PLANT SCIENCE TODAY
Vol 13(sp5): 01–10
https://doi.org/10.14719/pst.11182

eISSN 2348-1900

RESEARCH ARTICLE

Estimation of soil organic carbon under rice-fallow system using Sentinel-2 derivatives

Naorem Janaki Singh[1*], Priya Das[1], Deepak Kumar Dwivedi[1], Rajkumar Josmee Singh[2] & Basant Tamang[1]

[1]School of Natural Resource Management, College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University (Imphal), Umiam, Ri Bhoi, Shillong 793 103, Meghalaya, India

[2]School of Social Sciences, College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University (Imphal), Umiam, Ri Bhoi, Shillong 793 103, Meghalaya, India

*Correspondence email - naorem2005@gmail.com

Received: 07 August 2025; Accepted: 22 June 2026; Available online: Version 1.0: 28 September 2026

Cite this article: Naorem JS, Priya D, Deepak KD, Rajkumar JS, Basant T. Estimation of soil organic carbon under rice-fallow system using Sentinel-2 derivatives. Plant Science Today. 2026;13(sp5):01–10. https://doi.org/10.14719/pst.11182

Abstract

The evaluation of soil properties under a century-old traditional rice-fallow practice in highly fragile topography (i.e., hill slopes and lowlands) of North-east (NE) India in the Eastern Himalaya faces multiple challenges in conventional studies of soil properties. Alternatively, remote sensing offers a viable solution. An attempt was made to study the soil organic carbon (SOC) under century old traditional rice-fallow practice in highly fragile topography (i.e., hill slopes and lowlands) of Bhoirymbong, Meghalaya, NE India in Eastern Himalaya using remote sensing indices derived from Sentinel-2. The objective was to develop a suitable regression model for rice-fallow systems across different hill slopes in northeastern India in the Eastern Himalaya. A total of 100 composite soil samples from the 0–15 cm depth were collected from rice-fallow areas on various slopes (nearly level: 0–3 %, gentle: 3–8 %, moderate: 8–15 %, steep: 15–30 % and very steep: > 30 %) during November 2020. Each composite sample was prepared from 10 randomly collected soil samples within a 10 × 10 m area and analysed for SOC using standard protocols in the analytical laboratory of School of Natural Resource Management, College of Post Graduate Studies in Agricultural Sciences, Central Agricultural University, Imphal. Nearly synchronised, cloud-free (< 10 % cloud cover) Sentinel-2 data consistent with soil sampling period were accessed from the US Geological Survey (USGS) official website to derive 15 indices. Results showed that brightness index (BI), BI2 and the soil bareness index (SBI) were the most influential indices for SOC. The highest and lowest SOC values were found in gently undulating topography (2–5 % slope) and rolling topography (10–15 % slope), at 2.86 % and 1.17 %, respectively. After rice crop harvesting, the red-NIR-based indices did not show a significant difference, whereas the green-NIR-based indices were significantly different at $p < 0.05$. Hence, it is concluded that green-NIR-based indices are the most suitable for SOC prediction in NE India in the Eastern Himalaya.

Keywords: regression model; remote sensing indices; slope; soil organic carbon

Introduction

The Khasi tribe dwell in the fertile lowlands known as Bhois, who settled in the north-west of Khasi Hills under a kingdom known as the Nongpoh Syiemship between 1500 and 1600 A.D. The Bhois are the natives of the Ri Bhoi district, according to Rev. (Fr.) S. Sngi Lyngdoh. This Syiemship was relegated to the status of a Sardarship in the seventies of the 19th century i.e., probably in the years 1863—64 and thus came under the aegis of the Sixth Schedule to the Constitution of India (1). The livelihood of the tribal community, traditionally and customarily depends on rice as the staple food and the cultivation of the rice-fallow system under natural conditions without any addition of organic or inorganic plant nutritions. The rice-fallow system covers about 104800 hectares (ha) (37.84 % of the total cropped area) in Meghalaya, with an annual average production of 135000.50 metric tonnes (2). The average yield is 1290.6 kg ha^{-1}, which is 54 % lower than the national productivity (2390 kg ha^{-1}) (3). The nutrients reserved in the Pinus forest are the only source for nutrient supply to the crop. A long period (> 200 years) of rice–fallow in the traditional farming (i.e., natural farming) in difficult terrain may affect the soil organic carbon (SOC) and nutrient content of the soil. The SOC is affected by many factors, including climate (cold and wet), rates of chemical reaction (i.e., decomposition rate), topography, vegetation and management (4). Topography influences climate and the spatial and vertical distribution of SOC as well as organic matter (OM) decomposition (5). Mismanagement of land leads to erosion, acidification, pollution and loss of soil fertility (6). The SOC is a key indicator of soil health and also provides nutrients to plants, improves soil structure and water storage capacity (7—9). Hence, a precise estimation of SOC will be useful in understanding the influence of agro-ecosystem on the production of the rice-fallow system. However, conventional estimation in the laboratory is labour-intensive, expensive and time-consuming (10). The regular measurement of the SOC of hilly terrain with limited facilities is challenging. Alternatively, remote sensing is an option for SOC estimation and also has the advantage of covering large areas, providing real-time data and is possible for continuous surface information. This study estimates SOC from the combination of crop statistics, forest inventories, soil geographic databases and remote

sensing derivatives (11–12). Remote sensing provides a non-destructive, low-cost approach for monitoring spatio-temporal changes in SOC over large areas on a repetitive basis and also provides important environmental variables for determining SOC variation with changing topography and climate (13). Researchers have attempted to quantify SOC using remote sensing-derived indices such as normalised difference vegetation index (NDVI), modified soil-adjusted vegetation index (MSAVI), bare soil index (BSI), brightness index (BI), soil adjusted vegetation index (SAVI) and renormalised difference vegetation index (RDVI). Statistical models between SOC and remote sensing indices were established to obtain SOC (14–15). The indices derived from Sentinel-2 are very useful for soil properties, land management and agriculture studies (16–17). Considering these advantages, this study aimed to map the SOC of rice-fallow with the objectives of deriving different indices from Sentinel-2 data and developing SOC prediction models using remote sensing derivatives and the assumption that the surrounding pine forest at different slopes restores depleted SOC through crop-soil management.

Materials and Methods

Study area

Bhoirymbong, Ri Bhoi district, Meghalaya is the best representative of different topographical settings including lowland slopes (nearly level: 0–3 %), foothills (gentle slope: 3–8 %), mid-hills (moderate slope: 8–15 %; strong slope: 15–30 %) and steep slopes (> 30 %). The rice-fallow system is practised in all the slope positions and the Government of India has also declared Ri Bhoi an Aspirational district (Fig. 1 and 2).

