Classification of tropical trees growing in a sanctuary using Hyperion (EO-1) and SAM algorithm

Binal Christian and N. S. R. Krishnayya*
Ecology Laboratory, Botany Department, Faculty of Science, The M.S. University of Baroda, Vadodara 390 002, India

Tropical forests are one of the richest sources of biodiversity and are well known for their ecosystem services. There is a pressing need to monitor the rate and extent of changes in forest cover of countries like India for efficient planning and management leading to sustainable development. Imaging spectroscopy is one of the newer techniques adopted for species-level discrimination. Of the available sensors, spaceborne ones are cost-effective and are more appropriate for monitoring in countries like India. The present study aims at classifying tropical trees using Hyperion (EO-1) and SAM (Spectral Angle Mapper) algorithm. The study was conducted in the Shoolpaneshwar Wildlife Sanctuary (SWS), Narmada District, Gujarat, India. Hyperion data were obtained during October 2006 when the vegetation was lush green. Field survey was done coinciding with data acquisition time. The tree species identified for discrimination were Tectona grandis L., Dendrocalamus strictus Nees., Mangifera indica L., Madhuca indica J. F. Gmel. and Ficus glomerata Roxb. Hyperion data were preprocessed. Endmember spectra for each species were selected and used as library spectra for the classification. SAM was performed for the entire spectrum, VIS–NIR region (1–90 bands), SWIR-I region (103–136 bands), SWIR-II region (159–195 bands), 1–10 MNF and 1–15 MNF bands. Overall accuracy assessment (OAA), kappa coefficient and user’s and producer’s accuracy were calculated. SAM classification with 196 bands (full-spectra) of Hyperion data gave 51% OAA for the five tropical trees selected. The obtained OAA was appropriate looking at the pattern of vegetal cover and also of the sensor used. Partition analysis of the spectrum indicated superiority of VIS–NIR region for classification. SWIR-I and II did not fare well because of the biophysical state of vegetal cover. SAM showed the highest accuracy (59.57%) for spectra of 1–10 MNF bands. Higher accuracy using MNF band combination indicated the potential of MNF transformation to increase classification accuracy of tropical trees by reducing data dimensionality. Our study indicates that homogeneity in the vegetal cover is a critical aspect for classification in the tropical areas. We conclude that SAM is an appropriate method for classifying Hyperion data of the tropics. With the reported densities for Tectona and Dendrocalamus, Hyperion is found to be an appropriate sensor for monitoring.

Keywords: Hyperion (EO-1) data, MNF transformation, species-level classification, Spectral Angle Mapper, tropical forest.

FORESTS of the tropical zones constitute about half of the world’s forests and mostly occur in developing countries. In recent years tropical forests have received much attention because of their species richness, high standing biomass and global net primary productivity (NPP). Tropical forests consist of the world’s largest biodiversity and play an important role in the global terrestrial carbon budget. The structure, composition and functioning of these forests are undergoing rapid changes because of anthropogenic activities. Biotic pressure and widespread economic growth are altering the natural vegetal cover and putting tremendous pressure on the sustenance of the few leftover tropical forest covers in India. As a result, there is a lot of spatial and temporal variation in the reported values of species richness, composition and productivity. There is a pressing need to monitor the rate and extent of changes in the tropical forest cover of countries like India for efficient planning and management leading to sustainable development. Sustainable management of these systems requires spatial information of tree distribution. Remote sensing is one of the important tools to have a comprehensive understanding of species-level distribution in natural systems. Conventional multispectral remote sensing is not suitable for species-level discrimination. Imaging spectroscopy is one of the newer techniques adopted for species-level discrimination. The continuous narrow bands of spectral reflectance from 400 to 2500 nm range by imaging spectroscopy provide species-level separability. Many studies were conducted, mostly on temperate systems to understand species distribution. Few studies were carried out on the tropical systems. The prospect for tropical tree identification using hyperspectral imaging is intriguing. Availability of high spatial resolution satellite-borne hyper data

