

Bamboo Distribution Mapping using Hyperspectral Remote Sensing in Ri Bhoi District of North East India

Dhruval Bhavsar*, Kasturi Chakraborty, Sarma K.K, & Aggarwal S.P

North Eastern Space Applications Centre, Dept. of Space, Umiam, Meghalaya, India

*Corresponding Author's email: dhruval.bhavsar@nesac.gov.in

Abstract

Bamboo, renowned as the fastest-growing terrestrial plant, plays a pivotal role in enhancing carbon sequestration, ecosystem dynamics and livelihoods. India is the world's second-largest bamboo producer, with North Eastern states contributing over 50% of the total output in national bamboo production. The major challenge in mapping bamboo distribution is due to time-consuming traditional visual interpretation methods and differentiating it from mixed forest composition. This study aims to leverage hyperspectral remote sensing to enhance the accuracy of classifying bamboo amidst mixed forest compositions. The study area encompasses the Nongkhylllem reserve forest and its surroundings in Ri Bhoi district, Meghalaya, characterized by a predominance of bamboo and mixed moist deciduous forest. Combining spectral (vegetation indices) and texture-based layers, alongside aspect (Carto DEM), a comprehensive input set was prepared for the machine learning-driven random forest classifier. Bio-chemical based vegetation indices along with principal component layers were found to be most important variables for bamboo classification. The outcome of the classification process showed 72% accuracy, which has further scope of improvement using high spatial resolution hyperspectral data. Notably, around 22% of the study area was categorized as pure bamboo growth area, while 6% fell under the mixed bamboo class. This study underscores the value of hyperspectral data and advanced machine learning tools in accurately demarcating bamboo-rich regions. The outputs hold significant promise for sustainable management and Sustainable Development Goals (SDG). As India embarks on ambitious projects to elevate bamboo productivity, the hyperspectral remote sensing-based mapping of bamboo growing area, could be pivotal in steering smart governance strategies, aligning with the government's larger objectives.

Keywords PRISMA hyperspectral data, Bamboo, Multisource image database, Random Forest

Introduction

Bamboo, renowned as the fastest-growing terrestrial plant, plays a pivotal role in enhancing carbon sequestration, ecosystem dynamics and livelihoods. Bamboo is also referred as "Green Gold," because of its economic importance in diverse applications ranging from construction to handicrafts. Bamboo has been an essential resource for rural livelihoods and sustainable development. Bamboo is being used as a source of renewable energy, either as biomass for power generation or as a fuel for cooking and heating. Recognizing the immense potential of bamboo, the Indian government has undertaken various policy initiatives to promote its sustainable utilization and conservation. One of the most significant developments in this regard is the National Bamboo Mission (NBM, 2023), which was launched in 2018 as a component of the National Mission for Sustainable Agriculture

(NMSA). Considering the importance of bamboo resources, its mapping and monitoring is of utmost importance. The NBM aims to harness the potential of bamboo as a resource for enhancing livelihood opportunities and rural income, especially in the North East region of India. The region's topography, characterized by dense forests, hilly terrain, and abundant rainfall, provides an ideal environment for bamboo growth and boasts an extraordinary diversity of bamboo species. About 89 bamboo species out of 126 recorded in India under 16 genera grow naturally in different forest types of this region or are cultivated across its tropical and sub-tropical belts.

Remote sensing has been emerged as an important tool to map and monitor vegetation over a period of time using different techniques and method. Remote sensing satellites can provide images of the Earth's surface at different wavelengths, which can be used to identify and map bamboo forests. Several studies have been carried out using remote sensing technology to map and monitor, vegetation cover in North East India (Lele and Joshi, 2009; Roy et al., 2015; Roy and Joshi, 2022) while, very few attempts have been made to identify and map bamboo growing area in the region (Goswami et al., 2010). There are studies carried out globally and in India to map bamboo forest (Ghosh and Joshi, 2014; Du et al., 2018; Nfornekah et al., 2020) however, these studies mainly involve the multispectral satellite data using traditional classification techniques like visual interpretation and traditional classifiers. The major gap lies with such technique in bamboo mapping is the time consumption and poor classification accuracy. Hyperspectral remote sensing is one of the techniques which have been used extensively for discrimination of vegetation species with advance classifiers (Jha et al., 2019; Kishore et al., 2020). However, the information on bamboo forest mapping is lacking. The present study was aimed to utilize the hyperspectral remote sensing imagery to discriminate bamboo growing areas from the mixed vegetation composition. Machine learning algorithm was used for classification of bamboo forest.

Materials and Methods

Study Area: The present study was conducted in the Tropical Moist Deciduous Forest of North East India, situated between 25.97°N and 91.81°E to 25.78°N and 91.92°E in the Ri Bhoi district of Meghalaya (Figure 1). The study area encompasses the Nongkhyllem Reserve Forest, characterized by its hilly and rugged terrain with steep slopes. The study area also includes Nongpoh town, serving as a district headquarters. Due to disturbances in the forest caused by human activities in populated areas, bamboo growth has been observed. The major bamboo species in the study area include *Dendrocalamus strictus*, *Bambusa tulda*, and *Melocanna baccifera*. The climate of the area is tropical monsoon, featuring hot and wet summers and cool and dry winters. The average annual rainfall is approximately 2,500 mm (Forests & Environment Department, Government of Meghalaya, 2023).

Materials used: In this study, a comprehensive dataset was employed, comprising satellite data, ancillary data, and ground truth data to facilitate a thorough analysis. The spatial resolution of the Carto DEM (Digital Elevation Model) used in the study was 10 meters, providing detailed topographic information. The hyperspectral data utilized in the study was PRISMA (PRecursore IperSpettrale della Missione Applicativa) satellite data of March 6, 2022 downloaded from the PRISMA mission website (PRISMA, 2022). The specification of PRISMA data is shown in Table 1. Additionally, ancillary data in the form of the Forest Survey of India

(FSI) forest cover map was incorporated. Ground truth data played a crucial role, providing location information pertaining to the various species present in the study area. This multi-faceted approach, integrating diverse data sources, aimed to enhance the precision and reliability of the study's findings.

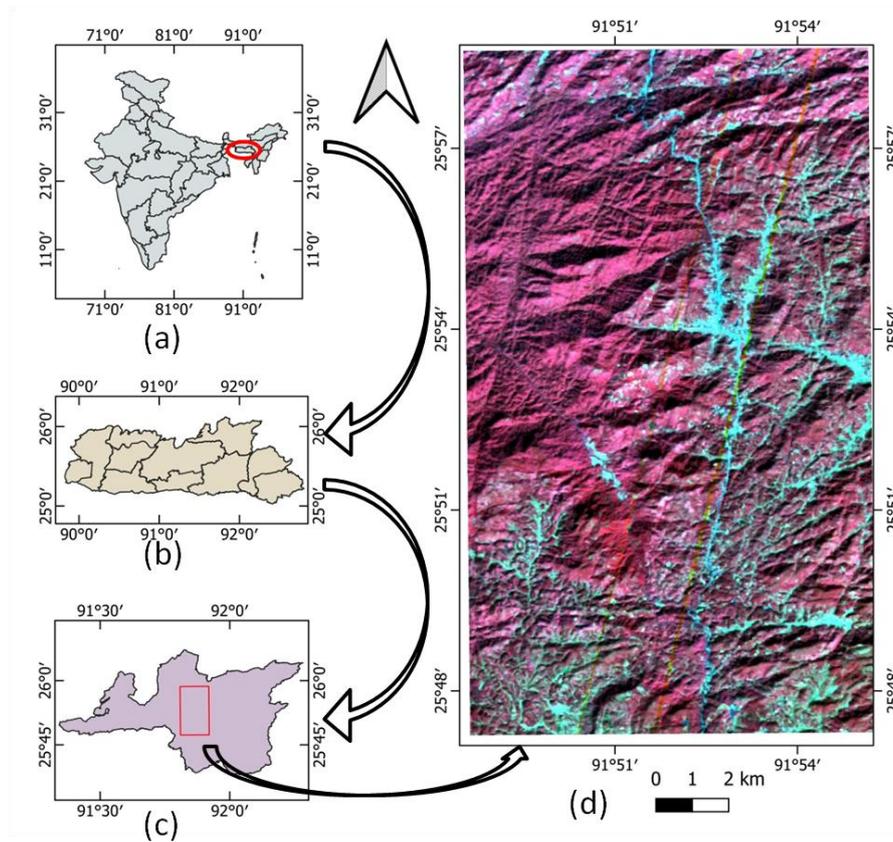


Fig. 1 Geographic location of study area (a) Meghalaya state in India (b) Districts in Meghalaya state (c) study area in Ri Bhoi district of Meghalaya (d) False Color Composite of PRISMA hyperspectral data.

