N., Gurung, J., Hatten, J. 2011. Total soil carbon in the coarse fraction and at depth. Forest Science; 57:11–18.

How to cite this article:

A.D.Abdellatif, Mohamed E. Abou-Kota and Ali R. A. Moursy (2020) ' Soil Properties Prediction on the Basis Reflectance Spectroscopy Using Plsr Modeling', *International Journal of Current Advanced Research,* 09(04), pp. 21972-21976.
DOI: http://dx.doi.org/10.24327/ijcar.2020. 21976.4327

****