Chapter 5

Assessment of mining activities on tree species and its diversity in hilltop mining areas

This chapter has originally been published as: **Kayet, N**., Pathak, K., Chakrabarty, A., Singh, C. P., & Chowdary, V. M., **(2020)**. Assessment of mining activities on tree species and diversity in hilltop mining areas using Hyperion and Landsat data, **Environmental Science and Pollution Research (Springer)**, SCI, Impact factor-4.22

5.1 Overview

The tree species and its diversity are two critical components to be monitored for sustainable management of forest as well as biodiversity conservation. Saranda forest is covered with different kinds of valuable trees (Sal and teak) and is rich in iron (Kiriburu, Meghahatuburu, Gua, and Chiria) depositions. This area covered with forest near the mining fields (buffer zone) is exhibiting high-stress conditions as described by dying and dry plant material, consequently affecting tree species and its diversity. Therefore, it is essential that the impact of mining on tree species and species diversity are adequately evaluated. This chapter emphasizes the description of mining activities in hilltop mining areas on tree species and its diversity using hyperspectral imagery and field survey data.

5.2 Data acquisition and pre-processing

5.2.1 Data used

The Hyperion (Hyperspectral) and Landsat-8 OLI (Multispectral) sensors satellite data were used for tree species identification and diversity mapping. Two satellite data sets, dated 16 Dec 2016 (Hyperion), and 9 Dec 2016 (Landsat OLI) were obtained from USGS (United States geological survey). Hyperion data were available only for the above-mentioned period that is why we have used Landsat OLI of that period. Hyperion sensor captures very narrow banded data (Hyperion tutorial handbook). Field-based tree species spectral data were acquired by the spectroradiometer instrument in the study area for marching with satellite imagery spectra. The species phytosociological observation data were collected from the Chaibasa forest office, Saranda forest, for tree species identification. For species biodiversity analysis (Shannon Index based), 18 plot data were collected from the study area. GPS (Global positioning system) has recorded the tree species and its diversity locations (latitude and longitude) of the study area. . The masked forest pixels and sample location were also shown on the map in Figure. 5.1. The secondary data were (base map, toposheet, mining plan, and forest survey data) obtained from different concerned state government offices.

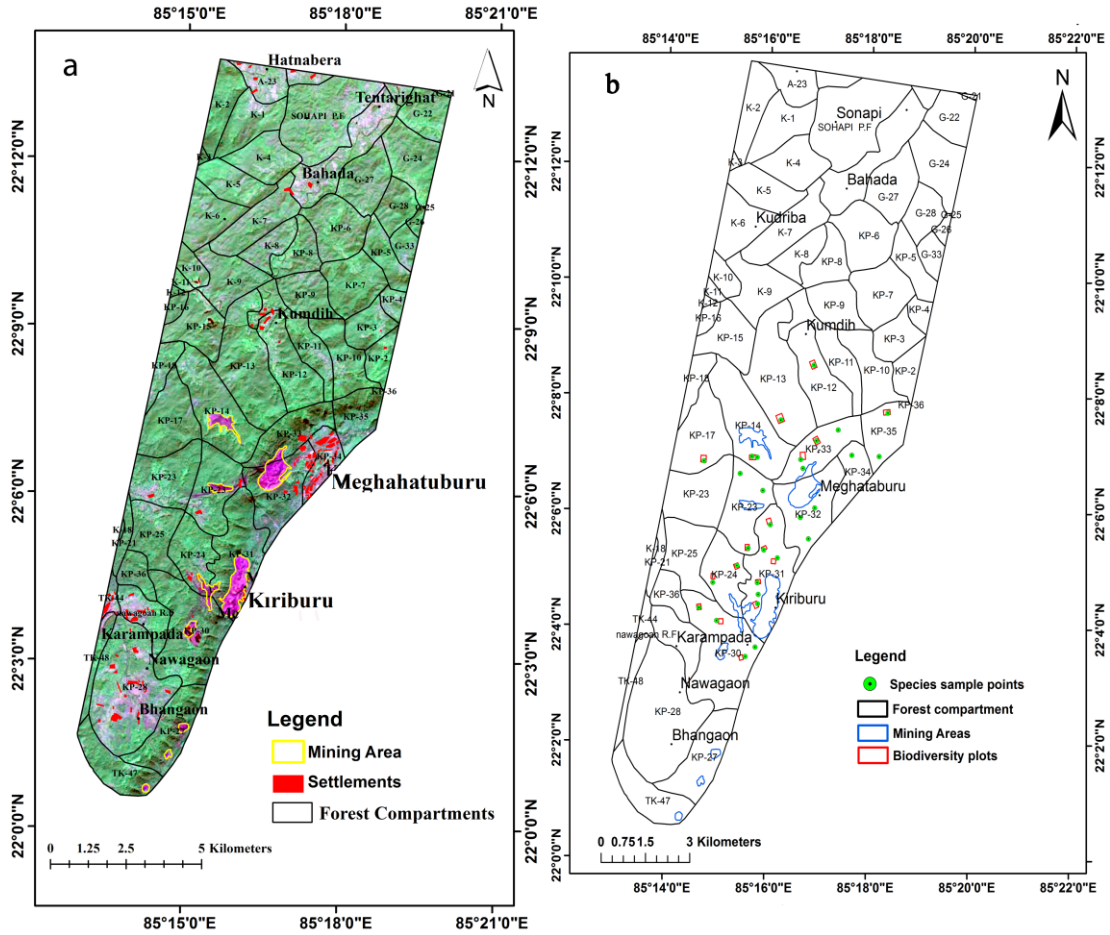

Figure 5.1: (a)Masking image used to eliminate forest pixels (left) , (b) Ground truth map indicating collected field tree species and biodiversity samples (Right)

5.2.2 Identification of forest cover pixels

Decision Tree (DT) algorithm was used for forest pixels identification. DT is a machine learning algorithm and a nonparametric classifier. Pal and Mather (2001) had used DT for mangrove species identification and classification. Five narrow-banded vegetation indices (MNDVI, MSI, NDNI, ARVI, and ARI1) were used as input of the Decision Tree (DT) classifier (Table.5.1).

Table 5.1: Narrow bands vegetation indices used for forest pixels identification

Narrow banded VIs	Index	Algorithms	Applications	References
Greenness	Modified Red Edge Normalized Difference Vegetation Index (MNDVI705)	$MRENDVI = \left(\frac{750\,nm \times 705\,nm}{(750\,nm + 705\,nm(2*445\,nm))}\right)$	Precision agriculture, forest monitoring and vegetation stress detection	Sims & Gamon, 2002
Canopy water content	Moisture Stress Index (MSI)	$MSI = \frac{P1599}{P819}$	Canopy stress analysis, Productivity and modeling fire hazard	Ceccato et al.2001

			condition analysis.	
Light use efficiency	Normalized Differences Nitrogen Index (NDNI)	$NDNI = (\log 1\ 1510nm) - (\log 1\ 1680nm) / (\log 1\ 1510nm) + (\log 1\ 1680nm)$	Estimate the amount of lignin in vegetation	Serrano et al.,2002
Greenness	Atmospherically Resistant Vegetation Index (ARVI)	$ARVI = ((NIR) - (2RED - BLUE) / (NIR) + (2RED - BLUE))$	Sensitive to changes in chlorophyll concentration	Curran., et al .1995
Leaf Pigment	Anthocyanin Reflectance Index 1 (ARI1)	$ARI1 = (1/550nm) - (1/700nm)$	Sensitive to anthocyanin amount in vegetation	Gitelson et al.,2001

DT derived threshold value was used for masking of forest pixels. The DT based identified forest pixels were justified by Google earth image (Geo-eye) and field survey data.

5.2.3 Local tree species of study area

Saranda forest comprises two main varieties of the forest, i.e., tropical moist deciduous and tropical dry deciduous and is also famous for the largest Sal forest of Asia. Teak and Sal trees are richly found in this region. Mostly the forest is covered by deciduous trees, and the most important species are Teak, Sal, Mangoes, Jamun, Piar, Akasmani, kusum, Mahua, Tilia, and Jackfruit, etc. The tree species attributes found in the study area shown in Table 5.2.

