Spectrum Science Inc.
Science Communication Initiative

International Journal of
Applied Remote Sensing

Volume 2 Number 1 (2008) 1-14

Endmembers discrimination in MODIS using spectral angle mapper and maximum likelihood algorithms

UTTAM KUMAR AND T. V RAMACHANDRA
Centre for Ecological Studies
Indian Institute of Sciences, Bangalore, India

ABSTRACT

Land cover (LC) describes the physical state of the earth's surface in terms of the natural environment (vegetation, waterbodies, etc.) and the human-made structures (built up). These features can be extracted from RS data of different spatial (Cartosat, IKONOS PAN, etc.), spectral (Multispectral, superspectral and hyperspectral) and temporal resolutions. LC mapping can be performed by processing these data into different themes or classes using supervised and unsupervised classification approaches. Robust algorithms are available for classifying multispectral data. However, classification of coarse spatial resolution data such as superspectral for discerning LC entities requires different strategies. This communication evaluates the suitability of algorithms such as Spectral Angle Mapper (SAM) and Gaussian Maximum Likelihood classifier (GMLC) for LC mapping using (i) MODIS 7 land bands product, (ii) Principal component and (iii) Minimum Noise Fraction components, of the MODIS 36 bands of 250 m to 1 km spatial resolution. The accuracy assessment shows that GMLC on MODIS bands 1 to 7 and SAM on the same 7 bands performed well (with 76% and 70% accuracy) in LC mapping compared to others. Validation reveals utility of these techniques for LC mapping from superspectral data at regional scale.

Keywords: Land cover, Superspectral, MODIS, Spectral Angle Mapper, Gaussian Maximum Likelihood Classifier, GRDSS

1.0 INTRODUCTION

Remote sensing (RS) techniques have been widely used since past four decades for inventorying, mapping and monitoring of natural resources. Diverse earth resources could be discriminated due to the dissimilarity among their spectral reflective properties. Recent trends in RS technologies have lead to the improvements in sensor design, which have helped in acquiring the data with better spatial and spectral resolutions. These data could be analysed using either visual interpretation or automated digital processing techniques Visual interpretation is based on interpretation keys such as tone, texture, shape, size and context. Compared to this, each individual pixel in an image is classified or grouped based on the spectral information in the digital image classification approaches. The supervised and unsupervised classification algorithms have been widely used for land use mapping exercises.

Algorithms based on supervised classification techniques, classify pixels considering the training site signatures as is done in Gaussian Maximum Likelihood classification (GMLC). This approach is based on probability density function associated with a particular signature (training site signatures). Pixels are assigned to most likely class based on a comparison of the posterior probability that it belongs to each of the signatures being considered. In contrast to this approach, Spectral Angle Mapper (SAM) compares each pixel in the image with every endmember for each class and assigns a ponderation value

ISSN : 0973-5860

between 0 (low resemblance) and 1 (high resemblance). Endmembers are taken directly from the image and compared with signatures based on field experiments. In this paper, we evaluate the suitability of superspectral data (different MODIS products) for LC mapping using GMLC and SAM techniques at a regional scale.

GMLC and SAM algorithms for classification of vegetation in southern Finland using AISA airborne imaging spectrometer of 1.1 m spatial resolution with 17 visual and near infrared bands data (Lumme, J. H., 2004) show GMLC classifier led to better results with an overall accuracy of 91%, but the results deteriorated under varying illumination and required more computation time. SAM was comparatively faster and they led to good classification results in poor illumination. Similar analysis with the hyperspectral data of 1 m spatial resolution and 20 spectral bands acquired by the AISA airborne imaging spectrometer onboard the NOMAD GAF-27 aircraft (Shafri, H. Z. M. *et al.*, 2007) show the relative performance of GMLC and SAM as 86%, and 49% respectively. The role of different spatial resolutions in vegetation studies using MIVIS airborne hyperspectral data (102 spectral channels: 20 channels in the VIS-NIR wavelength range (0.43-0.83 μm), 8 channels in the near-infrared (1.15-155 μm), 64 channels in the mid-infrared (1.98-2.47 μm) and 10 in the thermal infrared (8.18-12.7 μm)) in the forest area at different altitudes of 2000 m and 5000 m in Central Italy (Salvatori, R., 2003) revealed that GMLC has a better distinction among different vegetation units (with 99.4% on 4 m and 99.6% on 10 spatial resolution), while SAM had an overall accuracy of 69% with 4 m and 84 % on 10 m spatial resolution.

This endeavour evaluates the suitability of GMLC and SAM algorithms for LC (agriculture, built up (urban/rural), forest, plantation/orchard, waste land/barren rock/sheet rock and water bodies) mapping using superspectral data (MODIS) with spatial resolutions ranging between 250 m to 1 km and spectral resolution ranging from 7 to 36 bands with wavelength spread from 405 nm to 14.385 μm (covering visible and IR). This analysis was done using GRDSS (Geographic Resources Decision Support System) free and open source software (GRASS mirror site in India is at http://wgbis.ces.iisc.ernet.in/grass). It has the capabilities to capture, store, process, display, organize, and prioritize spatial and temporal data. It serves as a decision support system for decision making and resource planning, with functionalities for raster analysis, vector analysis, modelling and visualization (Ramachandra T. V. & Kumar, Uttam., 2004, Ramachandra T. V. *et al.*, 2004).

