

Influence of Bidirectional Reflectance Distribution Function in Estimating Basic Soil Properties Using Airborne Hyperspectral Data

Naveen Kumar Purushothaman^{1b}, Amanjit Premsagar, Mayank Raj, Israr Majeed^{2b}, N. Nagarjuna Reddy^{3b}, Lokesh Kumar Sinha, and Bhabani Sankar Das^{4b}

Abstract—Recent studies on hyperspectral remote sensing (HSR) have shown that the estimation accuracy of different vegetation characteristics improves when the HSR data are corrected for the bidirectional reflectance distribution function (BRDF) effects. Similar studies involving soil parameters are limited. Here, we used the BRDF-corrected HSR data collected using the airborne visible-infrared imaging spectrometer-next generation (AVIRIS-NG) sensor to estimate soil parameters over a 138-km² agricultural catchment. Surface soil samples were collected from 173 ground reference locations (GRLs) from this catchment to measure clay and sand contents, pH, electrical conductivity (EC), and soil organic carbon (SOC) contents. The BRDF correction was applied using the flexible BRDF (FlexBRDF) algorithm, and a polynomial unmixing approach was used to extract soil spectra from the corrected image. The BRDF correction successfully removed the shading effects and produced smooth transitions along the overlapping regions when multiple AVIRIS-NG images were mosaicked. Upon unmixing, soil spectra could be extracted at 140 GRLs when BRDF-corrected spectra were used, while uncorrected spectra produced soil spectra only for 114 GRLs. Chemometric models were validated using 109 common GRLs to compare estimation accuracy across laboratory-measured soil spectra (SS_{Lab}) and those obtained from unmixing of BRDF-corrected and uncorrected spectra. The coefficient of determination (R^2) values in the validation datasets ranged from 0.40 to 0.83 for both the BRDF-corrected and SS_{Lab} data, while the uncorrected spectra showed poor estimation accuracy (R^2 : 0.25–0.56). The resulting root-mean-squared error (RMSE) was reduced by 10% and 47% for the BRDF-corrected soil spectra compared to their uncorrected data. The BRDF-corrected and unmixed soil spectra were used to map soil properties at ~5-m spatial resolution for the entire catchment. Low SOC contents in the resulting maps adjoining the Ganges river flowing through our study site captured the topsoil loss typically observed from river banks. Thus, the BRDF-corrected HSR data not only improved the accuracy of soil estimates but also showed potential to identify vulnerable

areas needing precision management measures with high spatial resolution.

Index Terms—Memory-based learning (MBL), nonlinear polynomial unmixing, precision agriculture, soil erosion.

I. INTRODUCTION

HYPERSPECTRAL remote sensing (HSR) is an emerging alternative for testing soils and soil quality in large agricultural catchments [1], [2], [3]. Inspection of some of the key research publications on soil assessment via HSR approach (Table S1, supplementary materials) shows the wide range of conditions, under which HSR approach has been evaluated starting from small sampling densities [1] and small area of coverage [4] to as many as 325 soil samples collected over 420-km² study area [5], resulting in the coefficient of determination (R^2) values ranging from <0.20 to as large as 0.90. Several studies have data acquisition coinciding with bare soil (BS) conditions, while others have unmixing done to separate soil spectra from other prevailing endmembers such that the target soil property may be mapped for the whole study area [1]. Table S1 also shows that only a few studies implemented key preprocessing algorithms to enhance the image quality before any of the major data analytics, such as unmixing and chemometric modeling.

Restoration and denoising of hyperspectral imagery are increasingly emphasized in HSR data analysis [6]. Specifically, denoising and removal of stripping errors are needed to enhance HSR image quality. Sahadevan et al. [7] showed that the classification accuracy of an HSR image significantly improved when an edge-preserving bilateral filter was applied before classifying the image using a support vector machine (SVM). The filter preserved image details while smoothing each pixel, even in areas that were relatively homogenous. Similarly, the spectral reflectance of a pixel is also influenced by the bidirectional reflectance distribution function (BRDF), which arises from the variations in solar-surface-sensor geometry [8] in addition to the natural spectral variability of its endmembers. The BRDF effects may cause cross-track illumination gradients or reflectance inconsistencies, leading to the distortion of the spectral signatures of surface materials. These distortions significantly impact spectral analysis, reducing the accuracy of classification or regression models [9]. This challenge is particularly pronounced in agricultural landscapes, which are often partially or fully covered with photosynthetic and nonphotosynthetic vegetation. It is crucial to address

Received 11 January 2025; revised 10 March 2025; accepted 30 April 2025. Date of publication 12 May 2025; date of current version 27 May 2025. (Corresponding author: Bhabani Sankar Das.)

Naveen Kumar Purushothaman, Mayank Raj, and Bhabani Sankar Das are with the Department of Agricultural and Food Engineering, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India (e-mail: bsdas@agfe.iitkgp.ac.in).

Amanjit Premsagar is with the Department of Geological Sciences, University of Alabama, Tuscaloosa, Alabama 35401 USA.

Israr Majeed and N. Nagarjuna Reddy are with the International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Telangana 502324, India.

Lokesh Kumar Sinha, retired, was with the Defence Terrain Research Laboratory (DTRL), Defence Research and Development Organization, New Delhi 110011, India. He resides in Noida 201304, India.

This article has supplementary downloadable material available at <https://doi.org/10.1109/TGRS.2025.3569059>, provided by the authors.

Digital Object Identifier 10.1109/TGRS.2025.3569059

BRDF effects to improve the accuracy of HSR-based soil property assessment.

The BRDF effects are generally considered either by training a classification or regression model on each flight line (a single scene) and subsequently mosaicking resulting scenes into a single data product [1], [10] or by removing the BRDF effects from mosaicked scenes through preprocessing steps and then subjecting the whole hypercube for classification or regression analysis. Although the first approach can produce reasonable outcomes, a significant volume of training data is required, making large-scale applications challenging [11]. In contrast, the BRDF correction approach is more efficient and has gained widespread adoption due to its computational efficiency and ability to normalize reflectance across flight lines. The BRDF correction algorithms are generally either physical, empirical, or semi-empirical in nature. Both physical and empirical algorithms [12], [13], [14], [15] are computationally complex [16]. However, semi-empirical correction algorithms achieve a balance between computational complexity and operational feasibility [17]; and widely used semi-empirical BRDF correction algorithms, such as BREFCOR [18], aNBAR [19], RT-BRDF [20], and flexible BRDF (FlexBRDF) [21], use kernel-driven models (e.g., Ross-Thick and Li-Sparse) to normalize reflectance to nadir or bi-hemispherical geometries. These algorithms are shown to significantly improve the quality of HSR data across diverse land cover types [20], [22], although the model coefficients need to be estimated for specific land cover conditions [23].

Another challenge in using HSR for estimating soil properties is the presence of mixed pixels, which frequently occur in agricultural landscapes. After crop harvesting, pixels often contain a mix of BS, nonphotosynthetic vegetation (NPV) (e.g., crop residues), and green photosynthetic vegetation (GPV) (e.g., weeds) [1], [24], [25]. The spectral signatures of such mixed pixels complicate soil property estimation. Traditional methods often rely on thresholding to isolate BS spectra, which may lead to losing valuable data. Advances in spectral mixture analysis, including algorithms for estimating material abundances and their spectra [1], [26], [27], have partially addressed this issue. Few studies have attempted to apply either BRDF correction before spectral analysis or unmixing mixed pixels (Table S1). However, studies have rarely explored these two effects together to improve soil property estimation, especially in scenarios with multiple flight lines. The absence of readily available HSR data and the necessity for intricate preprocessing workflows pose additional barriers to the application of HSR in agricultural landscapes. Addressing these challenges may improve the utility of HSR for soil property mapping, especially for smallholder farms where accurate soil information is crucial for sustainable management. We hypothesize that integrating BRDF correction with spectral unmixing may enhance the utility of HSR data for mapping soil properties in mixed-pixel and multiflight-line scenarios. Thus, the main objective of this study is to evaluate how BRDF correction influences the estimation accuracy of soil properties in agricultural landscapes.