Table 5.2: Attributes of tree species at study area

S. No	Botanical Name	Common/Local Names	FSI Species Code
1	Shorea robusta	Sal	1096
2	Tectona grandis	Sagwan, Teak	1164
3	Syzygium cumini	Jamun, Jamoon, Piaman, Rajamun	1136
4	Madhuca latifolia	Mohwa, Lappa, Mahudo	759
5	Grewia tiliaefolia	Dhaman, Tada, Thadachiee, Chadichi	552
6	Gmelina arborea	Siwana, Gumari, Sivan, Gambhar	539
7	Ficus racemosa	Atti, Rumdi, Atthi	485
8	Ficus benghalensis	Figs, Wad or bat	477

9	Emblica officinalis	Amla, Aonla, Amlaki, Nellimara	410
10	Careya arborea	Kumbhi	215
11	Butea monosperma	Palas, Kakhar, Khakhara, Palasin	173
12	Albizzia odoratissima	Siris, Pullivage, Nellivega, Hiharu	56
13	Aegle marmelos	Bel, Billi, Bil, Belpatra	37
14	Acacia auriculiformis	Akasmani, Sona jhuri	6

5.2.4 Field survey and analysis

The leaf reflectance spectra of tree species were recorded by field-based spectroradiometer during the time of field survey. A total of twenty spectra corresponding to six different tree species were recorded, and the mean spectra of each tree species were used for analysis and classification. GPS has measured the longitude and latitude for each sample of tree species (SF.1). We have measured in 10×10 m2 plots in the field for Shannon Index analysis. A total of 18 plots were recorded during the field survey. The GPS position was acquired for the center of each diversity plot with the help of high-precision hand GPS. The species abundance cover, height, and habitat information were also acquired during field survey. The field survey photograph of tree species and its diversity are shown in Figure 5.2.

Figure 5.2: Spectro-radiometery field survey and laboratory analysis

5.2.5 Acquisition and pre-processing of field spectra

The spectroradiometer recorded the tree reflection spectra and their wavelength. This instrument recorded at spectral resolutions of VNIR (300 -1000nm) for 1.4 nm, NIR (1000-1700nm) for 2 nm, and SWIR (1700-2500nm) for 4 nm interval respectively. The different spectral wave ranges were resampled by the FWHM (Full width at half maximum) algorithm (Kayet et al., 2019). The spectra for different tree species were collected with the help of fiber optic source (300 to 2500nm) and 180o FOV (Field of view). For the measurement of white reference spectra, a standard reference panel (white) was used. Species leaf reflectance was measured with the help of a reflectance probe. The holder block of the reflectance probe was kept at sample distance 0-3/4″ and 90-degree angle was set. The raw field spectra of the study area were recorded by a spectroradiometer.

Pre-processing of spectra consisted of temperature drift correction, water absorption, noise bands removal, and spectral smoothing. The temperature drift errors were coming from 1001 & 1831 nm wavelength due to sensor detector changing (Lenhard et al., 2005). We have used a splice correction algorithm for temperature drift correction. The collected spectra had shown error of water vapor and noise (2350 to 2500, 1790 to 1960, and 1350 to 1460 nm wavelength) due to atmospheric components and instruments' self-generation (Staenz et al., 2002). We have just removed two types of spectral errors from wavelength bands. Some researchers have used linear and non-linear smoothing filter for spectral data smoothing. Savitzky-Golay algorithm based filter smoothing yields high accuracy (Savitzky & Golay 1964; Vaiphasa, 2006). So, we have used the Savitzky-Golay filter for spectral data smoothing. The average spectra of tree species were calculated after spectral smoothing. This spectral has been used for spectral library development and applied for classification.

5.2.6 Pre-processing of satellite data

Pre-processing correction (geometric, radiometric, and terrain) of Hyperion and Landsat 8-OLI data were done by image processing software. The Atmospheric correction was carried out by the FLAASH (Fast line-of-sight atmospheric analysis of the hypercubes) model in image processing software. The location of the study area in the hilly region induces a shadow effect on the satellite imagery. We have used a band ratio algorithm for shadow effect removal from satellite images. The projection of two images at WGS

(World geodetic system) 84 & zone 450 north, on UTM (Universal transverse mercator coordinate system) projection system were performed.

5.3 Methodology

5.3.1 Tree species discriminant analysis

For the band's selection, we have used Hyperion wavebands obtained from the discriminant analysis. This analysis found a set of prediction equations based on independent variables that have been used to classify individuals into groups (Somers et al., 2014). The discriminant analysis records the lowest Wilks lambda (L) values. The value of L lies between 0 to 1, with the value 1 or close to 1 indicates that the mean of the group is not different. Value of 0 or close to 0 indicates that the mean of the group is different. Green and Caroll developed the L statistic in 1978 (Equation.5.1).

$$L = \frac{|S_{effect}|}{|S_{effect}|+|S_{error}|} \quad \text{(Eq-5.1)}$$

Where, S_{effect} denotes a sum of squares matrix, and S_{error} denotes cross-products matrix. The classification of species was performed using selected spectral bands obtained from Wilk's lambda test.

5.3.2 Spectral separability analysis of tree species

For spectral characteristics of tree species, six different wavelength locations were selected for species spectral separability analysis. Jeffries-Matusita distance method is a method that was selected to estimate the spectral range for different species (Murakami et al., 2001). The value obtained from J-M method varies between 0 to √2. The value lying close to 0 indicates identical distribution whereas value close to √2 indicates dissimilar distribution. The equation.5.2 calculates the J-M distance method.

$$J - M_{ab} = \sqrt{2(1-e^{-d})}$$

$$d = \frac{1}{8}(\mu_a - \mu_b)^T \left(\frac{c_a+c_b}{2}\right)^{-1} (\mu_a - \mu_b) + \frac{1}{2} In \left(\frac{\left(\frac{1}{2}\right)|C_a+C_b|}{\sqrt{|C_a|\times C_b|}} \right) \quad \text{(Eq- 5.2)}$$

Where, $^{a\&b}$ are two target spectral signatures under comparison, $^{\mu}$ represents the average vector of spectral signature, C represents the covariance matrix of spectral signature, T represents the transposition role and |C| is the determinant of C (Richards and Xiuping, 2005). The selected end-members spectral wavebands of two datasets (Hyperion and Landsat OLI) were processed with J-M distance method for calculation of spectral seperability.

5.3.3 Data dimensionality and spectral similarity analysis

Atmospherically corrected Hyperion data were used in MNF (Minimum noise fraction) transformation for data dimensionality. MNF rotation transforms to determine the inherent dimensionality of image data, to segregate noise in the data, and to reduce the computational requirements for subsequent processing (Boardman and Kruse, 1994). We have analyzed noisy data in the MNF tool of image processing software, and outcome bands were used for the classification of tree species. The spectral analysis is based on spectral matching or similarity techniques. The satellite imagery-based derived end-member spectra were compared with field mean spectra using spectral similarity algorithms (Somers and Asner 2014). We have used SFF (Spectral feature fitting) algorithm for spectral similarity analysis. A high spectral similarity score denotes the closest match and exhibits maximum value.

5.3.4 Tree species classification and accuracy assessment

The tree species located in Saranda forest are homogeneous, so we have used the full pixel supervised classification methods. Some researchers have used supervised classification algorithms (SAM and MD) for full pixels classification based on trained data (Petropoulos et al., 2012; Richards and Jia, 2006). In the present study, supervised classification (SAM, SVM, and MD) algorithms have been used for full pixels classification for Landsat OLI, and Hyperion data based on training tree spectral data. The species classification accuracy matrixes were generated on the basis of ground locations spectra data. Equation.5.3 computes the accuracy of kappa statistic (K).

$$K = \frac{N \sum_{i-1}^{r} x_{ii} - \sum_{i-1}^{r} \left(x_{i} + x_{+i} \right)}{N^{2} - \sum_{i-1}^{r} \left(x_{i} + x_{+i} \right)} \qquad \text{(Eq-5.3)}$$

Where r denotes the number of rows, xii denotes the number of observation in the ith column and row. N indicates the total observations. The xi+ and x+i indicates the total number of observation in the ith row and column. A comparison was drawn between these algorithms on classified images based on accuracy assessment for the selection of the best classification algorithm.

