

A Novel Bare Soil Index for Enhancing the Mapping of Bare Soil Area – An Indicator of Urban Expansion

Smitha Prema Somanathan^{1*}, Maya Kesavan¹, Sudheer Kulamulla Parambath², Bindhu Vijayalekshmi Muraleedharan³, Sreelash Krishnan¹, Padmalal Damodharan¹

¹ National Centre for Earth Science Studies, Akkulam, Thiruvananthapuram, India

² Indian Institute of Technology Madras, Chennai, India

³ M/S Vassar Labs Pvt Ltd, Hyderabad, India

* Corresponding author's e-mail: pssmithaaa@gmail.com

ABSTRACT

Bare soil mapping is inevitable for the monitoring of soil resources, as it acts as a major indicator of urban development. Although remote sensing based indices are largely used for the delineation of bare soil, their effectiveness has always been a challenge due to overlapping spectral characteristics of the bare soil and built up areas. Recently, built up indices have been employed to increase the accuracy of bare soil mapping, which can induce discrepancies due to lesser sensitivity in smaller patches of bare soil areas in high density urban regions. In this context, modified normalized difference bare soil index (MNDBSI) is developed using a logical combination of index with additional constraint for the precise delineation of bare soil area using Sentinel 2 images. The qualitative and quantitative analyses adopted to evaluate the performance of the MNDBSI revealed an overall agreement value of above 95% and the improvement percentage of MNDBSI over the other indices compared in terms of spectral discrimination index and transformed divergence varied from 8% to 166% and 15% to 97%, respectively, across all study areas. The comparative analyses of results indicate that MNDBSI is an effective alternative to the existing indices for precise bare soil mapping which can be used as a promising indicator of urban expansion.

Keywords: bare soil index, built up areas, urban, bare land, soil mapping.

INTRODUCTION

The increased rate of urbanization has brought about unprecedented changes in the Earth's surface dynamics over the past few decades with drastic reduction in the area of bare soil around the world which may affect the sustainability of the regional ecosystem. The alarming rate of conversion from bare soil to artificial surfaces (He et al., 2020) can have long term impact on the natural environment with increased pressure on protected areas, along with increased urban flooding events (Wu et al., 2013; Du et al., 2015; Kim et al., 2016). The large scale mapping of bare soil areas with remote sensing techniques received greater popularity due to their simplicity, precision and cost effectiveness (Goudge et al., 2017;

Bouhennache et al., 2019; Gong et al., 2019) as well as proved to be a better alternative to traditional approaches of ground surveys and aerial photography which are costly, manpower intensive and time consuming.

Some indices employed combinations of red, NIR and thermal infrared bands (TIR) for the delineation of bare soil based on the ideology of higher temperature for urban areas (Koroleva et al., 2017; Li et al., 2017). The combination of NDVI with bare soil index (BSI) has been attempted by many researchers for more effective delineation of bare soil from vegetative cover in agricultural areas (Roy et al., 1996; Rikimaru et al., 2002; Wentzel, 2002; He et al., 2010; Stathakis et al., 2012). Ratio normalized difference soil index (RNDSI) has been proposed to delineate

bare soil by combining normalized difference of green and SWIR bands with the tasseled cap transformation (Deng et al., 2015). Although the performance of the normalized difference bareness index (NDBaI) and normalized difference bare land index (NBLI) is highly enhanced by the incorporation of thermal infrared band for the identification of bare soil (Zhao and Chen, 2005; Li et al., 2017), the lower spatial resolution (100 m in Landsat 8) has been its major disadvantage. Moreover, the nonavailability of thermal bands for some of the widely used satellite sensors limited their applicability. Though the modified normalized difference soil index (MNDSI) could overcome the resolution issues of using TIR band (Piyooosh et al., 2018), it has been effective only in hot and dry climate. The dry bare soil index (DBSI), which is an inverse of modified normalized difference water index (MNDWI) reduced by NDVI (Rasul et al., 2018), is effective in separating bare soil from vegetative areas; however, masking of built up areas was not effective. The modified bare soil index (MBI) (Nguyen et al., 2021) is effective in the separation of bare soil area from urban area during agricultural fallow period in tropical climatic region. The normalized difference bare soil index (NDBSI) has been effective in the delineation of bare land/soil from red brick which are mostly found in industrialized urban areas (Liu et al., 2022). Hyperspectral Bare Soil Index (HBSI) (Salas and Kumaran, 2023) achieved improved performance in bare soil classification compared to machine learning algorithms for classification.

Although the major challenge in the accurate classification of bare soil is due to the spectral similarity between bare soil and built up areas, which is further enhanced by the large variability and heterogeneity in the spectral characteristics of built up features (Gamba et al., 2003; Wang and Li, 2019). Recently, built-up indices have been largely relied upon for the mapping of bare soil which may lead to discrepancies in the identification of bare soil due to higher sensitivity of built up indices to urban features, rather than bare soil. Moreover, in a high density urban region, the saturation effect may bring in uncertainties related to bare soil mapping, as the built-up index may not be sensitive enough to identify small bare soil patches, especially when coarse resolution satellite imageries are used. In addition, the monitoring of bare soil surfaces is further limited due to spectral degradation, high density vegetative

areas, moisture content, surface roughness and deposition of undefined/foreign material (Ustin et al., 2004). As the bare soil indices rely completely upon spectral information, rather than contextual knowledge, the integration of additional information can enhance the accuracy of bare soil mapping, instead of employing built up indices for increasing mapping accuracy of bare soil areas. In this work, a new methodology which employs a logical combination of bare soil index with an additional constraint to eliminate built up area was proposed. This combined formulation which is termed as MNDBSI employs blue, red, NIR and SWIR1 bands of Sentinel-2 data to effectively separate bare land/soil from other types of land cover without the need for employing thermal band. The performance of MNDBSI in different environmental backgrounds and climate conditions is compared with widely used bare soil indices, such as BSI, DBSI and NDBSI. In addition, qualitative and quantitative evaluation of the indices are conducted to verify the level of performance of the new index.

STUDY AREA AND MATERIALS

To evaluate the performance of the proposed MNDBSI, one test and three validation sites with diverse environmental background and climate condition are selected. The location of all the four study areas is given in Figure 1. The Vembanad ecosystem, located in the state of Kerala, India, was selected as the test site. The latest cloud free Sentinel 2 images from Copernicus hub (<https://www.copernicus.eu/en/access-data/conventional-data-access-hubs>) provided at L1C level which are radiometrically and geometrically corrected, are used as the data sources to ensure the maximum number of valid pixels. The dates of the Sentinel 2 imageries used in the study are given in Table 1. The false color composite (FCC) of Sentinel 2 images is displayed in (RGB: bands 8-4-3) in order to easily interpret the commonly used types of land covers. Built up areas are displayed in various shades of cyan color, while bare land/soil is displayed as varying shades of pink/brown color and vegetation as light green color. The reference data for the quantitative evaluation of bare soil areas are interpreted based on field data collected and manually digitized bare soil map using high resolution images obtained from Google Earth Engine. The correctly identified

