

Constrained Linear Spectral Unmixing Technique for Regional Land Cover Mapping Using MODIS Data

Uttam Kumar ^{a,b}, Norman Kerle ^c, and Ramachandra T V ^{a,b,*}, *Senior Member, IEEE*

^a Energy Research Group, Centre for Ecological Sciences, Indian Institute of Science, Bangalore-560012.

^b Centre for Sustainable Technologies, Indian Institute of Science, Bangalore-560012, India.

^c International Institute for Geo-Information Science and Earth Observation (ITC), P.O. Box 6, 7500 AA Enschede, The Netherlands. E-mail: kerle@itc.nl

* Corresponding author: cestvr@ces.iisc.ernet.in

Abstract—Over the last few decades, there has been a significant land cover (LC) change across the globe due to the increasing demand of the burgeoning population and urban sprawl. In order to take account of the change, there is a need for accurate and up-to-date LC maps. Mapping and monitoring of LC in India is being carried out at national level using multi-temporal IRS AWiFS data. Multispectral data such as IKONOS, Landsat-TM/ETM+, IRS-1C/D LISS-III/IV, AWiFS and SPOT-5, etc. have adequate spatial resolution (~ 1m to 56m) for LC mapping to generate 1:50,000 maps. However, for developing countries and those with large geographical extent, seasonal LC mapping is prohibitive with data from commercial sensors of limited spatial coverage. Superspectral data from the MODIS sensor are freely available, have better temporal (8 day composites) and spectral information. MODIS pixels typically contain a mixture of various LC types (due to coarse spatial resolution of 250, 500 and 1000 m), especially in more fragmented landscapes. In this context, linear spectral unmixing would be useful for mapping patchy land covers, such as those that characterise much of the Indian subcontinent. This work evaluates the existing unmixing technique for LC mapping using MODIS data, using end-members that are extracted through Pixel Purity Index (PPI), Scatter plot and N-dimensional visualisation. The abundance maps were generated for agriculture, built up, forest, plantations, waste land/others and water bodies. The assessment of the results using ground truth and a LISS-III classified map shows 86% overall accuracy, suggesting the potential for broad-scale applicability of the technique with superspectral data for natural resource planning and inventory applications.

Index Terms—Remote sensing, digital image processing, superspectral, Geographic Information System.

I. INTRODUCTION

Land cover (LC) relates to the discernible Earth surface expressions, such as vegetation, geology, and hydrologic or anthropogenic features, and thus describes the Earth's physical state in terms of the natural environment and the man-made structures. Essentially, LC can have only one class or category at a given time and location, and can be mapped using suitable image data with spectral signatures. Land use is an expression of human uses of the landscape, e.g. for

residential, commercial, or agricultural purposes, and has no spectral basis for its unique identification. Thus it can not be explicitly derived from image data, but only inferred by visual interpretation or assessed in the framework of object-based contextual analysis. It involves both the manner in which the biophysical attributes of the land are altered and the intent underlying that alteration, and the purpose for which the land is used. LC changes induced by human and natural processes play a major role in global as well as at regional scale patterns of the climate and biogeochemistry of the Earth system. Variations in topography, vegetation cover, and other physical characteristics of the land surface influence surface-atmosphere fluxes of sensible heat, latent heat, and momentum of heated air particulates caused by conduction, convection and radiation, which in turn influence weather and climate [1].

Many remote areas of the world are now being opened to exploration and development, generating a growing demand for up-to-date knowledge of topography, LC and other geo-spatial information. By using as many data sources as possible, a more complete and accurate knowledge of a landscape can be obtained. Therefore, users are seeking to integrate a multitude of spatially referenced information into their management and decision-making systems, a step that is facilitated by the standardisation of digital formats and the rapidly expanding market of GIS (Geographic Information System). There is a growing need for a global standardised LC and land use mapping system, similar to the CORINE approach. Many classification systems and innumerable map legends exist, but in most cases even from the same country, they are incompatible with each other. This system would enable a variety of end-users to use the results for their specific application (e.g. from rural planning to energy planning) and would enable intercomparison of existing data and a harmonised approach of data collection in areas where this information is not available or obsolete, whilst minimising the data and processing cost. Therefore, LC mapping remains an important research field, one what has grown more sophisticated with more recent technical developments in object oriented analysis or ontology [2], [3], [4].

LC features such as vegetation, water, and soil are important components in regional planning and management for monitoring the dynamics associated with the Earth. LC changes induced by human and natural processes play a major role in the climate, hydrology and biogeochemistry at global as well as regional scales. LC features can be classified using remotely sensed satellite imagery of different spatial, spectral and temporal resolutions. Recently, LC mapping of the entire territory of China was done at a 1:1 million scale to understand the LC change, based on existing LC maps, field surveys, NOAA's (National Oceanic and Atmospheric Administration) AVHRR (Advanced Very High Resolution Radiometer) imagery and aerial photos. The CORINE LC map of the European Union includes 44 LC classes divided into 5 main categories (agricultural areas, artificial surfaces, forests & semi natural areas, wetlands and water bodies). It is based primarily on Landsat TM (Thematic Mapper bands 4, 5 and 7) data of different vegetation periods with additional information in the form of topographic maps and orthogonal photos [5], [6].

