Journal of Advanced Computer Applications
vol-01, issue-01, December, 2024

ISSN No.: XXXX-XXXX
Pages 109-118

JOURNAL OF ADVANCED COMPUTER APPLICATIONS

ISSN: XXXX-XXXX (Online)
Vol: 01 **Issue:** 01
Contents available at: https://www.swamivivekanandauniversity.ac.in/jaca/

Evaluation of organic carbon in Gangetic delta-soils using NIR

Somdeb Chandra[1] and Dilip Singh[1] *

[1] *Department of Electronics and Communication Engineering, Swami Vivekananda University, Barrackpore-700121, WB, INDIA, Email ID: somdebc@svu.ac.in**

ABSTRACT

An Internet of Things- (IoT-) based health monitoring systems are possibly enormously advantageous for The fertility of specific agricultural field normally depends on bio-chemical contents (e.g. Organized Carbon, Nitrogen, and Phosphorus) in soil. This study aimed to predict the organic carbon content estimation in soil by the noninvasive near infrared reflectance (NIR) spectroscopy method. Partial least squares (PLS) algorithm was conducted on the calibration of regression model on NIR spectra of 900-1700nm wavelength range. The number of PLS factors and the choice of preprocessing methods were optimized simultane-ously by leave –one-sample out cross-validation during the model calibra-tion. The efficacy of the model developed was evaluated by root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP) and correlation coefficient (R) which are 0.342, 0.1481and 0.942 respectively. Finally results showed that NIR spectroscopy with PLS algo-rithm could be used to analyze Nitrogen contents in soil with more than 87% accuracy.

Keyword: Near infrared reflectance (NIR) spectroscopy; Partial least squares (PLS); Soil Organic Carbon (SOC); Leave-one-sample out cross-validation.

I. INTRODUCTION

Soil, a crucial factor of the biosphere that plays a vital role in the biogeochemical, hydrological and erosion cycles and provides various important ecosystems goods and services [1-2], is one of the major terrestrial reservoirs of organic carbon [3]. Soil organic carbon (SOC) takes a key role to indicate the soil quality such as soil fertility and climate regulation and also is an essential key component for agricultural productivity [4]. In addition, SOC enriches the availability of nutrients and improves soil structure, physical properties, and water holding capacity [5]. This is because of its key functional roles in biological, physical, and chemical processes in the soil [6]. Therefore, quantitative and

---

* Authors for Correspondence

qualitative analysis of SOC is important as it enhances the understanding of ecosystem processes.

Soil organic carbon (SOC) can be normally found up to 1m soil depth [7-8] and often without considering SOC in the deeper soil [9-10]. Significant amounts can be estimated below 1m soil depth [9 11]. The subsoil between 1 and 3m depth can store between 760 and 1520Pg C. Deeper soil horizons may have high capacity to sequester SOC as the turnover time and chemical recalcitrance of SOC increase with depth [9]. Therefore, it is essential to predict soil data rapidly for environmental monitoring, soil quality assessment and precision agriculture and forestry [12-13]. The soil total nitrogen and organic matter content can reflect the crop growth supply ability which is an important index to measure the soil fertility. Therefore, the measurement of soil total nitrogen and organic matter content can generally understand soil fertility conditions [14]. But the traditional measurement method is getting soil samples back to analyze in the lab. Moreover, these sophisticate methods are costly, destructive and time consuming. Additionally, these methods require tedious and complex processing for samples, chemical rea-gents and also expert technicians to monitor the soil fertility testing [15-16].

On contrast, Near-infrared reflectance (NIR) spectroscopy [17] is a prompt, precise and non-destructive analytical method which could supplement or replace the cost-ly conventional soil analysis. Over the past few decades, however, NIR reflectance spectroscopy has rapidly developed to become a fast and robust analytical tool for many agricultural, pharmaceutical and food products [18]. The objective of this study is to predict the organic carbon in soil by noninvasive method to forecast soil fertility for agricultural production. The NIR spectral data are often calibrated with classical multivariate calibration analysis. Many spectral pre-processing techniques have been developed to reduce the effect of variations in spectral data that are not related to the chemical variations in sample.

