IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 23, Issue 1, Ser. I (Jan. – Feb. 2021), PP 14-18
www.iosrjournals.org

Spectral data analysis methods for soil properties assessment using remote sensing

Vipin Y. Borole,
Department.of Computer Science & Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India

Sonali B. Kulkarni
Department.of Computer Science & Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India

Abstract: *In the remote sensing techniques reflectance spectroscopy is one of most important technique is used for soil properties assessment, this article attempts to provide a review on fundamental concepts of reflectance spectroscopic techniques. In the spectroscopy technique, spectral signatures are more complicated to process. When proper methods are used in soil properties assessment it gives most accurate results. In this paper various process and methods are explained which are required for data collection and data processing of spectral data like sampling, preprocessing, analysis and classification. Their applications as well as exploring the role of Near-infrared reflectance spectroscopy that would be used for monitoring, mapping soil characteristics and analysis of soil properties. Spectroscopy techniques gives the fast and accurate results as compare to traditional methods because minimum soil sample preparation, no chemical analysis are required, minimum time require to process the acquired data. Moreover, this article explores the potentiality of predicting soil properties based on spectroscopic measurements.*

Keywords: *Spectroscopy, Soil, PLSR, PCA, LDA*

Date of Submission: 14-01-2021 Date of Acceptance: 29-01-2021

I. Introduction:

Remote sensing techniques for soil application have expanded ample concentration, because these techniques have been shown to generate faster and cheaper characterizations. The application of remote sensing techniques in soil studies started in the 1960s and developed for various applications, including fast and nondestructive quantification of soil properties [1]. In this domain spectroscopy technique is one of most important technique is used for soil properties assessment and it gives most accurate results. But the proper process and methods are required for data collection and data processing of spectral data like sampling, preprocessing, analysis and classification. but the number of algorithms, methods and procedure are used by many researchers. In the literature review the different methods and model are used for getting better results. Since traditional methods of soil analysis are costly and time-consuming, in recent decades researchers used novel methods such as Remote Sensing and reflectance spectroscopy to estimate soil properties One of the limiting factors in the evaluation of soil properties by spectroscopy is the identification of preprocessing methods in noise and error elimination and determining the appropriate regression model to estimate these properties. Therefore, proper regression and pre-processing methods are required to determine soil properties using soil reflectance.

Quantitative prediction of soil properties such as; salinity, organic carbon, soil moisture and heavy metals can be conducted using various calibration models – such models were developed depending on the measured soil laboratory analyses data. The most common used models are stepwise multiple linear regression (SMLR), partial least squares regression (PLSR), multivariate adaptive regression splines (MARS), principal component regression (PCR) and artificial neural networks (ANN). Those methods are required to quickly and accurately measure soil characteristics at field to improve soil management and conservation at local and regional scales. Visible-Near Infra-Red (VIS-NIR) has been recommended as a quick tool for mapping soil properties. Furthermore, VIS-NIR reflection spectroscopy reduces the cost and time, therefore has a great ability and potential use as a rapid soil analysis for both precision soil management and assessing soil quality [2]. When soil absorbs the energy from energy source then it reflects that radiation differently in various bands. The soil spectral signature is the radiation reflected as a function of the wavelength [3]. Recent researchers have indicated that visible and near-infrared, diffuse reflectance spectroscopy (VINR-DRS) and specific band from hyperspectral sensor data [4].

II. Spectral data acquisition using FieldSpec 4 Spectroradiometer:

FieldSpec 4 Spectroradiometer is scanned and absolute reflectance of samples is recorded for 350-2500 nm wavelength and collects spectra at the rate of 0.1 seconds per spectral scan. Spectral resolution is 3 nm @ 700 nm (350-1000 nm) and 10 nm/ 8 nm/ 30 nm @ 1400/2100 nm (100-2500 nm). Data sampling interval is 1.4 nm (350-100 nm) and 2 nm (1000-2500 nm) and spectral resolution of 3 nm at 700 nm and 10 nm at 1400 nm and 2100 nm. The total number of 2151 data points per spectrum using FieldSpec 4 Spectroradiometer (Analytical Spectral Devices Inc., Boulder, Colorado, USA). Its input is 1.5m permanent fiber optic cable and it is having 25 degree, 8 degree and 1 degree field-of-view (FOV). Fiber optic cable is flexible. FieldSpec 4 is a very flexible instrument, it can use in lab and on the field. It has some accessories like pistol grip tripod, lamp, reference panels, backpack, battery, AC power supply and laptop. The laptop is required to communicate the instrument through Ethernet wired interface or Wi-Fi Ethernet interface. It has some software such as RS3, ViewSpec Pro, Indico Pro and third party software. RS3 software is used for data collection and file saves as .asd extension [4-9].