The study area is located between 92°2'0"–92°5'30" E longitude and 25°43'–25°45' N latitude at an altitude of 1010 m above mean sea level. The soils are sandy loam to loam, with

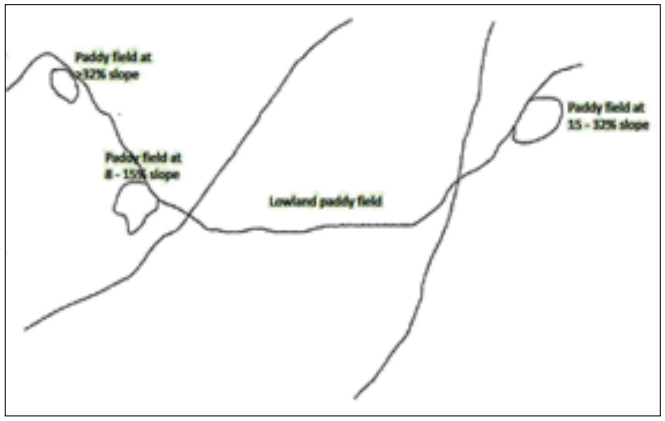


Fig. 1. Physiography of study site.

Source: QGIS, version 3.40.

shallow to medium depth, excessive drainage and extremely acidic. The soil taxonomic class is Typic Kandihumults (13). The average annual temperature ranges from 10 °C in January to 32 °C in the month of July and August. The monthly rainfall ranges from 0.9 cm in January to 23.9 cm in July (18).

Soil sampling and processing

A total of 100 composite soil samples from the 0–15 cm depth were collected from rice-fallow system at different slopes; lowland slopes (nearly level: 0–3 %), foothills (gentle slope: 3–8 %), mid-hills (moderate slope: 8–15 %, strong slopes: 15–30 %) and steep slopes (> 30 %) during 10–15 November, 2020. Each composite soil sample was prepared from 10 randomly collected soil samples within a 10 × 10 m area. The soil samples were air-dried, crushed and ground to pass through a 0.5 mm sieve for the determination of oxidisable organic carbon (19). The coordinates of the sampling locations were recorded using a handheld Global Positioning System (GPS) unit (GERMIN GPSMAP-64) with an accuracy of ± 3 m. About 75 % of the samples were used for SOC prediction and the remaining 25 % were

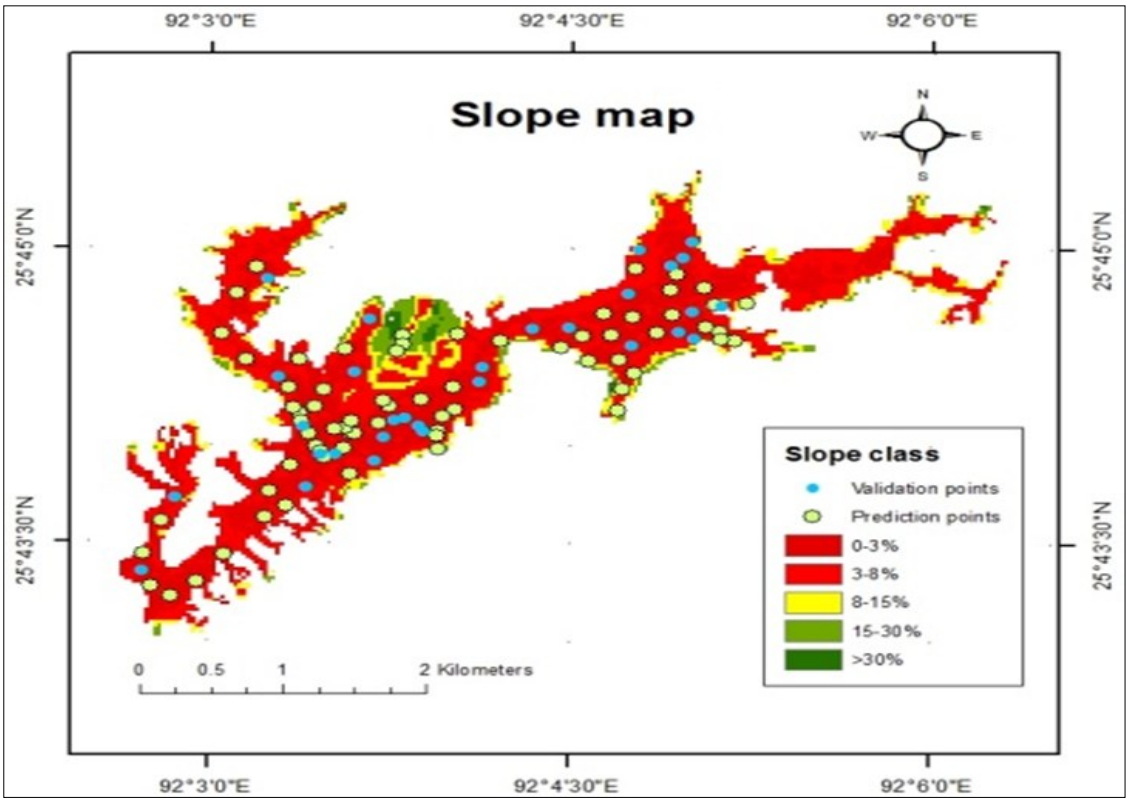


Fig. 2. Slope map of the study area generated from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) data.

Source: QGIS, version 3.40.

used for model validation.

Satellite image acquisition and processing

A cloud-free (< 10 % cloud cover) Sentinel-2 Multi Spectral Instrument (MSI) Level-2A product, acquired on 7th November, 2020, orthorectified in the UTM/WGS84 projection, with a 5-days revisit time and a 290 km swath, was downloaded from the ESA Copernicus Open Access Hub (formerly Sentinels Scientific Data Hub) (20). Sentinel-2 Level-2A (L2A) products provide Bottom-of-Atmosphere (BOA) reflectance, transformed from Level-1C Top-of-Atmosphere (TOA) data using the Sen2Cor processor for atmospheric correction. These images are fully orthorectified in the UTM/WGS84 projection, 100 × 100 km tiles. Atmospheric correction reduces the effects of aerosol scattering and absorption, ozone and water vapour, with outputs including reflectance, Aerosol Optical Thickness (AOT), Water Vapour (WV) and Scene Classification (SCL) maps. Fifteen different indices were derived for developing regression models of SOC estimation (Table 1). The bands B2 (490 nm; blue), B3 (560 nm; green), B4 (665 nm; red) and B8 (842 nm; near-infrared) have a 10 m spatial resolution. The indices were calculated using the raster calculator and extracted to each sampling point using the multi values to points tool in spatial analyst. The extracted values were stored in the attribute table and subsequently exported to an Excel (.xls) file using the conversion tool in ArcGIS 10.3. Slopes were derived from freely accessible 30 m Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) projected to WGS84 and EGM96, downloaded from the official USGS EarthExplorer website (21). The slopes were classified into five slope classes [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 (Fig. 2).