5.3.5 Species diversity estimation based on narrow banded VIs

Species diversity basically means the occurrence of different species of trees represented in a given community (Wang et al., 2003). Some researchers have used hyperspectral narrow banded VIs correlated with field measured Shannon Index (H) values for plant diversity mapping at the regional scale level (Peng et al., 2018; Dudley et al., 2015; Mapfumo et al., 2016). The H-index is a statistical method that classifies the species diversity by assuming that the sample represents all species (Peng et al., 2018). H-index is calculated by following Equation-5.4.

$$H = -\sum_{i=1}^{s} p_i \ln p_i \qquad \text{(Eq. 5.4)}$$

Where p represents the ratio (n/N), 'n' is the number of individual species and total number of different species is 'N'. The ln is the natural log, Σ is the sum of the calculations, and s denotes the different types of species. We have used 13 hyperspectral VIs (Table 5.3) extracted from Hyperion data correlated with Shannon Index (H) values for the estimation of tree species diversity in the study area. The best correlated (higher R^2 and lower RMSE) vegetation index was selected for this estimation.

Table 5.3: Estimation of species diversity based on Narrow banded VIs

VIs	Narrow bands Algorithms	Applications	References
Difference vegetation index (DVI)	$\rho NIR_{782} - \rho R_{675}$	This index distinguishes between soil and vegetation, but it does not account for the difference between reflectance and radiance caused by atmospheric effects or shadows.	Tucker et al., 1969
NDVI (Normalized Difference Vegetation Index)	$\frac{\rho NIR_{864} - \rho R_{660}}{\rho NIR_{864} + \rho R_{660}}$	This index is used because it has the ability to reduce many forms of multiplicative noise like sun illumination difference, cloud shadows, some atmospheric attenuation, some topographic variations that are present in multi-date imagery.	Baret, & Guyot ,1991
RVI (Ratio Vegetation Index)	$\frac{\rho R675}{\rho R782}$	This index is sensitive to photosynthetic rates in forest canopies, as green and red reflectance are strongly influenced by changes in leaf pigments.	Sripada et al.,2006

SAVI (Soil Adjusted Vegetation Index)	$\frac{(\rho NIR_{864}-\rho R_{660})(1+L)}{(\rho NIR_{864}+\rho R_{660}+L)}$ L=0.5 in this study	This index is widely used for minimizing the influence of soil brightness. It can be used to describe the dynamic soil-vegetation systems from satellite imagery.	Huete,et al, 1988
MSAVI (Modified Soil Adjusted Vegetation Index)	$2*R_{800}+1-(2*R_{800}+1)^2-8*(R_{800}-R_{670})^{1/2}$	This index is a simpler version of the MSAVI index. It reduces soil noise and increases the dynamic range of the vegetation signal. MSAVI2 is based on an inductive method that does not use a constant L value (as with SAVI) to highlight healthy vegetation.	Qi et al. 1994
TSAVI (Transformed Soil Adjusted Vegetation Index)	$\frac{a(\rho NIR_{864}-a\rho R_{660}-b)}{a\rho NIR_{864}+\rho R_{660}-ab+X(1+a^2)}$ a=slope of the soil line, 1.2 in this study b=soil line intercept, 0.06 in this study X=adjustment factor to minimize soil noise, 0.08 in this study	It is almost similar to SAVI to reduce the soil background effect, but it uses the parameter of the soil line. It is a modified form of SAVI to compensate for soil variability due to changes in solar elevation and canopy structure.	Baret and Guyot, 1991
NDVI705 (Red Edge Normalized Difference Vegetation Index)	$(\rho 750-\rho 705)/(\rho 750+\rho 705)$	The NDVI705 capitalizes on the sensitivity of the vegetation red edge to small changes in canopy foliage content, gap fraction, and senescence. Applications include precision agriculture, forest monitoring, and vegetation stress detection	Gitelson and Merzlyak 1994
PVI (Perpendicular Vegetation Index)	$\frac{1}{\sqrt{1+a^2}}(\rho NIR_{864}-a\rho R_{660}-b)$ a=slope of the soil line, 1.2 in this study, b=soil line intercept, 0.06 in this study	It is used to eliminate the difference in soil background and is most effective under conditions of low LAI, applicable for arid and semiarid regions.	Huete,et al, 1988
mNDVI705 (Modified Red Edge Normalized Difference Vegetation Index)	$(\rho 750-\rho 705)/(\rho 750+\rho 705-(2*\rho 445))$	It differs from the NDVI705 by incorporating a correction for leaf specular reflection. The mNDVI705 capitalizes on sensitivity of the vegetation red edge to small changes in canopy foliage content, gap fraction, and senescence. Applications include precision agriculture, forest monitoring, and vegetation stress detection.	Datt 1999
NLI(Non-Linear Index)	$\frac{(\rho^2 NIR_{864}-\rho R_{660})}{(\rho^2 NIR_{864}+\rho R_{660})}$	It is used for removing leaf angle distribution influence and view azimuth effect	Goel and Qin, 1994
mSR705(Modified Red Edge Simple Ratio Index)	$(\rho 750-\rho 445)/(\rho 750-\rho 445)$	It differs from the standard SR because it uses bands in the red edge and incorporates a correction for leaf specular reflection. Applications include precision agriculture, forest monitoring, and stressed vegetation detection.	Kycko et al.,2017
VOG1 (Vogelmann Red Edge Index 1)	$\frac{\rho 734-\rho 747}{\rho 715+\rho 726}$	This index is a narrowband reflectance measurement that is sensitive to the combined effects of foliage chlorophyll concentration, canopy leaf area, and water content. Applications include vegetation phenology (growth) studies, precision agriculture, and vegetation productivity modeling.	Vogelmann et al., 1993

5.3.6 Relationship between species diversity, distance from mines, and concentration of foliar dust

Saranda forest has some of the largest iron ore deposits of India. Mining activities are causing damage to tree species as well as its diversity. In this study, we have shown the relationship between species diversity and distance from mines with leaf dust

concentration. We have calculated distance from two mines (Kiriburu and Meghataburu) based on field survey points location using GPS measurement tool. PCE instrument was used for the collection of leaf dust at field location points (Kayet et al., 2019). We have then correlated three parameters (outcome species diversity values, distance from mines, and concentration of leaf dust values) for their relationship. The overall research flow chart has been shown in Figure 5.3.