In India, spatial accounting and monitoring of LC have been carried out at a national level at 1:250,000 scale, using multi-temporal IRS AWiFS (Indian Remote Sensing Satellite Advanced Wide Field Sensor) with 4 bands (Green, Red, NIR and SWIR) at 56 m resolution to address the spatial and temporal variability in cropping patterns and other LC classes. A decision tree classifier method was adopted to account for the variability of temporal datasets [7].

The above attempts are based on monotemporal remote sensing (RS) data with the analysis being done on an annual basis. Monitoring LC dynamics with time series satellite data would not be economical for regional or national level mapping with commercial data such as IRS LISS (Linear Imaging Self Scanner)-III, LISS-IV, SPOT or Landsat TM/ETM+ (Enhanced Thematic Mapper plus). This imposes a major limitation on the use of such data despite their high spatial resolution. RS data such as ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) are inexpensive and have a better spatial resolution, but are not regularly available for whole regions. MODIS (Moderate Resolution Imaging Spectroradiometer) data, with a spatial resolution of 250m to 1 km, have better spectro-temporal resolution (7 bands, and composite-data with Level 3 processing and 36 bands every 8 days, or every 1-2 day availability with Level 1B processing) can be downloaded freely and may be suitable for regional mapping and planning activities in many developing countries. Their frequent availability is especially useful to account for seasonal variations and changes in LC pattern.

In order to obtain these LC types, remotely sensed data are classified by identifying the pixels according to user-specified categories, by allocating a pixel to the spectrally maximally "similar" class, which is expected to be the class of maximum occupancy within the pixel [8]. MODIS based LC mapping

addresses large area coverage but is limited to classification of whole pixels [9], [10]. A variety of other classification methods exist, including spectral matched filter [11], mixture tune matched filtering and spectral angle mapper [12], which are appropriate when pixels do not contain mixtures of materials with correlated spectra, especially in higher spatial resolution data. However, MODIS pixels generally have the problem of spectral mixing. The mixed pixel problem is normally found at boundaries between two or more mapping units, or along gradients, etc. when the occurrence of any linear or small subpixel object takes place, and is usually dealt with by subpixel classifiers that assume either linear or non-linear mixtures [8] and [13]-[16]. It is often the case in RS that one wants to deal with identification, detection and quantification of fractions of the target materials for each pixel for diverse coverages in a region using unmixing approaches to discern the proportion of heterogeneity. A suitable way of extracting information is to estimate the composition of each pixel (proportion of the category contents) by spectral unmixing, i.e. soft classification techniques. The concepts of spectral unmixing emerged in the early 1970's [17], [18] and gained more prominence in 80's and 90's [8], [13]-[16] and [19]. During the last two decades, methods have been proposed ranging from modelling the component mixtures to solving the linear combinations to obtain abundances. Later, a number of techniques were developed to estimate and extract endmembers from the scenes or use spectral libraries.

In the present study, constrained linear spectral unmixing (CLSU) technique is applied on MODIS data to assess the suitability of the method for regional LC mapping with six LC classes. Linear Mixture Model (LMM), also known as a macro spectral mixture model, assumes no interaction between materials, and a pixel is treated as a linear combination of signatures resident in the pixel with relative concentrations [14]. This approach is different from the Spectral Angle Mapper (SAM) technique, where for the identification of pixel signature spectra only the angular information is used, which is based on the idea that an observed reflectance spectrum can be considered as a vector in a multidimensional space, and where the number of dimensions equals the number of spectral bands.

II. OBJECTIVE

Objectives of this study are: (i) LC mapping at regional level using superspectral MODIS data and (ii) to evaluate the suitability of CLSU to identify endmembers and generate abundance maps with MODIS (250 m spatial resolution).

III. ENDMEMBER EXTRACTION

Before modelling the linear mixture for unmixing, endmembers for the given study area have to be extracted. Various techniques, such as the Pixel Purity Index (PPI),

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 [16], [20]-[23]. In general, each algorithm finds appropriate spectra for endmembers. Orasis is capable of rapidly determining endmembers and unmixing large scenes. The N-FINDR method finds the set of pixels that define the simplex with the maximum volume potentially inscribed within the dataset. In the IEA algorithm, a series of constrained unmixing operation is performed, each time selecting as endmembers the pixels that minimize the remaining error in the unmixed image. CCA is based on the fact that some physical quantities, such as radiance and reflectance, are nonnegative. AMEE uses a morphological approach where spatial and spectral information are equally employed to derive endmembers. SAA is a method for constructing a simplex from a partition of the facets of the convex hull of a data cloud. The endmember extraction techniques used in this study are:

A. Pixel Purity Index (PPI)

This involves a dimensionality reduction using the Minimum Noise Fraction (MNF) transformation and the calculation of the PPI for each point in the image cube. This is accomplished by randomly generating lines in the N -dimensional space comprising a scatter plot of the MNF transformed data. All points in the space are then projected onto a line. After many repeated projections to different lines, those pixels above a certain threshold are declared "pure". There can be many redundant spectra in the pure pixel list. The actual endmember spectra are selected by a combination of review of the spectra themselves and through N -dimensional visualization. This provides an intuitive means to understand the spectral characteristics of materials [20].