Statistical analysis

Pearson's correlation coefficient (r) between SOC and the fifteen indices were calculated to generate a correlation matrix. The strength of the correlation was classified as weak when $0 < |r| < 0.3$, moderate when $0.3 < |r| < 0.7$ and strong when $0.7 < |r| < 1.0$. Regression analyses namely multiple linear regression-stepwise (MLR-stepwise) and partial least squares regression (PLSR) were performed. In MLR-stepwise, the indices were used as predictor variables and tested for their significance, while non-significant variables were removed. This process was repeated until no variable could be selected or rejected (22). Partial least squares regression was performed for datasets with multicollinearity, in which the dependent and independent variables were transformed into a new set of latent variables before ordinary linear regression was applied. Principal component analysis regression (PCA-R) is a combination of principal component analysis and linear regression. It is based on dimensionality reduction of highly correlated variables into principal components (PCs), which are determined by the highest eigenvector values and their corresponding eigenvectors. The performance of the regression models was evaluated by calculating the coefficient of determination (R^2) and the root mean square error (RMSE). The best-performing regression model was used for SOC prediction and mapping using the raster calculator tool in ArcGIS 10.5. One-way analysis of variance (ANOVA) was performed to test the significance of the predicted SOC among different slope classes using Gabriel's post-hoc test ($p < 0.05$), as the number of observations within each factor was not equal. All the statistical analyses were performed using XLSTAT (2022.1), Microsoft Excel and IBM SPSS Statistics version 21.0.

Model validation

The laboratory-analysed SOC (actual SOC) was cross-validated with the model-estimated SOC using R^2 and RMSE. The model with the lowest RMSE and highest R^2 values was considered the best and selected (15, 23). A lower RMSE value indicates that the observed data are closer to the line of best fit, thereby indicating a lower prediction error (24).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}i - yi)^2}{\sum_{i=1}^{n}(yi - \bar{y})^2};\ \mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - \hat{y}i)^2}$$

Where yi is the actual SOC, ŷi is the predicted SOC, $\bar{y}$ is the mean of the actual SOC and n is the number of samples. The schematic diagram of the detail work plan is presented in Fig. 3.

Results and Discussion

Slope effect on remote sensing indices

The statistical description of the remote sensing (RS) indices is presented in Table 2. The mean values of the indices are almost similar to the median values, indicating a normal distribution of data, which is mandatory for regression analysis. The SOC values of the study area varied from 1.17–2.88, with a mean of 2.22 and a standard deviation of 0.27. The highest and lowest SOC values were observed in gently undulating topography (2–5 % slope) and rolling topography (10–15 % slope), at 2.86 % and 1.17 %, respectively. The crop health indices namely NDVI, GNDVI, GRVI, EVI, WDVI and RDVI varied from 0.28–0.89 with a mean of 0.54, 0.40–0.77 with a mean of 0.53, 0.14–0.37 with a mean of 0.02, 0.52–2.78 with a mean of 1.23, 15–1156 with a mean of 415.10 and 16.99–49.31 with a mean of 31.78, respectively. The soil brightness indices such as BI, BI2 and soil bareness index (SBI) varied from 200.24–922.10, 135.81–615.51 and 1995.80–44490.6 with mean ± standard deviation values of 586.53 ± 118.24, 391.82 ± 78.68 and 2839.80 ± 406.38, respectively. Theoretically, the NDVI value ranges from -1 to +1; values from -1 to 0 indicate no vegetation to poor vegetation whereas +1 indicates very good vegetation. The surrounding forest of rice-fallow system at steep slope (> 30 % slope) resulted in the high NDVI value of 0.89 (Fig. 4), BI of 381.7, BI2 of 856.5 and SBI of 2742.5 (Table 3). Most of the rice-fallow system on steep slopes is less than 100 m^2 i.e., smaller than the 10 × 10 m spatial resolution of Sentinel-2. Therefore, spectral mixing of energy occurs in these small areas. The areas located on nearly level (0–3 %) land and foothills (gentle slope, 3–8 %) had an NDVI value of approximately 0.28 after rice harvesting. The moisture absorbed by the standing stubbles after harvesting at these slope positions resulted in lower NDVI values compared with those of the surrounding pine forest at the mid hills (moderate slope: 8–15 %, strong slope: 15–30 %) and steep slopes. The SBI value ranged widely from 1995.80–44490.6, with a mean of 2839.80, attributable to the wide variation in soil conditions, from exposed soil after crop harvesting to native forest. The SAVI and MSAVI minimise the effect of bare soil on vegetation health indices, i.e., NDVI. The SAVI and MSAVI2 after rice harvesting varied from 0.42–1.33 with a mean of 0.80 and from 0.88–1.88 with a mean of 1.39, respectively. After rice harvesting, the red-NIR-based indices were not significantly different and green-NIR-based indices were significantly different at $p < 0.05$. The absorption of the red spectrum and the reflection of NIR after crop harvesting were comparatively

Table 1. Various remote sensing indices used for development of soil organic carbon (SOC) prediction model

Sl. No.	Index	Definition
1	Normalised difference vegetation index (NDVI)	$\frac{NIR - RED}{NIR + RED}$
2	Green normalised difference vegetation index (GNDVI)	$\frac{NIR - GREEN}{NIR + GREEN}$
3	Soil adjusted vegetation index (SAVI)	$(1 + L) * \frac{NIR - RED}{NIR - RED + L}$ L= soil brightness correction factor (L=0.5)
4	Green red vegetation index (GRVI)	$\frac{GREN - RED}{GREEN + RED}$
5	Enhanced vegetation index (EVI)	$G * \frac{NIR - RED}{NIR + (\ _1 * RED - C_2 * BLUE) + L}$ C_1 and C_2= aerosol resistance coefficients; G= gain factor; L=canopy background adjustment Here, L=1, C1=6, C2=7.5 and G=2.5
6	Modified soil adjusted vegetation index (MSAVI)	$(1 + L) * \frac{NIR - RED}{NIR - RED + L}$ L= 1 – 2 * a * NDVI * WDVI $a = \frac{NIR}{RED}$
7	Weighted difference vegetation index (WDVI)	$WDVI = NIR - C * RED$ $C = \frac{NIR}{RED}$ C= soil line of bare soil
8	Brightness index (BI)	$\frac{\sqrt{(RED * RED) + (GREEN * GREEN)}}{2}$
9	Second brightness index (BI2)	$\frac{\sqrt{(RED * RED) + (GREEN * GREEN) + (NIR * NIR)}}{3}$
10	Redness index (RI)	$\frac{RED * RED}{GREEN * GREEN * GREEN}$
11	Vegetation index (V)	$\frac{NIR}{RED}$
12	Soil brightness index (SBI)	$\sqrt{(RED)^2 + (NIR)^2}$
13	Renormalised difference vegetation index (RDVI)	$\frac{NIR - RED}{\sqrt{NIR + RED}}$
14	Second modified soil adjusted vegetation ir. x (MSAVI2)	$\frac{2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)}}{2}$
15	Colour index (CI)	$\frac{RED - GREEN}{RED + GREEN}$

