Mapping and monitoring of soil organic carbon using regression analysis of spectral indices

Bullo Yami[1], N. J. Singh[1,*], B. K. Handique[2] and Sanjay Swami[1]

[1]College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University, Imphal–Umroi Road, Umiam 793 103, India
[2]North Eastern Space Applications Centre, Department of Space, Government of India, Umiam 793 103, India

The soil carbon sinking ability is dominantly controlled by local topographical settings, soil–crop management and traditional farming practices on which the food demand of the major population is dependent. The degradation of natural resources causing poor soil health is likely to strain the hilly and mountain ecosystem. This study aims to map soil organic carbon (SOC) of rice–fallow system under varying slopes and its changes during the past 20 years under traditional management practice using geospatial tools and techniques. Regression models of SOC were derived from remote sensing (RS)-based indices using multiple linear regression-stepwise (MLR-stepwise), partial least square regression (PLSR) and principal component analysis-regression (PCA-R). The MLR-stepwise model was found to be superior in performance with high R^2 (0.87) and least RMSE (0.026) compared to PLSR (R^2 = 0.71 and RMSE = 0.05) and PCA-R (R^2 = 0.27 and RMSE = 0.11) models for SOC prediction.

Keywords: Regression models, remote sensing, rice–fallow system, soil organic carbon, spectral indices.

THE heterogeneous traditional mountain communities living in close proximity are concerned with their indigenous farming practice and productivity. The traditional agricultural practice has been receiving attention for food security[1] and quality control for biodiversity conservation[2]. Soil degradation causes a decline in soil fertility and escalates soil erosion under varying slope gradients[3,4]. Soil organic carbon (SOC) indicates soil health[5] and provides nutrients to plants[6] and sustainable agroecosystem[7]. A small fraction of SOC is sensitive to changes in land management or environmental conditions[8,9]. Global issues related to carbon, i.e. carbon emission and decreasing soil carbon reserves under tremendous population pressure, can be minimized with judicious management. There is an urgent need to assess SOC in the hilly regions, as it is plagued by soil erosion and degradation[10]. There is scope for improving soil quality through soil nutrient pool and structure stability, enhancing cation exchange and water-holding capacity, as well as retaining toxic elements and microorganisms[11].

The ability to sink atmospheric carbon in the soil varies with topography, land use and management[12]. Topography-induced diverse local conditions within the same climatic zone lead to distinct SOC patterns and distribution. The sloppy land has a uniform variation of labile carbon fraction[13] and high SOC with vegetation[14]. On the other hand, labile carbon fraction is found in the order of upper slope > down slope > mid-slope regions[10]. In rice-cultivated areas, it is observed that the lower slope contains more SOC than the hilly slope[15]. Considering the importance of hilly topography, India's zero carbon emission policy can be achieved by properly managing hill SOC. Many studies suggested the enrichment of nutrients and SOC by incorporating crop residues into the soil[16]. However, crop residues and stubble are removed from the fields for fodder or fuel or burnt in the field. It is estimated that annually more than 5.3 billion tonnes of crop residues are produced worldwide, with increasing total production of cereal grains[17]. The concentration of total organic carbon in 10 years of continuous fertilizer has increased by 33% and treatment combination of fertilizer with wheat residues increases about 55% over the control[18].

Northeast India is one of the mega biodiversity hotspots. Rice is the dominant crop, where direct seeded rice is cultivated even up to a slope of 32%. Transplanted wet rice cultivation dominates the foothill regions with mostly a rice–fallow cropping system. In Meghalaya, NE India, indigenous tribal Khasi farmers follow a conventional farming system in the highly fragile undulating topography. The Ri-Bhoi district, popularly known as the 'Granary of Khasi hills', has fertile soil, and the local farmers don't sow seed but throw and wait for the harvest'[19]. However, crop productivity in such fragile land with traditional farming depends upon the natural enrichment of the soil. Timely checks of the fertility status of the soil, proper scientific interventions based on a quantitative assessment of SOC are important for sustaining crop production and providing scope for the potential of carbon sequestration[20]. Different methods have been used to estimate SOC in agricultural land. Most conventional methods for SOC estimation are based on the laboratory analysis of field soil samples[21–23], and aim for the maximum oxidation and recovery of carbon[24]. The sample point-based soil data are discrete in nature and incapable of providing continuous and complete information (i.e. spatial variability of SOC) of the entire study area. Numerous studies on the digital mapping of SOC have been conducted to address this issue

*For correspondence. (e-mail: naorem2005@gmail.com)

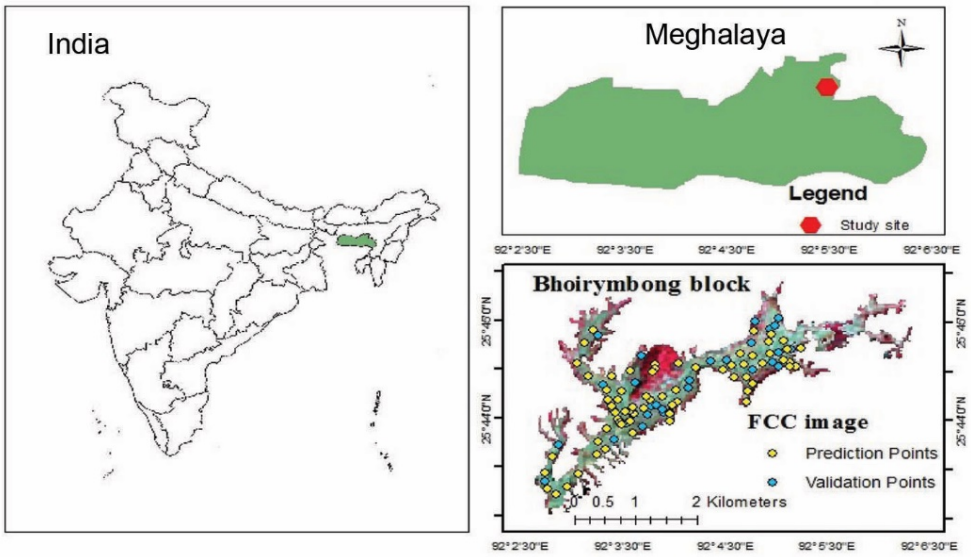

Figure 1. Study site with soil sampling locations.

based on spatial autocorrelation of the soil[25–28]. Many studies have proven that soil properties exhibit strong spatial dependence between neighbouring regions, and trend surface analysis, inverse distance weighted and geostatistical models have been successfully used to map and characterise the soil[29–31]. Martín *et al.*[32] used classical geostatistical tools to study temporal variability in agricultural soil. Similarly, Zhao *et al.*[33] studied the temporal variation of SOC in agricultural land on regional and continental scales using GIS-based spatial analysis and geostatistics tools. These methods merely rely on the correlation between soil sample points, limited by their geographical location. In other words, traditional geostatistical methods based on geospatial autocorrelation have two limitations: they are locally limited by sampling density[34,35] and ignore the role of environmental factors, resulting in inconsistent observations[36,37]. These methods encounter difficulty in describing the spatial distribution characteristics of SOC in complex terrains. The limitations of geostatistical approaches are overcome with indirect SOC estimation using remote sensing-based spectral indices[21,22]. Many researchers have estimated SOC content using the regression model built from the SOC data and remote sensing (RS) indices[23,38]. Yu *et al.*[39] suggested that climate parameters, mainly temperature and precipitation, management practice, and topographical settings such as slope, aspect and elevation influence SOC storage and variation. Also, soil properties significantly impact SOC stock and dynamics such as clay[40,41], pH[42] and alluvial deposition[43].

In this study, we assess the SOC storage of rice–fallow system under varying slope conditions with a regression approach taking SOC as the dependent variable and RS indices as predictors covering observations of the past 20 years (1999–2019). Considering the importance of SOC as a critical indicator of soil health in fragile hill ecosystems, geospatial information generated from the study is expected to support sustainable crop production and environmental planning.

Materials and methods

Description of the study area

The study was carried out in Bhoirymbong, Ri Bhoi district, Meghalaya, with a geographical extension of the study area from 91°20′30″ to 92°17′00″E longitude and 25°40′ to 26°20′N latitude, having an elevation of 1014 m amsl (Figure 1). Famers in Bhoirymbong follow rice–fallow cropping system covering an area of around 605 ha in varying slopes. Rice is transplanted as rainfed in June–July and harvested in October–November. The study area falls under the Eastern Himalayan Region (VII) Sub-Tropical Hill Zone (NEH-5) agro-climatic zone[44]. The soil is moderately acidic pH, fine, and excessively drained, varying from loamy to fine loamy Typic kandihumults. The area is confined to a moderately steep hill slope and is highly prone to erosion. The temperature ranges between 10°C and 30°C, and the average annual rainfall is about 1907.6 mm for 129 days a year.

