Cloud Publications
International Journal of Advanced Remote Sensing and GIS
2013, Volume 2, Issue 1, pp. 130-137, Article ID Tech-96
ISSN 2320 - 0243

Research Article **Open Access**

Comparison of MLC and FCM Techniques with Satellite Imagery in A Part of Narmada River Basin of Madhya Pradesh, India

Arun Mondal, Deepak Khare and Sananda Kundu

Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee, India

Correspondence should be addressed to Arun Mondal, arun.iirs@gmail.com

Publication Date: 3 July 2013

Article Link: http://technical.cloud-journals.com/index.php/IJARSG/article/view/Tech-96

Abstract Landuse and land cover are most important part which is linked with the environment and climate in various ways. This is also important for the modeling of greenhouse gas emissions, carbon balance etc., and is important for understanding the landscape features. The main objective of the present work is to reduce uncertainties in the landuse and land cover pattern. Remote sensing technique is extremely important for the classification purposes by empirical observation and algorithms. In case of present study, a part of Narmada river basin was taken where change in the landuse and landcover was assessed from the Landsat images of the year 2011 with two classification techniques of Maximum Likelihood Classification (MLC) and Fuzzy C-Mean (FCM). The major landuse classes are water body, built-up, vegetation, agricultural land and fallow land. The image has been digitally classified by both MLC and FCM algorithms which have been validated by the accuracy assessment process. The overall accuracy achieved by the FCM was about 84% while with MLC it was about 79%. The Survey of India toposheet was used as the base map for the purpose of geo-correction. FCM was found to be more accurate in comparison to MLC because of its soft classification technique.
Keywords *Landuse, Narmada River Basin, MLC, FCM, Accuracy*

1. Introduction

The landuse change intensity with respect to the population growth of the world has led to such consequences for the environment that resulted in the detail study of these transformations *(Wu et al., 2006)*. Landuse is considered as one of the basic supposition for understanding the global environment changes. Landuse and land cover mapping is an extremely important component, where different parameters are combined together. Remote sensing technology has helped for global information monitoring for the landuse and land cover in the form of spatial, spectral and temporal resolution. This technology has reduced the survey time, availability of latest maps, digital image classification, etc. Since 1972, Landsat satellite data have been used for digital classification for the preparation of landuse classes *(Townshend, 1992; Hall et al., 1995; Lu and Weng, 2007)*. Change in the landuse and land cover was analysed by many researchers like *He et al. (2000) and Zhang et al.*

(2002) who have used multi-temporal satellite images for classifying landuse in the city area of Beijing. Maximum likelihood classification (MLC) is very popular and is used in wide extent throughout the world for extraction of landuse classes *(Huang et al., 2002)*. The MLC classification is a parametric approach where selected classes of signatures are assumed in normal distribution. Some non-parametric types of classification techniques are also there which are considered to have better accuracy for deriving classes *(Kavzoglu and Reis, 2008)*. Some of these well-known non-parametric techniques are Fuzzy C-Mean, Decision Trees, Support Vector Machines (SVM), Artificial Neural Networks (ANN) etc. In fuzzy, a point may be assigned to a particular cluster with its degree of belonging to that particular cluster that is measured from the centroid point of the cluster. So problem of any hard classification can be avoided. Number of studies has been done with the application of fuzzy C-Mean. *Lucieer (2006)* used fuzzy C-Means classifier for estimation of the vegetation units. According to *Foody (1992)*, *Lucieer (2006)* and *Lu and Weng (2007)*, fuzzy technique of classification is important because it can overlap hard classes with the soft technique.

Present study compares the landuse classification techniques of MLC and Fuzzy C-Means to assess the classification accuracy using Landsat satellite imagery. Spatial accuracy was analyzed with the outputs of both the classification techniques in a part of Narmada river basin of Madhya Pradesh of India.

2. Study Area

The study area includes Burhanpur, Barwani, Sehore, Khandwa, Indore, Harda, Dewas, East Nimar and West Nimar districts. It is a part of Narmada river basin of Madhya Pradesh. The area stretches from 21°23'51" to 22°54'55"N and 75°21'41" to 77°20'53"E and covers about 52970 km^2 areas. The region experiences subtropical climate with hot dry summer and a cool dry winter. The rainfall average is about 1370 mm and it decreases from east to west (Figure 1).