*For correspondence. (e-mail: krish14@gmail.com)
is a significant limitation to advancement in this area. Tropical vegetation has unique features compared to temperate vegetation. Besides having larger diversity, the vegetal cover is not as uniform as is normally seen in the temperate region because of its rich diversity. This makes the discrimination process more challenging. Tropical trees in a deciduous forest can be discriminated by looking at phenological variations. Few reports are available for the discrimination of tropical vegetal cover using imaging spectroscopy. These studies were done using either laboratory spectra or airborne hyper data. For example, laboratory spectra (450–950 nm) have been used to discriminate 11 tropical rainforest tree species. Similarly, airborne HYDICE data have been used to classify seven tropical rainforest tree species at different spatial scales. Hence, there is a gap in the tropical tree species classification using spaceborne hyperspectral data. To the best of our knowledge there is no study on the classification of tropical trees using spaceborne hyperspectral data.

Many algorithms have been used for treespecies classification. Some of these such as ML (Maximum Likelihood), SAM (Spectral Angle Mapper), SCM (Spectral Correlation Mapper) and LDA (Linear Discriminant Analysis) have been optimized for distinguishing trees in temperate forests and also in tropical forests using airborne hyperspectral data. It is unclear how these algorithms will perform on a complex scene of tropical forests collected by spaceborne hyperspectral data (EO-1, Hyperion). Spaceborne sensors are cost-effective. They are more appropriate for vegetal cover monitoring in countries like India. Here we make an attempt to classify tropical trees using spaceborne hyperspectral data. The present study aims at using Hyperion to classify tropical trees based on the entire spectrum, spectrum partition analysis and spectra from MNF bands using SAM algorithm. The following points have been addressed.

- Extracting endmember spectra for tropical trees.
- Exploring the uniqueness of spectral regions for the classification of trees.
- Usability of MNF spectra coming from bands of high SNR (signal-to-noise) values for classification.
- To look at the importance of uniformity/homogeneity in tropical vegetal cover in affective accuracy assessment.

Materials and methods

Study area and image acquisition

The study was conducted in the Shoolpaneshwar Wildlife Sanctuary (SWS), Narmada District, Gujarat, India (Figure 1). SWS is one of the important naturally protected regions supporting sizeable biota. It is spread over an area of 675 km². Latitude and longitude of the study area are 21°29’–21°52’N and 73°29’–73°54’E respectively. Topography of the study area is undulated with continuous and discontinuous hilly tracts intermingled with valleys, streams and sporadic clearings for agriculture. The entire SWS is well drained by a large number of streams and rivulets during monsoon (July–October). In summer (March–June), many of these streams are dry as the flow of water ceases. At some places water remains stored in small pools on the rocky river beds, mainly of some use for herbivores. The area is important for its support to wildlife, tribal population and as a catchment area for local water bodies. There are ~127 tree species. Dominant species are Tectona grandis L. followed by Dendrocalamus strictus Nees. Both species are homogeneously distributed and occupy about 30% of the study area. Distribution of all other tree species is highly heterogeneous.

On 21 October 2006, Hyperion collected hyperspectral data in a 7.7 km by 42 km swath centred over the study area at nadir view (Figure 1). At the time of the satellite flight, the study area had less than 25% cloud cover and vegetation at the sanctuary comprised of lush green foliage corresponding to the post-monsoon season. Information was delivered as 16-bit calibrated radiance data.

![Figure 1](image_url). Location map of the Shoolpaneshwar Wildlife Sanctuary (SWS), Narmada District, Gujarat. Hyperion data FCC (RGB): 854, 650 and 559 nm with SWS boundary.
Field data collection

