Environ Earth Sci
DOI 10.1007/s12665-015-4383-x

ORIGINAL ARTICLE

Reflectance spectroscopic approach for estimation of soil properties in hot arid western Rajasthan, India

Priyabrata Santra[1] · Ravindra Singh[2] · M. C. Sarathjith[2] · N. R. Panwar[3] · Preeti Varghese[1] · B. S. Das[2]

Received: 27 May 2014 / Accepted: 1 April 2015

Abstract Periodic and regular assessment of land degradation in arid regions of the world is essential for implementing suitable corrective measures in time. Assessment of soil properties based on soil sampling from hot arid tracts followed by laboratory analysis is a formidable task. Reflectance spectroscopy appears to be an emerging technology for the assessment of soils in extreme environment. In this study, soil spectral library of 138 soil samples from hot arid western Rajasthan have been created in visible, near-infrared and short wave infrared (350–2500 nm) region of the electromagnetic spectrum along with the measurements of basic soil properties. Further, spectral reflectance-based algorithms have been developed for rapid assessment of soil resources of arid regions. Results showed that sand and clay content may be satisfactorily estimated from linear models involving principal components (PCs) or derived band reflectance as the input variables ($R^2 = 0.41$–0.43). Organic carbon (OC) content of soil was also found satisfactorily correlated with spectral data ($R^2 = 0.27$). Derived band reflectance corresponding to Operational Land Imager bands of Landsat-8 has been found best to predict soil properties. Soil OC content has been found to be best estimated by derived spectral band data corresponding to spectral bands of IRS-P6 satellite. Partial least square regression-based models were found even better than the PCs-based and band reflectance-based multiple regression models for estimating soil properties. Thus, the present study indicates that soil spectral reflectance data captured by remote sensing satellites may have a great potential for rapid assessment of soil resources in arid regions.

Keywords VIS–NIR–SWIR · Indian Thar Desert · Principal components · Band reflectance · Partial least square regression (PLSR) · Soil resource assessment

Introduction

Desertification processes are affecting the livelihoods of more than 2 billion people of the world in dry lands, which occupy nearly 41 % of the Earth's land area. In India, about 31.7 m ha of land lies under hot arid ecosystem, most of which again falls in western Rajasthan (62 %). The region is characterized by limited seasonal precipitation with erratic distribution. Mean annual rainfall in the region varies from 185 mm at Jaisalmer to more than 467 mm at Sikar. About 80–90 % of the annual rainfall is received during the southwest monsoon. It has been reported by Kar et al. (2009) that between 1982–1983 and 2005–2006 net irrigated and double cropped area in hot arid region has increased by 128 and 70 %, respectively. Such rapid and drastic change in land use pattern in cultivated area has a potential impact on native soil resources. Moreover, mapping effort of desertification under different land uses in the arid western Rajasthan revealed that 76 % area is affected by wind erosion, encompassing all the major land uses, mostly croplands and dunes/sandy areas. Under such

✉ Priyabrata Santra
priyabrata.iitkgp@gmail.com

1 Division of Agricultural Engineering for Arid Production Systems, ICAR-Central Arid Zone Research Institute, Jodhpur 342003, Rajasthan, India

2 Agricultural and Food Engineering Department, IIT Kharagpur, Kharagpur 721302, West Bengal, India

3 Division of Natural Resources and Environment, ICAR-Central Arid Zone Research Institute, Jodhpur 342003, Rajasthan, India

Published online: 12 April 2015

dynamic situation, regular monitoring of soil resources is essential for implementation of any corrective measures in future. However, assessment of soil resources through surveying remote desert tracts is not always feasible and therefore may rely on indirect estimation through remote sensing techniques.

Over the last few decades, the reflectance spectroscopy (RS) technique over the visible, near-infrared and short wave infrared (VNIR–SWIR) region (350–2500 nm) has emerged as a rapid non-invasive technique for the in situ estimation of soil properties (Ben-Dor et al. 2009). The technique has been successfully used for estimating soil organic matter content (Galvão and Vitorello 1998; Fox and Metla 2005), nitrogen content (Vagen et al. 2006), soil electrical conductivity (Shrestha 2006), cation exchange capacity (Fox and Metla 2005), iron content (Galvão and Vitorello 1998), soil colour (Mathieu et al. 1998), soil moisture content (Carlson et al. 1995; Gillies et al. 1997), soil carbonates (Lagacherie et al. 2008), and soil mineralogical composition (Clark 1999) among others. Recent studies have shown that proximal spectral reflectance may be used for estimating soil hydraulic properties (Santra et al. 2009). The RS approach has the following advantages: (1) less sample preparation (only drying and crushing), (2) non-invasive analysis, (3) no chemicals are required, (4) rapid measurement, (5) several soil properties can be estimated from a single scan, and (6) feasible technique in both laboratory and in situ conditions (ViscaraRossel et al. 2006).

Spectral signatures of soil in the VNIR region (350–2500 nm) are mainly due to electronic transition of atoms, overtones and combinations of the fundamental vibrations found in the mid-infrared region (2500–20,000 nm). The reflectance spectra of soil exhibits vibrational absorbance due to –OH functional group in minerals, and to –OH, –CH, and –NH organic functional groups in soil organic matter (ViscarraRossel and McBratney 1998; Reeves et al. 1999).The reflectance spectra of soils generally have three prominent absorption peaks at 1400, 1900 and 2200 nm. The absorption peaks at 1400 and 1900 nm are the water absorption bands (Leone and Sommer 2000), while the 2200 nm denotes the metal-hydroxyl stretching occurring due to clay mineral (Chabrillat et al. 2002).The other absorption bands are around 870 and 1000 nm for iron oxides, and between 2200 and 2500 nm for carbonates (Clark et al. 1990; Chang and Laird 2002).

Over the last decade, both space-borne and air-borne hyperspectral sensors have been used as a companion of several resource management systems. Successful implementation of VNIR–SWIR spectroscopic methods are needed for the successful exploitation of hyperspectral technology in managing soils from drylands. A basic requirement in the RS approach is the availability of robust relationships between soil property of interest and its corresponding reflectance spectra (Lagacherie et al. 2008). Generally, large databases on soil properties and soil reflectance spectra are required to develop such relationships. This paved the way for concept of the development of a global soil spectral library that can be used for local or regional scale predictions and thereby uplifting soil spectroscopy as a robust method for soil analysis and mapping (ViscarraRossel 2009). There is a need for evaluating the capability of RS approach locally, regional and with respect to soil types. With a large variation in soil properties across the country, there is a requirement for developing larger spectral libraries for different regions as the calibrations are soil specific. In addition the global calibrations are often less accurate in comparison with regional or local scale calibrations (Brown 2007; Stevens et al. 2010). In this regard, there is a necessity to perform soil specific calibrations at local scale for characterizing the soil.

Literature survey revealed that there are limited reports on predicting soil properties in arid regions or drylands through RS approach. Therefore, the present study aims at generating a soil spectral library of drylands in India with a view to develop the spectral reflectance-based proximal sensing technique for rapid estimation of different soil properties.

