VIS-NIR REFLECTANCE SPECTROSCOPY FOR ASSESSMENT OF SOIL ORGANIC CARBON IN A RICE-WHEAT FIELD OF LUDHIANA DISTRICT OF PUNJAB

Bhabani Prasad Mondal^{1*2}, Bharpoor S. Sekhon¹, Rabi N. Sahoo² and Priya Paul³

^{1*}Department of Soil Science, Punjab Agricultural University, Ludhiana, Punjab-141004,India-bpmondal1@gmail.com, bsekhon@pau.edu

²Division of Agricultural Physics, Indian Agricultural Research Institute (IARI), New Delhi-110012, India-rnsahoo-iari@gmail.com

³Division of Plant Physiology, Indian Agricultural Research Institute, New Delhi-110012, India-paulp3399@gmail.com

Commission III, WG III/10

KEY WORDS: PLSR, Reflectance Spectroscopy, Rice-Wheat, RMSEP, RPD, SOC

ABSTRACT:

Soil organic carbon (SOC) is a crucial indicator of soil fertility, maintaining soil health and sustaining the productivity of agroecosystem. Rapid, reliable and cost effective assessment of soil properties specially for SOC is important for monitoring soil fertility status along with soil health. Conventional chemical analysis of any soil property is hazardous, tedious and time consuming. So, the visible near infrared (VIS-NIR) reflectance spectroscopy can provide an effective alternative technique for rapid and ecofriendly measurement of soil properties. In view of this, a key soil fertility parameter SOC was examined through diffuse reflectance spectroscopy. Georeferenced surface soil samples (0-15cm) were collected from a rice-wheat field of the study area for both chemical and spectral analysis. A viable statistical technique namely partial least square regression (PLSR) technique were used to correlate the measured properties with soil reflectance spectra and for developing spectral model. The predictive performance of newly developed spectral model was evaluated through different reliable indices like root mean square of error of prediction (RMSEP), coefficient of determination (R²) and ratio of performance deviation (RPD). The result showed that the R² value for SOC is 0.44, RMSEP is 0.07 and the RPD value is 1.57 in the validation dataset. The RPD value indicating that SOC can be reliably predicted using the hyperspectral model or reflectance analysis. So, this hyperspectral modeling technique can be successfully employed for monitoring soil health as well as for sustainable agriculture.

1. INTRODUCTION

Soil organic carbon (SOC) is a key component of soil fertility and it plays a crucial role in affecting the plant growth and yield as an important energy source as well as a nutrient (Andrews et al., 2004). It plays a fundamental role in enhancing soil fertility, maintaining soil health and sustaining the productivity of the agro-ecosystem. However, the excessive chemical and inorganic chemical fertilization affects soil heath, reduces its fertility status, increases economic investment and leads to the development of fragile agro-ecosystem along environmental pollution (Savci, 2012). It is the high time to apply organic matter (OM) to soil, as carbon is a key component of OM along with other nutrients through reasonable fertilization for improving the soil quality, health and fertility status. Accurate and cost effective prediction of SOC is necessary for sustainable development of agro-ecosystems. Conventional chemical analysis for quantifying any soil property is destructive, tedious, time consuming and also hazardous to the environment and additionally these traditional analysis cannot meet the criteria for soil quality assessment for precision farming. So, in the recent decade, the visible nearinfrared (VIS-NIR) reflectance spectroscopy analysis approach is a potential and viable alternative to the traditional laboratory analysis for the accurate, non-destructive, cost-effective, rapid, reproducible and simultaneous assessment and successful prediction of several soil properties. This analytical technique is very much useful for soil spectral classification through