1.1 Spectral Angle Mapper (SAM)

In N dimensional multi-(or super) spectral space a pixel vector x has both magnitude (length) and an angle measured with respect to the axes that defines the coordinate system of the space (Richards *et al*, 1996). Only the angular information is used for identifying pixel spectra in the SAM technique based on the assumption that an observed reflectance spectrum is a vector in a multidimensional space, where the number of dimensions equals the number of spectral bands. It permits rapid mapping by calculating the spectral similarity between the image spectrums to reference spectra (Yuhas *et al.*, 1992; Kruse *et al.*, 1993; Van der Meer *et al.*, 1997; Crosta *et al.*, 1998; De Carvalho *et al*, 2000; Schwarz *et al*, 2001; Hunter *et al.*, 2002). It takes the arccosine of the dot product between the test spectrum "t" to a reference spectrum "r" as in equation 1.

$$a = \cos^{-1}\left(\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} t_i^2}\sqrt{\sum_{i=1}^{nb} r_i^2}} \right) \quad \text{..... (equation 1)}$$

where,

nb = the number of bands, t_i = test spectrum, r_i = reference spectrum

To compare two spectra, such as an image pixel spectrum and a library reference spectrum, the multidimensional vectors are defined for each spectrum and the angle between the two vectors is calculated. Smaller angles represent closer matches to the reference spectrum. If this angle is smaller than a given tolerance level, the spectra are considered to match, even if one spectrum is much brighter than the other (farther from the origin). Pixels farther away from the specified maximum angle threshold are not classified (Lillesand *et al.*, 2002). This is an easy and rapid method for mapping the spectral similarity of image spectra to reference spectra SAM represses the influence of shading effects to accentuate the target reflectance characteristics (De Carvalho *et al*, 2000) and hence considered a very powerful classification method and functions well even with scaling noise. SAM is invariant to unknown multiplicative scalings, and consequently, is invariant to unknown deviations that may arise from different illumination and angle orientation (Keshava, *et al*, 2002). There is a high likelihood that angular information alone will provide good separation, when the pixel spectra from the different classes are well distributed in feature space. However, the spectral mixture or mixed pixels pose problems due to the assumption that endmembers represent the pure spectra of a reference material. This could happen even with the medium spatial resolution (30 m) images. SAM technique fails if the vector magnitude is important in providing discriminating information, which happens in many instances (Richards *et al.*, 1996).

Girouard *et al*, (2004) validated SAM for geological mapping in Central Jebilet Morocco and compared the results between high and medium spatial resolution sensors, such as Quickbird and Landsat TM, respectively. The result showed that SAM of TM data can provide mineralogical maps that compare favourably with ground truth and known surface geology maps. Even though, Quickbird has a high spatial resolution compared to TM; its data did not provide good results for SAM because of low spectral resolution.

1.1 Gaussian Maximum Likelihood algorithm (GMLC)

Gaussian maximum likelihood classifier assumes that the distribution of the cloud of points forming the category training data is Gaussian (normally distributed) and classifies an unknown pixel based on the variance and covariance of the category spectral response patterns. This classification is based on probability density function associated with a particular signature (training site). Pixels are assigned to most likely class based on a comparison of the posterior probability that it belongs to each of the signatures being considered Under this assumption, the distribution of a category response pattern can be completely described by the mean vector and the covariance matrix. With these parameters, the statistical probability of a given pixel value being a member of a particular LC class can be computed (Lillesand *et al.*, 2002). GMLC can obtain minimum classification error under the assumption that the spectral data of each class is normally distributed. It considers not only the cluster centre but also its shape, size and orientation by calculating a statistical distance based on the mean values and covariance matrix of the clusters A major drawback of this classifier is that large number of computations is required to classify each pixel. This is particularly true when either a large number of spectral channels are involved or a large number of spectral classes are to be differentiated. In such cases, the GMLC is much slower computationally than other algorithms In addition, it is required to have the number of pixels per training set be between 10 times and 100 times as large as the number of sensor bands, in order to reliably derive class-specific covariance matrices.

An extension of the maximum likelihood approach is the *Bayesian classifier, which* applies two weighting factors to the probability estimate. First, the a priori probability, or the anticipated likelihood of occurrence for each class in the given scene is determined. Secondly, a weight associated with the cost of misclassification is applied to each class. Together, these factors act to minimize the

cost of misclassification, resulting in a theoretically optimum classification. Most maximum likelihood classifications are performed assuming equal probability of occurrence and cost of misclassification for all classes. If suitable data exist for these factors, the Bayesian implementation of the classifier is preferable

1.2 Endmember selection

The reflected spectrum of a pure feature is called a reference or *endmember* spectrum. Various techniques, such as the Pixel Purity Index (PPI), Scatter plot, Orasis (Optical real-time Adaptative Spectral Identification System), N-FINDR, Iterative Error Analysis (IEA), Convex Cone Analysis (CCA), Automated morphological endmember extraction (AMEE) and Simulated Annealing Algorithm (SAA) have been developed to extract endmember spectra automatically from remotely sensed data (Bateson, C A , 2000; Boardman, J., 1995; Plaza, A , 2004; Winter, M. E. and Winter, E. M , 2000).

The spectra are the points in an n-dimensional scatter plot, where *n* is the number of bands. The coordinates of the points in n-space consist of 'n' values that are simply the spectral reflectance values in each band for a given pixel. The distribution of these points in n-space was used to estimate the number of spectra and their spectral signatures, and provided an intuitive mean to understand the spectral characteristics of the land cover types. The spectral characteristics of the endmembers were analysed by plotting the endmembers and obtaining the transformed divergence matrix which showed a clear separability between the endmembers These endmembers were then rotated using the n-dimensional visualiser. Snapshots of the visualisation are shown in Figure 1. The endmembers are marked using white circles. It shows that the endmembers selected for the analysis are well separable and can be distinguished from each other.

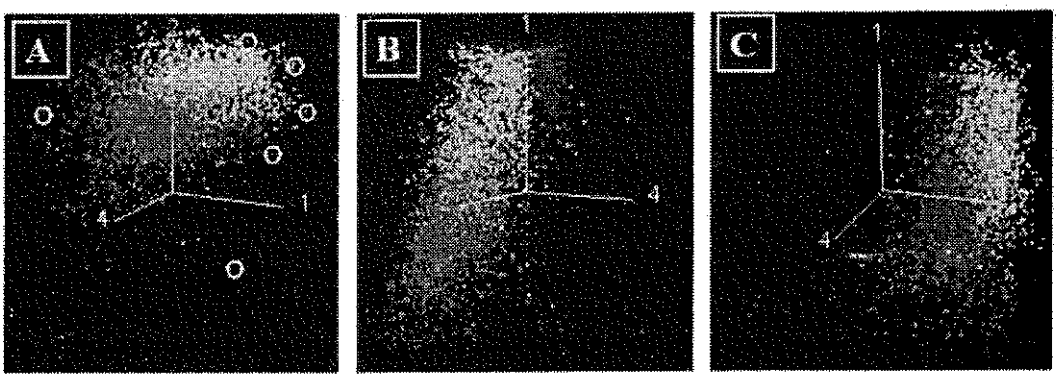

Figure 1: Three dimensional visualisation of the endmembers.