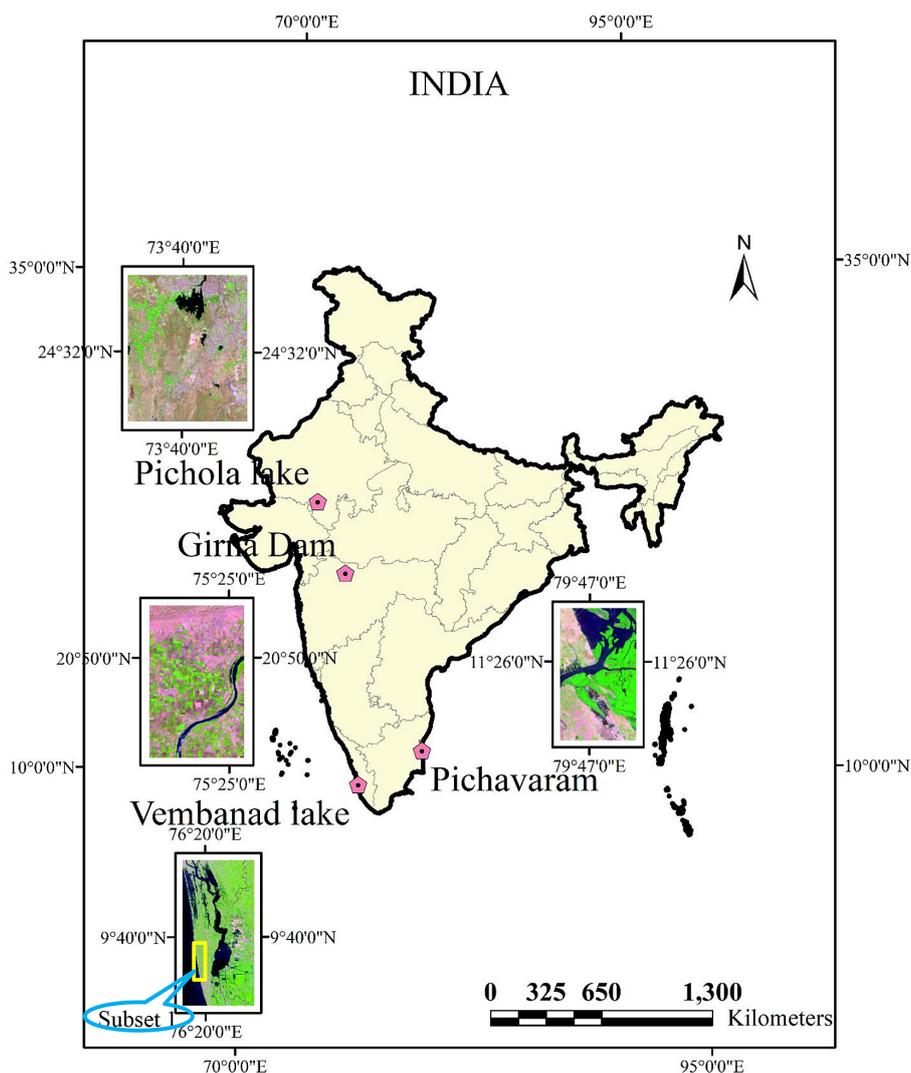


Figure 1. Location details of the selected study areas along with their false color composite

Table 1. Description of the four study sites

Study area	Climate type	Imagery date	Degree of built up area	Degree of bare soil area	Degree of vegetation area
Vembanad lake	Tropical monsoon	12-Feb-24	Medium	Low	High
Pichavaram mangrove	Tropical savannah	25-Jan-24	Low	High	Medium
Pichola lake	Dry arid desert	16-Feb-24	High	Medium	Low
Girna dam	Dry semi-arid	24-Jan-24	Low	Medium	High

bare soil areas are also compared with the results obtained from the classified image using maximum likelihood method to ensure minimum error in the reference data set.

METHODOLOGY

The methodology section covers the developmental procedure of the new bare soil index

(MNDBSI) and the assessment of the performance in comparison with the commonly used bare soil indices using quantitative and qualitative approaches.

Development of MNDBSI

The major built-up areas include buildings, parking lots, paved and gravel roads. The transformed land surfaces, such as stone, concrete, bricks, sand and asphalts also are considered as

built up areas (Liu et al., 2022; Zhao and Zhu, 2022). However, spectral characteristics of soil vary depending upon their composition, texture, moisture content and region of occurrence.

The box and whisker plots showing the variability in the reflectance of various common land covers in the study area are analyzed at first to arrive at a suitable combination of spectral bands for the formulation of index (Fig. 2). Figure 2 clearly indicates the close similarity in the reflectance pattern of built up areas in the visible to SWIR bands, which increases the complexity in the delineation of bare soil. As the reflectance pattern of built up areas does not show large variability among the short and long wavelength bands (Fig. 2), further analysis of scatter plot between the red and NIR (Fig. 3) is evaluated. From Figure 3, it can be seen that built up areas have distinct scattered distribution which lie exactly on the 1:1 trend line compared to bare soil areas. In turn, the area occupied by water and vegetation is just below and high above the 1:1 trend line, respectively, on red-NIR scatter plot which makes the scatter distribution unique. However, the scatter distribution of bare soil areas varies depending upon the type and nature of soil. For dry soil, the reflectance value of NIR band is higher than that of wet, dark and bright soil type (Fig. 3). Moreover, the reflectance

of dark soil is much lower for NIR band compared to other types of soils.

On the basis of the analysis of spectral responses (Fig. 2), soil has the highest reflectance in SWIR1 and lowest in blue band. In turn, built up areas showed lower reflectance for SWIR1 and higher reflectance for blue band. Moreover, the increase in the spectral reflectance from NIR to SWIR1 for bare lands is more than built up areas and this helps in the enhancement of bare soil features, if SWIR1 is employed as longer wavelength band in Equation 1. Hence, adopting normalized differences of SWIR1 and blue bands will achieve higher values for bare soil area compared to built-up areas. However, large variation between blue and SWIR1 bands exhibited by vegetation and water may bring uncertainties in the separation of bare soil. In addition, the difference in the reflectance value between red and NIR band for bare soil area is large compared to that of built up area (Fig. 2 and 3). Further, some built up areas have higher reflectance in NIR rather than SWIR bands (Zhao and Zhu, 2022), hence the incorporation of NIR band as short wavelength band in the index formulation may enhance the separation of bare soils from built up areas. To further increase the separability between soil and other land covers, an additional factor consisting of difference between