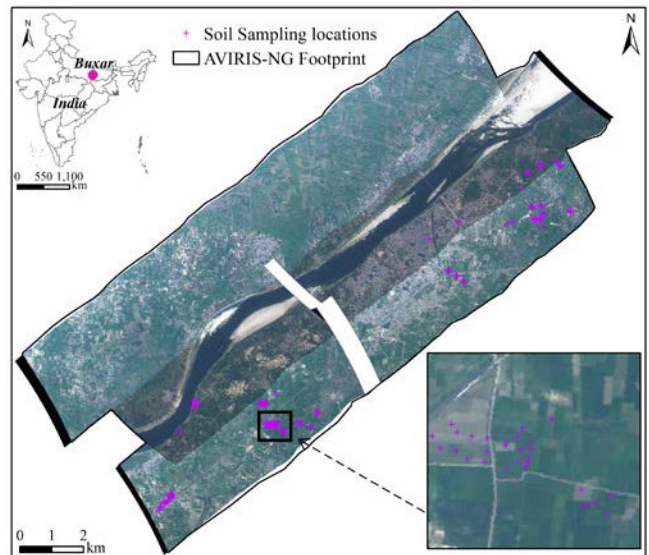


Fig. 1. Locations of soil samples collected from the Buxar region overlaid on the true color composite (B: 472 nm, G: 552 nm, and R: 642 nm) of the AVIRIS-NG BRDF-uncorrected image.

II. MATERIALS AND METHODS

A. Study Area and Soil Sampling

The study area (25.49° – 25.64° N and 83.86° – 84.03° E) covers two parallel strips of land (area: 138 km^2) adjoining the Ganges river flowing through the Buxar district in the state of Bihar, India (Fig. 1). Geomorphologically, the Buxar region comprises low lying northern plains and flat southern regions comprising diverse soil types ranging from recent alluvium to old alluvium [28]. Being a part of the southern Gangetic Plain, this site has a tropical climate with distinct seasonal variations with summer (from May to June) temperatures rising to as high 45°C while cooler winter (from January to February) temperatures dropping down to as low as 4°C . Monsoon rains prevail from June to September contributing to the district's annual precipitation of approximately 1021 mm with nearly 85% attributed to the southwest monsoon [28]. Agriculture forms the vital aspect of the region's landscape with crops such as paddy (*Oryza sativa*), maize (*Zea mays*), wheat (*Triticum aestivum*), lentils, and vegetables cultivated across three distinct seasons (kharif, rabi, and zaid).

Surface soil samples (depth: 0–15 cm) were collected from 173 ground reference locations (GRLs) during December 21–27, 2018. These sampling locations were primarily located in agricultural fields, which may also be visible from the land use and land cover (LULC) classification map of the study site (Fig. S1; supplementary materials). Each sample was air-dried, ground, and passed through a 2-mm sieve to obtain soil fractions for determining soil texture, soil organic carbon (SOC) content, pH, and electrical conductivity (EC). Soil textural fractions (clay and sand contents) were determined using the pipette method [29]. The chromic acid digestion method was followed to measure SOC content [30]. Soil pH and EC were measured in soil:water slurries (1:2 for pH and 1:2.5 for EC).

B. Collection of Airborne and Laboratory Spectra

During Phase 1 of the airborne visible-infrared imaging spectrometer-next generation (AVIRIS-NG) campaign, HSR data were collected over the Buxar region (site 150) on February 23, 2016, with the objective to map water constituents in the Ganges river. Three flight sorties were conducted using the ISRO B200 aircraft over this site between 11:39 AM and 12:00 PM IST to cover both the waterbody and the adjoining catchments on both sides of the Ganges river passing through this site. The AVIRIS-NG images contain spectral reflectance data at ~ 5 -nm spectral resolution over the 380–2510-nm (total 425 bands) wavelength region [31], [32]. The signal-to-noise ratio (SNR) ranged from more than 2000 at 600 nm to more than 100 at 2200 nm (95% accuracy) with 34° field of view (FOV) and 1-mrad instantaneous FOV. The flight altitude for B200 at this site ranged from 4 to 8 km, resulting in the ground sampling distance (GSD) varying from 4 to 8 m and swaths ranging from 3 to 6 km [33]. Unfortunately, no ground truth data were collected for soils during this campaign. Anticipating a similar campaign during Phase 2, we collected soil samples from this site during the scheduled campaign (December 21–27, 2018), although the campaign was not conducted during this period. With no AVIRIS-NG data available for this site during Phase 2 and with the premise that basic soil properties may change minimally over a three-year period [34], we used L2 HSR data collected during Phase 1 for estimating basic soil properties. Thus, we downloaded L2 data (atmospherically and geometrically corrected spectral reflectance data) collected during Phase 1 from the NASA site (<https://avirisng.jpl.nasa.gov/dataportal>) for our study site and considered it for estimating soil properties following BRDF correction. Before applying BRDF correction, spectral data for each pixel were extracted, and both the noisy bands (376.44–396.47 nm and 2404.95–2500.12 nm) and water vapor absorption bands (1348.12–1498.38 nm and 1803.91–1999.25 nm) were removed from each spectrum. This resulted in AVIRIS-NG spectra over 321 bands for each pixel, which we refer to as the raw AVS spectra.

In addition to raw AVS spectra, soil spectra were also collected in a dark room (Fig. S2; supplementary materials) in proximal sensing mode using a turntable fit with a halogen bulb and a spectroradiometer (Model: FieldSpec¹4 Hi-Res NG; Malvern Panalytical Ltd., USA). This spectroradiometer collects spectra over the 350–2500-nm spectral range with a spectral resolution of 3 nm over the visible- and near-infrared region and 6 nm over the shortwave infrared region. Soil samples were uniformly packed into a glass petri dish, and a thin glass plate was used to level the resulting soil surface to avoid soil compaction while packing [35], [36]. Each petri dish was then placed on the rotating turntable to collect spatially averaged reflectance spectra for each soil sample. A Spectralon¹ white reference panel (Labsphere, USA) was used to calibrate the spectroradiometer. Before the collection of soil spectra, the radiometer was warmed for a period of 1 h. The light source was optimized, and the white reference panel data were collected before starting the measurement and after

collecting soil spectra for every 30 samples. The 30 scans of spectral data were averaged for each sample to obtain a representative soil spectrum, which is referred to as laboratory soil spectra (SS_{Lab}).

C. BRDF Correction of AVIRIS-NG Data

For the BRDF correction, we used the FlexBRDF algorithm [21], which may be automated for correcting HSR data collected using AVIRIS-Classic, AVIRIS-NG, and National Ecological Observatory Network (NEON) over diverse land cover conditions. The FlexBRDF algorithm is integrated as a module within the Python (ver. 3.12.7) library *HyTools* (ver. 1.4.2) for processing hyperspectral image cubes [37]. The *HyTools* library consists of three submodules for correcting surface reflectance of HSR data, such as topographic illumination, BRDF, and sunlight glint corrections. Being part of the Indo-Gangetic Plain, the study site does not have much elevational variation (Fig. 1). Hence, we did not apply any topographic illumination correction, which is generally applied for HSR data collected from mountainous regions [38]. With sunlight glint correction applied to HSR data from waterbodies, we also did not use this for our image hypercubes. The FlexBRDF algorithm within the *HyTools* library uses a semi-empirical BRDF [21]

$$R_{\text{BRDF}}(\theta_v, \theta_s, \phi) = f_{\text{iso}} + f_{\text{geo}}K_{\text{geo}}(\theta_v, \theta_s, \phi) + f_{\text{vol}}K_{\text{vol}}(\theta_v, \theta_s, \phi) \quad (1)$$

where R_{BRDF} is the BRDF-corrected reflectance value, θ_v is the sensor zenith angle, θ_s is the solar zenith angle, and ϕ is the relative azimuth angle between the sun and the sensor. The weight fractions for isotropic, geometric, and volumetric components in (1) are represented by f_{iso} , f_{geo} , and f_{vol} , respectively. As proposed by Queally et al. [21], the Li-Sparse kernel was used for the geometric kernel function K_{geo} [39], while the Ross-Thick kernel was used for the volumetric kernel function K_{vol} [40]. Kernel values were estimated using ancillary datasets for θ_v , θ_s , and ϕ , which provide geometric measurements at the same pixel resolution as their reflectance products. Because both solar and viewing geometries can change depending on the type of vegetation structure [9], we applied individual BRDF models to respective land cover conditions, which required grouping of consistent vegetation structures in specific areas [41]. The FlexBRDF module stratifies vegetation into different groups with similar BRDF assumptions based on the normalized difference vegetation index (NDVI) and independently optimizes BRDF coefficients for each NDVI bin. The module also allows a user to set the number of bins, but the bin boundaries are dynamically set, and smoothing is applied at intermediate values of NDVI [41]. We tested 3, 9, 12, 15, and 18 NDVI bins to estimate BRDF coefficients. For our data, 18 NDVI bins offered the best classifications of the vegetation structure. As recommended by Queally et al. [21], all three flight scenes were grouped and mosaicked using the seamless mosaic feature of the mosaicking toolbox in the ENVI software (Exelis Visual Information Solutions, Boulder, Colorado) before doing the BRDF correction.