B. Scatter Plot

A scatter plot with the set of scene spectra shows the endmember spectra occurring at the extremities (at the corners of the plot) [20]. In two dimensions, pure endmembers fall at the two ends of the mixing line, while in the case of a three endmember mix, mixed pixels fall inside a triangle, four endmembers fall inside a tetrahedron, etc.

C. N -Dimensional Visualization

Pixels from the spectral bands are loaded into an n -dimensional scatter plot and rotated on the visualization tool until points or extremities on the scatter plot are exposed. These projections are marked using a region of interest (ROI) tool and are repeatedly rotated in lesser dimensions to determine if their signatures are unique. Mean spectra are then extracted for each ROI to act as endmembers for spectral unmixing. These endmembers are then used for subsequent classification and other processing.

IV. LINEAR SPECTRAL UNMIXING

Models designed to estimate class proportions (rather than a single class label) for individual pixels, while addressing the problem of mixed pixels and considering the spectral response from a mixture of classes, are referred to as mixture models [24], [25]. Such spectral mixture (SM) modelling was used on resampled 25 m Landsat-7 ETM+ data for a subpixel classification that achieved 87% accuracy for DN (Digital Number) values and 93% for radiance values [26]. The technique is useful for discerning information from data with low spatial resolution, and would thus be ideal for free MODIS data with large ground coverage [27]. The spectral radiance measured by the MODIS sensor consists of the radiances reflected by all materials present, thus the radiance can be summed in proportion to the sub-pixel area covered by each material, given that the endmembers are the reference spectra of each of the individual pure materials, and under the condition that these spectra are linearly independent. The sites of pure LC for each class (or component) of interest are identified, and their spectra are used to define endmember signatures as discussed in section III. The position of the spectral signature of an input pixel along this continuum indicates directly the percentage cover for each component [14]. Constrained Least-Squares method (CLSM) aids in computing $n-1$ variables with n simultaneous equations [19]. In the case of MODIS, seven bands were designed for LC mapping [28] and hence the maximum number of LC categories that can be obtained is only six.

Linear unmixing (and its variants - Multiple Endmember Spectral Mixture Analysis) has been used earlier for LC mapping using MODIS surface reflectance data of 250 m and 500 m spatial resolutions of Northern Africa [29] and the unmixed results were compared with high resolution classified maps that gave an overall classification accuracy of 54% with significant confusion between alluvial surfaces and regs (surface covering of coarse gravel/pebbles or boulders from which all sand and dust have been removed by wind and water; a stony desert), and between sandy and clayey surfaces and dunes. A second validation using 20 Landsat images in a stratified sampling scheme gave a classification accuracy of 70%, with confusion between dunes and sand sheets.

LC fractions derived from MODIS 1 km resolution data and MISR (Multi-angle Imaging SpectroRadiometer) using SM have been compared to results of a Bayesian-regularized artificial neural network (ANN), as well as with 30 m reflectance data, yielding a quantitative improvement over spectral unmixing of single-angle, multispectral data (Landsat-7/ETM+) [30].

V. MATERIALS AND METHODS

A. Data

MODIS data with atmospheric corrections, also known as the “MOD 09 Surface Reflectance 8-day L3 global” product at 250 m (band 1 and 2) and 500 m (bands 3 to 7) in Hierarchical Data Format (HDF), were downloaded from the Earth Observing System Data Gateway [31]. They are radiometrically corrected, fully calibrated and geolocated radiances at-aperture for all MODIS spectral bands and are processed to Level 3G. The spectral range is from 0.45 to 2.15 μm [28]. IRS - 1C/D LISS-III MSS (Multi Spectral Scanner) data in 3 bands (G, R and NIR, 0.52 to 0.86 μm) with a spatial resolution of 23.5 m were purchased from NRSA (National Remote Sensing Agency), Hyderabad.

B. Study area

The Kolar district in Karnataka State, India, located in the southern plain regions (semi arid agro-climatic zone) and extending over an area of 8238 km^2 between 77°21' to 78°35' E and 12°46' to 13°58' N, was chosen for this study (Fig. 1). Kolar is divided into 11 taluks for administration purposes. Rainfall occurs mainly during southwest and northeast monsoon seasons. The average population density of the district is about 209 persons/ km^2 [32]. The study area is mainly dominated by agricultural land, built up (urban/rural), evergreen/semi-evergreen forest, plantations/orchards, waste lands and water bodies. There are a few other LC classes (barren/rock/stone/others) that have very limited ground area proportions and are unevenly scattered among the major six classes, and were grouped under the waste land category.

