Automated assessment of the extent of mangroves using multispectral satellite remote sensing data in Google Earth Engine

Rupsa Sarkar¹, L. Gnanappazham²*, and A. C. Pandey¹
¹Central University of Jharkhand, Department of Geoinformatics, Brambe, Ranchi 835 222, India
²Indian Institute of Space Science and Technology, Thiruvananthapuram 695 547, India

This study on the automatic assessment of mangroves uses geometric, textural parameters and vegetation indices derived from Landsat 8 images utilizing the Google Earth Engine. The extent of Indian mangroves is estimated as 5581 sq. km for 2019, with an overall accuracy (OA) of 86% and kappa coefficient (k) of 0.77. Among the five regions studied, maximum OA was obtained for Mumbai (94%; k = 0.89) and minimum for Godavari (81.625%; k = 0.66). Such automated mapping will benefit effective mangrove monitoring and management with a near real-time accurate estimation of mangroves.

Keywords: Automated mapping, cloud platform, mangrove ecosystem, satellite data.

Mangroves are halophytes that grow in harsh tropical and subtropical conditions. They have a complex root system that expels the extra salt and adapts various structures. They also have a unique mode of reproduction in which the seeds germinate and grow while being attached to the parent tree in a viviparous manner. Wetlands are ecosystems inundated by water throughout the year or during some months of the year, constituting the most productive part of this ecosystem. Mangroves are also the prime nesting sites for hundreds of bird species.

According to the World Mangroves Atlas, mangrove forests are found in 123 countries and territories across the world, with roughly 70 species occupying a total area of 181,000 sq. km. India accounts for about 3% of this, including certain endangered species. Sonneratia griffithii and Heritiera fomes are endangered species with fewer seeds and slow growth.

Remote sensing is the viable approach to managing and monitoring large and dense vegetation. In the mangrove ecosystem, spectral properties of various plant species depend on their growth forms, density, soil type and other vegetative components that differ from each other and help in mapping, monitoring and integrating the temporal datasets into Geographic Information System (GIS). Remote sensing yields promising results for estimating biochemical and biophysical parameters of different species, making field sampling more efficient.

Various vegetation indices have been developed to evaluate forest cover better using spectral properties, such as enhanced vegetation index, advanced vegetation index, bare soil index, normalized difference wetland vegetation index, leaf area index and Canopy Shadow Index, which also gives the vegetation density. Normalized difference wetland vegetation index is commonly used in wetland studies. Mangrove forests of China have been mapped based on the greenness, leaf area index and tidal inundation, using enhanced vegetation index, modified NDWI and surface water index and normalized difference vegetation index. This has demonstrated the potential of using Google Earth Engine (GEE) for mangrove identification utilizing Landsat and Sentinel SAR datasets. Several measures, such as NDVI, EVI and NDWI derived to assess and monitor vegetation health; these are sensitive to the biophysical variables. Apart from the geometrical properties, vegetative spectra such as area and perimeter for different spectral bands have also been studied to understand vegetation. The three-angle index (TAI) concept uses a spectral curve of vegetation in the Vis–NIR region for the calculation of all three angles within a triangle and the length of the vectors.

Textural features were extracted from the spatial relationship of pixels and specific image properties such as homogeneity, contrast, dissimilarity, variance, correlation, etc. In 1970, texture analysis methods were used for land-use classification employing the power spectral method (PSM) to analyse a high-resolution B/W aerial photograph. One of the studies compared the relative merits of four algorithms, viz. spatial gray level dependence method (SGLDM), gray level run length method (GLRLM), PSM and gray level difference method (GLDM) for terrain-type classification. SGLDM texture measure gave the best overall classification accuracy.

Forest variables (age, top height, circumference, stand density and basal area) were derived using the texture features from IKONOS-2 imagery. In 2013, variance property (9 × 9 pixel metrics) was used for texture analysis to classify mangrove species in Malaysia using RapidEye satellite data. The accuracy of the textured data classification...
was found to be 84% in comparison to non-textured data classification, which was 80%. Expert intervention is very much important to decide on the threshold value of NDVI to discriminate the vegetation from non-vegetation and similarly to discriminate mangroves from non-mangroves using advanced or specific analyses\textsuperscript{19,20}. GEE is a cloud platform for big geospatial datasets, making it easy to access many spatial and temporal open data and analyse them on the same platform with high performance\textsuperscript{21}.

In this study, we automatically map the extent of mangroves in India and their discrimination from other vegetation using multispectral Landsat 8 ETM satellite data along with free and open elevation data. By using the potential of GEE, we have developed an index texture-based algorithm to facilitate the automated discrimination of mangroves using specific bands of the Landsat 8 data available as big geospatial data in GEE.

