Vol. 18, No. 1, pp. 88-98 (2018)
Journal of Agricultural Physics
ISSN 0973-032X
http://www.agrophysics.in

Research Article

Assessment of Soil Health Parameters using Proximal Hyperspectral Remote Sensing

SUJAN ADAK[1], K.K. BANDYOPADHYAY[1]*, R.N. SAHOO[1], T.J. PURAKAYASTHA[2], M. SHRIVASTAVA[3] AND N. MRIDHA[4]

[1]Division of Agricultural Physics, [2]Division of Soil Science and Agricultural Chemistry, [3]Centre for Environmental Science and Climate Resilient Agriculture
ICAR-Indian Agricultural Research Institute, New Delhi-110012
[4]National Institute of Research on Jute and Allied Fibre Technology, Kolkata-700040, West Bengal

ABSTRACT

Rapid and reliable assessment of soil health parameters is an essential requirement for sustainable management of the resources. Diffuse reflectance spectroscopy has emerged as a new tool to obtain both qualitative and quantitative information on soil properties and nutrient contents in a non-invasive manner from a single reflectance spectra. However, the potentials of various models to predict soil properties using spectral reflectance have not been fully explored. In this study, prediction of various properties of soil, collected from an ongoing field experiment on tillage, residue mulch and nitrogen interaction in maize-wheat cropping system were attempted from the soil spectral reflectance using four multivariate regression models *viz*., partial least square regression (PLSR), support vector regression (SVR), random forest (RF) and multivariate adaptive regression splines (MARS) using R software. Out of the 108 data points, $2/3^{rd}$ data was used for calibration of these models and $1/3^{rd}$ data was used for validation of these models. Among the four multivariate regression models using spectral reflectance, the RF model could account for 59, 48 and 54% variation in the observed sand, silt and clay content, respectively. The SVR model could account for 66% variation in the MWD whereas the RF model could account for 32% variation in WSA. The RF model could account for 81% variation for both the observed TOC and EC whereas the SVR model could account for 59% in the observed soil pH. The RF model could account for 61, 80 and 78% variation in the observed available nitrogen, phosphorus and potassium content, respectively. The prediction of soil biological parameters was poor. The RF model could account 44% variation in observed SMBC whereas the SVR model could account maximum 38% variation in observed DHA. Thus different chemical properties and selected physical properties of soil can be successfully assessed from spectral reflectance using different multivariate regression models.

Key words: Soil properties, Hyperspectral data, Multivariate regression models

Introduction

Rapid, reliable, cost effective and eco-friendly assessment of soil health has become a focus area in order to ensure their sustainable management. However, determination of soil health parameters at large scale is cumbersome and involves investment of money, manpower and time. Conventionally, determination of soil health parameters is performed under laboratory conditions. Among the soil health attributes, determination of soil chemical properties is based on wet chemistry with tedious and time-

*Corresponding author,
Email: kk.bandyopadhyay@gmail.com

consuming sample preparation and analyses steps whereas assessment of soil physical and biological attributes generally take a longer time than chemical attributes. Soil properties vary widely both in time and space (Minasny and Hartemink, 2011). Consequently, rapid and near-real time assessment of soil properties remains a formidable challenge despite decades of research and development in soil testing. Over the past few decades, remote sensing approaches provide some solution for rapid soil assessment (Vasques *et al*., 2010). These approaches are rapid, non-destructive and have large spatial coverage. In general, soils are opaque to most of the remote sensing methods. For example, microwave radiations penetrate only a few centimeters of the topsoil; whereas visible (VIS) and infrared radiations can barely penetrate through the soil surface. Remote sensing data have been used for soil classification, soil resources mapping (Ray *et al*., 2002; 2004), soil moisture assessment (Engman and Chauhan, 1995) and soil degradation (salinity) mapping (Metternicht and Zinck, 2003) among many others. Particularly, hyperspectral remote sensing (HRS) is emerging as a promising tool for its capability to measure the reflectance of earth surface features such as soil, water, vegetation, etc. at hundreds of contiguous and narrow wavelength bands. Availability of that large pool of spectral information offers an opportunity to estimate multiple soil attributes from the same reflectance spectra with greater specificity than their multispectral counterpart.