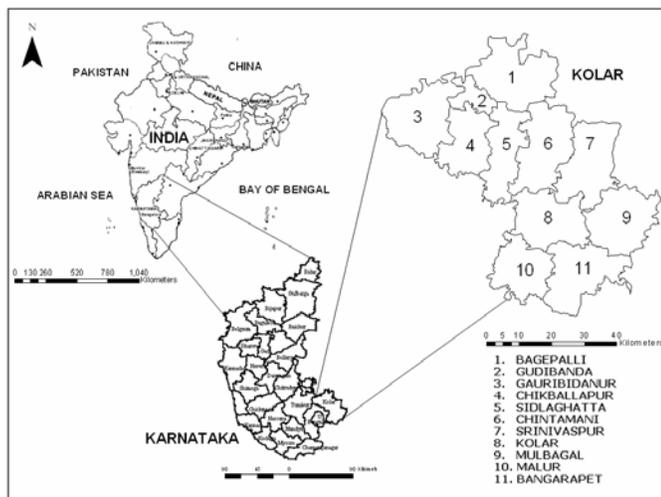


Fig. 1. Study area – Kolar district, Karnataka State, India.

C. Methods

The study involved creation of base layers, such as district, taluk and village boundaries, road network, etc. from the Survey of India (SOI) topographic maps of scale 1:250000 and 1:50000. The LISS-III bands were geo-registered using ground control points (GCPs). LISS-III bands were resampled to 25 m, which helped in pixel level comparison of abundance maps obtained by unmixing MODIS data with LISS-III classified map. This was followed by cropping and mosaicing of data corresponding to the study area from the image scenes. Supervised classification was performed on LISS-III MSS data using a Gaussian Maximum Likelihood Classifier (GMLC) followed by accuracy assessment. The MODIS data were geo-corrected with an error of 7 meters with respect to LISS-III images. The 500 m resolution bands 3 to 7 were resampled to 250 m using nearest neighbourhood technique (with Polyconic projection and Evrst 1956 as the datum). Minimum Noise Fraction (MNF) components were derived from the 7 bands to reduce noise and computational requirements for subsequent processing. Endmembers were extracted directly from the data without using existing spectral libraries through: (a) Pixel Purity Index (PPI) with the MNF components, (b) Scatter Plot, (c) N-Dimensional Visualization and (d) Collected training data - training polygons ($\geq 250 \times 250$ m) of homogenous patches corresponding to MODIS pure pixels were collected in the study area, thus enabling direct selection of assumed pure pixels from the images. The spectral characteristics of the endmembers were analysed by plotting them and analysing their separability using a Transformed Divergence matrix. It shows that the endmembers selected for the analysis are separable and can be distinguished from each other. The abundance maps were generated via constrained linear unmixing of MODIS data. Accuracy assessment was done for the abundance maps: LC percentages were compared at boundary level and at pixel level with a LISS-III classified map and also with ground truth data which is discussed later (Accuracy assessment).

VI. RESULTS

A. LC analysis using LISS-III MSS data

The class spectral characteristics for six LC categories using LISS-III MSS bands 2, 3 and 4 were obtained from the training pixels spectra to assess their inter-class separability and the images were classified with training data uniformly distributed over the study area collected with pre calibrated GPS (Fig. 2). This was validated with the representative field data (training sets collected covering the entire district and a detailed validation in Chikballapur Taluk covering ~ 15% of the study area) and the LC statistics are given in Table I. Producer's, user's, and overall accuracy computed are listed in Table II. A

TABLE I
LAND COVER DETAILS USING LISS -III MSS

Class	Area (%)
agriculture	19.03
built up land (urban/rural)	17.13
evergreen /semi-evergreen forest	11.41
plantations/orchards	10.96
wasteland/others: barren/rocky/stony	40.39
water bodies	01.08
Total (%)	100.00

TABLE II
ACCURACY OF LC CLASSIFICATION USING LISS-III MSS FOR
CHIKBALLAPUR TALUK

Category	Producer's Accuracy (%)	User's Accuracy (%)	Overall accuracy
agriculture	94.21	84.54	95.63%
built up	96.47	83.11	
forest	94.73	96.20	
plantation	92.27	91.73	
waste land/others	97.49	89.88	
water bodies	96.13	98.33	

κ (k) statistics of 0.95 was obtained indicating that the classified outputs are in good agreement with the ground conditions to the extent of 95%.

Classification errors can occur when the signal of a pixel is ambiguous, perhaps as a result of spectral mixing, or due to overlap of spectral reflectance (as in the case of certain agriculture and horticulture crops) or when the signal is produced by a cover type that is not accounted for in the training process. Another possible source of error may be due to the temporal difference in training data collection and image acquisition.

B. Spectral unmixing of MODIS data

Here each pixel spectrum of an image is modeled as a linear combination of a finite set of known components (or endmembers) given by (1)

$$r_i = \sum_{j=1}^n (a_{ij}x_j) + e_i \quad (1)$$

where, r_i = Spectral reflectance of a pixel in i^{th} spectral band containing one or more components; a_{ij} = Spectral reflectance of the j^{th} component in the pixel for i^{th} spectral band; x_j = Proportion value of the j^{th} component in the pixel; e_i = Error term for the i^{th} spectral band; $j = 1, 2, 3, \dots, n$ (Number of components assumed; in this study $n=6$); $i = 1, 2, 3, \dots, m$ (Number of spectral bands for the sensor system; in case of MODIS $m=7$). In principle, according to constrained assumptions, the proportions of abundance for each class range between 0 and 1, are non-negative and add to one as give by (2).