**Study area**

As this study aimed to discriminate the mangroves from other adjoining vegetation types for any given multispectral satellite image of any coastal region, we have assessed the extent of the entire Indian coast. India has 4975 sq. km of mangroves, with West Bengal having 42.45% of the mangrove cover, followed by Gujarat (23.66%) and Andaman and Nicobar Islands (12.39%)\textsuperscript{22}. The distribution of mangroves on the east and west coasts is uneven. Most mangroves (60%) are found on the east coast compared to only 14% along the west coast due to nutrient-rich deltas and suitable terrain in the former\textsuperscript{23}.

The five wetlands of India under study, viz. Gulf of Kutch (Gujarat), Thane Creek (Maharashtra), Godavari River delta (Andhra Pradesh), Bhitarkanika (Odisha) and Sundarbans (West Bengal) (Figure 1) have different climatic conditions. These regions were used as test sites to finalize the process of discriminating mangroves, as they represent the major mangroves of India.

Sundarbans is the world’s largest mangrove ecosystem spreading over India and Bangladesh, covering 10,000 sq. km with diverse mangrove species and other flora and fauna due to the rich nutrients received from the rivers Ganga and Brahmaputra\textsuperscript{24}. Bhitarkanika is the second largest mangrove ecosystem in India after Sundarbans. It flourishes in the deltaic region, formed by the rich alluvial deposits of Brahmani, Baitarani, Maipura and Dhamra rivers. About 62 species of mangroves can be found here, including three species of *Rhizophora*, *Heritiera* and *Avicennia* each and four of *Bruguiera*.

Kutch has the single largest (88%) patch of mangroves in Gujarat and the west coast of India. It has the maximum tidal amplitude of 5.54 m, dominated by single-species stands of *Avicennia marina* and a mangrove associate *Urochondra setulosa*\textsuperscript{25,26}. Godavari Basin is a major basin in Andhra Pradesh. Its delta is known for its fertile soil, which is favourable for cultivating paddy and sugarcane. The Godavari River is connected to Kakinada Bay by two major rivers, the Corangi and the Gaderu.

---

**Figure 1.** Map of the study area showing false colour composite of five mangroves – three along the east coast and two along the west coast used for methodology development. **a**, Gulf of Kutch, Gujarat; **b**, Bhitarkanika, Odisha; **c**, Sundarbans, West Bengal; **d**, Thane Creek, Maharashtra; **e**, Godavari Delta, Kakinada, Andhra Pradesh,
Data used and methodology

In this automation study, three types of data were used.

(i) Landsat 8 multispectral data (five bands with the mid wavelength of 0.56 µm, 0.655 m, 0.865 m, 1.61 m and 2.19 m and spatial resolution of 30 m; https://www.usgs.gov/) available from the GEE big data platform for the period between November 2019 and February 2020 for the entire coastal stretch of India were considered.

(ii) CartoDEM (version 3 R1) downloaded from ISRO’s Bhuvan website (https://bhuvan.nrsc.gov.in/) covering the Indian coast with a spatial resolution of 30 m was merged and uploaded in the GEE workspace. Digital elevation model (DEM) is a contributing factor as the mangroves are distributed till the tidal inundation from the coast.

(iii) Ground truth, Google Earth Image and field knowledge of the five sample study regions helped extract only the mangroves from the selected parameters and validate the methodology proposed.

The automation methodology included: (i) cloud masking of the selected data, (ii) extraction of textural parameters, indices and spectral geometry features, (iii) rule-based extraction of mangroves using backward feature elimination process, (iv) filtering for salt–pepper noise and (v) accuracy assessment (Figure 2).

(i) Bit-wise masking of clouds was used to remove cloud-covered regions from image collection over the study area with the help of the QA (quality assessment) band. QA bits indicate the pixels which are affected by surface conditions such as cloud cover/cloud shadow or sensor contamination, on which cloud masks are developed according to specifications (https://www.usgs.gov/land-resources/nli/landsat/landsat-collection-1-level-1-quality-assessment-band) to avoid anomalous results.

(ii) In GEE, GLCM computes the 12 metrics (how often pairs of pixels with specific values in a specified spatial relationship occur in an image) proposed in 1973 (ref. 27). NIR (band 5) sensitive to vegetation and shortwave infrared (SWIR) band (band 6 of L8 OLI) sensitive to wetland features were used to derive the GLCM parameters in order to differentiate mangroves from terrestrial land 28. The 24 GLCM parameters of both bands initially derived using a $5 \times 5$ kernel were (a) angular second momentum, (b) sum average, (c) contrast, (d) dissimilarity, (e) entropy, (f) correlation, (g) sum entropy, (h) homogeneity, (i) difference entropy, (j) variance, (k) sum variance and (l) difference variance.