¹Registered trademark.

D. Ground Cover Condition and Unmixing of AVIRIS-NG Data

Our sampling locations at the selected GRLs contained varying levels of BS, GPV, and NPV fractions. Because our HSR data acquisition and soil sampling dates did not coincide and we did not directly quantify these three fractions, we relied on the AVIRIS-NG data for evaluating BS fractions such that soil spectra for each GRL may be extracted through unmixing. To estimate different soil cover fractions, we first estimated NDVI and cellulose absorption index (CAI) using AVIRIS-NG data. Fig. S3 (supplementary materials) shows estimated CAI and NDVI values for the three mosaicked AVIRIS-NG scenes of the study area. This figure clearly shows that several patches in our study site had NDVI values > 0.4 with wide variations in both NDVI and CAI values. Estimated CAI and NDVI values for the GRLs ranged from -3.2 to 1.5 and from 0.11 to 0.83 , respectively; 97 out of 173 GRLs had NDVI values > 0.4 and 38 GRLs had CAI values > 0 . These results suggest that almost half of the GRL containing pixels had either standing crops and/or a mix of green and dry vegetation when HSR data were collected. Thus, a three-endmember system (BS, NPV, and GPV) may be needed for the unmixing of AVS spectra before soil properties may be estimated.

The linear polynomial mixing (LPM) approach [1], [42] was used to unmix HSR data. Specifically, the linear extended algorithm [43] was utilized to unmix the AVS spectra (x_i) at the i th pixel. With three endmembers, x_i may be written as the sum of three linear terms (s_1 , s_2 , and s_3 representing spectra for BS, GPV, and NPV, respectively) and their four polynomial mixture terms [1], [42]

$$x_i = a_1(i)s_1 + a_2(i)s_2 + a_3(i)s_3 + a_{1,2}(i)s_1s_2 + a_{1,3}(i)s_1s_3 + a_{2,3}(i)s_2s_3 + a_{1,2,3}(i)s_1s_2s_3 \quad (2)$$

where i ranges from 1 to N pixels in the HSR image. Equation (2) with the first three terms constitutes the three-endmember linear mixture (LM) model. With the additional four terms, the spectral signature of each AVIRIS-NG pixel is considered to be a linear polynomial mixture of the spectral signatures of individual endmembers. In the matrix form, (2) may be written to implement the nonnegative matrix factorization (NMF) methods

$$X = AS = A_a S_a + A_b S_b + A_c S_c. \quad (3)$$

The matrix X in (3) represents the collection of AVS spectra (x_i) for different pixels, A represents the mixing coefficient matrix containing respective abundance values (a_1 , a_2 , and a_3), and S represents the source matrix of endmember spectra

$$A = [A_a \quad A_b \quad A_c]$$

$$S = \begin{bmatrix} S_a \\ S_b \\ S_c \end{bmatrix}.$$

The coefficient matrix containing A_a , A_b , and A_c is defined as

$$A_a = \begin{bmatrix} a_1(1) & a_2(1) & a_3(1) \\ \vdots & \vdots & \vdots \\ a_1(N) & a_2(N) & a_3(N) \end{bmatrix}$$

$$A_b = \begin{bmatrix} a_{1,2}(1) & a_{1,3}(1) & a_{2,3}(1) \\ \vdots & \vdots & \vdots \\ a_{1,2}(N) & a_{1,3}(N) & a_{2,3}(N) \end{bmatrix}$$

$$A_c = \begin{bmatrix} a_{1,2,3}(1) \\ \vdots \\ a_{1,2,3}(N) \end{bmatrix}.$$

Similarly, the source matrix containing S_a , S_b , and S_c may be defined as

$$S_a = [s_1 \quad s_2 \quad s_3]^T$$

$$S_b = [s_1 \odot s_2 \quad s_1 \odot s_3 \quad s_2 \odot s_3]^T$$

$$S_c = [s_1 \odot s_2 \odot s_3]^T$$

where $\odot$ represents the element-wise multiplication. Equation (3) was solved by minimizing the Frobenius norm [43]

$$J = \frac{1}{2} \|X - AS\|_F^2 \quad (4)$$

along with the condition that the abundance fractions are nonnegative and sum to one for each pixel. More details about the LPM algorithm and the pseudocode may be found in Majeed et al. [1] and Majeed and Das [3].

Although the LPM algorithm was originally created for the blind source separation of potential endmembers from the HSR data [42], we supervised the algorithm by providing reference spectra for the three endmembers of BS, GPV, and NPV and along with their corresponding initial abundance values, instead of random initialization of the algorithm. For the BS, SS_{Lab} values for soil samples collected from all the GRLs were averaged to obtain the reference BS spectra. Because we did not directly measure GPV or NPV spectra, the NPV reference spectra from Majeed et al. [1] were used in our study; we assumed that the paddy stubbles at both our sites would have similar spectral characteristics. For the reference GPV spectra, we first selected locations with dense vegetation based on the true color composite of the AVIRIS-NG image [1]. The AVS spectra of these pixels were then averaged to obtain the reference GPV spectra. Initial abundance values of each endmember were obtained from the linear unmixing algorithm [26] using the hyperspectral image processing toolbox in MATLAB [44]. Equation (3) was then solved to extract soil spectra from BRDF-corrected and uncorrected AVS spectra through the LPM algorithm. Hereinafter, unmixed soil spectra derived from BRDF-corrected and uncorrected AVS spectra are referred to as SSB_c and SSB_{uc} , respectively.

E. Soil Spectral Preprocessing and Chemometric Modeling

We estimated five soil parameters using three spectral data sources. Soil spectra measured in laboratory condition (SS_{Lab}) served as the spectral data source having the highest SNR and spectral resolution. Unmixing of BRDF-corrected and uncorrected AVIRIS-NG data provided the other two sources of soil spectra: SSB_c and SSB_{uc} . Before chemometric modeling, several preprocessing techniques were applied to each

spectrum for capturing significant spectral features. First, spectral reflectance values were converted to spectral absorbance values. Resulting absorption spectra were then smoothed using the second-order Savitzky–Golay smoothing method [45]; the span length was set to 11 nm to improve SNR for each absorption spectrum [46]. Similarly, the effects of light scattering were minimized by transforming absorbance values into standard normal variates for each spectrum [47]. These preprocessing steps were implemented using the *prospectr* package (ver. 0.2.6) [48] from RStudio software (ver. 4.3.1) [49].

For the chemometric modeling, soil spectra and corresponding soil property datasets were first divided into 75% calibration and 25% validation datasets using the *caret* package (ver. 6.0–94) [50] in RStudio. Four commonly used chemometric models were then evaluated for their ability to estimate each soil property. These include the partial-least-squares regression (PLSR), cubist [51], support vector regression [52], and memory-based learning (MBL) algorithm [53]. We chose to present only the results of the MBL algorithm because of its superior performance compared to other considered models in what follows. A flowchart showing the modeling steps is shown in Fig. S4 (supplementary materials).

The MBL algorithm is a spectrum-based learner approach and is associated with case-based reasoning processes [53]. Generally, a local model is developed for each soil separately in the MBL approach instead of constructing a general or global model. In this approach, a set of k -nearest neighbors to each soil is first selected from the calibration dataset (i.e., reference library) based on the Mahalanobis distances in the principal component (PC) space. A local model is then fit using the selected dataset using a specific chemometric model. For our MBL approach, we used the weighted average PLSR model [54]. Briefly, multiple models are first built by considering a minimum of three PCs to a maximum of 20 PCs. The predicted value for a given soil parameter for the selected soil is then obtained as the weighted average of all the predicted values from multiple models generated for different sets of PCs [54], [55]. The sequence of k -nearest neighbors ranged from 5 to the total number of observations in the calibration dataset in steps of 5. The MBL algorithm generally outperformed all other competitive models, with three out of five soil properties showing close alignment with the observed data, as may also be seen in the Taylor diagram of Fig. 2.

