Diffuse Reflectance Spectroscopic Approach for the Characterization of Soil Aggregate Size Distribution

M. C. Sarathjith
Bhabani S. Das*
Hitesh B. Vasava
Biswajita Mohanty
Anand S. Sahadevan
Agricultural and Food Engineering
Indian Institute of Technology
Kharagpur, WB 721302
India

Suhas P. Wani
Kanwar L. Sahrawat
ICRISAT
Patancheru
Hyderabad 502324
India

Assessment of soil structure and soil aggregation remains a challenging task. Routine methods such as dry- and wet-sieving approaches are generally time consuming and tedious, which calls for a robust, fast, and nondestructive method of soil aggregate characterization. Over the last two decades, diffuse reflectance spectroscopy (DRS) has emerged as a rapid and noninvasive technique for soil characterization. Combined with chemometric and data-mining algorithms, it provides an effective way of measuring several soil attributes and has the added advantage of being amenable to a remote sensing mode of operation. The objective of this study was to determine if the DRS approach could be used as a rapid, noninvasive technique to estimate soil aggregate characteristics. The DRS approach was examined for the estimation of soil aggregate characteristics such as the geometric mean diameter and two statistical parameters of the lognormal aggregate size distribution (ASD) functions using 910 soil samples from India representing three important soil groups. Results showed that the geometric mean diameter and the median aggregate size parameter provided excellent predictions, with ratio of performance deviation (RPD) values ranging from 1.99 to 2.28. The RPD value for the standard deviation of the ASD ranged from 1.36 to 1.72, suggesting moderate prediction. It was further observed that soil aggregates influence the incident electromagnetic radiation on soils primarily in the visible region and to some extent the shortwave- and near-infrared regions. Electronic transitions of Fe-bearing minerals, clay minerals, and C–H functional groups of organic matter may be responsible for modifying the spectral reflectance from soils in addition to the self-shadowing effects of surface roughness. The results of this study suggest that the chemometric approach may be combined with DRS to estimate soil aggregate size characteristics.

Abbreviations: ASD, aggregate size distribution; DRS, diffuse reflectance spectroscopy; EC, electrical conductivity; FD, first derivative; GMD, geometric mean diameter; NIR, near infrared; OC, organic carbon; PLSR, partial least square regression; R-VIP, product of absolute regression coefficient and absolute variable importance for projection after normalization; RPD, ratio of performance deviation; SWIR, shortwave infrared; VIS, visible.

Soil structure is much studied in different geoscience fields because of its role in providing the fundamental pore–solid network through which mass and energy are transported. The development of a pore fabric in soil is even considered as the fundamental process of regolith formation (Graham et al., 2010; Brantley, 2010). Recently, soil aggregation to form a specific soil structure has been linked with the sequestration of specific C fractions (Stamati et al., 2013). Although the importance of soil structure has long been recognized, its characterization remains a challenging task. The geometric mean diameter (GMD) and probabilistic frameworks of aggregate size distribution (ASD) functions (Fieller and Flenley, 1992) are, by far, the only quantitative descriptors of soil structure.

Experimentally, the aggregate size distribution is still measured using dry- and wet-sieving techniques that are generally time consuming and tedious, although efforts to use computer tomography, image analysis, and laser diffraction methods have met with limited success. A robust, fast, and nondestructive method of soil structural characterization is needed for this important soil attribute.

In the last few decades, diffuse reflectance spectroscopy (DRS) has emerged as a rapid and noninvasive technique for the estimation of several soil properties (Ben-Dor et al., 2009), including the soil aggregate-forming factors such as organic matter (Ben-Dor et al., 1997), clay content (Sørensen and Dalsgaard, 2005), and Fe oxide content (Richter et al., 2009). These aggregate-forming factors serve as chemical chromophores in soil (Ben-Dor et al., 2009) and are responsible for the specific spectral response of a soil to incident electromagnetic radiation. Thus, we hypothesized that the ASD may be estimated from spectral reflectance via these spectrally active aggregate-forming soil chromophores.

The effect of the soil particle size on the spectral reflectance has long been recognized (Bowers and Hanks, 1965; Salisbury and Hunt, 1968). Increasing albedo with decreasing particle size is generally attributed to the self-shadowing nature of the aggregates (Baumgardner et al., 1985). Wu et al. (2009) applied a wavelet transformation approach to remove the effects of soil roughness from reflectance spectra. Böttcher et al. (2012) showed that the DRS approach may be used to study the structural attributes of surface soils under field conditions. To the best of our knowledge, no studies have been reported to estimate ASD characteristics using soil spectral reflectance by the rapidly emerging chemometric methods such as partial least square regression (Viscarra Rossel and Behrens, 2010), among others. Thus, the objective of this study was to evaluate the potential of DRS for the estimation of soil aggregate size characteristics using samples from three broad soil groups of India.