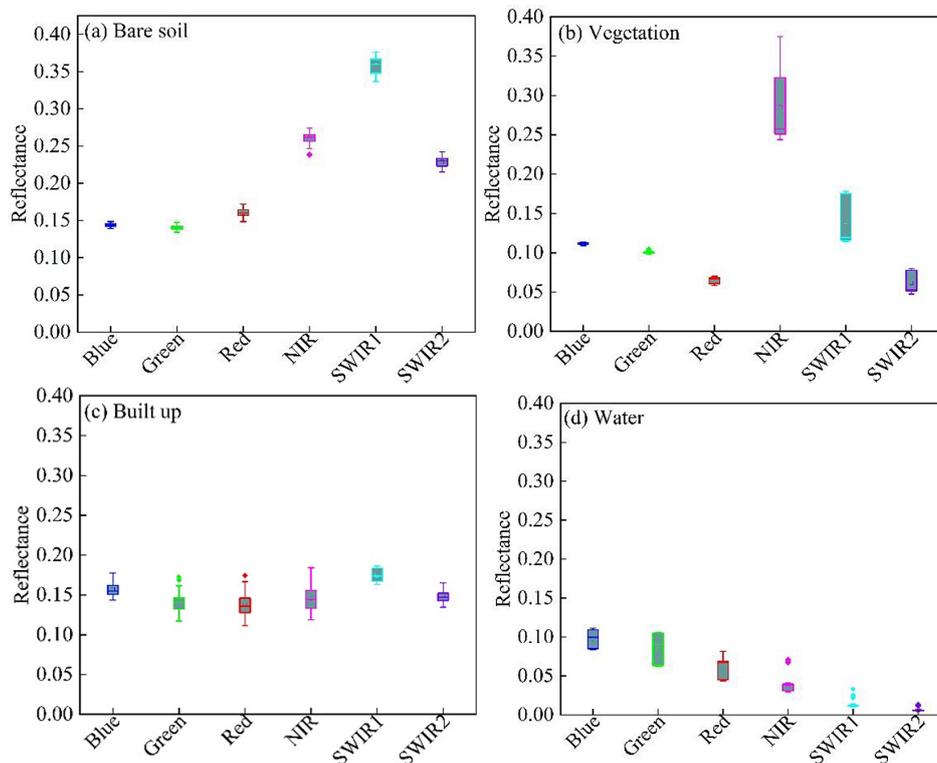


Figure 2. Box and whisker plot showing spectral characteristics of four types of land covers

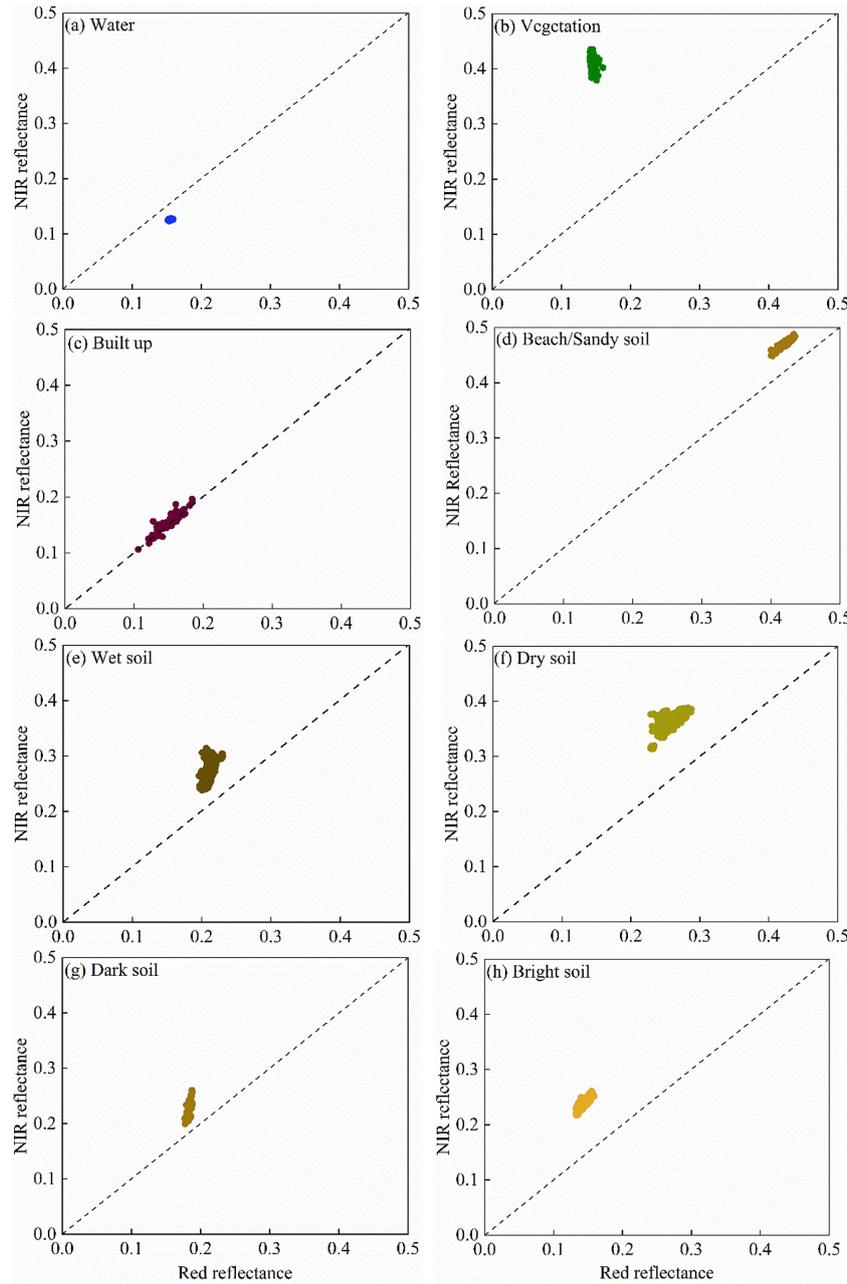


Figure 3. Scatter plot between red and NIR reflectance for (a) water (b) vegetation (c) built up areas (d) beach/sandy soil (e) wet soil (f) dry soil (g) dark soil (h) bright soil

the sum of reflectance of red and NIR bands with that of blue band is subtracted from the normalized difference ratio of SWIR1 and blue bands (Equation 1). Thereafter, optimal threshold value (β_T), is estimated using Otsu's algorithm (Lee et al., 1990; Liao et al., 2001; Sezgin et al., 2004) and a constraint, 'k' has been set (Equation 2) to ensure the complete separation of bare soil from built-up areas. The 'k' value of one represents bare soil and sparse vegetation, while zero value of 'k' represents built up areas. The value of 'k' is fixed based on two parameters (k_1, k_2), as in Equation 3

and 4. The final values of MNDBSI are evaluated using the Equation 5.

$$MNDBSI^* = \frac{SWIR1 - Blue}{SWIR1 + Blue - (NIR + Red - Blue)} \quad (1)$$

$$k = \text{if}(k_1 < k_2, \text{then } 0 \text{ else } 1) \quad (2)$$

$$k_1 = 1 - \frac{Red}{NIR} \quad (3)$$

$$k_2 = \frac{Red}{NIR} - 0.5 \quad (4)$$

$$MNDBSI = \text{if}((MNDBSI^* > \beta_T) \& (k = 1)) \text{ then } \text{abs}(MNDBSI^*) \text{ else } MNDBSI^* \quad (5)$$

where: k , k_1 and k_2 are parameters and b_T represents optimal threshold value. The values of *MNDBSI* range from -1 to 1, while positive values represent bare soil and negative values represent non bare soil areas.

Performance assessment of MNDBSI

The performance assessment of the bare soil map derived is conducted using reference data developed using high resolution satellite images from the Google Earth engine. The test samples for bare soil, built up, water and vegetation are chosen in such a way that the samples are well distributed and include 2700, 1595, 928 and 907 points for bare soil, built up, water and vegetation, respectively. For the quantitative analysis of the performance of MNDBSI, mapping accuracies, F1 score, Spectral Discrimination Index (SDI) and Transformed Divergence (TD) are evaluated. The mapping accuracy is assessed using overall agreement (OA), allocation disagreement (AD) and quantity disagreement (QD) (Pontius Jr and Millones, 2011; Pickard et al., 2017). SDI is used to quantify the degree of discrimination between the histogram for target (bare soil) and non-target area (non-bare soil) by measuring the difference between mean of bare soil and non-bare soil which is normalized by the sum of their standard deviations (Kaufman and Remer (1994). A good level of separability between target and non-target areas is indicated by SDI values greater than 1 (Pereira, 1999) and the calculation of SDI values (Deng et al., 2015) are based on Equation 6.

$$SDI = \frac{|\mu_1 - \mu_2|}{\sigma_1 + \sigma_2} \quad (6)$$

where: μ_1 and μ_2 represent mean index values of bare soil and non-bare soil area, and σ_1 and σ_2 are corresponding standard deviations of the index values of bare land/soil and non-bare soil area, respectively.