Materials and methods

Study area

The study was carried out in western Rajasthan, India, which occupies 62 % area of hot arid ecosystem in India consisting of twelve districts of Rajasthan, e.g., Jaisalmer, Barmer, Jalore, Pali, Jodhpur, Bikaner, Nagaur, Sikar, Churu, Ganganagar, Hanumangarh, Jhunjhunu and Sikar (Fig. 1). In western Rajasthan, soils are majorly classified under two soil order; Entisol and Aridisol, which constitutes about 61 and 38 % of the total area of western Rajasthan, respectively. Torripsamments is the major great groups under Entisol covering 87.7 % area, whereas Haplocambids are the major great group under Aridisol covering 71.1 % area. Soils under Entisols are generally found in those parts of western Rajasthan where high aeolian activities are observed. Average soil depth of soil profiles under Entisols is 118 cm and bulk density of these soils is 1.58 Mg m^{-3}. Surface horizon is rich in sand content in comparison to subsurface horizons. Soil texture is sandy with average sand content of 89 %. In contrast to Entisols, Aridisols are dominant in those parts of arid regions where aeolian activity is comparatively less. Average

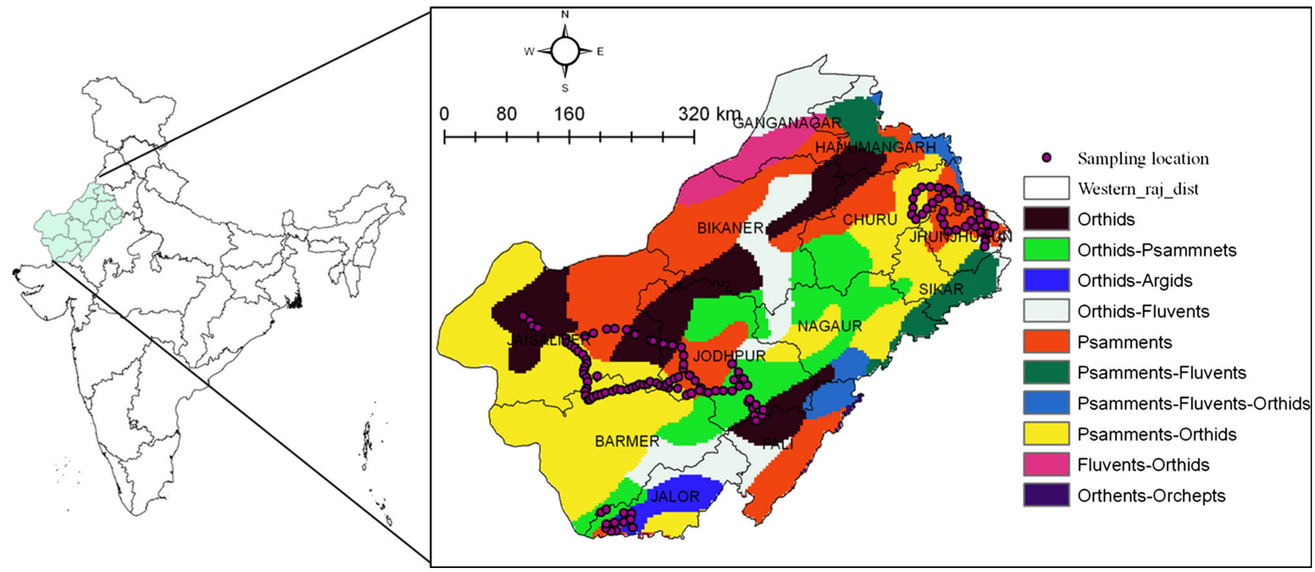

Fig. 1 Location of study area in India; sampling locations on soil maps with district boundary of western Rajasthan has been shown in *right frame of the figure*

depth of this type of soil is 100 cm with well-demarcated horizons. Concretions of calcite below the soil profile are a common feature of these soils. Average bulk density of these soils is 1.43 Mg m^{-3} whereas average sand content is lesser than Entisol and is around 70 %.

Soil sampling and laboratory analysis

Surface soil samples from 138 locations in western Rajasthan spreading over 20 million ha area were collected in four sampling campaigns covering the major soil types of the region (Fig. 1). Since the area is very large, we could not collect samples from each districts in western Rajasthan; however, the collected samples represented two dominant soil orders Aridisols and Entisols found in Rajasthan. Limited budget was the major constraint of collecting well distributed samples throughout western Rajasthan. However, the sampled locations covered Orthids, Psamments, Psamments–Orthids and Orthids–Psamments, which are the major suborders found in western Rajasthan. Thus, the sampling strategy adopted in this study included major soil resources of western Rajasthan.

Collected soils were air-dried, ground manually, and transferred through a 2 mm sieve and stored in plastic containers for laboratory analysis. The processed soils were used to determine pH, electrical conductivity (EC), soil organic carbon (OC) contents and particle size distribution using respective standard procedures. Soil pH was measured potentiometrically in soil–water slurry in the ratio 1:2 using an electronic pH meter equipped with combination electrodes that contain hydrogen ion sensitive electrode and a reference electrode. Prior to sampling routine calibration of the instrument was performed with buffer solutions of known pH values. Electrical conductivity of soil was determined 1:2.5 soil to water ratio using a Thermo Orion electrical conductivity meter. Chromic acid digestion method, commonly known as the Walkley and Black method, was used to determine SOC contents (Walkley and Black 1934). Proportion of sand (%), silt (%) and clay (%) in the soil was determined by international pipette method (Gee and Bauder 1986).

Measurement of soil spectral reflectance

Soil reflectance spectra was acquired using a portable spectroradiometer (Model: FieldSpec® 3, Analytical Spectral Devices, CO, USA) equipped with a high intensity contact probe at about 1 nm resolution over a wavelength range of 350–2500 nm. About 50 g of soil was taken in an aluminium soil moisture can (5 cm diameter) and the surface was levelled with a rubber cork used as a mallet. Four reflectance spectra were taken for each soil sample from different quadrants over the central area of the container. To optimize the measurements spectral reflectance from a Spectralon standard white reference panel (99 % reflectance, Labsphere) was acquired prior to that for soil samples. The scan number was set to 30 for obtaining an average of 30 scans for each spectrum. In this fashion, the reflectance spectra of each soil sample was acquired and saved with the help of RS^3 software associated with the radiometer.

Extraction of band reflectance

Band reflectance was derived from raw spectral reflectance data keeping in view of future translation of developed algorithm to remote sensing platform. The spectral bands of the LISS-III, LISS-IV and AWiFS camera of IRS-P6 satellite and Operational Land Imager (OLI) sensor on-board Landsat-8 satellite were selected for estimating band reflectance values from $R(\lambda)$ obtained from the spectroradiometer. The details of spectral bands of above said cameras of IRS-P6 and Landsat-8 are given in Table 1. The following relationship was used for estimating the band reflectance:

$$R_i = \frac{\sum_{\lambda u_i}^{\lambda l_i} R(\lambda)\varphi_i(\lambda)}{\sum_{\lambda u_i}^{\lambda l_i} \varphi_i(\lambda)} \tag{1}$$

Where R_i is the calculated ith band reflectance, λu_i is the upper boundary of band i, λl_i is the lower boundary of band i, $R(\lambda)$ is the reflectance for wavelength i, $\varphi(\lambda)$ is the spectral response function of band i.