Environmental and topographical variables

In the study, auxiliary variables like climate (temperature and rainfall) and land use (pine forest and rice–fallow system) were considered as constant, while the variable factors were soil attributes, elevation and slope (Figure 2). The elevation and slope were derived from the freely accessible 30 m Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) projected to WGS84 and EGM96

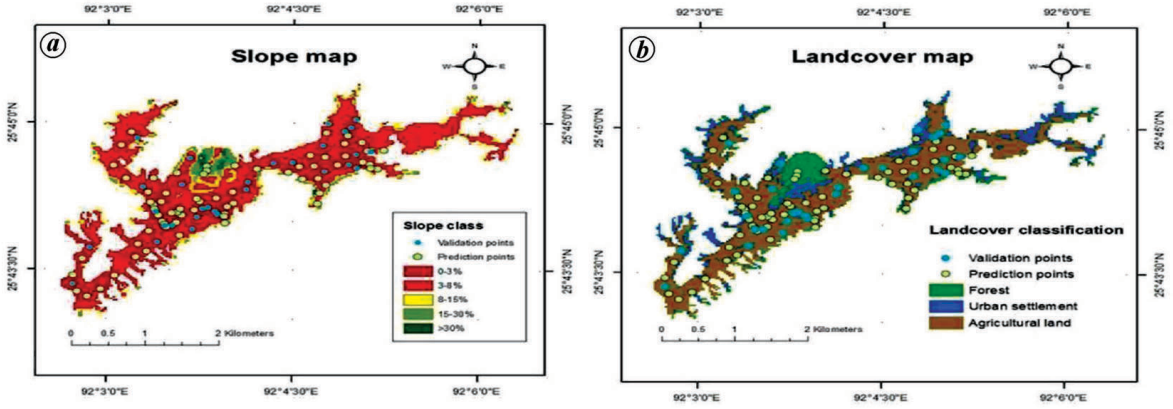

Figure 2. (***a***) Slope map and (***b***) land cover map of the study area.

Table 1. Detailed information of Landsat data

					Resolution			
					Spectral (μm)			
					Blue	Green	Red	NIR
Year	Satellite image	Acquisition date	Sensor	Spatial (m)	(0.45–0.51)	(0.52–0.60)	(0.63–0.68)	(0.85–0.89)
2019	Landsat 8	November	OLI	30	B2	B3	B4	B5
2009	Landsat 5	November	TM	30	B1	B2	B3	B4
1999	Landsat 7	December	ETM+	30	B1	B2	B3	B4
2014	SRTM		Slope, DEM	30				

(Earth Gravitational Model 1996) from the official website of USGS Earth Explorer (https://earthexplorer.usgs.gov). The DEM file was reprojected and reclassified by importing a coordinate system of Landsat data to ensure similar pixel size. The slopes were classified into five classes, viz. nearly level (0–3%), gentle slope (3–8%), moderate slope (8–15%), strong slope (15–30%) and steep slope (>30%) using the spatial analyst tool in ArcGIS 10.5 (Figure 2)[45]. Using Iso-cluster unsupervised classification, three distinct land-cover maps were generated based on grouping similar raster values for (a) agriculture, (b) forest and (c) urban settlement (Figure 2).

Remote sensing indices

Orthorectified WGS84 and UTM Zone 46 N datum, <10% cloud-free Landsat data for 1999 (Enhanced Thematic Mapper, ETM+), 2009 (Thematic Mapper, TM) and 2019 (Optical Land Imager (OLI) level-2 C2; path 136/row 042) with acquisition on 4 December 1999, 4 November 2009 and 16 November 2019 respectively were accessed from the website of USGS Earth Explorer USGS on 11 November 2021. The satellite data were nearly synchronized with soil sampling time. Table 1 presents the sensor characteristics. The shape file of the study area was prepared from the Google Earth Pro image and projected with the same projection parameters of Landsat data in ArcGIS 10.5. The study site was delineated and extracted with all the bands of Landsat data using the shape file. Top of atmosphere (TOA) reflectance was more consistent than surface reflectance[46]. In spatial SOC prediction using Landsat TM level-2 data, the digital number was converted to radiance and then to reflectance[47]. The atmospheric and radiometric correction of each band was converted to TOA reflectance using the following equation[48]

$$\rho_\lambda = M_p Q_{cal} + A_p/\sin(V_{se}),$$

where ρ_λ is the TOA planetary reflectance, M_p the band-specific multiplicative rescaling factor, A_p the band-specific additive rescaling factor, Q_{cal} the quantized and calibrated standard product pixel values and V_{se} is the sun elevation.

A total of 15 different RS indices as a predictor (independent) variables in model construction were derived, which include normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), renormalized difference vegetation index (RDVI), optimized soil adjusted vegetation index (OSAVI), modified soil adjusted vegetation index (MSAVI), infrared percentage vegetation index (IPVI), enhanced vegetation index (EVI2), redness index (RI), hue index (HI) and coloration index (CI). The selected soil indices were bare soil index (BSI) and normalized difference soil index (NDSI). Other indices considered were normalized difference water index (NDWI) and soil organic carbon index (SOCI; ref. 49; Table 2 and Figure 3). NDVI

Table 2. Remote sensing indices used in this study

Variables	Source/resolution/formula	Reference
Auxiliary		
Slope	SRTM (30 m)	
Elevation	SRTM (30 m)	
Vegetation indices		
NDVI	NDVI = (NIR – Red)/(NIR + Red)	89
SAVI	SAVI = (1 + L) ∗ (NIR – Red)/(NIR – Red + L)	60, 90
RDVI	RDVI = (NIR – Red)/√(NIR + Red)	91
OSAVI	OSAVI = (NIR – Red)/(NIR + Red + 0.16)	92
MSAVI	MSAVI = (1 + L) ∗ (NIR – Red)/(NIR – Red + L) L = 1 – 2 ∗ a ∗ NDVI ∗ WDVI a = NIR/Red	93
IPVI	IPVI = NIR/(NIR + R)	
EVI2	EVI2 = 2.5[(NIR – Red)/(NIR + 2.4Red + 1)]	94
Color indices		
RI	RI = Red^2/(Blue ∗ $Green^3$)	95, 96
HI	HI = (2Red – Green – Blue)/(Green – Blue)	95, 96
CI	CI = (Red – Green)/(Red + Green)	95, 96
Soil indices		
BSI	BSI = [(SWIR + Red) – (NIR + Blue)]/[(SWIR + Red) + (NIR + Blue)]	97
NDSI	NDSI = (middle IR – Green)/(middle IR + Green)	98, 99
Others		
NDWI	NDWI = (Green – NIR)/(Green + NIR)	100
SOCI	SOCI = Blue/(Red ∗ Green)	49

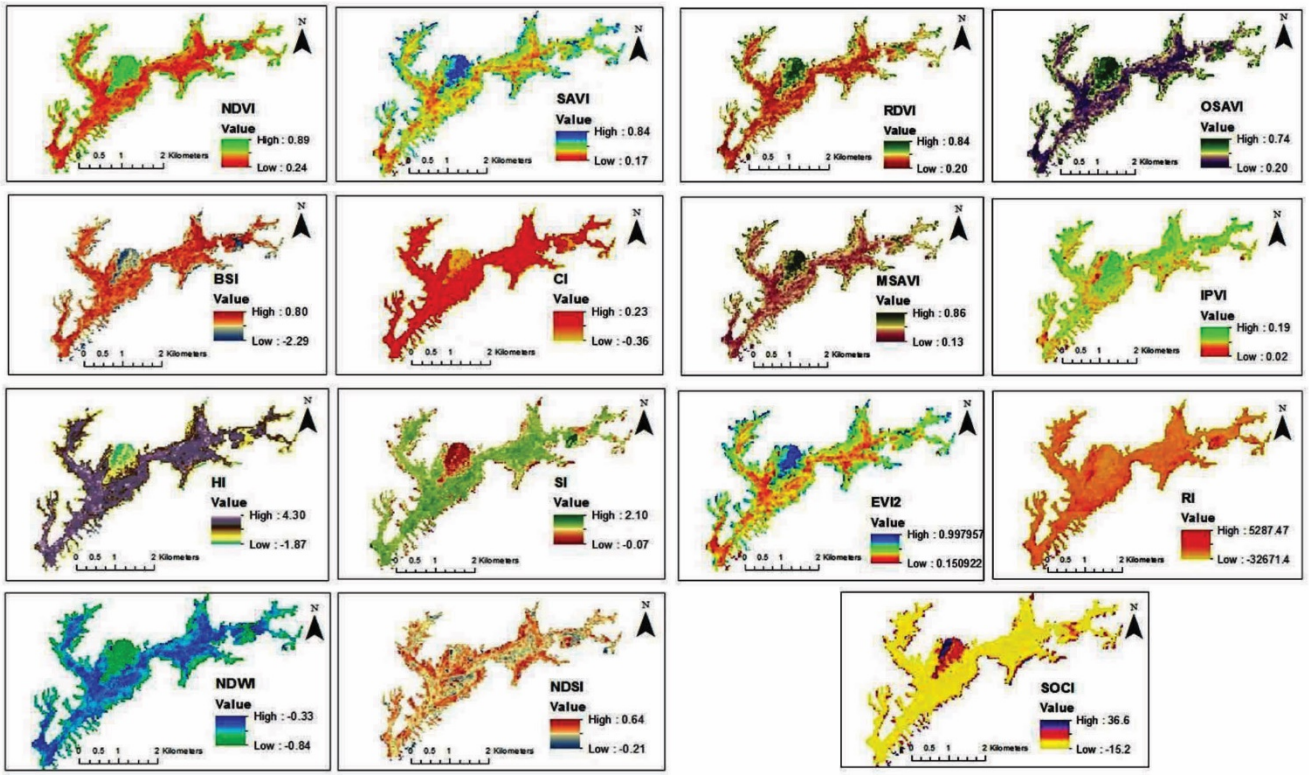

Figure 3. The remote sensing (RS)-derived (15) indices.