identifying the major absorption feature in the spectral signal or reflectance spectra at specific wave band and predicting the soil attributes through constructing regression model. The assessment of SOC through reflectance spectroscopy mainly depends on overtones, fundamental vibrations of individual molecular bond, absorption of incidental radiation at different wave bands due to presence of organic matter and physical and chemical chromophores (Stenberg et al., 2010). Clark et al., 1990 exhibited broad sensitive band of soil organic matter extending from visible to near-infrared as a consequence of overtone and combined absorption of different bonds like C-H, N-H and O-H bonds present in OM. Various statistical techniques like principal component regression (PCR) (Chang et al., 2001), multiple linear regression (MLR) (Vasques et al., 2008), partial least square regression (PLSR) (Vasques et al., 2008) have been applied to hyper-spectral data to model and predict the SOC. PLSR is the most popular and commonly used linear calibration technique for spectral modeling of SOC due to its capacity to model the linear relationship between soil parameter and raw spectra through resolving the multicollinearity exist within the raw spectral data. Several spectral preprocessing techniques like smoothening, continuum removal, normalization spectral derivative etc. have been applied to transform the soil spectra, remove nose and extract useful spectral information and features for improving the prediction accuracy of SOC using hyper-spectral data. Different spectral indices like root mean square of error of prediction (RMSEP), coefficient of determination (R2) and ratio of performance deviation (RPD) were used by several researchers for comparing and evaluating the predictive performance of different spectral models. (Sarathjith et al., 2016, Yu et al., 2016). A high prediction accuracy indicates a good agreement between measured and predicted parameter under study specially in validation dataset. Chang et al. (2001) categorized the prediction capability of the spectral model based on the value of RPD and this categorization stated that if the RPD value of any soil property were more than 2.0, the prediction of that property would be excellent by the model. The RPD value in between 1.4 to 2.0 indicating the prediction of soil property with an acceptable accuracy and the RPD value less than 1.4 indicates poor prediction of the property by the existing model. The primary goal of any prediction model is to achieve good prediction accuracy with respect to any soil parameter.

The objective of this present study is (i) to investigate the predictive ability of the VIS-NIR reflectance spectroscopic technique for SOC in a rice-wheat field of the study area and (ii) to evaluate the predictive performance of the newly developed spectral model based on PLSR algorithm through using different indices like R², RMSEP and RPD.

2. STUDY AREA

The experiment was conducted at Ladian village of Ludhiana district of Punjab, India and the geo-referenced surface soil samples (0-15 cm depth) were collected from a rice-wheat field of the study area using global positioning system (GPS). This land use system (30⁰ 57' 330"N and 75⁰ 47' 675"E) followed rice-wheat cropping system for the past ten years in the same field. The study site showed high organic carbon (OC) content and fine (almost clayey) texture (confirmed by particle size distribution analysis, done by International pipette method given by Piper, 1966).

The purpose of this study was to evaluate the potential of VIS-NIR reflectance spectroscopy for the assessment and spatial prediction of SOC of that major land use system of the study site.

3. MATERIALS AND METHODS

After collection of soil samples, they were air dried properly under shaded condition, grounded in wooden log, and passed through 2 mm sieve for preparing fine samples. The air-dried samples were used rather than moist samples for getting optimal and accurate result in spectral modeling. Then the prepared and processed samples were divided into two parts; one part is for laboratory chemical analysis and another part for spectral analysis using ASD Field spec-spectroradiometer (350-2500 nm).

3.1 Chemical Analysis

Soil organic carbon (SOC) was determined chemically following the wet chromic acid oxidation protocol given by Walkley and Black in 1934. First, the organic matter content of the soil was digested with the mixture of potassium dichromate (K₂Cr₂O₇) and concentrated sulphuric acid (H₂SO₄). After the reaction excess dichromate was reduced through titrating with ferrous ammonium sulphate, (FAS) and the reduced dichromate corresponded to the organic carbon content in the soil sample.

3.2. Spectral analysis

Reflectance spectra (spectral range: 350-2500 nm) of the airdried and grounded soil samples was taken with the help of ASD Field spec spectroradiometer (Analytical Spectral Devices Inc., Boulder, Colorado, USA) with a sampling interval of 1 nm. Each sample was distributed homogeneously on the black colored polythene petridis and the spectral reading was obtained with a contact probe (Contact probe, Analytical Spectral Devices Inc., Boulder, Colorado, USA) having own light source ensuring direct contact with the soil sample to exclude the noise. A white spectralon was used to calibrate the instrument by taking thirty white reflectance spectra before initiating the spectral measurement of the sample. Averaging thirty spectra for each sample helped to minimize the noise of the instrument.