MATERIALS AND METHODS
Study Area and Soil Samples

This study was performed using 247 red and 249 black soil samples from the state of Karnataka and 414 soil samples from the states of Odisha and West Bengal in India. Soil samples from Karnataka were collected as a part of a large-scale effort to enhance agricultural productivity across Karnataka by ICRISAT, Patancheru, Hyderabad, India. Black and red soils are classified as Vertisols and Alfisols, respectively. The soils of Odisha and West Bengal were collected as part of the building of a spectral library of Indian soils at the Indian Institute of Technology, Kharagpur, India. Soils from these two states are primarily lateritic, with a mix of coastal alluvium and soils from hilly terrain of the Eastern Ghat mountain ranges. These soils are a collection of Alfisols, Inceptisols, and Entisols, with limited samples falling under the Ultisol and Oxisol soil orders (Bhattacharyya et al., 2013). This set of soils is referred to here as *laterite soils* and are treated as a mixed pool, while the Alfisols and Vertisols of Karnataka are treated as specific soil groups.

Determination of Basic Soil Properties and Aggregate Size Distribution Characteristics

Soil samples were air dried, ground, and sifted through a 2-mm sieve and were stored for subsequent analyses. Among the basic soil properties, soil organic carbon (OC) content was determined using the chromic acid digestion method (Walkley and Black, 1934), and soil separates were estimated using the pipette method (Gee and Bauder, 1986). Soil pH and electrical conductivity (EC) were determined in a 1:2.5 soil/water solution.

The soil ASD was determined by dry sieving 100- to 200-g soil samples in a stack of eight sieves (1.18-, 0.3-, 0.2-, 0.18-, 0.125-, 0.09-, 0.075-, and 0.053-mm nominal diameters). A pan kept at the bottom of the sieve stack collected the fraction that passed through the bottom sieve (0.053 mm) and thus constituted the ninth soil fraction corresponding to <0.053-mm diameter. Soil samples retained in each sieve and the pan were weighed to estimate the mass fraction of soil aggregates (w_i) corresponding to the ith sieve diameter, which served as the lower limit of the aggregate size (aggregate diameter d). For instance, samples retained in the top 1.18-mm sieve were assigned the diameter of >1.18 mm and represented samples in the range of >1.18 to 2 mm in diameter. Samples passing through the bottommost sieve represented the finest fraction in our study and were arbitrarily assigned an aggregate diameter of 0.005 mm.

For modeling aggregate size attributes, the ASD data were analyzed in two different ways. First, the GMD for each soil sample was estimated using the geometric mean of the different size fractions:

$$\mathrm{MGD} = \frac{\sum w_i \ln(d_i)}{\sum w_i} \qquad [1]$$

We also estimated the mean weight diameter (MWD). Because both these parameters were highly correlated, only GMD data were used for subsequent analysis. In the second approach, data pairs of d_i and w_i were fitted to a lognormal ASD function (Buchan, 1989):

$$f(\ln d) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{(\ln d - \ln d_\mathrm{m})^2}{2\sigma^2}\right] \qquad [2]$$

where d_m is the mean aggregate diameter and σ is the standard deviation of the frequency distribution f. Based on this assumption, the cumulative aggregate mass fraction may be expressed as (Buchan et al., 1993; Hwang and Choi, 2006)

$$W(\ln d) = 0.5\,\mathrm{erfc}\left[\frac{\ln(d/d_\mathrm{m})}{\sigma\sqrt{2}}\right] \qquad [3]$$

The parameters d_m and σ in Eq. [3] were estimated using the SOLVER function in Microsoft Office Excel. The coefficient

of determination (R^2) values were found to be >0.97 for 906 soil samples out of the total of 910 sets of ASD data, suggesting that the log-transformed aggregate size was normally distributed. The R^2 values were >0.92 for the remaining four soils.

Collection of Spectral Reflectance Data

A portable spectroradiometer (FieldSpec3 FR, Analytical Spectral Devices) equipped with a contact probe (10-mm spot size) was used for spectral reflectance acquisition across the wavelength range of 350 to 2500 nm, covering the visible (VIS), near-infrared (NIR) and shortwave-infrared (SWIR) regions. About 50 g of soil was placed in an aluminum moisture box (10-cm diameter), and the soil surface was leveled (Mouazen et al., 2010) with a rubber cork used as a mallet. A spectrum from each quadrant of the moisture box was acquired by keeping the contact probe at the respective positions so as to have four reflectance spectra per soil sample. For each soil sample, a reference spectrum was also collected using a 9.2-cm diameter Spectralon white reference panel (Labsphere). In the data preparation stage, a third-order Savitsky–Golay filtering algorithm with a span length of 9 nm was used to smooth each soil spectrum (Vasques et al., 2010). Later, the four smoothed spectra were averaged to generate a representative spectral signature of each soil sample.

Development of Chemometric Model for Estimating Aggregate Size Distribution Parameters