was used to monitor and predict the agricultural yield and biomass[50–52]. It was determined by the ratio of the normalized difference between recorded canopy reflectance in the red (600–700 nm) and near-infrared (750–1300 nm) bands[53]. NDVI values ranged between –1 and +1; positive values indicated vegetative or highly reflecting surfaces,

while negative values indicated non-vegetative or non-reflective surfaces[54]. It was sensitive to a small range of variations in spectral reflectance from vegetation[55,56]. The normalization reduces the effect of atmospheric and background soil[57] and also reduces the impact of degradation of satellite calibration[58,59]. The results can be improved by employing other vegetation indices such as SAVI, RDVI, OSAVI and MSAVI. SAVI was developed to minimize the influence of soil using a correction factor (L)[60]. The value of L ranged from 0 (high-density vegetation) to 1 (scattered vegetation, very high soil reflection), with 0.5 as the intermediate L value. SAVI was equal to NDVI with $L = 0$. SAVI was modified to MSAVI by replacing its correction factor (L) with a factor (L) of the product NDVI and weighted difference vegetation index (WDVI; WDVI = NIR – a ∗ red) as $L = 1 - 2a$ NDVI × WDVI, where a is the ratio of NIR/red. RDVI uses the difference between near-infrared and red wavelengths, indicating healthy vegetation. It is insensitive to the effects of soil and sun-viewing geometry. OSAVI is based on SAVI and uses a standard value of 0.16 for the canopy background adjustment factor, which provides greater soil variation for low vegetation cover. OSAVI is useful in areas of sparse vegetation (visible soil through the canopy). IPVI is functionally the same as NDVI, but computation is faster, with values ranging from 0 to 1. EVI is the outcome of a combination of blue (to correct for aerosol influences in the red band), red and NIR band reducing the atmospheric soil background effect by adjusting the factor L. RI defines the redness intensity of sand-grain coating with hydrates of iron oxide, i.e. goethite (FeOOH) and hematite (Fe_2O_3)[61], and during intense weathering liberates iron attached to clay minerals[62]. BSI is the normalized index of blue, red, NIR and SWIR bands to characterize soil variation, i.e. uncovered by grass, woodchips, live ground, artificial turf, etc. and its value ranges from –1 to +1. SWIR and red bands are used to quantify the soil mineral composition, while blue and NIR bands enhance vegetation[63]. A comparison between BSI and NDVI was carried out to estimate the barren areas. NDSI was derived from the normalized combination of middle infrared (2080–2350 nm) and green (520–600 nm), improving soil information and suppressing impervious surface area and vegetation. NDWI effectively measures the moisture content and water body, which are separated from soil and vegetation. It is calculated using green (maximum reflection from water surface) and NIR (maximum reflection from vegetation and soil) combination having values from 0.2 to 1 for water, 0.0 to 0.2 for flooding or humidity, –0.3 to 0.0 for moderate drought (soil and vegetation) and –1 to –0.3 for drought (soil and vegetation). SOCI was developed by a combination of blue, green and red bands with SOC data. This index is comparable to estimations from SWIR/NIR, i.e. 1608/833 nm ratio having a high correlation with SOC ($R^2 = 0.98$). The SWIR (1300–2500 nm) and NIR (700–1300 nm) portions are sensitive to SOC since the reflectance decreases with increasing SOC[64].

RS indices were estimated using a raster calculator to generate regression equations for predicting the dependent variable, i.e. SOC. The RS indices, slope and land cover for each sampling point were extracted using spatial analyst tools. Figure 4 presents the flowchart of SOC mapping. SOC was validated with 25% of the total samples.

Soil sampling and analysis

The coordinates of soil sample locations were recorded with a hand-held global positioning system (GARMIN, GPSMAP-64) having an accuracy of ±3 m. A total of 100 composite soil samples were collected at 0–20 cm depth after rice harvesting in November. Each composite sample was prepared from the ten randomly collected samples within 1 ha area, following the standard procedure[65,66]. Immediately, the samples were brought to the laboratory, air-dried, crushed, mixed thoroughly and sieved through 0.5 mm sieves for SOC estimation by the Walkley and Black wet oxidation method[67]. About 75% of the sample population was used as SOC predictors, and the remaining 25% for model validation. The model outputs were extrapolated for the subsequent years (1999–2019).

Statistical analysis

Pearson's correlation coefficient between SOC and the 15 RS indices was calculated at a 0.05 significance level using IBM SPSS package v26.2. Different regression models, viz. multiple linear regression-stepwise (MLR-stepwise), partial least square regression (PLSR) and principal component analysis regression (PCA-R), were used for SOC prediction. The model performance was evaluated by calculating the coefficient of determination (R^2) and root mean square error (RMSE). For the MLR-stepwise model, the predictor variables were selected one by one based on the performance of significance tests and the non-significant variables were removed. This process was run until no variable could be selected or rejected[68]. The general equation for MLR is as follows

$$Y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2, \ldots, \alpha_k x_k, \quad (1)$$

where Y is the dependent variable, $x_{1,2,\ldots,k}$ are the different independent variables, α_0 is a constant and $\alpha_{1,2,\ldots,k}$ is the regression coefficient for the corresponding independent variables.

PLSR is considered more robust when there is multicollinearity between the dependent and independent variables by projecting into new space instead of the hyperplane used in ordinary linear regression. It has been extensively studied to estimate SOC with various spectral indices[69]. The general equation for PLSR is as follows

$$X = TP\hat{} + E, \quad (2)$$

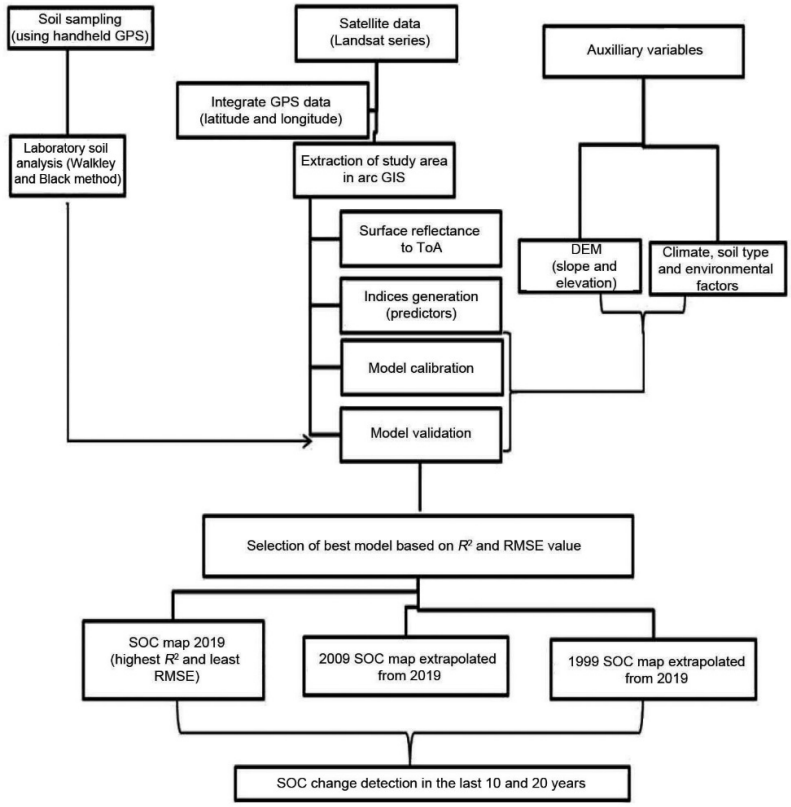

Figure 4. Schematic diagram of the detailed work plan.

$$Y = UQ^{\wedge} + F, \tag{3}$$

where X and Y are the predicted and response matrices respectively; T and U are the new projected matrices for X and Y respectively, $P^{\wedge}$ and $Q^{\wedge}$ are the orthogonal loading matrices; and E and F are the error terms for the predicted and response matrices respectively.

PCA-R is a combined mathematical method of principal component analysis and linear regression. It is based on the dimension reduction of variables having a high correlation to principal components (PCs), which is determined by the highest eigenvector value. Standardization of the data is done as follows

$$Y^* = (Y - \bar{y})/\mathrm{SD}, \tag{4}$$

where Y^* is the standardized data, Y the original data, $\bar{y}$ the mean, and SD is the standard deviation of Y. PCs are calculated based on the following matrix equation

$$\begin{vmatrix} \mathrm{PC}_1 \\ \mathrm{PC}_2 \\ \mathrm{PC}_3 \\ \cdots \\ \mathrm{PC}_n \end{vmatrix} = \begin{vmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ C_{31} & C_{32} & \cdots & C_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{vmatrix} \begin{vmatrix} Y_1^* \\ Y_2^* \\ Y_3^* \\ \cdots \\ Y_n^* \end{vmatrix}, \tag{5}$$

where $\mathrm{PC_n}$ is the n PC of the dataset, C_{nn} is the loading coefficient and Y_n^* are the standardized variables. The general equation showing the relation between the covariance matrix, eigenvectors and principal axes is as follows

$$C\mathrm{PC_n} = \alpha_n \mathrm{PC_n}, \tag{6}$$

where C is the covariance matrix and α_n is the nth largest eigenvalue of C.

The best-performing regression model for SOC prediction was selected using a raster calculator tool in ArcGIS 10.5 on the basis of R^2 and RMSE. The level of significance of the predicted SOC was tested with a one-sample t-test. The significant difference of SOC within varying slopes was analysed with one-way ANOVA using Post-hoc Gabriel's test ($P < 0.05$). Finally, the change detection of SOC during the past 10 (2009) and 20 (1999) years from the base year (2019) was evaluated with a sample t-test taking the mean SOC value of each year. All the predictors and other auxiliary variables were described numerically. All the statistical analyses were performed using XLSTAT (2022.1) package, MS-Excel and IBM SPSS package v26.2 (Table 2).