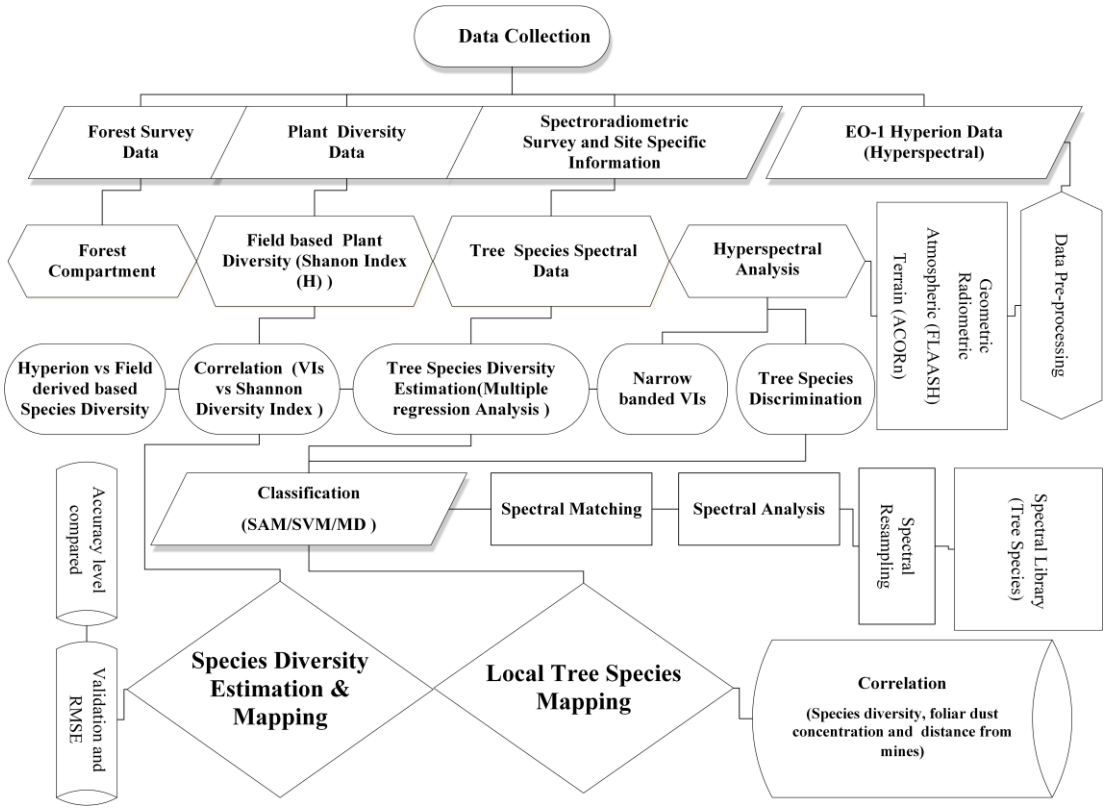

Figure.5.3: Research flowchart of tree species classification and its diversity estimatio

5.4 Results and Discussion

5.4.1 Tree species discrimination

The tree species discrimination result is displayed in Table 5.4. The value of Wilks' lambda ranged between 0 to 0.0099. The smaller value indicates that the group's mean of the wavelength bands are different and have high separability between different tree species. From this analysis, 21 optimal wavebands were obtained. From 21 bands, 07 bands fall in the VIR region, 08 bands in the NIR region, and 06 bands in the SWIR region. These wavebands were used for tree species analysis and its classification.

Table 5.4: Wilk's lambda values for 21 optimal selected wavebands

Wavelength (nm)																					Wilk's lambda
1326.05																					0.099
599.79	1326.053																				0.013
599.79	1326.053	1749.791																			0.004
599.79	993.1709	1326.053	1749.791																		0.002
599.79	993.1709	1326.053	1336.15	1749.791																	0.001
599.79	993.1709	1326.053	1336.15	1749.791	2304.713																0
599.79	660.848	993.1709	1326.053	1336.15	1749.791	2304.713															0
599.79	660.848	993.1709	1326.053	1336.15	1749.791	1780.087	2304.713														0
599.79	660.848	993.1709	1023.398	1326.053	1336.15	1749.791	1780.087	2304.713													0
599.79	660.848	993.1709	1023.398	1326.053	1336.15	1749.791	1780.087	1981.86	2304.713												0
599.79	660.848	993.1709	1023.398	1134.38	1326.053	1336.15	1749.791	1780.087	1981.86	2304.713											0
559.09	599.7959	660.848	993.1709	1023.398	1134.38	1326.053	1336.15	1749.791	1780.087	1981.86	2304.713										0
559.09	599.7959	660.848	993.1709	1023.398	1134.38	1326.053	1336.15	1477.431	1749.791	1780.087	1981.86	2304.713									0
559.0944	599.7959	660.848	993.1709	1023.398	1134.38	1326.053	1336.15	1477.431	1679.204	1749.791	1780.087	1981.86	2304.713								0
559.0944	599.7959	660.848	993.1709	1023.398	1124.283	1134.38	1326.053	1336.15	1477.431	1679.204	1749.791	1780.087	1981.86	2304.713							0
559.0944	589.6205	599.7959	660.848	993.1709	1023.398	1124.283	1134.38	1326.053	1336.15	1477.431	1679.204	1749.791	1780.087	1981.86	2304.713						0
559.0944	589.6205	599.7959	660.848	721.8994	993.1709	1023.398	1124.283	1134.38	1326.053	1336.15	1477.431	1679.204	1749.791	1780.087	1981.86	2304.713					0
518.3937	559.0944	589.6205	599.7959	660.848	721.8994	993.1709	1023.398	1124.283	1134.38	1326.053	1336.15	1477.431	1679.204	1749.791	1780.087	1981.86	2304.713				0
518.3937	559.0944	589.6205	599.7959	660.848	721.8994	993.1709	1023.398	1124.283	1134.38	1326.053	1336.15	1477.431	1679.204	1739.695	1749.791	1780.087	1981.86	2304.713			0
518.3937	559.0944	579.4455	589.6205	599.7959	660.848	721.8994	993.1709	1023.398	1124.283	1134.38	1326.053	1336.15	1477.431	1679.204	1739.695	1749.791	1780.087	1981.86	2304.713		0
518.3937	559.0944	579.4455	589.6205	599.7959	660.848	721.8994	993.1709	1023.398	1124.283	1134.38	1275.661	1326.053	1336.15	1477.431	1679.204	1739.695	1749.791	1780.087	1981.86	2304.713	0

5.4.2 Spectral separability of tree species

The J-M distance method based spectral separability values were derived from Hyperion and Landsat 8-OLI satellite imagery (Table 5.5). The values thus obtained by J-M-distance method from Hyperion data ranged between 1.25 to 1.87, which indicates that it has high spectral separability between tree species. The value ranged between 1.107 to 1.392 indicates that it has moderate to low spectral separability between tree species. The spectral separability value of different tree species derived from Hyperion data is higher than Landsat 8 OLI data.

Table 5.5: JM distance values for Hyperion (a) & Landsat (b) images based on training sample values

(a)Hyperion	Sal	Teak	Akasmani	Mohwa	Palash	Bot
Sal	-	1.77	1.14	1.79	1.69	1.71
Teak	1.77	-	1.70	1.87	1.72	1.80
Akasmani	1.14	1.70	-	1.71	1.48	1.46
Mohwa	1.79	1.87	1.71	-	1.25	1.70
Palash	1.69	1.72	1.48	1.25	-	1.38
Bot	1.71	1.80	1.46	1.70	1.38	-
(b) Landsat	Sal	Teak	Akasmani	Mohwa	Palash	Bot
Sal	-	1.34	1.10	1.37	1.39	1.33
Teak	1.34	-	1.24	1.32	1.37	1.31
Akasmani	1.10	1.24	-	1.33	1.39	1.32
Mohwa	1.37	1.32	1.33	-	1.29	1.37
Palash	1.39	1.37	1.39	1.29	-	1.31
Bot	1.33	1.31	1.32	1.37	1.31	-

5.4.3 Data dimensionality and similarity

After performing data dimensionality, the eigen values lay between 103.88 to 1.07 (Appendix.2). The first 34 MNF bands had shown good result and exhibited better

spectral information. These bands were used for tree species classification. The spectral similarity result (field spectra vs. Hyperion image spectra) is shown in Table 5.6.

Table 5.6: Spectral similarity values between Hyperion image and ground reflectance spectra

S. No	Species Botanical Name	Common/Local Names	SAM Score
1	Shorea robusta)	Sal	0.81
2	Tectona grandis	Teak	0.78
3	Acacia auriculiformis	Akasmani	0.71
4	Ficus benghalensis	Bot or wad	0.68
5	Madhuca latifolia	Mohwa	0.63
6	Butea monosperma	Palash	0.69

The similarity scores indicated that spectral similarity ranged between high to medium. The spectral similarity score for Sal and Teak trees were found highest than the other trees. Sal and Teak trees covers around 65% of the study area (FSI report, 2015). The spectral variations of different tree species in the study area are shown in Figure.5.4.