Proximal soil sensing refers to metering and data processing technology that allows *in situ* determination of physical, chemical, and other soil characteristics while placing sensor systems in close proximity to the soil being evaluated. Spectral reflectance of soil samples collected in the laboratory using contact probe can be used for assessing different soil attributes. Proximal Diffuse Reflectance Spectroscopy (DRS) in the Visible (VIS), Near-Infrared (NIR), and Shortwave-Infrared (SWIR) regions (350–2500 nm) forms the basis of Hyperspectral remotesensing. When soil is exposed to electromagnetic energy, chemical bonds of different molecules vibrate at characteristic frequencies, which is captured in the spectral reflectance of soil. Energy absorbed, reflected and scattered in the process may, therefore, be related to specific wavelengths (Ben-Dor and Banin, 1995). In particular, the specificity (reflectance at characteristic wavelength) allows for the assessment of different soil attributes once spectral reflectance is known and a relationship between the spectral feature and soil attribute is known a priori. Thus, spectral signatures are often considered as inherent soil properties that vary across different soils (Ben-Dor *et al*., 2009). So diffuse reflectance spectroscopy has emerged as a new tool to obtain both qualitative and quantitative information on soil properties and nutrient contents in a non-invasive manner from a single reflectance spectra. Laboratory-scale studies have clearly shown that the DRS approach may be used for estimating several soil properties such as soil texture (Bilgili *et al*., 2010), electrical conductivity (EC) (Shrestha, 2006), cation exchange capacity (CEC) (Fox and Metla, 2005), organic carbon (OC) content (Galvao and Vitorello, 1998; Singh *et al*., 2013), nutrient content such as nitrogen (N) (Vågen *et al*., 2006), phosphorus (P), potassium (K) (Mouazen *et al*., 2007), iron (Fe) content (Galvao and Vitorello, 1998), soil moisture content (Carlson *et al*., 1995), carbonates (Lagacherie *et al*., 2008) and hydraulic properties (Santra *et al*., 2009). Recently, the DRS approach has been shown to be successfully used for estimating some of the soil physical parameters (median aggregate diameter and standard deviation of lognormal aggregate size distribution function of soils (Sarathjith *et al*., 2014). Machine learing techniques are used for prediction of different soil and plant parameters using their spectral reflectance at characteristic wavelengths. Silva and ten Caten (2016) reported that sand content (R^2 = 0.81), clay content (R^2 = 0.80) and less satisfactory for silt content (R^2 = 0.70) can be predicted by PLSR model. Artificial neural networks (ANN) using a spectrum (400–1100 nm) was found as a precise detector of SOM (R^2 = 0.86) (Daniel *et al*., 2003) whereas a support vector machine regression (SVMR) and a successive projections algorithm (SPA) model

(SPASVMR model) have been used for improving the accuracy of soil organic carbon (SOC) which has resulted from integrating the laboratory-based visible and near-infrared (VIS/NIR, 350–2500 nm) spectroscopy of soils (Peng *et al*., 2014). Nawar *et al*. (2014) reported better prediction of soil salinity using MARS (R^2 = 0.73, RMSE = 6.53, and RPD = 1.96), than PLSR model (R^2 = 0.70, RMSE = 6.95, and RPD = 1.82). Moreover, the authors emphasized that MARS gives very good results for prediction of soil salinity, especially under high salinity levels. Mohamed *et al*. (2016) reported that SMLR model can be used for estimation of concentrations of heavy metals with high accuracy with R^2 of 0.82, 0.75 and 0.65 for Cr, Mn and Cu, respectively.

Tillage and crop residue mulch treatments modify the physical, chemical and biological properties of soil. Increase in water stable aggregates and mean weight diameter (MWD) of water stable aggregates (WSA) under no tillage than under conventional tillage has been reported by Abid and Lal (2008). Since soil aggregates are protected under no till practices, it is expected to contain higher SOC storage than conventional tillage system. Hati *et al.* (2014) found that SOC content was the highest in NT (8.6 g kg^{-1}) and the lowest in CT (6.5 g kg^{-1}) in 0-5 cm soil layer where as it was higher in RT and MB than in CT in the same soil layer under soybean-wheat cropping system. Surface retention or incorporation of crop residue and belowground biomass under NT and CT decreased pH of the soil. But, pH decline was more pronounced under NT than CT at 0-2.5 cm whereas there was no significant effect of tillage on pH throughout 2.5-20 cm profile (Mrabet *et al*., 2001). Sharma *et al*. (2011) found that conservation tillage increased the soil respiration and soil microbial biomass carbon by 81.1 and 104 %, respectively as compared to conventional tillage during wheat cultivation. However assessment of these changes in soil properties using remote sensing techniques has not been attempted.

In this study, an attempt was made to estimate some selected physical, chemical and biological properties of soil using DRS technology. The objective of this study was to examine the suitability of different multivariate regression models (partial least square regression (PLSR), support vector regression (SVR), random forest (RF) and multivariate adaptive regression spline (MARS)) based on spectral reflectance to estimate the selected soil attributes under different tillage, residue mulch and nitrogen management practices.

Materials and Methods

Study area and soil sampling

Field experiments were conducted during 2017-18 in the MB-4C farm of ICAR-Indian Agricultural Research Institute, New Delhi (28° 35′N latitude, 77°12′E longitude and at an altitude of 228.16 m above mean sea level) in an ongoing long term field experiment (since 2014) under maize-wheat cropping system. The experiment was laid out a split-split plot design, with two levels of tillage as main plot factor (Conventional tillage (CT) and No Tillage (NT), two levels of residue as subplot factor (with residue @ 5 t ha^{-1} (R+) and without residue (R0), and three levels of nitrogen as sub-sub plot factor (50% (N50), 100% (N100) and 150% (N150) of the recommended dose of nitrogen), replicated three times.

Soil samples from 0-5, 5-15 and 15-30 cm soil depths were collected using a bucket type soil core sampler from three random locations in each plot after harvest of wheat during 2016-17. Each sample was divided into three parts: the first part was stored in a refrigerator for determination of soil microbial biomass carbon (SMBC) and the second part was dried in shade, processed and passed through 2-mm sieve for analysis of soil chemical parameters and spectral reflectance measurement. The third part, passed through 8-mm sieve and retained in 4-mm sieve was used for aggregate analysis.