A chemometric approach was followed to estimate the ASD parameters from the spectral reflectance values using a series of modeling approaches including partial least square regression (PLSR), principal component regression, and regression using discrete wavelet transformation (DWT) of the reflectance spectra. Before the application of these approaches, each soil aggregate attribute (response variable) was tested for normality using the Kolmogorov–Smirnov test at the 5% significance level. Wherever necessary, the data were suitably transformed such that the transformed response variable showed a normal distribution. Each soil spectrum was also clipped to retain the reflectance values in the wavelength range of 400 to 2447 nm (predictor variables) because the tails of the reflectance spectra are prone to noise. The resulting spectra were subjected to different transformations: absorbance, standard normal variate, first derivative (FD), Kubelka–Munk response, and discrete wavelets. Because soil spectra are generally multi-collinear, principal component analysis (PCA) and DWT were used for dimensionality reduction and subsequent regression analysis. It was observed, however, that the PLSR model in conjunction with the first-derivative transform of the soil spectra provided the best performance in estimating the ASD parameters. Hence, the results of the regression of the PCA and DWT parameters are not included in the subsequent discussion. The principal components were combined with the multiple linear regression model to remove outliers. The MatLab function *rcoplot* was used to implement this step. The method consisted of plotting the residuals (= the difference between the observed and predicted values of the output variable obtained from the regression model) with

their corresponding 95% confidence intervals termed as error bars for each observation. When the error bars for a given observation did not intersect the zero residual line, that observation was treated as an outlier. For testing the robustness of the modeling approach, the soils were divided into calibration and validation data sets in the ratio 3:1, as suggested by Viscarra Rossel and Lark (2009). Validation data comprised every fourth observation after sorting the respective soil property in ascending order. The two-parameter *t*-test and two-parameter *F*-test at the 5% level of significance were used to ensure the similarity of mean and variance between calibration and validation data sets, respectively.

The leave-one-out cross-validation approach (Viscarra Rossel, 2007) was used to select the number of latent variables or components in the PLSR analyses. The root mean squared error (RMSE), R^2, and ratio of performance deviation (RPD) were used as the model evaluation indices:

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n}\left[\left(\hat{Y}_i - Y_i\right)^2 \Big/ n\right]} \qquad [4]$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}_i\right)^2} \qquad [5]$$

$$\text{RPD} = \left(\frac{\sigma}{\text{RMSE}}\right)_{\text{Validation}} \qquad [6]$$

where Y is the observed response variable, $\hat{Y}_i$ is the predicted response variable, $\overline{Y}$ is the mean of the observed response variable, n is the number of soils involved in the analysis, and σ is the standard deviation of the observed values in the model validation data set. Based on the RPD statistics, the prediction accuracy of the model was categorized into accurate (RPD > 2), moderate (1.4 < RPD < 2), and poor (RPD < 1.4), as classified by Chang et al. (2001). Although the PLSR coefficients may be used for identifying significant wavelength bands (spectral features) responsible for the predictability of the soil aggregate characteristics, we used a combination of regression coefficient (Vasques et al., 2010) and variable importance for projection (Viscarra Rossel, 2008) for feature selection. Recently, Teófilo et al. (2009) suggested that a combination of the absolute value of such vectors after normalization may be used for feature selection. The product of the absolute regression coefficient and the absolute variable importance for projection after normalization (hereafter referred to as R-VIP) was used for the selection of the most relevant wavelengths for prediction. Large R-VIP values suggest that the corresponding wavelength is an important spectral feature in predicting the specified soil attribute.

RESULTS AND DISCUSSION
Descriptive Statistics for Basic Soil Properties and Aggregate Size Characteristics

Descriptive statistics of basic soil properties, GMD, and the mean and standard deviation for the lognormal ASD functions are shown in Table 1. Generally, Vertisols are fine textured and

Table 1. Descriptive statistics of basic soil properties and aggregate size distribution characteristics.

Soil attribute†	Vertisols Mean	Vertisols Range	Alfisols Mean	Alfisols Range	Laterite soil Mean	Laterite soil Range
	Basic soil properties					
Sand, %	65.6 (15)‡	21.7–87.8	78.1 (10)	50.9–91.6	45.4 (40)	2.0–92.0
Clay, %	14.6 (36)	4.3–43.4	12.4 (53)	2.1–34.3	30.5 (44)	0.0–74.1
pH	8.54 (6)	6.00–9.60	6.67 (21)	4.30–9.50	6.13 (16)	3.98–9.22
EC, μS/cm	0.6 (318)	0.1–28.2	0.4 (67)	0.1–0.9	419 (164)	30–8460
OC, %	0.40 (38)	0.11–0.93	0.38 (36)	0.11–0.91	0.80 (44)	0.01–2.12
	Aggregate size distribution characteristics					
GMD	0.31 (25)	0.17–0.52	0.21 (25)	0.13–0.45	0.19 (28)	0.05–0.32
d_m	0.54 (35)	0.22–1.19	0.33 (35)	0.17–0.91	0.28 (38)	0.09–0.63
σ	0.99 (12)	0.62–1.33	0.98 (12)	0.68–1.29	1.08 (25)	0.45–1.76

† EC, electrical conductivity; OC, organic C; GMD, aggregate mean diameter; d_m and σ, median and standard deviation, respectively, of the lognormal aggregate size distribution function.

‡ Values in parentheses are the coefficients of variation (%).

clayey in nature and Alfisols are dominantly coarse textured and less clayey. The lateritic soil pool had a wide range of soil properties because these soils were collected from a large geographical area representing different agroecosystems. Alfisols and Vertisols were distinctly different with regard to their clay mineralogy, with Vertisols being montmorillonitic and Alfisols kaolinitic. The lateritic soil group was primarily kaolinitic in the coastal belt and had a mixture of smectites in the hilly regions. The average OC for both the Karnataka soils was found to be low and similar, while in the lateritic pool it was high, being from a mixed pool of soils with a wide range of OC values (Table 1). The average pH underlined the slightly acidic nature of the Alfisols and lateritic soils as well as the alkaline nature of the Vertisol samples.