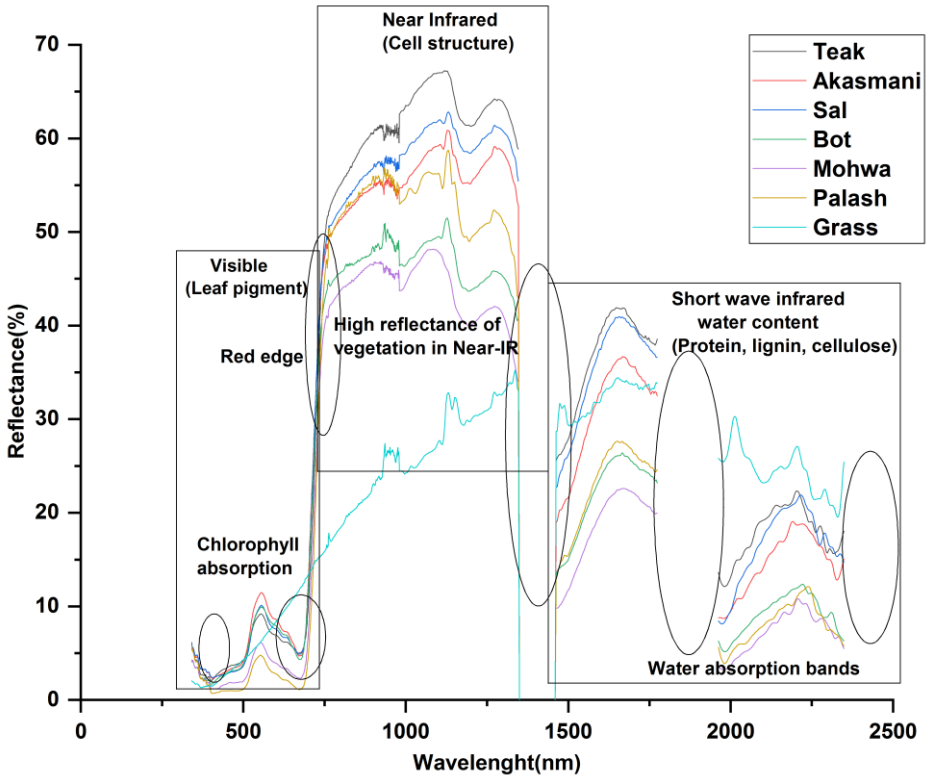

Figure 5.4: Visual comparison of resampled field average reflectance spectra for different tree species at the study area

5.4.4 Tree species classification and accuracy assessment

We have classified tree species of the study area into six different categories based on SVM, SAM, and MD algorithms using Hyperion and Landsat 8 OLI. The enlarged view of the mines and its surrounding region classified by the SVM algorithm on Hyperion data is shown in Figure. 5.5.

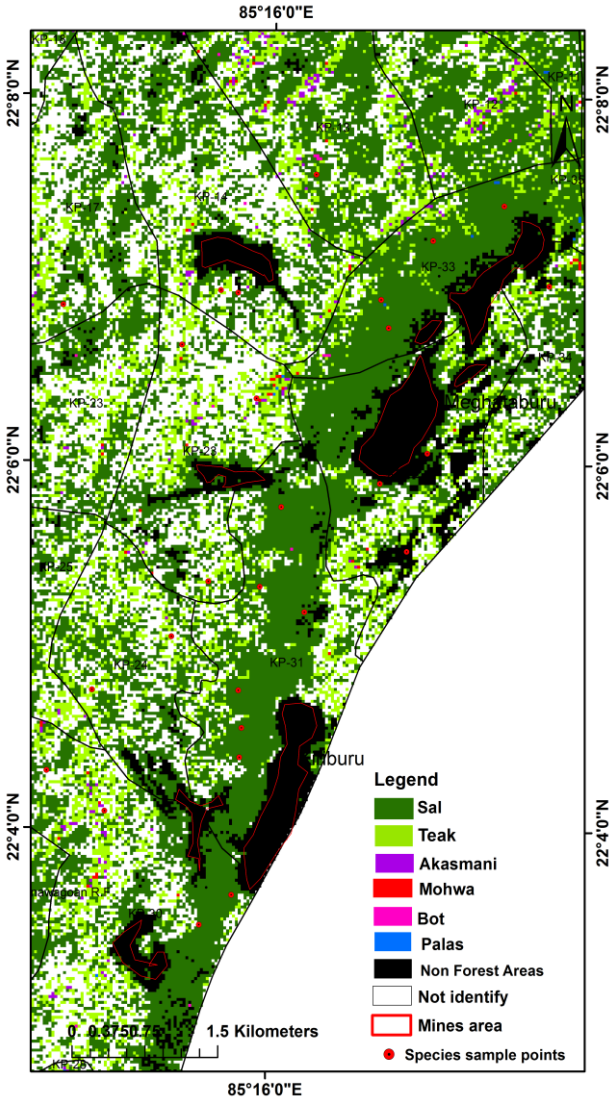

Figure 5.5. Spatial distribution of tree species mapped by SVM algorithm. based on Hyperion data

Sal and Teak trees covered most of the area. These trees were located at higher altitudes (700 to 900 meters) on the hilly side of the study region. Other trees are dominant at lower altitude (300 to 400 meter), northeast, and southeast parts of the study region. Classification accuracy estimation based on ground species spectra data shown that Hyperion image- based SVM algorithm provided better accuracy results (overall

accuracy= 85.16, kappa=0.78), than SAM algorithm (overall accuracy=7828, kappa=0.76) and MD algorithm (overall accuracy=75.58, kappa=0.73). Also, Landsat 8OLI image-based species classifications carried out by SVM algorithm; show an overall accuracy of 68.71% and a Kappa statistic of 0.66. The accuracy comparison (Hyperion based SVM, SAM, MD and Landsat 8 OLI based SVM) matrix is shown in Table 5.7.

Table 5.7: Accuracy assessment results of (a) SVM on Hyperion (b) MD on Hyperion, (c) SAM on Hyperion, and (d) SVM on Landsat

(a)	Sal	Teak	Akasmani	Mohwa	Palash	Bot	Total	UA
Sal	11	0	2	0	0	1	14	88.68
Teak	3	9	0	0	5	0	17	85.53
Akasmani	0	0	5	5	0	0	10	78.22
Mohwa	0	0	2	6	0	0	8	83.19
Palash	0	2	0	0	7	0	9	75.47
Bot	0	0	0	0	0	11	11	76.18
Total	14	11	9	11	12	12	69	
PA	89.53	83.76	81.29	82.11	84.23	83.95		
Overall accuracy: 85.16%, kappa statistics: 0.78								
(b)	Sal	Teak	Akasmani	Mohwa	Palash	Bot	Total	UA
Sal	10	0	1	0	0	1	12	78.55
Teak	1	11	0	0	1	1	14	80.78
Akasmani	0	0	8	2	0	3	13	79.11
Mohwa	0	0	2	7	0	0	9	65.83
Palash	0	2	0	0	9	0	11	81.45
Bot	0	0	0	3	0	8	11	73.27
Total	11	13	11	12	10	13	70	
PA	87.19	88.53	84.27	83.95	78.76	85.61		
Overall accuracy: 75.58%, kappa statistics: 0.73								
(c)	Sal	Teak	Akasmani	Mohwa	Palash	Bot	Total	UA
Sal	11	0	0	0	0	0	11	81.12
Teak	2	9	0	0	1	0	12	74.73
Akasmani	0	0	9	0	0	1	10	79.22
Mohwa	0	1	1	7	0	0	9	83.64
Palash	0	0	0	0	8	0	8	75.67
Bot	0	0	0	4	0	10	14	87.48
Total	13	10	10	11	9	11	64	
PA	87.43	85.92	84.28	76.38	78.84	80.47		
Overall accuracy: 79.55%, kappa statistics: 0.75								
(d)	Sal	Teak	Akasmani	Mohwa	Palash	Bot	Total	UA
Sal	12	0	0	0	0	1	13	79.53
Teak	0	8	0	1	1	0	10	73.48
Akasmani	0	0	11	1	0	1	13	
Mohwa	1	0	0	6	0	0	7	71.79
Palash	0	1	0	0	8	0	9	62.73
Bot	0	0	0	4	1	9	14	78.15
Total	13	9	11	12	10	11	66	
PA	80.44	75.18	78.59	68.15	74.72	75.27		
Overall accuracy: 68.71%, kappa statistics: 0.66								

5.4.5 Species diversity estimation and mapping

We have correlated 13 VIs with field measured Shannon index values. The regression analysis results (SSE, R^2, Adj. R^2, and RMSE) is shown in Table 5.8. The NDVI705 had shown best linear fitting (R^2 =0.76, RMSE= 0.04)) with Shannon index values.