Here, the most suitable pre-processing techniques e.g SNV and MSC was used to rectify the scattering effect [19]. After scattering correction PLS regression method [20-21] was used to develop a model with optimized factors which could predict the sucrose concentration in leaves of rice plant.

II. EXPERIMENT

A) Soil Sampling

Soil samples were collected from 19 different geographical places of West Bengal, India by Centre for Development of Advanced Computing (CDAC), Kolkata, India and chemical analysis was made by Vivekananda Institute of Biotechnology, West Bengal, India. The samples were analyzed for organic-carbon concentration in per-centage. The experiments were conducted during the months of November 2016. The soil samples were also brought to the laboratory of Jadavpur University for NIR spectra acquisition. NIR spectrophotometer (Model: Stellar

NetInc, USA) was used for spectrum analysis of same samples and the technique for the measurement of NIR was standardized inside the laboratory in dark chamber.

B) Chemical Analysis

The chemical analysis was done by using Walkley and Black method. The reference was taken from Soil Chemical Analysis, M.L.Jackson (1967) Pentice Hall of India, New Delhi.

C) Spectra Acquisition

The NIR spectra of soil samples were collected in reflectance mode. The wave-length range of NIR spectrometer is from 900 nm to 1700 nm. The data points in the spectra were acquired in 1.75 nm intervals. The spectrometer thus produced 457 spectral points for each sample. Five replicate measurements were taken for each sample thus in total 457 x 19 x 5 spectral data points were obtained.

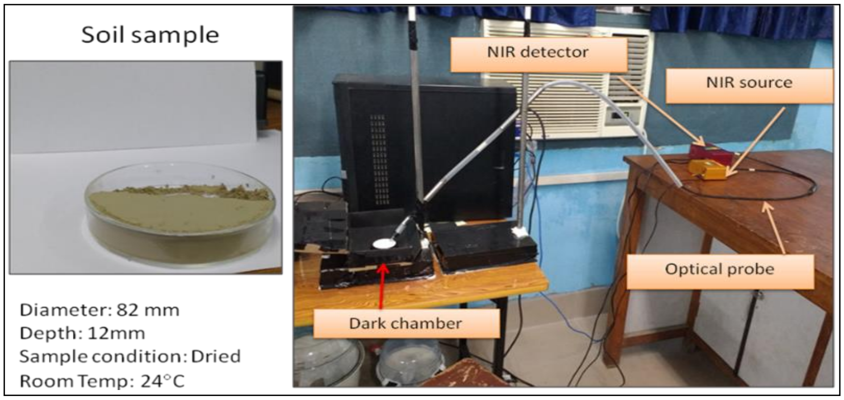

Fig. 1: Experimental Setup of NIR Spectrometry

D)Pre-processing Methods

In this study, several standard preprocessing techniques were investigated, but the Standard Normal Variate (SNV) and Multiplicative Scattering Correction (MSC) techniques were adopted due to its satisfactory results in other soild samples i.e rice, tea etc.

The soil samples usually has a rough surface due to it's presence in granules form and that causes scattering of light. The scattering effect may play a role for variability of the spectra from sample to sample. While in Fig. 2, the raw spectral profile of sample stems are shown, the spectra after SNV and MSC pre-processing are shown in Fig. 3 and 4, respectively.

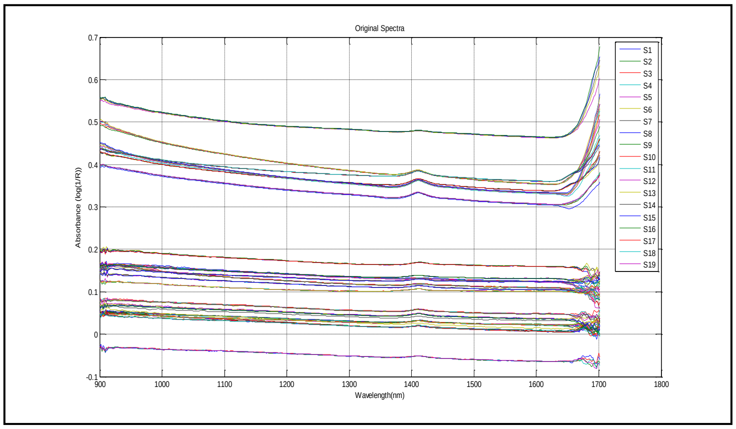

Fig. 2: Graphical representation of the near infrared wavelength vs Absorbance of the original data without any processing