Laboratory analysis of soil samples

Hydrometer method (Bouyoucos, 1962) was used to determine sand, silt and clay percentage for each sample. The aggregate size distribution of soil (mean weight diameter and water stable aggregate) was determined by wet sieving method using Yodder's apparatus (Yodder, 1936). Soil pH and EC were determined in 1:2.5 soil-water suspension using combined electrode for pH and Conductivity Bridge for EC (Jackson, 1973). Total organic carbon was determined by automatic elemental analyser (*Vario EL, Elementar Analysen systeme GmbH*, Hanau, Germany). Available N (Subbiah and Asija, 1956), available P (ascorbic acid blue color method; Watanabe and Olsen, 1965) and available K (flame photometer method; Hanway and Heidel, 1952) in the soil samples were determined as per standard procedures. The microbial biomass carbon (MBC) in soil was determined by fumigation extraction method as described by Jenkinson and Powlson (1976) and Dehydrogenase activity in soils was determined as per Klein *et al.* (1971) method.

Soil reflectance measurements

The reflectance spectra of soil samples (air-dried, crushed and 2 mm sieved) were collected using spectroradiometer with contact probe. The soil was only illuminated by a constant light source inside the contact probe (Contact Probe, Analytical Spectral Devices, Boulder, CO) after calibration of sensor using a white spectral panel. The Viewspec Pro software of the instrument has been set to process reflectance at 1 nm interval. Spectral reflectance was derived as the ratio of reflected radiance to incident radiance estimated by a calibrated white reference.

Spectroscopic Data Pre-Processing

In order to boost the predictive power of multivariate calibration models, spectral data are often preprocessed prior to data analysis as variation in the predictor variables that is unrelated to response variable may reduce the predictive ability of the models. The aim of pre-processing is to reduce the effects of random noise and improve signal to noise ratio. In the present study the data was filtered using Savitzky-Golay filter (Savitzky and Golay, 1964).

Multivariate Techniques

Soil and plant parameters were assessed from the spectral reflectance using different multivariate regression models like partial least square regression (PLSR), multivariate adaptive regression splines (MARS), support vector regression (SVR), and random forest (RF) through R software.

Accuracy assessment

Calibration of the four multi linear regression models (PLSR, MARS, SVM and RF) for soil properties was done using 2/3rd of the total (108) soil samples (training data set). Validation of those models was done using 1/3rd of the total soil samples (testing data set) which were not used for model calibration.

The accuracy of the model prediction was tested using different statistical indices like R^2, normalized root mean squared error (nRMSE) and ratio of performance to deviation (RPD).

The root mean square error (RMSE) was used to calculate the error between the estimated and measured results.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_i - O_i\right)^2} \quad \ldots(1)$$

The normalized RMSE is computed as RMSE as a percentage of the observed mean value.

$$nRMSE = (RMSE/\bar{O}) \times 100 \quad \ldots(2)$$

where, P_i is predicted value, O_i is observed value, $\bar{O}$ is observed mean and n is number of samples. nRMSE (%) shows the relative difference between the predicted and observed data. The prediction is considered excellent if the nRMSE < 10 %, good if 10–20 %, fair if 20–30 %, poor if > 30 % of the observed mean (Jamieson *et al.*, 1991).

The appropriateness of a prediction was also evaluated by the ratio of performance to deviation (RPD), which is the ratio between the standard deviation (SD) of the validation sample set and standard error of prediction (SEP) (Willimas and Norris 1987):

$$RPD = SD/SEP \quad \ldots(3)$$

$$SEP = \sqrt{\frac{1}{N-1}\sum_{n-1}^{1}(Pi - Oi)^2} \quad \ldots(4)$$

Where P_i is predicted value, O_i is observed value and n is number of samples.

Chang *et al*. (2001) classified prediction accuracies into accurate (RPD > 2), moderate (1.4 < RPD < 2), and poor (RPD < 1.4), although such a classification rule is still being debated (Bellon-Maurel *et al*., 2010).

Results and Discussion

Soil reflectance spectra

Reflectance spectra of soil samples as influenced by tillage, residue and nitrogen management are presented in Fig. 1. There was a continuous increase in reflectance with the increase in wavelength. There were distinct dips in the soil reflectance spectra at 1400 and 1900 nm due to water absorption band and at 2200 nm due to hydroxyl group of soil. Averaged over residue and nitrogen management, the highest soil reflectance was recorded at no tillage at 15-30 cm soil depth and lowest soil reflectance was recorded in no tillage at 5-15 cm soil depth. Averaged over tillage and nitrogen management, highest soil reflectance was recorded in no residue mulch (R0) treatment at 15-30 cm soil depth and lowest soil reflectance was recorded in no residue mulch treatment (R0) at 5-15 cm soil depth. Averaged over tillage and residue management, highest soil reflectance was recorded in N150 treatment at 15-30 cm soil depth and lowest soil reflectance was recorded in N100 treatment at 5-15 cm soil depth. In general it was observed that soil reflectance increased for lower depth soil and in the absence of crop residue mulch. Lower organic carbon content in soil at lower depth under no mulch treatment might be responsible for the higher spectral reflectance at this depth. Saxena *et al.* (2003) reported that the soil spectral reflectance decreased with increase in organic carbon content of soil.