III. Pre-processing:

ViewSpec Pro software is used for data analysis and pre-process data. It has the same tools like reflectance, absolute reflectance, log 1/ T, log 1/ R, 1st and 2nd derivative, and parabolic correction and so on. It can export .asd file into the ASCII code (text file) and shows the latitude and longitude coordinates of the field location (GPS) [10].

3.1. Near Infrared spectral data corrections

The near infrared spectroscopy can be employed to determine soil fertility parameters and certainly can be used in many field of multidisciplinary area of research and applications. The five pillars of NIRS comprise the skills of spectroscopy, calibration modelling or called as chemometrics, NIR instruments design and construction, standard chemical analysis and data transformation. All of them simultaneously necessary to make proper use of infrared technology in agricultural applications [10].

Acquired spectra data may contain irrelevant background information and noises which can affect and interfere soil fertility properties and other quality attributes information. Interfering spectra data such as light scattering and other random noises due to overheated sensors, instrumental parts and changes on physical sample properties need to be reduced or even eliminated in order to obtain more accurate, robust and stable calibration models database [12]. Therefore, it is strongly recommended to correct and enhance spectra data before establishing and constructing calibration models.

The first stage in spectra correction can be made by mean-centered (MC). This method is simply preferable due its effectively and focuses on differences between observations spectra data. Another commonly used spectra correction method is smoothing which can improve the visual aspect of the NIR spectra data. Further, normalization is also commonly used as spectra enhancement method. It can normalize data based on the mean and peak spectrum [13].

Multiplicative scatter correction (MSC) and standard normal variate (SNV) are the two most popular spectra correction and enhancement techniques [11]. MSC is used to compensate for additive and multiplicative effects in the spectral data caused by physical effects. This attempted to remove the effects of scattering by linearizing each spectrum to an ideal spectrum of the spectra data which is corresponds to the average spectrum [22], [23]. The SNV seek to normalize each individual spectrum of the spectra data to zero mean and unit variance. Apart from the different algorithm, obtained calibration models resulted from MSC and SNC spectra data are more-less similar. Spectra corrections can also be coupled among those methods to generate better calibration and prediction performances.

3.2. Principal Component Analysis (PCA):

The PCA is a method that transforms the original data into a set of new uncorrelated variables called principal components (PCs), thereby potentially reducing the number of variables. The value of the PCA is that the importance of each wavelength in each PC can be determined by the magnitude of the eigenvectors or factor loading, as the higher [14,15].

3.3. Linear Discriminant Analysis (LDA):

Linear Discriminant Analysis (LDA) is simple, mathematically robust and often produces models whose accuracy is good. It is based upon the concept of searching for a linear combination of variables (predictors) that best separates the two classes (targets). LDA is a transform-based method which attempts to minimize the ratio of within-class scatter to be between class scatter. The mathematical formulation involved in the theory of LDA is explained in the following sections. A within-class scatter matrix defines the scatter of samples around their respective class centers (means) [22].

IV. Prediction MODELS

Soil fertility properties and other quality attributes information are suppressed in the spectra data. In order to be able to reveal those information, prediction models must be established through a process called calibration. The methods adopted regression either as linear or non-linear regression approach from which spectra data (X-data) and actual measured quality parameters (Y-data) were regressed. Sample dataset used in calibration phase must be representative of the present and of future prediction samples [19]. It means that all expected sources of variability must be considered in both calibration and validation sample datasets.

4.1 Principal Component Regression (PCR) and Partial Least Square Regression (PLSR)

The two most common regression approaches used in calibration models development are principal component regression (PCR) and partial least square regression (PLSR). Both of them continue to be the workhorses for regression in near infrared spectroscopy applications [18]. The initial and essences of PCR and PLSR are linear methods which assuming linear relationship of the modelled soil quality parameters and fertility properties as a function of infrared spectra data variations. PCR and PLSR requires latent variables (LVs) in constructing calibration models which known as principal component (PC) in PCR, and factors in PLSR respectively [20].

