

## Linear Mixture Model Applied to the Land-Cover Classification in an Alluvial Plain Using Landsat TM Data

M. A. Mohammed-Aslam<sup>1\*</sup>, Rokhmatuloh<sup>2</sup>, Z. E. Salem<sup>3</sup>, and Ts. Javzandulam<sup>2</sup>

<sup>1</sup>Department of Post-Graduate Studies and Research in Geology, Government College, Kasaragod, Vidyanagar, Kerala 671123, India

<sup>2</sup>Center for Environmental Remote Sensing (CEReS), Chiba University, 1-33 Yayoi, Inage, Chiba 263-8522, Japan

<sup>3</sup>Geology Department, Faculty of Science, Tanta University, Tanta, 31527, Egypt

**ABSTRACT.** The accurate delineation of the different pixel information is required for many remote sensing applications. However, the complexity of land cover makes the classification process difficult when using traditional methods, especially in areas where the heterogeneity is pronounced. In this paper, a Linear Mixture Model (LMM) approach is applied to classify the land cover in an alluvial tract using Thematic Mapper (TM) imagery around Talakad, parts of Mysore, Mandya and Chamarajanagar districts, Karnataka, India, in respect of five classes, viz.: sand, sparse vegetation, settlements, vegetation, and water. Fraction images of these classes were generated from Landsat TM image by un-mixing the image using LMM. This study indicates that the LMM approach is a promising method for distinguishing successional land cover in alluvial plain, where thick vegetation is noticed, using TM data. It gave better classification accuracy than traditional techniques did. The outputs of fraction images showed the high capability of LMM to extract many features. This was not possible with maximum likelihood classification method in spite of an overall accuracy of 98.83%, which was particularly not so efficient in extracting the vegetation and water bodies. The land cover units contained in the area of this alluvial tract were not picked up properly in the maximum likelihood classification.

*Keywords:* Alluvial plain, land cover classification, Landsat TM, Linear Mixture Model

### 1. Introduction

One of the challenges face by remote sensing scientists is the lack of a suitable methodology for extracting information from the satellite data to estimate the proportionate land cover occupied in a single pixel. In the traditional classification techniques, the signature from the combined effect of the spectra of all pixels within a training set for a given feature is obtained. Therefore, the resulting signature contains the contributions of all materials present in the training set (Huguenin et al., 1997). The extraction of information from a single pixel is known as the "Mixture Modeling". Several researches were carried out previously (Cross et al., 1991; Quarmby et al., 1992; Dickinson, 1995; Oleson et al., 1995; Atkinson et al., 1997; DeFries et al., 1997; Sellers et al., 1997) in this technique for resolving the mixture problem, to distinguish different classes from one another. Such techniques were widely applied to the coarse and moderate resolution data. The LMM evaluates each pixel spectrum as a linear combination of a set of endmember spectra (Adams et al., 1995). It has been used for land cover/land use studies (Ustin et al., 1996; Cochrane and Souza, 1998; Aguiar et al., 1999; Ustin et al., 1999; Theseira et al., 2002). The difficulty of accurately classifying the different classes is related to many factors. The heterogeneity of the land cover is a major problem, due to the limitation of

the spatial resolution, i.e., a single Landsat TM pixel often contains more than one land-cover class. For example, even within the same vegetation class there is often great variation in plant species, plant geometry, and vegetation structure (Lu et al., 2003). Therefore, the traditional classifiers could not effectively handle the mixed pixel problem (Lu et al., 2003). In order to unmix the pixels into different proportions of the endmembers from the mixed pixel, some methods have been developed (Ichoku and Karnieli, 1996; Bastin, 1997; Erol, 2000). The field of remote sensing has evolved substantially, and is currently far more diverse in its goals, methods, and conclusions.

The objective of this work is to make use of the effectiveness of the LMM in classifying the land classes in and around an alluvial landform, since such a study improves understanding of a variety of units. This work aims at studying and implementing a methodology to practise the land cover classification in and around alluvial landforms. The LMM method has been adopted to improve understanding of geographic distribution and percentage of land cover units to reduce the uncertainty of regional spatial heterogeneity.