3.3 Preprocessing of spectral data

Several mathematical calculations and procedures are involved in preprocessing of spectral data for transforming spectra, removing noise and other physical variations due to scattering of light from spectra and to identify the features of interest i.e. the diagnostic features in the spectra before applying the calibration model. The preprocessing techniques include smoothening, averaging, resampling, normalization, derivation and other spectral transformations for obtaining the best result from the spectral data. (Vasques et al., 2008, Nawar et al., 2016 & Dotto et al., 2018). The huge volume of spectral data points (2151 points in the spectral range of 350 to 2500 nm) was reduced by resampling using the window size of 10 nm. Then the first order derivative of the 10 nm raw reflectance spectra was done by applying Savitzky-Golay (Savitzky and Golay, 1964) first derivative with a first order polynomial equation or transformation in Unscrambler 10 software (CAMO software, Norway).

3.4 Statistical modeling (descriptive statistics and multivariate technique) and evaluation indices

Classical or descriptive statistical analysis was carried out using SAS software. Several parameter of classical statistics like mean, median, mode, percent coefficient of variation (% CV), skewness and kurtosis was calculated for characterizing the SOC in the study area by descriptive statistics.

One of the important and most commonly used statistical modeling approach namely partial least square regression (PLSR) model has been implemented to develop the spectral model with the dataset. Two third of the whole sample was kept for calibrating the model and rest one third for validating the spectral model. This model development was done using PLSR algorithm in Unscrambler 10 software.

The prediction accuracy of the newly developed spectral model was evaluated through three different statistical indices: coefficient of determination (R²) (Eqn.1), root mean square error of prediction (RMSEP) (Eqn.2) and ratio of performance deviation (RPD) (Eqn.3). These evaluation indices are described in detail with mathematical form (equation).

$$R^{2} = 1 - \left(\frac{\sum_{l}^{n}(Y_{pred} - Y_{mear})^{2}}{\sum_{l}^{n}(Y_{l} - Y_{mean})^{2}}\right)$$
(1)

$$RMSEP = \sqrt{\frac{\sum_{i}^{n}(Y_{pred} - Y_{meas})^{2}}{n-1}}$$
(2)

$$RPD = SD/RMSEP$$
(3)

Where Y $_{pred}$ = predicted values,

Y mean = mean of measured values,

 Y_{meas} = measured values,

n = number of measured or predicted values (number of samples),

SD = standard deviation.

 R^2 is most commonly used as a measure of goodness of fit of the model to give the percent variance provided by the model. RMSEP measure the difference between observed and predicted values in the model and smaller difference indicates a good model accuracy for prediction. RPD also indicates the prediction accuracy by the model.

4 RESULTS AND DISCUSSIONS

4.1 Descriptive statistics

The descriptive statistical parameters of SOC is represented in Table 1. In rice-wheat system, it hovered around a mean value of 0.87 % and ranged from 0.74 to 1.05 % with a value of CV of 10.38 %. The SOC content in this cropping system was high due to higher amount of clay and silt content in soil and also the external addition of organic matter to the soil.

Statistical parameters	Values
Mean	0.87
Minimum	0.74
Maximum	1.05
Median	0.88
CV%	10.38
Skewness	-0.85
Kurtosis	3.27

Table 1. Descriptive statistics of SOC in rice-wheat cropping system.

4.2 Characteristics of spectral reflectance curve of soil samples

The spectral reflectance characteristics curve showed variation due to its various chemical, mineralogical, physical composition along with varying level of moisture and organic matter content of the soil samples. The reflectance spectra (Figure 1) showed the characteristics absorption features at around 1400, 1900 and 2200 nm due to presence of moisture and these are known as water absorption bands (mainly 1400 and 1900 nm) and the hydroxyl (-OH) absorption band at 2200 nm indicating the decrease in the values of reflectance around these specific wavelengths. Clark, (1999) reported that soil spectra could be characterized by the water absorption bands at 1400 and 1900 nm and dominant hydroxyl absorption band at 2200 nm. Dematte et al. (2017) mentioned the same findings due to vibration of water molecules and presence of hydroxyl groups and they also mentioned the presence of kaolinite and other phyllosilicate minerals at 2200 nm band.

The raw reflectance spectra of 10 nm and first derivative spectra of soil sample were represented in the Figure 1 and 2.