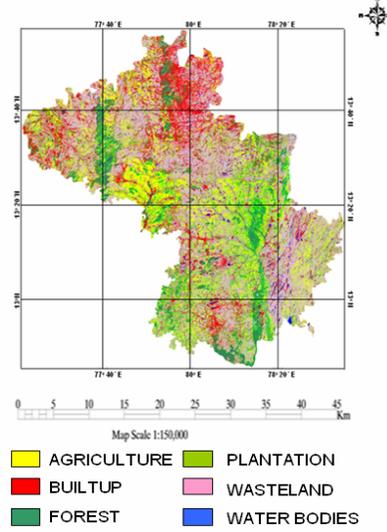


Fig. 2. Supervised classification of high resolution LISS-III image.

$$\sum_{j=1}^n (x_j) = 1 \quad (2)$$

Visual inspection as well as accuracy assessment of abundance images corresponding to each endmember showed that the CLSU algorithm maps LC categories spatially and proportionally similar to supervised classified image of LISS-III. The proportions of the endmembers in these images (Fig. 3) range from 0 to 1, with 0 indicating absence of the endmember and increasing value showing higher abundance.

Bright pixels represent higher abundance of 50% or more stretched from black to white. Errors were found due to confusion between agriculture and horticulture (plantation) in the central regions and built up and wasteland in northern regions of the district. Overall, the distribution and abundance values of other classes are comparable within $\pm 6\%$ to the classified outputs of LISS-III. There are many areas where proportions of agriculture are properly identified, mainly in the western central portion of the study area. However, there is also underestimation of agriculture in the south-central region due to errors of commission and omission.

C. Accuracy assessment

1) *Comparison with LISS-III MSS*: Table III provides the difference in LC percentage between the abundance maps generated taluk wise using CLSU and classified LISS-III data. Hence, further analysis was carried out at pixel level to understand the sources of these differences.

2) *Pixel level assessment*: A pixel of MODIS (250 m) corresponds to a kernel of 10 x 10 pixels of LISS-III spatially (Fig. 4). Endmembers generated along with abundance maps were validated pixel by pixel with the

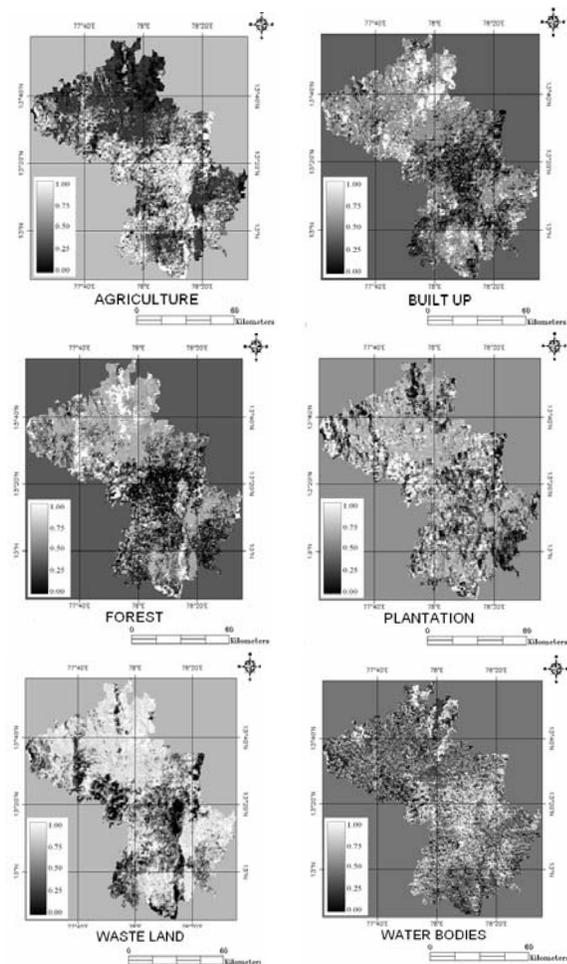


Fig. 3. Gray scale abundance maps for agriculture, built up, evergreen/semi-evergreen forest, plantations/orchards, wasteland and water bodies.

TABLE III
LC AREA DIFFERENCE (IN %) BETWEEN FRACTION IMAGES
OBTAINED FROM MODIS DATA AND LISS-III DATA

Taluk	agri-culture	built up	forest	plantation	waste land	water bodies
1	+0.4	-4.3	+0.2	+1.4	-1.5	+3.7
2	+0.2	0.0	-2.4	-0.3	+0.2	+2.1
3	-2.8	+0.4	-0.7	+0.5	+0.8	+1.5
4	-4.4	-4.2	+3.3	0.0	+4.2	+1.2
5	-0.9	-0.7	+0.2	+2.0	-3.6	+3.0
6	-3.5	+1.9	+3.7	+2.4	-5.9	+1.4
7	+3.4	-3.8	+4.1	-2.4	-4.1	+2.8
8	-0.4	-2.5	+0.8	+0.8	0.0	+1.1
9	-1.0	+1.4	+1.9	+0.6	-3.4	+0.5
10	-2.9	-0.7	+1.5	+1.5	-1.9	+2.3
11	+1.2	-0.4	-2.7	+2.9	+3.4	+3.2
District	-2.0	-1.3	-1.8	-1.7	+4.8	+2.0

1. Bagepalli, 2. Bangarpet, 3. Chikballapur, 4. Chintamani, 5. Gauribidanur, 6. Gudibanda, 7. Kolar, 8. Malur, 9. Mulbagal, 10. Sidlaghatta, 11. Srivasapur classified LISS-III image for Chikballapur taluk. Endmember for a particular category was considered as pure when its abundance value was greater than 90%.