The area wherein the LMM technique has been applied is characteristic of the alluvial landform and much of this zone is intensely irrigated. The variability in heterogeneity of the alluvial land, vegetation cover and water are important in evaluating the vegetation dynamics, and the area available for water irrigation etc. However, the complexity created while

\* Corresponding author: aslam090@yahoo.co.in

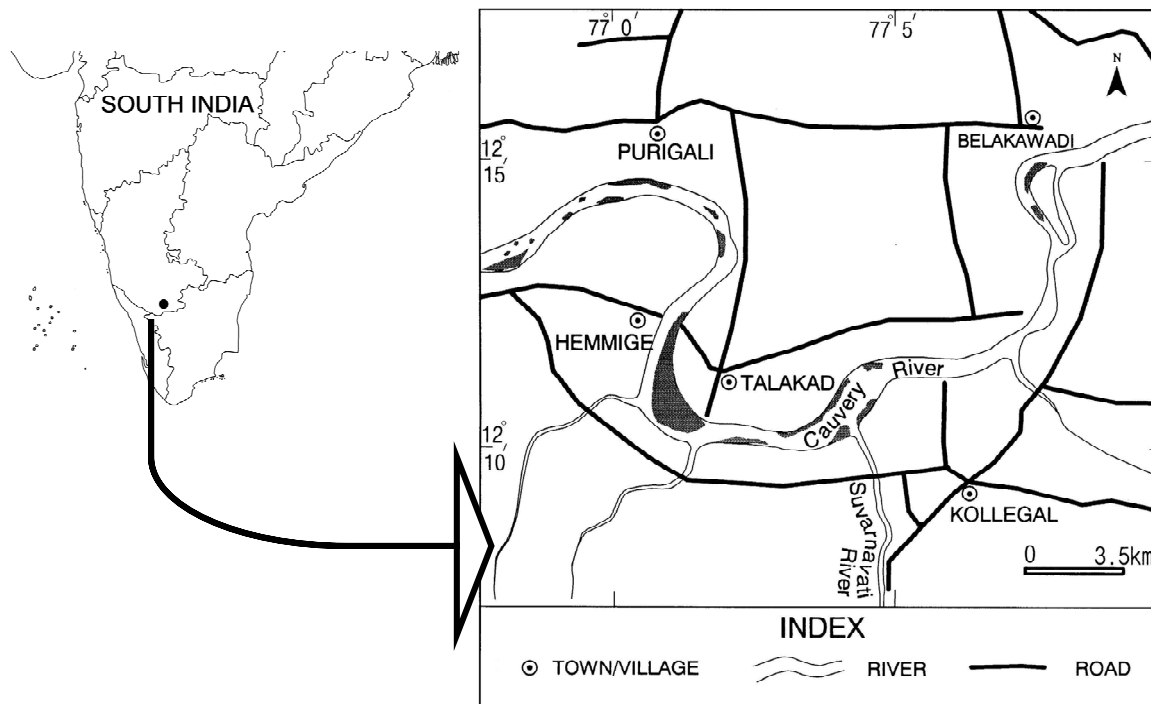


Figure 1. Map showing the study area.

making a smooth transition from one class to another renders the classification difficult when using traditional approaches such as the maximum likelihood classifier (MLC). Invariably, the traditional classification makes only broad distinctions between different classes. It has been difficult to accurately distinguish stages of fraction of the landforms. In this paper, a linear mixture model (LMM) approach is employed to classify the alluvial land cover using Landsat Thematic Mapper (TM) imagery.

## 2. The Study Area

This article focuses on the application of LMM to classify land cover patterns within an intensely irrigated alluvial plain, around Talakad, situated within parts of Mandya, Mysore and Chamarajanagar districts, Karnataka state, located within the upper Cauvery river basin, India. Such an area is characterized by the presence of palaeochannels with good hydrological potential (Mohammed-Aslam and Balasubramanian, 2001). It is lying between North latitudes  $12^{\circ}7'30''$  and  $12^{\circ}17'30''$  and East longitudes of  $76^{\circ}57'30''$  and  $77^{\circ}10'$  (Figure 1). It forms one of the typical alluvial landforms of the South India. The area of study comprises of paddy-fields and sugarcane cultivation, as well as grass-lands. They require large scale irrigation for cultivation.