Descriptive statistics

Descriptive statistics of 13 soil parameters (108 data set) *viz*., sand, silt, clay, mean weight diameter, water stable aggregate, total organic

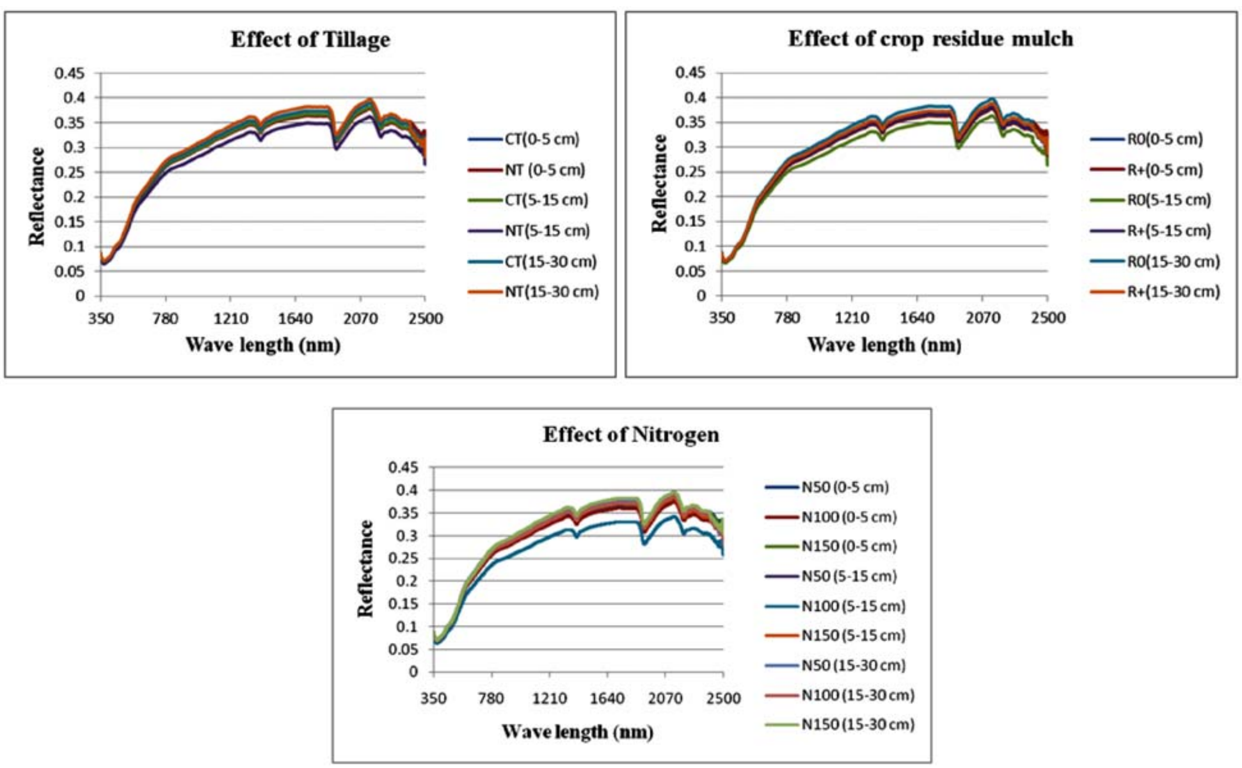

Fig 1. Soil reflectance spectra at 0-5, 5-15 and 15-30 cm soil depth as influenced by tillage, residue and N management

carbon, pH, electrical conductivity, available nitrogen, available phosphorus, available potassium, soil microbial biomass carbon, dehydrogenase activity pooled over all treatments for 0-5, 5-15 and 15-30 cm soil depths is presented in Table 1. The coefficient of variance of soil parameters ranged from 3.51 (sand) to 45.5% (dehydrogenase activity).

Calibration of the models

Calibration of four multivariate regression models *viz.*, partial least square regression (PLSR), multivariate adaptive regression splines (MARS), support vector regression (SVR), and random forest (RF) based on spectral reflectance was carried out using R software and their relative performance was judged by comparing the nRMSE and the R^2 (Table 2). During calibration the RF model performed best for prediction of sand, silt and clay content where it could account 93, 92 and 93% variation in the observed sand, silt and clay with an nRMSE of 1.6, 2.29 and 4.1%, respectively. During calibration. the SVR model performed best for perdiction of MWD with a R^2 and nRMSE value of 0.81 and 8.76%, respectively whereas for the calibration of WSA prediction RF model performed best with a R^2 and nRMSE value of 0.95 and 5.91%, respectively. During calibration, the RF model performed best for prediction of TOC, pH and EC where it could account 94, 91 and 91% variation in the observed TOC, pH and EC with an nRMSE of 2.69, 1.77 and 15.25%, respect-ively. During calibration, the RF model performed best for prediction of available N, P, K, SMBC and DHA where it could account 91, 92, 93, 94 and 93% variation in the observed available N, P, K, SMBC and DHA with an nRMSE of 4.52, 11.22, 15.76, 17.87 and 17.79%, respectively. So RF model performed marginally better than the other models for predicting different soil parameters with relatively lower nRMSE and higher R^2. Generally, the machine learning methods (RF, SVR, MARS) were found to be more accurate than PLSR using the RSME of calibration for assessing model performance (Zakaria and Shabri, 2012; Bricklemyer *et al.*, 2007).