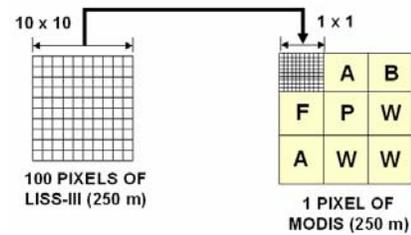


Fig. 4. 10 x 10 pixels of LISS-III equals to 1 x 1 pixel of MODIS spatially.

A total of 51 pixels with respect to the six LC classes were verified, and the result is tabulated in Table IV. The validation indicates that 35 pure and 14 mixed were correctly classified. Two pixels were misclassified; belong to water bodies which constitute < 1% of the total area.

3) *Assessment based on the field data of proportion LC:* The abundance maps were also verified through field investigations using calibrated (with known benchmarks and digital SOI topographic maps) GPS. In the case of mixed pixels, proportional LC were mapped and overlaid on the abundance map, and producer's and user's accuracy (Tables V and VI), overall accuracy and the kappa statistics were computed.

VII. DISCUSSION

Sustainable management of natural resources entails spatio-temporal information related to various surface features such as vegetation, water, and soil. In this regard, LC mapping using temporal RS data helps in capturing the dynamics associated with these LC components [33]. In principle, LC mapping can be done with any multispectral dataset, but currently available data have limitations such as poor spatial and temporal resolutions, or high spatial but low spectral resolution (IKONOS PAN, CARTOSAT-II) or moderate spatial and spectral resolutions (IRS LISS-III/IV, SPOT, Landsat TM/ETM+) as well as limited coverage, reduce the suitability of the available data types. Most of these RS data are commercial and many scenes would be required if mapping is to be carried out for larger regions. Even though some of the data (such as ASTER) are free, they are not available for all parts of the world and at variable temporal resolution. In this context, high temporal resolution MODIS data are freely available and are useful in providing insights into LC dynamics.

The present study evaluated the scope of CLSU technique for LC mapping using low spatial resolution MODIS data. Endmembers and abundance maps for the six LC categories were generated using MODIS data through linear unmixing. The proportion of the classes in each pixel of MODIS was computed and compared with the classified image of better spatial resolution and also through field data.

TABLE IV
VALIDATION OF LC CLASSES IN CHIKBALLAPUR

Class	Number of pixels identified	Pure pixels	Mixed pixels	Wrongly classified
agriculture	23	18	5	0
built up	6	4	2	0
forest	2	1	1	0
plantation	11	7	4	0
waste land	6	4	2	0
Water	3	1	0	2
Total	51	35	14	2

TABLE V
PRODUCER'S ACCURACY (IN PERCENTAGE)

Percentage→ Classes ↓	20 %	40 %	60 %	80 %	100 %
agriculture	50	83	100	100	100
built up	100	50	-	-	100
forest	100	100	100	-	100
plantation	63	88	100	100	100
waste land	33	100	33	100	100
water bodies	100	100	-	-	100

TABLE VI
USER'S ACCURACY (IN PERCENTAGE)

Percentage→ Classes ↓	20 %	40 %	60 %	80 %	100 %
agriculture	66	100	100	100	100
built up	50	100	-	-	100
forest	60	75	100	-	100
plantation	63	88	100	100	100
waste land	50	75	100	100	100
water bodies	100	100	-	-	100

Overall Accuracy = 86 %

Kappa (KHAT value) = 85 %

The results show that five of the six identified classes exhibit high interclass variability, allowing linear spectral unmixing. The variability of the pixels that represent the local pure classes is also responsible for the uncertainty of the mixture proportions. Probability density functions and parametric representation of the constituent pure classes did represent local conditions evident from the post classification validation. Assigning the categories to endmembers in consultation with the field data proved to be critical as it helped to discern the adjacency effect that exists between contrasting features such as forest and plantation, agriculture and plantation, etc. [34], [35]. This highlights that LMM, with improved estimates of proportional abundance values, has greater scope compared to other hard classification techniques in handling coarse spatial resolution data. The spectra of the different endmembers were modelled by linear equations. This introduces another form of quality assurance, as these equations can be inverted, taking the endmember spectra to reconstruct the spectra of each pixel from the original image. Divergences between the original and reconstructed image then give idea of the goodness of fit of the model, and provide insights to which bands add more to the

errors. When no more recognizable patterns are found and the error is overall small, it can be deduced that a near perfect model has been reached.