Established calibration models database need to be evaluated and quantified by means of validation, either as cross validation or independent validation. Prediction performances can be evaluated using these following statistical parameters: the coefficient of determination (R2), coefficient of correlation (r) between predicted and measured soil fertility properties, prediction error which is defined as the root mean square error in calibration (RMSEC), root mean square error in cross validation (RMSECV), and the residual predictive deviation (RPD), defined as the ratio between standard deviation (SD) of the actual measured fertility parameters (N, P, K, pH, sand, silt, clay and SOC), and the RMSE of respective predicted soil properties. The higher value of RPD, the greater probability of models to predict quality parameters or chemical concentrations of soil samples dataset accurately [21].

Beside using linear regression approaches, calibration models database can also be constructed using non-linear regression methods such support vector machine regression (SVMR) or artificial neural networks (ANN) regression based. Both SVMR and ANN are more flexible calibration methods since they can handle linear and nonlinear relationship between the near infrared spectra data and corresponding soil fertility parameters or other chemical constituents of biological materials [22].

4.2. Support vector machine (SVM):

Support vector machine (SVM) is a particular class of algorithm, characterized by the use of kernels in its regression and classification approaches. At the beginning of initial development, the SVM method was normally employed for classification problems, but nowadays it also has been applied for regression purposes. In kernel-based methods like SVM, calibration model is carried out in a space of non-linearly transformed input data and it defined by the kernel function [16].

For the efficiency and simplification purposes, spectra data of near infrared spectroscopy can be subjected firstly onto the principal component analysis, then, score value of first seven latent variables were used as input data instead of all infrared spectra values. Generally, calibration models database generated from the non-linear regression approaches achieved better prediction performances than linear regression one [17].

Partial least squares regression (PLSR) and random forest (RF) methods were used to estimate soil salinity and based on their research results, the RF model had better performance than the PLSR model [23].

In the context of considering the complicated relationship between soil properties and their reflectance, it is necessary to use the spectroscopic method to determine the best statistical model for spectral analysis in different regions [11].

Prediction of different soil characteristics using spectral reflections depends on statistical models that explain the relationship between them, most common used models are stepwise multiple linear regression (SMLR), partial least squares regression (PLSR), multivariate adaptive regression splines (MARS), principal component regression (PCR) and artificial neural networks (ANN). SMLR is a statistical method of regressing multiple variables while simultaneously removing those that aren't important. The choice of predictive variables is carried out by an automatic. The variable that considered for addition to or subtraction from the set of explanatory variables in each step is based on a form of a sequence of F-tests or t-tests [24].

4.3. PLS and PCR Regression:

PLS is a statistical method that finds a linear regression by projecting the predicted variables and the observable variables to a new space. PLS regression is today most widely used in chemometrics, sensometrics, and other related areas.

PCR is a regression method that considers regressing the dependent variable on a set of independent variables based on a standard linear regression model, but uses PCA for estimating the unknown regression

coefficients in the model. Instead of regressing the dependent variable on the explanatory variables directly, the principal components of the explanatory variables are used as regressors. PCR is some kind of a normalized procedure [25].

The principal components with the higher variances are selected as the regressors. The major use of PCR lies in overcoming the multicolinearity problem which arises when two or more of the explanatory variables are close to being collinear. PCR can result in dimension reduction through substantially lowering the effective number of parameters characterizing the underlying model. PCR can lead to efficient prediction with the appropriate selection of the principal components to be used for regression [26].

V. Conclusion:

In the remote sensing domain, spectroscopy techniques gives the better results. But it require appropriate methods and process for analysis of data. This study assessed the efficiency of spectroscopy and its techniques for rapid and inexpensive determination of soil properties parameters. This study covered the soil data acquisition using Field Spectroscopy in the form of spectral signatures then creation of spectral database of soil and preprocessing of spectral data. The spectroscopic technique is nondestructive, rapid, and low cost techniques. It is also useful for creation of real time spectral database. In the soil spectral data analysis soil sample collection and instrument setup are important then preprocessing, analysis, classification techniques are required for getting the accurate results of estimation of soil parameters using spectroscopy techniques. For the spectral analysis required methods are explain in this paper. This study is helpful to researcher for analysis of soil parameter using spectral data and process the data using appropriate methods in minimum time.