Validation of the models

The model performance statistics for the validation of different soil properties using 1/3rd of total dataset is presented in Table 3. It was observed that among these four models, RF model performed best for prediction of sand, silt and clay content. It could account for 59, 48 and 54% variation in the observed sand, silt and clay content with an nRMSE of 1.97, 4.13 and 6.76% and RPD of 1.51, 1.34 and 1.36, respectively. However, this model could be considered as a moderate predictor of sand whereas poor for silt and clay based on the RPD values. Low variability in texture within a field due to different

Table 1. Descriptive statistics of soil properties (pooled over treatments and depths)

Parameters	No	Min	Max	Range	Mean	Std. Dev.	CV	Std. Error
Sand (%)	108	52.00	61.79	9.79	55.71	1.95	3.51	0.19
Silt (%)	108	26.29	35.40	9.11	31.45	1.86	5.86	0.18
Clay (%)	108	9.55	15.80	6.25	12.64	1.22	9.64	0.12
MWD (mm)	108	0.51	1.29	0.78	0.90	0.18	20.19	0.02
WSA (%)	108	32.11	75.31	43.20	56.43	7.73	13.69	0.74
TOC (%)	108	0.53	0.95	0.42	0.75	0.08	10.95	0.01
pH	108	6.88	8.49	1.61	7.86	0.33	4.25	0.03
EC (ds m^{-1})	108	0.19	0.93	0.42	0.75	0.08	10.95	0.01
Av-N (mg kg^{-1})	108	39.00	78.40	39.40	55.66	6.92	12.43	0.67
Av-P (mg kg^{-1})	108	2.00	9.00	7.00	3.90	1.39	35.63	0.13
Av-K (mg kg^{-1})	108	91.00	487.50	396.50	185.27	77.70	41.94	7.48
SMBC (µg g^{-1} soil)	108	105.80	438.50	332.70	238.98	93.92	39.30	9.04
DHA (µg TPF g^{-1} soil day^{-1})	108	3.09	21.64	18.55	10.09	4.59	45.47	0.44

Table 2. Calibration of soil parameters using multivariate regression models

Parameters	PLSR		SVR		RF		MARS	
	R^2	nRMSE	R^2	nRMSE	R^2	nRMSE	R^2	nRMSE
Sand (%)	0.55	2.50	0.84	1.5	0.93	1.6	0.76	1.78
Silt (%)	0.5	4.33	0.82	2.72	0.93	2.29	0.61	3.72
Clay (%)	0.34	15.01	0.72	5.17	0.92	4.10	0.59	6.23
MWD (mm)	0.50	14.14	0.81	8.76	0.83	10.18	0.50	13.92
WSA (%)	0.36	11.46	0.80	6.37	0.95	5.91	0.66	8.13
TOC (%)	0.70	5.80	0.99	1.02	0.94	2.69	0.89	3.49
pH	0.70	2.52	0.79	1.97	0.91	1.77	0.65	2.44
EC (ds m^{-1})	0.86	13.91	0.91	11.67	0.91	15.25	0.71	20.01
Available N (mg kg^{-1})	0.5	8.60	0.83	5.14	0.91	4.52	0.73	6.19
Available P (mg kg^{-1})	0.6	18.83	0.8	13.19	0.92	11.22	0.7	11.73
Available K (mg kg^{-1})	0.6	24.44	0.9	12.65	0.93	15.76	0.8	16.99
SMBC (µg g^{-1} soil)	0.37	34.82	0.77	21.13	0.94	17.87	0.47	30.69
DHA (µg TPF g^{-1} soil day^{-1})	0.52	32.34	0.69	25.06	0.93	17.79	0.68	25.19

management practices may be the cause for such performance of the models. With respect to prediction of mean weight diameter (MWD), SVR model performed the best and it could account for 66% variation in the observed MWD with an nRMSE of 12.8% and RPD of 1.65. However, the RF model could account for maximum 32.3% variation in the observed WSA with an nRMSE of 10.4% and RPD of 1.19. For MWD and WSA predication, most of the models performed poorly. Use of disturbed soil sample (sieved) for capturing spectral signature may be the reason for poor prediction as MWD and WSA are determined from undisturbed soil aggregates. The RF model performed best for prediction of soil EC, TOC, available P and available K. It could