Area under study is a gently undulating table land with a few isolated hills. The eastern part of the area is characterized by two major hills forming the part of Biligirirangana Betta

and Mahadeshwara Betta with elevations ranging from 830 m to 914 m AMSL (Above Mean Sea Level). A large number of tanks of different dimensions are located throughout the study area and are used for storing rainwater for irrigation purposes. The perennial flow of the Cauvery River and its drainage network helped the land to bloom as the fertile agricultural field right from the pre-historic time. The alluvial plains are typical and major geomorphic units of this area. They are important from the groundwater exploration point of view. Groundwater occurs at shallow depths in this region. Anomalously wide lateral extent of the alluvial plain is clearly indicative of the existence of palaeochannels. The Peninsular gneiss is the major rock-type forming the aquifer in plains and is the major basement rock (Ramakrishnan et al., 1976). Inclusions within the gneisses are mainly mafic and ultra-mafic and can be of any dimensions ranging from fragments to large schist belts. The general trend of schist belts in this region is NNE-SSW direction. A few patches of granulitic rocks are observed in the eastern margin of this area.

The climatic conditions of this region are quite congenial for those working in the fields as well as those who are engaged in common agriculture. The year may be divided into four seasons. The summer season is from March to the end of May, which is followed by the southwest monsoon season lasting to the end of September. October and November may be termed as the post monsoon or retreating monsoon season. The period from December to February is the dry season with generally clear bright weather.

### 3. Methodology

The basic hypothesis behind the linear mixture model is that the image spectra are the result of mixtures of surface materials, shade and clouds. The mixtures in the image are described as linear combinations of their respective spectra. Linear Mixture Model (LMM) is appropriate for a large number of the land scenes and is based on a simple mathematical description for unmixing the given mixture (Tsolmon, 2003). The linear mixture modeling technique basically applies the multiple linear regression technique. This provides pure end-member reflectance for each band as well as fraction images for each land cover class. The spectral response of each image pixel in every spectral band can be considered as a linear combination of the response of each fraction (or end-member) contained in the mixture. So, every pixel exhibits information about the respective proportion and spectral response of the different components present.

The method suggested by Tsolmon (2003) is followed for LMM based feature extraction for the alluvial tract. The basic concept and the mathematical model of LMM are expressed as Equation (1). In the linear spectral model, each pixel will be considered to be a linear combination of the “pure” or end-member spectra of the objects occupied in the pixel area, weighted by their respective fractional abundance. Many techniques are used to measure the adequacy of a multiple regression model. These are based on the analysis of variance, as described in Montgomery and Runger (1994). Equation (2) estimates the proportion of each component in a pixel by minimizing the sum of squares of errors. For solving the LMM, Conditional Gradient Method (CGM), an optimization technique was followed (Tsolmon, 2003), which was developed for assumed (*n*) components in a pixel:

$$\begin{aligned}
 R_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + e_1 \\
 R_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + e_2 \\
 &\dots\dots\dots \\
 R_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + e_m
 \end{aligned}
 \tag{1}$$

$$E(x) = \sum_{i=1}^m e_i^2 = \sum_{i=1}^m (R_i - \sum_{j=1}^n a_{ij}x_j)^2 \rightarrow \min
 \tag{2}$$

Subject to:

$$\begin{aligned}
 x_1 + x_2 + x_3 + \dots + x_n &= 1, \\
 x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, \dots, x_n &\geq 0,
 \end{aligned}$$

where  $R_i$  is radiometric response for a pixel in spectral band  $i$ , unit: watts/(meter squared  $\times$  ster  $\times$   $\mu\text{m}$ );  $a_{ij}$  is spectral response of mixture component  $j$  in spectral band  $i$ ;  $x_j$  is the proportion of  $j$  in a pixel;  $e_j$  is error term in spectral band  $i$ ;  $m$  is number of spectral bands;  $n$  is number of components (classes) within a pixel.