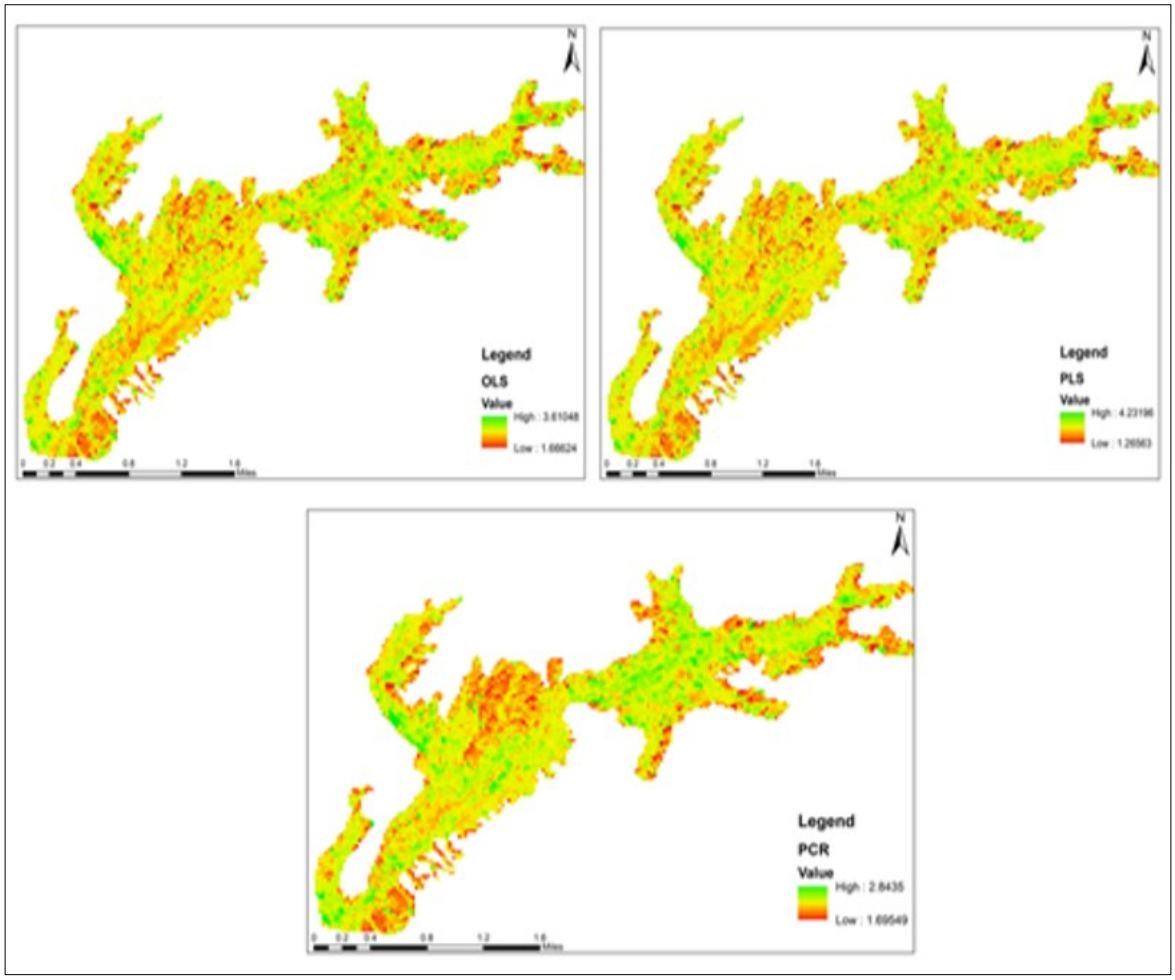


Fig. 3. The soil organic carbon map created using the predicted soil organic carbon of ordinary least square regression, partial least square regression and principal competent regression.

OLS: ordinary least square regression; PLSR: partial least square regression; PCR: principal competent regression.

Source: QGIS, version 3.40.

Table 2. The statistics of Sentinel-2 data derived indices

Indices	Minimum	Maximum	Mean ± Std. deviation	Median	Skewness	Kurtosis
NDVI	0.28	0.89	0.54 ± 0.09	0.52	0.98	6.86
GNDVI	0.40	0.77	0.53 ± 0.06	0.52	1.45	7.42
SAVI	0.42	1.33	0.80 ± 0.13	0.79	0.98	6.86
GRVI	0.14	0.37	0.02 ± 0.07	0.01	1.73	9.04
EVI	0.52	2.78	1.23 ± 0.33	1.16	2.04	9.72
MSAVI	0.42	1.33	0.80 ± 0.13	0.79	0.98	6.86
WDVI	15.00	1156.00	415.10 ± 277.19	381.00	0.37	2.13
BI	200.24	922.10	586.53 ± 118.24	577.86	0.32	4.98
BI2	135.81	615.51	391.82 ± 78.68	385.98	0.33	4.97
V	1.00	16.00	3.08 ± 1.77	3.00	4.84	32.48
SBI	1995.80	4490.60	2839.80 ± 406.38	2793.20	1.00	4.98
RDVI	16.99	49.31	31.78 ± 5.59	30.85	0.67	4.54
MSAVI2	0.88	1.88	1.39 ± 0.14	1.38	0.13	6.28
CI	0.37	0.14	-0.02 ± 0.07	-0.01	-1.73	9.04

NDVI: normalised difference vegetation index; GNDVI: green normalised difference vegetation index; SAVI: soil adjusted vegetation index; GRVI: green red vegetation index; EVI: enhanced vegetation index; MSAVI: modified soil adjusted vegetation index; WDVI: weighted difference vegetation index; BI: brightness index; BI2: second brightness index; V: vegetation index; SBI: soil brightness index; RDVI: renormalised difference vegetation index; MSAVI2: second modified soil adjusted vegetation index; CI: colour index.