Table 1. Specification of PRISMA hyperspectral sensor (Coppo et al., 2019).

Sr. No.	Description	Value
1	Scene size	30 Km
2	Pixel size nadir	30 M
3	FOV	2.4 Degrees
4	Spectral range—VNIR	400–1010 Nm
5	Spectral range—SWIR	920–2505 Nm
6	Spectral range—PAN	400–700 Nm
7	Spectral resolution—VNIR	≤13 Nm
8	Spectral resolution—SWIR	≤14.5 Nm
9	Spectral resolution—PAN	≤13.5 Nm
10	Number of spectral bands— VNIR	66
11	Number of spectral bands— SWIR	174
12	Spatial resolution—VNIR-SWIR	30 m/px
13	Spatial resolution—PAN	5 m/px
14	SNR—VNIR	>160 (>450 at 650 nm)
15	SNR—SWIR	>100 (>360 at 1550 nm)
16	SNR—PAN	>240
17	Absolute radiometric accuracy	Better than 5%

Methodology: The study harnessed Level 2D PRISMA data, which is a geocoded atmospheric corrected product. Leveraging the prismaread package (Busetto, 2020) within the R environment developed for converting and importing the hyperspectral data, both the hyperspectral and panchromatic bands were extracted. The study area of interest was then clipped from these datasets. The extraction process involved an intricate fusion of information from the Forest Survey of India (FSI) forest cover map and the Normalized Difference Vegetation Index (NDVI) derived from hyperspectral data. This dual-source approach enabled the isolation of the vegetated area, which subsequently is the focal point for further analysis. For the mapping of bamboo growing area, input data is very crucial. In the present study different parameters viz. dimensionality reduction, Vegetation Index, texture, topography and spectral properties were considered and used for bamboo mapping.

Dimensionality Reduction: As hyperspectral data has lot of information which is also embedded with the noise. Recognizing the vastness of hyperspectral data, Principal Component Analysis (PCA) was employed for dimensionality reduction. It is the most commonly used statistical approach for variable selection in highly correlated system (Navalgund and Ray, 2011). This step was crucial in distilling essential information from the entire dataset while retaining its inherent complexity.

Vegetation Index: A spectrum of vegetation indices, reflective of both plant structure and biochemical components, were meticulously applied. The rationale behind each index's selection and its significance in characterizing vegetation dynamics within the study area were considered. A detailed list of these indices is thoughtfully presented in Table 2. All the VI layers were preceded for the correlation analysis to reduce the redundancy in the data.

Table 2. List of Vegetation indices used in the preset study.

Sr. No.	Vegetation Index	Description	Formula	Reference
1	EVI	Enhanced Vegetation Index	$2.5 * ((R800 - R670) / (R800 - (6 * R670) - (7.5 * R475) + 1))$	Huete et al., 1997
2	gNDVI	Green NDVI	$(R750 - R550) / (R750 + R550)$	Datt, 1998
3	hNDVI	hyperspectral NDVI	$(R827 - R668) / (R827 + R668)$	Oppelt, 2002
4	LCI	Leaf Chlorophyll Index	$(R850 - R710) / (R850 + R680)$	Datt, 1999
5	LWVI1	Leaf Water Vegetation Index	$(R1094 - R983) / (R1094 + R983)$	Galvo et al., 2005
6	MCARI	Modified Chlorophyll Absorption in Reflectance Index	$((R700-R670) - 0.2*(R700 - R550))*(R700 / R670)$	Daughtry et al., 2000
7	MCARI1	Modified Chlorophyll Absorption in Reflectance Index	$MCARI1=1.2*(2.5*(R800-R670)-1.3*(R800-R550))$	Haboudane et al., 2004
8	MCARI2	Modified Chlorophyll Absorption in Reflectance Index	$MCARI2=1.5*(2.5*(R800-R670)-1.3*(R800-R550)) / ((2*R800+1) ^2 - (6*R800-5*R680^0.5)-0.5) ^ 0.5$	Haboudane et al., 2004
9	MSAVI	Modified Soil Adjusted Vegetation Index	$0.5 * (2 * R800 + 1) - ((2 * R800 + 1) ^2 - 8 * (R800 - R670)) ^ 0.5$	Qi et al., 1994
10	MSI	Moisture Stress Index	$R1600/R817$	Hunt et al., 1989
11	MTVI	Modified Triangular Vegetation Index	$1.2 * (1.2 * (R800 - R550) - 2.5 * (R670 - R550))$	Haboudane et al., 2004
12	MTVI2	Modified Triangular Vegetation Index	$1.5*(1.2*(R800-R550)-2.5*(R670-R550)) / ((2*R800+1) ^2 - (6*R800 - 5*(R670 ^ 0.5)) -0.5) ^ 0.5$	Haboudane et al., 2004
13	NDLI	Normalized Difference Lignin Index	$(\log(1/R1754) - \log(1/R1680)) / (\log(1/R1754) + \log(1/R1680))$	Serrano et al., 2002
14	NDNI	Normalized Difference Nitrogen	$(\log(1/R1510) - \log(1/R1680)) /$	Serrano et

15	NDRE	Index Normalized Difference Red Edge Index	$(\log(1/R1510) + \log(1/R1680)) / (R800 - R680) / (R800 + R680)$	al., 2002 Sims & Gamon, 2002
16	NDVI	Normalized Difference Vegetation Index	$(R800 - R680) / (R800 + R680)$	Datt, 1999
17	NPCI	Normalized total Pigment to Chlorophyll Index	$(R680 - R430) / (R680+R430)$	Peñuelas et al., 1994
18	OSAVI	Optimized Soil Adjusted Vegetation Index	$(1 + 0.16) * (R800 - R670) / (R800 + R670 + 0.16)$	Rondeaux et al. 1996
19	PRI	Photochemical Reflectance Index	$(R531 - R570) / (R531 + R570)$	Gamon et al., 1992
20	RDVI	Renormalized Difference Vegetation Index	$(R800 - R670) / (R800 + R670) ^ 0.5$	Roujean & Breon 1995
21	SAVI	Soil Adjusted Vegetation Index	$(1 + 0.5) * (R800 - R670) / (R800 + R670 + 0.5)$	Huete, 1988
22	SIPI	Structural Independent Pigment Index	$(R445 - R800) / (R670-R800)$	Peñuelas & Filella, 1999
23	TCARI	Transformed Chlorophyll Absorption in Reflectance Index	$3 * ((R700 - R670) - 0.2 * (R700 - R550) * (R700/R670))$	Haboudane et al., 2002
24	TVI	Triangular Vegetation Index	$0.5 * (120 * (R750 - R550) - 200 * (R670 - R550))$	Broge & Leblanc, 2000

Texture Analysis: To capture spatial intricacies, the panchromatic band underwent a detailed texture analysis was carried out using Gray-Level Co-occurrence Matrix (GLCM). It analyses an image's texture by calculating pairs of pixels with specific values and spatial relationship. Statistical measures, including mean, correlation, variance, homogeneity, contrast, entropy, second moment, and dissimilarity, were computed, providing a nuanced understanding of the textural features within the imagery.