2.0 Materials and methods

2.1 Study area

The study area, Kolar is located in Karnataka state, India. It lies in the plain regions (semi arid agro-climatic zone) extending over an area of 8238 sq km. between 77°21' to 78°35' E and 12°46' to 13°58' N (shown in figure 2) For administrative purposes, it is divided into 11 taluks (or administrative boundaries /blocks/units). The distribution of rainfall is during southwest and northeast monsoon seasons. The average population density of the district is about 2 09 persons/hectare. The district forms part of northern extremity of the Bangalore plateau. The area is often subjected to recurring drought. The

rainfall is both scanty and erratic in nature. The district is devoid of significant perennial surface water resources leading to limited ground water potential. The terrain has a high runoff due to less vegetation cover contributing to erosion of top productive soil layer leading to poor crop yield. Out of about 280,000 hectares of land under cultivation, 35% is under well and tank irrigation (Ramachandra *et al.*, 2005).

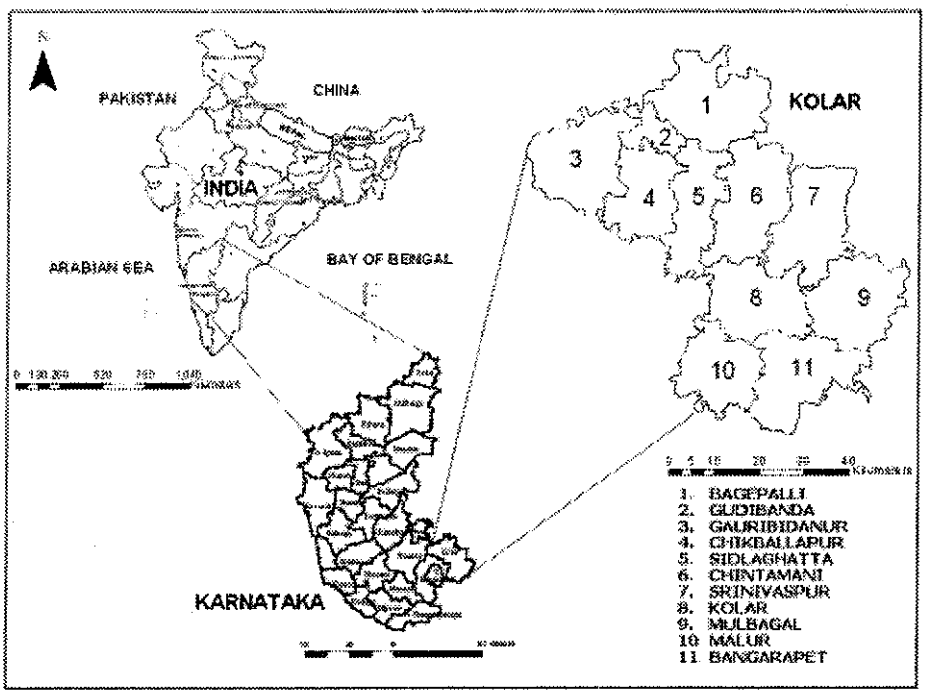

Figure 2: Study area – Kolar district, Karnataka State, India.

2.2 Data

MODIS data were downloaded from the Earth Observing System (EOS) Data Gateway (http://edcimswww.cr.usgs.gov/pub/imswelcome/). These data sets are known as "MOD 09 Surface Reflectance 8-day L3 global" product with spatial resolutions 250 m (bands 1 and 2) and 500 m (bands 3 to 7). The MODIS Surface-Reflectance Product (MOD 09) is computed from the MODIS Level 1B land bands 1, 2, 3, 4, 5, 6, and 7 (centered at 648 nm, 858 nm, 470 nm, 555 nm, 1240 nm, 1640 nm, and 2130 nm, respectively). The product is an estimate of the surface spectral reflectance for each band as it would have been measured at ground level if there were no atmospheric scattering or absorption (http://modis.gsfc.nasa.gov/data/dataprod/dataproducts.php?MOD_NUMBER=09). Each MODIS Level-1B data product (Guenther *et al*, 1998), contains the radiometrically corrected, fully calibrated and geolocated radiances at-aperture for all 36 MODIS spectral bands at 1 km resolution http://daac.gsfc.nasa.gov/MODIS/Aqua/product_descriptions_modis.shtml#rad_geo). These data are broken into granules approximately 5-min long and stored in Hierarchical Data Format (HDF). Bands 1 to 36 MODIS data "MOD 02 Level-1B Calibrated Geolocation, Data Set" were downloaded from EOS Data Gateway (http://edcimswww.cr.usgs.gov/pub/imswelcome/). The Level 1B data set contains calibrated and geolocated at-aperture radiances for 36 bands generated from MODIS Level 1A sensor counts

(MOD 01) The radiances are in W/(m^2 μm sr). In addition, Earth BRDF may be determined for the solar reflective bands (1-19, 26) through knowledge of the solar irradiance (e.g., determined from MODIS solar-diffuser data and from the target-illumination geometry). Additional data are provided, including quality flags, error estimates, and calibration data (http://modis.gsfc nasa gov/data/dataprod/dataproducts php? MOD_NUMBER=02). The Indian Remote Sensing (IRS) Satellites-1C/1D LISS-III (Linear Imaging Self-Scanning Sensor III) MSS (Multi Spectral Scanner) having 3 bands (G, R and NIR) data with a spatial resolution 23.5 m were procured from NRSA, Hyderabad. The main sources of primary data are from field (using GPS), the Survey of India (SOI) toposheets of 1:50,000, 1:250,000 scale and the secondary data were collected from the government agencies (Directorate of census operations, Agriculture department, Forest department and Horticulture department) etc.