Principal component analysis of spectra

Principal component analysis (PCA) involves a mathematical procedure that transforms a number of possibly correlated variables into a number of uncorrelated variables called principal components, related to the original variables by an orthogonal transformation. In this study, PCA of spectra was carried out to reduce the dimension of raw data from 2151 (350–2500 nm) to a limited number for better realization and further analysis. Principal components (PCs) describing cumulatively about 95 % of the total variation in data, were considered in final spectral data with reduced dimensions. The eigenvalue of each PC indicates the amount of variation explained. To establish the physical significance of the PCs, relationship between original variables and eigenvectors, commonly referred to as factor loadings (Hair et al. 1995), were checked. The original variables with large loadings on a given PC determine its physical significance. Scores for major PCs across soil samples were then calculated for further analysis.

Development of linear regression model

Models were developed using the multiple linear regression (MLR) equation of the form given below to estimate soil properties using soil spectral data:

$$Y = a_0 + \sum_{k=1}^{k} a_k X_k \tag{2}$$

where Y is the dependent variable (soil properties), X_k is the kth independent variable (spectral data), a_0, a_1, a_2, …, a_k

Table 1 Spectral bands of LISS-III, LISS-IV and AWiFS camera onboard RS-P6 satellite and of Operational Land Imager (OLI) onboard Landsat-8 satellite

Satellite	Camera	Band	Band region	Spectral width (nm)	Spatial resolution (m)
IRS-P6	LISS-III	B2	Green	520–590	23.5
		B3	Red	620–680	
		B4	NIR	770–860	
		B5	SWIR	1550–1700	
	LISS-IV	B2	Green	520–590	5.8
		B3	Red	620–680	
		B4	NIR	770–860	
	AWiFS	B2	Green	520–590	56
		B3	Red	620–680	
		B4	NIR	770–860	
		B5	SWIR	1550–1700	
Landsat-8	Operational Land Imager (OLI)	Band 1	Coastal aerosol	430–450	30
		Band 2	Blue	450–510	30
		Band 3	Green	530–590	30
		Band 4	Red	640–670	30
		Band 5	Near infrared (NIR)	850–880	30
		Band 6	SWIR 1	1570–1650	30
		Band 7	SWIR 2	2110–2290	30
		Band 8	Panchromatic	500–680	15
		Band 9	Cirrus	1360–1380	30

are regression coefficients and k is the number of independent variables in the regression equation. Three sets of spectral data were used to develop the linear model. First set consisted of major PCs obtained from PCA. Second set consisted of four band reflectance data corresponding to spectral bands of IRS-P6 satellite. Third set consisted of five band reflectance data corresponding to spectral bands of OLI onboard Landsat-8. The MLR equations were developed using the *lm* function of R-package. Before, running the *lm* function, stepwise procedure with both forward and backward approach was applied to each set of independent variable to select significant input variables. Finally, developed linear model was validated using tenfold cross validation approach and prediction accuracy of soil properties was calculated.

Partial least square regression (PLSR)

The PLSR method is similar to PCA, except that both predictor and response variables are used to build vectors with the greatest predictive power. The PLSR algorithm integrates the compression and regression steps and it selects successive orthogonal factors that maximize the covariance between predictor and response variables (ViscaraRossel and Behrens 2010). In this study, apart from linear model using PCs as input, PLSR model was also developed, keeping in view of the slight advantage of this technique over PCA because it takes on account the relationship between target variable and inputs while reducing the total data into few components. Since, PLSR technique reduces the data dimension considering the target variable and input relationship, for each target variable the component set will be different. Optimal number of components was decided by checking the root mean squared error (RMSE) of prediction of cross validation with different number of components. Optimal number of components corresponds to first local minima. Finally, the developed PLSR model with optimum number of components were used for tenfold cross validation and corresponding prediction accuracy of soil properties were measured.

Results and discussion

Soil reflectance spectra

Soil reflectance spectra of 138 surface soils from arid western Rajasthan are depicted in Fig. 2a. Wide variation in spectra across soil samples may be observed, specifically, in the SWIR region. Peak reflectance varied from as low as 0.3–0.6. Two distinct absorption features near 1400 and 1900 nm are characteristics of water absorption features of these soil spectra. All these spectra depicted in Fig. 2a represent different land use situations in western Rajasthan covering open sand dunes, stabilised sand dunes, wastelands, open scrubs, grasslands and cultivated area etc.

Selected soil reflectance spectra representing two major land forms in western Rajasthan, open dunes and arid grasslands dominated with Sewan grass (*Lasiurus sindicus*) are shown in Fig. 2b. It has been observed that from the figure that the overall soil spectral brightness in arid grasslands is higher than that in open dunes, which may be due to the difference in particle size distribution and carbon content. Sand content in sand dune site and arid grassland site was 96 and 90 %, respectively, whereas OC content was more (0.28 %) in sand dune site than arid grassland site (0.22 %). However, the specific analysis of these samples revealed that inorganic carbon content in the form of $CaCO_3$ was more (0.36 %) in arid grassland site than sand dune site (0.28 %), which ultimately resulted in brightness of the spectra in case of arid grasslands.

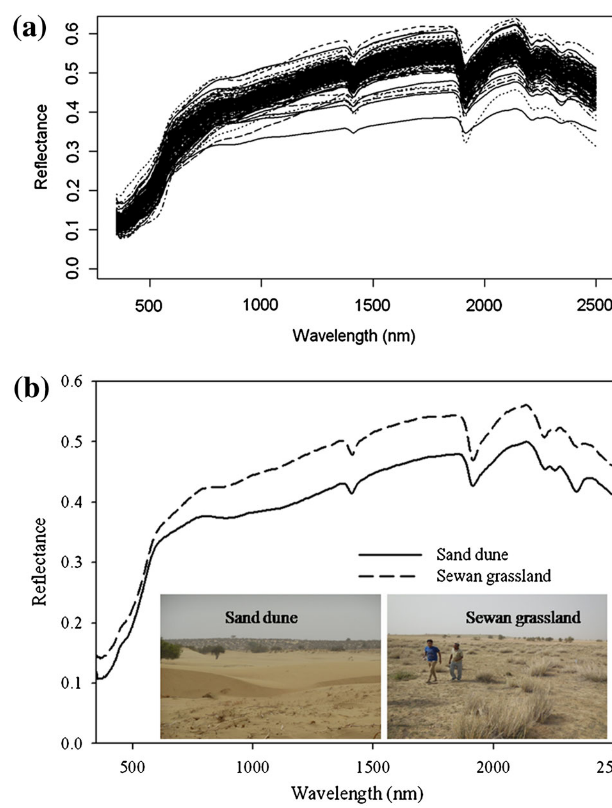

Fig. 2 Soil spectral signatures in arid western Rajasthan; **a** measured reflectance spectra of 138 surface soil samples and **b** selected reflectance spectra for open dune and arid grassland situation from Jaisalmer, Rajasthan (field views of sampling locations of the selected spectra are shown in the *bottom two frames*)

Soil properties

Descriptive statistics

Descriptive statistics of soil properties are shown in Table 2. Soil pH is generally found more than 7 with an average value of 7.9. Such soil reactions are characteristics of desert soil. Soil EC ranged widely from as low as 6–805 mS m^{-1}.