Field survey was done coinciding with the data acquisition time. The selected plot of the study area (328 columns, 322 lines) was criss-crossed to locate the homogenous cover of species, at least matching the size of the Hyperion pixel (30 m x 30 m). Like in any tropical area heterogeneity in species distribution is unique to the study area. Excepting for Tectona and Dendrocalamus, all other species showed mixed occurrence. Hence establishing an appropriate quadrat for one species became critical. The size of each quadrat laid down was 30 m x 30 m, matching the Hyperion resolution. Keeping this in mind a quadrat was selected for a species when it occupied ≥60% of the area. Based on this distribution pattern five species were selected for classification. The tree species identified for discrimination were T. grandis L., D. strictus Nees., Mangifera indica L., Madhuca indica J. F. Gmel. and Ficus glomerata Roxb. Phytosociological data such as dbh (diameter breast height), height, density and spread of canopy were recorded from each quadrat. Quadrats laid down were randomly spread across the selected area. At each quadrat the tree density was inversely proportional to dbh and spread of the canopy. The density of Tectona was 75 trees with a dbh of 0.3–0.8 m, in Dendrocalamus 37 clumps were with 25–80 stem density, in Mangifera the density was 1–7 with a dbh of 1.55–3.2 m, in Madhuca the density was 1–5 with a dbh of 1.75–2.9 m, and in Ficus the density was 1–2 with a dbh of 5.25–6.70 m. Distance between quadrats was dependent on the type and distribution of vegetal cover. For Tectona and Dendrocalamus 3–4 quadrats were marked together as their distribution was relatively homogenous. For others, the distance between quadrats ranged from 250 to 500 m. Tectona spread over a patch size of 100 m x 100 m. Dendrocalamus showed coverage of 50 m x 80 m. For the others, maximum occupancy as a homogenous patch was 30 m x 30 m. A total of 141 quadrats were laid for the selected tropical trees. Latitude and longitude of these quadrats were measured using Leica GPS. These were referred to as GCPs (Ground Control Points) on the Hyperion image. Tectona and Dendrocalamus have high economic value in these parts of the world. Other selected trees are used by the locals as NTFP (Non Timber Forest Products).

Preprocessing of Hyperion data

In preprocessing, spectral (excluding uncalibrated and overlapping bands) and spatial subsets were generated within the scene. The image was masked to exclude water bodies and to minimize under- or over-correction in subsequent processing steps. Calibrated radiance data were converted to surface reflectance using Atmospheric Correction Now (ACORN) v5.0 program of ImSpec LLC.24

A ‘destreaking’ program25 was applied to minimize the striping artifact present in the Hyperion data.

Geometric correction

Geometric correction was done using ALI (Advanced Land Imager) image obtained at the same time. It was rectified to a common projection (geographic spheroid WGS 84) using well-identified GCPs. The resultant rectified ALI image had 0.5 pixels RMSE (Root Mean Square Error). Finally, the atmospherically corrected subset of Hyperion was geo-registered with nearest-neighbour resampling to the geo-registered image of ALI and the resultant RMSE was 0.2 pixels.

MNF transformation

Hyperspectral data consume large amounts of memory leading to longer processing time. Furthermore, information in the hyperspectral data is redundant to a large extent, making it difficult to extract anything useful. To overcome these difficulties, data reduction and enhancement were performed by MNF transformation using ENVI 3.6. The MNF transform is essentially two principal component transformations. The first transformation, based on an estimated noise covariance matrix, decorrelates and rescales the noise in the data. The second step is a standard principal component transformation which creates several new bands containing majority of the information.26 MNF transformation was performed on all 165 bands after removing 0-value bands. Only the first 15 MNF bands were used for subsequent analysis as they contained high values of SNR.

Pixel Purity Index

All remotely sensed images contain different levels of information coming from a wide array of pixels. This would be more common in the tropical cover similar to that of our study area. In hyperspectral analysis it is useful to separate purer pixels from highly mixed ones in order to reduce the number of pixels to be analysed for endmember separation and identification. The PPI is a way of finding the most spectrally pure pixels in images.27,28 Hence, PPI was performed on selected MNF bands. It helps in filtering the data as well as in choosing the correct endmember. Different MNF band combinations were selected such as 1–2, 1–3 up to 1–15 to extract the purest pixels for each species. PPI was confirmed by 10,000 iterations. Each MNF band combination has different number of pure pixels. Field observations were overlaid for each species. The number of purest pixels and hits for each species were observed and data value/number of iterations (the number of times that pixel was recorded as
an extreme) of each of these pixels was recorded. Band combination was different for different species. For example, the higher number of hits for Tectona and Dendrocalamus was found in 1–5 MNF band combination. While for the others it was in 1–10 MNF band combination. Hence, 1–5 MNF band combination was used as an unique band combination to select correct endmember spectra of Tectona and Dendrocalamus for subsequent analysis. For other species, 1–10 band combination was used to select endmember spectra.