A comparison of the mean values of the ASD function parameters using a two-sample *t*-test at the 5% level of significance revealed that the Vertisols had larger mean aggregate diameters

Table 2. Correlation coefficients between aggregate size distribution (ASD) characteristics and basic soil properties of pH, electrical conductivity (EC), and organic C (OC), sand, and clay contents.

ASD characteristic†	Basic soil properties pH	EC	OC	Sand	Clay
	Vertisols				
GMD	−0.06	0.22**	0.13	−0.50***	0.52***
d_m	−0.08	0.23**	0.15	−0.49***	0.52***
σ	−0.14	0.19*	0.21*	−0.37***	0.36***
	Alfisols				
GMD	−0.13	0.04	−0.02	−0.54***	0.52***
d_m	−0.14	0.04	−0.02	−0.56***	0.53***
σ	0.08	0.06	0.05	−0.03	−0.04
	Laterite soil				
GMD	0.38***	0.03	−0.01	−0.45***	0.44***
d_m	0.40***	0.06	0.05	−0.46***	0.44***
σ	0.14*	0.10	0.19**	−0.35***	0.29***

* Significant at P < 0.01.

**Significant at P < 0.001.

***Significant at P < 0.0001.

† GMD, geometric mean diameter; d_m and σ, median and standard deviation, respectively, of the lognormal ASD function.

(GMD and d_m) than the remaining two soil groups. The parameter σ for the ASD functions appeared to be similar across all three soil groups. The coefficient of variation for this parameter in the Alfisols and Vertisols was low, suggesting that these two soils were more homogenous than the lateritic soil group. A major focus of this study was to examine the efficacy of the DRS approach in estimating the soil aggregate characteristics. Thus, the heterogeneity in the lateritic soil group and relatively homogenous nature of the Alfisols and Vertisols offered diverse data sets for testing this DRS approach.

The frequency distributions of most of the soil aggregate attributes were relatively less skewed. Simple logarithmic transformation yielded normal distributions for these parameters. The Kolmogorov–Smirnov test at the 5% significance level suggested that the GMD in the Alfisols and the parameter d_m in the Alfisols and laterites were lognormally distributed; these parameters in the remaining data sets were normally distributed. Correlation coefficients of the GMD, d_m, and σ with basic soil properties showed that both d_m and GMD were strongly correlated with the sand and clay contents in each soil group (Table 2). This result also held for σ except for the Alfisols. Both OC and EC showed very little correlation with these structural attributes. The lack of correlation with the OC contents even at the 99% confidence level may be a result of low OC contents in these soils, which suggests that OC contents may be playing a limited role in soil aggregation in these soils.