Organic carbon content has been found very low for most of the soil samples with an average content of 0.19 %. Average sand content of soil samples from western Rajasthan is 90.1 %, representing the dominance of aeolian activity in the region, a typical feature of arid regions. Consequently, clay and silt contents of soils are very low, 5.4 and 4.5 %, respectively. Histogram of soil properties are depicted in Fig. 3. From the histogram plot, it is clear that only soil pH is normally distributed, whereas other soil properties are skewed. Electrical conductivity is negatively skewed with almost half of the soil samples having EC of <200 mS m^{-1}; similarly, sand content is highly negatively skewed with most of the soil samples having value >80 %. Organic carbon contents of majority of soil samples have been found in low category (<0.5 %). Except few soil samples, silt content has been found <10 %. Clay content has been found within a range of 0–15 % except for some outliers.

Correlation matrix

Correlation among soil properties is presented in Table 3. Organic carbon content appears to be significantly correlated with EC ($r = 0.37$**), sand content ($r = -0.49$**), and clay content ($r = 0.44$**).

Before developing the linear model for predicting soil properties, correlations of them with PCs and band reflectance were checked. Correlation between soil properties and PCs is presented in Table 3. PC1 showed significant correlation with OC and clay content. PC2 showed significant correlation with OC content only. PC3 showed significant relation with particle size distribution especially sand and clay content. From the correlation table, it is noted that pH and EC were not significantly correlated with any of three PCs.

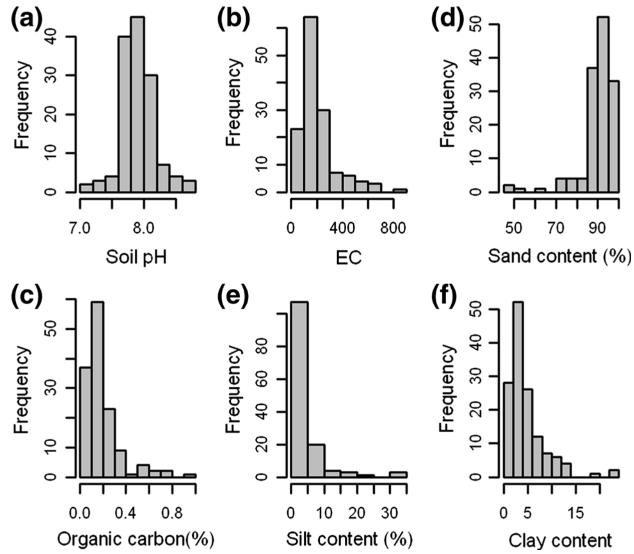

Fig. 3 Histogram of soil properties; **a** pH, **b** electrical conductivity (mS m^{-1}), **c** organic carbon content (%), **d** sand content (%), **e** silt content (%) and **f** clay content (%)

Correlation between soil properties and derived band reflectance from spectral data corresponding to spectral bands of LISS-III, LISS-IV, and AWiFS camera onboard IRS-P6 is given in Table 3. Spectral band reflectance in red (B3) and NIR (B4) region have been found significantly correlated with OC, sand, silt and clay content. Soil pH and EC were not found significantly correlated with spectral band reflectance. Moreover, spectral band reflectance in SWIR region (B5) was not correlated with soil properties.

Correlation between soil properties and derived band reflectance from spectral data corresponding to spectral bands of OLI camera on-board Landsat-8 is given in Table 3. Spectral band reflectance in red region (Band 4) has been found significantly correlated with OC, sand, silt and clay content. Band reflectance in green (Band 3) and NIR region (Band 5) were also found significantly correlated with OC, sand and clay contents. Soil pH and EC were not found significantly correlated with any band reflectance. Moreover, spectral band reflectance in SWIR region (Band 6 and 7) were not found correlated with most soil properties except the negative correlation between Band 7 and clay content ($r = 0.22$*).

Table 2 Descriptive statistics of soil properties from western Rajasthan, India

Soil properties	Minimum	Mean	Maximum	Standard deviation
pH	7.02	7.90	8.73	0.27
Electrical conductivity (EC) (mS m^{-1})	6	209	805	136
Organic carbon content (%)	0.01	0.19	0.94	0.15
Sand content (%)	46.0	90.1	98.0	8.53
Silt content (%)	0	4.5	34.4	5.45
Clay content (%)	0	5.4	23.40	3.98

Table 3 Correlation matrix of soil properties and different spectral signatures

Soil properties	pH	EC	OC	Sand	Silt	Clay
pH	1					
EC	0.10	1				
OC	0.04	0.37**	1			
Sand	0.04	−0.14	−0.49**	1		
Silt	−0.09	0.11	0.45**	−0.93**	1	
Clay	0.03	0.14	0.44**	−0.87**	0.63*	1
PC1	0.07	−0.03	0.21*	0.12	−0.06	0.18*
PC2	0.14	0.13	0.28**	−0.08	0.07	0.08
PC3	0.10	0.02	0.13	−0.64**	0.53**	0.64**
IRS P6-B2	0.10	−0.11	−0.33**	0.22**	−0.14	−0.27**
IRS P6-B3	0.10	−0.13	−0.45**	0.32**	−0.26**	−0.34**
IRS P6-B4	0.13	−0.13	−0.43**	0.24**	−0.18*	−0.26**
IRS P6-B5	0.04	0.02	−0.1	−0.02	0.07	−0.05
Landsat 8-Band 3	0.10	−0.11	−0.33**	0.22**	−0.15	−0.27**
Landsat 8-Band 4	0.10	−0.13	−0.45**	0.33**	−0.26**	−0.34**
Landsat 8-Band 5	0.13	−0.13	−0.38**	0.17*	−0.12	−0.21*
Landsat 8-Band 6	0.04	0.02	−0.10	−0.02	0.07	−0.05
Landsat 8-Band 7	0.02	−0.01	−0.11	0.15	−0.07	−0.22*

Landsat 8-Band 3, 4, 5, 6 and 7 are derived band reflectance corresponding to spectral bands of Operational Land Imager (OLI) on-board Landsat-8

* Significant at $p < 0.05$, ** significant at $p < 0.01$; PC—principal components (PCs) of soil reflectance spectra in VIS–NIR–SWIR region; B2, B3, B4 and B5 are band reflectance corresponding to spectral bands of LISS-III, LISS-IV and AWiFS camera on-board IRS-P6

Principal components of reflectance spectra

Variance explained by major principal components showed that three major PCs explain about 95 % variation in dataset. Therefore, original spectral dataset of 2151 dimension may be reduced to three prominent dimensions of the three principal components. Loading factor, indicating the weight for each wavelength reflectance, is depicted in Fig. 4. It may be seen that the first principal component has positive weight for the maximum portion of VIS–NIR–SWIR wavelength region except for some lower values in the visible region, and thus may be represented as the overall brightness of the spectra. More is the value of first PC; higher will be the brightness or height of the spectra from *x*-axis. Second major principal component has negative but higher loadings in the visible region and has been changed to positive higher values in SWIR region and, thus, may be considered to represent slope of the reflectance spectra. Third principal component has higher loadings in the wavelength regions with absorption features and, thus, may be considered to represent the absorption features of the spectra.