2.3 Methods

The methods adopted in the analysis involved:

1. Creation of base layers like district boundary, district with taluk and village boundaries, road network, drainage network, mapping of water bodies, etc. from the SOI toposheets of scale 1:250000 and 1:50000.
2. Identification of ground control points (GCP's) and geo-correction of MODIS (MOD 09 Surface Reflectance 8-day L3 global Products) band 1 and 2 (spatial resolution 250 m) and bands 3 to 7 (spatial resolution 500 m) and MOD 02 Level-1B Calibrated Geolocation Data Set band 1 to band 36 (spatial resolution 1 km).
3. Resampling of MODIS bands 3 to 7 (MOD 09 product) and MODIS bands 1 to 36 (MOD 02 product) to 250 m using nearest neighbourhood technique for easy processing, overlaying and comparison for analysis consistency.
4. Reprojection from Sinusoidal projection to Polyconic projection with Evrst 1956 as the datum, followed by masking of the study area
5. Principal Component Analysis (PCA) of the MODIS 36 bands and derivation of Principal Components
6. Derivation of Minimum Noise Fraction (MNF) of the MODIS 36 bands. It helped in to determine the inherent dimensionality of the data, to reduce noise and computational requirements for subsequent processing.
7. Classification of MODIS data using SAM and GMLC
8. Extraction of high resolution LISS-III bands from NRSA dataset (which were in BIL format), identification of ground control points (GCP's) and geo-correction of the bands through resampling followed by cropping and mosaicing of data corresponding to the study area.
9. Generation of FCC (False Colour Composite) and identification of training sites (heterogeneous patches) on FCC. Heterogeneous patches correspond to various land features. Training sites are chosen so as to represent all land categories and uniformly spread all over the study area
10. Collection of attribute information from field corresponding to the chosen training sites using GPS.
11. The training polygons collected from the field were overlaid on the FCC of the image and a region of interest (ROI) was created, thus enabling direct selection of assumed pure pixels (endmembers)
12. Scatter plots of bands helped in locating some of the purest endmembers by taking the extreme corner pixels. Finally, these spectra obtained by different methods (ROI and scatter plots) were merged into six classes
13. Supervised Classification of LISS-III MSS data using GMLC.
14. Accuracy Assessment of LISS-III MSS classified data.
15. Accuracy assessment of MODIS classified data (using SAM and GMLC) and comparison with MSS classified data (pixel by pixel).

3.0 RESULTS AND DISCUSSION

3.1 Classification of MODIS data using SAM

Class spectral characteristics plot for the six classes across the first seven bands, PC's and MNF components of the 36 bands showed distinct pattern for respective categories. The similar pattern was observed in the transformed divergence matrices. The MODIS data were classified using SAM technique. At an angle of 0.5 radians, the classified output of the MODIS bands 1 to 7 represented the 6 land cover classes. The first five PC's were then classified using SAM (the spectral angle set for classes are: 5 for agriculture, 5 for built up, 0.05 for forest, 0.02 for plantation, 3 for forest and 0.02 for water bodies) after iterations of tuning the angle between the class spectra and the reference spectra to avoid misclassifications. The five MNF components were also classified using the SAM. At an angle of 5 radians, all the classes represented the land cover classes well (except water, which was classified at an angle of 0.2 radians). The classified maps obtained using SAM are as shown in figure 3 (C), (D), and 4 (A).

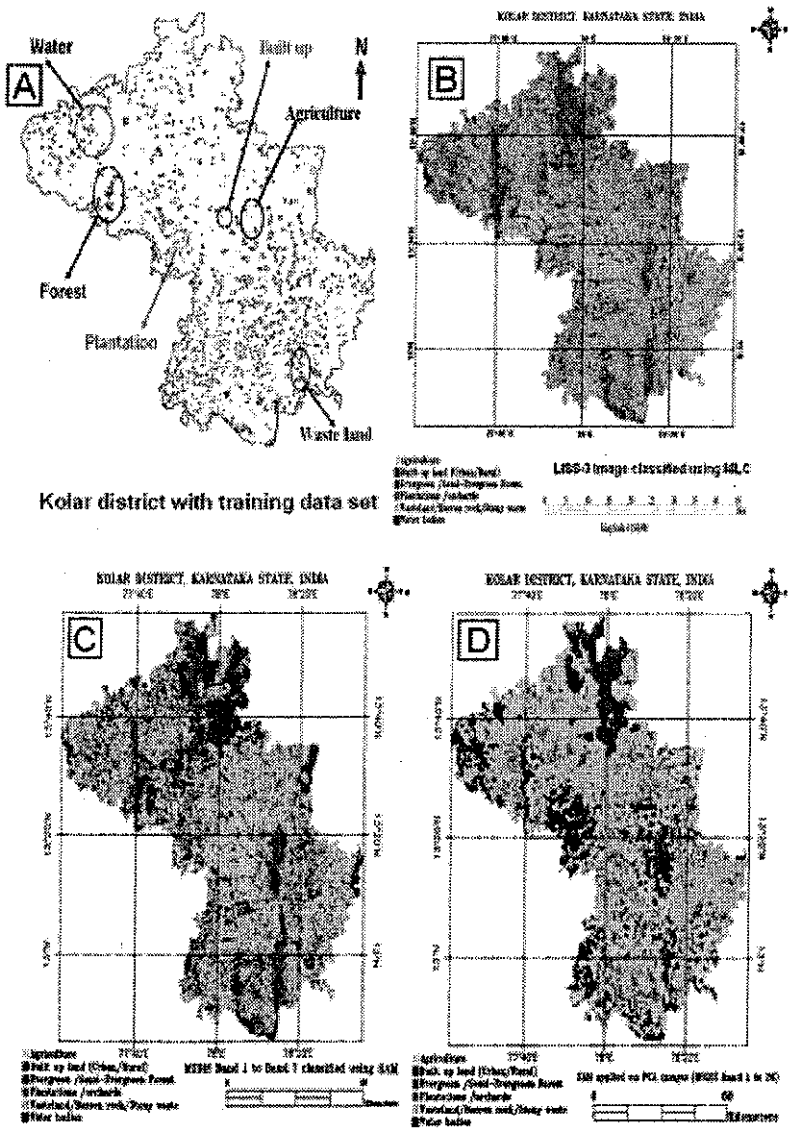

Figure 3: (A) Distribution of training data set over the district (B) MLC on LISS-III MSS (C) SAM on MODIS bands 1 to 7 (D) SAM on PCs.