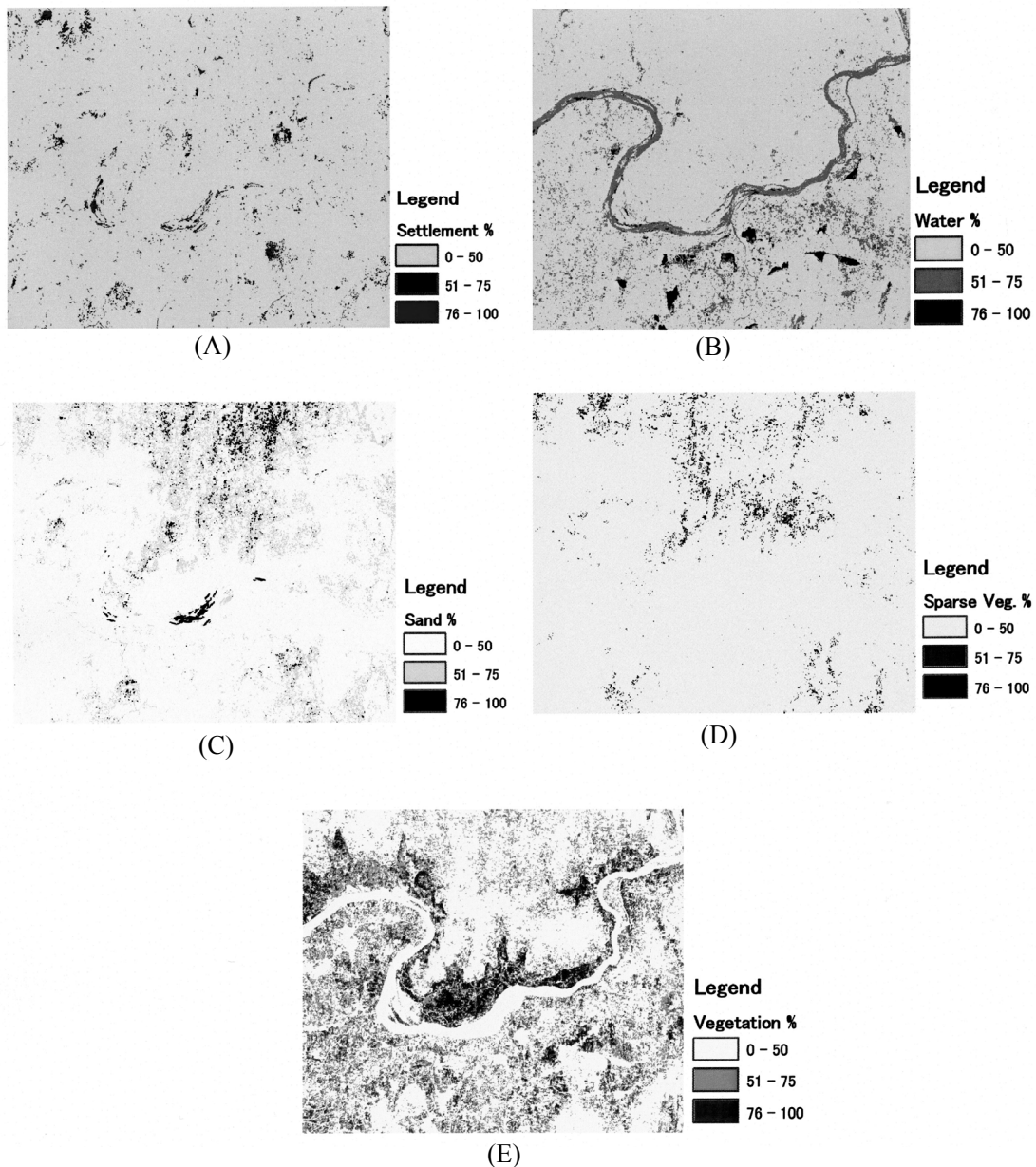
The pixel-encoded radiance is affected by the radiometric errors at the moment of image acquisition. These errors are

mainly resulted by atmospheric conditions, sensor gains, solar azimuth, earth-sun distance, viewing angle, scene illumination and satellite observation angles, the effect of the Bidirectional Reflectance Distribution Function (BRDF) of the sensed surface, and sensor band spectral response functions. Distortions to the image can be caused by position, size and orientation of a pixel to be altered during the acquisition process. The reasons of such errors include Earth rotation during the acquisition stage, platform speed and altitude variations, land curvature, change in topographical elevation, sensor's parameters, etc. These combine to produce significant band dependent radiometric differences, confounding the interpretation of both temporal and spatial data sets. When using Landsat satellite imagery to map land cover features, it is desirable to remove these effects. Such an improvement is required for any study, including LMM on a per-pixel basis. In the case of linear mixture model the electromagnetic energy interacts with a single component before being reflected by the surface.