Prediction of Aggregate Size Distribution Parameters

Figure 1 shows the mean reflectance spectra for the Alfisol samples of the first quartile, interquartile, and third quartile. Corresponding GMD values were also averaged and are shown in this graph. This figure shows that the spectral reflectance of a soil decreases with an increase in aggregate size, as has been observed in earlier studies (Baumgardner et al., 1985; Bänninger and Flühler, 2004). Larger aggregate size increases the
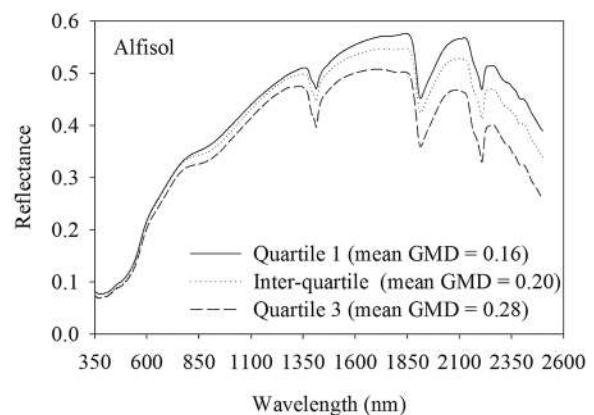
Fig. 1. Mean reflectance in each quartile of the geometric mean diameter (GMD).
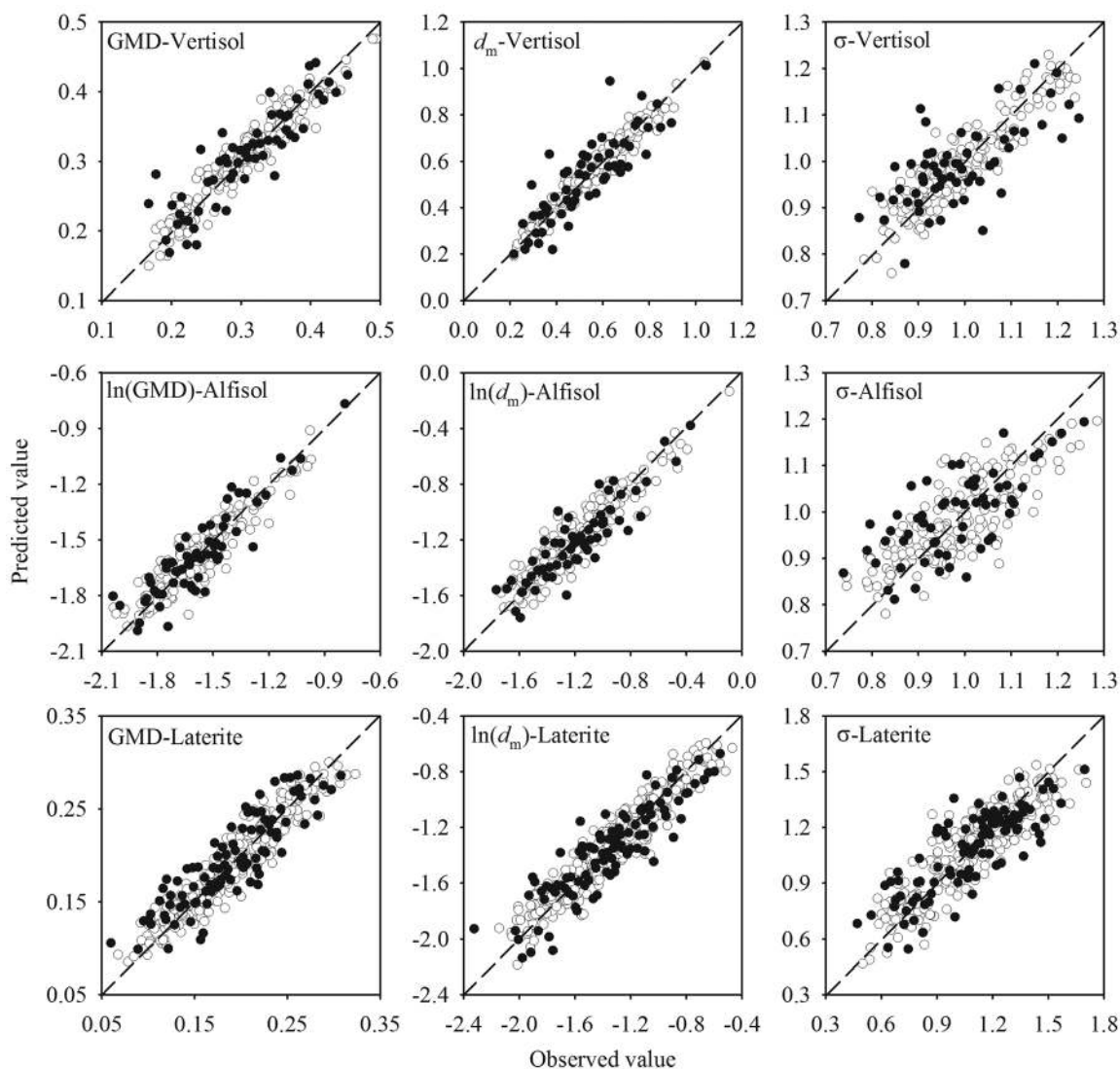
Fig. 2. Observed vs. predicted plot of the aggregate size distribution parameters geometric mean diameter (GMD) and the median and standard deviation of the lognormal aggregate size distribution function (d_m and σ, respectively). Open and closed circles represent the calibration and validation data sets, respectively.

roughness of a soil. Wu et al. (2009) applied a wavelet transformation approach to remove the effects of soil roughness on reflectance spectra. In this study, the pattern of decreasing reflectance with increasing aggregate size fraction has been exploited to develop a chemometric algorithm for estimating the aggregate size parameters.

Examination of the pretreatment approaches showed that the FD approach of transforming the reflectance spectra yielded the lowest Akaike information criterion values (data not shown) compared with all other transformations. Therefore, modeling results on the PLSR calibration and validation on the FD-transformed spectra are summarized in Fig. 2 and Table 3. Optimum numbers of latent variables used in the regression and the transformation responsible for the normality of the ASD characteristics across different soil groups are also included in the table. Even though RMSE is an unbiased estimator, it was less useful for comparing the performance of modeling results because some of the ASD characteristics were transformed into the logarithmic scale to ensure normality of the predicted

Table 3. Regression statistics for the prediction of aggregate size distribution (ASD) characteristics.

ASD characteristics†	LV‡	Calibration			Validation			
		n	R^2	RMSE	n	R^2	RMSE	RPD§
			Vertisols					
GMD	11	176	0.92	0.02	59	0.78	0.03	2.14
d_m	14	177	0.96	0.04	60	0.72	0.10	1.91
σ	11	174	0.82	0.04	59	0.45	0.08	1.36
			Alfisols					
GMD¶	9	174	0.85	0.08	59	0.80	0.11	2.28
d_m ¶	8	172	0.86	0.11	58	0.78	0.14	2.14
σ	8	175	0.64	0.07	59	0.47	0.08	1.38
			Laterite soil					
GMD	15	299	0.90	0.02	100	0.75	0.03	1.99
d_m ¶	14	296	0.90	0.11	99	0.79	0.17	2.18
σ	15	292	0.80	0.11	98	0.66	0.15	1.72

† GMD, geometric mean diameter; d_m and σ, median and standard deviation, respectively, of the lognormal ASD function.

‡ Number of latent variables.

§ Ratio of performance deviation.