The accuracy of the chemometric models was evaluated using both the root-mean-squared error (RMSE) and R^2 values

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (6)$$

where Y_i and $\hat{Y}_i$ are the measured and predicted response variables, respectively, at the i th location; $\bar{Y}$ is the mean value of Y_i ; and N is the number of GRLs. We also estimated the bias values as the difference between the mean value of predicted and measured soil properties. The MBL modeling

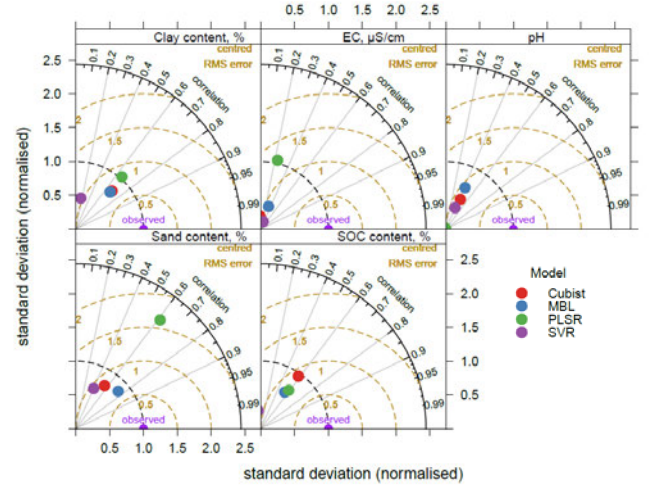


Fig. 2. Taylor diagram for selected soil properties estimated by different chemometric models using the AVIRIS-NG BRDF-uncorrected spectra.

work was carried out in RStudio version 4.3.1 software [49] using the *resemble* package [55].

III. RESULTS AND DISCUSSION

A. Characteristics of the Soil Properties and Their Correlation Structure

Table I shows the summary statistics for the five soil parameters measured at different GRLs. Soil samples collected at different sampling locations showed medium to coarse-textured soils (Fig. S5; supplementary materials) with the average clay and sand contents of 20.5% and 50.4%, respectively. About 43% of collected samples had loamy texture, followed by sandy loam (27% of samples) and clay loam (14% of samples) textures. Such textural variations in the study area may be because of the deposition of alluvium from the Ganges river [28]. Soil samples also had medium to high SOC contents with the first and third quartile values of 0.49% and 0.78%, respectively. Soil pH varied between 6.63 and 7.51, with the average value of 7.51 (Table I). The recent alluvial soils from the tributary of the Ganges river of the Buxar region generally show medium to high soil fertility with neutral soil pH [28]. Generally, the study area remains nonsaline with EC values $<2000 \mu\text{S cm}^{-1}$ with a range of $61\text{--}743 \mu\text{S cm}^{-1}$. Overall, the soils of the study area show a wide range of coefficient of variation (CV), with the lowest CV of 4% for soil pH to the highest CV of 40% for EC.

The Pearson correlation coefficient (r) between different soil properties (Table S2) showed a strong positive correlation between SOC and clay contents ($r = 0.50$), which is typically observed for the Indian soils [56]. All three soil chromophores (spectrally active soil parameters), such as sand, clay, and SOC contents, were highly correlated with each other. Among soil nonchromophores (spectrally inactive soil parameters), soil pH showed a moderate and strong negative correlation with SOC ($r = -0.34$) and clay contents ($r = -0.53$), respectively. This shows that the soil pH may be efficiently estimated using HSR data. However, there was only a moderate correlation between soil chromophores and EC, which would make its estimation difficult using HSR data.

TABLE I
DESCRIPTIVE STATISTICS FOR MEASURED BASIC SOIL PROPERTIES OF BUXAR

Soil Properties	Min	Max	Q1	Q3	Mean	CV, %
Sand content, %	23.3	82.1	41.0	59.2	50.4	26
Clay content, %	7.47	40.4	14.6	26.2	20.5	38
SOC content, %	0.12	2.13	0.49	0.78	0.66	37
pH	5.38	8.17	6.63	7.51	7.04	9
EC, $\mu\text{S cm}^{-1}$	61	743	139	248	206	46

CV: Coefficient of variation; Min: Minimum; Max: Maximum; SD: Standard Deviation; SOC: soil organic carbon; EC: electrical conductivity

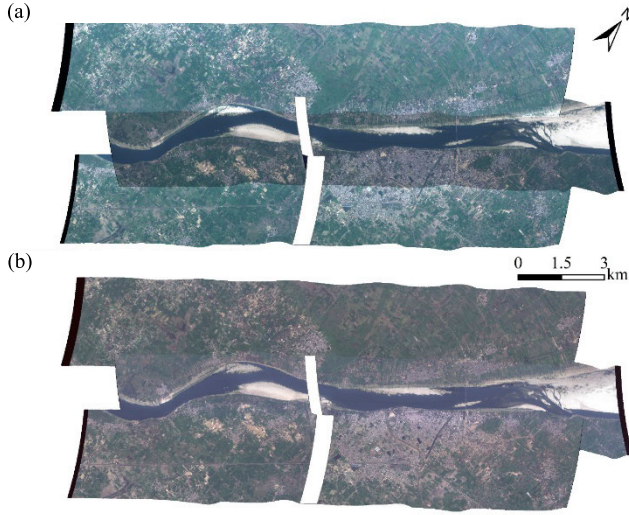


Fig. 3. True color composite images (B: 472 nm, G: 552 nm, and R: 642 nm) over the Buxar region for BRDF (a) uncorrected and (b) corrected AVIRIS-NG image.

B. Performance of BRDF Correction Algorithm

Fig. 3 shows the true color composited AVIRIS-NG images before and after BRDF correction. The BRDF-uncorrected and mosaicked images [Fig. 3(a)] show the effects of BRDF in the form of cross-track brightness gradients along the overlapping regions between adjacent flight lines. The BRDF-corrected image [Fig. 3(b)] minimizes these effects and shows minimum discontinuity between two mosaicked hypercubes. Fig. 4 shows the mean reflectance values as a function of view zenith angles (VZAs) ranging from 0° (nadir view) to $+20^\circ$ estimated at specific wavelengths of different soil constituents: 420, 480, and 660 nm for goethite [57], [58], [59]; 682 nm for hematite [58], [60]; 2169 nm for organic matter [61]; and 2200 nm for clay minerals [62], [63]. The differences in mean reflectance values estimated from the uncorrected and BRDF-corrected data were minimal up to about 6° VZA but increased significantly at higher zenith angles (Fig. 4). Specifically, a significant BRDF effect may be seen at 420 nm (goethite), where the relative change in reflectance between the nadir view and maximum VZA reached 75%. Other selected wavelengths show variations ranging from 33% to 51%. After applying the BRDF correction algorithm, these differences were reduced to 59% at 420 nm and to less than 33% for other wavelengths. Colgan et al. [64] also reported stronger BRDF effects over the visible region than the NIR region.

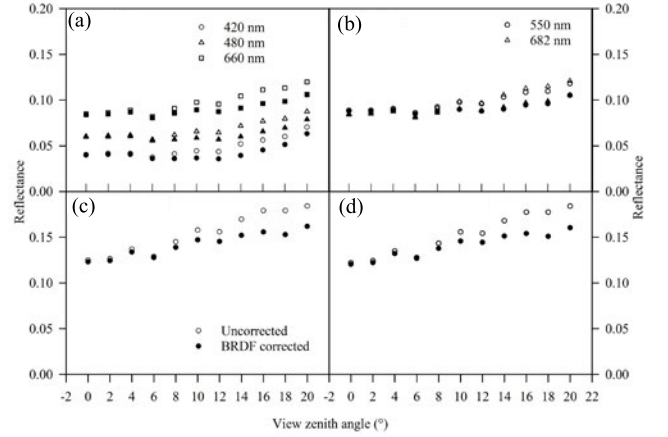


Fig. 4. Average spectral reflectance versus VZA for BRDF-corrected (filled symbols) and uncorrected (unfilled symbols) AVIRIS-NG data at wavelengths specific to goethite (420, 480, and 660 nm), hematite (550 and 682 nm), organic matter (OM: 2169 nm), and clay minerals (2200 nm). (a) Goethite. (b) Hematite. (c) OM (2169 nm). (d) Clay minerals (2200 nm).

C. Unmixing of AVS Spectra

Unmixing of AVS spectra extracted for the 173 GRLs yielded soil spectra at 140 GRLs when the BRDF-corrected hypercube was considered for unmixing. However, only 114 GRLs showed nonzero BS fractions, yielding 114 unmixed soil spectra. Interestingly, five of these GRLs did not show any BS component when the BRDF-corrected AVIRIS-NG image was used for unmixing. All of these five GRLs showed BS fractions ranging from 1.7 to 3.7%; however, resulting unmixed soil spectra had vegetation-linked spectral features. When BRDF-corrected mixed spectra were unmixed, all these five locations showed no BS fractions. Therefore, we considered unmixed soil spectra for 109 common GRLs for analyzing how BRDF correction influences the estimation accuracy of soil properties in the HSR approach.