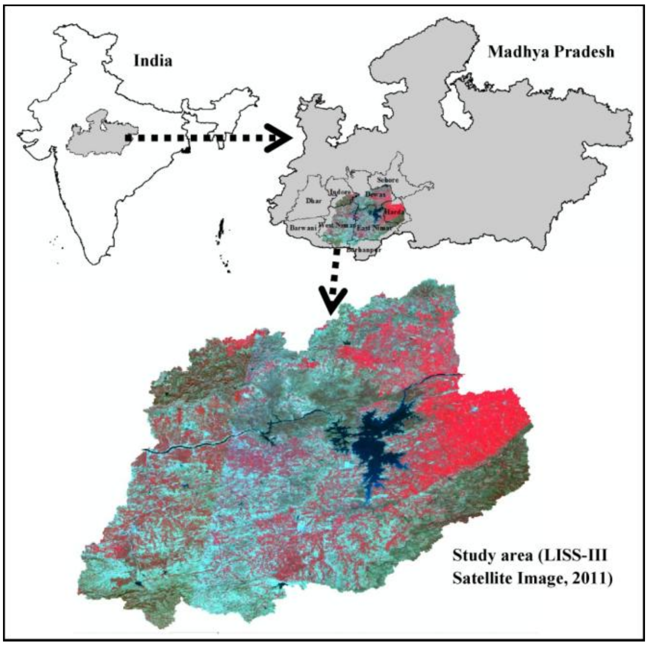

Figure 1: *Study Area*

3. Data and Methods

For the purpose of classification, Landsat TM satellite imagery was used for applying the methods of Maximum Likelihood Classification (MLC) and Fuzzy C-Mean algorithms. The Landsat TM data of 2011 was used which was corrected geometrically and radiometrically for removing the errors.

Landuse classification of images with the MLC and Fuzzy C-Means were done and accuracy assessment was applied to compare both the techniques.

3.1. Maximum Likelihood Classification (MLC) Technique

MLC is considered as a process to determine known class distribution as the maximum for a given statistic *(Scott & Symons, 1971)*. It is one of the most widely used processes in remote sensing. Here, a pixel having maximum likelihood is assigned to the corresponding class. If there are m numbers of predefined classes, then the class a posteriori probability is given as

$$P(k|x) = \frac{P(k)P(k|x)}{\sum_{i=1}^{m} P(i)P(k|i)} \quad (1)$$

where P (k) represents the prior probability of class k and P(x | k) is considered as the conditional probability for observing x from the class k (probability density function). For the normal distributions, the likelihood function, P(x | k), is given as following

$$L_k(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \left|\sum_k\right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_k)^T \sum_k^{-1} (x-\mu_k \right) \quad (2)$$

where x = (x1, x2 . . . xn)^T stands for the vector of a pixel with n number of bands; $L_k(x)$ is the likelihood membership function of x belonging to class k and μk = (μ_k1 μ_k2 . . . μ_kn)^T represents the mean of the kth class;

Kth class: $$\sum_k = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdot & \cdot & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdot & \cdot & \sigma_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_{n1} & \sigma_{n2} & \cdot & \cdot & \sigma_{nn} \end{pmatrix}$$ is the variance covariance matrix of the class k. (3)

3.2. Supervised Fuzzy C-Mean Method

Jain and Dubes (1988) give many types of existing clustering techniques where images have been classified by applying supervised classification technique with fuzzy logic. The fuzzy technique gives a natural model where a pixel may have partial membership value corresponding to many land cover categories. This algorithm operates on the basis of iterative minimization of the objective function.

$$J_m(U,v) = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^m \left\| y_k - v_i \right\|_A^2 \quad (4)$$

Where,

$Y = \{Y_1, Y_2 \ldots\ldots Y_N\} \subset R^n$ = the data,
c = number of clusters in Y; $2 \le c \le n$,
m = weighting exponent; $1 \le m < \infty$
U = fuzzy c-partition of Y; $U \in M_{fc}$
$v = (v_1, v_2 \ldots\ldots v_c)$ = vectors of centers,
$v_i = (v_{i1}, v_{i2}, .., v_{in})$ = center of cluster i,
$\| \|_A$ = induced A-norm on R", and
A = positive-definite (n × n) weight matrix.