¶ Variables subjected to natural logarithm transformation.
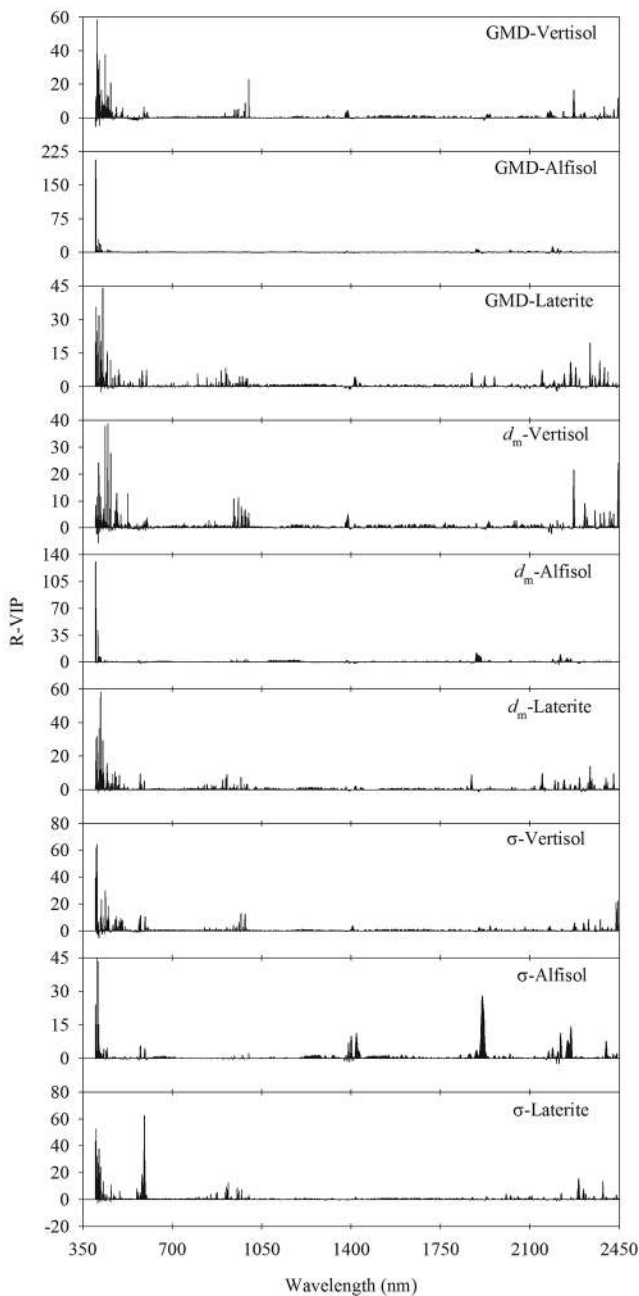
Fig. 3. Significant wavelengths for prediction of aggregate size distribution characteristics geometric mean diameter (GMD) and the median and standard deviation of the lognormal aggregate size distribution function (d_m and σ, respectively); R-VIP is the product of the absolute regression coefficient and the absolute variable importance for projection after normalization.

response variables. Hence, the RPD values may be used for comparing the PLSR statistic. Table 3 clearly shows that the PLSR approach provided accurate estimation of the GMDs across the three soil groups. High RPD values for the mean aggregate diameter d_m suggest that it was accurately estimated in the Alfisols and lateritic soils and moderately estimated in the Vertisols. The scenario for σ was poor in the Vertisols and Alfisols and moderate in the lateritic group. Poor predictability of σ may be related to the smaller coefficient of variation for this parameter in the Alfisols and Vertisols compared with the

lateritic soils. Nevertheless, compact distribution of observed and predicted values of all the aggregate size parameters along the 1:1 line in Fig. 2 suggests that the PLSR approach with the FD of the spectra may be a robust method of estimating the ASD parameters.

Although all the soil aggregate-forming soil constituents are active soil chromophores and it may be obvious that the parameters for the ASD ought to be well predicted by the DRS approach, soil aggregates have also been shown to directly influence the albedo (Baumgardner et al., 1985; Bänninger and Flühler, 2004). The R-VIP plot in Fig. 3 shows that the VIS part of the electromagnetic spectra is the most important wavelength region for the prediction of all the ASD parameters across the three soil groups used in this study. Chemically, these wavelength regions are known to be associated with the electronic transition of soil Fe oxides (Viscarra Rossel and Behrens, 2010), as summarized in Table 4. Thus, the ASD spectral signatures in the visible range may get confounded with the presence of extractable Fe in soils. Figure 3 also shows high R-VIP values around 900 nm (NIR region) and 2200–2300 nm (SWIR regions). The NIR regions are associated primarily with the electronic transitions of the Fe-bearing minerals, while the SWIR regions are associated with the presence of clay mineral signatures and C–H functional groups. Thus, these results suggest that soil aggregates influence incident electromagnetic radiation directly as physical chromophores and indirectly as chemical chromophores through their constituents.