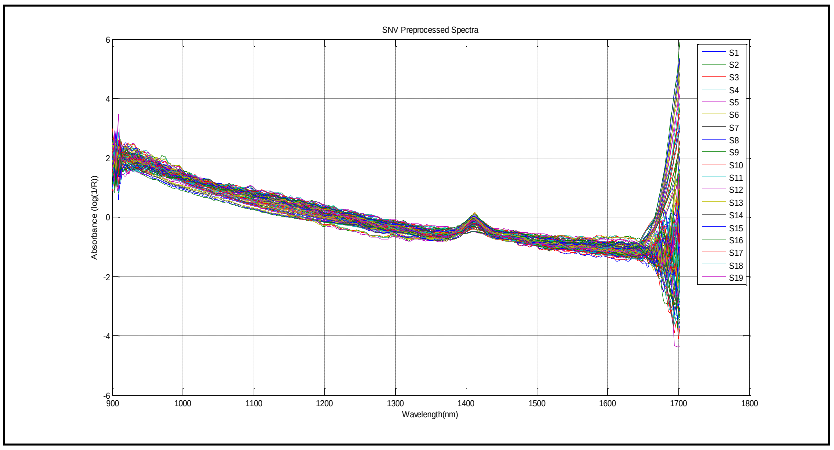

Fig. 3: Graphical representation of the near infrared wavelength vs Absorbance of the data pre-processed using standard normal variate

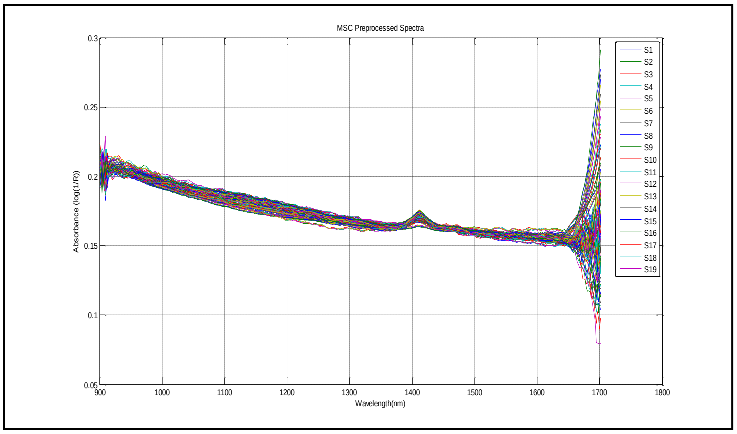

Fig. 4: Graphical representation of the near infrared wavelength vs Absorbance of the data pre-processed using multiplicative scatter correction

E) Data Analysis Method

Partial least squares (PLS) regression is a statistical method that bears some relation to principal components regression. It finds a linear regression model by projecting prediction of the training set which explains the maximum multi-dimensional variance direction in the Y space. Here PLS regression has been used to predict the reference sucrose content values of the testing set.

Predicted variables and the observable variables to a new space, PLS is used to find the fundamental relations between two matrices (X and Y), i.e. a latent variable approach to build a model of the covariance structures in these two spaces.

F) Software

Preprocessing techniques and PLS regression algorithms were implemented in MatlabV10.0 (Mathworks Co., USA).

III. RESULT AND DISCUSSION

A) Spectra Investigation

Scrutinized from the spectra of the original data (Fig. 1a), it is revealed that some intensive spectral peaks were present mainly in 934-1044 nm, 1100-1232nm and 1400-1600nm spectral region. These intensive peaks are caused by the stretch or deformation vibration of the hydric groups (such as C–H and O–H). The molecular structure of soil organic carbon, which is responsible for this type of spectral observation, contains many hydric groups. NIR spectra in the region of 924-1045nm, 1100-1222nm and 1390-1600nm contain much chemical information about organic carbon and therefore, selected for calibration of PLS model. Thus, for each spectrum, we have considered only 262 data points in this analysis.