A cloud free portion of Landsat TM scene with path 144 and row 52, acquired on 9<sup>th</sup> November, 1999 was selected for the LMM based classification of land cover around Talakad. The accuracy and quality of fraction images greatly depend on the proper selection of sufficient number of training classes. The pixel values which shows the best representation of each class in mixtures were selected and used as the input for LMM, to extract the fraction images from Landsat TM data for the main land cover types in the study area. The “best represented” pixels in this context refer to the sites that were characterized by 80 to 100% of the respective land covers. Knowledge on the proportion of corresponding surface spatial extension is clearly of primary importance in land cover studies. Note that some land cover units in reality can be a mixture of a few other surfaces. Over natural areas, to delineate the surface parameters correctly, is much more complex due to the high variability and diversity of the vegetation types.

The LMM applied in this study was formulated as convex programming problem coded using visual C++ language and solved numerically using Conditional Gradient Method (Tsolmon, 2003) by deriving components  $x_1, x_2, x_3, x_4$  and  $x_5$  corresponding to sand, sparse vegetation, settlements, vegetation, and water respectively within a pixel. The solving of the linear mixture model yielded the fraction images. From the resulting fraction images of each class, the main land cover types were further isolated from the others by classifying them as “classed” and “non-classed”. Subsequently, each pixel was associated with percentage values from 0 to 100 for sand, sparse vegetation, settlements, vegetation, and water from the respective derived images of LMM. These land cover classes are deemed sufficient to cover almost all the variability of the study area. The resultant fraction images were further reclassified as 0 - 50%, 51 - 75% and 76 - 100% groups to improve their quality for a quick evaluation.

The total number of end-members is limited by the spectral dimensions of the image data set used in mixture modeling. Pure end-member reflectance has been calculated by applying the linear mixture modeling.



**Figure 2.** Fraction images for components (A) Settlement, (B) Water, (C) Sand, (D) Sparse vegetation, and (E) Vegetation generated by applying LMM to TM image acquired on November 9, 1999.

#### 4. Results and Discussion

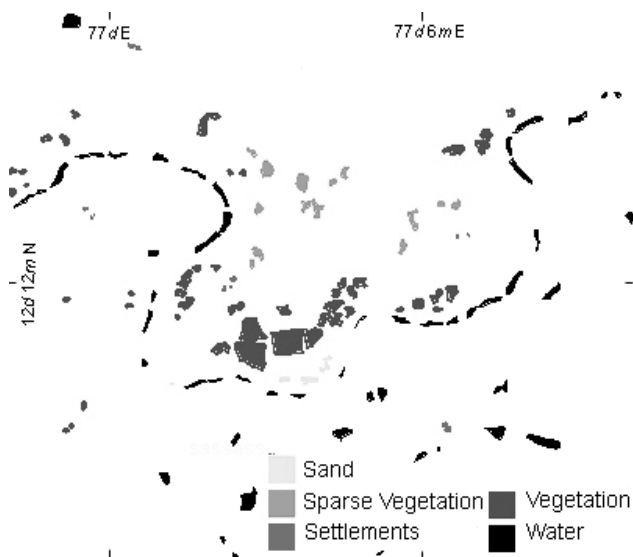
Success of the spectral mixture approach rests on the selection and identification of end-member reflectance bundles that encompass the actual spectral variability of the dominant land cover types (Adams et al., 1995). The response of each pixel in any spectral wavelength can be considered as a linear combination of the responses of each component, which is assumed to be in the mixture. Land cover classification is one of the primary objectives of the use of satellite data. Opti-

mal conditions for discriminating land cover types are based on the information on land cover characteristics and image quality.

During the linear mixture modeling, equations have been solved for the unknowns (the fractions) for each pixel. As a result, the model produces slightly smoother images while compared with true images. Due to mathematical constraints the number of fractions that can be estimated is limited to the number of bands input into the model.