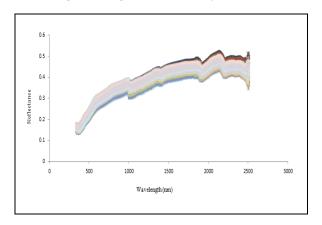


Figure 1. Reflectance spectra (10 nm) of the soil samples

The first derivative spectra showed major diagnostic characteristics in the spectral curve in the form of either positive or negative peaks in the curve and these were related to the SOC content of the soil. So many diagnostic peaks in the first derivative spectra were very much useful for characterizing the SOC as they contained more useful information than raw spectra.

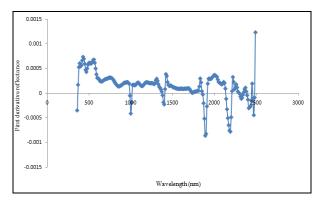


Figure 2. First derivative reflectance spectra of soil sample

It was found that with increasing organic carbon content in soil, there was a decrease in the spectral reflectance due to the presence and complexity of so many functional groups in the organic matter and this make it spectrally active in NIR region and the same result was reported by Bowers and Hanks (1965). Different complex composition of organic matter like cellulose, hemicelluloses, pectic substances and several functional groups like hydroxyl (-OH), carboxyl (-COOH), amine (-NH₂) groups etc. along with varying moisture content are mainly responsible for creating absorption bands in the spectral reflectance characteristics curve.

4.3 Correlation of reflectance spectra with SOC

The statistical correlation between SOC and raw reflectance spectra was carefully studied and this correlation showed both positive and negative correlation at different wavelength in the spectral range of 350 to 2500 nm. SOC was negatively correlated up to 1700 nm and then positively correlated up to 2500 nm and this correlation was showed in Figure 3.

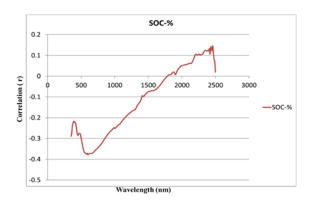


Figure 3. Correlation of SOC with reflectance spectra at different wavelengths

4.4. Partial least square regression (PLSR) modeling

The most popular linear multivariate technique i.e. PLSR has been utilized for developing the spectral model considering the relationship between SOC and soil reflectance spectra. The PLSR modeling technique is most popular due to easy derivation and interpretation of this model and its insensitivity towards the data having multi-collinearity.

PLSR model was utilized to calibrate the SOC with spectral reflectance data of first derivative in the calibration dataset. The statistical components of the calibration dataset like no of factors, R² and RMSE values etc. for SOC were taken under consideration for calibration purpose. In calibration dataset, R² and RMSE values were 0.53, 0.06 respectively, 0.41, and 0.06 respectively when internal cross validation (leave one out validation) was done using the same samples. SOC exhibited moderately higher R² value and relatively lower RMSE value. From the calibration dataset, it was clear that the SOC was moderately well calibrated with the spectral model. The better calibration could be obtained through getting more variation in spectra of soil sample and by taking more number of samples for analysis. The outliers' samples should be removed for improving the prediction accuracy of the calibrating model. The scatter plot of measured vs. predicted values of SOC was depicted in the Figure 4.

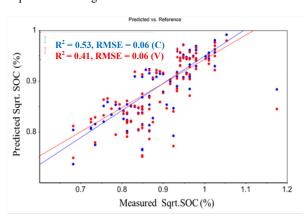


Figure 4. Scatter plot of SOC in calibration dataset

The statistical indices of the validation dataset like R², RMSEP and RPD were employed to assess the prediction accuracy of the newly developed spectral model for SOC prediction. Greater R² and RPD values and lower RMSEP value are always desirable for a good prediction model. Chang et al. (2001) gave the category of prediction based on RPD values and the

category was RPD>2.0 (good prediction), RPD =1.4-2.0 (satisfactory prediction) and RPD <1.4 (poor prediction). Our study revealed the lower R^2 and RMSEP and moderately higher RPD values for SOC ($R^2_{\rm val}$ = 0.44, RMSEP =0.07 and RPD = 1.57) in the validation dataset of the model. Based on the classification of Chang et al. (2001), the RPD value for SOC was in between 1.4 to 2.0 in our study; so our spectral model could satisfactorily predict this parameter through PLSR modeling technique. The scatter plot between measured and predicted (square root transformation) values of SOC in validation dataset was represented in Figure 5.