B) Calibration of model

The performance of a PLS algorithm depends significantly on PLS factors and the treatment of spectra. The data obtained from the spectrometer were processed using two pre-processing techniques as mentioned earlier section. Pre-processed data were subjected to PLS regression algorithm to build the calibration model. During the model calibrations following steps were followed:

i) Comparing the result with different pre-processing techniques.

ii) Performance evaluation of PLS.

iii) Finalize the PLS component by comparing the maximum accuracy or least root mean square error of cross validation (RMSECV).

The performance of the final PLS model was evaluated in terms root mean square error of cross-validation (RMSECV), the root mean square error of prediction (RMSEP) and the correlation coefficient (R). For RMSECV, a leave-one-sample-out cross-validation was performed: the spectrum of one sample of the training set was removed and a PLS model was built with the remaining spectra of the training set. The left-out sample was predicted with this model and the procedure was re-peated by leaving out each of the samples of the training set.

C) Results of calibration model

The number of PLS factors were optimized by determining the lowest root mean square error cross-validation (RMSECV) by raw spectra and two different pre-processing methods, SNV and MSC.

Fig. 5(a) represents the comparison of RMSECV of sample-one-out cross-validation results for 19 samples obtained by raw spectra analysis and SNV and MSC pre-processed spectral analysis for component number 1 to 20.

The PLS algorithm is usually a linear regression, and the PLS model with low number of PLS factors was not enough to explain the complex correlation between the NIR spectral data and organic carbon content. The lowest RMSECV of 0.342 and lowest residual variance of 2.63 x 10^{-2} were obtained by raw spectral analysis when 11 PLS components were taken. The lowest RMSECV of 0.446 and lowest residual variance of 4.00 x 10^{-2} with the SNV preprocessing method was obtained when the number of PLS components was 12 and 10 respectively. On the other hand the lowest RMSECV of 0.460 and lowest residual variance of 4.30 x 10-2 with the MSC preprocessing method was obtained when the number of PLS components was 5 and 10 respectively. The comparative bar charts for RMSECV and residual variance with different components are given in Fig. 5(a) and 5(b).

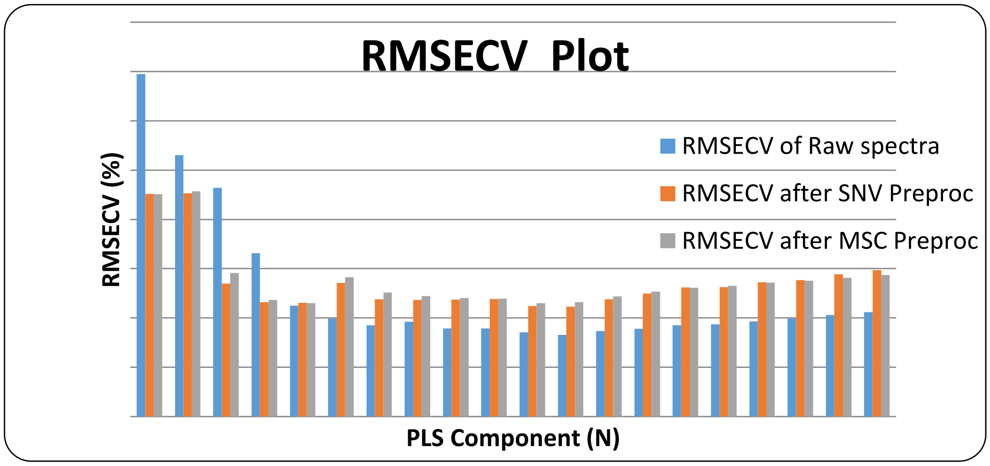

Fig 5(a). RMSECV comparison of sample-one-out cross validation results for 19 samples obtained by Raw, SNV and MSC for component number 1 to 20.