Scores of principal components

Scores of three major principal components clearly showed that soil samples are well separated from each other in four distinct quadrants in PC1 vs PC2 plot and PC1 vs PC3 plot, except few outliers. PC1 score generally varied from −100 to 100 whereas PC2 and PC3 varied from −40 to 40 to −20 to 20, respectively. Sample with PC1 value of −197.37 may be considered as an outlier, and examined in the dataset, which revealed that the soil sample was collected from an area dominated with red sandstone rocks, and thus was totally different from other soils. A nearby sample from this outlier sample also showed very small value of −111.47. Soil samples collected from Churu district located at north-eastern part of western Rajasthan have shown positive values for both PC1 and PC2 and, thus, are distinctly separated from soils collected from Jalore district having negative PC1 score but positive PC2 score.

Hyperspectral algorithm for estimation of soil properties

Linear model for estimating soil properties from spectral signatures

Keeping in view the correlation between soil properties and PCs of reflectance spectra, linear regression models were developed. However, before developing the model stepwise regression was run in both forward and backward approach and significant PCs were selected to develop the model. In case of OC as target variable, all PCs were found

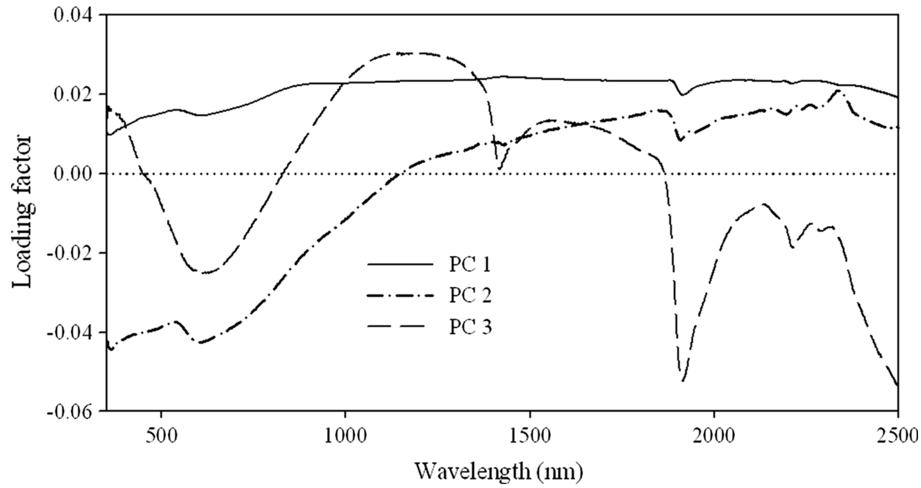

Fig. 4 Loading factor of three major principal components derived from raw reflectance spectra in VIS–NIR–SWIR region

significant in the stepwise procedure whereas PC1 and PC3 were found significant for sand content, silt content and clay content. Linear regression models developed with selected significant inputs from stepwise procedure are presented in Table 4. Models for predicting sand and clay content were found better with R^2 value of 0.41 and 0.43 than for predicting OC content and silt content with R^2 value of 0.12 and 0.27, respectively. Since the correlation of spectral data with pH and EC was very poor, the model performance was also very poor as reflected in R^2 value of the developed model and thus may be neglected.

Stepwise analysis of soil properties as target variable to estimate from IRS-P6 bands revealed that B2 and B3 were selected in most cases whereas band reflectance in SWIR region (B5) has been additionally selected as significant input for sand and clay content (Table 4). Similar to the model performance with PCs, pH and EC were not able to predict satisfactorily from IRS-P6 band data. The predictive performance of sand, silt and clay was better with PCs of spectral data in VIS–NIR–SWIR as input than the band reflectance corresponding to IRS-P6 bands. However, it is interesting to note here that OC can be better predicted from band reflectance corresponding to IRS-P6 bands ($R^2 = 0.27$) than from PCs of spectral data in VIS–NIR–SWIR ($R^2 = 0.12$).

Developed linear models for estimating soil properties from derived band reflectance corresponding to OLI camera on-board Landsat-8 are given in Table 4. Similar to findings mentioned above, models for predicting OC, sand, silt, and clay content from derived OLI band data were found satisfactory and even better than from PCs of spectral data and derived IRS-P6 bands whereas model performance for pH and EC was very poor. It is interesting to note here that band reflectance in SWIR regions (Band 6 and 7) have been found as significant inputs for estimating sand, silt and clay contents.

Cross validation of linear model

Developed models were cross validated using tenfold cross validation approach, where whole dataset was divided into ten segments randomly and models were developed from combined data of nine segments and then tested on remaining one segment.

Observed and predicted values of soil properties using all type of developed spectral models (see Table 4) were plotted and RMSE values were checked. The best predictions of soil properties were observed corresponding to models with Landsat-8 OLI band reflectance data as a predictor variable and are shown in 1:1 plot (Fig. 5). Since, the R^2 values of the models for soil pH and EC were very poor, cross validation was not performed on these two soil properties.

In case of PC-based model, RMSE value for organic carbon content was observed 0.15, whereas for sand, silt and clay content it was 6.75, 4.76 and 3.02, respectively. RMSE values of predicted soil properties using model involving derived spectral band data corresponding to available spectral bands of IRS-P6 were 0.13, 7.84, 5.05 and 3.75, respectively for OC, sand, silt and clay content. Comparatively, RMSE value for OC was lower than PCs-based model whereas it was higher for sand, silt and clay content than PCs-based model. RMSE values of predicted soil properties by Landsat 8 OLI band based model were lower than PCs—based and IRS-P6 band—based models developed in this study and were found 0.13, 6.58, 4.68 and 3.03 for OC, sand, silt and clay contents, respectively. Observed and predicted values of soil properties using derived OLI band based model in Fig. 5 shows that except for few soil samples with high silt, clay and OC contents, most of the points lies around 1:1 line. It indicates that derived band data corresponding to OLI bands of Landsat-8 data are better than other spectral information tested here to

Table 4 Developed spectral algorithms for estimating soil properties using principal components of soil reflectance spectra in VIS–NIR–SWIR region, derived IRS-P6 band reflectance and Landsat-8 OLI band reflectance