TD is a widely used quantitative estimator to evaluate the separability between classes (Tolpekin et al., 2009) which takes values between 0 and 2 and the separability between classes increases as TD values increases. Further, for the quantitative analyses of the performance of MNDBSI, frequency histogram and box charts of the index values are examined. Subsequently, mapping accuracies are also estimated to ascertain the effectiveness of MNDBSI.

RESULTS AND DISCUSSION

The quantitative and qualitative evaluations of MNDBSI are conducted to evaluate the effectiveness of delineation of bare soil areas.

Visual analyses of MNDBSI

The bare soil maps derived from Sentinel 2 images for the four indices are shown in Figure 4. The visual difference in the results of soil maps indicate that BSI and DBSI classified more areas as bare soil compared to MNDBSI and NDBSI for all sites (Fig. 4 (a3) to (e3)), except for Pichavaram, which is a mangrove area with less interference from built up features. The bare soil area delineated are more for BSI and DBSI in Vembanad Lake with large misclassifications of built up areas as bare soil (Fig. 4 (b1) to (e1)). This clearly brings out the inability of BSI and DBSI in the elimination of built up areas from bare soil areas in an urban area. From the visual comparison of the images of Pichola lake area (Fig. 4 (a2) to (e2)), it can be inferred that misclassification of BSI and DBSI occurs only in the built-up areas, while that of NDBSI includes both built up areas and lake water. The major misclassification of NDBSI arises due to water areas which are high density/sedimented water, such as sea water, sediment laden river/lake water. This is because the value of constraint factor in the formulation of NDBSI becomes greater than zero for non-bare soil area, which results in high positive value for high density/stagnant waters along with the positive values of bare soil. In the case of cultural command areas (CCA) of Girna dam (Fig. 4 (a4) to (e4)), the performance of BSI, DBSI and NDBSI is similar to that of Pichola lake. NDBSI showed misclassifications of built up areas, in addition to that of water in dry semi-arid (Girna dam) and dry arid desert (Pichola lake) climate. However, the performance of NDBSI in suppressing built up areas has been reasonably good and comparable to MNDBSI in the case of Vembanad lake which belongs to tropical climate.

Further a detailed comparison between the MNDBSI and other three indices for subset1 of test site (location shown in Figure 1) is given in Figure 5. For subset1 site, which represents sandy soils of beach area where built up area coexists with vegetation, BSI and DBSI have misclassifications of sandy soils in the beach as bare soil. NDBSI has misclassifications of water area and wet beaches as bare soil. This is mainly due to the

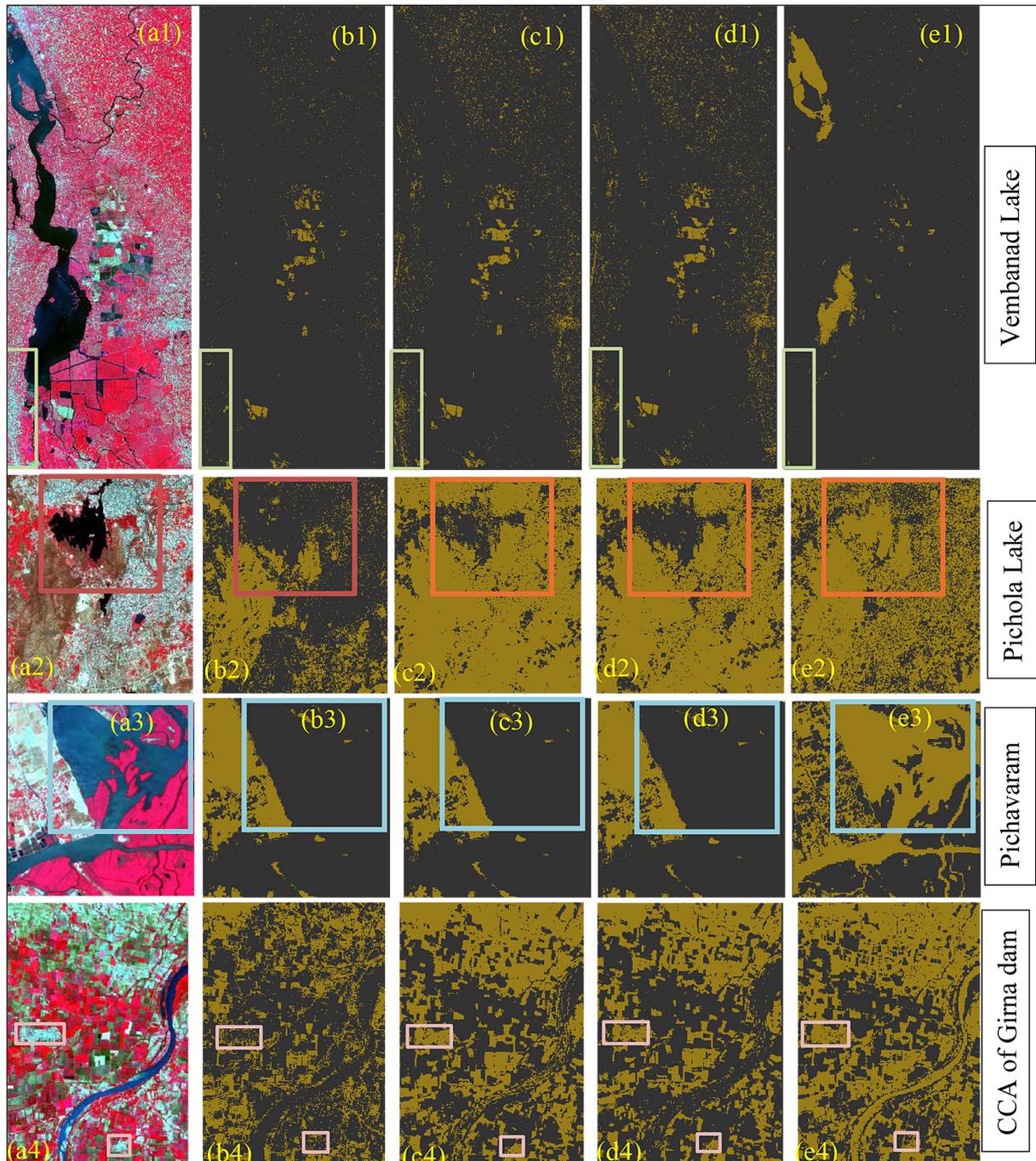


Figure 4. Comparison of (a) false color composite and images of indices (b) MNDBSI (c) BSI (d) DBSI and (e) NDBSI (classified bare soil is indicated as mustard or dark yellow color and black color represents non bare soil area. The colored rectangular boxes highlight the areas misclassified as bare soil. CCA represents culturable command area)

close resemblance of spectral reflectance between high albedo (reflection) surfaces, such as sandy beaches and built-up areas. However, MNDBSI outperforms the other three indices in the elimination of built-up areas and sandy areas. The present study is effective in achieving spectral separability of bare soil and sandy/saline beach areas which has not been studied much in previous studies.