Table 3. Prediction of soil parameters using multivariate regression models

Parameters	PLSR			SVR			RF			MARS		
	R^2	nRMSE	RPD	R^2	nRMSE	RPD	R^2	nRMSE	RPD	R^2	nRMSE	RPD
Sand (%)	0.12	2.94	1.04	0.42	2.75	1.11	0.59	1.97	1.51	0.43	2.50	1.24
Silt (%)	0.37	4.21	1.17	0.47	4.36	1.23	0.48	4.13	1.34	0.37	4.91	1.02
Clay (%)	0.25	8.38	1.10	0.41	7.32	1.26	0.54	6.76	1.36	0.30	8.42	1.10
MWD (mm)	0.32	17.76	1.19	0.66	12.80	1.65	0.61	13.54	1.56	0.43	16.33	1.29
WSA (%)	0.18	11.36	1.10	0.20	11.73	1.06	0.32	10.40	1.19	0.22	12.08	1.02
TOC (%)	0.61	7.21	1.57	0.77	6.14	1.85	0.81	4.92	2.30	0.70	6.22	1.82
pH	0.57	3.35	1.34	0.59	2.85	1.57	0.53	3.08	1.46	0.44	3.53	1.27
EC(ds m^{-1})	0.77	19.52	1.86	0.68	21.35	1.70	0.81	19.16	1.89	0.54	25.58	1.42
Available N (mg kg^{-1})	0.38	10.56	1.20	0.65	8.28	1.54	0.61	8.19	1.55	0.56	10.04	1.28
Available P (mg kg^{-1})	0.58	18.59	1.52	0.67	17.49	1.61	0.80	13.73	2.05	0.48	22.97	1.23
Available K (mg kg^{-1})	0.53	31.74	1.44	0.55	30.72	1.49	0.78	22.76	2.00	0.70	25.49	1.79
SMBC (µg g^{-1} soil)	0.21	30.03	1.11	0.38	28.97	1.15	0.44	25.66	1.30	0.21	31.19	1.07
DHA (µg TPF g^{-1} soil day^{-1})	0.27	41.70	1.17	0.19	46.33	1.04	0.35	39.15	1.23	0.38	39.09	1.23

account 81, 94, 80 and 78% variation in the observed soil EC, TOC, available P and available K with an nRMSE of 19.16, 4.9, 13.73 and 22.76% and RPD of 1.89, 2.30, 2.05 and 2.00, respectively. Based on the RPD values, the RF model performed as accurate predictor for TOC, available P and available K whereas other models performed as moderate predictors. In general, many studies reported that RF served as better predictor compared to support vector machine (SVM) for most of the soil parameters (Ließ *et al*., 2016; Siegmann and Jarmer, 2015; Ma *et al*., 2016; Fassnacht *et al*., 2014). Rossel *et al*. (2010) reported that RF had better prediction accuracy compared to stochastic gradient boosting (SGB). The SVR model performed best and could account 59 and 65 % variation in observed pH and available N with nRMSE of 2.85 and 8.28% and RPD of 1.57 and 1.54, respectively. Hitziger and Ließ (2014) found that SVM and SGB were superior to the RF in prediction of some soil properties (sand and clay). Among the four models, the RF model performed best for the prediction of SMBC and could account 44 % variation in observed SMBC with nRMSE of 25.66% and RPD of 1.30 whereas MARS model could account 38 % variation in observed SMBC with nRMSE of 39.09% and RPD of 1.23. Thus, all models performed as poor predictor for soil biological parameters. Biological properties like SMBC and DHA are determined on fresh or field moist soil samples whereas the spectra was taken in from processed soil samples. So the spectral reflectance could not capture changes in these biological properties. Reeves *et al*. (2000) observed only moderate accuracies for prediction of four enzymes (arylsulfatase, dehydrogenase, phosphatase and urease) with a coefficient of determination (R^2) ranging from 0.43 to 0.77 in the cross-validation. Thus in this study, RF model performed better than SVR for prediction of most of the parameters. However, SVR model performed better than MARS and PLSR. Nonparametric models such as RF, SVM and SGB have been found superior to MLR due to their ability to handle non-linear relations and multi-source data (Hahn and Gloaguen, 2008; Bricklemyer *et al*., 2007; Wålinder, 2014).

Conclusions

The present study concluded that the multivariate regression model, random forest (RF) could satisfactorily assess sand, silt, clay, TOC, EC, available P, K content of soil whereas support vector regression (SVR) model could satisfactorily assess soil pH and available N using soil reflectance spectra. So different chemical properties and selected physical properties of soil can be satisfactorily assessed from diffuse reflectance spectra using different multivariate regression models. The prediction of soil biological parameters using spectral reflectance spectra was not satisfactory.

Acknowledgements

This work is a part of the M.Sc. degree research. The first author acknowledges the help received from ICAR in the form of Junior Research Fellowship during the study period. The logistic support received from the Director, IARI during the study period is thankfully acknowledged.

References

Abid, M. and Lal, R. 2008. Tillage and drainage impact on soil quality: I. Aggregate stability, carbon and nitrogen pools. *Soil and Tillage Research* **100(1-2)**: 89-98.

Bellon-Maurel, V., Fernandez, E., Palagos, B., Roger, J.M. and McBratney, A. 2010. Prediction of soil attributes by NIR/MIR spectroscopy. Coming back to statistics fundamentals to improving prediction accuracy. *TrAC - Trends in Analytical Chemistry* **29(9)**: 1073-1081.

Ben-Dor, E. and Banin, A. 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Sci. Soc. Am. J.* **59**: 364-372.