Acknowledging the nuanced relationship between vegetation growth and geographic factors, an aspect layer, derived from cartographic Digital Elevation Model (DEM), was incorporated. This layer played a pivotal role in discerning species distribution and ecological variations within the study area. Given the inherent differences in spatial resolutions among diverse data layers, a meticulous resampling process ensued. All layers were standardized to 5m resolution, the highest among them, to ensure uniformity and prevent loss of critical information during subsequent analyses. Post-resampling, the layers were methodically stacked to create an integrated and harmonized database. This database formed the foundation for subsequent classification and mapping endeavors. The classification strategy was multifaceted, exploring various combinations of datasets. This included spectral reflectance, PCA, visible (Vis) layer, PCA and Vis integration, and PCA+Vis+texture layers. The objective was to discern the most informative layers conducive to accurate bamboo mapping.

Real-world validation was ensured through the collection of ground truth (GT) data, encompassing diverse species and vegetation types. This dataset was then judiciously split into a 70:30 ratio, with 70% earmarked for model training and the remaining 30% for testing and validation. From the GT data, pure pixels in the image were identified using Pixel Purity Index (PPI). The output of the PPI was analyzed using n-Dimensional visualize and endmember of the spectrally pure pixels were identified. The endmembers were subsequently used for the classification purpose. Diverse classification algorithms, such as Spectral Angle Mapper (SAM), Support Vector Machine (SVM), and Random Forest (RF),

were systematically employed. The bootstrap method facilitated a robust assessment of each classifier, with the ultimate selection being guided by the minimization of classification errors. This rigorous approach ensured the selection of the most suitable classifier and best dataset for bamboo mapping. The accuracy of the bamboo mapping results was meticulously evaluated using the reserved 30% of the ground truth data. This step provided a robust validation of the classification outcomes and affirmed the reliability of the selected classifier for accurate bamboo distribution mapping within the study area. The broad outline of the methodology is presented in Figure 2.

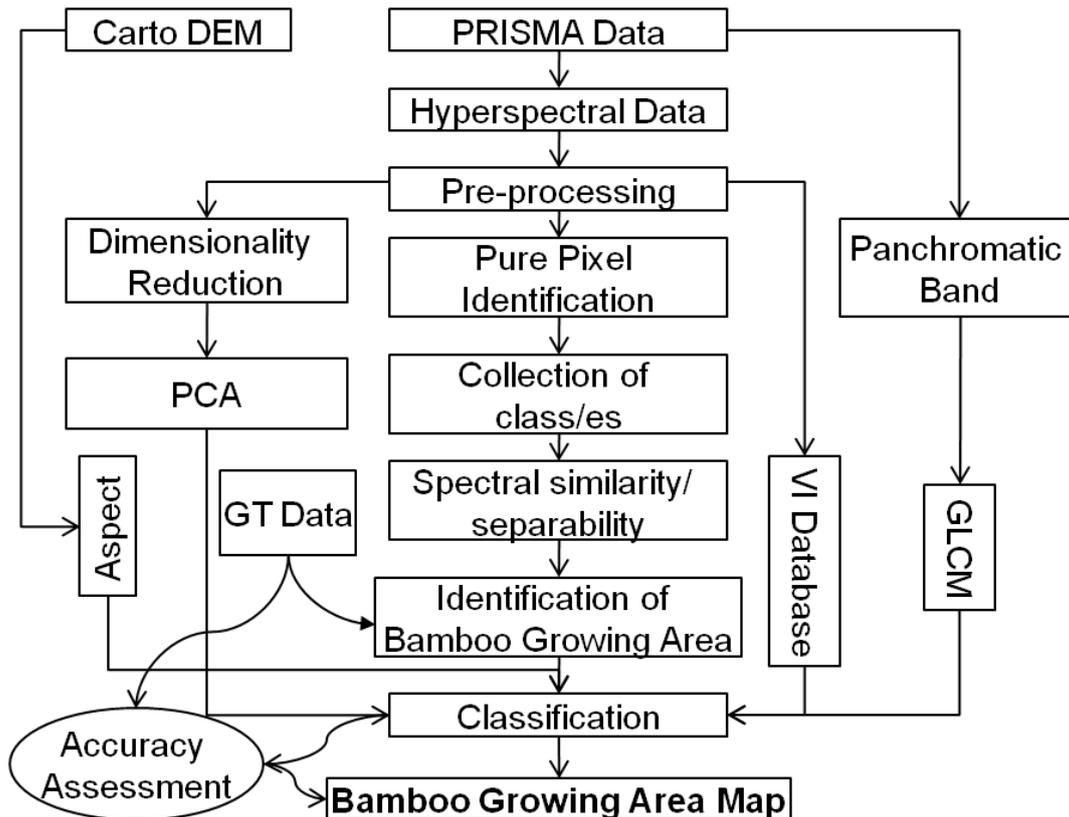


Fig. 2 Broad outline of the methodology used in the study.

Results

The several outcomes and finding generated during the study is as follows:

Hyperspectral Dimensionality Reduction: Principal Component Analysis (PCA) plays a crucial role in visualizing hyperspectral data to extract detailed spectral information. In the current study, the results of PCA revealed significant information within the first two bands. The Eigenvalue curve, depicted in Figure 3, demonstrates a sharp decline, approaching zero, beyond the second band. Examining Figure 4a and 4b, which illustrate PC1 and PC2 respectively, a distinct set of features emerges for each band. PC1 distinctly accentuates vegetated areas, showcasing its effectiveness in capturing spectral signatures related to vegetation. Conversely, PC2 accentuates non-vegetated regions, predominantly highlighting roads, settlements, and wastelands. The observed diminishing Eigenvalues after the second band suggest that subsequent principal components contribute minimally to the overall variance of the hyperspectral data. This highlights the efficiency of limiting the analysis to

the first two principal components for meaningful insights into the spectral characteristics of the data.

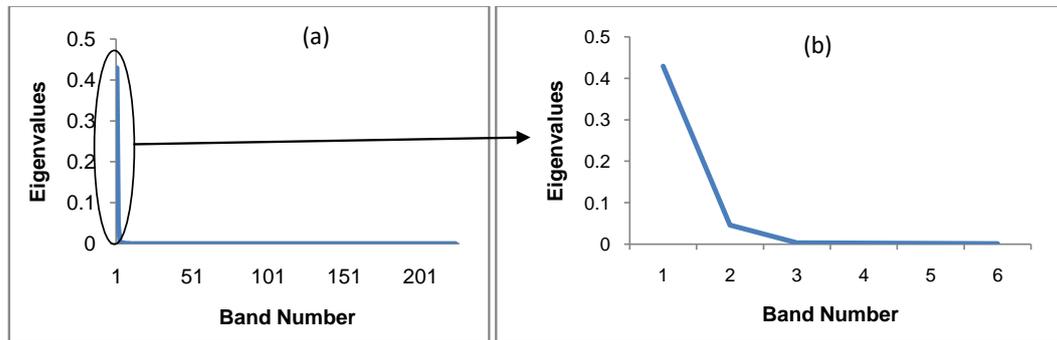


Fig. 3 Spectral curve of Eigen values derived from PCA run on the hyperspectral data (a) full range spectra with all the bands and (b) highlighted and enhanced spectra where high Eigen values are present up to 2 bands.

Identification of pure pixel: Ground truth data were systematically collected within the study area to identify various Land Use Land Cover (LULC) classes, with corresponding locations meticulously recorded. The diverse LULC classes encompassed in this study include forest, bamboo, plantation, agriculture, water, urban, and non-forest categories. Notably, the identified forest in our study is of the semi-evergreen type, further categorized into two classes: forest type 1 and forest type 2, based on distinct species composition and dominance. Forest type 1 comprises species such as *Terminalia chebula*, *Adina cordiafolia*, *Aesculus assamica*, and *Tectona grandis*, with *Adina cordiafolia* and *Albizia lebeck* being the predominant species. On the other hand, forest type 2 consists of species like *Schima wallichii*, *Shorea robusta*, *Gmelina arborea*, *Tectona grandis*, *Lagerstroemia parviflora*, and *Albizia lebeck*, with *Shorea robusta* as the dominant species. The plantation category in the study area primarily consists of *Areca catechu*, which is prevalent but coexists with agriculture and urban land use classes.