The following constraints are fulfilled by the membership values,

$$0 \leq u_{ik} \leq 1; \text{Where } i \in \{1, \ldots . C\}; k \in \{1, \ldots \ldots N\} \tag{5}$$

$$\sum_{i=1}^{C} u_{ik} = 1; \qquad k \in \{1, \ldots, N\} \tag{6}$$

$$\sum_{k=1}^{N} u_{ik} > 0; \qquad i \in \{1, \ldots, C\} \tag{7}$$

The objective function is the sum of square of Euclidean distances from each input sample and its corresponding cluster centre, and these distances are weighted by the memberships of fuzzy. It is given in the following equations:

$$\hat{v}_i = \left[\sum_{k=1}^{N} u_{ik}^{m} y_k\right] / \sum_{k=1}^{N} u_{ik}^{m} \tag{8}$$

$$\hat{u}_{ik} = 1 / \sum_{j=1}^{C} \left[\|y_k - v_i\| / \|y_k - v_j\|\right]^{2/(m-1)} \tag{9}$$

The membership value of each class is dependent on its distance to the centre of the corresponding cluster. Smaller will be the influence of samples having small membership values with greater value of m *(Bezdek et al., 1984)*.

3.3. Accuracy Assessment

Number of methods for accuracy assessment can be found. One common technique is error matrix *(Congalton, 1991; Foody, 2002)*. In this study, overall, producer's and user's accuracy have been calculated. Producer's accuracy is found by dividing the accurate number of sampling points in one class by the total number of points taken from the reference data. For user's accuracy, correct classified points are divided by the total number of points which are already been classified in that particular class.

3.4. Kappa Coefficient

Kappa is a discrete multivariate technique that is used for the measurement of the accuracy of maps and error matrices derived by remote sensing. The Khat statistic is given as

$$\hat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} . x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} . x_{+i})} \tag{10}$$

here 'r' is the number of rows in the matrix X_{ij} showing number of observations in row 'i' and column 'j', and X_{i+} and X_{+i} are the marginal totals for the 'i' and I respectively, and N gives the total number of observation *(Lillesand and Kiefer., 2000)*.

Change in the land use and comparison of accuracy of two methods are represented in the given land use classes.

4. Results and Analysis

4.1. Accuracy in Classification

Total of 120 points were taken for accuracy assessment in case of both the classification results. Producer and user accuracy of water body, built-up, dense vegetation, grass land and fallow land are 80%, 77.78%, 76.67%, 83.02%, 73.91% and 66.67%, 77.78%, 79.31%, 81.48%, 77.27% of MLC technique respectively. In Fuzzy C-Mean tecnique, producer and user accuracy of water body, built-up, vegetation, agricultural land and fallow land are 83.33%, 91.67%, 84.62%, 85.42%, 78.57% and 83.33%, 84.62%, 81.48%, 85.42%, 84.62% respectively (Table 1).

Table 1: *Error Matrix*

1990					
Class Name	**Reference Totals**	**Classified Totals**	**Number of Correct Points**	**Producers Accuracy (%)**	**Users Accuracy (%)**
MLC (2011)					
Water body	5	6	4	80.00	66.67
Built-up	9	9	7	77.78	77.78
Dense Vegetation	30	29	23	76.67	79.31
Grass land	53	54	44	83.02	81.48
Fallow land	23	22	17	73.91	77.27
Total	120	120	95	………..	………..
Fuzzy C-Mean					
Water body	6	6	5	83.33	83.33
Built-up	12	13	11	91.67	84.62
Dense Vegetation	26	27	22	84.62	81.48
Grass land	48	48	41	85.42	85.42
Fallow land	28	26	22	78.57	84.62
Total	120	120	101	………..	………..

Kappa statistics of the following landuse in MLC are 0.65, 0.76, 0.72, 0.67, 0.72 and in Fuzzy C-Mean they are 0.82, 0.83, 0.76, 0.76 and 0.80 recpectively. The overall kappa statistics and overall accuracy for MLC are 0.70 and 79.17% and in Fuzzy C-Mean they are 0.78 and 84.17% respectively. Thus difference in overall accuracy is 5% between two methods (Table 2 and Table 3).