References:

[1]. Zejian Lei, Mingyin Yao, Muhua Liu, Qiulian Li and Hanping Mao. Comparison between Fertilization N, P, K and No Fertilization N, P, K in Paddy Soil by Laser Induced Breakdown Spectroscopy. IEEE Intelligent Computation Technology and Automation. 2011. 1; 363-366.

[2]. E.S. Mohamed, A.M. Saleh, A.B. Belal, Abd_Allah Gad , "Application of near-infrared reflectance for quantitative assessment of soil properties", The Egyptian Journal of Remote Sensing and Space Sciences, 21, 2018 pg. 1–14

[3]. Vipin Y. Borole, Sonali B. Kulkarni, Pratibha R. Bhise, "Effect of fertilizers on soil properties for different crops in pre-monsoon season using spectroradiometer for raver tehsil of jalgaon district", International Journal of Scientific and Technology Research, Volume 9, Issue 2, February 2020, Pg. 844-849

[4]. Ramdas D. Gore, Sunil S. Nimbhore, Bharti W.Gawali, "Understanding Soil Spectral Signature Though RS and GIS Techniques", International Journal of Engineering Research and General Science Volume 3, Issue 6, pp. 866-872, November-December, 2015

[5]. Bhise Pratibha.R, Kulkarni Sonali.B "Remote Sensing and Data Mining Techniques Applied on Soil Characteristics Data Classification", IOSR Journal of Computer Engineering (IOSR-JCE), pp. 83-91

[6]. P.R. Bhise, S.B. Kulkarni, "Evaluation of Soil Physical/Chemical Parameters for Agriculture Production in Vaijapur Taluka Using VNIR-SWIR Reflectance Spectroscopy", International Journal of Computer Sciences and Engineering, Vol.-6, Issue-12, Dec 2018, pp. 43-48

[7]. Vipin Y. Borole, Sonali B. Kulkarni, "Soil quality assessment for analyzing the effect of chemical fertilizers on agriculture field using Spectroradiometer: A review", International Conference on Electrical, Communication, Electronics, Instrumentation and Computing (ICECEIC), IEEE, 2019

[8]. vipin y. Borole, sonali b.kulkarni, pratibha r. Bhise, "soil properties assessment in surface and Subsurface using spectroradiometer for Raver tehsil of jalgaon district", journal of critical reviews, vol 7, issue 18, aug 2020, pg. 2487-2495

[9]. Vipin Y. Borole, Sonali B. Kulkarni, Pratibha R. Bhise, "Soil spectral signature analysis for influence of fertilizers on two differen crops in raver Tahshil", International Journal of Recent Technology and Engineering, Volume-8, Issue-3, Sep-2019, pp. 659-663

[10]. Agus Arip Munawar , Yuswar Yunus , Devianti Devianti , Purwana Satriyo,"Calibration models database of near infrared spectroscopy to predict agricultural soil fertility properties", Data in Brief (2020), doi: https://doi.org/10.1016/j.dib.2020.105469

[11]. Mansour Chaternour, Ahmad Landi, Ahmad Farrokhian Firouzi, Aliakbar Noroozi, Hosseinali Bahrami,"Spectral behavior modeling of soil texture over dust center of Khuzestan Province using hyperspectral images and Random Forest (RF) model", J. Adv. Appl. Geol., Vol 9 (4), Winter 2019, Pg. 466-479

[12]. ZHANG Peng-peng, XU Shou-zhen, ZHANG Guo-juan, PU Xiao-zhen, WANG Jin, ZHANG Wangfeng, "Carbon cycle in response to residue management and fertilizer application in a cotton field in arid Northwest China", JIA, ScienceDirect, Elsevier, 2019, Pg. 1103-119

[13]. Okoffo E. D., Ofori A., Nkoom M., Bosompem O.A., " Assessment of the physicochemical Characterristics of soils in major Cocoa producing areas in the dormaa west district of Ghana", International Journal of Scientific & Technology Research , Vol. 5, issue 02, feb. 2016, Pg. 62-68