To assess the accuracy of the unmixing approach, we estimated spectral angle mapper (SAM) values between the unmixed soil spectra and the SS_{Lab} -derived reference soil spectra. The estimated mean (μ) and standard deviation (σ) of the SAM values (Table II) suggest that the unmixed soil spectra extracted from the BRDF-corrected image have lower μ (0.121) and σ (0.188) values than those extracted from the uncorrected image ($\mu = 0.157$ and $\sigma = 0.218$). Thus, the unmixed soil spectra extracted from the BRDF-corrected AVIRIS-NG image were more similar to the

TABLE II

MEAN (μ) AND STANDARD DEVIATION OF SAM VALUES FOR SOIL SPECTRA EXTRACTED USING UNMIXING METHOD FROM BOTH UNCORRECTED AND BRDF-CORRECTED AVIRIS-NG IMAGES

Parameters	Uncorrected	BRDF corrected
μ	0.157	0.121
σ	0.215	0.188

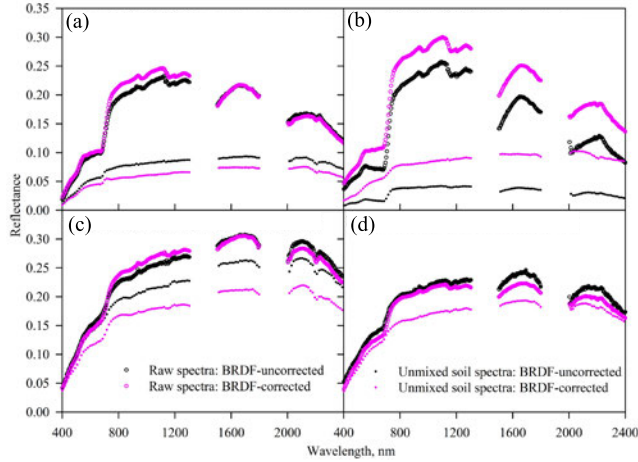


Fig. 5. Raw AVIRIS-NG spectra (circles) and soil spectra extracted using the nonlinear unmixing method (dots) for selected (a) and (b) locations with high and (c) and (d) low GPV fractions from uncorrected and BRDF-corrected images.

reference soil spectra compared to their uncorrected counterparts. This may be because the shading and other BRDF effects influenced both the shape and magnitude of resulting soil spectra upon unmixing when correction measures were not implemented.

Fig. 5 shows AVS spectra before BRDF correction (black circles) and after correction (pink circles), and their corresponding unmixed soil spectra: SSB_{uc} (black dots) and SSB_c (pink dots) for four selected GRLs; corresponding endmember abundance fractions for both the BRDF-corrected and uncorrected spectra are shown in Table S3. For the nonoverlapping regions [Fig. 5(a) and (c)], the AVS spectra before and after BRDF correction were relatively more similar to each other than their counterparts in the overlapping regions [Fig. 5(b) and (d)]. The presence of a high GPV fraction accentuated this difference. Specifically, the shape of the BRDF-corrected AVS spectra was identical to its uncorrected AVS spectra over the visible region [Fig. 5(a)] for the nonoverlapping region, while the BRDF correction resulted in higher reflectance values than its uncorrected AVS spectra for the entire visible to near-infrared (VNIR) region (400–2400 nm) in the overlapping case [Fig. 5(b)]. Unmixed soil spectra showed contrasting characteristics. For the scenarios where BRDF effects are less pronounced (e.g., the nonoverlapping regions in the mosaicked hypercubes), the unmixed soil spectra showed greater reflectivity when the AVS spectra were corrected for the BRDF effects (i.e., SSB_c represented with pink dots in Fig. 5) than the uncorrected ones (i.e., SSB_{uc} represented with black dots in Fig. 5) for both high [Fig. 5(a)] and low [Fig. 5(c)]

vegetation areas. With increased reflectivity in the AVS spectra upon BRDF correction in the overlapped region, SSB_c also showed increased reflectivity over the entire VNIR wavelength region [Fig. 5(b)]. Moreover, the characteristic red edge effect over the visible part is also no longer visible in SSB_c [pink dots in Fig. 5(b)]. Increase in the overall reflectance values after BRDF correction at high GPV location indicates the removal of the shading effects known for high vegetation areas [64]. Inspection of resulting unmixed soil spectra suggests that the magnitude of soil reflectance decreased with the increase in soil pH, as expected [65]. The unmixed soil spectra at location 3 [Fig. 5(c)] with a soil pH of 5.58 was most reflective, while the spectra from location 1 [Fig. 5(a)] showed less reflection values with the soil pH of 7.95. Soil textural fractions were similar across all these sites, with clay contents ranging from 12.1% [location 4, Fig. 5(d)] to 22.3% [location 3, Fig. 5(c)]. The SOC contents for the high GPV locations [Fig. 5(a) and (b)] were similar (0.3–0.4%) and were higher than those at the low GPV locations (SOC contents: 0.6–0.66%). Even with high SOC contents, the sites having low GPV fractions showed high reflectance values because of high Fe and Al oxides and hydroxides typically found in low pH soils [66]. The unmixed soil spectra for both the GRLs showed characteristic metal–OH stretching observed for Si–OH at 2200 nm [62], [67] as expected.

Fig. S6 (supplementary materials) shows the fractional abundance maps for BS, NPV, and GPV from the BRDF-corrected image along with a ternary plot of the extracted fractional abundance values for all GRLs. Abundance maps show a typical inverse relationship between BS and vegetation cover. Specifically, greenish patches (high BS fraction) in the BS fraction map, white patches (no green vegetation) in the GPV abundance map, and light greenish brown patches (moderate dry crop residue) in the NPV abundance map indicate the presence of a mix of BS and rice stubble, as seen in rice-growing areas. The ternary plot shows that the BS fractional abundance for GRLs ranges from 0 to 0.80. Overall, the unmixing approach helped in extracting BS.

To summarize, the unmixed soil spectra derived from the BRDF-corrected mixed spectra showed greater similarity to the reference soil spectra, had limited vegetation-linked features, and were consistent with the typical behavior expected because of the presence or absence of specific soil chromophores. The BRDF correction also yielded unmixed soil spectra for more GRLs than their uncorrected counterparts. This suggests that there is a need to make BRDF correction for obtaining different endmember spectra through unmixing approaches, specifically when vegetation is present in the FOV of the sensor.

D. HSR-Based Soil Property Estimation and Mapping

Table III lists performance statistics when the MBL approach was used to estimate soil properties using three spectral data sources (SS_{Lab} , SSB_{uc} , and SSB_c). Soil pH, clay, and SOC contents showed a higher estimation accuracy in the SS_{Lab} data ($R^2 = 0.83$), followed by sand content ($R^2 = 0.63$). In contrast, the SSB_{uc} data showed low to moderate estimation

TABLE III

PERFORMANCE OF THE MBL APPROACH IN THE VALIDATION DATASETS FOR DIFFERENT SOIL PARAMETERS USING SOIL SPECTRA COLLECTED UNDER LABORATORY CONDITION (SS_{Lab}) AND SOIL SPECTRA DERIVED FROM UNMIXING OF AVIRIS-NG SPECTRA BEFORE BRDF CORRECTION (SSB_{uc}) AND AFTER BRDF CORRECTION (SSB_c)

Data sources	Metrics	Sand content, %	Clay content, %	SOC content, %	pH	EC, $\mu S\ cm^{-1}$
SS_{Lab}	R^2	0.63	0.83	0.83	0.83	0.40
	RMSE	8.10	3.19	0.10	0.28	61.6
	Bias	0.78	-0.35	0.01	-0.02	-14.0
SSB_{uc}	R^2	0.56	0.46	0.30	0.18	0.25
	RMSE	9.52	6.21	0.16	0.70	59.3
	Bias	-0.07	1.03	-0.01	-0.08	9.54
SSB_c	R^2	0.72	0.83	0.43	0.80	0.41
	RMSE	7.50	3.70	0.14	0.37	53.3
	Bias	0.07	-0.20	-0.01	-0.12	6.30
% decrease in RMSE		21	41	10	47	10

SOC: soil organic carbon; EC: electrical conductivity; R^2 : coefficient of determination; RMSE: root-mean-squared error; BRDF: bidirectional reflectance distribution function; AVIRIS-NG: airborne visible-infrared imaging spectrometer-next generation.

accuracy for these properties (R^2 range: 0.18–0.56), while the SSB_c data showed moderate to high estimation accuracy with R^2 values ranging from 0.43 to 0.83. Specifically, soil properties such as pH, clay, and sand contents showed a high estimation accuracy with R^2 values > 0.72 . Moreover, the SSB_c spectra showed 10%–47% decrease in the RMSE values as compared to SSB_{uc} spectra (Table III). The performance of the chemometric models for estimating the soil EC parameter was generally poor across all the spectral data sources. This may be because of its poor correlation with soil chromophores (Table S2). Moreover, the estimation of SOC content based on the AVS data sources showed lower accuracy than the SS_{Lab} data, with R^2 values decreasing from 0.83 (SS_{Lab}) to 0.30 (SSB_{uc}). This discrepancy may have resulted from the mismatch between the soil sampling and the AVIRIS-NG data acquisition dates.