Table 2: *Kappa Statistics*

Class Name	MLC (2011)	Fuzzy C-Mean (2011)
Water body	0.65	0.82
Built-up	0.76	0.83
Vegetation	0.72	0.76
Agricultural land	0.67	0.76
Fallow land	0.72	0.80

Table 3: *Overall Kappa Statistics and Overall Accuracy*

Category	MLC (2011)	Fuzzy C-Mean (2011)
Over all kappa statistics	0.70	0.78
Over all accuracy (%)	79.17	84.17

4.2. Statistics of MLC and Fuzzy C-Mean

Water body area from MLC and Fuzzy C-Mean techniques are about 884 km² and 930 km² respectively. Built-up areas are differing in 3.32% in two methods while difference in area of vegetation is 1.64%. Agricultural land and fallow land is showing difference of 5.72% and 3.82% respectively (Table 4, Figure 2 and Figure 3).

Table 4: *Distribution of Area of Different Landuse*

	MLC		Fuzzy C-Mean	
Classes	**2011 (Area in km²)**	**2011 (%)**	**2011 (Area in km²)**	**2011 (%)**
Water	883.94	4.30	930.42	4.53
Built up	1471.68	7.16	2154.32	10.48
Vegetation	5004.31	24.34	4665.85	22.70
Agricultural land	9518.64	46.30	8341.78	40.58
Fallow Land	3679.42	17.90	4465.63	21.72
Total	20558	100	20558	100

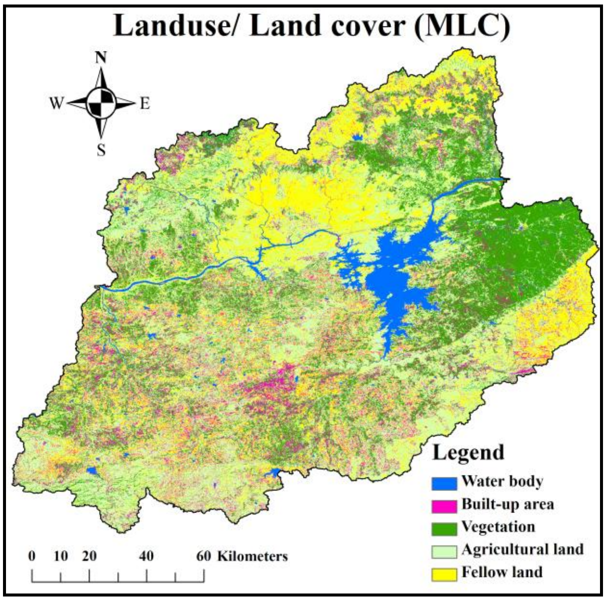

Figure 2: *Landuse/ Land Cover Maps from MLC*

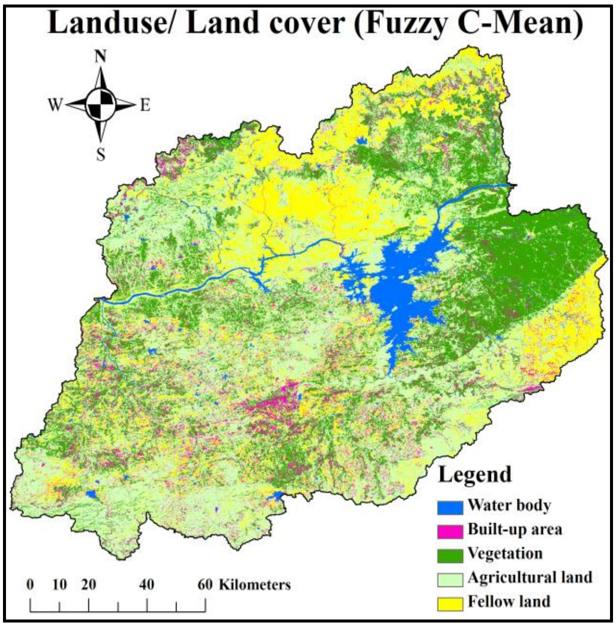

Figure 3: *Landuse/ Land Cover Maps from Fuzzy C-Mean*

Spatial similarity from the outputs of MLC and Fuzzy C-Mean was carried out for each of the landuse classes. 83% of the area has matched i.e. 83% of area is same for both the type of classifications; only 17% of the area has not matched or is different (Figure 4).