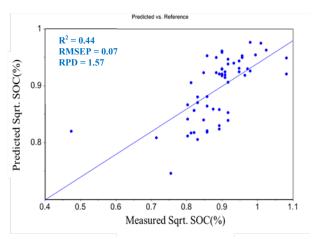


Figure 5. Scatter plot of SOC in validation dataset

The prediction capability of the spectral model for the prediction of any soil parameter depends on the relationship of that parameter with spectrally active compounds like clay, OM content etc, and other physical and chemical chromophores present in the soil as these possess primary response in near infrared region of the spectrum. Our study produced a similar type of result to a small extent with a study conducted by Araujo et al. (2014), who improved the prediction performance of the model with a large tropical VIS-NIR spectral library of Brazilian soil with R^2_{val} = 0.60 and RMSEv of 0.55 % for soil organic matter by taking large number of soil samples (7172 samples) and by applying PLSR algorithm. Several workers like Viscarra Rossel and Behrens (2010) found higher R² value in the PLSR prediction model ($R^2 = 0.82$ and RMSE = 0.96). Similar higher results were also obtained by Kuang et al. (2015). But our study exhibited moderately lower R² value; RMSE value was much lower than the other findings. Lower RMSE value is a good indication of better performance of the model. Both high and low values of evaluation indices for SOC prediction through spectral modeling approach using PLSR technique were reported by several studies. Lower R² value might be due to less variation in spectral data, higher spatial heterogeneity of the studied parameter under natural condition and the collection of soil samples from farmers' field under natural condition rather than green house or laboratory or any other controlled conditions. Our prediction through PLSR modeling could be improved by creating large variation in the spectra and by taking large spectral library of that area. Airdried ground soil samples were used for PLSR modeling with VNIR spectra for getting higher prediction accuracy rather than moist samples as the air-dried samples could exclude the influence of moisture to get optimal result. Several study reported the usefulness of using air-dried samples in spectral modeling (Jia et al., 2017). The smaller RMSE value (in spite of having lower R² value), indicated that the spectral model could predict the SOC with some uncertainty. Actually, differences in management within the same land use type caused uncertainty

in prediction and the variability in SOC concentration within the same land use influenced the accuracy of the prediction model (Jia et al., 2017). In spite of having variability in SOC concentration, the study showed that limited variability in the spectral dataset and it might be due to the relatively smaller size of the studied samples and these all could limit the potential performance of the PLSR model. Better performance of the PLSR model could also be obtained through increasing the number of dataset (calibration and validation dataset) for SOC prediction.

5 CONCLUSION

The study was mainly focused on the rapid and accurate assessment of SOC, an important soil fertility parameter through VIS-NIR reflectance spectroscopy instead of laborious and time-consuming traditional chemical analysis. Though the R² value was lower, lower RMSEP or RMSE and moderately higher RPD values confirmed the potential of the spectroscopic technique for spatial prediction of SOC in that cropping system. This study also confirmed the efficiency and accuracy of PLSR modeling in quantifying the SOC in rice-wheat land use system. Other multivariate techniques could be exploited further to find the best technique for predicting SOC in that land use system through spectral modeling approach. One of the potential limitation of this spectral approach for SOC prediction is the spatial heterogeneity and variability of soil under natural ecosystem, which results poor or moderate performance of the model. Therefore, large number of soil samples are required to increase the number of both calibration and validation dataset and additionally building of a good and comprehensive spectral library at a local scale is needed for improving the performance of the spectral model in quantifying SOC at a particular area.

ACKNOWLEDGEMENT

The first author thanks the Indian Council of Agricultural Research (ICAR) for providing Junior Research Fellowship (JRF) for conducting the research work during M.Sc. degree (Soil Science) at Punjab Agricultural University (PAU), Ludhiana, and Punjab, India. The first and second authors are thankful to Indian Agricultural Research Institute (IARI) for providing spectroradiometer for spectral analysis.