To distinguish these classes in the data, Minimum Noise Fraction (MNF) analysis was executed, and the results are portrayed in Figure 5(a). Five MNF bands exhibited minimal noise, featuring high eigenvalues, while the remainder demonstrated a subsequent decrease. Leveraging these MNF bands, Pixel Purity Index (PPI) was applied, and endmembers (pure pixels) were visualized in the feature space using n-Dimensional visualizer, as depicted in Figure 5(b). The color-coded endmembers effectively illustrate the separability among different classes. Notably, non-vegetated classes such as non-forest/barren, water bodies, and urban features demonstrated distinct separations. Additionally, forest type 1 and forest type 2 exhibited clear separability using the identified endmembers. The pure bamboo class in the feature space displayed a close association with agriculture and mixed bamboo. Furthermore, the mixed bamboo classes were observed to be linked with forest type 1, aligning with the ground truth data. In the classification process, the endmembers assigned to a particular class served as the training input, ensuring a robust and accurate classification of the hyperspectral data.

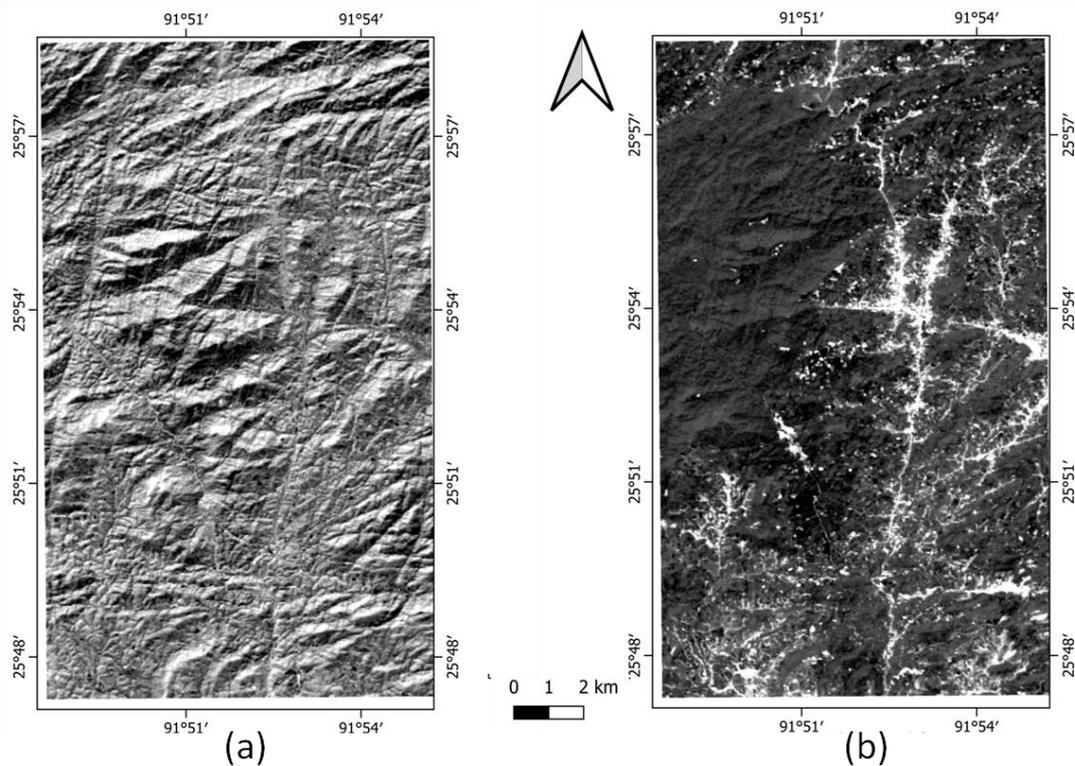


Fig. 4 Principal components generated from hyperspectral data (a) PC 1 highlighting vegetation (b) PC 2 – non vegetated area appearing in white color.

Vegetation Index Database: In the current study, a total of 24 vegetation indices (VIs) were computed and subsequently correlated to assess redundancy among them. The exclusion criteria for redundancy were set at a correlation value exceeding 0.5. The correlation coefficients between pairs of layers were meticulously analyzed, and layers displaying multiple high correlations with others were omitted from further processing. The correlation heatmap, presented in Figure 6, numerically designates VI layers from 0 to 23, with diagonal values indicating perfect correlation (1) as they represent the same layer. The results revealed that 10 out of the 24 layers exhibited high correlation and were consequently excluded from subsequent analyses. The 14 remaining VIs, namely hNDVI, NDVI, MTVI, NDRE, EVI, MSI, LCI, MCARI, TCARI, NDLI, NDNI, SIPI, SAVI, and OSAVI, were identified as less correlated or unique. Figure 7 illustrates a false-color composite generated based on these shortlisted VIs, providing an overview of different feature classes within the study area. From the VI generated FCC, non-vegetated areas are distinctly highlighted in blue, attributed to the Multispectral Stress Index (MSI), which effectively accentuates stressed regions. The variation in green tones is particularly noteworthy, influenced by the red-edge band, a crucial element for discriminating between different plant species.

Classification and Accuracy Assessment: In the current study, a supervised classification technique was used, involving the testing of different combinations of classifiers and input layers. The classifiers were applied to the training data using a 10-fold cross-validation process. Evaluating the spectral reflectance of features from hyperspectral data, the Spectral Angle Mapper (SAM) exhibited the highest accuracy at 53.01% compared to other methods.

For Principal Component Analysis (PCA), the Support Vector Machine (SVM) achieved the highest accuracy, reaching 49.45%. Notably, the Random Forest classifier demonstrated the highest accuracy when the number of bands was increased, incorporating Vegetation Indices (VIs), a combination of PCA and VI, and a combination of PCA, VI, aspect and texture. The most accurate classification, reaching 72.19%, was achieved by combining texture, vegetation indices, aspect and PCA, utilizing the Random Forest classifier. A detailed breakdown of the accuracy of different classifiers using various input layers is provided in Table 3.

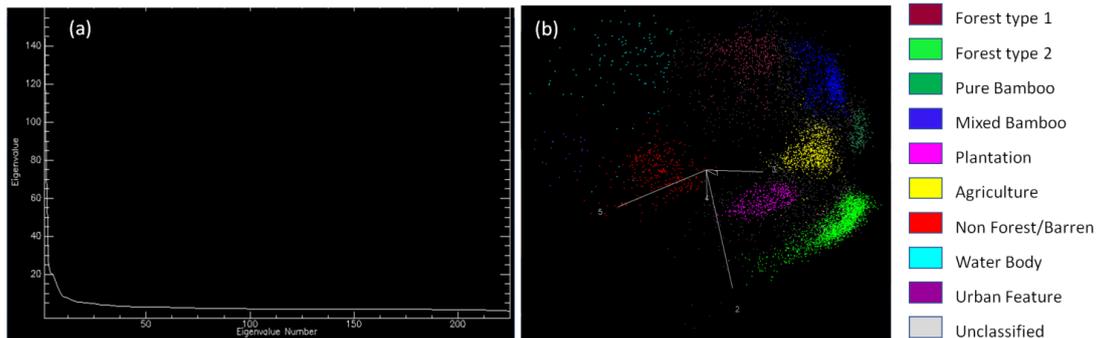


Fig. 5 (a) Minimum Noise Fraction curve derived from hyperspectral data (b) Endmember (pure pixel) visualization in the feature space and its separability among the classes.