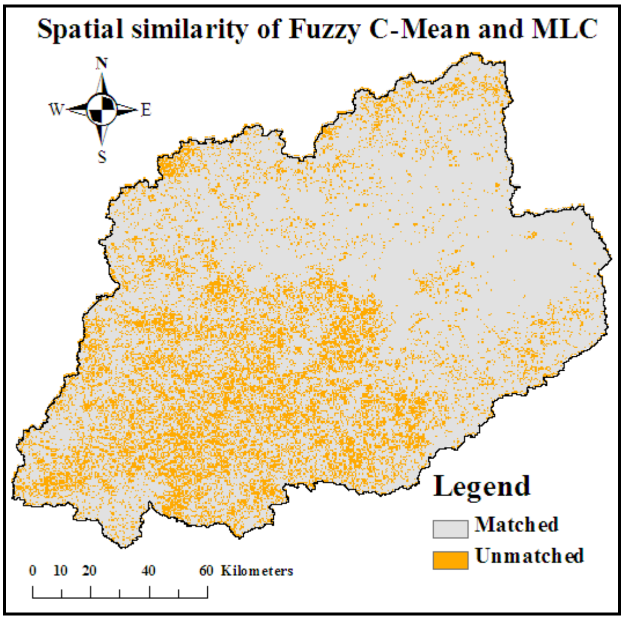

Figure 4: *Spatial Similarity of MLC and Fuzzy C-Mean*

5. Conclusion

Two different types of classification like MLC and Fuzzy C-mean techniques have been applied in a part of Narmada river basin in Madhya Pradesh to identify the more accurate method. Verification with ground observation was done with handheld GPS (Global Positioning System) after classifications to assess the level of accuracy. Level of accuracy has been found to be better in case of Fuzzy C-Mean algorithm than MLC, although the difference in area of the landuse classes is less. 83% of similarity in area was observed in both the methods with only 17% of difference. Nonetheless, Fuzzy C-Mean is considered to be superior to MLC landuse classification because of better accuracy achieved.

Acknowledgement

The authors thankfully acknowledge the United States Geographical Survey (USGS) and Survey of India (SOI) for providing the Landsat satellite imagery and topographical map for the purpose. Authors are also thankful to CSIR for financial support.

References

Bezdek et al. *The Fuzzy C-Means Clustering Algorithm.* Computers & Geosciences. 1984. 10 (2-3) 191-203.

Congalton. *A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data* Remote Sensing of Environment. 1991. 37; 35−46.

Foody G.M. *Status of Land Cover Classification Accuracy Assessment.* Remote Sensing of Environment. 2002. 80; 185−201.

Foody G.M. *A Fuzzy Sets Approach to the Representation of Vegetation Continua from Remotely Sensed Data -An Example From Lowland Heath.* Photogrammetric Engineering and Remote Sensing. 1992. 58; 221-225.

Hall F.G., et al. *Status of Remote Sensing Algorithms for Estimation of Land Surface State Parameters.* Remote Sensing of Environment. 1995. 51; 138-156.

He C.Y., et al. *A Study on Landuse/Cover Change in Beijing Area.* Geographical Research. 2000. 20 (6) 679–687.

Huang C., et al. *An Assessment of Support Vector Machines for Land Cover Classification.* International Journal of Remote Sensing. 2002. 23; 725-749.

Jain A., et al., 1988: *Algorithms for Clustering Data.* Prentice-Hall, Englewood Cliffs, NJ.

Kavzoglu T., et al. *Performance Analysis of Maximum Likelihood and Artificial Neural Network Classifiers for Training Sets with Mixed Pixels.* GIScience and Remote Sensing. 2008. 45; 330-342.

Lillesand T.M., et al., 2000: *Remote Sensing and Image Interpretation.* 4th Ed. John Wiley & Sons, Inc., New York.

Lu D., et al. *A Survey of Image Classification Methods and Techniques for Improving Classification Performance.* International Journal of Remote Sensing. 2007. 28; 823-870.

Lucieer A., 2006: *Fuzzy Classification of Sub-Antarctic Vegetation on Heard Island Based On High-Resolution Satellite Imagery*. IEEE Geoscience and Remote Sensing Symposium, 2777-2780.

Scott A.J., et al *Clustering Methods Based on Likelihood Ratio Criteria.* Biometrics. 1971. 27 (2) 387-397.

Townshend J.R.G. *Land Cover.* International Journal of Remote Sensing. 1992. 13; 1319-1328.

Wu Q., et al. *Monitoring and Predicting Land Use Change in Beijing Using Remote Sensing and GIS.* Landscape and Urban Planning. 2006. 78; 322-333.

Zhang Q., et al. *Urban Built-Up Land Change Detection with Road Density and Spectral Information From Multi-Temporal Landsat TM data.* International Journal of Remote Sensing. 2002. 23 (15) 3057-3078.