Nevertheless, even though the number of spectral bands in MODIS is higher compared to LISS-III, with the CLSM [36], only six categories can be discerned (number of bands available minus one). The condition of identifiability (the number of classes should be one less than the number of bands) can be solved by a two step process provided that many spectral endmembers are available. A subset with a prefixed number of endmembers that optimally decompose the candidate pixel is first selected by a Gram-Schmidt orthogonalisation process [37]. This restricted subset is then used in conventional LMM. The final result is the decomposition of the scene into all end-members considered while reducing the residual errors. Also, this approach fails in heavily fragmented landscapes or small isolated areas of high contrasting nature. This can be addressed through image fusion techniques [38] of low spatial resolution (MODIS) with high resolution (LISS-III/Landsat-TM/ETM+). To account for contributions from the neighbourhood pixels at the same time, non-linear mixture models are suitable for unmixing of coarse spatial resolution data.

VIII. CONCLUSION

Spectral linear mixture analysis provides an efficient mechanism for the classification of superspectral RS data (e.g. from MODIS). It aims to identify a set of reference signatures (endmembers) that can be used to model the reflectance spectrum at each pixel of the original image. These endmembers are extracted from the images using techniques such as PPI or scatter plots. Thus, the modelling is carried out as a linear combination of a finite number of ground components and their reflectance spectra. An abundance map helps in estimating the proportion of each endmember in a pixel. The performance of linear unmixing technique for identification of endmembers, interclass variability and presence of adjacency effect using 250-500 m MODIS data was evaluated. MODIS LC outputs are comparable to LISS-III classification results, evident from validation, as most endmembers were correctly classified while mixed pixels were within 10-15% of the values obtained in LISS-III classification. The MODIS data tested here have an advantage over alternative high or medium resolution datasets, given their coverage, high temporal resolution, cost-free availability, and utility for LC mapping as shown in this study. The challenges associated with this data type are (i) georegistration of the pixels when no or limited identifiable GCPs exist, (ii) misclassification of pixels due to LC mixtures, and (iii) mapping of LC in heavily fragmented areas with highly contrasting nature. One way of overcoming problems (ii) and (iii) is to use MODIS data with high spatial resolution images.

Even, in case MODIS ceases to function, the technique would still be applicable for datasets with low spatial resolution.

ACKNOWLEDGMENT

We are grateful to the Ministry of Science and Technology, Government of India for the financial assistance.