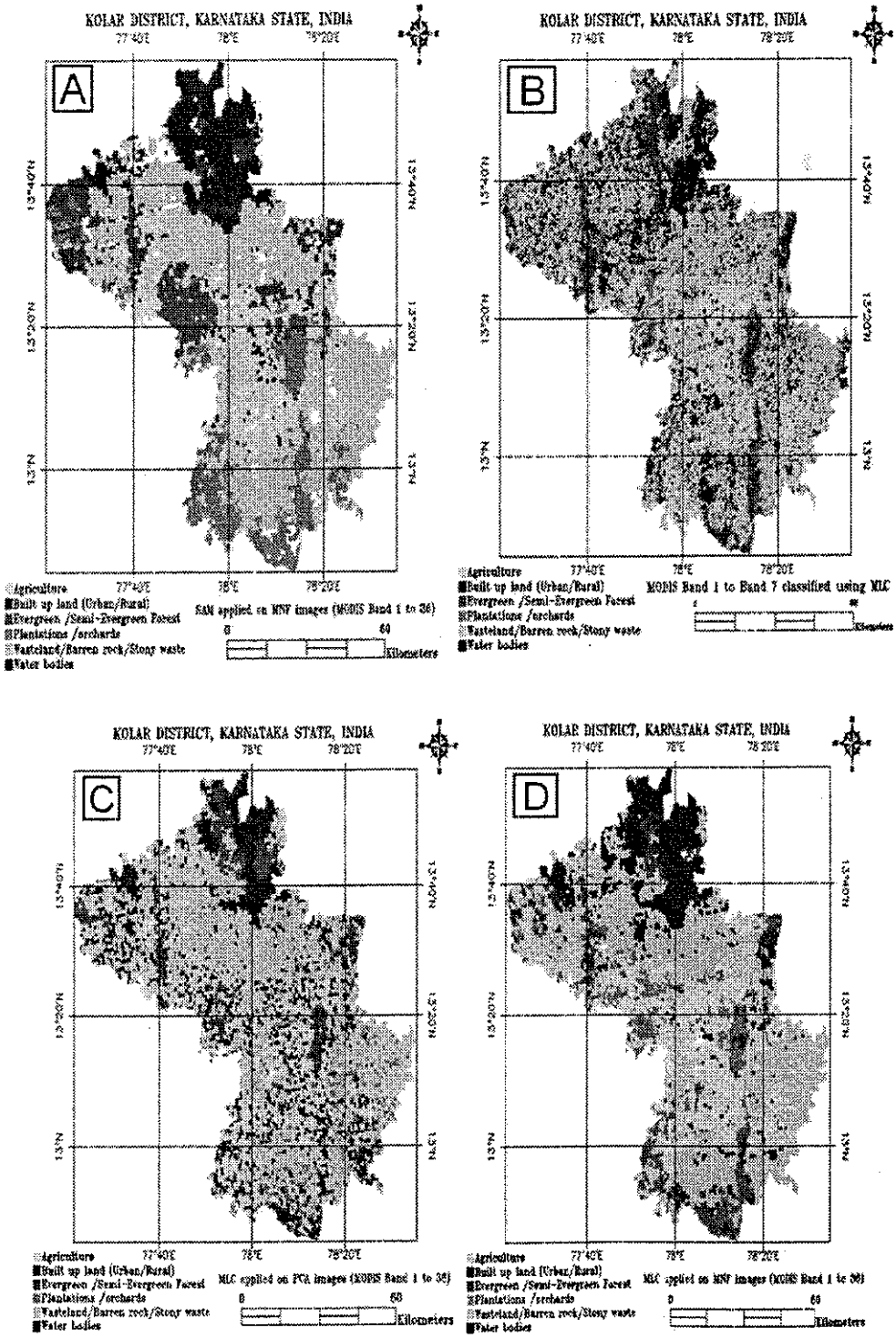

Figure 4: (A) SAM on MNF components (B) MLC on MODIS bands 1 to 7 (C) MLC on PC's (D) MLC on MNF components.

3.2 Classification of MODIS data using GMLC

FCC was created using MODIS bands 4 (Green), 1 (Red) and 2 (NIR). The heterogeneous patches were identified in the FCC as training regions for attribute data collection

MODIS bands 1 to 7 were classified by GMLC. At a threshold of 0.001, these 7 bands gave a good representation of the land cover classes. Similarly, PC's of 36 bands were classified by GMLC maintaining a threshold of 0 001. The pixels in the MNF components were not very distinct and were clustered into sub groups (comprising of two or three pixels). Though, the class separation was possible, yet it was difficult to classify each pixel based on signature, since the image was slightly pixelated. The MNF components were also classified at a threshold of 0 001 as shown in figure 4 (B), (C) and (D)

3.3 Classification of LISS-III MSS data

The class spectral characteristics for the six land cover classes considering LISS-III MSS bands 2, 3 and 4 were generated to see separability among classes. The transformed divergence matrix also helped in distinguishing different classes. FCC was generated from the LISS-III MSS data. The heterogeneous patches (training polygons) were chosen for the field data collection. This data with other ancillary data were used for classifying LISS-III MSS data using GMLC. Care was taken to see that these training sets are uniformly distributed representing / covering the study area as shown in figure 3 (A). The supervised classified image shown in figure 3 (B) was validated by field visit and by overlaying the training sets used for classification. The land cover statistics are listed in table 1. Classified LISS-III data aided as a reference high resolution image (with a spatial resolution of 23.5 m). Table 1 provides a comparative statistics of land cover using MODIS and LISS data with different classification approaches

Table 1: Comparative analysis of LC classes (%) obtained from LISS-III MSS (using MLC) and MODIS classification on Surface Reflectance Bands (1 to 7), Principal Components and MNF components of MODIS bands 1 to 36 using SAM and MLC.