Model validation

Laboratory analysed SOC, i.e. actual SOC variables (Table 3), were cross-validated with model-estimated SOC using

R^2 and RMSE (Table 4). The model with the least RMSE and highest R^2 values was considered ideal and selected[23,70]. The lower RMSE value indicates more concentrated data around the line of best fit, i.e. less error[71].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{j=1}^{n}(y_i - \bar{y})^2}; \quad \text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

where y_i is the actual SOC, $\hat{y}_i$ the predicted SOC, $\bar{y}$ the mean of actual SOC and n is the number of samples. Figure 4 is a schematic diagram of the detailed work plan.

Results and discussion

Descriptive statistics

Table 5 and Figure 3 present the descriptive statistics of SOC and RS indices. The data were observed to follow a normal distribution upon removing the outliers; hence, no transformation was performed. SOC content ranged between 1.24% and 2.87% with a mean value of 2.23%, highest (2.87%) at strongly sloping (15–30%) and lowest (1.24%) at gently sloping category (3–8%). This difference in SOC is associated with varying slope positions and steepness of topography which may result in the microclimate (local climatic) and typical geological processes, leading to distinguishing vegetation types. Consequently, uneven SOC distribution occurs along the slope with different vegetation and is also caused by the anthropogenic activities of soil–crop management. The spatial distribution of SOC in the forest and vineyard soils is significantly influenced by the slope and aspect; the highest SOC content (1.09%) was observed in north and north-facing soils and lowest in the southeast[72,73]. The predicted SOC mean (2.56%) was not significantly different from the actual SOC mean (2.24; t-value < 1.96; $P < 0.05$).

NDVI value of 0.53 ± 0.05 indicates moderate vegetation due to the standing rice stubble, which is cut approximately 30 cm above the soil. The fields with litter from pine forests in strongly sloping (15–30%) areas also contribute to high NDVI. This results in a low BSI value of –0.08 ± 0.16. An intermediate SAVI value of 0.45 ± 0.05 indicates neither dense vegetation nor very high soil reflection. The SAVI is confirmed with MSAVI, in which the soil brightness correction factor (L) is introduced.

The indices such as SI (0.44 ± 0.03), RDVI (0.43 ± 0.05) and OSAVI (0.43 ± 0.04) have almost the same value, indicating the presence of rice stubbles. IPVI is 0.11 ± 0.01, which is low, indicating high soil reflection. The reflection of red band from the soil may be high, resulting in a larger denominator (NIR + Red) value of IPVI (IPVI = NIR/(NIR + Red)). The RI value of 177.45 ± 54.05 indicates the presence of sufficient iron oxide in the soil. RI defines the redness intensity of sand-grain coating with hydrates of iron oxide, i.e. goethite (FeOOH) and hematite (Fe_2O_3)[61], and during intense weathering, it liberates iron attached to clay minerals[62]. NDSI of 0.21 ± 0.06 improves soil information and suppresses the impervious surface area and vegetation. NDWI value of –0.58 ± 0.04 within the range –0.3 to 0.0 indicates moderate drought (soil and vegetation) after harvesting of the crop, as a high SOCI value (2.95 ± 0.48) is sensitive to SOC content[64]. Table 5 and Figure 3 also present other indices such as HI (1.56 ± 0.30), CI (0.07 ± 0.03) and EVI2 (0.47 ± 0.06).

Table 3. Ground control points used for model validation

Latitude (N)	Longitude (E)	Elevation (m)	Soil organic carbon (%)	Slope class
25.743917	92.06117	895	1.83	3
25.7432	92.0725	878	2.33	1, 2
25.7447	92.0835	891	1.44	2
25.7486	92.082	889	2.6	2
25.7506	92.0835	909	2.14	1
25.745184	92.085483	–	1.37	1
25.742968	92.082547	–	2.15	1
25.743261	92.075019	887	2.05	1, 2
25.747319	92.054105	876	2.09	1, 2
25.74992	92.079743	–	2.58	3
25.749278	92.082819	894	2.52	1, 2
25.746225	92.079093	884.24	2.23	1, 2
25.739056	92.054908	878.095	2.27	1
25.734806	92.056611	868.03	2.31	2
25.7325	92.057722	868.945	1.94	1
25.732444	92.058833	869.25	1.87	1, 2
25.733972	92.062167	868.945	2	1
25.735333	92.062917	873.52	1.87	1
25.735583	92.063639	870.165	1.81	1
25.7344	92.0649	879	2.33	1
25.7349	92.0646	873	2.37	1
25.7418	92.0793	886	2.36	2, 3
25.742358	92.083616	879	2.49	2
25.738624	92.068782	–	2.22	1
25.739857	92.069044	874	2.31	1, 2
25.731949	92.061549	876.7	2.57	1, 2
25.739458	92.060098	887	2.57	1, 2
25.729715	92.056877	890	2.63	1, 2
25.728756	92.047725	877	2.29	2
25.722532	92.045532	869	2.12	2

Note: 1, Nearly level (0–3%); 2, Gentle slope (3–8%); 3, Moderate slope (8–15%); 4, Strong slope (15–30%) and 5, Steep slope (>30%).

Effect of slope on SOC and RS indices

The effect of slope on SOC is statistically insignificant due to low SOC variability (CV < 15%), which is related to a non-uniform sample size of each slope position. This may be due to the fact that the highest sample size (70%) was from nearly level (0–3%) and gentle slope (3–8%), while 10% from moderate slope (8–15%), strong slope (15–30%) and steep slope (>30%). It followed Gabriel's posthoc test

Table 4. Validation of regression model-predicted SOC with actual SOC data (validation data)

Regression models	Validation R^2	Validation RMSE
MLR-stepwise SOC (%) = 1.47 + 0.131 * SOCI + 0.801 * NDVI	0.68	0.038
PLSR SOC (%) = 1.59 + 0.027 * SOCI + 0.23 * NDVI + 0.143 * SAVI – 0.05 * BSI – 0.18 * SI + 0.0002 * RI + 0.12 * RDVI + 0.24 * SAVI – 0.33 * NDWI + 0.11 * NDSI + 0.12 * MSAVI – 0.68 * IPVI – 0.02 * HI + 0.11 * EVI2 – 0.19 * CI	0.5	0.05

Table 5. Descriptive statistics of SOC and remote sensing (RS) indices 2019

Parameters	Mean	Standard deviation	Kurtosis	Skewness	Minimum	Maximum
SOC	2.23	0.09	0.36	–0.06	1.98	2.45
SOC (100)	2.24	0.22	7.39	–1.09	1.28	2.87
NDVI	0.57	0.1	1.68	1.36	0.4	0.86
SAVI	0.48	0.09	3.42	1.73	0.36	0.84
BSI	–0.16	0.27	3.27	–1.4	–1.22	0.31
SI	0.42	0.08	4.38	–2.03	0.12	0.53
RI	173.57	56.51	–0.12	0.27	50.85	320.18
RDVI	0.46	0.09	4.09	1.84	0.34	0.84
OSAVI	0.46	0.08	2.25	1.52	0.33	0.74
MSAVI	0.48	0.1	2.91	1.69	0.35	0.85
HI	1.37	0.59	2.71	–1.67	–0.58	2.27
CI	0.04	0.09	5.66	–2.36	–0.29	0.13
NDWI	–0.6	0.06	0.72	–1.01	–0.77	–0.5
IPVI	0.11	0.02	3.32	1.37	0.09	0.18
NDSI	0.21	0.06	–0.61	–0.08	0.07	0.34
EVI2	0.5	0.12	4.65	1.98	0.36	1
SOCI	3.43	1.4	10.24	2.83	1.81	11.02

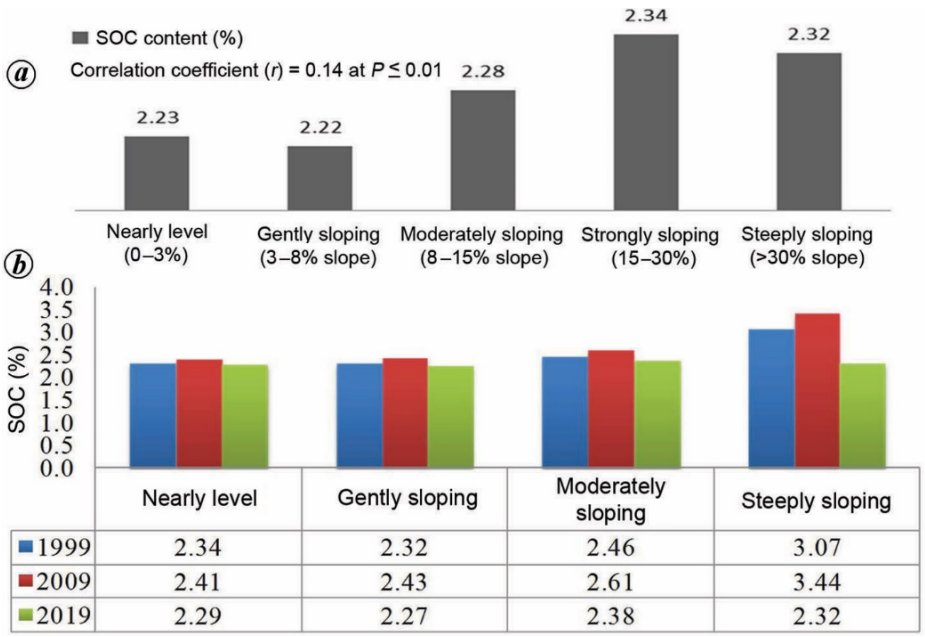

Figure 5. ***a***, Actual (field data) soil organic carbon (SOC; %) content at varying slope. ***b***, MLR-stepwise-predicted and PLSR-predicted SOC (%).