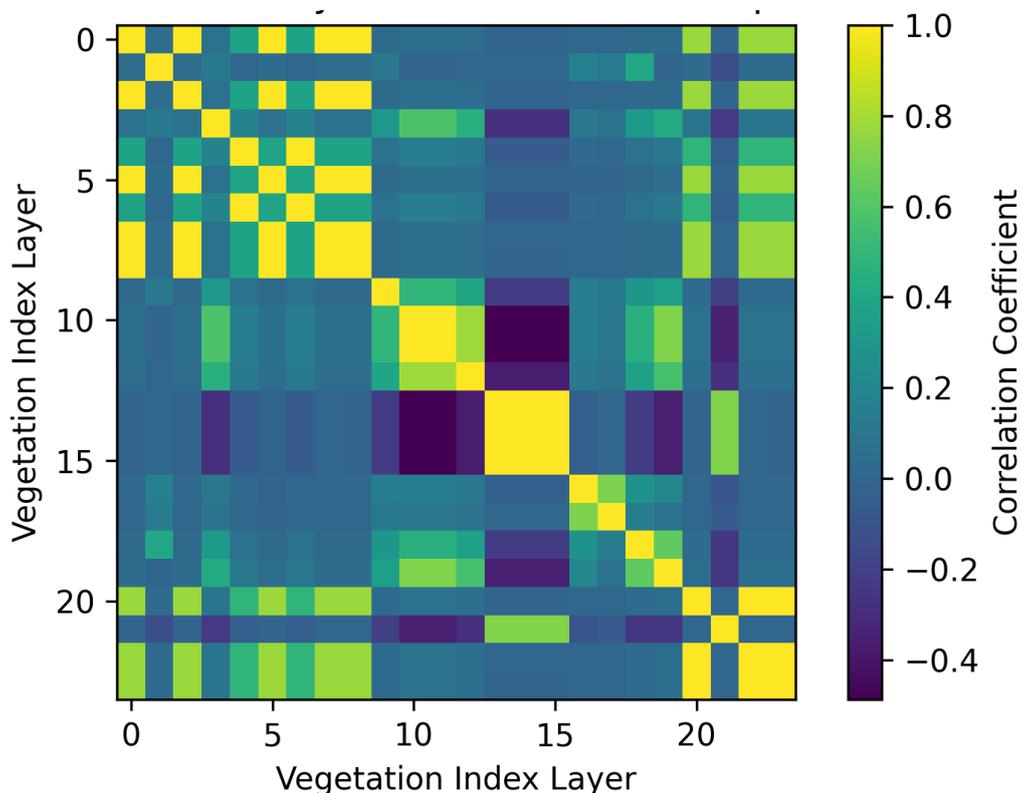


Fig. 6 Vegetation index correlation matrix for 24 layers (0 to 23). The vegetation index corresponds to each number is 0 (gNDVI), 1 (hNDVI), 2 (TVI), 3 (NDVI), 4 (RDVI), 5 (PRI), 6 (MTVI), 7 (MTVI2), 8 (LWVI1), 9 (NDRE), 10 (EVI), 11 (MSI), 12 (LCI), 13 (MCARI), 14 (MCARI1), 15 (MCARI2), 16 (TCARI), 17 (NDLI), 18 (NDNI), 19 (SIPI), 20 (NPCI), 21 (SAVI), 22 (MSAVI), 23 (OSAVI).

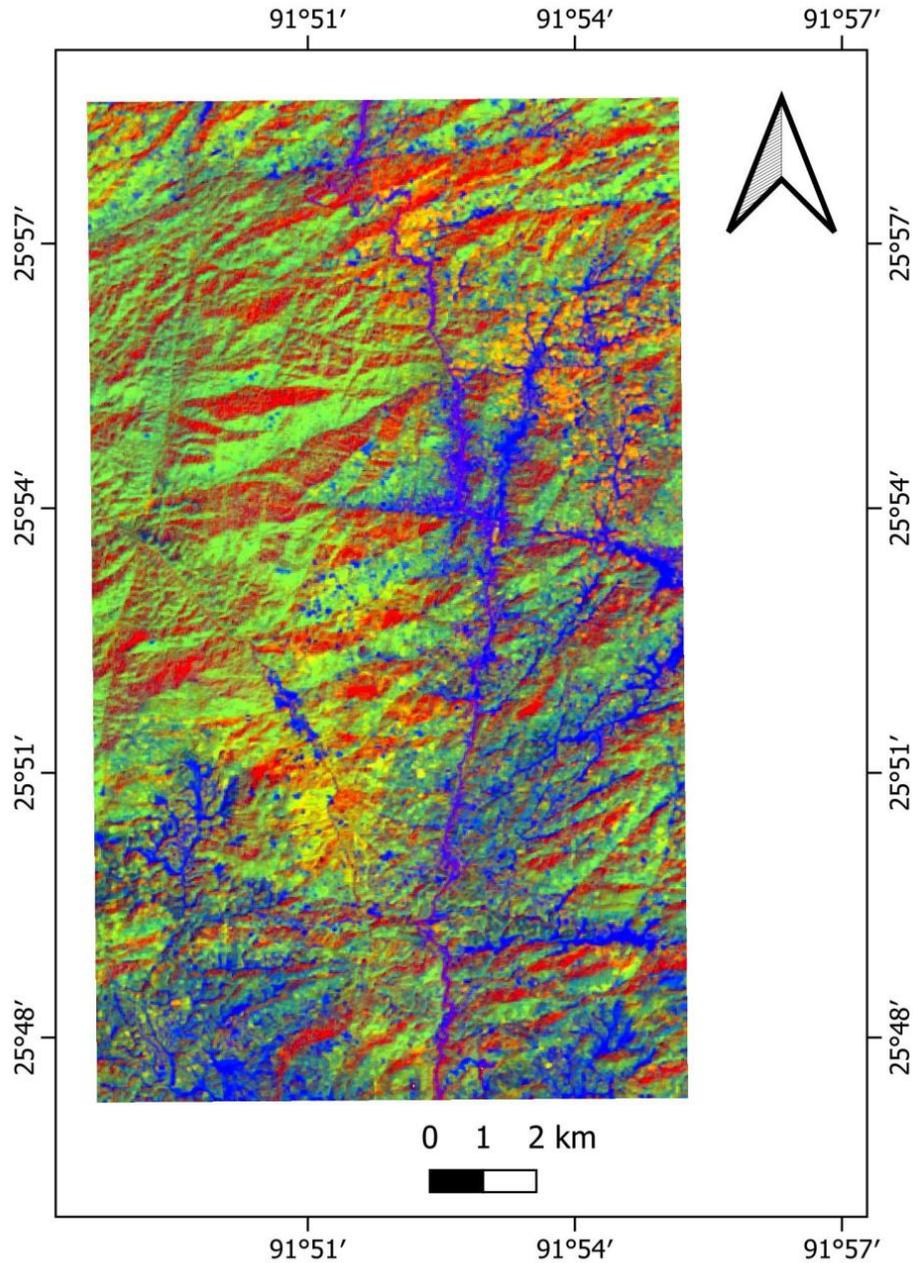


Fig. 7 False Color Composite (FCC) of study area using different vegetation index (R: NDVI, G: NDRE, B: MSI).

Table 3. Accuracy (%) of different classifiers using various input from multisource database.

Sr. No.	Input Layer	Classifier		
		SAM	SVM	RF
1	Spectral (Hyperspectral)	53.01	51.52	46.21
2	PCA	42.58	49.45	44.12
3	VI database	58.69	61.16	63.36
4	PCA+VIs	63.45	65.94	68.28
5	PCA+ VIs + texture + aspect	65.85	70.45	72.19

Using the trained Random Forest classifier, the mapping of different Land Use Land Cover (LULC) classes was performed. The resulting classified map, encompassing all classes, is presented in Figure 8. The map illustrates that bamboo-growing areas are predominantly located in the reserve forest area, while mixed bamboo classes are mapped in the vicinity of urban and disturbed areas. The lower portion of the study area is primarily composed of agricultural land, with Forest Type 2 dominated by *Shorea robusta*. The largest area in the study area is occupied by Forest Type 1 (83.56 km²), followed by Forest Type 2 (59.99 km²). Bamboo-growing areas cover 22.85 km², while mixed bamboo occupies 5.85 km². The comprehensive distribution of different land cover classes in the study area is summarized in Table 4. These classification results not only provide a detailed understanding of the spatial distribution of various land cover types but also highlight the effectiveness of incorporating texture, vegetation indices, and PCA in improving classification accuracy, particularly when utilizing the Random Forest classifier.