Classes	LISS - III	MODIS Surface Reflectance bands (Bands 1 to 7)		MODIS derived PCs (36 bands)		MODIS derived MNF Components (36 bands)	
		SAM	MLC	SAM	MLC	SAM	MLC
Agriculture (%)	19.03	17.88	20.76	19.78	25.93	19.55	26.63
Built up (Urban/Rural) (%)	17.13	20.55	15.54	11.46	16.66	12.20	14.33
Evergreen/ Semi-Evergreen Forest (%)	11.41	8.37	10.41	10.73	10.78	14.30	7.29
Plantation/ orchards (%)	10.96	11.81	13.19	13.16	8.85	13.07	11.12
Waste land/Barren Rock / Stony waste (%)	40.39	40.92	39.11	44.47	37.16	40.21	40.27
Water bodies (%)	1.08	00.47	00.99	0.40	0.62	00.67	00.36
Total (%)	100.00	100.00	100.00	100.00	100.00	100.00	100.00

3.4 Accuracy Assessment

3.4.1 Accuracy Assessment of LISS-III classified map

The accuracy assessment was done by collecting field data for Chikballapur taluk (which constitutes 10% of the study area – Kolar district) The producer's, user's accuracy and overall accuracy corresponding to the various categories were computed, along with the error matrices for supervised classified MSS data of LISS-III, which is summarised in table 2, 3 and 4. The LISS-III supervised classification accuracy assessment gave a *kappa* (k) value of 0.95 indicating that an observed classification is in agreement to the order of 95 percent.

Table 2: User's Accuracy of classified LISS-III and MODIS Data of Chikballapur taluk

Algorithms	Agriculture	Built up	Forest	Plantation	Waste land	Water bodies
LISS-III (MLC)	**94.21**	**96.47**	**94.73**	**92.27**	**97.49**	**96.13**
SAM (B1 to B7)	46.63	71.71	84.05	61.18	**87.53**	45.61
SAM (PCA)	53.44	65.31	88.32	54.97	73.99	61.45
SAM (MNF)	57.45	49.94	53.22	63.07	81.05	**70.34**
MLC (B1 to B7)	**84.01**	**84.77**	94.21	51.33	84.33	51.49
MLC (PCA)	71.22	69.99	89 05	**86.48**	67.33	64.41
MLC (MNF)	43.33	61.02	**91.23**	61.22	85.11	41.22

Table 3: Producer's Accuracy of LISS-III and MODIS classified Data of Chikballapur

Algorithms	Agriculture	Built up	Forest	Plantation	Waste land	Water bodies
LISS-III (MLC)	**84.54**	**83.11**	**96.20**	**91.73**	**89.88**	**98.33**
SAM (B1 to B7)	63.49	81.53	71.42	27.99	51.61	70.01
SAM (PCA)	**86.79**	**86.21**	24.52	67.05	33.09	44.07
SAM (MNF)	47.66	31.89	31.42	51.01	45.51	67.22
MLC (B1 to B7)	76.49	81.66	**73.64**	36.31	76.82	89.35
MLC (PCA)	53.21	69.78	48.26	**69.91**	**92.20**	91.11
MLC (MNF)	71.66	56.77	45.35	66.13	45.77	**99.00**

Table 4: Overall Accuracy

Techniques	Overall Accuracy
MLC on LISS-III	**95.63**
SAM on MODIS Surface reflectance bands (B1 to B7)	69.41
SAM on MODIS derived PCs (36 bands)	35.22
SAM on MODIS derived MNF Components (36 bands)	49.27
MLC on MODIS Surface reflectance bands (B1 to B7)	**75.99**
MLC on MODIS derived PCs (36 bands)	30.44
MLC on MODIS derived MNF Components (36 bands)	42.22

3.4.2 Accuracy Assessment of MODIS classified Maps

Accuracy assessments of classified data were done 1) using field data, 2) comparing LC percentage area for different taluks across various techniques and data, and 3) pixel to pixel comparison with the classified LISS-III data

3.4.2.1 Accuracy Assessment using field data

User's, producer's and overall accuracy assessment of the MODIS classified maps were done with the field data and the results are listed in tables 2, 3 and 4 respectively. Accuracy assessment of MODIS classified maps were also performed at two spatial scales – at the administrative boundary level (taluk) and at the pixel level.

3.4.2.2 Comparison based on land cover class percentage area

LC statistics were computed for all taluks pertaining to each classification algorithm and is given in table 5 The analysis showed that MLC on MODIS band 1 to 7 was better than all other techniques for mapping agriculture followed by SAM on MNF components whereas SAM on MODIS B1 to B7 performed poorly on this class. MLC on MODIS band 1 to 7 is also good for mapping built up areas, forest, plantation and waste land On the other hand, MLC on PC's was good for mapping water bodies

Table 5: Land cover statistics for Chikballapur

Algorithms	Agriculture	Built up	Forest	Plantation	Waste land	Water bodies
MLC (LISS-III)	28.07	11.51	17.06	11.56	31.04	0.76
SAM (B1 to B7)	48.16	21.38	12.97	14.91	35.38	0.5
SAM (PCA)	17.4	8.28	15.37	14.96	43.82	0.16
SAM (MNF)	30.73	0	28.74	4.02	36.21	0.3
MLC (B1 to B7)	28.31	11.71	16.35	10.4	32.72	0.51
MLC (PCA)	40.27	7.09	13 06	4.89	34.01	0.68
MLC (MNF)	23.72	3.48	14.62	16.1	41.87	0.21

3.4.2 Pixel to pixel analysis with LISS-3 MSS classified image

Classified data of MODIS and LISS-III MSS were compared on a pixel by pixel basis for accuracy assessment of pure (homogenous) pixels One pixel of MODIS spatially corresponds to 121 pixels (that is approximately equal to 258.5 m) of LISS-III (figure 5). The error matrix was generated with user's accuracy, producer's accuracy and overall accuracy for the taluk and is listed in table 6, 7 and 8.