Table 5:8 Relationship between narrow banded VIs & Shannon Index based species diversity

Narrow banded VIs	DVI	NDVI	RVI	mNDVI 705	TSAVI	NDVI 705	PVI	SAVI	NLI	mSR 705	VOG 1	MSR	TC greenness
SSE	0.56	0.07	0.08	0.30	0.37	0.35	0.32	0.34	0.39	0.108	0.43	0.13	0.13
R^2	0.43	0.71	0.52	0.47	0.29	0.76	0.31	0.28	0.39	0.43	0.26	0.52	0.37
Adj R^2	0.35	0.68	0.45	0.45	0.25	0.73	0.22	0.19	0.3	0.35	0.17	0.46	0.29
RMSE	0.19	0.07	0.07	0.14	0.15	0.04	0.14	0.15	0.16	0.08	0.16	0.09	0.09

SSE= Sum squared error, R^2= Coefficient of Determination, RMSE= Root Mean Square Error

Since, NDVI 705 correlated well with waned chlorophyll content (Kumar et al., 2015), so we have used this index for diversity estimation. The linear regression plot between narrow banded VIs and species diversity is shown in Figure. 5.6.

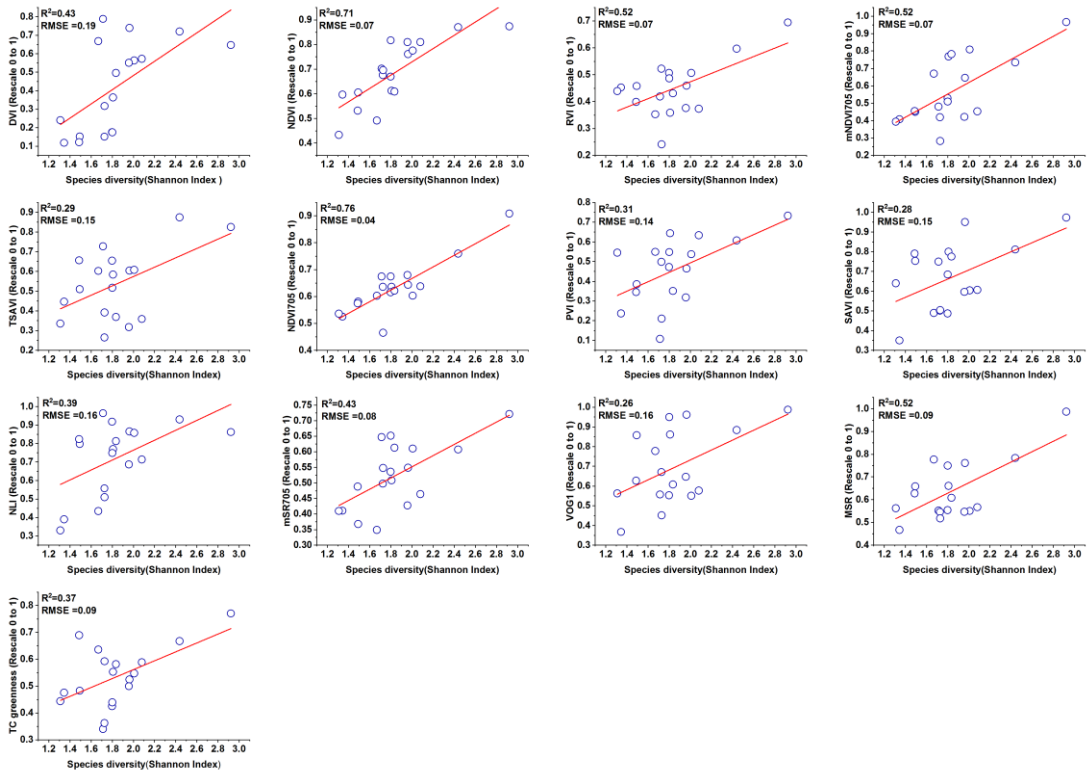

Figure 5.6 Regression between Hyperspectral narrow banded VIs and field measured Shannon Index of 18 sampling plots

Enlarged view of the species diversity map for the mines and its surrounding region is shown in Figure. 5.7.

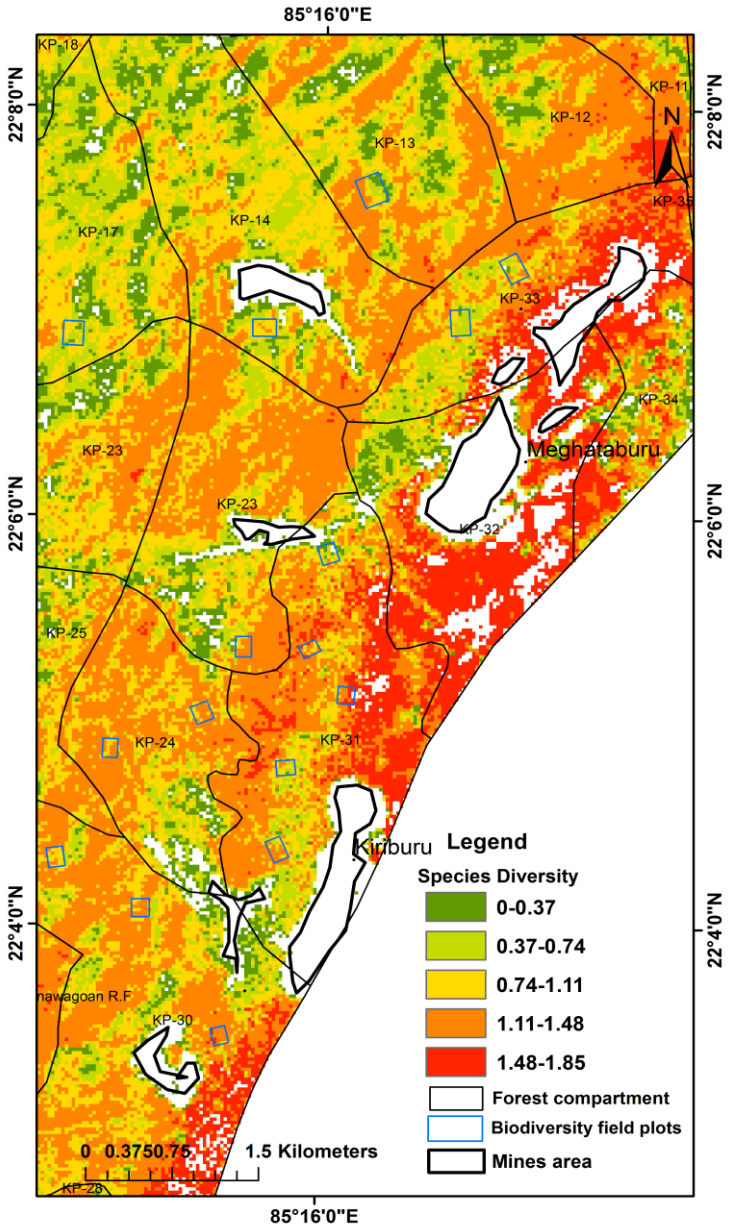

Figure 5.7. Species diversity mapped by Shannon Index based on narrow banded VIs

The linear regression between fields measured Shannon Index, and Hyperion derived Shannon Index gave the R^2 value of 0.72 and RMSE value of 0.15 (Figure. 5.8). The correlation between Hyperion and field derived Shannon index had shown better relationship (R^2 0.68).