($P > 0.05$; $r = 0.14$) analysis (Figure 5 and Table 6). Wilding[74] reported similar results, while, Hamzehpour *et al.*[75] found low SOC variation in slope attributed to a small sample size taken from <8% slopy agricultural land. There was a positive correlation between slope and SOC due to the surrounding pine forests, with similar findings reported

by Singh and Benbi[20], and Wang *et al.*[76]. A small SOC value (1.24%) at the lower slope (0–8%) was observed compared to that at other slopes, i.e. 2.85% at the steep slope (15–30%). Similarly, Singh *et al.*[77] found that SOC content at 32% slope land (Pyllun) > 12% slopy land owing to the presence of pine trees which contain waxy-coated, needle-shaped leaves that undergo slow decomposition resulting in high particulate organic matter (POM). The SOC content at 32% slope had decreased from the uppermost part of the slope ($L1$) to the lowest part of the slope ($L3$)[78]. The highest POM content was observed at 32% slope of Pyllun (rice–fellow) than in the intermountain valley (Bhoirymbong). Similarly, Singh and Benbi[20] also found high SOM at hilltops (62% slope) > backslope and bottomland (42%) and middle slope (9%). However, other forms of SOC, such as light fraction organic matter, dissolved organic carbon (DOC), microbial biomass carbon (MBC) and hot water soluble carbon, were high at the lower slope (i.e. 6%) positions and minimum at 32% slope of Paham (Nongpoh)[78].

High values of vegetation indices such as SOCI, NDWI, NDVI, SAVI, OAVI, MSAVI, EVI2 and IPVI were observed at higher slope positions due to quick growing pastures after harvesting of rice along with the effect of surrounding forests. However, soil base indices, viz. spectral colour indices like BSI, CI, HI, NDSI, RI and SI have high values soil at lower slope (0–8%) due to soil exposure after crop harvested (Figure 3). Wang *et al.*[76] reported low index values with low SOC content in agricultural areas with bare soil and sparse vegetation.

Table 6. Multiple comparisons of slope and SOC (dependent variable) using post-hoc Gabriels's test ($P = 0.05$)

Slope (I)	Slope class (J)	Mean difference (I–J)	Standard error	Sigma
1	2	0.01131	0.06251	1
	3	–0.05387	0.09043	1
	4	–0.10854	0.10739	0.969
	5	–0.08187	0.16825	1
2	1	–0.01131	0.06251	1
	3	–0.06518	0.07612	0.985
	4	–0.11985	0.09566	0.811
	5	–0.09318	0.16101	0.998
3	1	0.05387	0.09043	1
	2	0.06518	0.07612	0.985
	4	–0.05467	0.11584	1
	5	–0.028	0.17377	1
4	1	0.10854	0.10739	0.969
	2	0.11985	0.09566	0.811
	3	0.05467	0.11584	1
	5	0.02667	0.18317	1
5	1	0.08187	0.16825	1
	2	0.09318	0.16101	0.998
	3	0.028	0.17377	1
	4	–0.02667	0.18317	1

Note: 1, Nearly level (0–3%); 2, Gentle slope (3–8%); 3, Moderate slope (8–15%); 4, Strong slope (15–30%) and 5, Steep slope (>30%).

SOCI ($r = 0.87$) and NDVI ($r = 0.72$) were strongly correlated with SOC ($r > 0.7$) as NIR is associated with carbon–hydrogen bonds of SOC[79]. The indices such as SAVI ($r = 0.38$), OSAVI ($r = 0.58$), MSAVI ($r = 0.35$), NDSI ($r = 0.46$) and EVI2 ($r = 0.34$) were moderately correlated ($0.3 < r < 0.7$) with SOC, while BSI ($r = –0.58$), SI ($r = –0.40$), HI ($r = –0.43$), CI ($r = –0.42$), NDWI ($r = –0.69$) and IPVI ($r = –0.48$) were negatively correlated with SOC content. RI ($r = –0.15$), RDVI ($r = 0.11$) and slope ($r = 0.14$) were poorly correlated ($r < 0.3$) with SOC (Table 7). Moderate to strong correlation is attributed to strong absorption in red and NIR bands, similar to the findings in earlier studies[49,80]. The darker soil colour having high SOC is strongly correlated with soil colour indices. This result is in line with the findings of Son *et al.*[81] and Nuarsa *et al.*[82].

Model prediction and validation

Among the selected statistical approaches, the performance of MLR-stepwise ($R^2 = 0.87$ and RMSE = 0.026) was found to be the best compared to PLSR ($R^2 = 0.71$ and RMSE = 0.05) and PCA-R ($R^2 = 0.27$ and RMSE = 0.11) (Table 8). Table 4 and Figure 6 depict field validation results using the actual field data. Among the indices as dependent variables, NDVI and SOCI were strongly correlated with SOC irrespective of slope position ($r \geq 0.70$), whereas other indices were moderately correlated with SOC content ($0.3 < r < 0.7$). With the MLR-stepwise method, predictor variables were selected based on the significance test. This process was run until no variable could be selected or rejected[68]. Finally, the stage was reached where most of the influencing factors of NDVI and SOCI were included in the model for SOC prediction under varying topographical conditions. Studies have reported better performance of MLR-stepwise over the PLSR method, where the inclusion of multicollinearity variables in the regression model reduces its performance[70,76].

Change detection of SOC

The MLR-stepwise model was used for SOC change detection for 10 (2009–19) and 20 (1999–2019) years. About –0.17% and –0.09% decrease in SOC was observed during the past 10 and 20 years respectively (Table 9 and Figure 7). Traditional farming practices without using organic or inorganic nutrient sources coupled with a monoculture farming system may be the reason for declining SOC over time. Meghalaya received less rainfall (1205 mm) during 2009 than the average annual rainfall of 2173 mm (refs 83, 84). The less decomposition and leaching loss with low precipitation may be attributed to higher SOC content in 2009 under similar land management[85]. In 1999 and 2019, increasing precipitation and temperature enhanced the decomposition rate lowering SOC concentration[86]. Laxminarayana

Table 7. (Left) Correlation between SOC and RS indices as well as slope. (Right) Effect of slope on SOC and RS indices

Slope	SOC	Slope	1–3%	3–8%	8–15%	15–30%	>30%
SOC	1	NDVI	0.56	0.53	0.66	0.8	0.73
NDVI	0.72**	SAVI	0.48	0.45	0.55	0.71	0.68
SAVI	0.38**	BSI	–0.14	–0.08	–0.39	–0.72	–0.36
BSI	–0.58**	SI	0.45	0.44	0.37	0.22	0.26
SI	–0.40**	RI	193.72	174.5	189.59	119.02	65.04
RI	–0.15	RDVI	0.45	0.43	0.52	0.69	0.66
RDVI	0.11	OSAVI	0.45	0.42	0.52	0.66	0.61
OSAVI	0.58**	MSAVI	0.47	0.44	0.56	0.74	0.69
MSAVI	0.35**	HI	1.54	1.56	0.93	–0.07	0.04
HI	–0.43**	CI	0.07	0.07	–0.02	–0.18	–0.17
CI	–0.42**	NDWI	–0.6	–0.58	–0.65	–0.73	–0.66
NDWI	–0.69**	IPVI	0.11	0.11	0.11	0.14	0.15
IPVI	–0.48**	NDSI	0.21	0.21	0.23	0.2	0.14
NDSI	0.46**	EVI2	0.5	0.46	0.58	0.8	0.76
EVI2	0.34**	SOCI	3.05	3.03	4.45	6.65	4.67
SOCI	0.87**						
Slope	0.14**						

**Correlation is significant at the 0.01 level.

Table 8. Regression model for SOC prediction using RS variables

Model	Models	R^2	RMSE
MLR-stepwise	SOC (%) = 1.47 + 0.131 * SOCI + 0.801 * NDVI	0.87	0.026
PLSR	SOC (%) = 1.59 + 0.027 * SOCI + 0.23 * NDVI + 0.143 * SAVI – 0.05 * BSI – 0.18 * SI + 0.0002 * RI + 0.12 * RDVI + 0.24 * SAVI – 0.33 * NDWI + 0.11 * NDSI + 0.12 * MSAVI – 0.68 * IPVI – 0.02 * HI + 0.11 * EVI2 – 0.19 * CI	0.71	0.054
PCA-R	SOC (%) = 2.22 + 0.06 * PC1 + 0.0078 * PC2	0.27	0.11

Note: MLR-stepwise, Multiple linear regression-stepwise; PLSR, Partial least square regression; PCA-R, Principal component analysis regression.