Identification of important variable and bamboo growing area mapping: The classification results revealed that the integration of multisource data is valuable for discriminating bamboo classes. To further enhance the study's precision, a variable importance analysis was conducted, as illustrated in Figure 9. The analysis identified PC 2, MCARI, NDRE, and TCARI as the most crucial variables for effectively segregating the classes. This underscores the significance of leaf pigments, particularly chlorophyll, in the discrimination of vegetation types. The wavelengths utilized in these indices (550nm, 670nm, 680nm, 700nm, 800nm) further emphasize the importance of specific spectral bands in distinguishing bamboo classes.

Table 4. Area (km²) under different LULC classes of the study area.

Sr. No.	LULC Class	Area
1	Forest type 1	83.56
2	Forest type 2	59.99
3	Bamboo	22.85
4	Mixed Bamboo	5.85
5	Plantation	5.93
6	Agriculture	43.08
7	Water body	2.57
8	Urban	6.87
9	Non forest/Barren	9.30

Moreover, structural indices such as MTVI, hNDVI, along with the Moisture Stress Index (MSI), demonstrated notable importance in vegetation discrimination. The inclusion of these indices signifies the role of structural characteristics and stress indicators in differentiating between various vegetation types. Additionally, the mean of texture analysis emerged as an important factor in bamboo mapping, providing insights into the spatial patterns and arrangements of bamboo vegetation.

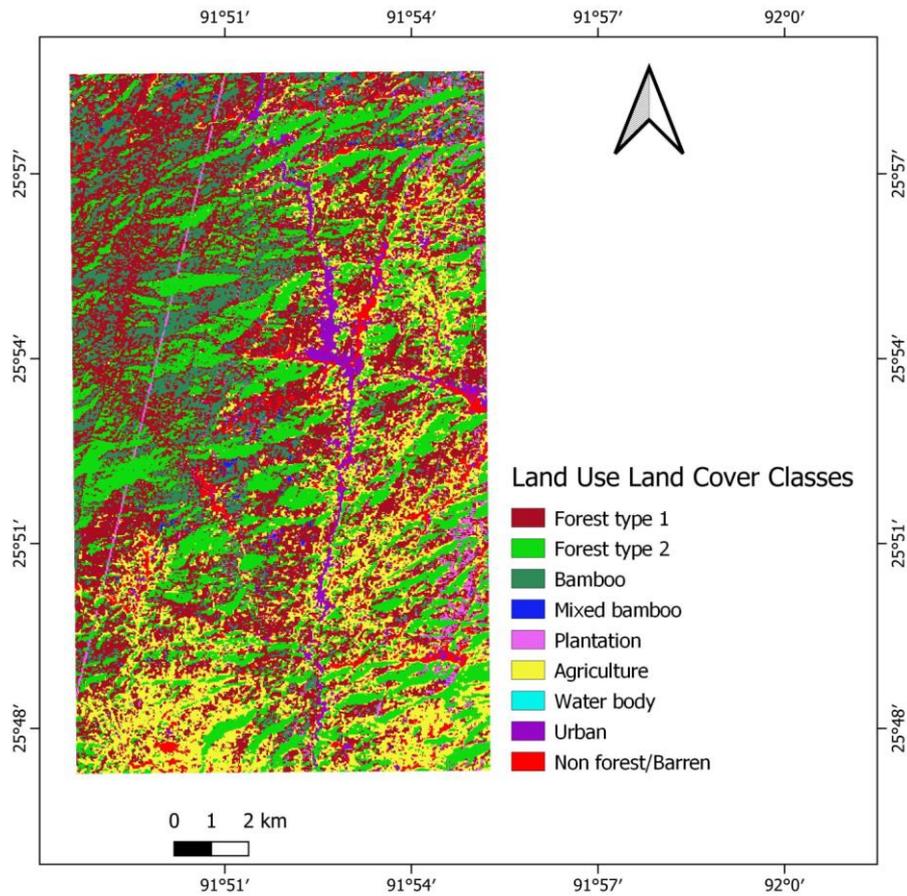


Fig. 8 Classified LULC map using combined database of PCA, VI, aspect and texture (GLCM).

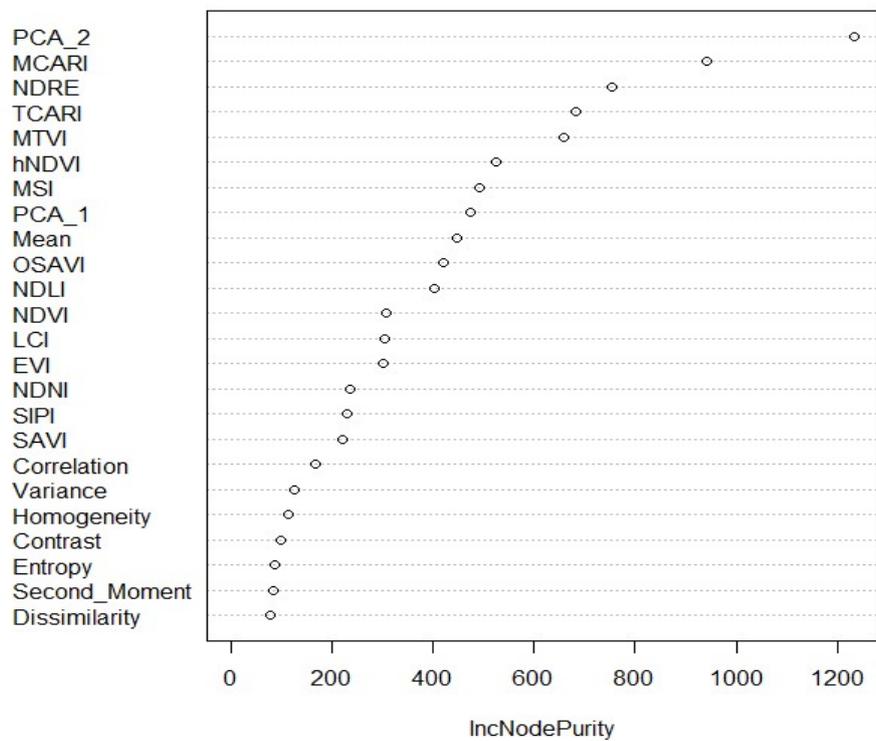


Fig. 9 Importance of each input variables for classification using random forest classifier.

Utilizing the important variables identified PC 2, MCARI, NDRE, TCARI, MTVI, hNDVI, PC 1, Mean of texture analysis, OSAVI, and NDLI; each with a node purity exceeding 400, representing one-third of the total node purity, a probability map for bamboo areas was generated using a Random Forest classifier. This analysis was specifically applied to areas classified as bamboo and mixed bamboo. The probability of bamboo presence in each pixel was computed. To assess the accuracy of the generated bamboo probability map, a field survey was conducted within the mapped bamboo probability areas. Validation was performed by generating and validating a total of 30 random points on the ground. The accuracy of the bamboo probability map was determined to be 82%. Figure 10 presents a high-resolution Google Earth image alongside field photos, showcasing the validation process and affirming the accuracy of the classified bamboo probability map when compared with ground truth data.