Statistical comparison

The quantitative assessment of the performance of indices is conducted using box and whisker plot as shown in Figure 6. MNDBSI has highest values for bare soil followed by built up, water and vegetation (Fig. 6a) respectively. Though there is overlap between water, vegetation and built up

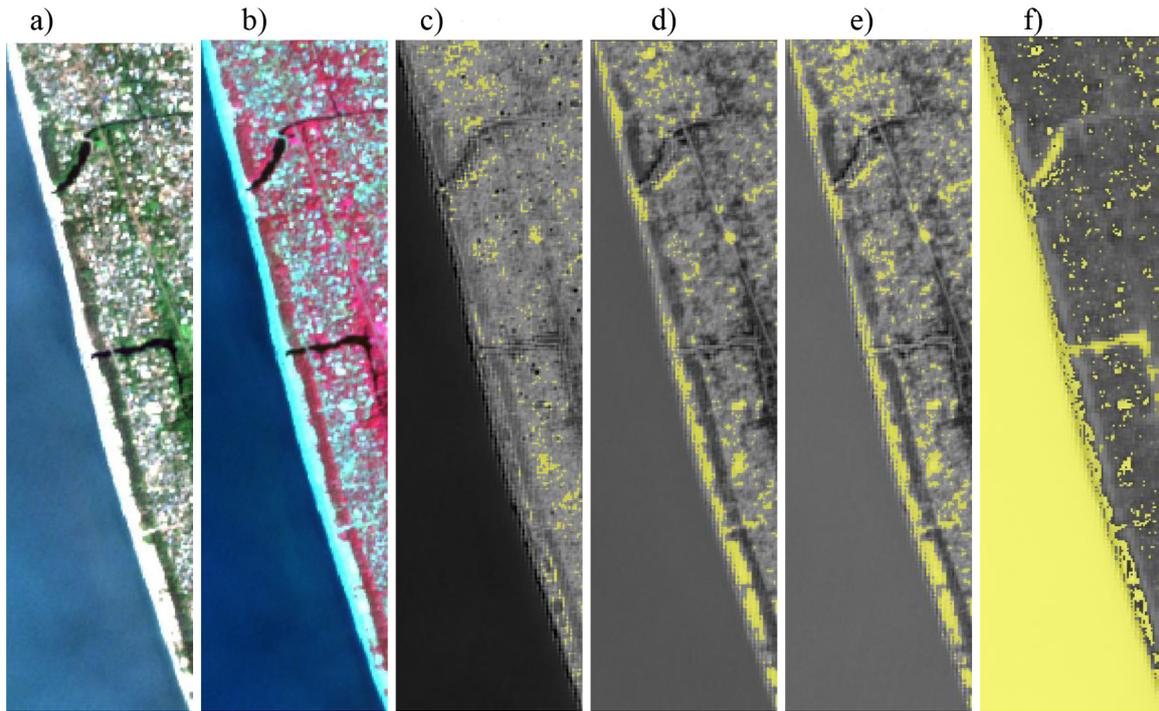


Figure 5. Bare soil map from subset 1 of test site (a) true color of Sentinel 2 (b) false color composite of Sentinel 2 image in Vembanad lake along the seaside with wide beaches having prominent sandy soil derived using (c) MNDBSI (d) BSI (e) DBSI (f) NDBSI. The yellow color indicates bare soil area delineated by the corresponding indices

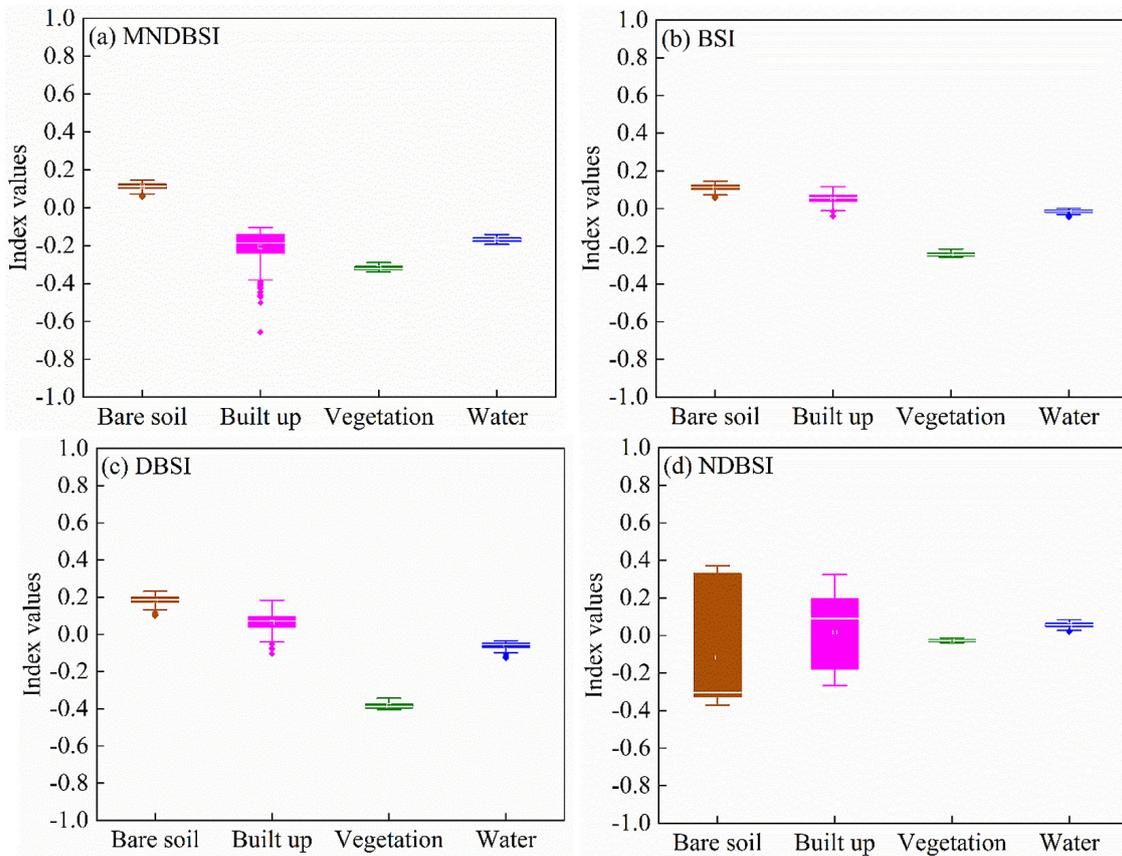


Figure 6. Box and whisker plot of the index values (a) MNDBSI (b) BSI (c) DBSI (d) NDBSI for different types of land cover