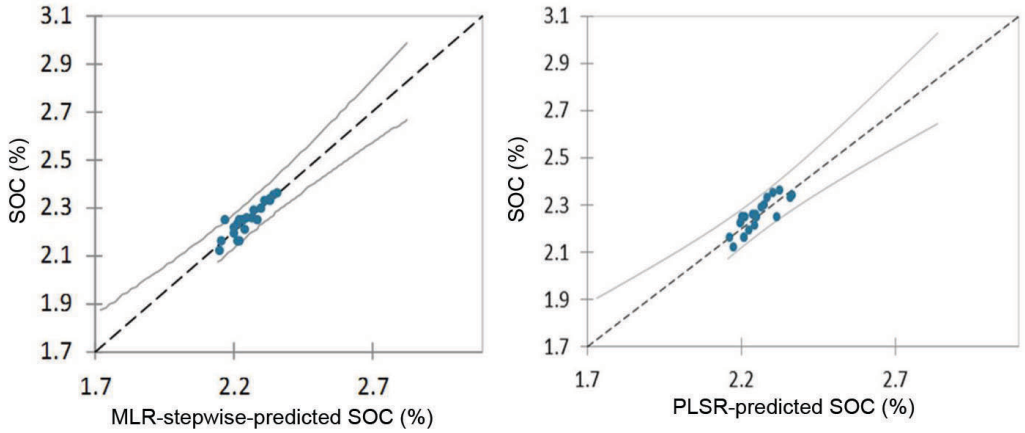

Figure 6. Model accuracy by plotting between actual SOC (%) and MLR-stepwise and PLSR-predicted SOC (%).

and Bharati[87] reported 2.16–2.77% SOC in rice fields in 2009. The fine, loamy texture had higher SOC[88].

Conclusion

Challenges of SOC quantification with limited laboratory facilities and the complexity of hilly terrain can be overcome by adopting advanced geospatial tools and techniques. Spatial maps of SOC and quantifying its changes due to long-term cultivation (rice–fallow) will be beneficial for managing soil fertility in the study area. The remote sensing-based indices could be promising variables for SOC prediction and preparing temporal SOC distribution maps. The terrain complexity of the hilly region creates micro-climatic zones resulting in uneven vegetation and biomass

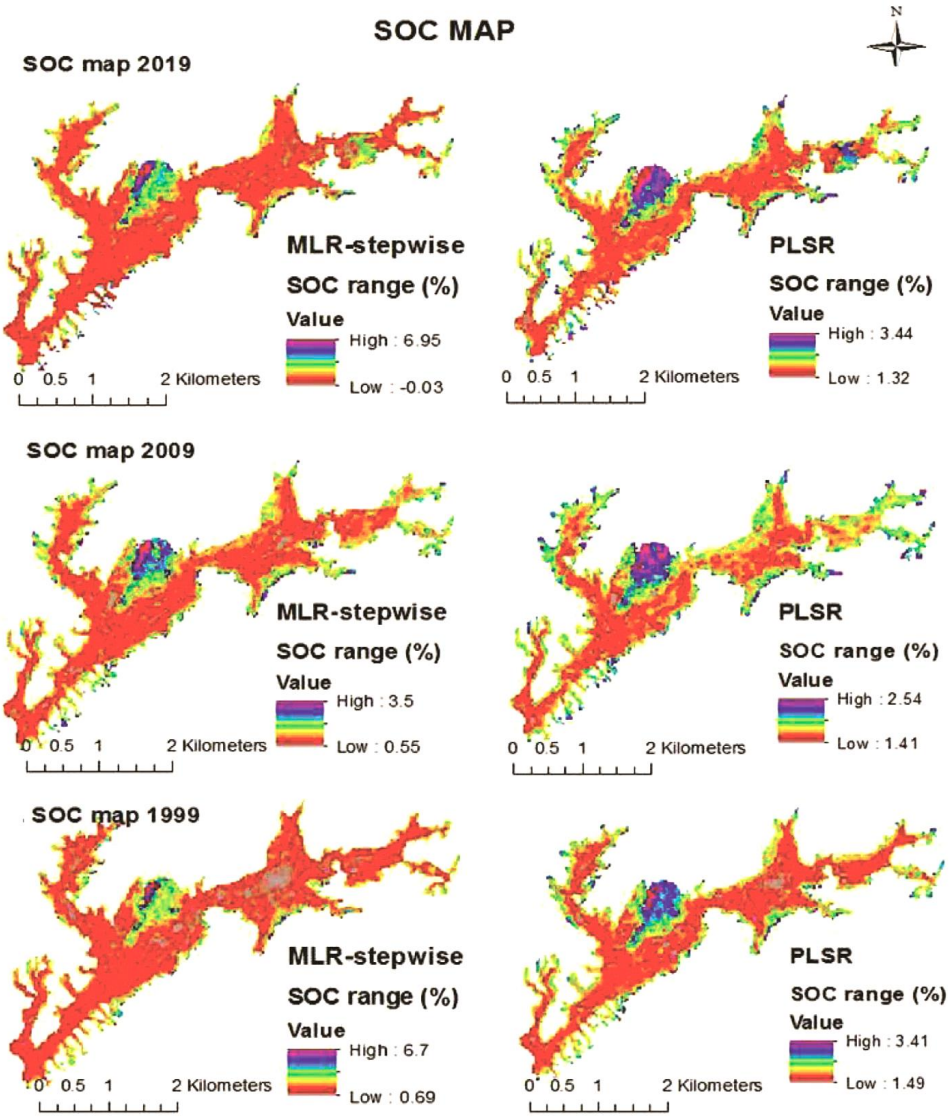

Figure 7. Mapping of SOC for 2019, 2009 and 1999 using MLR-stepwise and PLSR models.

Table 9. Changes of SOC during the last 10 (2009–19) and 20 (1999–2019) years

One-sample test Test value (2019-predicted SOC) = 2.56			
Year	*t* value	Degree of freedom (d*f*)	Sigma (two-tailed)
2009	1.292	4	0.266
1999	1.023	4	0.364

accumulation. Under such conditions, remote sensing-based information along with digital elevation models is the right combination for SOC prediction.

1. Rajasekaran, B. and Whiteford, M. B., Rice–crop production system: The role of indigenous knowledge in designing food security policies. *Food Policy*, 1992, **18**(3), 237–247.
2. Gadgil, M., Berkes, F. and Folke, C., Indigenous knowledge for biodiversity conservation. *Ambio*, 1993, **22**, 151–156.
3. Kala, C. P., Traditional ecological knowledge on characteristics, conservation and management of soil in tribal communities of Pachmarhi Biosphere Reserve, India. *J. Soil Sci. Plant Nutr.*, 2013, **13**(1), 201–214.
4. Sanders, D., Soil conservation. In *Land Use, Land Cover and Soil Sciences* (ed. Willy, H. V.), UNESCO, Eolss Publishers, Oxford, UK, 2004, **4**, 1–21.
5. Arunrat, N., Kongsurakan, P., Sereenonchai, S. and Hatano, R., Soil organic carbon in sandy paddy fields of northeast Thailand: a review. *Agronomy*, 2020, **10**(8), 1061.