SUMMARY

Diffuse reflectance spectroscopy is a rapidly emerging technology for noninvasive characterization of soils. Combined with the chemometric and data-mining algorithms, it provides an effective methodology for measuring several soil attributes; it has the added advantage of being amenable to the remote sensing mode of operation. Soil structure is an important integrative property of soils, and soil aggregation characteristics are quantitative descriptors of soil structure. Because aggregation is controlled by many spectrally active soil constituents, we tested the popular PLSR algorithm in the DRS approach to estimate soil aggregate characteristics such as the GMD and two statistical parameters of the lognormal ASD functions using 910 soil samples representing three important soil groups of India. The results show that all three parameters may be satisfactorily estimated using the DRS approach. The R-VIP plots showed further that soil aggregates influence the incident electromagnetic radiation on soils primarily in the visible region and to some extent the SWIR and NIR regions. Electronic transitions of Fe-bearing minerals, clay minerals, and C–H functional groups of organic matter may be responsible for modifying the spectral reflectance from soils in addition to the self-shadowing effects of surface roughness as proposed by Baumgardner et al. (1985). The results of this study suggest that the chemometric approach may be combined with DRS to estimate soil aggregate size characteristics.

Table 4. Assignment of functional groups for aggregate size distribution (ASD) characteristics for most relevant wavelengths (λ).

$\lambda_{Alfisol}$	$\lambda_{Vertisol}$	$\lambda_{Laterite}$	$\lambda_{Reported}$	Constituent	Possible functional group assignment	Reference[†]
				Geometric mean diameter		
400–420	400–500	400–500	400–500	Fe oxides	electronic transitions (ET) of goethite, hematite	[2]
	872		870	Fe oxides	ET band of Fe^{3+}	[3]
		891	884	Fe oxides	ET of hematite	[2]
		913	920	Fe oxides	ET of goethite	[2]
	1000		1000	organics	$4\nu1$ asymmetric–symmetric doublet	[2]
		2148	2142	organics	$4\nu1$ of polysaccharides	[2]
2189			2193	organics	3ν aromatic C=C of lignin/starch/wax/tannins	[1]
	2272		2275	organics	amide II of protein	[1]
		2281	2279	organics	$3\nu1$ of aliphatics	[2]
		2336	2336	clay	$3\nu3$ of carbonate	[2]
		2374	2372	clay	kaolinite	[3]
	2445		2450	clay	illite (poorly defined)	[2]
				Median of the lognormal ASD function		
400–416	400–500	400–500	400–500	Fe oxides	ET of goethite, hematite	[2]
	527		529	Fe oxides	ET of hematite	[2]
		898	900	Fe oxides	transition bands of Fe^{2+} and Fe^{3+}	[3]
	941		940	Fe oxides	$2\nu1 + \nu3$ of water	[2]
	1000		1000	organics	$4\nu1$ asymmetric–symmetric doublet	[2]
		2149	2142	organics	$4\nu1$ of polysaccharides	[2]
		2198	2200	clay	stretch of montmorillonite, illite	[3]
2221			2216	clay	illite	[3]
		2234	2230	clay	$\nu1 + \delta b$ of smectite	[2]
	2272		2275	organics	amide II of protein	[1]
		2295	2300	organics	3ν of CH_2, CH_3	[1]
	2316		2312	clay	kaolinite	[3]
		2336	2336	clay	$3\nu3$ of carbonate	[2]
	2377		2376	clay	kaolinite	[3]
	2446		2450	clay	illite (poorly defined)	[2]
				Standard deviation of the ASD function		
400–414	400–500	400–460	400–500	Fe oxides	ET of goethite, hematite	[2]
	570		550	Fe oxides	ET band of Fe^{2+} of hematite	[3]
		913	920	Fe oxides	ET of goethite	[2]
1390			1395	clay	$2\nu1a$ of kaolin doublet	[2]
1401			1400	Fe oxides	$2\nu1$ of hydroxyl	[2]
1402			1400	Fe oxides	$2\nu1$ of hydroxyl	[2]
1913			1915	Fe oxides	$\nu2+\nu3$ of water	[2]
2220			2216	clay	illite	[3]
		2310	2308	clay	kaolinite	[3]
		2310	2309	organics	$\nu + \nu4$ of CH stretch	[3]
	2376		2376	clay	kaolinite	[3]
		2386	2386	organics	$4\nu1$ of carbohydrates	[2]
	2444		2450	clay	illite (poorly defined)	[2]

† [1] Ben-Dor et al. (1997); [2] Viscarra Rossel and Behrens (2010); [3] Bayer et al. (2012).

REFERENCES

Bänninger, D., and H. Flühler. 2004. Modeling light scattering at soil surfaces. IEEE Trans. Geosci. Remote Sens. 42:1462–1471. doi:10.1109/TGRS.2004.828190

Baumgardner, M.F., L.F. Silva, L.L. Biehl, and E.R. Stoner. 1985. Reflectance properties of soils. Adv. Agron. 38:1–44. doi:10.1016/S0065-2113(08)60672-0

Bayer, A., M. Bachmann, A. Muller, and H. Kaufmann. 2012. A comparison of feature-based MLR and PLS regression techniques for the prediction of three soil constituents in a degraded South African ecosystem. Appl. Environ. Soil Sci. 10:1454–1463.