By applying LMM for feature identification, the sandy area, settlements, water bodies and vegetation fractions of the satellite image (Figure 2) were effectively extracted. It is a classification technique which derives sub-pixel information instead of the maximum likelihood classification, where every pixel is assigned to only one class. Therefore, a classification should not assign a pixel to a single class, but proportionally to several classes in areas of higher order of heterogeneity. The signatures of the individual land cover classes lie at extreme ends of a continuum in spectral feature space. The position of the spectral signature of an input pixel along this continuum indicates directly the percentage proportion for each component in the linear mixture model. It has been difficult to accurately distinguish the different types of land cover units in areas that are within one pixel of satellite images. This linear mixture model is widely used in unmixing of pixels, that the signatures present in a pixel are linearly super positioned. The mixture unmixing is based on the fact that the spectral signature of satellite data essentially results from a combination of a small number of pure components with characteristic spectra.

The derived fraction images for the main land cover types were compared with ground truth data. The outputs of fraction images showed the high capability of LMM to extract many features, which were not possible with the traditional classification approach. Figure 3 shows the ground truth information used for generating the maximum likelihood classification.



**Figure 3.** The map showing the details of ground-truth information used for land cover classification.

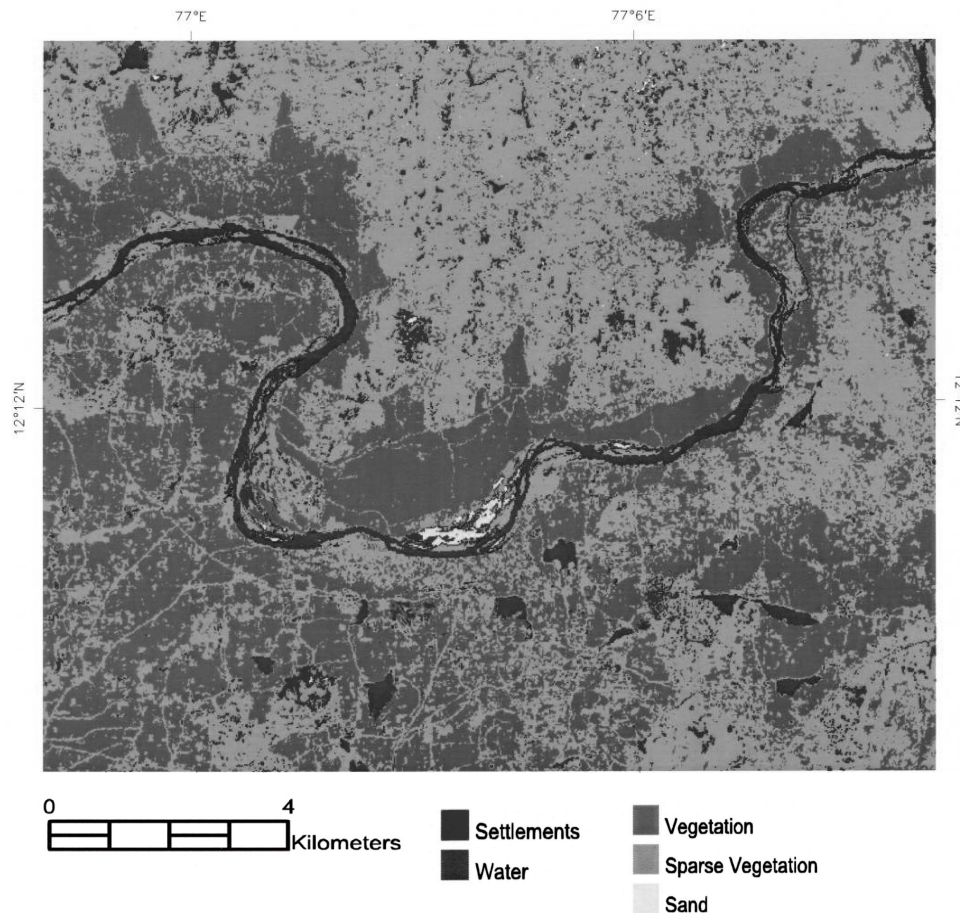
Since the use of moderate resolution imagery increases the proportion of mixed pixels in relation to the alluvial area, the estimates with a maximum likelihood classification become less accurate. The maximum likelihood classifier uses statistical descriptions of the class samples or training sets to