REFERENCES

- [1] J.R. Jensen, *Remote Sensing of Environment: An Earth resource perspective*, Pearson Education, 2003.
- [2] G. Camara, A.M.V. Monteiro, J. Paiva, and R. C. M. Souza, "Action-Driven Ontologies of the Geographical Space. In: M.J. Egenhofer and D.M. Mark (Editors)," *GIScience*, AAG, Savannah, GA, 2000.
- [3] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information," *Photogrammetry and Remote Sensing*, vol. 58, pp. 239-258, 2004.
- [4] H. Sun, S. Li, W. Li, Z. Ming, and S. Cai, "Semantic-Based Retrieval of Remote Sensing Images in a Grid Environment," *IEEE Geoscience and Remote Sensing Letters*, vol. 2 (4), pp. 440-444, 2005.
- [5] EEA & ETC/LC, Corine LC Technical Guide, 1999. Available: http://etc.satellus.se/the_data/Technical_Guide/index.htm
- [6] M. Torma, and P. Harma, "Accuracy of CORINE LC Classification in Northern Finland," *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04, Proceedings, 2004 IEEE International*, vol. 1, pp. 227-230, 20-24 Sept. 2004.
- [7] Natural Resources Census, "National Landuse and LC Mapping Using Multitemporal AWiFS Data (LULC – AWiFS)," Project Manual, Remote Sensing & GIS Applications Area, National Remote Sensing Agency, Department of Space, Government of India, Hyderabad, April 2005.
- [8] P. Bosdogianni, M. Petrou, and K. Josef, "Mixture Models with Higher Order Moments," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 35(2), pp. 341-353, 1997.
- [9] Y. J. Kaufman, and D. Tanre, "Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 30 (2), pp. 261-270, 1992.
- [10] A. Strahler, D. Muchoney, J. Borak, M. Fried, S. Gopal, E. Lambin, and A. Moody, "MODIS LC Product, Algorithm Theoretical Basis Document (ATBD) Version 5.0, MODIS LC and Land-Cover Change," 1999.
- [11] A. D. Stocker, I. S. Reed, X. Yu, "Multi-dimensional signal processing for electrooptical target detection," *Proc. SPIE International Society for Optical Eng*, vol. 1305, pp. 1-7, 1990.
- [12] R. H. Yuhua, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," *Summaries 3rd Annual JPL Airborne geoscience Workshop*, vol. 1, pp. 147-149, 1992.
- [13] J. T. Kent, and K. V. Mardia, "Spectral classification using fuzzy membership models," *IEEE Trans. Pattern Anal. Machine Intell*, vol. 10(4), pp. 659-671, 1986.
- [14] J. B. Adams, M. O. Smith, and P. E. Johnson, "Spectral mixture modelling: A new analysis of rock and soil types at the Viking Lander 1 site," *Journal of Geophysical Research*, vol. 91, pp. 8098 – 8112, 1986.
- [15] J. J. Settle, N. A. Drake, "Linear mixing and the estimation of ground cover proportion," *International Journal of Remote Sensing*, vol. 14 (6), pp. 1159-1177, 1993.
- [16] J. Boardman, "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, 1995.
- [17] H. M. Horwitz, R.F. Nalepka, P. D. Hyde, and J.P. Morgenstern, "Estimating the proportions of objects within a single pixel resolution elements of a multispectral scanner," *In Proc, 7th Int. Symp, Remote Sensing of Environment (Ann Arbor, MI)*, pp. 1307-1320, 1971.
- [18] D. M. Detchmendy, and W.H. Pace, "A model for spectral signature variability for mixtures," *Remote Sensing of Earth Resources*, vol 1 F, Shahrokhi, Ed. Tullahoma TN: Univ. Tennessee Press, pp.596-620, 1972.
- [19] Y. E. Shimabukuro, and A. J. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29(1), pp. 16-20, 1991.
- [20] M. E. Winter, and E. M. Winter, "Comparison of Approaches for Determining End-members in Hyperspectral Data," *IEEE Transaction on Geoscience and Remote Sensing*, vol 3, pp. 305-313, 2000.
- [21] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember Bundles: A New Approach to Incorporating Endmember Variability into Spectral Mixture Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38 (2), pp. 1083-1094, 2000.
- [22] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A Quantitative and Comparative Analysis of Endmember Extraction Algorithms From Hyperspectral Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42(3), pp. 650-663, 2004.
- [23] J. Settle, "On the effect of Variable Endmember Spectra in the Linear Mixture Model," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 44 (2), pp. 389-396, 2006.
- [24] L.R. Inverson, E. A. Cook, and R. L. Graham, "A technique for extrapolation and validating forest cover across large regions: calibrating AVHRR data with TM data," *International Journal of Remote Sensing*, vol. 10, pp. 1805-1812, 1989.
- [25] N. Quarmbay, J. Townshend, J. Settle, K. White, M. Milnes, T. Hindle, and N. Silleos, "Linear mixture modelling applies to AVHRR data for crop area estimation," *International Journal of Remote Sensing*, vol. 13(3), pp. 415-425, 1992.
- [26] C. Palaniswami, A. K. Upadhyay, and H. P. Maheswarappa, "Spectral mixture analysis for subpixel classification of coconut," *Current Science*, vol. 91(12), pp. 1706-1711, 2006.
- [27] A. M. Cross, J. J. Settle, N. A. Drake, and R. T. M. Paivinin, "Subpixel measurement of tropical forest cover using AVHRR data," *International Journal of Remote Sensing*, vol. 12(5), pp. 1119-1129, 1991.
- [28] <http://modis.gsfc.nasa.gov/data/dataproduct/>
- [29] J. A. C. Ballantine, G. S. Okin, D. E. Prentiss, and D. A. Roberts, "Mapping North African landforms using continental scale unmixing of MODIS imagery," *Remote Sensing of Environment*, vol. 97 (4), pp. 470-483, 2005.
- [30] B. H. Braswell, S. C. Hagen, S. E. Frothing, W. A. Salas, "A multivariable approach for mapping sub-pixel LC distributions using MISR and MODIS: Application in the Brazilian Amazon region," *Remote Sensing of Environment*, vol. 87 (2-3), pp. 243-256, 2003.
- [31] <http://edcimswww.cr.usgs.gov/pub/imswelcome/>
- [32] T. V. Ramachandra and G. R. Rao, "Inventorying, mapping and monitoring of bioresources using GIS and remote sensing," *Geospatial Technology for Developmental Planning*, Allied Publishers Pvt. Ltd. New Delhi., pp: 49-76, 2005.
- [33] T.V. Ramachandra and Uttam Kumar, "Geographic resources decision support system for land use, land cover dynamics analysis," *Proceedings of the FOSS/GRASS Users Conference - Bangkok, Thailand*, .12-14 Sept 2004.
- [34] http://gisws.media.osaka-cu.ac.jp/grass04/papers.php?first_letter=T
- [35] L. Zhu, and R. Tateishi, "Application of Linear Mixture Model to tile series AVHRR data," *Paper presented at the 22nd Asian Conference on Remote Sensing*, Singapore, 5 – 9 November, 2001.
- [36] M. M. Dunder, and D. Landgrebe, "A Model-Based Mixture-Supervised Classification Approach in Hyperspectral Data Analysis," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 40 (12), pp. 2692-2699, 2002.
- [37] R. Tateishi, Y. Shimazaki, and P. D. Gunin, "Spectral and temporal linear mixing model for vegetation classification," *International Journal of Remote Sensing*, vol. 25(20), pp. 4208-4218, 2004.
- [38] F. Maselli, "Multiclass spectral decomposition of remotely sensed scenes by selective pixel unmixing," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 36 (5), pp. 1809-1820, 2002.
- [39] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A Comparative Analysis of Image Fusion Methods," *IEEE Transactions of Geoscience and Remote Sensing*, vol. 43 (6), pp. 1391-1402, 2005.