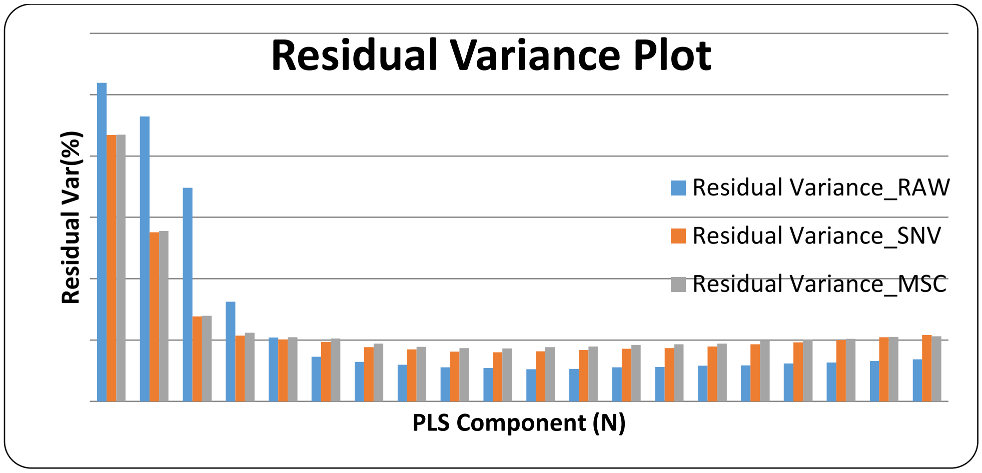

Fig. 5(b): Residual variance comparison of sample-one-out cross validation results for 19 samples obtained by Raw, SNV and MSC for component number 1 to 20.

However, as the lowest values of RMSECV and residual variance are obtained with the raw spectral processing method and 11 PLS components, this model is considered for the prediction. In this paper, out of 19 samples, 15 samples were considered for the final model building and 4 samples out of 19 samples, which were not included in final training set, considered for testing set. Table 1 shows the prediction results of 4 test samples by raw spectral analysis method with component number 11.

Table I: The Prediction Results Of 4 Test Samples By Raw Spectra Analysis With Component Number 11

Serial No	**Ref Organic Carbon value**	**Predicted Organic Carbon Value**	**% Accu**	**Avg Acc**	**RMSEP**	**R**
1	0.38	0.407	93.007	87.871	0.1481	0.942
2	1.93	1.634	84.675			
3	0.66	0.775	82.595			
4	1.09	1.186	91.206			

From the Table 1, it was found that the average accuracy of prediction was 87.87% in case of raw spectral analysis with component no. 11. It was also observed that the root-mean square error of prediction (RMSEP) of 4 samples with component number 11 was sufficiently low i.e 0.148 and correlation coefficient (R) was 0.942 which exhibits good correlation between training and testing set.

D) CONCLUSION

The overall results demonstrate that OC in soil can be predicted with more than 87% of accuracy. Predictive model was successfully built by means of PLS with raw spectra. The model was optimized by root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP), and correlation coefficient (R). This study demonstrated that NIR measurements could be used for measurement of soil fertility and other organic matters like Nitrogen (N), Phosphorus (P), and Potassium (K) etc. Also, this procedure may be used as an analytical tool to predict Carbon concentration in soil replacing the conventional chemical method.

ACKNOWLEDGMENT

The research work has been carried out in collaboration with Centre for Development of Advanced Computing (CDAC), Kolkata, India and Vivekananda Institute of Biotechnology, Sri Ramkrishna Ashram, West Bengal, India.