Ben-Dor, E., Chabrillat, S., Demattê, J.A.M., Taylor, G.R., Hill, J., Whiting, M.L. and Sommer, S. 2009. Using imaging spectroscopy to study soil properties. *Remote Sensing Environ*. **113**: S38-S55.

Bilgili, A.V., van Es, H.M., Akbas, F., Durak, A. and Hively, W.D. 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *J. Arid Environ*. **74**: 229-238.

Bouyoucos, G.J. 1962. Hydrometer method improved for making particle size analysis of soils. *Agronomy Journal* **54**: 464-465.

Bricklemyer, R.S., Lawrence, R.L., Miller P.R. and Battogtokh, N. 2007. Monitoring and verifying agricultural practices related to soil carbon sequestration with satellite imagery. *Agriculture, Ecosystems and Environment* **118(1)**: 201-210.

Carlson, T.N., Gillies, R.R. and Schmugge, T.J. 1995. An interpretation of methodologies for indirect measurement of soil water content. *Agric. For. Meteorol*. **77**: 191-205.

Chang, C., Laird, D.A., Mausbach, M.J. and Hurburgh Jr., C.R. 2001. Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of soil properties. *Soil Science Society of America Journal* **65**: 480-490.

Daniel, K.W., Tripathi, N.K. and Honda, K. 2003. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Soil Research* **41(1)**: 47-59.

Engman, E.T. and Chauhan, N. 1995. Status of microwave soil moisture measurements with remote sensing. *Remote Sensing Environ.* **51(1)**: 189-198.

Fassnacht, F.E., Hartig, F., Latifi, H., Berger, C., HernaÂndez, J., Corvalaan, P. and Koch, B. 2014. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment* **154**: 102-114.

Fox, G.A. and Metla, R. 2005. Soil property analysis using principal components analysis, soil line, and regression models. *Soil Sci. Soc. Am. J.* **69**: 1782-1788.

Galvao, L.S. and Vitorello, I. 1998. Role of organic matter in obliterating the effects of iron on spectral reflectance and colour of Brazilian tropical soils. *Int. J. Remote Sensing* **19**: 1969-1979.

Hahn, C. and Gloaguen, R. 2008. Estimation of soil types by non linear analysis of remote sensing data. *Nonlinear Processes in Geophysics* **15(1)**: 115-126.

Hanway, J.J. and Heidel, H. 1952. Soil analysis methods as used in Iowa State College, Soil Testing Laboratory, *Iowa State College Bull*. **57**: 1-131.

Hati, K.M., Singh, R.K., Mandal, K.G., Bandyopadhyay, K.K., Somasundaram, J., Mohanty, M., Sinha, N.K., Chaudhary, R.S. and Biswas, A.K. 2014. Conservation tillage effects on soil physical properties, organic carbon concentration and productivity of soybean-wheat cropping system. *Journal of Agricultural Physics* **14(2)**: 121-129.

Hitziger, M. and Ließ, M. 2014. Comparison of three supervised learning methods for digital soil mapping: Application to a complex terrain in the Ecuadorian Andes. *Applied and Environmental Soil Science* 4331-4334.

Jamieson, P.D., Porter, J.R. and Wilson, D.R. 1991. A test of the computer simulation model ARCWHEAT1 on wheat crops grown in New Zealand. *Field Crops Research* **27(4)**: 337-350.

Jenkinson, D.S. and Powlson, D.S. 1976. The effects of biocidal treatments on metabolism in soil-V. A method for measuring soil biomass. *Soil Biology and Biochemistry* **8**: 209-213.

Klein, D.A., Loh, T.C. and Goulding, R.L. 1971. A rapid procedure to evaluate the dehydrogenase activity of soils low in organic matter. *Soil Biology and Biochemistry* **3**: 385-387.

Lagacherie, P., Baret, F., Feret, J.B., Madeira Netto, J. and Robbez-Masson, J.M. 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sensing Environ.* **112**: 825-835.

Ließ, M., Schmidt, J. and Glaser, B. 2016. Improving the spatial prediction of soil organic carbon stocks in a complex tropical mountain landscape by methodological specifications in machine learning approaches. *PloS one*. **11(4)**: e0153673.

Ma, W., Tan, K. and Du, P. 2016. Predicting soil heavy metal based on Random Forest model. In Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International. pp. 4331-4334.

Metternicht, G.I. and Zinck, J.A. 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sensing Environ.* **85(1)**: 1-20.

Minasny, B. and Hartemink, A.E. 2011. Predicting soil properties in the tropics. *Earth-Science Reviews* **106(1-2)**: 52-62.

Mohamed, E.S., Ali, A.M., El Shirbeny, M.A., El Razek, A.A.A. and Savin, I.Y. 2016. Near infrared spectroscopy techniques for soil contamination assessment in the Nile Delta. *Eurasian Soil Science* **49(6)**: 632-639.

Mouazen, A.M., Maleki, M.R., De Baerdemaeker, J. and Ramon, H. 2007. On-line measurement of some selected soil properties using a VIS–NIR sensor. *Soil Till. Res.* **93**: 13-27.