REFERENCES

Andrews, S.S., Karlen, D.L., Cambardella, C.A., 2004. The soil management assessment framework: a quantitative soil quality evaluation method. *Soil Science Society of America Journal*. 68, pp. 1945–1962. http://dx.doi:10.2136/sssaj2004.1945.

Araújo, S.R., Wetterlind, J., Demattê, J.A.M., Stenberg, B. 2014. Improving the prediction performance of a large tropical Vis-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. *European Journal of Soil Science*, 65, pp. 718–729. http://dx.doi.org/10.1111/ejss.12165.

Bowers, S.A. and Hanks, R. J., 1965. Reflectance of radiant energy from soils. *Soil Science Society of America Journal*, 100, pp. 130-138.

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, pp. 480–490. http://doi:10.2136/sssaj2001.652480x.

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G., & Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysics Research*, pp. 12653–12680. https://doi.org/10.1029/JB095iB08p12653.

Clark, R.N., 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. In: Rencz, N. (Ed.), Remote Sensing for the Earth Sciences: *Manual of Remote Sensing*. John Wiley & Sons, New York, pp. 3–52.

Demattê, J.A.M., Horák-Terra, I., Beirigo, R.M., Terra, F. da S., Marques, K.P.P., Fongaro, C.T., Silva, A.C., & Vidal-Torrado, P., 2017. Genesis and properties of wetland soils by VIS-NIR-SWIR as a technique for environmental monitoring. *Journal of Environmental Management*. 197, pp. 50–62. http://dx.doi.org/10.1016/j.jenvman.2017.03.014.

Dotto, A.C., Dalmolin, R.S.D., ten Caten, A. and Grunwald, S., 2018. A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. *Geoderma*, 314, pp. 262-274. https://doi.org/10.1016/j.geoderma.2017.11.006.

Jia, X., Chen, S., Yang, Y., Zhou, L., Yu, W. & Shi, Z., 2017. Organic carbon prediction in soil cores using VNIR and MIR techniques in an alpine landscape. *Scientific reports*, 7(1), pp.2144. http://doi: 10.1038/s41598-017-02061.

Kuang, B., Tekin, Y., Mouazen, A.M., 2015. Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. *Soil & Tillage Research*, 146 (Part B), pp. 243–252. http://dx.doi.org/10.1016/j.still.2014.11.002.

Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., Mouazen, A.M., 2016. Estimating the soil clay content and organic matter by means of different calibration methods of Vis-NIR diffuse reflectance spectroscopy. *Soil & Tillage Research*, 155, pp. 510–522. https://doi.org/10.1016/j.still.2015.07.021.

Piper, C.S., 1966. Soil and Plant Analysis. Hans Publisher, Bombay. pp. 1-356.

Savci, S., 2012. An agricultural pollutant: chemical fertilizer. *International Journal of Environmental Science and Development.* 3(1), pp. 77-79. https://DOI: 10.7763/IJESD.2012.V3.191.

Savitzky, A., & Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedure. *Analytical Chemistry*, 36(8), pp. 1627-1639. https://*DOI*: 10.1021/ac60214a047.

Sarathjith, M., Das, B., Wani, S., Sahrawat, K.L., Gupta, A., 2016. Comparison of data mining approaches for estimating soil nutrient contents using diffuse reflectance spectroscopy. *Current Science*, 110 (6), pp. 1031–1037. ISSN 0011-3891.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, 107, pp. 163–215. https://doi.org/10.1016/S0065-2113(10)07005-7.

Vasques, G.M., Grunwald, S., Sickman, J.O. 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. *Geoderma* 146, pp. 14–25. https://doi.org/10.1016/j.geoderma.2008.04.007.

Rossel, R.V. and Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158(1-2), pp. 46-54. https://doi.org/10.1016/j.geoderma.2009.12.025.

Walkey A, Black IA. 1934. An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science*, 37, pp. 29-37.

Yu, X., Liu, Q., Wang, Y., Liu, X., Liu, X., 2016. Evaluation of MLSR and PLSR for estimating soil element contents using visible/near-infrared spectroscopy in apple orchards on the Jiaodong peninsula. *Catena*, 137, pp. 340–349. https://doi.org/10.1016/j.catena.2015.09.024.