Model type	Model equation	R^2
PCs of hyperspectral soil reflectance based model	pH = 7.90 × PC1 − 0.002 × PC2	0.01
	EC = 209.23 + 1.049 × PC2	0.01
	OC = 0.192 − 0.0008 × PC1 + 0.002 × PC2 + 0.002 × PC3	0.12
	Sand = 90.15 + 0.025 × PC1 − 0.537 × PC3	0.41
	Silt = 4.46 + 0.284 × PC3	0.27
	Clay = 5.40 − 0.017 × PC1 + 0.252 × PC3	0.43
Derived IRS-P6 band reflectance based model[a]	pH = 7.70 + 4.66 × B2 − 10.76 × B3 + 11.81 × B4 − 4.07 × B5	0.02
	EC = 278 − 1319 × B4 + 882 × B5	0.02
	OC = 1.11 +3.82 × B2 − 5.64 × B3	0.27
	Sand = 66.3 − 304.5 × B2 + 605.7 × B3 − 366.3 × B4 + 88.1 × B5	0.20
	Silt = 11.53 + 157.52 × B2 − 264.82 × B3 + 102.10 × B4	0.17
	Clay = 18.19 + 109.65 × B2 − 255.16 × B3 + 175.8 × B4 − 49.42 × B5	0.16
Derived Landsat-8 OLI band reflectance based model[b]	pH = 7.32 + 1.42 × Band 5	0.01
	EC = 422 + 1240 × Band 4 − 3945 × Band 5 + 4217 × Band 6 − 2320 × Band 7	0.05
	OC = 1.12 + 3.72 × Band 3 − 5.56 × Band 4	0.27
	Sand = 52.8 − 168.5 × Band 3 + 316.1 Band 4 − 129.1 × Band 5 − 434.9 × Band 6 + 480.5 × Band 7	0.44
	Silt = 22.54 + 102.21 × Band 3 − 147.35 × Band 4 + 266.63 × Band 6 − 253.86 × Band 7	0.32
	Clay = 23.67 − 45.91 × Band 4 + 251.89 × Band 6 − 252.77 × Band 7	0.44

[a] Derived band reflectance corresponding to IRS-P6 bands of LISS-III, LISS-IV and AWiFS camera: B2 = 520–590 nm, B3 = 620–680 nm, B4 = 770–860 nm, B5 = 1550–1700 nm

[b] Derived band reflectance to Landsat-8 OLI bands: Band 3 = 530–590 nm, Band 4 = 640–670 nm, Band 5 = 850–880 nm, Band 6 = 1570–1650 nm, Band 7 = 2110–2290 nm

predict soil properties. It also indicates that remotely captured band reflectance in spectral bands of Landsat-8 satellite may be used for estimation of soil properties in future, specifically for arid regions or drylands of the world.

PLSR based estimation of soil properties

Apart from linear models involving PCs and derived band reflectance data, PLSR-based models were also developed to estimate soil properties. Although the PLSR technique is quite similar with PC analysis in regrouping the data with reduced dimensions capturing the most variations in dataset, it has the additional advantage of selecting components as per target variable. Leave-out-one cross validation approach with ten components as initial guess was used to obtain optimal number of components. Optimal number of components in PLSR model was decided corresponding to first local minima of RMSE estimate.

Scores of major components was plotted to see the percentage variation of data explained by respective component. For example, in case of OC content first major component explained 71.8 % variation in dataset. Overall, 97.7 % variation in dataset was explained by four major components. Here, it is notable that component 3 explaining 3.4 % variation in dataset was lower than 4.1 % variation explained by component 4. Yet, it has been ranked three because this component had greater role in estimation of OC than the component 4. Thus, the relative importance of components in relation to target variable has been considered in PLSR analysis, which is generally ignored in PCA.

Cross validation results of PLSR model are presented in Fig. 6 as 1:1 plot of measured vs predicted value of soil properties along with RMSE of prediction. RMSE values for OC, sand, silt, and clay content were observed 0.14, 6.28, 4.47, and 3.01, respectively. Performance of PLSR model was better than PCs-based linear model as well as derived band data-based models for estimation of sand, silt and clay content as shown by lower values of RMSE than the linear models. Similar to linear models, PLSR models have also shown poor performance to estimate soil pH and EC.

Proximally derived relationship to satellite images

Translating the developed laboratory-derived relationship between soil properties and proximally-measured spectral reflectance characteristics to remote sensing platform depends on several factors such as the spectral consistency of

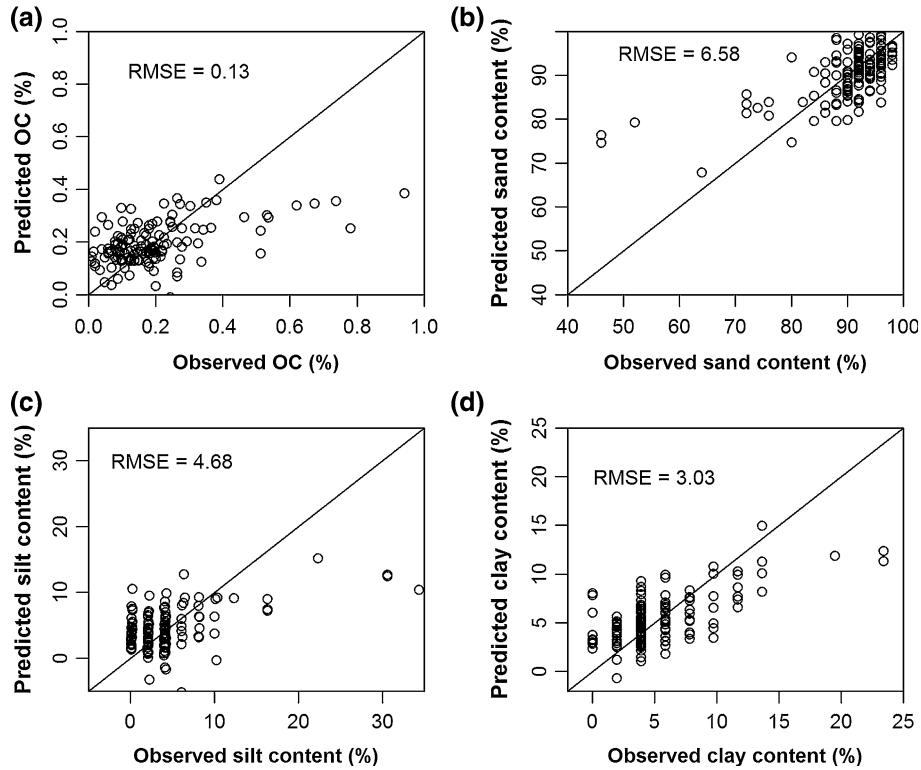

Fig. 5 Observed and predicted values of soil properties from derived band reflectance corresponding to OLI onboard Landsat-8 satellite; **a** organic carbon content (%), **b** sand content (%), **c** silt content (%) and **d** clay content