Fig. 6 shows scatter plots of observed versus predicted values for soil parameters. The data points for soil pH and clay and sand contents from the SSB_c data closely align with the 1:1 line, similar to the SS_{Lab} data, indicating improvements in the estimation accuracy of these soil properties. A significant improvement is also seen for SOC content in the SSB_c spectra, as shown in Fig. 6(c). These chemometric modeling results suggest that BRDF correction is a crucial preprocessing step for the airborne-based or satellite-based HSR data, particularly when estimating soil properties across multiple flight line scenes.

Validated MBL models were then applied to estimate selected soil parameters across the entire study area using pixel-wise SSB_c spectra. Maps for clay and SOC contents over the 138-km² study area at a 5-m spatial resolution are shown in Fig. 7; corresponding maps of sand content and pH are presented in Fig. S7 (supplementary materials). As expected, typical inverse relationships between clay and sand contents are observed with pink patches in the sand content map coinciding with greenish patches in the clay content map. Notably, most agricultural fields near the banks of the Ganges river show low SOC contents ($<0.36\%$). This may be attributed to runoff from agricultural lands, possibly eroding the top fertile organic soil layer. Low (pinkish patches)

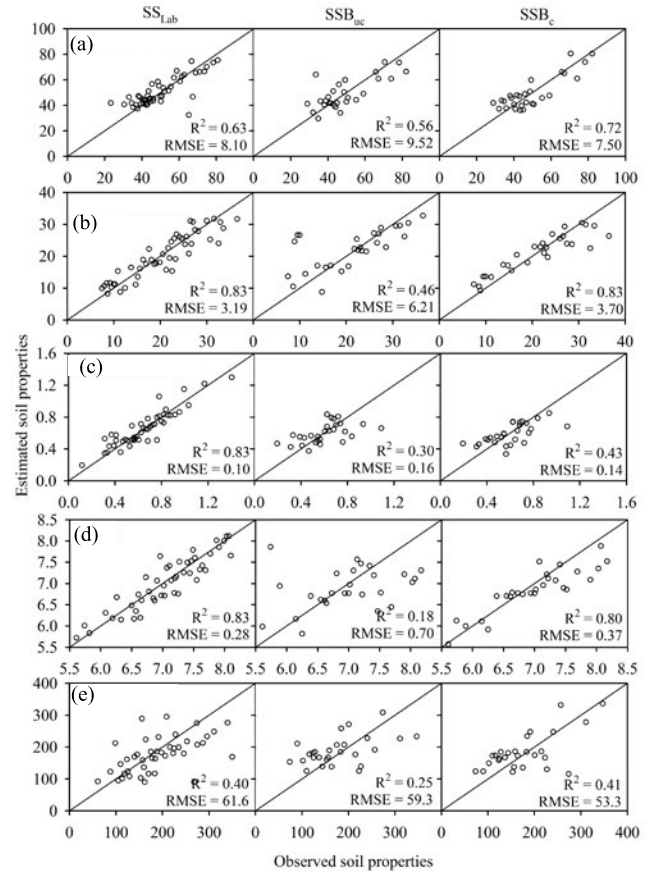


Fig. 6. Observed and estimated soil properties using different spectral data sources [laboratory soil spectra (SS_{Lab}), soil spectra derived from unmixing of AVIRIS-NG spectra before BRDF correction (SSB_{uc}) and after BRDF correction (SSB_c)]. (a) Sand content, (b) clay content, (c) SOC content, (d) pH, and (e) EC, $\mu S\ cm^{-1}$.

to medium (light yellowish patches) SOC content in these agricultural landscapes highlights the need for better nutrient management practices to sustain crop production. Thus, these HSR-derived soil property maps yield spatially continuous measures for mapped parameters across a large agricultural landscape with high spatial resolution (~ 5 m) and estimation accuracies relevant for agricultural management.

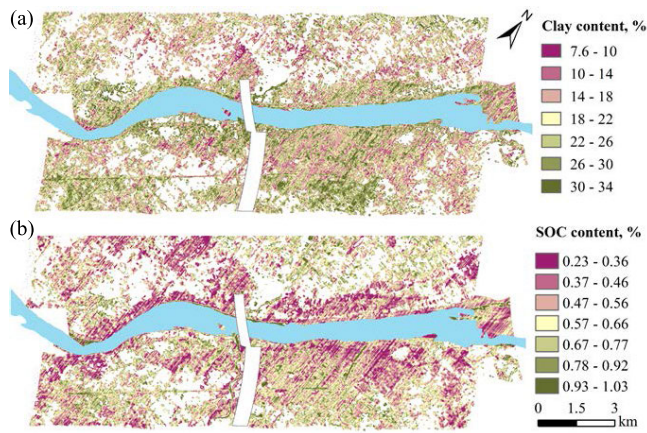


Fig. 7. Estimation maps of (a) clay and (b) SOC contents developed using the BRDF-corrected unmixing AVIRIS-NG image across the Buxar region.

IV. CONCLUSION

Although HSR data have the potential to produce soil maps for large agricultural catchments, raw HSR data often contain BRDF effects in the form of cross-track brightness gradients and shading effects, specifically when multiple images need to be mosaicked. To correct for such effects, we applied the recently developed FlexBRDF algorithm on the AVIRIS-NG data to improve the estimation accuracy of selected soil properties for a 138-km² agricultural catchment adjoining the Ganges river in Bihar. The BRDF correction algorithm produced HSR data with minimal cross-track brightness gradients and completely removed shading effects from the mosaicked images. Moreover, a greater number of GRLs showed BS fractions when BRDF-corrected spectra were used for unmixing. Because soil spectra could be unmixed for 140 GRLs using the BRDF-corrected images and 114 GRLs using the uncorrected images, we considered only 109 GRLs for which soil spectra were available from all three sources: lab-measured soil spectra and unmixed soil spectra from both BRDF-corrected and uncorrected mixed spectra for chemometric modeling. Results showed that the MBL model calibrated with BRDF-corrected data achieved significantly improved accuracy with R^2 values as high as 0.83 and RMSE reductions of 10%–47% relative to uncorrected data. High estimation accuracy was observed for clay and sand contents and soil pH with R^2 values exceeding 0.72. Spatial maps generated at a high spatial resolution (~5 m) revealed patches of vulnerable soil zones, such as those having low SOC contents, along both sides of the Ganges river. With elevational gradients, topsoil erosion frequently reduced SOC contents along river banks, which could be captured with high spatial resolution maps derived from HSR data. Therefore, a correct preprocessing step such as the BRDF approach not only improves the performance of both the unmixing and chemometric algorithms to accurately estimate soil properties but also indirectly assists in identifying vulnerable areas that may require immediate management interventions. High spatial resolution soil maps from HSR data yield soil parameter maps vital for precision agriculture and sustainable land management. While the results highlight the importance of BRDF correction, there is significant potential for further improvement by addressing other confounding

factors such as soil moisture. Future research can also focus on developing integrated correction frameworks and leveraging advancements in machine learning to enhance the scalability of hyperspectral data. With continued innovation, this approach has the potential to generate soil property maps across diverse and dynamic agricultural landscapes, which is much needed for implementing precision agricultural practices.

ACKNOWLEDGMENT

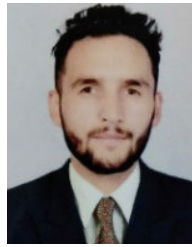
Bhabani Sankar Das thankfully acknowledges the erstwhile Defence Terrain Research Laboratory, Defence Research and Development Organization (DRDO), New Delhi, for facilitating soil collection at their study site. Naveen Kumar Purushothaman acknowledges the Prime Minister's Research Fellowship for his Ph.D. research.