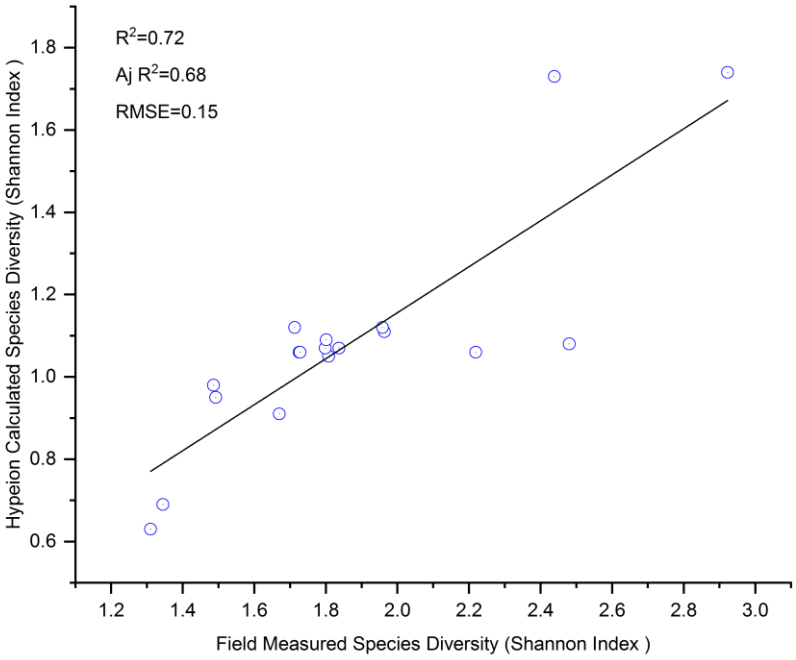

Figure 5.8. Regression between Hyperion imagery derived by Shannon index and field measured Shannon index

5.4.6 Relationship between species diversity, distance from mines and foliar dust concentration

For each sample point, values of species diversity, distance from either mines (Kiriburu and Meghataburu), and foliar dust concentration are shown in Table 5.9. Those values were used for correlations analysis using three different correlation methods (Spearman, Pearson, and Kendall). The correlation results thus obtained by the abovementioned methods are shown in Table 5.10 (for Meghahatuburu mine) and Kiriburu mine).

Table.5.9 Distance from Kiriburu and Meghataburu mines of tree species diversity with their foliar dust concentration

Sample Plots ID	Species Diversity (Shannon Index)	Foliar Dust (gm/m^2)	Kiriburu Mine from Distance (m)	Meghahatuburu Mine from Distance(m)
1	2.38	1.39	3,206	2,364
2	1.61	2.13	2,858	1,452
3	2.42	1.15	3,136	3,403
4	1.92	1.91	4,126	3,369
5	1.55	1.93	3,915	5,017
6	1.91	14.05	1,105	964
7	1.74	2.12	5,192	1,375
8	2.18	1.43	6,156	3,633
9	1.83	2.014	4,057	3,418
10	1.40	13.51	720	1,192

11	1.34	16.92	327	1,728
12	1.80	2.19	2,553	3,542
13	1.72	3.29	1,971	2,836
14	1.82	2.37	1,233	2,640
15	2.81	1.23	2992	4,672
16	1.93	4.15	1,111	2,463
17	1.99	4.85	1463	1,886
18	1.80	3.89	1,467	2,095

Table 5.10: Spearman, Pearson and Kendall correlation matrix amongst species diversity, foliar dust concentration and mines distance to Meghahatuburu and Kiriburu.

Spearman			Distance(m) from Meghahatuburu Mine
	Species Diversity (Shannon Index)	Foliar dust (gm/m^2)	
Species Diversity (Shannon Index)	1.00	-0.58	0.24
Foliar dust (gm/m^2)	-0.58	1.00	-0.67
Distance(m) to Meghahatuburu Mine	0.24	-0.67	1.00
Pearson			Distance(m) to Meghahatuburu Mine
	Species Diversity (Shannon Index)	Foliar dust (gm/m^2)	
Species Diversity (Shannon Index)	1.00	-0.46	0.36
Foliar dust (gm/m^2)	-0.46	1.00	-0.59
Distance(m) to Meghahatuburu Mine	0.366	-0.59	1.00
Kendall			Distance(m) to Meghahatuburu Mine
	Species Diversity (Shannon Index)	Foliar dust (gm/m^2)	
Species Diversity (Shannon Index)	1.00	-0.43	0.15
Foliar dust (gm/m^2)	-0.43	1.00	-0.50
Distance(m) to Meghahatuburu Mine	0.15	-0.50	1.0000
Spearman			Distance(m) from Kiriburu Mine
	Species Diversity (Shannon Index)	Foliar dust (gm/m^2)	
Species Diversity (Shannon Index)	1.00	-0.58	0.38
Foliar dust (gm/m^2)	-0.58	1.00	-0.83
Distance(m) from Kiriburu Mine	0.38	-0.83	1.00
Pearson	Species Diversity (Shannon Index)	Foliar dust (gm/m^2)	Kiriburu Mine to Distance(m)
Species Diversity (Shannon Index)	1.00	-0.46	0.35
Foliar dust (gm/m^2)	-0.46	1.00	-0.65
Distance(m) from Kiriburu Mine	0.35	-0.65	1.00
Kendall	Species Diversity (Shannon Index)	Foliar dust (gm/m^2)	Distance(m) to Kiriburu Mine
Species Diversity (Shannon Index)	1.00	-0.43	0.25
Foliar dust (gm/m^2)	-0.43	1.00	-0.66
Distance(m) from Kiriburu Mine	0.25	-0.66	1.00

The correlations results thus obtained show that there exists a good negative correlation between foliar dust concentration, species diversity, and the distance from mines (Figure.5.9).

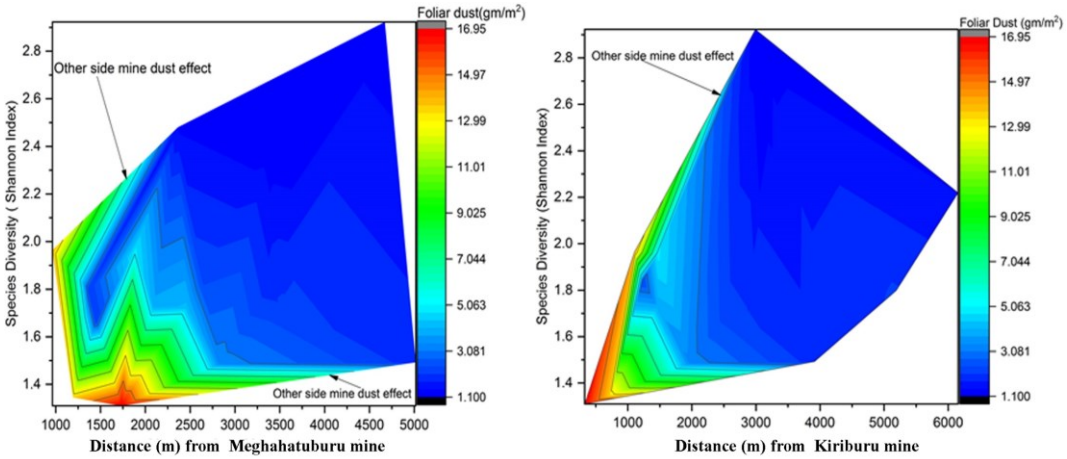

Figure 5.9. The relation amongst species diversity indices (Shannon Index) , distance from mines (Kiruburu and Meghataburu) and foliar dust concentration

5.4.7 Discussion

As per the result obtained in this study, we could infer that, Hyperspectral (Hyperion) data has more capability in tree species mapping and diversity assessment when coupled with field spectral data, than any other multispectral data (Landsat). Some researchers studied on tree species classification and diversity estimation based on hyperspectral and multispectral data at a fine-scale level. Dalponte et al., 2014 had studied on tree crown and classification using airborne hyperspectral data in boreal forest area. They had shown that hyperspectral data has better accuracy for tree species classification than other multispectral data. Shen & Cao, (2017) worked on tree species classification using hyperspectral and Lidar data in Subtropical forest area. They had used random forest classification algorithm to differentiate five tree species and provided a relatively higher accuracy (85.4%). This study has displayed a step-wise discrimination test for the identification of wavebands, which is significant for tree species classification. As obtained from the tree species discrimination analysis, 21 different spectral wavebands were selected for tree species classification, of which six belongs to the visual infrared region; eight to the near-infrared, and seven to shortwave infrared region (Table.5.11).