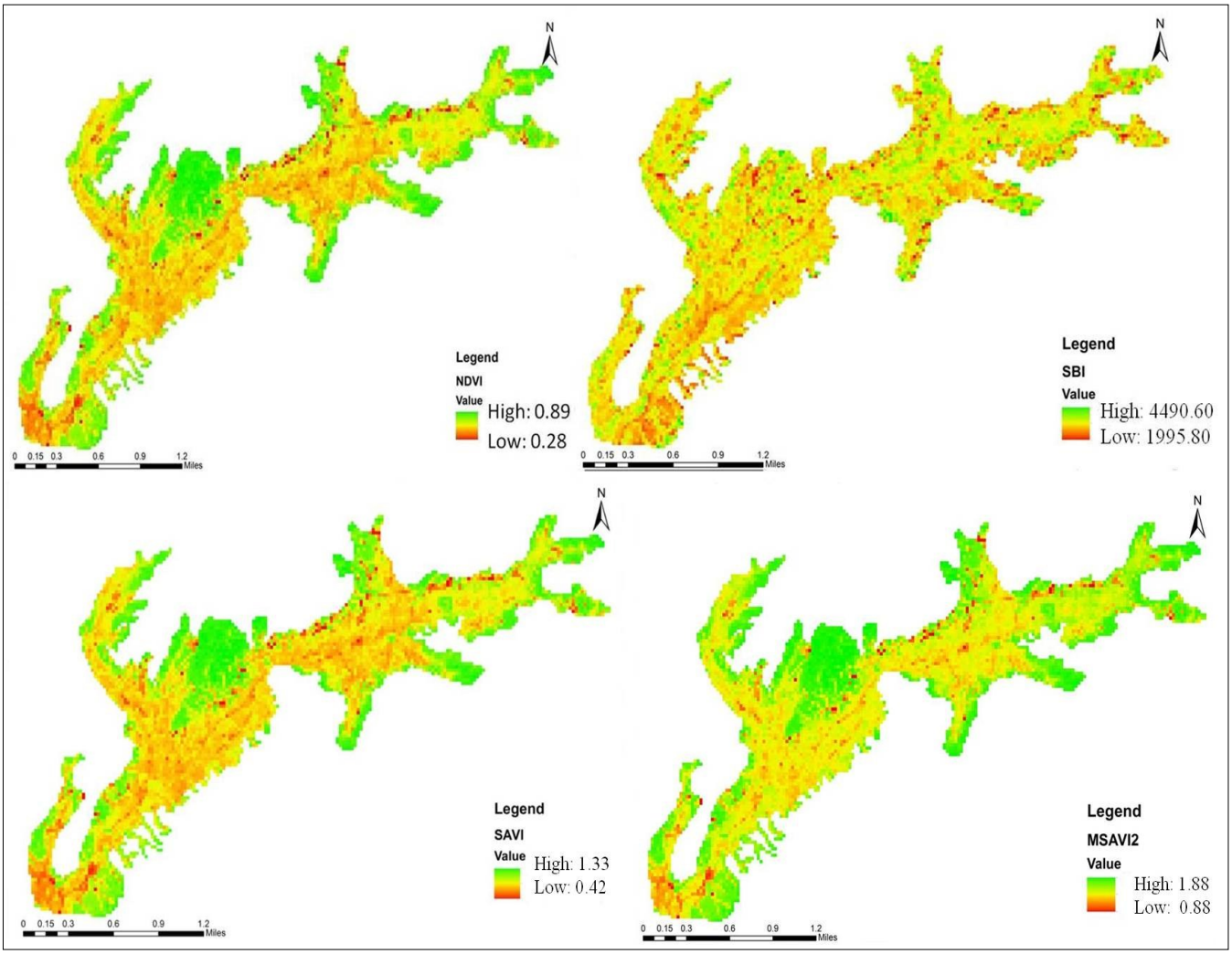


Fig. 4. The rice-fallow system, vegetations and topographical effect on various remote sensing indices.

NDVI: normalised difference vegetation index; SBI: soil bareness index; SAVI: soil adjusted vegetation index; MSAVI: modified soil-adjusted vegetation index.

Table 3. Effect of slope on soil organic carbon and remote sensing (RS) indices

Slope (%)	SOC (%)	NDVI	GNDVI	SAVI	GRVI	EVI	MSAVI	WDVI	BI	BI2	SBI	RDVI	MSAVI2
Nearly level (0—3 %)	2.19	0.52	0.52	0.78	0.01	1.18	0.78	435.8	602.6	402.5	2815.5	30.88	1.36
Foot hill (3—8 %)	2.13	0.58	0.56	0.87	0.04	1.39	0.87	366.2	543.9	363.5	2918.1	34.56	1.46
Mid hills (8—15 % & strong slope, 15—30 %)	1.99	0.68	0.60	1.02	0.16	1.95	1.02	273.9	426	690.3	2721.3	37.98	1.60
Steep slope (> 30 %)	1.86	0.72	0.63	1.09	0.21	2.07	1.09	241.2	381.7	856.5	2742.5	40.36	1.67

SOC: soil organic carbon; NDVI: normalised difference vegetation index; GNDVI: green normalised difference vegetation index; SAVI: soil adjusted vegetation index; GRVI: green red vegetation index; EVI: enhanced vegetation index; MSAVI: modified soil adjusted vegetation index; WDVI: weighted difference vegetation index; BI: brightness index; BI2: second brightness index; SBI: soil brightness index; RDVI: renormalised difference vegetation index; MSAVI2: second modified soil adjusted vegetation index.

low whereas there was a wide difference between the reflection of green and NIR i.e., GNDVI (0.40–0.77, with a mean of 0.53) and GRVI (0.14–0.37, with a mean of 0.02), resulting in significance at $p < 0.05$.