Ben-Dor, E., S. Chabrillat, J.A.M. Dematte, G.R. Taylor, J. Hill, M.L. Whiting, and S. Sommer. 2009. Using imaging spectroscopy to study soil properties. Remote Sens. Environ. 113:S38–S55. doi:10.1016/j.rse.2008.09.019

Ben-Dor, E., Y. Inbar, and Y. Chen. 1997. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. Remote Sens. Environ. 61:1–15. doi:10.1016/S0034-4257(96)00120-4

Bhattacharyya, T., D.K. Pal, C. Mandal, P. Chandran, S.K. Ray, D. Sarkar, et al. 2013. Soils of India: Historical perspective, classification and recent advances. Curr. Sci. 104:1308–1323.

Böttcher, K., C. Gläßer, and S.J. Mooney. 2012. Examining the relationship between soil structure and soil reflectance using soil pore structure characteristics obtained from image analysis. Remote Sens. Lett. 3:557–565. doi:10.1080/01431161.2011.640958

Bowers, S.A., and R.J. Hanks. 1965. Reflection of radiant energy from soils. Soil Sci. 100:130–138. doi:10.1097/00010694-196508000-00009

Brantley, S.L. 2010. Rock to regolith. Nat. Geosci. 3:305–306.

doi:10.1038/ngeo858

Buchan, G.D. 1989. Applicability of the simple lognormal model to particle size distribution in soils. Soil Sci. 147:155–161. doi:10.1097/00010694-198903000-00001

Buchan, G.D., K.S. Grewal, and A.B. Robson. 1993. Improved models of particle-size distribution: An illustration of model comparison techniques. Soil Sci. Soc. Am. J. 57:901–908. doi:10.2136/sssaj1993.03615995005700040004x

Chang, C.W., D.A. Laird, M.J. Mausbach, and C.R. Hurburgh. 2001. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. Soil Sci. Soc. Am. J. 65:480–490. doi:10.2136/sssaj2001.652480x

Fieller, N.R.J., and E.C. Flenley. 1992. Statistics of particle size data. Appl. Stat. 41:127–134. doi:10.2307/2347623

Gee, G.W., and J.W. Bauder. 1986. Particle-size analysis. In: A. Klute, editor, Methods of soil analysis. Part 1. Physical and mineralogical methods. SSSA Book Ser. 5. SSSA, Madison, WI. p. 383–412.

Graham, R.C., A.M. Rossi, and K.R. Hubbert. 2010. Rock to regolith conversion: Producing hospitable substrates for terrestrial ecosystems. GSA Today 2:4–9. doi:10.1130/GSAT57A.1

Hwang, S., and S. Choi. 2006. Use of a lognormal distribution model for estimating soil water retention curves from particle-size distribution data. J. Hydrol. 323:325–334. doi:10.1016/j.jhydrol.2005.09.005

Mouazen, A.M., B. Kuang, J. De Baerdemaeker, and H. Ramon. 2010. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. Geoderma 158:23–31. doi:10.1016/j.geoderma.2010.03.001

Richter, N., T. Jarmer, S. Chabrillat, C. Oyonarte, P. Hostert, and H. Kaufmann. 2009. Free iron oxide determination in Mediterranean soils using diffuse reflectance spectroscopy. Soil Sci. Soc. Am. J. 73:72–81. doi:10.2136/sssaj2008.0025

Salisbury, J.W., and G.R. Hunt. 1968. Martian surface materials: Effect of particle size on spectral behavior. Science 161:365–366.

Sørensen, L.K., and S. Dalsgaard. 2005. Determination of clay and other soil properties by near infrared spectroscopy. Soil Sci. Soc. Am. J. 69:159–167. doi:10.2136/sssaj2005.0159

Stamati, F.E., N.P. Nikolaidis, S. Banwart, and W.E.H. Blum. 2013. A coupled carbon, aggregation, and structure turnover (CAST) model for topsoils. Geoderma 211–212:51–64. doi:10.1016/j.geoderma.2013.06.014

Teófilo, R.F., J.P.A. Martins, and M.M.C. Ferreira. 2009. Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression. J. Chemom. 23:32–48. doi:10.1002/cem.1192

Vasques, G.M., S. Grunwald, and W.G. Harris. 2010. Spectroscopic models of soil organic carbon in Florida, USA. J. Environ. Qual. 39:923–934. doi:10.2134/jeq2009.0314

Viscarra Rossel, R.A. 2007. Robust modelling of soil diffuse reflectance spectra by bagging–partial least squares regression. J. Near Infrared Spectrosc. 15:39–47. doi:10.1255/jnirs.694

Viscarra Rossel, R.A. 2008. ParLeS: Software for chemometric analysis of spectroscopic data. Chemom. Intell. Lab. Syst. 90:72–83. doi:10.1016/j.chemolab.2007.06.006

Viscarra Rossel, R.A., and T. Behrens. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158:46–54. doi:10.1016/j.geoderma.2009.12.025

Viscarra Rossel, R.A., and R.M. Lark. 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. 60:453–464. doi:10.1111/j.1365-2389.2009.01121.x

Walkley, A.J., and I.A. Black. 1934. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci. 37:29–38. doi:10.1097/00010694-193401000-00003

Wu, C., A.R. Jacobson, M. Laba, and P.C. Baveye. 2009. Accounting for surface roughness effects in the near-infrared reflectance sensing of soils. Geoderma 152:171–180. doi:10.1016/j.geoderma.2009.06.002