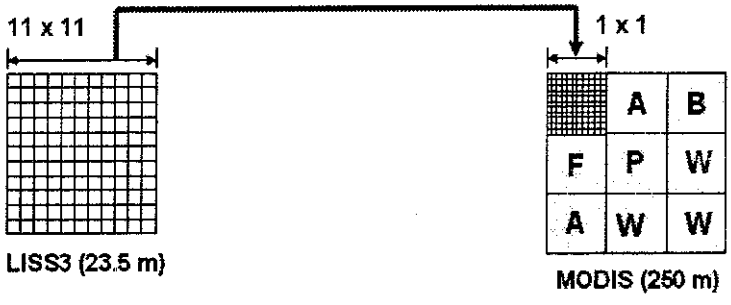

Figure 5: 11 x 11 pixels of LISS-III compares spatially to 1 x 1 pixel of MODIS

Table 6: User's Accuracy obtained from pixel to pixel analysis with LISS-III image comparison for Chikballapur taluk

Algorithms	Agriculture	Built	Forest	Plantation	Waste land	Water
SAM (B1 to B7)	37	29	**75**	64	61	49
SAM (PCA)	17	31	71	64	79	51
SAM (MNF)	21	**45**	39	61	88	33
MLC (B1 to B7)	**69**	39	**72**	65	85	51
MLC (PCA)	36	24	50	48	82	49
MLC (MNF)	40	45	53	**69**	**92**	**55**

Table 7: Producer's Accuracy obtained from pixel to pixel analysis with LISS-III image comparison for Chikballapur taluk

Algorithms	Agriculture	Built up	Forest	Plantation	Waste land	Water bodies
SAM (B1 to B7)	41	55	31	54	59	41
SAM (PCA)	21	41	29	66	**74**	37
SAM (MNF)	29	55	45	61	61	42
MLC (B1 to B7)	45	56	**66**	**73**	66	**55**
MLC (PCA)	**65**	49	32	61	66	39
MLC (MNF)	58	**57**	41	57	60	48

Table 8: Overall Accuracy.

Techniques	Overall Accuracy
SAM on MODIS Surface reflectance bands (B1 to B7)	56.66
SAM on MODIS derived PCs (36 bands)	51.38
SAM on MODIS derived MNF Components (36 bands)	47.38
MLC on MODIS Surface reflectance bands (B1 to B7)	**65.77**
MLC on MODIS derived PCs (36 bands)	53.02
MLC on MODIS derived MNF Components (36 bands)	41.55

Further, within the LC parameter, errors are generated either due to commission or omission when the signal of a pixel is ambiguous, perhaps as a result of spectral mixing, or when the signal is produced by a cover type that is not accounted for in the training process. With respect to a particular class, errors of omission occur when pixels of that class are assigned wrong labels; errors of commission occur when other pixels are wrongly assigned the label of the class considered. These errors are a normal part of the classification process, which can be minimized.

4.0 DISCUSSION

Hard classification technique perform well when the actual ground cover is heterogeneous with high spatial resolution data (i.e. LISS-III IRS data). However, classification accuracies depend on adequate number of training sites, their distribution, and overlap of season (if not day) of training data collection and RS data acquisition. MODIS classified image with coarse spatial resolution had many pixels that were misclassified as is clear from the Accuracy Assessment (table 2, 3 and 4).

Classification accuracy at different spatial scales (administrative boundary level and pixel level) revealed that GMLC on MODIS bands 1 to 7 is superior to other techniques for mapping agriculture as well as built up followed by SAM on MNF for mapping agriculture. Also, a certain algorithm may be good for mapping a "particular class" but at the same time may not be equally good for mapping "all other classes" – GMLC on MODIS band 1 to 7 could map agriculture, built up, forest, plantation and waste land well. However, water bodies could be well mapped with GMLC on PC's compared to other techniques.

Pre-processing techniques such as PCA and MNF on MODIS data had different effects on the accuracy of the hard classification algorithms. SAM and GMLC performed badly on the PC's and the MNF components when assessing accuracy with ground truth. Though, the class separability was very

good, yet it was difficult to classify each pixel based on signature, since the image was slightly pixelated. However, at the pixel level, the two techniques performed moderately better on PC's but relatively poor on MNF components maintaining the same level of accuracy as obtained by accuracy assessment using ground truth in the range of 42% to 49%. The pixels in the MNF components were not very distinct and were clustered into sub groups comprising of two or three pixels, leading to inaccurate classification results. The interesting point here is that both the techniques performed well at various spatial scales on MODIS band 1 to 7 products. This reveals that highly preprocessed MOD 09 data (level 3) takes care of all the atmospheric disturbances whereas the 36 band data, MOD 02 at level 1B requires further preprocessing to actually represent a good estimate of the surface spectral reflectance as it would have been measured at ground level without atmospheric scattering or absorption.

The accuracy assessment showed that GMLC on MODIS bands 1 to 7 has highest overall accuracy followed by SAM on the same 7 bands (table 4). However, some errors may have occurred since the signal of the pixel is ambiguous, perhaps as a result of spectral unmixing, or when the signal is produced by a cover type that has not been accounted for in the training process. Also, with coarser resolution, due to mixed pixels, chances of high accuracies in hard classifications decline.