Prediction of soil organic carbon using remote sensing indices

The SOC content showed a significant negative correlation with slope (r = -0.303) and RS indices such as NDVI (r = -0.01), GNDVI (r = -0.09), SAVI (r = -0.01), EVI (r = -0.07), MSAVI (r = -0.01), V (r = -0.11), MSAVI2 (r = -0.004) and CI (r = -0.03) (Table 4). This may be due to the exposure of land after rice harvesting. At these locations, which are normally situated on strong to steep slopes, direct seeding is practised and plants grow sparsely; consequently, the stubble remaining after harvesting is insufficient to cover the soil. The indices, namely BI (r = 0.28), BI2 (r = 0.26), RI (r = -0.27) and SBI (r = 0.33) were significantly correlated with SOC at $p \leq 0.01$, whereas the other indices were poorly correlated with SOC. Previous study showed that, the coarse-sand (2000–200 µm) and clay (< 2 µm) fractions increased with prolonged rice cultivation, whereas the aggregate size fractions were dominated by fine-sand (200–20 µm) and silt (20–2 µm) fractions across the chronosequence (25). The SOC was highly enriched in the coarse-sand fractions (40–60 g kg^{-1}) and moderately enriched in the clay fractions (20–25 g kg^{-1}) but was depleted in the silt fractions (10 g kg^{-1}). This is associated with the high reflection of the red wavelength and low NIR reflection at crop harvest, while humus, clay and moisture content also affect spectral reflectance. The soil of the study site belongs to the Typic

Table 4. Pearson correlation among the soil organic carbon (SOC) and remote sensing indices

	OC (%)	NDVI	GNDVI	SAVI	GRVI	EVI	MSAVI	WDVI	BI	BI2	RI	V	SBI	RDVI	MSAVI2	CI
OC (%)	1															
NDVI	-0.01	1														
GNDVI	-0.09	0.92**	1													
SAVI	-0.01	1.00**	0.92**	1												
GRVI	0.03	0.90**	0.69**	0.90**	1											
EVI	-0.07	0.96*	0.87**	0.96**	0.92**	1										
MSAVI	-0.01	1.00**	0.92**	1.00**	0.90**	0.96**	1									
WDVI	0.1	-0.36**	-0.31**	-0.36**	-0.32**	-0.32**	0.36**	1								
BI	0.26**	-0.65**	-0.74**	-0.65**	-0.46**	-0.69**	-0.65**	0.33**	1							
BI2	0.26**	-0.65**	-0.74**	-0.65**	-0.46**	-0.69**	-0.65**	0.33**	1.00**	1						
RI	-0.27**	-0.43*	-0.12	-0.43**	-0.64**	-0.38**	-0.43**	0.13	-0.30**	-0.30**	1					
V	-0.11	0.85**	0.83**	0.85**	0.85**	0.88**	0.85**	-0.41**	-0.59**	-0.59**	-0.34**	1				
SBI	0.33**	0.26**	0.12	0.26**	0.33**	0.15	0.26**	0.06	0.54**	0.54**	-0.75**	0.16	1			
RDVI	0.13	0.92**	0.79**	0.92**	0.86**	0.83**	0.92**	-0.27**	-0.32**	-0.32**	-0.64**	0.73**	0.62**	1		
MSAVI2	-0.004	0.99**	0.89**	0.99**	0.88**	0.93**	0.99**	-0.38**	-0.64**	-0.64**	-0.44**	0.79**	0.27**	0.92**	1	
CI	-0.03	-.90**	-0.69**	-.90**	-1.0**	-0.92**	-0.90**	0.32**	0.46**	0.46**	0.64**	-0.85**	-0.33**	-0.86**	-0.88**	1

OC: organic carbon; NDVI: normalised difference vegetation index; GNDVI: green normalised difference vegetation index; SAVI: soil adjusted vegetation index; GRVI: green red vegetation index; EVI: enhanced vegetation index; MSAVI: modified soil adjusted vegetation index; WDVI: weighted difference vegetation index; BI: brightness index; BI2: second brightness index; RI: redness index; V: vegetation index; SBI: soil brightness index; RDVI: renormalised difference vegetation index; MSAVI2: second modified soil adjusted vegetation index; CI: colour index.

Kandihumults. The SOC content of 2 : 1 clay-type soils was significantly higher (14 ± 0.5 g kg^{-1} dry soil) than that of 1 : 1 clay type soils (10 ± 0.7 g kg^{-1}) (26). Clay minerals have a high electric charge, which plays a special role in stabilising the organic compounds (27). The values of some indices decreased because of strong absorption and low reflectance. The thick pine forest on steep slopes provides resistance to soil erosion and protects soil aggregates, resulting in SOC accumulation. In addition, the waxy coated, needle shaped pine leaves have a high cellulose and lignin content and do not decompose quickly under favourable climatic conditions. Earlier studies reported that root C/N ratios play an important role in a large proportion of SOC (28). Differences in rooting systems contribute more carbon to the soil and high SOC accumulation depends on supplementing crop nitrogen requirements through the surrounding vegetation cover (deposition from the pine forest). The effect of indices followed the order SBI > RI > BI > BI2 for SOC (29). Areas with a high SBI had low SOC, whereas those with a low SBI had high SOC. Similar trends for SBI were identified previously, which agrees with the present findings (30).

Prediction of soil organic carbon

Using the most influential indices, SOC prediction models were constructed (Table 5). All the prediction models had very weak coefficients of determination (R^2) i.e., 0.123, 0.142, 0.122 and 0.137 for SLR, MLR, OLS and PLS, respectively, with a similarly low RMSE of 0.06. The models explained between 12.3 % and 14.2 % of the variation in SOC, indicating very poor predictability. The adjusted R^2 of the MLR model (0.106) differed from the corresponding R^2 (0.142) by approximately 4 % and the root mean square error (RMSE) was 0.066 which is very good compared with the range of the data studied (23). The adjusted R^2 for the stepwise regression (OLS) model was 0.118 which was only 0.142 % lower than its corresponding R^2 value (0.122), indicating no significant difference between R^2 and adjusted R^2. The SOC accumulation at different slope positions was found in the following order; nearly level (0–3 %) (2.19 % SOC) > foothills (3–8 %) and mid-hills (8–15 % and strong slope, 15–30 %) > steep slopes (> 30 %) (1.86 % SOC) (Table 3). Very labile and labile carbon may be transported from hilltops and sloppy lands and accumulate in the low-lying areas, thereby increasing the SOC content (31). At higher altitudes, SOC may decrease through runoff, erosion, oxidation and distortion of the ratio of carbon fractions. The SOC ligand exchange bonding of SOC with Fe^{3+} and Al^{3+} (cation bridges) in acid soils is very strong. The SOC storage was influenced by site factors (mainly bulk density and soil type), land use and climate, which contributed to the variance explained by the models (32). The remote sensing indices such as WDVI, BI, BI2, RI and SBI, showed higher values in hilly areas (> 30 % slope) than in nearly level areas (< 2 % slope), reflecting higher SOC under hilly conditions.