Mrabet, R., Saber, N., El-Brahli, A., Lahlou, S. and Bessam, F. 2001. Total, particulate organic matter and structural stability of a Calcixeroll soil under different wheat rotations and tillage systems in a semiarid area of Morocco. *Soil and Tillage Research* **57(4)**: 225-235.

Nawar, S., Buddenbaum, H., Hill, J. and Kozak, J. 2014. Modeling and mapping of soil salinity with reflectance spectroscopy and Landsat data using two quantitative methods (PLSR and MARS). *Remote Sensing* **6(11)**: 10813-10834.

Peng, X., Shi, T., Song, A., Chen, Y. and Gao, W. 2014. Estimating soil organic carbon using VIS/NIR spectroscopy with SVMR and SPA methods. *Remote Sensing* **6(4)**: 2699-2717.

Ray, S.S., Singh, J.P., Das, G. and Panigrahy, S. 2004. Use of high resolution remote sensing data for generating site-specific soil management plan. *Int. Arch. Photogramm. Remote Sensing Spatial Inf. Syst. B.* **35(7)**: 127-131.

Ray, S.S., Singh, J.P., Dutta, S. and Panigrahy, S. 2002. Analysis of within field variability of crop and soil using field data and spectral information as a pre-cursor to precision crop management. *Int. Arch. Photogramm. Remote Sensing Spatial Inf. Syst. C.* **34(7)**: 302-307.

Reeves III, J.B., McCarty, G.W. and Meisinger, J. 2000. Near infrared reflectance spectroscopy for the determination of biological activity in agricultural soils. *J. Near Infrared Spectrosc.* **8**: 161-170.

Rossel, R.V. and Behrens, T. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **158(1-2)**: 46-54.

Santra, P., Sahoo, R.N., Das, B.S., Samal, R.N., Pattanaik, A.K. and Gupta, V.K. 2009. Estimation of soil hydraulic properties using proximal spectral reflectance in visible, near-infrared, and shortwave-infrared (VIS–NIR–SWIR) region. *Geoderma* **152**: 338-349.

Sarathjith, M.C., Das, B.S., Vasava, H.V., Mohanty, B., Anand, S.S., Wani, S.P. and Sahrawat, K.L. 2014. Diffuse reflectance spectroscopic approach for the characterization of soil aggregate size distribution. *Soil Sci. Soc. Am. J.* **78**: 369-376.

Savitzky, A. and Golay, M.J.E. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* **36(8)**: 1627-1639.

Saxena, R.K., Vermal, K.S., Srivastava, R., Shiwalkar, A.A. and Londhel, S.L. 2003. Spectral reflectance properties of some dominant soils occurring on different altitudinal zones in Uttaranchal Himalayas. *Agropedology* **13(2)**: 35-43.

Sharma, P., Singh, G. and Singh, R.P. 2011. Conservation tillage, optimal water and organic nutrient supply enhance soil microbial activities during wheat (*Triticum aestivum L.*) cultivation. *Brazilian Journal of Microbiology* **42(2)**: 531-542.

Shrestha, R. 2006. Relating soil electrical conductivity to remote sensing and other soil properties for assessing soil salinity in Northeast Thailand. *Land Degrad. Dev*. **17**: 677-689.

Siegmann, B. and Jarmer, T. 2015. Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data. *International Journal of Remote Sensing*. **36(18)**: 4519-4534.

Silva, E.B., ten Caten, A., Dalmolin, R.S.D., Dotto, A.C., Silva, W.C. and Giasson, E. 2016. Estimating Soil Texture from a Limited Region of the Visible/Near-Infrared Spectrum. In *Digital Soil Morphometrics*. pp. 73-87.

Singh, K., Murphy, B. and Marchant, B. 2013. Towards cost-effective estimation of soil carbon stocks at the field scale. *Soil Res.* **1**: 485-490.

Subbiah, B.V. and Asija, G.L. 1956. A rapid procedure for the determination of available nitrogen in soils. *Current Science* **25**: 259-260.

Vågen, T.G., Shepherd, K.D. and Walsh, M.G. 2006. Sensing landscape level change in soil fertility following deforestation and conversion in the highlands of Madagascar using VIS–NIR spectroscopy. *Geoderma* **133(3-4)**: 281-294.

Vasques, G.M., Grunwald, S. and Harris, W.G. 2010. Spectroscopic models of soil organic carbon in Florida, USA. *Journal of Environmental Quality* **39(3)**: 923-934.

Wålinder, A. 2014. Evaluation of logistic regression and random forest classification based on prediction accuracy and metadata analysis: Linnaeus University: Sweden. pp. 44.

Watanabe, F.S. and Olsen, S.R. 1965. Test of ascorbic acid method for determining phosphorus in water and sodium bicarbonate extracts of soil. *Proc. Soil Sci. soc. Am.* **29**: 677-678.

Williams, P. and Norris, K. 1987. Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists, Inc.

Yodder, R.E. 1936. A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses. *Journal of the American Society of Agronomy* **28**: 337-351.

Zakaria, Z.A. and Shabri, A. 2012. Streamflow forecasting at ungaged sites using support vector machines. *Applied Mathematical Sciences*. **6(60):** 3003-3014.

Received: April 20, 2018; Accepted: May 22, 2018