map image data into groups of land units so that each pixel is assigned to its most likely class. Even though such a type of classification is practical in many situations, two drawbacks were identified: spatial information in the data is ignored and the mixed pixels are treated as “pure” elements. The classes obtained from the maximum likelihood classification were not quite efficient in extracting the vegetation and water bodies. Even though an overall accuracy of 98.83% and Kappa coefficient of 0.98 were obtained, the land cover units comprised were not picked up appropriately in the maximum likelihood classification (Figure 4). For example, it is evident from the proportionate percentage of vegetation that the vegetation fraction of 51 to 75% and 76 to 100% constitute the intensely irrigated land, which is located in alluvial land. The vegetation fraction upto 50%, are mostly constituted by other classes (in this case, settlement, sand or bare land) with less than 50% of vegetation in these pixels. Similarly, the wet condition of land in the southern part of the study area, wherein the water spreading is widely used for cultivating paddy and sugarcane, was identified by picking up the pixels that were sensitive to water in the 51 to 75% group, which was not picked up by the maximum likelihood classification. Therefore, study of these classes is very important, as they constitute very sensitive parameters for the hydrological cycle of alluvial landforms. The relatively poor performance of traditional classification with respect to the LMM method is pronounced because of the combined reflectance which was resulted due to a variety of components present within the single pixel, even though the overall accuracy of the maximum likelihood classification is within the acceptable limit. Since the occurrence and pattern of land cover within alluvial landforms is very complicated, it is very difficult to distinguish them very clearly under normal classification methods.

The use of the satellite data with poor spatial resolution inevitably entails some misclassification, either due to the failure to detect small sub-pixel objects, or due to the mixed spectral response of edge pixels lying along the contact between two land cover types. This is due to the fact that the value of each pixel is the composite spectral signature of the land cover types present. The interpretation of this satellite data can be improved greatly by an understanding of the spectral components within each pixel. For the current application, the linear mixture modeling technique has been proved to be very effective and useful, mainly because it produced pure end-member reflectance next to the sub-pixel proportions.

## 5. Conclusions

LMM approach is a suitable method for distinguishing successional alluvial zones in the alluvial land cover using TM data. LMM technique included both modeling approach and empirical estimations of land cover components. This method is highly suitable for regions where the single pixel is comprised of multiple components. The supervised classification has shown 98.83% accuracy. The results from fraction images are comparable with that of supervised classification. Over natural areas, the problem of classifying the surface parame-



**Figure 4.** Land cover classification performed by Maximum Likelihood Classifier for TM image acquired on November 9, 1999.

ters is much more complicated due to the high variability and diversity of the vegetations and their thickness.

In the alluvial plain adjacent to the river, mixed pixels in the Landsat TM images are common because of the abundant vegetation species and the limitations of spatial resolution of the remotely sensed data in a heterogeneous condition. This research indicates that the linear mixture model approach is a good method in distinguishing the mixed land cover, especially in alluvial plains.

LMM of the remotely sensed data is a vital component for extracting the sub-pixel heterogeneity in characterization of land cover. The LMM as sub-pixel classification technique could be utilized to maximize the accuracy of any land cover especially in the intensive agricultural areas, like the alluvial landforms.

**Acknowledgments.** We are thankful to the anonymous reviewers for their constructive comments.

## References

- Adams, J.B., Sabol, D.E., Kapos, V., Filho, R.A., Roberts, D.A., Smith, M.O. and Gillespie, A.R. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sens. Environ.*, 52, 137-154.
- Aguiar, A.P.D., Shimabukuro, Y.E. and Mascarenhas, N.D.A. (1999). Use of synthetic bands derived from mixture models in the multispectral classification of remote sensing images. *Int. J. Remote Sens.*, 20, 647-657.
- Atkinson, P.M., Cutler, M.E.J. and Lewis, H. (1997). Mapping sub-pixel proportional land cover with AVHRR imagery. *Int. J. Remote Sens.*, 18(4), 917-935.
- Bastin, L. (1997). Comparison of fuzzy c-means classification, linear mixture modeling and MLC probabilities as tools for unmixing coarse pixels. *Int. J. Remote Sens.*, 18, 3629-3648.
- Cochrane, M.A. and Souza-Jr., C.M. (1998). Linear mixture model classification of burned forests in the eastern Amazon. *Int. J. Remote Sens.*, 19, 3433-3440.
- Cross, A.M., Settle, J.J., Drake, N.A. and Paivinen, R.T.M. (1991). Subpixel measurement of tropical forest cover using AVHRR