Table 5. Pearson correlation between the soil organic carbon (SOC) and slope (%) of the topography

	SOC (%)
OC (%)	1
Slope (%)	-0.303

Model validation analysis

The accuracy of these models was evaluated using RMSE of cross-validation and the R^2 as the performance indicators (Table 6 and Fig. 5). The RMSE and R^2 values varied from 0.170–0.198 % and 0.533–0.563, respectively, across all models. A smaller RMSE value and a greater R^2 for the predictive model indicate a better-suited model for prediction. The presence of mixed pixels from rice and the surrounding vegetation on steep, sloping land reduces spectral information and model predictability. Mixed pixels from rice and the surrounding grasses affect the accuracy of the analysis (32, 33). The PLSR chooses the optimal latent variables (LVs) and extracts the number of indices equal to the number of LVs. The R^2 and RMSE of the regression methods indicated the equivalence of PLSR.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}i - yi)^2}{\sum_{i=1}^{n} (yi - \bar{y}i)^2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (yi - \hat{y}i)^2}$$

Where, y_i: actual value; $\hat{y}_i$: predicted value at site i; n: number of samples; $\bar{y}i$: mean of the dependent variable.

Table 6. The soil organic carbon (SOC) prediction models

Sl. no.	Indices	Model	R^2	Adjusted R^2	RMSE
1	SLR	OC (%) = 1.535 + 0.0002 × SBI	0.123		
2	MLR	OC (%) = 2.05 + 0.26 × BI – 70.89 × RI + 0.01× SBI – 0.39 × BI2	0.142	0.106	0.066
3	OLS	OC (%) = 1.66 + 2.90E – 04 × BI - 42.69 × RI + 1.36E - 04 × SBI	0.122	0.118	0.261
4	PLSR	OC (%) = 1.41 + 5.82E – 04 × BI-13.29 × RI + 1.42E – 04 × SBI	0.137		0.261

SLR: simple linear regression; MLR: multiple linear regression; OLS: ordinary least square regression; PLSR: partial least square regression; OC: organic carbon; RMSE: root mean square error.

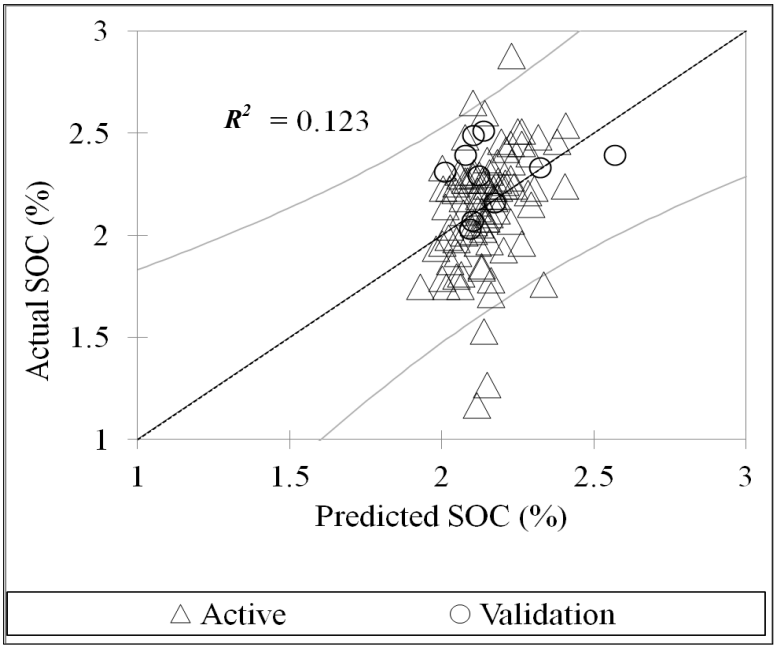


Fig. 5. Validation of model (multiple linear regression) predicted soil organic carbon with actual soil organic carbon.

SOC: soil organic carbon.

Conclusion

It is concluded that digital data, including remote sensing or space-borne data, are an alternate solution for soil organic carbon (SOC) mapping and estimation, considering the topological complexity of the study area. Fine spatial resolution data (< 10 m) are recommended for SOC mapping to support the sustainable management of rice-fallow systems in the highly fragile complex terrain of the Himalayan range. After rice harvesting, the absorption of red spectrum and reflection of NIR are comparatively low whereas the reflection of green spectrum is high. Therefore, green-band-based remote sensing (RS) indices such as the soil brightness indices (BI and BI2), the green red vegetation index (GRVI) and the green normalised difference vegetation index (GNDVI) are suitable for SOC estimation. The developed regression models had very low predictive power due to the non-uniformity of the sample size. The minimum sample size from the steep slope (> 30 %) position was due to the unavailability of long-duration rice-fallow system areas at steep slope positions, their very small and scattered distribution and the inaccessibility of soil sampling sites in the difficult terrain of the Himalayas. Therefore, for similar complex landscapes, simple spectral index-based regression models may not be sufficient and future work should explore the fusion of multiple datasets using artificial intelligence and machine learning-based algorithms. Among the models, the Multiple Linear Regression (MLR) model was the best predictor of SOC under Eastern Himalayan conditions:

$$SOC\ (\%) = 2.05 + 0.26 \times BI - 70.89 \times RI + 0.01 \times SBI - 0.39 \times BI2$$

Acknowledgements

The authors express their sincere gratitude to the competent authority of Central Agricultural University (Imphal), Lamphelpat, Manipur and the Director of the North Eastern Space Applications Centre (NESAC), Department of Space, Government of India, Umiam for their support and assistance throughout the study.

Authors' contributions

NJS undertook conceptualisation, funding acquisition, data curation, formal analysis and writing the original draft. PD was responsible for conceptualisation, data curation and writing the original draft. DKD participated in conceptualisation, formal analysis and writing the original draft. RJS contributed to writing, review and editing. BT was involved in writing, review and editing, as well as funding acquisition. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None