5.0 CONCLUSIONS

This paper evaluates SAM and GMLC algorithms for classifying two different products of the MODIS data (MOD 09 surface reflectance and MOD 02 Level-1B Calibrated Geolocation Data Set). Accuracy assessment was done using ground truth obtained from field, at a administrative boundary level and was also compared with the high resolution LISS-III classified map on a pixel by pixel basis. The results obtained from accuracy assessment showed that MLC on MODIS 7 bands was good for land cover mapping at regional scale followed by SAM on the same 7 bands.

ACKNOWLEDGEMENT

We thank the Ministry of Science and Technology, Government of India and Indian Institute of Science for the financial assistance and infrastructure support.

REFERENCES

BATESON, C. A., ASNER, G. P., AND WESSMAN, C. A., (2000), Endmember Bundles: A New Approach to Incorporating Endmember Variability into Spectral Mixture Analysis. IEEE Transactions on Geoscience and Remote Sensing, 38(2): 1083-1094.

BOARDMAN, J., (1995), Analysis, understanding and visualization of hyperspectral data as a convex set in n-space. International SPIE symposium on Imaging Spectrometry, Orlando, Florida, pp. 23-36.

CROSTA, A.P., SABINE, C. AND TARANIK, J.V., (1998). Hydrothermal Alteration Mapping at Bodie, California, using AVIRIS Hyperspectral Data, Remote Sensing of Environment, 65:309-319.

DE CARVALHO, O.A. AND MENESES, P.R., (2000). Spectral Correlation Mapper (SCM); An Improvement on the Spectral Angle Mapper (SAM): Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication, 18, pp. 9.

GIROUARD, G., AND BANNARI, A., HARTI, A. E., AND DES ROCHERS, A., (2004). Validated Spectral Angle Mapper Algorithm for Geological Mapping: Comparative Study between Quickbird and Landsat-TM. www.isprs.org/istanbul2004/comm4/papers/432 p df.

GUENTHER, B., GODDEN, G. D., (1998). Pre-launch algorithm and data format for level 1 calibration products for the EOS-AM1 MODIS, IEEE Transaction on Geoscience and Remote Sensing, 36(4):1142–1151

HUNTER, E.L. AND POWER, C H., (2002) An Assessment of Two Classification Methods for Mapping Thames Estuary Intertidal Habitats Using CASI Data, International Journal of Remote Sensing, 23(15):2989-3008.

KESHAVA, N. AND MUSTARD, J.F., (2002). Spectral Unmixing, IEEE Signal Processing, 19(1):44.57.

KRUSE, F.A., BOARDMAN, J.W., LEFKOFF, A.B., HEIDEBRECHT, K.B., SHAPIRO, A.T., BARLOON, P.J., AND GOETZ, A.F.H., (1993). The Spectral Image Processing System (SIPS) – Interactive Visualization and Analysis of Imaging Spectrometer Data, Remote Sensing of Environment, 44:145-163.

LILLESAND, T.M. AND KIEFER, R. W., (2002). Remote Sensing and Image Interpretation, Fourth Edition, John Wiley and Sons, ISBN 9971-51-427-3.

LUMME, J. H., (2004), Classification of Vegetation and Soil using Imaging Spectrometer data. Proceedings of the XXth ISPRS Congress, 12-23 July 2004 Istanbul, Turkey, URL: http://www.isprs.org/istanbul2004/index.html.

PLAZA, A., MARTINEZ, P., PEREZ, R., AND PLAZA, J., (2004), A Quantitative and Comparative Analysis of Endmember Extraction Algorithms From Hyperspectral Data. IEEE Transactions on Geoscience and Remote Sensing, 42(3): 650-663

RAMACHANDRA, T.V., AND KUMAR, UTTAM., (2004). Geographic Resources Decision Support System for Landuse/Landcover dynamics analysis: GRASS Free and Open Source Software conference proceedings, Bangkok, Thailand http://gisws.media.osaka-cu.ac.jp/grass04/viewabstract.php?id=5

RAMACHANDRA, T.V., KUMAR, UTTAM., PRASAD, S. N., AND VAISHNAV B., (SEPTEMBER-OCTOBER, 2004). Geographic Resources Decision Support System – an open source GIS, Geospatial Today: 52-59.

RICHARDS, J.A., AND JIA, X., (1996). Remote Sensing Digital Image Analysis - An Introduction, Third Edition, Springer.

SALVATORI, R., GRIGNETTI, A., CASACCHIA, R., AND MANDRONE, S., (2003), The Role of Spatial Resolution in Vegetation Studies by Hyperspectral Airborne Images. Acts - Workshop on Airborne Remote Sensing for geophysical and environmental applications, Rome from 14 to 16 April 2003. URL: www.epa.gov/esd/land-sci/srsv/images/salvatori-casacchia.pdf

SCHWARZ, J. AND STAENZ, K., (2001). Adaptive Threshold for Spectral Matching of Hyperspectral Data, Canadian Journal of Remote Sensing, 27(3):216-224.

SHAFRI, H. Z. M., SUHAILI, A., AND MANSOR, S., (2007), The Performance of Maximum Likelihood, Spectral Angle Mapper, Neural Network and Decision Tree Classifiers in Hyperspectral Image Analysis. Journal of Computer Science, 3(6): 419-423.

VAN DER MEER, F., VASQUEZ-TORRES, M., AND VAN DIJK, P.M., (1997). Spectral Characterization of Ophiolite Lithologies in the Troodos Ophiolite Complex of Cyprus and its Potential in Prospecting for Massive Sulphide Deposits, International Journal of Remote Sensing, 18(6):1245-1257.

WINTER, M. E., AND WINTER, E. M., (2000), Comparison of Approaches for Determining End-members in Hyperspectral Data. IEEE Transaction on Geoscience and Remote Sensing, 3: 305-313.

YUHAS, R.H., GOETZ, A.F.H., AND BOARDMAN, J.W., (1992). Discrimination Among Semi-Arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm. Summaries of the 4th JPL Airborne Earth Science Workshop, JPL Publication, 92(41):147-149.