Table 5.11: Wilk's lambda selected wavebands and their significant for tree species

S.L No	Wavelength (nm)	Cause of Absorption	Leaf Chemicals	Reference
1	518	Electron Transition	Chlorophyll b	Curran et al., 1991
2	559	N-H stretch	Nitrogen	Curran, 1989
3	579	Electron Transition	Nitrogen	Asner,2008
4	589	Electron transition	Protein	Curran, 1989

5	599	N-H stretch	Nitrogen ,Protein	Lucas and Curran, 1999
6	660	Electron Transition	Chlorophyll b	Vyas et al.,2011
7	721	N-H stretch 1st overtone	Protein and Nitrogen	Curran, 1989
8	993	H-bend, 1st overtone	Starch	Kumar et al.,2001
9	1023	C-H stretch	Protein ,Water content	Thenkabail 2002
10	1124	H bend, 1st overtone	Water content	Vyas et al.,2011
11	1134	H stretch, C-H deformation	Moisture absorption	Serrano et al.,2002
12	1275	H bend, 1st overtone	Moisture absorption	Serrano et al.,2002
13	1326	H stretch, C-H deformation	Moisture absorption	Serrano et al.,2002
14	1336	N-H Bend, 1st overtone	Water	Kumar et al.,2001
15	1477	H stretch 1st overtone-	sugar	Lucas and Curran, 1999
16	1679	N-H stretch 1st overtone	Protein, Nitrogen, Starch	Thenkabail et al.,2004
17	1739	C-H stretch	Protein	Sobhan, 2007
18	1749	C-H stretch 1st overtone	Cellulose, sugar, starch	Lucas and Curran,1999
19	1780	C-H stretch	starch	Sobhan, 2007
20	1981	N-H asymmetry	Protein	Thenkabail et al.,2004
21	2304	N-H Stretch/C-H stretch/C-H bend, 2nd overtone	Protein , Nitrogen	Lucas and Curran,1999

Vyas et al. (2011) studied on tree species discrimination analysis, and they found 22 wavebands, of which seven falls in VIR, eight in NIR, and six bands in the SWIR region. Peerbhay et al., (2013) worked on tree species discrimination analysis in Natal, South Africa.. They found a total of 27 wavebands (8 –VIR, 12 -NIR, & 7- SWIR) from discrimination analysis, and they used those bands for tree species classification. In this work, the result obtained from J-M distance method had shown that Hyperion data -based species spectral separability value (1.25 to 1.87) was higher than Landsat 8 OLI data (1.10 to 1.39). Puletti et al., (2016) had applied the J-M method for spectral separability analysis of tree species. They found that the spectral separability value-obtained from hyperspectral data (1.17- 1.93) was higher than multispectral data (1.20-1.67). Hao et al. (2014) had used Landsat data for spectral separability analysis of tree species based on the J-M distance method. They found that the spectral separability value lay between 1.27 to 1.73 for different tree species. Some previous studies have reported that the tree species classification performed on hyperspectral data had shown better result than multispectral data. This study has shown that tree species classification based on hyperspectral data (85.16%) provided better classification accuracy than multispectral data (68.71 %,). Vyas et al., (2014) had compared species classification accuracy based on Hyperion (Accuracy 85.25%) and Landsat ETM data (Accuracy 65.25%) in Western Himalaya region, India. Lim et al., (2019) studied on tree species classification using Hyperion and Sentinel-2

satellite imagery in South Korea and China and compared the accuracy level also (Hyperion- 67% and Sentinel-2 -51%). In the study, we have used hyperspectral VIs data for species diversity estimation based on Shannon Index values. NDVI705 has shown best correlated value (R^2 =0.72) with field-based Shannon Index data as it has good sensitivity to chlorophyll content, leaf pigment, canopy structure, and canopy water content (Gitelson et al., 2005; Croft et al., 2014). So we have used the NDVI705 index for species diversity estimation. Other vegetation indices were not matched perfectly with field-based Shannon Index due otlow canopy structure, canopy water content and chlorophyll content in the study area (Tuominen et al., 2009; Sims et al., 2002). Some researchers had shown that SD and CV NDVI were best correlated with Shannon Index values for plant diversity estimation (Peng et al., 2018; Peng et al., 2019). Onyia et al., (2018) studied plant diversity in Oil polluted regions usingNDVVI (normalized difference vegetation vigour index) on hyperspectral data. They found that NDVVI was best correlated with Shannon Index values. In this study, we have correlated Hyperion and field derived Shannon Index values for result validation. The correlation results show that R^2 is 0.72, and RMSE is 0.15. These values are not matched well due to noise content in the hyperspectral data, and forest canopy problem in the study area. Jha et al., (2019) had performed correlation between AVIRIS -NG (Airborne visible/infrared imaging spectrometer -next generation) and field measured Shannon diversity Index values and found that R^2 was 0.86. Onyia et al., (2019) had correlated two species diversity results (Hyperion and Shannon Index diversity) and obtained a R2 value of 0.67.. In this study, the correlation between species diversity, foliar dust concentration, and distance from mines had shown a strong negative relationship. Kayet et al. (2019) showed a better negative relationship between forest health, distance from mines, and foliar dust deposition. Tuominen et al. (2009) had shown a clear negative relationship between leaf reflectance and trees distance from mines.

This study involved the tree species classification and diversity estimation. Some errors obtained in the study are shown in regression analysis graph. Many reasons are contributing to the error in tree species classification and diversity estimation. Hyperion data exhibits higher noise ratio and get affected by atmospheric components. It could have induced some error to the study results (Shaw et al., 2003). The spatial resolution of the Hyperion image is 30m, so the mixed pixel problem arose for species classification and

diversity estimation (Lee & Lathrop, 2005). The spectroradiometer instrument collected some self-generated noise during field spectra collection. It may haveeffect on the results (Vaiphasa et al., 2006). Due to the location of the study area on the hills, the satellite imagery gets infected with shadow error (Adler et al., 2001). Forest canopy can induce the problem of image spectral segregation (Ustin et al., 2004). The study area has a canopy density cover of about 30 to 40 %.

5.5 Summary

The potential utility of hyperspectral data demonstrated the in discriminating the tree species and classifying the tree species diversity in hilltop mining areas. The pre-processing of 242 Hyperion (narrow bands) spectral wavebands resulted into 145 corrected spectral wavebands. The 21 spectral wavebands were selected through discrimination analysis (Wilk's Lambda test). The SVM, SAM, and MD algorithms were applied for tree species classification based on field spectra data. We have identified six species (Sal, Teak, Akasmani, Mohwa, Palash, and Bot) in the study area at the spatial level. The hyperspectral vegetation indices (VIs) were used to estimate species diversity based on field measured Shannon Diversity Index. Regression analysis between Hyperion imagery derived from Shannon index and field measured Shannon index have been done for validation purposes. As well as show the relationship among species diversity and foliar dust concentration as a function of distance from mines. The methodology adopted by us can also be applied to other forest areas in the vicinage of the mines, and it could serve as the base for future work for forest management and geo-environmental planning.