Endmember spectra for each species were selected manually from the developed PPI image. The number of iterations of GCP locations of each species was extracted. Pixel locations having maximum number of iterations were considered as the endmembers for the respective species. Spectra of these pixels were used as library spectra of selected tree species for further analysis.

**Spectral Angle Mapper**

The SAM algorithm determines the spectral similarity between two spectra by calculating the angle between the spectra, treating them as vectors in a space with dimensionality equal to the number of bands\(^{2,30}\). SAM compares the angle between the spectrum vectors of a known class with each pixel vector of an unknown class in n-dimensional space. SAM was performed using the selected endmembers of each species as different classes. To minimize false detection in this study, no threshold value was used in the classification. It was performed for the entire spectrum, VIS + NIR region (1–90 bands), SWIR-I region (103–136 bands), SWIR-II region (159–195 bands), 1–10 MNF and 1–15 MNF bands.

Hits from the collected GCPs were used for accuracy assessment. Two measures of classification accuracy (user’s and producer’s accuracy), overall accuracy assessment (OAA) and kappa coefficient were calculated\(^{20,31}\).

### Results and discussion

The results of our study are given in Tables 1–3 and Figures 2 and 3. The results indicate the intricacies and limitations involved in the extraction of information for tropical vegetation from Hyperion sensor. Data processing showed the advantages of MNF transform followed by the development of PPI. These transformations enabled to narrow down sizably the actual pixels to be classified. They also helped in the extraction of endmember spectra from field-based measurements for each species. Steps taken here for classification helped in getting better OAA for tropical trees, which is important. Heterogeneity in the vegetation cover has largely influenced the accuracy assessment. Analysis of specific spectral regions indicated that spectra of MNF bands with high SNR values fared better compared to the entire spectrum. SAM proved to be a good algorithm for classifying tropical trees.

SAM classification with 196 bands (full spectra) of the Hyperion data gave 51% OAA for the five tropical trees selected. Obtained OAA is appropriate looking at the pattern of vegetal cover and also of the sensor used. Another limitation for OAA could be from the redundancy of data coming from the entire spectral range itself. Similar values of OAA were reported earlier using airborne data\(^{22}\). Our accuracy values are better for a space-borne sensor. Partition analysis across the spectrum was done for three regions (VIS–NIR, SWIR-I and SWIR-II) instead of four regions as reported in earlier findings\(^2\). The VIS–NIR region was taken together to give emphasis for red edge variations as the canopy is lush green. OAA was maximum at VIS–NIR, indicating its superiority in classification. OAA for the other two regions (SWIR-I and II) was less. This is in contrast to the earlier results\(^2\) where crown scale separability was better at SWIR-II. In this study crowns of trees were with green foliage, having the highest chlorophyll content in a growing period. Hence VIS–NIR region gave a better canopy-level separability compared to the other two regions. Importance of VIS–NIR region was earlier highlighted for vegetal

<table>
<thead>
<tr>
<th>Band combination</th>
<th>Angle</th>
<th>OAA</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>196 bands</td>
<td>None</td>
<td>51.06</td>
<td>0.477</td>
</tr>
<tr>
<td>VIS–NIR (1–90 bands)</td>
<td>None</td>
<td>54.60</td>
<td>0.450</td>
</tr>
<tr>
<td>SWIR-I</td>
<td>None</td>
<td>33.33</td>
<td>0.251</td>
</tr>
<tr>
<td>SWIR-II</td>
<td>None</td>
<td>26.24</td>
<td>0.222</td>
</tr>
<tr>
<td>1–15 MNF bands</td>
<td>None</td>
<td>55.31</td>
<td>0.467</td>
</tr>
<tr>
<td>1–10 MNF bands</td>
<td>None</td>
<td>59.57</td>
<td>0.519</td>
</tr>
</tbody>
</table>