6. Zhang, Y. *et al.*, Prediction of soil organic carbon based on Landsat 8 monthly NDVI data for the Jianghan Plain in Hubei Province, China. *Remote Sensing*, 2019, **11**, 1683; doi:10.3390/rs11141683.
7. Meetei, T. T., Kundu, M. C. and Devi, Y. B., Long-term effect of rice-based cropping systems on pools of soil organic carbon in farmer's field in hilly agroecosystem of Manipur, India. *Environ. Monit. Assess.*, 2020, **192**(4), 1–17.
8. Wu, J. S. and Xiao, H. A., Measuring the gross turnover time of soil microbial biomass C under incubation. *Acta Pedol. Sin.*, 2004, **41**(3), 401–407.
9. Lui, D. *et al.*, Spatial distribution of soil organic carbon and analysis of related factors in cropland of the black soil region, Northeast China. *Agric. Ecosyst. Environ*., 2006, **113**, 73–81.
10. Li, Z. W., Nie, X. D., Chen, X. L., Lu, Y. M., Jiang, W. G. and Zeng, G. M., The effects of land use and landscape position on labile organic carbon and carbon management index in red soil hilly region, sourthern China. *J. Mt. Sci.*, 2015, **12**(3); doi:10.1007/s11629-013-2964-2.
11. Bayer, C. and Mielniczuk, J., Capítulo 2: Dinâmica da matéria orgânica. In *Fundamentos da Matéria orgânica no solo: Ecossistemas Tropicais e Subtropicais* (eds Santos, G. A. *et al.*), 07–18, Porto Alegre, Metropole, Brazil, 2008.
12. Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J. and Skjemstad, J. O., Visible, near infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 2006, **131**, 59–75.
13. Shi, P. *et al.*, Land-use types and slope topography affect the soil labile carbon fractions in the loess hilly–gully area of Shaanxi, China. *Arch. Agron. Soil Sci.*, 2020, **66**(5), 638–650.
14. Zhang, X. *et al.*, Which slope aspect and gradient provides the best afforestation-driven soil carbon sequestration on the China's Loess plateau? *Ecol. Eng.*, 2020, **147**, 105782.
15. Jendoubi, D., Liniger, H. and Ifejika Speranza, C., Impacts of land use and topography on soil organic carbon in a Mediterranean landscape (north-western Tunisia). *Soil*, 2019, **5**(2), 239–251.
16. Rengel, Z., The role of crop residues in improving soil fertility. In *Nutrient Cycling in Terrestrial Ecosystems*, Springer, Berlin, Germany, 2007, pp. 183–214.
17. Jiang, D., Zhuang, D. and Huang, Y., Crop residues as an energy feedstock: availability and sustainability. In *Sustainable Bioenergy Production*, CRC Press, Boca Raton, Florida, USA, 2014, pp. 236–249.
18. Zhang, L. *et al.*, Soil labile organic carbon fractions and soil enzyme activities after 10 years of continuous fertilization and wheat residue incorporation. *Sci. Rep.*, 2020, **10**, 11318; https://doi.org/10.1038/s41598-020-68163-3.
19. *The Meghalayan*, Dispatches from the farmlands of Ri-bhoi; https://themeghalayan.com/dispatches-from-the-farmlands-of-Ri-Bhoi (accessed on 6 April 2022).
20. Singh, P. and Benbi, D. K., Soil organic carbon pool changes in relation to slope position and land-use in Indian lower Himalayas. *Catena*, 2018, **166**, 171–180.
21. Angelopoulou, T., Tziolas, N., Balafoutis, A., Zalidis, G. and Bochtis, D., Remote sensing techniques for soil organic carbon estimation: a review. *Remote Sensing*, 2019, **11**(6), 676; https://doi.org/10.3390/rs11060676.
22. Vohland, M., Ludwig, B., Seidel, M. and Hutengs, C., Quantification of soil organic carbon at regional scale: benefits of fusing Vis-NIR and MIR diffuse reflectance data are greater for *in situ* than for laboratory-based modelling approaches. *Geoderma*, 2022, **405**, 115426.
23. Bhunia, G. S., Shit, P. K. and Pourghasemi, H. R., Soil organic carbon mapping using remote sensing techniques and multivariate regression model. *Geocarto Int.*, 2019, **34**(2), 215–226.
24. Partyka, T. and Hamkalo, Z., Estimation of oxidizing ability of organic matter of forest and arable soil. *Zemdirbyste-Agric.*, 2010, **97**, 33–40.
25. Wu, H. Y., Zeng, F. P., Song, T. Q., Peng, W. X., Li, X. H. and OuYang, Z. W., Spatial variations of soil organic carbon and nitrogen in peak-cluster depression areas of Karst region. *Plant Nutr. Fert. Sci.*, 2009, **15**, 1029–1036.
26. Zhang, L., Gao, P., Wang, C., Liu, S. and Li, X., Spatial distribution of soil organic carbon in the forestland of the Yaoxiang small watershed in central and southern Shandong Province. *Sci. Soil Water Conserv*., 2015, **13**, 83–89.
27. Kumar, S., Lal, R. and Liu, D., A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*, 2012, **189**, 627–634.
28. Chabala, L. M., Mulolwa, A. and Lungu, O., Application of ordinary kriging in mapping soil organic carbon in Zambia. *Pedosphere*, 2017, **27**, 338–343.
29. Lu, F., Zhao, Y., Huang, B. and Wang, J., Comparison of predicting methods for mapping the spatial distribution of topsoil organic matter content in cropland of hailun. *J. Soil Sci*., 2012, **43**, 662–667.
30. Xu, E. and Zhang, H., Multi-scale analysis of kriging interpolation and conditional simulation for soil organic matters in newly reclaimed area in Yili. *Soils*, 2013, **45**, 91–98.
31. Zhao, D., Zhao, H., Rao, J. and Gao, X., Analysis of the spatial distribution pattern of cultivated land quality and the influential factors based on trend-surface. *Res. Soil Water Conserv.*, 2015, **22**, 219–223.
32. Martín, J. A. R. *et al.*, Soil organic carbon stock on the Majorca Island: temporal change in agricultural soil over the last 10 years. *Catena*, 2019, **181**, 104087.
33. Zhao, M. S., Qiu, S. Q., Wang, S. H., Li, D. C. and Zhang, G. L., Spatial–temporal change of soil organic carbon in Anhui Province of East China. *Geoderma Reg.*, 2021, **26**, e00415.
34. Guo, L., Linderman, M., Shi, T. Z., Chen, Y. Y., Duan, L. J. and Zhang, H. T., Exploring the sensitivity of sampling density in digital mapping of soil organic carbon and its application in soil sampling. *Remote Sensing*, 2018, **10**, 27.
35. Lin, Y., Zhu, A., Qin, C., Li, B. and Pei, T., A soil sampling method based on representativeness grade of sampling points. *Acta Pedol. Sin*., 2011, **48**, 938–946.
36. Liu, Y., Guo, L., Jiang, Q., Zhang, H. and Chen, Y., Comparing geospatial techniques to predict soc stocks. *Soil Till. Res*., 2015, **148**, 46–58.
37. Malone, B. P., Jha, S. K., Minasny, B. and McBratney, A. B., Comparing regression-based digital soil mapping and multiple-point geostatistics for the spatial extrapolation of soil data. *Geoderma*, 2016, **262**, 243–253.
38. Žížala, D., Minařík, R. and Zádorová, T., Soil organic carbon mapping using multispectral remote sensing data: prediction ability of data with different spatial and spectral resolutions. *Remote Sensing*, 2019, **11**(24), 2947.
39. Yu, H., Zha, T., Zhang, X., Nie, L., Ma, L. and Pan, Y., Spatial distribution of soil organic carbon may be predominantly regulated by topography in a small revegetated watershed. *Catena*, 2020, **188**, 104459.
40. Zhong, Z. *et al.*, Relationship between soil organic carbon stocks and clay content under different climatic conditions in Central China. *Forests*, 2018, **9**(10), 598.
41. Churchman, G. J., Singh, M., Schapel, A., Sarkar, B. and Bolan, N., Clay minerals as the key to the sequestration of carbon in soils. *Clays Clay Miner.*, 2020, **68**(2), 135–143.
42. Zhang, H., Kang, J., Xu, X. and Zhang, L., Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: a case study of Yi'an County, Heilongjiang Province, China. *Comput. Electron. Agric*., 2020, **176**, 105618.
43. Ferugoson, B., Lukens, W. E., El Masri, B. and Stinchcomb, G. E., Alluvial landform and the occurrence of paleosols in a humid–subtropical climate have an effect on long-term soil organic carbon storage. *Geoderma*, 2020, **371**, 114388.