REFERENCES

- [1] I. Majeed, N. K. Purushothaman, P. Chakraborty, N. Panigrahi, H. B. Vasava, and B. S. Das, "Estimation of soil and crop residue parameters using AVIRIS-NG hyperspectral data," *Int. J. Remote Sens.*, vol. 44, no. 6, pp. 2005–2038, Mar. 2023.
- [2] T. Paz-Kagan et al., "Mapping the spectral soil quality index (SSQI) using airborne imaging spectroscopy," *Remote Sens.*, vol. 7, no. 11, pp. 15748–15781, Nov. 2015.
- [3] I. Majeed and B. S. Das, "Large-scale mapping of soil quality index in different land uses using airborne hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5507812, doi: 10.1109/TGRS.2024.3360334.
- [4] H. Bartholomeus, G. Epema, and M. Schaepman, "Determining iron content in Mediterranean soils in partly vegetated areas, using spectral reflectance and imaging spectroscopy," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 9, no. 2, pp. 194–203, May 2007.
- [5] A. Stevens et al., "Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy," *Geoderma*, vol. 158, nos. 1–2, pp. 32–45, Aug. 2010.
- [6] P. Ghamisi et al., "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.
- [7] A. S. Sahadevan, A. Routray, B. S. Das, and S. Ahmad, "Hyperspectral image preprocessing with bilateral filter for improving the classification accuracy of support vector machines," *J. Appl. Remote Sens.*, vol. 10, no. 2, Apr. 2016, Art. no. 025004.
- [8] D. P. Roy et al., "A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance," *Remote Sens. Environ.*, vol. 176, pp. 255–271, Apr. 2016.
- [9] D. J. Jensen, M. Simard, K. C. Cavanaugh, and D. R. Thompson, "Imaging spectroscopy BRDF correction for mapping Louisiana's coastal ecosystems," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1739–1748, Mar. 2018.
- [10] M. Wietecha, L. Jełowicki, K. Mitelsztedt, S. Miścicki, and K. Stereńczak, "The capability of species-related forest stand characteristics determination with the use of hyperspectral data," *Remote Sens. Environ.*, vol. 231, Sep. 2019, Art. no. 111232.
- [11] W. Jia, Y. Pang, and R. Tortini, "The influence of BRDF effects and representativeness of training data on tree species classification using multi-flightline airborne hyperspectral imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 207, pp. 245–263, Jan. 2024.
- [12] C. Walthall, "A study of reflectance anisotropy and canopy structure using a simple empirical model," *Remote Sens. Environ.*, vol. 61, no. 1, pp. 118–128, Jul. 1997.
- [13] M. I. Mishchenko, J. M. Dlugach, E. G. Yanovitskij, and N. T. Zakharova, "Bidirectional reflectance of flat, optically thick particulate layers: An efficient radiative transfer solution and applications to snow and soil surfaces," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 63, nos. 2–6, pp. 409–432, Sep. 1999.
- [14] G. Roberts, "A review of the application of BRDF models to infer land cover parameters at regional and global scales," *Prog. Phys. Geography: Earth Environ.*, vol. 25, no. 4, pp. 483–511, Dec. 2001.
- [15] B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*. Cambridge, U.K.: Cambridge Univ. Press, 2012, doi: 10.1017/CBO9781139025683.

- [16] F. Kizel and Y. Vidro, "Bidirectional reflectance distribution function (BRDF) of mixed pixels," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 3, pp. 195–200, Jun. 2021.
- [17] M. Vögtli, D. Schläpfer, M. C. Schuman, M. E. Schaepman, M. Kneubühler, and A. Damm, "Effects of atmospheric, topographic, and BRDF correction on imaging spectroscopy-derived data products," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 109–126, 2024.
- [18] D. Schläpfer, R. Richter, and T. Feingersh, "Operational BRDF effects correction for wide-field-of-view optical scanners (BREFCOR)," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1855–1864, Apr. 2015.
- [19] J. Weyermann, A. Damm, M. Kneubühler, and M. E. Schaepman, "Correction of reflectance anisotropy effects of vegetation on airborne spectroscopy data and derived products," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 616–627, Jan. 2014.
- [20] W. Jia, Y. Pang, R. Tortini, D. Schläpfer, Z. Li, and J.-L. Roujean, "A kernel-driven BRDF approach to correct airborne hyperspectral imagery over forested areas with rugged topography," *Remote Sens.*, vol. 12, no. 3, p. 432, Jan. 2020.
- [21] N. Queally et al., "FlexBRDF: A flexible BRDF correction for grouped processing of airborne imaging spectroscopy flightlines," *J. Geophys. Res., Biogeosci.*, vol. 127, no. 1, Jan. 2022, Art. no. JG006622, doi: 10.1029/2021JG006622.
- [22] J.-L. Roujean, M. Leroy, and P.-Y. Deschamps, "A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data," *J. Geophys. Res., Atmos.*, vol. 97, no. D18, pp. 20455–20468, Dec. 1992.
- [23] Z. Wang and L. Liu, "Correcting bidirectional effect for multiple-flightline aerial images using a semiempirical kernel-based model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4450–4463, Sep. 2016.
- [24] D. A. Roberts, M. O. Smith, and J. B. Adams, "Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data," *Remote Sens. Environ.*, vol. 44, nos. 2–3, pp. 255–269, May 1993.
- [25] J. P. Guerschman, M. J. Hill, L. J. Renzullo, D. J. Barrett, A. S. Marks, and E. J. Botha, "Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 hyperion and MODIS sensors," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 928–945, May 2009.
- [26] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [27] F. Kizel and J. A. Benediktsson, "Hyperspectral and spatially adaptive unmixing for an analytical reconstruction of fraction surfaces from data with corrupted pixels," in *Handbook of Pattern Recognition and Computer Vision*, C. H. Chen, L. F. Pau, and P. S. Wang, Eds., Singapore: World Scientific, 2020, pp. 209–230, doi: 10.1142/9789811211072_0011.
- [28] Central Ground Water Board. (2013). *Ground Water Information Booklet Buxar District, Bihar State*. Accessed: Feb. 12, 2024. [Online]. Available: https://www.cgwb.gov.in/old_website/District_Profile/Bihar/Buxar.pdf
- [29] G. W. Gee and J. W. Bauder, "Particle size analysis," *Methods Soil Anal. 1 Phys. Mineralogical Methods*, vol. 5, pp. 383–411, Jan. 1986.
- [30] A. Walkley and I. A. Black, "An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method," *Soil Sci.*, vol. 37, no. 1, pp. 29–38, 1934.
- [31] L. Hamlin et al., "Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments," in *Proc. Aerosp. Conf., Big Sky, MT, USA*, Mar. 2011, pp. 1–7.
- [32] A. K. Thorpe et al., "Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG)," *Remote Sens. Environ.*, vol. 179, pp. 104–115, Jun. 2016.
- [33] B. K. Bhattacharya et al., "An overview of AVIRIS-NG airborne hyperspectral science campaign over India," *Current Sci.*, vol. 116, no. 7, p. 1082, Apr. 2019.
- [34] A. A. Berhe, "Drivers of soil change," in *Developments in Soil Science*, M. Busse, C. P. Giardina, D. M. Morris, and D. S. Page-Dumroese, Eds., Elsevier, 2019, pp. 27–42, doi: 10.1016/b978-0-444-63998-1.00003-3.
- [35] A. M. Mouazen, B. Kuang, J. De Baerdemaeker, and H. Ramon, "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*, vol. 158, nos. 1–2, pp. 23–31, Aug. 2010, doi: 10.1016/j.geoderma.2010.03.001.
- [36] N. M. Dhawale, V. I. Adamchuk, S. O. Prasher, R. A. V. Rossel, and A. A. Ismail, "Evaluation of two portable hyperspectral-sensor-based instruments to predict key soil properties in Canadian soils," *Sensors*, vol. 22, no. 7, p. 2556, Mar. 2022.
- [37] Chlus, Queally, and Townsend. (2021). *Python, Environmental Spectroscopy Lab*. Accessed: May 5, 2023. [Online]. Available: <https://github.com/EnSpec/hytools>
- [38] C. Wang et al., "Quantification of wetland vegetation communities features with airborne AVIRIS-NG, UAVSAR, and UAV LiDAR data in peace-athabasca delta," *Remote Sens. Environ.*, vol. 294, Jun. 2023, Art. no. 113646.
- [39] X. Li, F. Gao, L. Chen, and A. H. Strahler, "Derivation and validation of a new kernel for kernel-driven BRDF models," *Proc. SPIE*, vol. 3868, pp. 368–379, Dec. 1999.
- [40] F. Maignan, F.-M. Bréon, and R. Lacaze, "Bidirectional reflectance of Earth targets: Evaluation of analytical models using a large set of spaceborne measurements with emphasis on the hot spot," *Remote Sens. Environ.*, vol. 90, no. 2, pp. 210–220, Feb. 2004.
- [41] E. Greenberg et al., "An improved scheme for correcting remote spectral surface reflectance simultaneously for terrestrial BRDF and water-surface sunglint in coastal environments," *J. Geophys. Res., Biogeosci.*, vol. 127, no. 3, Mar. 2022, Art. no. e2021JG006712.
- [42] I. Meganem, Y. Deville, S. Hosseini, P. Déliot, X. Briottet, and L. T. Duarte, "Linear-quadratic and polynomial non-negative matrix factorization; application to spectral unmixing," in *Proc. 19th Eur. Signal Process. Conf.*, Aug. 2011, pp. 1859–1863.
- [43] I. Meganem, Y. Deville, S. Hosseini, P. Déliot, and X. Briottet, "Linear-quadratic blind source separation using NMF to unmix urban hyperspectral images," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1822–1833, Apr. 2014.
- [44] *The MathWorks*, document R2021b, Natick, MA, USA, 2021.
- [45] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.
- [46] W. Ng et al., "Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra," *Geoderma*, vol. 352, pp. 251–267, Oct. 2019.
- [47] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc.*, vol. 43, no. 5, pp. 772–777, Jul. 1989.
- [48] A. Stevens and L. Ramirez-Lopez. (2022). *An Introduction to the Prospectr Package*. [Online]. Available: <https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html>
- [49] R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. [Online]. Available: <https://www.R-project.org/>
- [50] M. Kuhn, "Building predictive models in R using the caret package," *J. Stat. Softw.*, vol. 28, no. 5, pp. 1–26, Jan. 2008, doi: 10.18637/jss.v028.i05.
- [51] R. A. Viscarra Rossel et al., "A global spectral library to characterize the world's soil," *Earth-Sci. Rev.*, vol. 155, pp. 198–230, Apr. 1, 2016.
- [52] A. Ahmadi, M. Emami, A. Daccache, and L. He, "Soil properties prediction for precision agriculture using visible and near-infrared spectroscopy: A systematic review and meta-analysis," *Agronomy*, vol. 11, no. 3, p. 433, Feb. 2021.
- [53] L. Ramirez-Lopez, T. Behrens, K. Schmidt, A. Stevens, J. A. M. Dematté, and T. Scholten, "The spectrum-based learner: A new local approach for modeling soil VIS-NIR spectra of complex datasets," *Geoderma*, vols. 195–196, pp. 268–279, Mar. 2013.
- [54] J. S. Shenk, M. O. Westerhaus, and P. Berzaghi, "Investigation of a LOCAL calibration procedure for near infrared instruments," *J. Near Infr. Spectrosc.*, vol. 5, no. 4, pp. 223–232, 1997.
- [55] L. Ramirez-Lopez et al. (2024). *Resemble: Regression and Similarity Evaluation for Memory-based Learning in Spectral Chemometrics: R Package Version 2.2.3*. [Online]. Available: <https://CRAN.R-project.org/package=resemble>
- [56] N. N. Reddy et al., "Legacy data-based national-scale digital mapping of key soil properties in India," *Geoderma*, vol. 381, Jan. 2021, Art. no. 114684.
- [57] D. M. Sherman and T. D. Waite, "Electronic spectra of Fe³⁺ oxides and oxide hydroxides in the near IR to near UV," *Am. Mineral.*, vol. 70, pp. 1262–1269, Dec. 1985.
- [58] A. C. Scheinost, A. Chavernas, V. Barrón, and J. Torrent, "Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify Fe oxide minerals in soils," *Clays Clay Minerals*, vol. 46, no. 5, pp. 528–536, Oct. 1998.