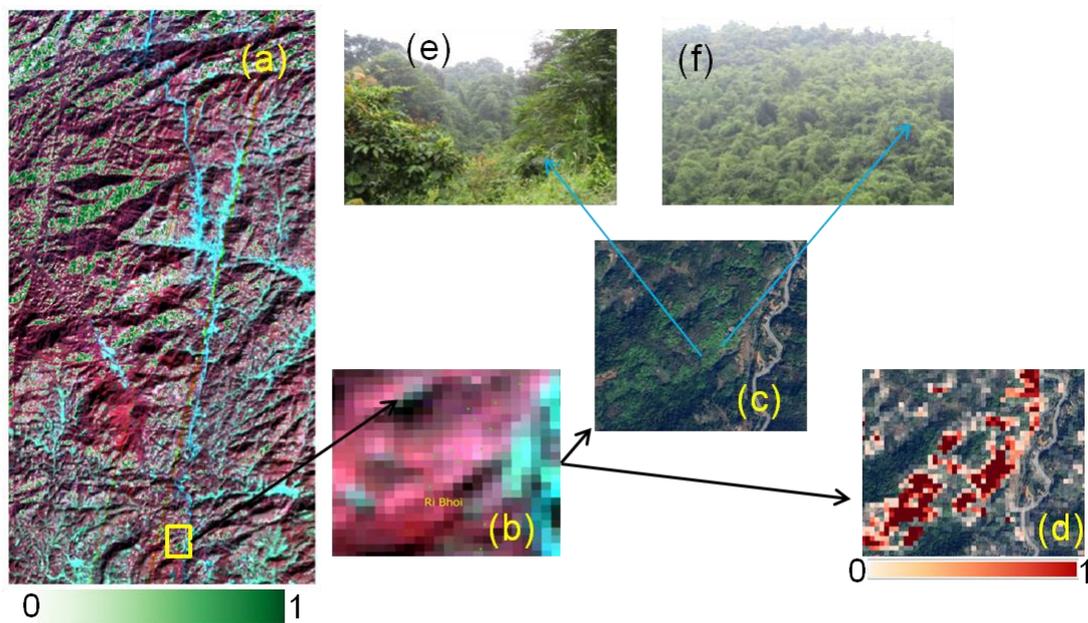


Fig. 10 (a) Probability of bamboo growing area in the study area. (b) enhanced view of PRISMA hyperspectral data where field data was collected. (c) High resolution Google Earth image of the enhanced area. (d) classified bamboo probability map from 0 to 1 in the enhanced view. (e) field photo of the mixed bamboo composition. (f) pure bamboo forest patch photograph as taken from the field and can be seen in the satellite data.

This comprehensive analysis not only emphasizes the relevance of spectral features, such as leaf pigments and structural characteristics, but also highlights the importance of incorporating texture analysis for accurate bamboo discrimination. The study's findings contribute valuable insights into the key variables influencing classification outcomes and pave the way for refining and optimizing future vegetation mapping efforts.

Discussion

Bamboo conservation is important from the point of view of carbon sequestration and dependence of locals on it for livelihood. Many scholars have tried to map and document bamboo resources of north east region and have emphasized on conservation of bamboo in north eastern region. Kharlyngdoh et al. (2016) have documented important bamboo species

of North east India and suggested the conservation of *Schizostachyumdullooa*, an important bamboo species which has a spread of 2,678 ha area in Nongkhylllem area in Ri Bhoi district. Bamboo in many areas has been considered invasive also in North East India as it spreads and grows rapidly. The reason for more spread of bamboo in Ri Bhoi district could be due to high disturbance which was reported by Roy and Tomar (2000). The present study focused on utilizing hyperspectral remote sensing imagery to discriminate bamboo growing areas from mixed vegetation compositions, employing a machine learning algorithm for classification. The study underscores the limitations of traditional mapping techniques in bamboo identification and mapping, emphasizing the need for advanced technologies. Hyperspectral remote sensing, with its ability to capture detailed spectral information, emerges as a valuable tool. The study addresses the gap in the literature by highlighting the limited use of hyperspectral data for bamboo mapping, advocating for its superior capabilities in discriminating vegetation species (Tamang et al., 2022).

Hybrid classification method using PCA, visual interpretation and GLCM was found to have highest accuracy for classification of Bamboo and mixed forest. PCA is an appropriate tool for hyperspectral image analysis as hyperspectral images have high dimensional data and only few components are actually important (Shabna and Ganesan 2014). The application of PCA in this study not only elucidates crucial information within the initial two bands but also aids in the differentiation of features such as vegetated and non-vegetated areas. These findings contribute to a comprehensive understanding of the hyperspectral data and its potential applications in environmental monitoring and land use classification (Gambardella et al., 2021). Further, to improve the classification, GLCM was used which added texture feature to the spectral feature. GLCM and PCA were also observed to be best combination of hyperspectral image classification with SVM classifier (Ding et al., 2020). The finding of the present study also showcases the highest accuracy of PCA classification using SVM classifier. In the present study a meticulous analysis of 24 vegetation indices was carried out, correlating them to reduce redundancy. This refined selection of VIs not only minimizes redundancy but also enhances the discriminative power of the analysis (Jopia et al., 2020). The chlorophyll pigment was found to be very useful in the discrimination of bamboo from other vegetation classes. The red-edge region in combination with other visible and near infrared region was very crucial in classification. Various studies (Prosperre et al., 2014; Zulfa et al., 2020; Johnson et al., 2023) have highlighted the importance of chlorophyll in discrimination of vegetation.

A supervised machine learning classification technique was found to provide more accurate bamboo growing area. The random forest classifier was found to have highest accuracy for bamboo mapping which has also been reported by various researchers around the world (Ghosh et al., 2014; Chen et al., 2018; You et al., 2020). The accuracy achieved by different classifiers, highlighting the efficacy of the Random Forest classifier when incorporating texture, vegetation indices, aspect, and PCA. The mapped LULC classes, including bamboo-growing areas and mixed bamboo, demonstrated the classifier's effectiveness in capturing the spatial distribution of different land cover types. The discussion highlights the achieved accuracy of 82%, affirming the reliability of the classification outcomes and the selected classifier for accurate bamboo distribution mapping within the study area.

Conclusions

In this rapidly changing world, bamboo holds a unique position as a sustainable, renewable, and versatile resource. The study successfully demonstrated the effectiveness of PRISMA hyperspectral remote sensing data, coupled with machine learning algorithms, in discriminating bamboo forests in the North East India region. The integration of vegetation indices, texture analysis, and PCA proved crucial for achieving accurate classification, with Random Forest emerging as the most effective classifier. Future research could explore the temporal aspect of bamboo mapping, considering seasonal variations in spectral signatures. Additionally, incorporating advanced machine learning techniques and deep learning models could further enhance classification accuracy. Ground truth data collection could be expanded to cover a broader range of bamboo species, contributing to a more comprehensive understanding of bamboo distribution.

The study's findings have implications for sustainable bamboo management, providing valuable insights for conservation efforts, livelihood enhancement, and policy formulation. As the National Bamboo Mission aims to harness bamboo's potential for rural income and sustainable development, the study's methodology and results contribute to the ongoing efforts in promoting the sustainable utilization of bamboo resources in the region.

Acknowledgements

The authors express their gratitude to ISRO/DOS for generously sponsoring the project and Italian Space Agency for providing the PRISMA hyperspectral data. Additionally, sincere thanks are extended to the Forest and Environment Department, Government of Meghalaya, for granting permission to conduct the fieldwork.

References

- Broge, N. H., & Leblanc, E. (2000). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote sensing of environment*, 76(2), 156-172.
- Busetto, L. Prismaread: An R Package for Importing PRISMA L1/L2 Hyperspectral Data and Convert Them to a More User-Friendly Format—v0. 1.0. 2020.
- Chen, Y., Li, L., Lu, D., & Li, D. (2018). Exploring bamboo forest aboveground biomass estimation using Sentinel-2 data. *Remote Sensing*, 11(1), 7.
- Coppo, P., Brandani, F., Faraci, M., Sarti, F., & Cosi, M. (2019). Leonardo spaceborne infrared payloads for earth observation: SLSTRs for Copernicus Sentinel 3 and PRISMA hyperspectral camera for PRISMA satellite. *Multidisciplinary Digital Publishing Institute Proceedings*, 27(1), 1.
- Datt, B. (1998). Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+ b, and total carotenoid content in eucalyptus leaves. *Remote Sensing of Environment*, 66(2), 111-121.
- Datt, B. (1999). Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *International Journal of Remote Sensing*, 20(14), 2741-2759.
- Daughtry, C. S., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey Iii, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote sensing of Environment*, 74(2), 229-239.
- Ding, H., Xu, L., Wu, Y., & Shi, W. (2020). Classification of hyperspectral images by deep learning of spectral-spatial features. *Arabian Journal of Geosciences*, 13, 1-14.
- Du, H., Mao, F., Li, X., Zhou, G., Xu, X., Han, N., ... & Zhou, Y. (2018). Mapping global bamboo forest distribution using multisource remote sensing data. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 11(5), 1458-1471.
- Forests & Environment Department, Government of Meghalaya, (2023). https://megforest.gov.in/forest_reserved.html (accessed on 25 July 2023).