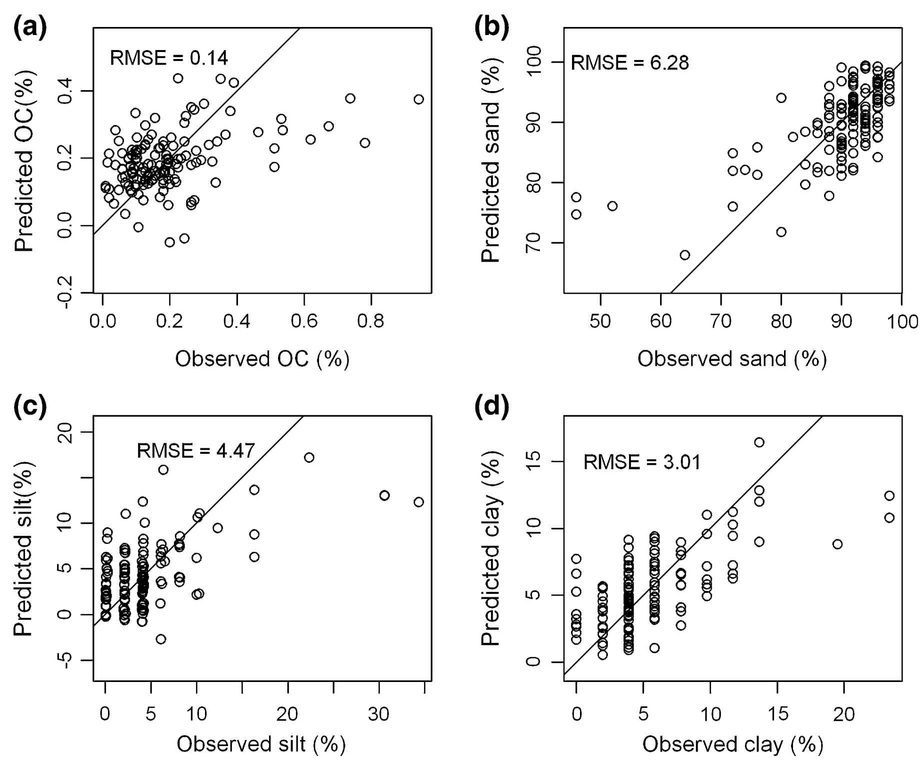

Fig. 6 Observed and predicted values of soil properties using partial least square regression (PLSR) analysis; **a** organic carbon content (%), **b** sand content (%), **c** silt content (%) and **d** clay content

satellite images, spectral resolution, atmospheric degradation of spectral behavior, surface roughness, soil moisture content, spatial resolution, the presence of gravels on surface, land surface composition, etc. To demonstrate this transferability, we estimated a few soil properties using the OLI band reflectance from the Landsat-8 scene (path 142, row 49) of 19th June 2013 downloaded from the earth explorer website (http://earthexplorer.usgs.gov/) and the

Table 5 Correlation of satellite measured OLI band reflectance of Landsat-8 and soil electrical conductivity and pH

Soil property	Band 1 (coastal aerosol)	Band 2 (blue)	Band 3 (green)	Band 4 (red)	Band 5 (NIR)	Band 6 (SWIR 1)	Band 7 (SWIR 2)
pH	−0.09	−0.02	−0.02	−0.03	−0.09	−0.19	−0.23
EC	−0.08	−0.11	−0.18	−0.18	−0.15	−0.10	−0.16

developed reflectance-based model (Table 5) from proximal sensing as a possible example. Figure 7 shows the estimated sand content for Shergarh Tehsil of Jodhpur district in western Rajasthan.

Soil pH and EC could not be satisfactorily predicted from soil spectral reflectance data in the present study although better estimation methods of these soil properties have been reported by previous researchers (Shrestha 2006; Lagacherie et al. 2008; Farifteh et al. 2007; Nawar et al. 2014). Most of these studies have reported the soil salinity assessment by using Landsat band reflectance or few derived indices such as normalized difference salinity index, salinity index, brightness index, etc. from satellite-measured band reflectance and not from proximally measured reflectance data. Therefore, we also tried to estimate soil EC and pH from OLI band reflectance of Landsat-8 using Landsat scene of 19th June 2013 (path 149, row 42). Land surface reflectance corresponding to sampling points located within the scene was extracted from Landsat image. Table 5 shows the correlation coefficients estimated using the measured and extracted EC and pH values from the Landsat-based surface maps. Negative correlation for EC shown in this table was also observed by Nawar et al. (2014) in contrast to the positive correlation between measured EC and Landsat-ETM+ derived EC observed by Shreshta et al. (2006). Stepwise regression of band reflectance and soil properties showed that OLI Band 2 and OLI Band 3 reflectance data are significant to predict soil EC (mS m^{-1}) with following relationship:

$$EC = 447^{**} + 266612^{**} \times \text{Band } 2 - 18671^{**} \times \text{Band } 3 \ (\text{Adj } R^2 = 0.20)$$

(* and ** represents the significance at $p < 0.05$ and $p < 0.01$, respectively).

The prediction performance of the developed model to estimate soil pH from OLI band reflectance was poor:

$$pH = 8.337^{***} + 5.795 \times \text{Band } 4 - 5.430^{*} \times \text{Band } 7 (R^2 = 0.08)$$

(* and *** represents the significance at $p < 0.05$ and $p < 0.001$, respectively).

This may be because of the heterogeneity of land surface with scattered vegetation in most of the sampling points and small sample size of the collected soils in this study. However, an improvement in prediction performance of EC and pH was observed while Landsat-8 OLI band reflectance was used instead of derived band reflectance to develop the linear model. Similarly, improved prediction of sand, silt, clay, and OC was also observed with R^2 value of 0.70, 0.54, 0.58, and 0.54, respectively.

Conclusion

Soil samples ($N = 138$) from western Rajasthan were collected and analysed in laboratory to generate soil spectral library at VIS–NIR–SWIR (350–2500 nm) region of the electromagnetic spectrum along with the measurements of basic soil properties. Subsequently, soil spectral information was related with soil properties to develop spectral reflectance-based algorithms for the rapid assessment of soil resources of arid regions. Overall, following observations and/or conclusions may be made from this study:

1. A soil spectral library consisting of 138 spectra of surface soil representing different land use situation of western Rajasthan covering Jaisalmer, Barmer, Jodhpur, Pali, Churu and Jalore was developed. Along with these spectral signatures, basic soil properties e.g. pH, electrical conductivity, organic carbon content, sand content, silt content and clay content were measured and added to soil spectral library.
2. Principal components of raw reflectance spectra were calculated. Three major principal components were identified to explain the total variation in spectra. These principal components represented overall brightness, slope of VIS–NIR–SWIR region and the absorption features of the spectra, respectively. Band reflectance corresponding to spectral bands available in LISS-III, LISS-IV and AWiFS camera onboard IRS-P6 and spectral bands available in OLI onboard Landsat-8 was also derived from measured reflectance spectra.
3. Linear regression models were developed to relate soil properties with PCs and derived band reflectance. Sand and clay content of arid western Rajasthan were satisfactorily estimated from linear models involving PCs as the input variables ($R^2 = 0.41$–0.43). Organic carbon content was also found satisfactorily correlated