[14]. Anserd J. Foster, Vijaya Gopal Kakani, Jianjun Ge, Mark Gregory & Jagadeesh Mosali, "Discriminant analysis of nitrogen treatments in switchgrass and high biomass sorghum using leaf and canopy-scale reflectance spectroscopy", International Journal of Remote Sensing, 37:10,2016, Pg. 2252-2279, DOI: 10.1080/01431161.2016.1171926

[15]. Suzana Romeiro Araujo, Jose Alexandre Melo Dematte, Henrique Bellinaso, " Analysing the effect of applying agricultural lime to soils by VNIR spectral sensing:a quantitative and quick method", International Journal of Remote sensing, Tylor &Francis, Vol.34, no. 13, Pg. 4570-4584

[16]. Huan Yu, Bo Kong, Guangxing Wang, Rongxiang Du, Guangping Qie, "Prediction Of Soil Properties Using A Hyperspectral Remote Sensing Method", Archives Of Agronomy And Soil Science, Taylor & Francis Group, 2017

[17]. Naxin Cui, Min Cai, Xu Zhang, Ahmed A. Abdelhafez, Li Zhou, Huifeng Sun, Guifa Chen, Guoyan Zou, Sheng Zhou, "Runoff loss of nitrogen and phosphorus from a rice paddy field in the east of China: Effects of long-term chemical N fertilizer and organic manure applications", Global Ecology and Conservation, Elsevier,22 , March 2020, pg.1-12

[18]. Mansour Chaternour, Ahmad Landi, Ahmad Farrokhian Firouzi, Aliakbar Noroozi, Hosseinali Bahrami, "Spectral behavior modeling of soil texture over dust center of Khuzestan Province using hyperspectral images and Random Forest (RF) model", Vol 9 (4), Oct 2019, Pg.466-479

[19]. Pratibha R.Bhise , Sonali B.Kulkarni, "Estimation of Soil Macronutrients From Spectral Signatures Using Hyperspectral Non-Imaging Data", International Conference on Electrical, Communication, Electronics, Instrumentation and Computing (ICECEIC), IEEE, 2019
[20]. Pratibha R.Bhise , Sonali B.Kulkarni, "Review on Analysis and Classification Techniques of Soil Study in Remote Sensing and Geographic Information System", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 6, Issue 1, pp.124-138
[21]. Pratibha R.Bhise , Sonali B.Kulkarni, Vipin Y. Borole.," Preprocessing and statistical analysis of soil parameters using conventional laboratory techniques and non-imaging spectral techniques for Vaijapur taluka", International Journal of Recent Technology and Engineering, Volume 8, Issue 2, July 2019, Pg. 3092-3096
[22]. Tiezhu Shi, Junjie wang, Huizeng Liu, Guofeng Wu, "Estimating leaf nitrogen concentration in heterogeneous crop plants from hyperspectral reflectance", International Journal of Remote Sensing, Vol. 36, No. 18, Oct 2015, Pg.4652-4667
[23]. Nataliia Ridei, Vita Strokal, Maryna Strokal, "Soil Quality classes for characterizing the soil potential to produce biologically and ecologically valuable crops", Archives of Agronomy and soil Science, Vol. 58, No. S1. Oct 2012 Pg. 219-225
[24]. GUO Wei, Mathias N Andersen, QI Xue-bin, LI Ping, LI Zhong-yang, FAN Xiang-yang, ZHOU Yuan, "Effects of reclaimed water irrigation and nitrogen fertilization on the chemical properties and microbial community of soil", Journal of Integrative Agriculture, Elsevier, Volume16, issue 3, page no, 679–690, 2017
[25]. Offer Rozenstein, Tarin Paz-Kagan, Christoph Salbach, and Arnon Karnieli, "Comparing the Effect of Preprocessing Transformations on Methods of Land-Use Classification Derived From Spectral Soil Measurements", IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Page 1-12

Vipin Y. Borole, et. al. "Spectral data analysis methods for soil properties assessment using remote sensing." *IOSR Journal of Computer Engineering (IOSR-JCE)*, 23(1), 2021, pp. 14-18.