- [59] B. Stenberg, R. A. V. Rossel, A. M. Mouazen, and J. Wetterlind, "Visible and near infrared spectroscopy in soil science," *Adv. Agronomy*, vol. 2010, pp. 163–215, Jan. 2010, doi: [10.1016/S0065-2113\(10\)07005-7](https://doi.org/10.1016/S0065-2113(10)07005-7).
- [60] J. A. M. Demattê and F. da Silva Terra, "Spectral pedology: A new perspective on evaluation of soils along pedogenetic alterations," *Geoderma*, vols. 217–218, pp. 190–200, Apr. 2014.
- [61] K. W. Daniel, N. K. Tripathi, and K. Honda, "Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of lop buri (Thailand)," *Soil Res.*, vol. 41, no. 1, p. 47, 2003.
- [62] R. N. Clark, A. J. Gallagher, and G. A. Swayze, "Material absorption band depth mapping of imaging spectrometer data using a complete band shape least-squares fit with library reference spectra," in *Proc. 3rd Annu. JPL Airborne Geosci. Workshop*, vol. 54, 1990, pp. 176–186.
- [63] J. L. Post and P. N. Noble, "The near-infrared combination band frequencies of dioctahedral smectites, micas, and illites," *Clays Clay Minerals*, vol. 41, no. 6, pp. 639–644, 1993.
- [64] M. Colgan, C. Baldeck, J.-B. Féret, and G. Asner, "Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data," *Remote Sens.*, vol. 4, no. 11, pp. 3462–3480, Nov. 2012.
- [65] M. Wan et al., "Estimation of soil pH using PXRF spectrometry and vis-NIR spectroscopy for rapid environmental risk assessment of soil heavy metals," *Process Saf. Environ. Protection*, vol. 132, pp. 73–81, Sep. 2019.
- [66] E. Ben-Dor, "Quantitative remote sensing of soil properties," in *Advances in Agronomy*, vol. 75. New York, NY, USA: Academic, Jan. 2002, pp. 173–243.
- [67] L. Vaculíková and E. Plevová, "Identification of clay minerals and micas in sedimentary rocks," *Acta Geodynamica et Geomaterialia*, vol. 2, no. 2, p. 163, Apr. 2005.



Israr Majeed received the Ph.D. degree in agricultural and food engineering from IIT Kharagpur, Kharagpur, India, in 2024.

Currently, he is an Associate Scientist at the International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India. His research is focused on improving agricultural production using proximal and remote sensing approaches.



N. Nagarjuna Reddy received the Ph.D. degree in agricultural and food engineering from IIT Kharagpur, Kharagpur, India, in 2024.

Currently, he is an Associate Scientist at the International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India. He works on digital soil mapping, remote sensing, and landscape management.



Naveen Kumar Purushothaman received the M.Tech. degree in agricultural systems and management from the Department of Agricultural and Food Engineering, IIT Kharagpur, Kharagpur, India, in 2022, where he is currently pursuing the Ph.D. degree.

His research interests include multispectral and hyperspectral remote sensing and digital soil mapping.



Amanjit Premsagar received the M.Tech. degree in agricultural systems and management from the Department of Agricultural and Food Engineering, IIT Kharagpur, Kharagpur, India, in 2024. She is currently pursuing the Ph.D. degree with the Department of Geological Sciences, University of Alabama, Tuscaloosa, AL, USA.

Her research focuses on water quality assessment using remote sensing techniques.



Mayank Raj received the B.Tech. degree in agricultural engineering from the University of Agricultural Science, Raichur, Karnataka, India, in 2021, and the M.Tech. degree in agricultural systems and management from the Department of Agricultural and Food Engineering, IIT Kharagpur, Kharagpur, India, in 2023.



Lokesh Kumar Sinha received the Ph.D. degree in structural geology from the Banaras Hindu University, Varanasi, Uttar Pradesh, India, in 1988.

He is the former Director of Defence Terrain Research Laboratory (DTRL) and Snow and Avalanche Study Establishment (SASE), New Delhi, India. He works on satellite image processing, digital terrain analysis, and deformation monitoring.



Bhabani Sankar Das received the Ph.D. degree in agronomy from Kansas State University, Manhattan, KS, USA, in 1996.

He works as a Professor at the Department of Agricultural and Food Engineering, IIT Kharagpur, Kharagpur, India. He uses proximal and remote sensing approaches and geohydrological modeling for agricultural decision making.