References

1. Ri Bhoi District. Government of Meghalaya. Ri Bhoi District website [Internet]. 2023 [cited 2023].
2. Bhalerao AK, Kumar B, Singha AK, Jat PC, Bordoloi R, Deka BC. Ri-Bhoi district inventory of agriculture. Umiam: ICAR-Agricultural Technology Application Research Institute, ICAR Zone-III; 2015. p. 1–86.
3. National Food Security Mission. A status note on rice in India [Internet]. 2016.
4. Ontl TA, Schulte LA. Soil carbon storage. Nat Educ Knowl. 2012;3(10):35.
5. Tshering D, Odeh IOA, Field DJ, Baillie IC. Digital soil mapping of soil organic carbon stocks under different land use and land cover types in montane ecosystems, eastern Himalayas. For Ecol Manag. 2014;318:91–102. https://doi.org/10.1016/j.foreco.2014.01.003
6. Van Breemen N, Buurman P. Soil formation. Dordrecht: Kluwer Academic Publishers; 2002. https://doi.org/10.1007/0-306-48163-4
7. Parras-Alcántara L, Lozano-García B, Keesstra SD, Cerdà A, Brevik EC. Long-term effects of soil management on ecosystem services and soil loss estimation in olive grove top soils. Sci Total Environ. 2016;571:498–506. https://doi.org/10.1016/j.scitotenv.2016.07.016
8. Herrick JE, Wander M. Relationships between soil organic carbon and soil quality in cropped and rangeland soils: the importance of distribution, composition and soil biological activity. In: Lal R, Kimble JM, Follett RF, Stewart BA, editors. Soil processes and the carbon cycle. Boca Raton: CRC Press; 1998. p. 405–25. https://doi.org/10.1201/9780203739273-28
9. Schoonover JE, Crim JF. An introduction to soil concepts and the role of soils in watershed management. J Contemp Water Res Educ. 2015;154(1):21–47. https://doi.org/10.1111/j.1936-704X.2015.03186.x
10. Vågen TG, Winowiecki L. Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. Environ Res Lett. 2013;8(1):015011. https://doi.org/10.1088/1748-9326/8/1/015011
11. Campbell G, Phinn SR, Dekker AG, Vittorio E. Remote sensing of water quality in an Australian tropical freshwater impoundment using matrix inversion and MERIS images. Remote Sens Environ. 2011;115(9):2402–14. https://doi.org/10.1016/j.rse.2011.05.003
12. Kumar S. Soil organic carbon mapping at field and regional scales using GIS and remote sensing applications. Adv Crop Sci Technol. 2013;1:e105. https://doi.org/10.4172/2329-8863.1000e105
13. National Bureau of Soil Survey and Land Use Planning (NBSS&LUP). Meghalaya soils. Nagpur: National Bureau of Soil Survey and Land Use Planning in collaboration with Directorate of Agriculture, Meghalaya; 1996.
14. Ranjan R, Jha A, Ramu N, Nain AS. Soil organic carbon estimation using remote sensing in Tarai region of Uttarakhand. Ann Plant Soil Res. 2015;17(Special Issue):361–4.
15. Bhunia GS, Kumar SP, Pourghasemi HR. Soil organic carbon mapping using remote sensing techniques and multivariate regression model. Geocarto Int. 2019;34(2):215–26. https://doi.org/10.1080/10106049.2017.1381179
16. Byrd KB, Ballanti L, Thomas N, Nguyen D, Holmquist JR, Simard M, et al. A remote sensing-based model of tidal marsh aboveground carbon stocks for the conterminous United States. ISPRS J Photogramm Remote Sens. 2018;139:255–71. https://doi.org/10.1016/j.isprsjprs.2018.03.019
17. Wang J, Xiao X, Bajgain R, Starks P, Steiner J, Doughty RB, et al. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. ISPRS J Photogramm Remote Sens. 2019;154:189–201. https://doi.org/10.1016/j.isprsjprs.2019.06.007
18. Central Ground Water Board. Ground water information booklet: Ri-Bhoi district, Meghalaya. Guwahati: Ministry of Water Resources, Government of India; 2013. Technical Report Series D No. 48/2011–12. p. 1–14.
19. Chan KY, Bowman A, Oates A. Oxidizable organic carbon fractions and soil quality changes in an Oxic Paleustalf under different pasture leys. Soil Sci. 2001;166:61–7. https://doi.org/10.1097/00010694-200101000-00009
20. European Space Agency. Sentinel-2 user handbook [Internet]. Paris: ESA; 2015. p. 1–64.
21. United States Geological Survey. Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) [Internet].
22. He T, Wang J, Lin Z, Cheng Y. Spectral features of soil organic matter. Geospat Inf Sci. 2009;12(1):33–40. https://doi.org/10.1007/s11806-009-0160-x
23. Jaber SM, Lant CL, Al-Qinna MI. Estimating spatial variations in soil organic carbon using satellite hyperspectral data and map algebra. Int J Remote Sens. 2011;32(18):5077–103. https://doi.org/10.1080/01431161.2010.494637
24. Somvanshi SS, Kunwar P, Tomar S, Singh M. Comparative statistical analysis of the quality of image enhancement techniques. Int J Image Data Fusion. 2018;9(2):131–51. https://doi.org/10.1080/19479832.2017.1355336
25. Lin YW, Deng BC, Xu QS, Yun YH, Liang YZ. The equivalence of partial least squares and principal component regression in the sufficient dimension reduction framework. Chemom Intell Lab Syst. 2016;150:58–64. https://doi.org/10.1016/j.chemolab.2015.11.003
26. Garrido FE, Hidalgo C, Paz F, Etchevers J, Merino C, Báez YA. Carbon saturation in the silt and clay particles in soils with contrasting mineralogy. Terra Latinoam. 2016;34:311–9.
27. Yan Y, Yang H. Interactions of clay minerals with biomolecules and protocells complex structures in the origin of life: a review. Adv Funct Mater. 2024;34:2406210. https://doi.org/10.1002/adfm.202406210
28. Dietzel R, Liebman M, Archontoulis S. A deeper look at the relationship between root carbon pools and the vertical distribution of the soil carbon pool. SOIL Discuss. 2017:1–30. https://doi.org/10.5194/soil-2017-5
29. Gholizadeh A, Zizala D, Saberioon M, Borůvka L. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. Remote Sens Environ. 2018;218:89–103. https://doi.org/10.1016/j.rse.2018.09.015
30. Kumar P, Pandey PC, Singh BK, Katiyar S, Mandal VP, Rani M, et al. Estimation of accumulated soil organic carbon stock in tropical forest using geospatial strategy. Egypt J Remote Sens Space Sci. 2016;19(1):109–23. https://doi.org/10.1016/j.ejrs.2015.12.003
31. Nath AJ, Sileshi GW, Laskar SY, Pathak K, Reang D, Nath A, et al. Quantifying carbon stocks and sequestration potential in agroforestry systems under divergent management scenarios relevant to India's nationally determined contribution. J Clean Prod. 2020;274:124831. https://doi.org/10.1016/j.jclepro.2020.124831
32. Strahler AH, Boschetti L, Foody GM, Friedl MA, Hansen MC, Herold M, et al. Global land cover validation: recommendations for evaluation and accuracy assessment of global land cover maps. Luxembourg: Office for Official Publications of the European Communities; 2006. p. 1–60.
33. Nuarsa IW, Nishio F, Hongo C. Relationship between rice spectral and rice yield using MODIS data. J Agric Sci. 2011;3(2):80–8. https://doi.org/10.5539/jas.v3n2p80

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc

See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Copyright: © The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited (https://creativecommons.org/licenses/by/4.0/)

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.