areas for MNDBSI, the degree of separation between bare soil and built up area is sufficiently large to enhance delineation of bare soil which is the priority of the study. Although the overlap between bare soil and built up areas for DBSI is lesser than BSI (Fig. 6b and c) DBSI has greater potential in suppressing built up areas than BSI. However, BSI and DBSI showed good level of separability between water and vegetation from bare soil, which indicates that BSI and DBSI can be effectively used in the delineation of bare soil in wetlands and agricultural areas rather than in urban areas. In contrast, NDBSI with a boarder range of bare soil values (-0.4 to 0.4) which showed more confusion in the elimination of water and vegetation along with built up areas (Fig. 6d), has limited ability of mapping bare soil. The frequency histogram analysis of MNDBSI clearly shows a good level of separation between the distributions of water bodies, built up and vegetation areas from bare soil (Fig. 7a) which is also supported by high values of SDI (> 1) in Table 2. Although the distribution of DBSI follows same pattern as that of BSI (Fig. 7b and c), the distribution of built up areas is wider and shorter peaked compared to that of BSI. The high degree of overlap between bare

soil area and water observed for NDBSI (Fig. 7 d), indicates the confusion between bare soil and water and the need for masking of water areas prior to the delineation of bare soil. Visually, MNDBSI separates the types of land cover into two, while the other three indices into four categories of land cover. Further, the separability of bare soil from other land covers evaluation by TD and SDI values indicates good level of separability of bare soil from other land cover for MNDBSI with values of 2.0 and above for the test site (Table 2). The lower values of SDI and TD for BSI, DBSI and NDBSI specifically for the urban / built up areas can be of major concern as the priority of these indices is to delineate bare soil from other types of land cover. A TD value greater than 1.0 is obtained for all validation sites and indices except for NDBSI (Table 3). However, the validation sites which have the influence of water bodies such as Pichola Lake and Girna dam, has SDI values lower than 1.0 for NDBSI. In turn, Pichavaram region, with minimal interference from built up features, has SDI value greater than 1.0 for DBSI and NDBSI. The statistical analysis reveals the better performance of MNDBSI with SDI and TD values consistently greater than 1.0 for all validation sites (Table 3).

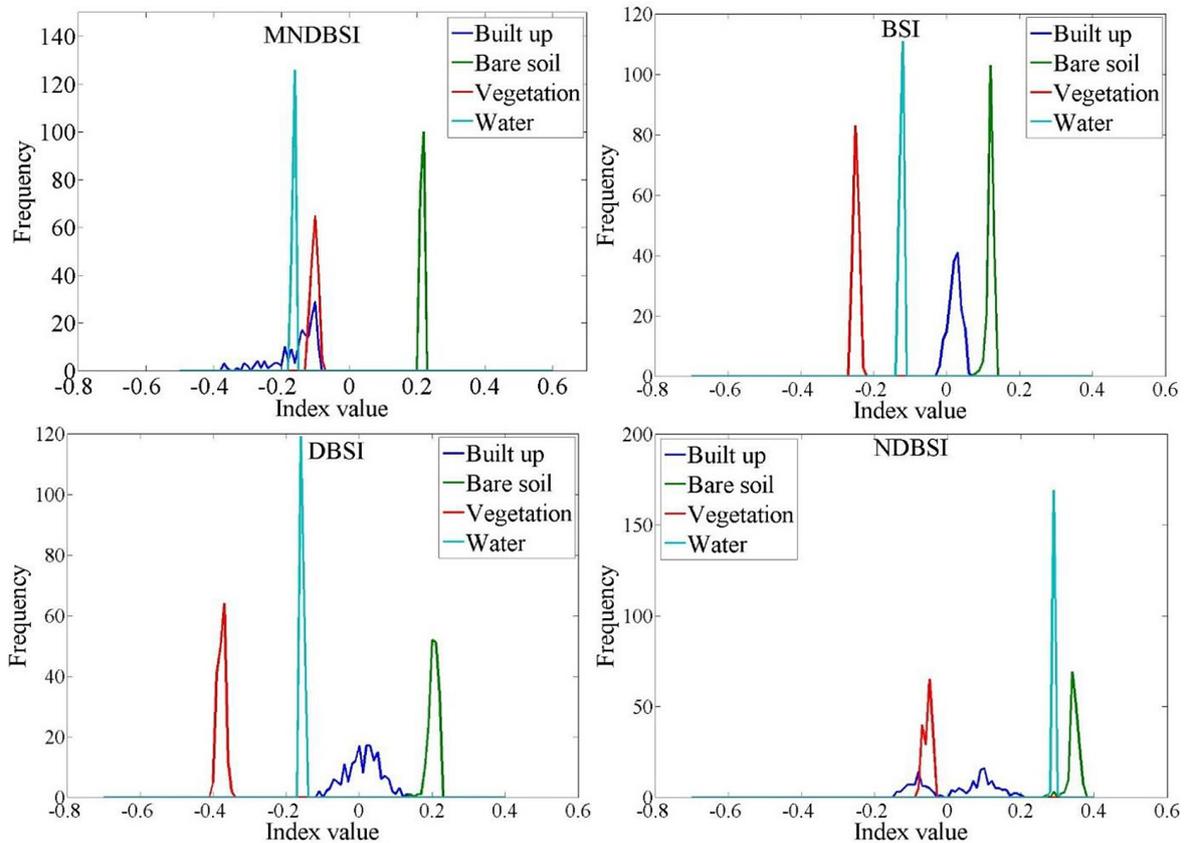


Figure 7. Frequency histogram of (a) MNDBSI, (b) BSI (c) DBSI and (d) NDBSI

Table 2. Statistics of different types of land cover for test site

Metrics	Land cover types			
	Bare Soil	Urban	Vegetation	Water
MNDBSI				
Minimum	-0.04	-0.48	-0.28	-0.22
Maximum	0.18	-0.02	-0.14	-0.13
Mean	0.08	-0.17	-0.21	-0.17
Standard deviation	0.03	0.09	0.03	0.01
SDI		2.05	4.71	5.3
TD		2	2	2
BSI				
Minimum	-0.03	-0.05	-0.32	-0.06
Maximum	0.16	0.11	-0.21	0.01
Mean	0.09	0.05	-0.27	-0.01
Standard deviation	0.02	0.03	0.02	0.01
SDI		0.9	7.76	3.37
TD		0.73	2	2
DBSI				
Minimum	-0.1	-0.12	-0.43	-0.17
Maximum	0.24	0.16	-0.33	-0.01
Mean	0.12	0.05	-0.38	-0.05
Standard deviation	0.04	0.05	0.02	0.01
SDI		0.82	8.45	3.44
TD		0.68	2	2
NDBSI				
Minimum	-0.4	-0.3	-0.38	0.58
Maximum	0.44	0.36	-0.27	0.64
Mean	0.38	0.18	-0.34	0.62
Standard deviation	0.1	0.17	0.02	0.01
SDI		0.77	5.84	-2.17
TD		0.61	2	0.04

Effect of seasonality on bare soil area mapping

The separability metrics (Table 4) indicates that the performance of MNDBSI is good ($SDI > 1.0$) during all the months from November through March, which includes post monsoon, winter and start of the summer season. The best performance of MNDBSI is in the month of March (summer season) and lowest in November (post monsoon season). The performance of NDBSI is poor ($SDI < 0.5$) for all the four months. However, the performance of BSI and DBSI is good ($SDI > 1.0$) during March. The comparison of TD values (Table 4) indicates that for the month of March, BSI and DBSI have TD values greater than 1.0 for urban/built up areas. Moreover, TD values are greater than 1.0 for vegetative areas and around

water bodies for all the compared months and indices except for NDBSI. However, NDBSI gave low to negative values only in the areas around water bodies. The TD values of MNDBSI are consistently above 1.0 for all the four months and types of land cover. Hence, seasonality analysis confirmed that MNDBSI has good level of robustness in the delineation of bare soil from other types of land cover.