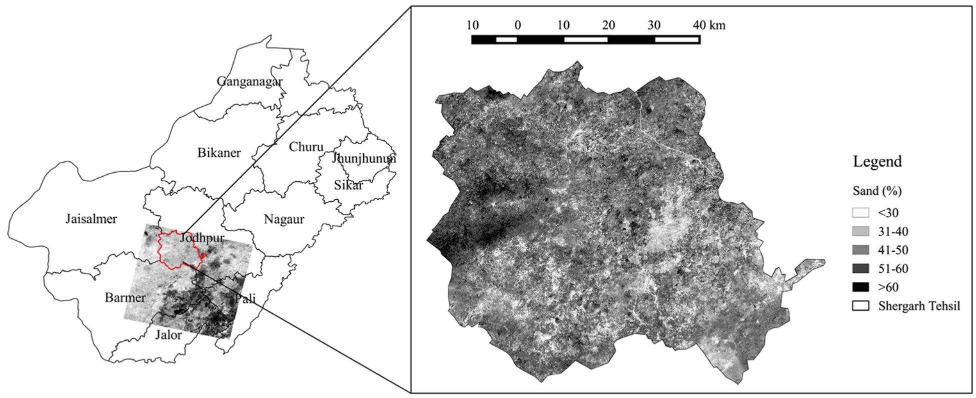

Fig. 7 Estimation of sand content in Shergarh Tehsil of Jodhpur District in western Rajasthan from OLI band reflectance of Landsat-8

with spectral data ($R^2 = 0.27$); however, soil pH and electrical conductivity could not be satisfactorily predicted using the spectral reflectance data.

4. Among different spectral data derived from raw reflectance spectra, band reflectance corresponding to OLI bands of Landsat-8 has been found to predict soil properties specifically sand and clay content. Organic carbon content of soils from arid western Rajasthan has been found to be best estimated by spectral bands of IRS-P6 satellite with predicted R^2 value of 0.27. Overall, derived band data corresponding to spectral bands of Lansat-8 satellite has been found satisfactory to estimate soil properties. Considering this performance, it is suggested to include Landsat-8 data in digital soil mapping approach of soil properties, specifically where availability of legacy soil data is limited and scattered.
5. PLSR model was also developed to relate soil properties with reflectance spectra. The PLSR-based models were found better than the PC-based multiple regression models for estimating soil properties from reflectance data. RMSEs of predicted organic carbon content (%), sand content (%) and clay content (%) were found to be 0.14, 4.47 and 3.01, respectively.

From the above findings, it may be concluded that soil spectral information has a great potential for the rapid assessment of soil resources. It has a special relevance in arid region for translation of ground-based spectral algorithm to remote sensing platform, since abstraction of soil reflectance by canopy vegetation and atmospheric cloud is negligible in the arid ecosystem. Moreover, robust algorithms relating soil resources with spectral data may be developed in future involving raw reflectance spectra as well as other secondary products.

Acknowledgments Financial support from Department of Science and Technology (DST), Govt of India under Fast Track Scheme for Young Scientist (SR/FTP/ES-60/2011) is highly acknowledged to carry out the present study. Technical support of Sh. Nand Kishore, ICAR-CAZRI Jodhpur during sampling and further processing the samples in the laboratory is also appreciated.

References

Ben-Dor E, Chabrillat S, Demattê JAM, Taylor GR, Hill J, Whiting ML (2009) Using imaging spectroscopy to study soil properties. Remote Sens Environ 13(1):S38–S55

Brown DJ (2007) Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. Geoderma 140:444–453

Carlson TN, Gillies RR, Schmugge TJ (1995) An interpretation of methodologies for indirect measurement of soil water content. Agric For Meteorol 77(3–4):191–205

Chabrillat S, Goetz AFH, Krosley S, Olsen HW (2002) Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. Remote Sens Environ 82:431–445

Chang CW, Laird DA (2002) Near-infrared reflectance spectroscopic analysis of soil C and N. Soil Sci 167:110–116

Clark RN (1999) Spectroscopy of rocks and minerals, and principles of spectroscopy. In: Rencz AN (ed) Remote sensing for the earth sciences: manual of remote sensing. American Society for Photogrammetry and Remote Sensing, USA, pp 3–58

Clark RN, King TVV, Klejwa M, Swayze GA, Vergo N (1990) High spectral resolution reflectance spectroscopy of minerals. J Geophys Res 95:12653–12680

Farifteh J, van dan Meer F, Atzberger C, Carranza EJM (2007) Quantitative analysis of salt affected soils reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). Remote Sens Environ 110:59–78

Fox GA, Metla R (2005) Soil property analysis using principal components analysis, soil line, and regression models. Soil Sci Soc Am J 69(6):1782–1788

Galvão LS, Vitorello Í (1998) Role of organic matter in obliterating the effects of iron on spectral reflectance and colour of Brazilian tropical soils. Int J Remote Sens 19:1969–1979

Gee GW, Bauder JW (1986) Particle-size analysis. In: Klute A (ed) Methods of soil analysis part I, physical and mineralogical methods. Soil Science Society of America, Madison, pp 383–412

Gillies RR, Carlson TN, Cui J, Kustas WP, Humes KS (1997) A verification of the 'triangle' method for obtaining surface soil water content energy fluxes from remote measurements of the normalized difference vegetation index (NDVI) and surface radiant temperature. Int J Remote Sens 18(15):3145–3166

Hair JF Jr, Anderson RE, Tatham RL, Black WC (1995) Multivariate data analysis. Prentice Hall, Englewood Cliffs

Kar A, Garg BK, Singh MP, Kathju S (2009) Trends in arid zone research in India. Central Arid Zone Research Institute, Jodhpur

Lagacherie P, Baret F, Feret JB, Madeira Netto J, Robbez-Masson JM (2008) Estimation of soil clay and calcium carbonate using laboratory, field, and airborne hyperspectral measurements. Remote Sens Environ 112(3):825–835

Leone AP, Sommer S (2000) Multivariate analysis of laboratory spectra for the assessment of soil development and soil degradation in the Southern Apennines (Italy). Remote Sens Environ 72:346–359

Mathieu R, Pouget M, Cervelle B, Escadafal R (1998) Relationship between satellite based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment. Remote Sens Environ 66:17–28

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

Reeves JB, McCarty GW, Meisenger JJ (1999) Near-infrared diffuse reflectance spectroscopy for the analysis of agricultural soil. J Near Infrared Spectrosc 7:179–193

Santra P, Sahoo RN, Das BS, Samal RN, Pattanaik AK, Gupta VK (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

Shrestha RP (2006) Relating soil electrical conductivity to remote sensing and other soil properties for assessing soil salinity in northeast Thailand. Land Degrad Dev 17(6):677–689

Stevens A, Udelhoven T, Denis A, Tychon B, Lioy R, Wesemael B (2010) Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. Geoderma 158:32–45

Vagen T-G, Shepherd KD, Walsh MG (2006) Sensing landscape level change in soil fertility following deforestation and conversion in the highlands of Madagascar using Vis–NIR spectroscopy. Geoderma 133:281–294

ViscaraRossel RA, Behrens T (2010) Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158:46–54

ViscaraRossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO (2006) Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131:59–75

ViscarraRossel RA (2009) The Soil Spectroscopy Group and the development of a global soil spectral library. NIR News 20(4):14–15

ViscarraRossel RA, McBratney AB (1998) Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content. Geoderma 85:19–39

Walkley A, Black IA (1934) An examination of Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci 37:29–38