44. Department of Agriculture, Agriculture contingency plan for district RiBhoi. Directorate of Agriculture, Government of Meghalaya. 2014, pp. 1–3.
45. Sehgal, J. L., Sys, C., Stoops, G. and Tavernier, R., Morphology, genesis and classification of two dominant soils of the warm temperate and humid region of the central Himalayas. *J. Indian Soc. Soil Sci.*, 1985, **33**(4), 846–857.
46. Chen, J. and Zhu, W., Comparing Landsat-8 and Sentinel-2 top of atmosphere and surface reflectance in high latitude regions: case study in Alaska. *Geocarto Int.*, 2021, **37**(9), 1–20; doi:10.1080/10106049.2021.1924295.
47. Mondal, A., Khare, D., Kundu, S., Mondal, S., Mukherjee, S. and Mukhopadhyay, A., Spatial soil organic carbon (SOC) prediction by regression kriging using remote sensing data. *Egypt. J. Remote Sensing Space Sci.*, 2017, **20**(1), 61–70.
48. Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R. and Evangelista, P. H., A survival guide to Landsat preprocessing. *Ecology*, 2017, **98**(4), 920–932.
49. Thaler, E. A., Larsen, I. J. and Yu, Q., A new index for remote sensing of soil organic carbon based solely on visible wavelengths. *Soil Sci. Soc. Am. J.*, 2019, **83**(5), 1443–1450.
50. Singh, R. P. *et al.*, Retrieval of wheat leaf area index using price approach based on inversion of canopy reflectance model. *J. Indian Soc. Remote Sensing*, 2005, **33**(2), 307–313.
51. Ray, S. S., Sood, A., Panigrahy, S. and Parihar, J. S., Derivation of indices using remote sensing data to evaluate cropping systems. *J. Indian Soc. Remote Sensing*, 2005, **33**(4), 475.
52. Kumari, M. and Sarma, K., Changing trends of land surface temperature in relation to land use/cover around thermal power plant in Singrauli district, Madhya Pradesh, India. *Spat. Inf. Res.*, 2017, **25**(6), 769–777.
53. Nageswara, P. P. R., Shobha, S. V., Ramesh, K. S. and Somashekhar, R. K., Satellite-based assessment of agricultural drought in Karnataka state. *J. Indian Soc. Remote Sensing*, 2005. **33**(3), 429–434.
54. Gbolo, P., Gerla, P. J. and Vandeberg, G. S., Using high resolution, multispectral imagery to assess the effect of soil properties on vegetation reflectance at an abandoned feedlot. *Geocarto Int.*, 2015, **30**(7), 1–17.
55. Rouse, J. W., Haas, R. H., Schelle, J. A., Deering, D. W. and Harlan, J. C., Monitoring the vernal advancement or retrogradation of natural vegetation. NASA/GSFC, Type III, Final Report, NASA, Greenbelt, MD, USA, 1974, p. 371.
56. Tucker, C. J., Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing Environ.*, 1979, **8**(2), 127–150.
57. Lillesaeter, E., Spectral reflectance of partly transmitting leaves: laboratory measurements and mathematical modeling. *Remote Sensing Environ.*, 1982, **12**, 247–254.
58. Holben, B. N., Kaufman, Y. J. and Kendall, J. D., NOAA-11 AVHRR visible and near-IR in flight calibration. *Int. J. Remote Sensing*, 1990, **11**(8), 1511–1519.
59. Holben, B. and Fraser, R. S., Red and near-infrared sensor response to off-nadiir viewing. *Int. J. Remote Sensing*, 1984, **5**(1), 145–160.
60. Huete, A. R., A soil-adjusted vegetation index (SAVI). *Remote Sensing Environ.*, 1988, **25**, 295–309.
61. Pye, K. and Tsoar, H., *Aeolian Sand and Sand Dunes*, Springer, Berlin, Germany, 2009, p. 458.
62. Mayhew, S. and Penny, A., *The Concise Oxford Dictionary of Geography*, University Press, Oxford, UK, 1992, p. 250.
63. Geo University, Spectral indices with multispectral satellite data; https://www.geo.university›pages›blo (accessed on 18 December 2022).
64. Bartholomeus, H., Schaepman, M., Kooistra, L., Stevens, A., Hoogmoed, W. and Spaargaren, O., Spectral reflectance based indices for soil organic carbon quantification. *Geoderma*, 2008, **145**, 28–36; doi:10.1016/j.geoderma.2008.01.010.
65. Zhou, W., Han, G., Liu, M. and Li, X., Effects of soil pH and texture on soil carbon and nitrogen in soil profiles under different land uses in Mun River Basin, Northeast Thailand. *Peer J.*, 2019, **7**, e7880.
66. Page, K. L., Dang, Y. P. and Dalal, R. C., The ability of conservation agriculture to conserve soil organic carbon and the subsequent impact on soil physical, chemical, and biological properties and yield. *Front. Sustain. Food Syst.*, 2020, **4**, 31.
67. Nelson, D. W. and Sommers, L. E., Total carbon, organic carbon, and organic matter. In *Methods of Soil Analysis: Part 3 – Chemical Methods*, SSA Book Series, Madison, 1996, vol. 5, pp. 961–1010.
68. He, T., Wang, J., Lin, Z. and Cheng, Y., Spectral features of soil organic matter. *Geo-spat. Inf. Sci.*, 2009, **12**(1), 33–40.
69. Lin, L., Wang, Y., Teng, J. and Wang, X., Hyperspectral analysis of soil organic matter in coal mining regions using wavelets, correlations, and partial least squares regression. *Environ. Monit. Assess.*, 2016, **188**(2), 1–11.
70. Jaber, S. M., Lant, C. L. and Al-Qinna, M. I., Estimating spatial variations in soil organic carbon using satellite hyperspectral data and map algebra. *Int. J. Remote Sensing*, 2012, **32**(18), 5077–5103.
71. Somvanshi, S. S., Kunwar, P., Tomar, S. and Singh, M., Comparative statistical analysis of the quality of image enhancement techniques. *Int. J. Image Data Fusion*, 2018, **9**(2), 131–151.
72. Rasool, S. N., Gaikwad, S. W. and Talat, M. A., Relationship between soil properties and slope segments of Sallar Wullarhama watershed in the Liddar catchment of Jammu and Kashmir. *Asian J. Eng. Res.*, 2014, **2**(2), 1–10.
73. Jaksic, S., Ninkov, J., Milic, S., Vasin, J., Zivanov, M., Jaksic, D. and Komlen, V., Influence of slope gradient and aspect on soil organic carbon content in the region of Nis, Serbia. *Sustainability*, 2021, **13**(15), 8332.
74. Wilding, L. P., Spatial variability: its documentation, accommodation and implication to soil surveys. In *Soil Spatial Variability.* Proceedings of a Workshop of the ISSS and the SSA (eds Nielsen, D. R. and Bouma, J.), Las Vegas NV, Wageningen, USA, 30 November–1 December 1985, pp. 166–187.
75. Hamzehpour, N., Shafizadeh-Moghadam, H. and Valavi, R., Exploring the driving forces and digital mapping of soil organic carbon using remote sensing and soil texture. *Catena*, 2019, **182**, 104141.
76. Wang, S., Zhuang, Q., Jin, X., Yang, Z. and Liu, H., Predicting soil organic carbon and soil nitrogen stocks in topsoil of forest ecosystems in Northeastern China using remote sensing data. *Remote Sensing*, 2020, **12**(7), 1115.
77. Singh, N. J., Kudrat, M. and Jain, K., Effect of land use and topography on spatial distribution of soil organic carbon in semi-arid subtropical ecosystems in Uttar Pradesh, India. *Int. J. Ecol. Environ. Sci.*, 2014, **40**, 189–197.
78. Challam, L. L., Singh, N. J., Debbarma, K., Ray, L. I. and Swami, S., Effect of topographical settings on distribution of soil organic carbon fractions in rice ecosystem of North East India. *Int. J. Ecol. Environ. Sci.*, 2016, **42**(5), 129–135.
79. Viscarra Rossel, R. V. and Behrens, T., Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 2010, **158**(1–2), 46–54.
80. Castaldi, F., Palombo, A., Santini, F., Pascucci, S., Pignatti, S. and Casa, R., Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. *Remote Sensing Environ.*, 2016, **179**, 54–65.
81. Son, N. T., Chen, C. F., Chen, C. R., Minh, V. Q. and Trung, N. H., A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agric. For. Meteorol.*, 2014, **197**, 52–64.
82. Nuarsa, I. W., Nishio, F. and Hongo, C., Relationship between rice spectral and rice yield using MODIS data. *J. Agric. Sci.*, 2011, **3**(2), 80.
83. Soil and Water Conservation Department, Integrated Watershed Management Project, Detailed project report, Govt of Meghalaya,

Umsning C & RD block, 2009; https://megsoil.gov.in›iwmp›2009-10
84. IMP, Rainfall statistics of India, India Meteorological Department, Report no. MoES/IMD/HS/RAINFALL REPORT/30, 2019.
85. Pulak, G., Bhagwat, P. P., Satpute, U. S., Menon P., Prasad, A. K., Sable, S. T. and Advani, S. C., Observed rainfall variability and changes over Meghalaya state. India Meteorological Department, Pune.
86. Arunrat, N., Pumijumnong, N. and Hatano, R., Predicting local-scale impact of climate change on rice yield and soil organic carbon sequestration: a case study in RoiEt Province, Northeast Thailand. *Agric. Syst.*, 2018, **164**, 58–70.
87. Laxminarayana, K. and Bharali, S., Distribution of inorganic N fractions and N availability indices in the rice soils of Meghalaya. *ORYZA – Int. J. Rice*, 2010, **47**(2), 128–135.
88. Gabarrón-Galeote, M. A., Trigalet, S. and van Wesemael, B., Effect of land abandonment on soil organic carbon fractions along a Mediterranean precipitation gradient. *Geoderma*, 2015, **249**, 69–78.
89. Rouse Jr, J. W., Haas, R. H., Schell, J. A. and Deering, D. W., Monitoring vegetation systems in the Great Plains with ERTS. NASA Special Publication, Greenbelt, MD, USA, NASA Goddard Space Flight Center, 1974, vol. 351, p. 309.
90. Escadafal, R., Belghith, A. and Ben-Moussa, H., Indices spectraux pour la télédétction de la dégradation des milieux naturels en Tunisie aride. In Actes du 6eme Symp. Int. sur les mesures physiques et signatures en teledetection, Val d'Isère (France), 17–24 January 1994, pp. 253–259.
91. Roujean, J. L. and Breon, F. M., Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing Environ.*, 1995, **51**(3), 375–384.
92. Mokarram, M., Roshan, G. and Negahban, S., Landform classification using topography position index (case study: salt dome of Korsia-Darab plain, Iran). *Model. Earth Syst. Environ.*, 2015, **1**(4), 1–7.
93. Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H. and Sorooshian, S., A modified soil adjusted vegetation index. *Remote Sensing Environ.*, 1994, **48**(2), 119–126.
94. Jiang, Z., Huete, A. R., Kim, Y. and Didan, K., 2-band enhanced vegetation index without a blue band and its application to AVHRR data. In *Remote Sensing and Modeling of Ecosystems for Sustainability IV,* SPIE, 2007, vol. 6679, pp. 45–53.
95. Madeira, J., Bedidi, A., Cervelle, B., Pouget, M. and Flay, N., Visible spectrometric indices of hematite (Hm) and goethite (Gt) content in lateritic soils: the application of a Thematic Mapper (TM) image for soil-mapping in Brasilia, Brazil. *Int. J. Remote Sensing*, 1997, **18**(13), 2835–2852.
96. Pouget, J. P., Jozefowicz, M. E., Epstein, A., Tang, X. and MacDiarmid, A. G., X-ray structure of polyaniline. *Macromolecules*, 1991, **24**(3), 779–789.
97. Jamalabad, M. and Abkar, A., Forest canopy density monitoring using satellite images. In 20th ISPRS Congress on International Society for Photogrammetry and Remote Sensing, Istanbul, 2004, Turkey, pp. 12–23.
98. Deng, Y., Wu, C., Li, M. and Chen, R., RNDSI: a ratio normalized difference soil index for remote sensing of urban/suburban environments. *Int. J. Appl. Earth Obs. Geoinf.*, 2015, **39**, 40–48.
99. Rogers, A. S. and Kearney, M. S., Reducing signature variability in unmixing coastal marsh Thematic Mapper scenes using spectral indices. *Int. J. Remote Sensing*, 2004, **25**(12), 2317–2335.
100. McFeeters, S. K., The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sensing*, 1996, **17**(7), 1425–1432.

ACKNOWLEDGEMENTS. We thank the authorities of College of Post Graduate Studies in Agricultural Sciences, Umiam and Central Agricultural University, Imphal for extending the required permission and providing the necessary support to undertake this study. We also thank the anonymous reviewers for their valuable suggestions that helped improve the manuscript.

Received 16 August 2022; re-revised accepted 31 January 2023

doi: 10.18520/cs/v124/i12/1431-1444