REFERENCE

1) Brevik, E.C., Cerdà, A., Mataix-Solera, J., Pereg, L., Quinton, J.N., Six, J., Van Oost, K., The interdisciplinary nature of soil. Soil 1, 117–129 (2015).
2) Keesstra, S.D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., Montanarella, L., Quinton, J.N., Pachepsky, Y., van der Putten, W.H., Bardgett, R.D., Moolenaar, S., Mol, G., Jansen, B., Fresco, L.O., The significance of soils and soil science towards realization of the United Nations sustainable development goals. Soil 2, 111–128 (2016).
3) Lal, R., Soil carbon sequestration to mitigate climate change. Geoderma 123 (1–2), 1–22 (2004).
4) L. M. Chabala, A. Mulolwa, and O. Lungu, "Application of Ordinary Kriging in Mapping Soil Organic Carbon in Zambia," Pedosphere, vol. 27, no. 2, pp. 338-343, Apr, 2017.
5) M. Knadel, R. Gislum, C. Hermansen, Y. Peng, P. Moldrup, L. W. de Jonge, and M. H. Greve, "Comparing predictive ability of laser-induced breakdown spectroscopy to visible near-infrared spectroscopy for soil property determination," Biosystems Engineering, vol. 156, pp. 157-172, Apr, 2017.
6) Cotching W E, Oliver G, Downie M, Corkrey R, Doyle R B. Land use and management influences on surface soil organic carbon in Tasmania. Soil Res. 51: 615–630 (2013).
7) Bonfatti, B.R., Hartemink, A.E., Giasson, E., Tornquist, C.G., Adhikari, K., Digital mapping of soil carbon in a viticultural region of Southern Brazil. Geoderma 261, 204–221(2016).
8) Ellert, B.H., Janzen, H.H., VandenBygaart, A.J., Bremer, E., Measuring change in soil organic carbon storage. In: Carter, M.R., Gregorich, E.C. (Eds.), Soil Sampling and Methods of Analysis. CRC Press, Boca Raton, FL, pp. 25–38(2008).
9) Lorenz, K., Lal, R., The depth distribution of soil organic carbon in relation to land use

and management and thepotential of carbon sequestration in subsoil horizons. Adv. Agron. 88, 35–66 (2005).

10) Sombroek, W.G., Nachtergaele, F., Hebel, A., Amounts, dynamics and sequestering of carbon in tropical and subtropical soils. Ambio 22 (7), 417–426(1993).

11) Batjes, N.H., Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 47 (2), 151–163(1996).

12) Cohen, M.J., Prenger, J.P., DeBusk, W.F., Visible-Near infrared reflectance spectroscopy for rapid, non-destructive assessment of wetland soil quality. Journal of Environmental Quality 34, 1422–1434(2005).

13) Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131, 59–75(2006).

14) Y. Huang, D. Wang and J. Duan, Mod. Instruments, 5 (2011) 13.

15) W. Z. Lu, H. F. Yuan and G. T. Xu, Chin Petrochemical Press, 2007.

16) J. Zhang, X. D. Chen and Z. G. Piao, Optics Precision Eng, 16 (2008) 986.

17) B. Stenberg, R. A. V. Rossel, A. M. Mouazen, and J. Wetterlind, "VISIBLE AND NEAR INFRARED SPECTROSCOPY IN SOIL SCIENCE," Advances in Agronomy, Vol 107, vol. 107, pp. 163-215, 2010.

18) Somdeb Chanda, Ashmita De, Bipan Tudu, Rajib Bandyopadhyay, Ajanto Kumar Hazarika, Santanu Sabhapondit, B.D.Baruah, Pradip Tamuly, Nabarun Bhattachryya“Prediction of Polyphenol Content in Tea Leaves Using NIR Spectroscopy” , 2016 International Conference on Intelligent Control Power and Instrumentation (ICICPI), 21st-23rd October 2016. DOI: 10.1109/ICICPI.2016.7859672. http://ieeexplore.ieee.org/document/7859672/.

19) M. Dhanoa, S. Lister, R. Sanderson, and R. Barnes, “The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra”, J. Near Infrared Spectrosc., vol. 2, pp. 43-45, 1994

20) Bhuyan, L. P., Hussain, A., Tamuly, P., Gogoi, R. C., Bordoloi, P. K., Hazarika, M.: Chemical characterization of CTC black tea of northeast India: Correlation of quality parameters with tea tasters’ evaluation. Journal of the Science of Food and Agriculture, 89, 1498–1507, (2009).

21) Sorol, N., Arancibia, E., Bortolato, S., Alejandro, C.: Visible/near infrared-partial least-squares analysis of Brix in sugar cane juice, A test field for variable selection methods. Olivieri hemometrics and Intelligent Laboratory Systems, 102, 100–109(2010).