- Galvao, L. S., Formaggio, A. R., & Tisot, D. A. (2005). Discrimination of sugarcane varieties in Southeastern Brazil with EO-1 Hyperion data. *Remote sensing of Environment*, 94(4), 523-534.
- Gambardella, C., Parente, R., Ciabrone, A., & Casbarra, M. (2021). A principal components analysis-based method for the detection of cannabis plants using representation data by remote sensing. *Data*, 6(10), 108.
- Gamon, J. A., Penuelas, J., & Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of environment*, 41(1), 35-44.
- Ghosh, A., & Joshi, P. K. (2014). A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high-resolution WorldView 2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, 26, 298-311.
- Goswami, J., Tajo, L., & Sarma, K. K. (2010). Bamboo resources mapping using satellite technology. *Current Science*, 650-653.
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote sensing of environment*, 90(3), 337-352.
- Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote sensing of environment*, 81(2-3), 416-426.
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote sensing of environment*, 25(3), 295-309.
- Huete, A. R., Liu, H., & van Leeuwen, W. J. (1997, August). The use of vegetation indices in forested regions: issues of linearity and saturation. In *IGARSS'97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing-A Scientific Vision for Sustainable Development (Vol. 4, pp. 1966-1968)*. IEEE.
- Hunt Jr, E. R., & Rock, B. N. (1989). Detection of changes in leaf water content using near-and middle-infrared reflectances. *Remote sensing of environment*, 30(1), 43-54.
- Jha, C. S., Singhal, J., Reddy, C. S., Rajashekar, G., Maity, S., Patnaik, C., ... & Martinez, M. H. (2019). Characterization of species diversity and forest health using AVIRIS-NG hyperspectral remote sensing data. *Current Science*, 116(7), 1124-1135.
- Johnson, R. M., Orgeron, A. J., Spaunhorst, D. J., Huang, I. S., & Zimba, P. V. (2023). Discrimination of weeds from sugarcane in Louisiana using hyperspectral leaf reflectance data and pigment analysis. *Weed Technology*, 37(2), 123-131.
- Jopia, A., Zambrano, F., Pérez-Martínez, W., Vidal-Páez, P., Molina, J., & De la Hoz Mardones, F. (2020). Time-series of vegetation indices (VNIR/SWIR) derived from Sentinel-2 (A/B) to assess turgor pressure in kiwifruit. *ISPRS International Journal of Geo-Information*, 9(11), 641.
- Kharlyngdoh, E., Adhikari, D. & Barik, S.K. (2016) Modelling the Distribution of a Few Lesser-known Bamboo Species of Meghalaya and Determining Areas for Their Conservation. In *Biodiversity and Environmental Conservation Discovery Publishing House Pvt. Ltd., New Delhi – 110002*, 202-216.
- Kishore, B. S. P. C., Kumar, A., Saikia, P., Lele, N., Pandey, A. C., Srivastava, P., ... & Khan, M. L. (2020). Major forests and plant species discrimination in Mudumalai forest region using airborne hyperspectral sensing. *Journal of Asia-Pacific Biodiversity*, 13(4), 637-651.
- Lele, N., & Joshi, P. K. (2009). Analyzing deforestation rates, spatial forest cover changes and identifying critical areas of forest cover changes in North-East India during 1972–1999. *Environmental monitoring and assessment*, 156, 159-170.
- National Bamboo Mission (NBM), Ministry of Agriculture and Farmers Welfare, Government of India (2023). <https://www.bambooonline.org/mission.php> (accessed on 14 March, 2023)
- Navalgund, R. R., & Ray, S. S. (Eds.). (2011). *Hyperspectral Data, analysis techniques, and applications*. Bishen Singh Mahendra Pal Singh.
- Nfornekah, B. N., Rene, K., Martin, T., Louis, Z., Cedric, C., & Armand, T. (2020). Assessing the spatial distribution of bamboo species using remote sensing in Cameroon. *Journal of Ecology and the Natural Environment*, 12(4), 172-183.
- Oppelt, N. (2002). *Monitoring of plant chlorophyll and nitrogen status using the airborne imaging spectrometer AVIS (Doctoral dissertation, lmu)*.
- Peñuelas, J., & Filella, I. (1998). Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in plant science*, 3(4), 151-156.

- Peñuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J., & Field, C. B. (1994). Reflectance indices associated with physiological changes in nitrogen-and water-limited sunflower leaves. *Remote sensing of Environment*, 48(2), 135-146.
- PRISMA, (2022). ASI PRISMA data web portal. <http://www.prisma-i.it/> (accessed on 06 August 2022)
- Prospere, K., McLaren, K., & Wilson, B. (2014). Plant species discrimination in a tropical wetland using in situ hyperspectral data. *Remote sensing*, 6(9), 8494-8523.
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote sensing of environment*, 48(2), 119-126.
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote sensing of environment*, 55(2), 95-107.
- Roujean, J. L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote sensing of Environment*, 51(3), 375-384.
- Roy, P. S., & Joshi, P. K. (2002). Forest cover assessment in north-east India--the potential of temporal wide swath satellite sensor data (IRS-1C WiFS). *International Journal of Remote Sensing*, 23(22), 4881-4896.
- Roy, P. S., & Tomar, S. (2000). Biodiversity characterization at landscape level using geospatial modelling technique. *Biological conservation*, 95(1), 95-109.
- Roy, P. S., Roy, A., Joshi, P. K., Kale, M. P., Srivastava, V. K., Srivastava, S. K., ... & Kushwaha, D. (2015). Development of decadal (1985–1995–2005) land use and land cover database for India. *Remote Sensing*, 7(3), 2401-2430.
- Serrano, L., Penuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote sensing of Environment*, 81(2-3), 355-364.
- Shabna, A., & Ganesan, R. (2014). HSEG and PCA for Hyper-spectral Image Classification. In 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 42-47, July, IEEE.
- Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote sensing of environment*, 81(2-3), 337-354.
- Tamang, M., Nandy, S., Srinet, R., Das, A. K., & Padalia, H. (2022). Bamboo mapping using earth observation data: A systematic review. *Journal of the Indian Society of Remote Sensing*, 50(11), 2055-2072.
- You, S., Zheng, Q., Lin, Y., Zhu, C., Li, C., Deng, J., & Wang, K. (2020). Specific bamboo forest extraction and long-term dynamics as revealed by Landsat time series stacks and Google Earth Engine. *Remote Sensing*, 12(18), 3095.
- Zulfa, A. W., Norizah, K., Hamdan, O., Zulkifly, S., Faridah-Hanum, I., & Rhyma, P. P. (2020). Discriminating trees species from the relationship between spectral reflectance and chlorophyll contents of mangrove forest in Malaysia. *Ecological Indicators*, 111, 106024.

Citation

Bhavsar, D., Chakraborty, K., Sarma, K.K., Aggarwal, S.P. (2024). Bamboo Distribution Mapping using Hyperspectral Remote Sensing in Ri Bhoi district of North East India. In: Dandabathula, G., Bera, A.K., Rao, S.S., Srivastav, S.K. (Eds.), *Proceedings of the 43rd INCA International Conference, Jodhpur, 06–08 November 2023*, pp. 138–154, ISBN 978-93-341-2277-0.

Disclaimer/Conference Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of INCA and/or the editor(s). The editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.