Accuracy assessment of the bare soil indices

The results of the accuracy assessment show that the bare soil maps generated have an OA of 97.7%, 86.1% and 85.7% for MNDBSI, BSI and DBSI, respectively (Table 5) for the test site. NDBSI exhibited consistently low values of OA for all the four sites. This is majorly due to the

Table 3. Separability metrics for the validation sites

Index	Metrics	Land cover types		
		Urban	Veg	Water
Pichavaram				
MNDBSI	SDI	1.83	6.69	7.88
	TD	1.71	2	2
BSI	SDI	0.7	1.02	0.31
	TD	0.65	1.01	0.2
DBSI	SDI	1.42	6.17	7.98
	TD	1.33	2	2
NDBSI	SDI	1.2	2.04	-8
	TD	1.01	1.47	0
Pichola lake				
MNDBSI	SDI	1.82	4.64	6.41
	TD	1.68	1.98	1.99
BSI	SDI	1.68	11.71	9.13
	TD	1.42	2	2
DBSI	SDI	1.36	13.5	9.41
	TD	1.21	2	2
NDBSI	SDI	0.83	2.23	0.86
	TD	0.67	1.89	0.7
Girna dam				
MNDBSI	SDI	1.47	3.04	7.21
	TD	1.38	2	2
BSI	SDI	1.13	6.32	4.33
	TD	1.08	2	2
DBSI	SDI	1.06	6.79	5.55
	TD	1.01	2	2
NDBSI	SDI	0.97	2.84	-2.55
	TD	0.51	1.85	0.08

lower efficiency in masking water area rather than the separation of bare soil from built up areas. MNDBSI exhibited the effectiveness in mapping of bare soil with over 95% overall agreement and F1 score value above 0.93 for all sites. The mapping accuracy of NDBSI varies from 51.3% to 78.1% with F1 score from 0.41 to 0.72. Overall, NDBSI exhibited limited skill in the mapping of bare soil areas with water bodies. The mapping accuracy of BSI ranges from 85.0% to 92.2% while that of DBSI varies from 82.0% to 91.2%. The results of accuracy assessment indicate that the performance of BSI and DBSI are closely matching. To confirm the similarity between indices, correlation coefficient is evaluated (Table 6). The correlation between BSI and DBSI is found to be 0.98. MNDBSI has the least correlation of 0.24 with NDBSI and highest correlation of 0.88 with BSI. The percentage of improvement of

MNDBSI over the other three indices in terms of SDI and TD are also calculated. The results indicated that MNDBSI improves SDI and TD values by 8% to 166% and 15% to 97%, respectively, over the other three indices. Hence, MNDBSI can be used as an effective alternative for the existing bare soil indices for precise soil mapping. However, the major limitation of this study is in the analysis of monthly variability of bare soil area which has been limited due to the non-availability of cloud free images during rainy season.

CONCLUSIONS

The MNDBSI proposed in the present work employs blue, red, NIR and SWIR1 bands of Sentinel 2 to separate built up areas and other types of land covers from bare soil areas under

Table 4. Seasonal variability analysis of the indices

Month	Index	Metric	Land cover types		
			Urban	Veg	Water
November	MNDBSI	SDI	1.02	3.74	7.93
		TD	1.41	2	2
	BSI	SDI	-1.37	5.02	4.36
		TD	0	2	2
	DBSI	SDI	-1.44	4.72	3.9
		TD	0	2	2
NDBSI	SDI	-1.17	4.05	-6.95	
	TD	0	1.9	0	
December	MNDBSI	SDI	1.74	6.71	7.13
		TD	1.65	2	2
	BSI	SDI	0.09	7.46	9.44
		TD	0.01	2	2
	DBSI	SDI	-0.09	7.03	9.07
		TD	0	1.98	2
NDBSI	SDI	0.23	4.13	-2.5	
	TD	0.2	1.99	0	
January	MNDBSI	SDI	1.72	5.39	6.78
		TD	1.65	2	2
	BSI	SDI	1.11	7.4	3.37
		TD	0.95	2	1.88
	DBSI	SDI	0.85	6.78	3.21
		TD	0.95	2	1.85
NDBSI	SDI	0.12	0.99	-0.08	
	TD	0.23	0.5	0	
March	MNDBSI	SDI	2.47	7.02	8.15
		TD	2	2	2
	BSI	SDI	1.27	5.59	6.84
		TD	1.01	1.98	2
	DBSI	SDI	1.57	5.27	8.63
		TD	1.18	2	2
NDBSI	SDI	-0.13	0.54	-1.42	
	TD	0	0.2	0	

different climate conditions and environment. An additional constraint using ratio of red and NIR bands is incorporated in the MNDBSI to maximize the suppression of built up areas. The qualitative evaluation of the indices demonstrates the better performance of MNDBSI, compared to the other three indices in suppressing built up areas along with vegetation and water. The quantitative evaluation reveals that MNDBSI has overall best performance across different sites considered in this study. The overall agreement values of above 95% is obtained for MNDBSI across all study areas. In addition, the improvement percentage of MNDBSI over the other indices in terms of

SDI and TD ranges from 8% to 166% and 15% to 97%, respectively, across all study areas. This study could effectively improve the precision in the separation of built up from bare soil areas, even without the incorporation of thermal bands which may not be available for many commonly available optical satellite imageries. The results indicate that MNDBSI can be reliably used for differentiating built up and bare land from other land use classes in tropical, arid and semi-arid climates. As the spatiotemporal representation of bare soil surface indicates natural or human activities, along with many interdisciplinary applications related to soil resource mapping, MNDBSI

Table 5. Summary of accuracy assessments of sites

Index	Overall agreement	Quantity disagreement	Allocation disagreement	F1 score
Vembanad lake				
MNDBSI	97.7	0.8	1.5	0.97
BSI	86.1	2.9	11	0.83
DBSI	85.7	11.9	2.4	0.82
NDBSI	75.5	21.9	2.6	0.72
Pichola lake				
MNDBSI	97.2	1.1	1.7	0.96
BSI	85	6.5	8.5	0.87
DBSI	82	16.8	1.2	0.85
NDBSI	65.3	22.2	12.5	0.63
Girna dam				
MNDBSI	96.2	0.8	2.9	0.94
BSI	92.2	6.8	1.2	0.93
DBSI	91.2	7.4	1.4	0.9
NDBSI	78.1	20.7	1.2	0.71
Pichavaram				
MNDBSI	95.2	3.4	1.4	0.93
BSI	90.2	8.7	1.5	0.88
DBSI	89.7	9.8	0.5	0.85
NDBSI	51.3	48.1	0.5	0.41

Table 6. Correlation matrix of the four indices

Index	MNDBSI	BSI	NDBSI	DBSI
MNDBSI	1	0.88	0.24	0.82
BSI	0.88	1	0.47	0.98
NDBSI	0.24	0.47	1	0.47
DBSI	0.82	0.98	0.47	1

can aid in the development of policies and assessment of changes in land-use and soil management. Hence, MNDBSI can be effectively used as an indicator of urban development.