- data. *Int. J. Remote Sens.*, 6, 1159-1177.
- Dickinson, R.E. (1995). Land processes in climate models. *Remote Sens. Environ.*, 51, 27-38.
- DeFries, R., Hansen, M., Steininger, M., Dubayah, R., Sohlberg, R. and Townshend, J. (1997). Subpixel forest cover in central Africa from multisensor, multitemporal data. *Remote Sens. Environ.*, 60, 228-246.
- Erol, H. (2000). A practical method for constructing the mixture model for a spectral class. *Int. J. Remote Sens.*, 21, 823-830.
- Huguenin, R.L., Karaska, M.A., Blaricom, D.V. and Jensen, J.R. (1997). Subpixel classification of Bald Cypress and Tupelo Gum trees in Thematic Mapper imagery. *Photogramm. Eng. Remote Sens.*, 63, 717-725.
- Ichoku, C. and Karnieli, A. (1996). A review of mixture modeling techniques for sub-pixel land cover estimation. *Remote Sens Rev.*, 13, 161-186.
- Lu, D., Moran, E. and Batistella, M. (2003). Linear mixture model applied to Amazonian vegetation classification. *Remote Sens. Environ.*, 87, 456-469.
- Mohammed Aslam, M.A. and Balasubramanian, A. (2001). Identification of Palaeochannels around Cauvery river near Talakad, Karnataka, India, using Remote sensing data, *J. Ind. Soc. Remote Sens.*, 29(4), 237-242.
- Montgomery, D.C. and Runger, G.C. (1994). *Applied Statistics and Probability for Engineers*, John Wiley & Sons, New York.
- Oleson, K.W., Sarlin, S., Garrison, J., Smith, S., Privette, J.L. and Emery, W.J. (1995). Unmixing multiple land-cover type reflectances from coarse spatial resolution satellite data. *Remote Sens. Environ.*, 54, 98-112.
- Quarmby, N.A., Townshend, J.R.G., Settle, J.J., White, K.H., Milnes, M., Hindle, T.L. and Silleos, N. (1992). Linear mixture modeling applied to AVHRR data for crop estimation. *Int. J. Remote Sens.*, 13, 415-425.
- Ramakrishnan, M., Viswanathan, M.N. and Swaminath, J. (1976). Basement cover relationships of Peninsular gneiss with high grade schists and green stone belts of southern Karnataka. *J. Geol. Soc. India*, 17, 97-111.
- Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Mooney, H.A., Nobre, C.A., Sato, N., Field, C.B. and Henderson-Sellers, A. (1997). Modeling the exchanges of energy, water and carbon between continents and atmosphere. *Sci.*, 275, 502-509.
- Theseira, M.A., Thomas, G. and Sannier, C.A.D. (2002). An evaluation of spectral mixture modeling applied to a semi-arid environment. *Int. J. Remote Sens.*, 23, 687-700.
- Tsolmon, R. (2003). *Methodology to Estimate Coverage and Biomass of Boreal Forests using Satellite Data*, unpublished Ph.D. Dissertation, Graduate School of Science and Technology, Chiba University, Japan, pp. 152.
- Ustin, S.L., Hart, Q.J., Duan, L. and Scheer, G. (1996). Vegetation mapping on hardwood rangelands in California. *Int. J. Remote Sens.*, 17, 3015-3036.
- Ustin, S.L., Smith, M.O., Jacquemoud, S., Verstraete, M. and Govaerts, Y. (1999). Geobotany: Vegetation mapping for earth sciences, in A.N. Rencz (Ed.), *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*, New York, volume 3, pp. 189-233.