In the automation process, three basic indices to differentiate vegetation, mangroves and water bodies, namely (i) enhanced vegetation index (EVI), (ii) normalized difference wetland vegetation index (NDWVI), and (iii) normalized difference water index (NDWI) 29,30 were used. NDWVI uses two narrow channels centred near the infrared and the SWIR bands. The SWIR band reflectance of mangroves is...
significantly lower than that of terrestrial vegetation, thus enhancing that enhances the discrimination of mangrove forests from terrestrial vegetation and making NDWVI better at segregating mangroves than NDVI.

Figure 3 shows the average reflectance values of mangroves, forests and agricultural vegetation at different wavelengths/bands (blue, green, red, NIR(A), SWIR1(B) and SWIR2(C)) of Landsat 8. Geometric properties like area, distance and shape are proven features to quantify the biochemical properties of vegetation and minerals present in the soil. It can be seen from Figure 3 that up to the red band, the geometry of the spectral signature lies closeby and after the NIR band, the geometry of the spectra varies significantly for all three types of vegetation. Thus, the spectral geometric parameters, namely length and area made by A, B and C and the angle between AB and BC were considered for further analysis.

(iii) At first, correlation coefficient between the parameters was used to reduce them into 12 (NDWVI, EVI and NDWI for indices; length, angle and area for geometric properties; contrast, dissimilarity, variance, sum variance and sum average for GLCM of SWIR1 and DEM). These 12 were further reduced into four essential parameters (EVI, elevation, NDWVI and dissimilarity) using backward feature elimination process (Figure 2). The multiplicative product of these four parameters with the threshold specified in Table 1 resulted into a binary layer, wherein 0 represents absence of mangroves and 1 represents presence of mangroves.

(iv) Salt-and-pepper errors were removed using morphological operations on the resultant binary layer. The contiguous mangrove pixels with less than eight counts were removed.

(v) OA and kappa coefficient of the mangroves were estimated using the proposed method by setting up fishnet grids for the entire coast of India (Figure 4). A landward buffer zone of up to 10 km was developed for the entire coast of India that was modified to accommodate the extent of mangroves by visual inspection. The zone was divided into grids of size 3 \times 3 km so that a sample area for validation covers 100 pixels \times 100 pixels of Landsat 8 data with 30 m resolution. Around 1000 random sample grids were selected, of which 200 were chosen manually to remove complete terrestrial land. These 200 sample points were verified for the presence and absence of mangroves using ground truth from field data and Google Earth images, and percentage was calculated for the entire coast of India.

### Results and discussion

To map the mangroves automatically without intervention from the user, parameters such as texture, indices and spectral geometry features derived from Landsat 8 OLI multispectral data and DEM were subjected to filtering by (i) visual analysis, (ii) correlation and (iii) BFE to separate the parameters of the highest importance, analysed and validated...
Figure 4. Random sample points (green dots) on regular grids (red) over the Indian coast for the validation of automated mapping.

Figure 5. Indices, geometrical properties and elevation of the Thane Creek, Maharashtra. a, Length; b, area; c, angle; d, EVI; e, NDWVI; f, NDWI; g, elevation.
for accuracy by generating random sample points on regular grids generated over the Indian coast.

**Elevation**

Mangroves are distributed along the Indian coast and mostly spread to the tidal inundation region. Maximum tidal amplitude was in the coastal states of Gujarat and West Bengal for about 11 m (ref. 33) and 6 m (ref. 34). In the present study, when the threshold of elevation was chosen as 11 m, many of the true mangroves that grew taller than 10 m in height (though they lie within 1 to 2 m ground level from the mean sea level) were also removed when SRTM data were used, whereas CartoDEM, the elevation data from Bhuvan, was able to retain such tall mangroves.

**Indices**

EVI, NDWVI and NDWI were the indices derived in this analysis, resulting in values ranging from –1 to 1. EVI helped discriminate vegetation from non-vegetation. Similar to NDVI, EVI also highlighted vegetation by values greater than 0.02 (Figure 5d). NDWVI was derived to make use of sensitivity of wetlands and thus mangroves to SWIR band. The same could be seen from Figure 5e showing the separation of the mangroves from neighboring features. NDWI was used to verify whether it could discriminate mangrove wetlands from water bodies which could not be resolved by NDWVI (Figure 5f). The threshold value for EVI was greater than 0.02, while for NDWVI and NDWI, was less than 0.