Acknowledgements

This study is funded by Kerala State Council for Science, Technology and Environment (KSCSTE), Council (P) Order No. 156/2021/KSCSTE. We are grateful to the Director, National Centre for Earth Science Studies, Ministry of Earth Science, Government of India for the facilities and support.

REFERENCES

1. Ali, A., and Nayyar, Z.A. 2021. A Modified Built-up Index (MBI) for automatic urban area extraction from Landsat 8 Imagery. *Infrared Physics &*

Technology. 116, 103769.
 2. Bouhennache, R., Bouden, T., Taleb-Ahmed, A., Cheddad, A. 2019. A new spectral index for the extraction of built-up land features from Landsat 8 satellite imagery. *Geocarto International.* 34(14), 1531–1551.
 3. Deng, Y.B., Wu, C.S., Li, M., Chen, R.R. 2015. RNDSI: A ratio normalized difference soil index for remote sensing of urban/suburban environments. *Int J Appl Earth Obs.* 39, 40–48.
 4. Du, S., Shi, P., Van Rompaey, A., Wen, J. 2015. Quantifying the impact of impervious surface location on flood peak discharge in urban areas. *Natural Hazards.* 76(3), 1457–1471.
 5. Gamba, P., Acqua, F.D. 2003. Increased accuracy multiband urban classification using a neuro-fuzzy classifier. *Int. J. Remote Sens.* 24, 827–834.
 6. Gong, P., Li, X., Zhang, W. 2019. 40-Year (1978–2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Science Bulletin.* 64(11), 756–763.

7. Goudge, T.A., Russell, J.M., Mustard, J.F., Head, J.W., Bijaksana, S.A. 2017. 40,000 Yr Record of Clay Mineralogy at Lake Towuti, Indonesia: Paleoclimate Reconstruction from Reflectance Spectroscopy and Perspectives on Paleolakes on Mars. *GSA Bull.* 129, 806–819.
8. Hardi, T., Repask'a, G., Veselovský, J., Vilinova, K. 2020. Environmental consequences of the urban sprawl in the suburban zone of Nitra: An analysis based on landcover data. *Geographica Pannonica.* 24(3), 205–220.
9. He, C., Shi, P., Xie, D., Zhao, Y. 2010. Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. *Remote Sens. Lett.* 1, 213–221.
10. Kaufman, Y.J., Remer, L.A. 1994. Detection of forests using mid-ir reflectance – an application for aerosol studies. *IEEE T Geosci Remote.* 32(3), 672–683.
11. Kim, H., Jeong, H., Jeon, J., Bae, S., 2016. The impact of impervious surface on water quality and its threshold in Korea. *Water.* 8(4), 111.
12. Koroleva, P.V., Rukhovich, D.I., Rukhovich, A.D., Rukhovich, D.D., Kulyanitsa, A.L., Trubnikov, A.V., Kalinina, N.V., Simakova, M.S. 2017. Location of bare soil surface and soil line on the RED–NIR spectral plane. *Eurasian Soil Sc.* 50, 1375–1385.
13. Li, H., Wang, C., Zhong, C., Su, A., Xiong, C., Wang, J., Liu, J. 2017. Mapping urban bare land automatically from landsat imagery with a simple index. *Remote Sensing.* 9, 249.
14. Liao, P.S., Chen, T.S., Chung, P.C. 2001. A fast algorithm for multilevel thresholding, *J. Inform. Sci. Eng.* 17(5) 713–727.
15. Lee, H., Park, R.H. 1990. Comments on An optimal multiple threshold scheme for image segmentation, *IEEE Trans. Syst. Man Cybern.* 20(3) 741–742.
16. Liu, Y., Meng, Q., Zhang, L., Wu, C. 2022. NDBSI: A normalized difference bare soil index for remote sensing to improve bare soil mapping accuracy in urban and rural areas, *Catena.* 214, 106265.
17. Nguyen, C.T., Chidthaisong, A., Kieu Diem, P., Huo, L.Z. 2021. A modified bare soil index to identify bare land features during agricultural fallow-period in Southeast Asia using Landsat 8. *Land.* 10, 231.
18. Pereira, J.M.C. 1999. A comparative evaluation of NOAA/AVHRR vegetation indexes for burned surface detection and mapping. *IEEE T Geosci Remote.* 37(1), 217–226.
19. Pontius Jr, R.G., Millones, M. 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment, *Int. J. Remote Sens.* 32(15), 4407–4429.
20. Pickard, B., Gray, J., Meentemeyer, R. 2017. Comparing quantity, allocation and configuration accuracy of multiple land change models. *Land.* 6, 52.
21. Piyooosh, A.K., Ghosh, S.K. 2018. Development of a modified bare soil and urban index for Landsat 8 satellite data. *Geocarto Int.* 33, 423–442.
22. Rasul, A., Balzter, H., Ibrahim, G.R.F., Hameed, H.M., Wheeler, J., Adamu, B., Ibrahim, S., Najmaddin, P.M. 2018. Applying built-up and bare-soil indices from Landsat 8 to cities in dry climates. *Land.* 7, 81.
23. Rikimaru, A., Roy, P.S., Miyatake, S. 2002. Tropical forest cover density mapping. *Trop. Ecol.* 43, 39–47.
24. Roy, P.S., Sharma, K.P., Jain, A. 1996. Stratification of density in dry deciduous forest using satellite remote sensing digital data—An approach based on spectral indices. *J. Biosci.* 21, 723–734.
25. Salas, E.A.L., Kumaran, S.S. 2023. Hyperspectral Bare Soil Index (HBSI): Mapping soil using an ensemble of spectral indices in machine learning environment. *Land.* 12, 1375.
26. Sezgin, M., Sankur, B. 2004. Survey over image thresholding techniques and quantitative performance evaluation, *J. Electron. Imaging.* 13(1) 146–166,
27. Stathakis, D., Perakis, K., Savin, I. 2012. Efficient segmentation of urban areas by the VIBI. *Int. J. Remote Sens.* 33, 6361–6377.
28. Tolpekin, V., Stein, A. 2009. Quantification of the effects of land-cover-class spectral separability on the accuracy of markov-random-field-based super resolution mapping. *IEEE Trans. Geosci. Remote Sens.* 47, 3283–3297.
29. Ustin, S.L., Roberts, D.A., Gamon, J.A., Asner, G.P., Green, R.O. 2004. Using imaging spectroscopy to study ecosystem processes and properties. *Bioscience.* 54, 523–534.
30. Wang, Y., Li, M. 2019. Urban impervious surface detection from remote sensing images: A review of the methods and challenges. *IEEE Geoscience and Remote Sensing Magazine.* 7 (3), 64–93.
31. Wentzel, K. 2002. Determination of the overall soil erosion potential in the Nsikazi District (Mpumalanga Province, South Africa) using remote sensing and GIS. *Can. J. Remote Sens.* 28, 322–327.
32. Wu, K.Y., Ye, X.Y., Qi, Z.F., Zhang, H. 2013. Impacts of land use/land cover change and socioeconomic development on regional ecosystem services: The case of fast growing Hangzhou metropolitan area, China. *Cities.* 31, 276–284.
33. Zhao, H., Chen, X. 2005. Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. *Int. Geosci. Remote Sens. Symp.* 3, 1666–1668.
34. Zhao, Y., Zhu, Z. 2022. ASI: An artificial surface Index for Landsat 8 imagery. *International Journal of Applied Earth Observation and Geoinformation.* 107, 102703.