### Table 2. User’s accuracy for each class in SAM (threshold value – none) using different band combinations

<table>
<thead>
<tr>
<th>Class</th>
<th>VIS–NIR</th>
<th>1–15 MNF</th>
<th>1–10 MNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tectona</td>
<td>82.35</td>
<td>74.50</td>
<td>70.58</td>
</tr>
<tr>
<td>Dendrocalamus</td>
<td>40.54</td>
<td>48.64</td>
<td>51.35</td>
</tr>
<tr>
<td>Mangifera</td>
<td>38.88</td>
<td>44.44</td>
<td>44.44</td>
</tr>
<tr>
<td>Madhuca</td>
<td>22.22</td>
<td>31.57</td>
<td>36.84</td>
</tr>
<tr>
<td>Ficus</td>
<td>31.25</td>
<td>31.25</td>
<td>43.75</td>
</tr>
</tbody>
</table>

### Table 3. Producer’s accuracy for each class in SAM (threshold value – none) using different band combinations

<table>
<thead>
<tr>
<th>Class</th>
<th>VIS–NIR</th>
<th>1–15 MNF</th>
<th>1–10 MNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tectona</td>
<td>53.84</td>
<td>60.31</td>
<td>85.71</td>
</tr>
<tr>
<td>Dendrocalamus</td>
<td>57.69</td>
<td>58.06</td>
<td>82.60</td>
</tr>
<tr>
<td>Mangifera</td>
<td>61.53</td>
<td>33.33</td>
<td>77.77</td>
</tr>
<tr>
<td>Madhuca</td>
<td>40</td>
<td>72.7</td>
<td>72.72</td>
</tr>
<tr>
<td>Ficus</td>
<td>35.71</td>
<td>27.77</td>
<td>70</td>
</tr>
</tbody>
</table>
analysis and also for subtropical tree species. Our results are in conformity with these reports.

OAA is minimal for spectra of the SWIR-II region. This is analogous to the earlier findings where poor representation of SWIR-II region was reported during the discrimination of six southern tree species. SWIR-I and SWIR-II regions are related to the expression of water absorption features and ligno-cellulose absorption features which may be expressed when high fractions of non-photosynthetic woody tissues are exposed to the sensor, such as species with low leaf area index (LAI). In this study canopy reflectance is dominated by green leaves (LAI > 3). The results of our study are in tune with other reports. Figures 2 and 3 show classified images from different spectra using SAM as a classifier. Images from spectra of MNF bands are much sharper compared to the others (Figure 3 b and c). Similar advantage in classification using MNF bands has been reported. Higher
SNR values are mainly responsible for sharper and distinctly classified images. SAM showed the highest accuracy (59.57%) for spectra of 1–10 MNF bands. Similar OAA values were also reported for classifying age classes of Douglas fir and Norway spruce using SAM as a classification algorithm\(^\text{20}\). Higher accuracy using MNF band combination indicated the potential of MNF transformation to increase classification accuracy of tropical trees by reducing data dimensionality.

Homogeneity in the vegetal cover is a critical aspect for classification in the tropical areas. The study area has an interesting distribution of vegetation. Of the ~127 tree species, Tectona and Dendrocalamus are the most dominant, having large homogeneous patches of size 100 m × 100 m and 50 m × 80 m respectively. Other species have heterogeneous distribution not even occupying a 40 m × 40 m plot. This is aptly reflected in the OAA figures. Tectona and Dendrocalamus showed higher accuracies. Tectona had the highest accuracy level because of the kind of its distribution as well as its crown shape (Figure 4). Though Dendrocalamus also had similar pattern of distribution, accuracy levels are different. We attribute this to its crown shape. The spread is relatively less with dissected appearance giving way to the influence of background radiation. Accuracy assessment levels for Mangifera, Madhuca and Ficus are modest owing to their pattern of occurrence. At most of the quadrats laid, their occupancy ranged between 65% and 90%. Here also crown shape (Figure 4) influenced the classification. Better spread of canopy as seen in Mangifera allowed it to fare better.

Finally we conclude that SAM is an appropriate method for classifying Hyperion data in the tropics. MNF transformation followed by PPI enabled in better end-member detection. The VIS–NIR region of the spectra is more appropriate for classifying tropical trees with thick canopy. Spectra from MNF bands fared equally better in feature extraction. With the reported densities for Tectona and Dendrocalamus, Hyperion is found to be an appropriate sensor for monitoring. Work is in progress to explore other suitable classification algorithms.


ACKNOWLEDGEMENT. We thank the Indian Space Research Organization, Bangalore, for financial assistance.

Received 23 June 2008; revised accepted 6 April 2009