**Textural parameters**

Twelve grey-level co-occurrence matrix parameters were derived for NIR and SWIR1 bands (Figure 6). However, dissimilarity, variance, contrast, sum average and sum variance of the SWIR1 band were chosen based on the correlation measure. The remaining parameters did not show any marked difference between mangroves and other adjoining features.
Table 2. Accuracy assessment of backward feature elimination (trials with accuracy only above 50%)

<table>
<thead>
<tr>
<th>Trials/parameters</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>NDWI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DEM</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>NDWVI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Dissimilarity</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Variance</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Contrast</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Sum variance</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Sum average</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Area</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Length</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Angle</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>59.47</td>
<td>59.47</td>
<td>59.47</td>
<td>78.94</td>
<td>81.05</td>
<td>81.05</td>
<td>83.16</td>
<td>84.21</td>
<td>78.42</td>
<td>82.11</td>
<td>82.63</td>
<td>86.05</td>
<td>81.05</td>
</tr>
</tbody>
</table>

Figure 7. Accuracy of the estimates of the proposed model to automatically map the mangroves of the five study areas.

Geometric parameters

Mangroves could be identified using the parameter length and differentiated from adjoining terrestrial vegetation with brighter shade (Figure 5a) and using the parameter area with still bright shade (Figure 5b) while parameter angle has darker shade to distinguish mangroves from other vegetation categories. The threshold value for length was greater than 0.15, for area greater than 0.001, and for angle, it was less than $179.6^\circ$.

Each parameter has its own significance in distinguishing mangroves, and the proposed method helps fix a standard threshold for each parameter and combines them to find suitable parameters in discriminating mangroves from neighbourhood land-cover categories.

Automated mangrove mapping based on rule set

The first three columns in Table 2 show the accuracy levels when one or more geometric parameters are considered. Besides the indices and elevation, adding two textural parameters, viz. dissimilarity and variance, yielded more accuracy (83.16%); the removal of variance further improved the accuracy to 84.21% (Table 2; columns 7 and 8). Continuing the elimination of individual indices resulted in a maximum of 86.05% accuracy among all probable combinations when NDWI was removed, retaining EVI and NDWVI along with elevation and dissimilarity. Thus the resulting four optimum parameters discussed were EVI, elevation, dissimilarity and NDWVI (Table 2; column 12).

While considering the accuracy of the methodology for the five study regions individually, maximum accuracy was obtained for Thane Creek (OA 94.4%, $k = 0.886$) and minimum for Godavari delta (OA 81.25%, $k = 0.6364$) (Figure 7). The reduction in accuracy in the Godavari delta could be due to the similarity in spectral behaviour of mangroves with the deltaic wetland paddy being cultivated close to the marshy wetland region.

Figure 8 shows the essential parameters over all five study regions. Elevation values below 11 m retains coastal area under tidal influence thus mangroves also and removes the other region. EVI above 0.2 separates vegetation from water bodies, built-up land, wasteland and other in-land features. Dissimilarity separates wetlands, and land-cover features with similar textural properties (including water bodies and agricultural fields) from the others, which have different textural properties with values below 150. NDWVI differentiates between wetlands and other inundated land-cover
features (like agriculture). Dissimilarity is a subset of NDWVI for Mumbai, but the reverse is true for Bhitarankanika and showing an entirely different pattern for the Gulf of Kutch (Figure 8, last two columns).

The proposed algorithm was used to assess mangroves of the entire Indian coast, which resulted in a 4975 sq. km area (Figure 9). The accuracy for the overall Indian coast was found to be 86.05%, i.e. 26 out of 190 points were ‘false’, and 164 out of 190 points were ‘true’ in their ‘accuracy’. The kappa coefficient value of 0.722 could be considered as ‘substantial’.

**Conclusion**

This study contributes to the automatic mapping of mangroves from Landsat 8 OLI multispectral satellite images.
available in GEE, a big data platform. It shows that a combination of binary data derived using a single threshold from Cartosat DEM, NDWVI and EVI is useful in mapping the extent of mangroves and distinguishing them from other terrestrial and associated vegetation on the entire coast of India with significantly better accuracy. Adding one of the textural parameters of SWIR (band 6), viz. dissimilarity, improves the accuracy from 84.21% to 86.05%. Geometrical parameters were not good for extracting mangroves, though they could distinguish them from other vegetation types using site-specific thresholds. Since the kappa coefficient of 0.722 is substantial, the study can be extended to other mangroves of Southeast Asia to arrive at a preliminary view. In Participatory Mangrove Management in a Changing Climate (eds DasGupta, R. and Shaw, R.), Disaster Risk Reduction Springer, Tokyo, Japan, 1997.


Received 18 October 2022; revised accepted 1 May 2023

doi: 10.18520/cs/v125i3/299-308