Accuracy Assessment of Land Cover /Land Use Mapping Using Medium Resolution Satellite Imagery

Paliwal M.C &. Katiyar S.K

Abstract— Classification of satellite imagery is very important for the assessment of its accuracy. In order to determine the accuracy of the classified image, usually the assumed-true data are derived from ground truth data using Global Positioning System. The data collected from satellite imagery and ground truth data is then compared to find out the accuracy of data and error matrices are prepared. Overall and individual accuracies are calculated using different methods. The study illustrates advanced classification and accuracy assessment of land use/land cover mapping using satellite imagery. IRS-1C-LISS IV data were used for classification of satellite imagery. The satellite image was classified using the software in fourteen classes namely water bodies, agricultural fields, forest land, urban settlement, barren land and unclassified area etc. Classification of satellite imagery and calculation of accuracy was done by using ERDAS-Imagine software to find out the best method. This study is based on the data collected for Bhopal city boundaries of Madhya Pradesh State of India [6].

Keywords—Resolution, Accuracy assessment, Land use mapping, Satellite imagery, Ground truth data, Error matrices.

I. INTRODUCTION

FOR classification of satellite imagery remote sensing output has to be assessed accurately. Without accuracy assessment, output/results has less significance. In remote sensing derived land use or land cover maps it is important to quantitatively evaluate classification accuracy. In order to determine classification accuracy, it is necessary to check whether the output map meets certain predetermined classification accuracy criteria.

A most common and typical method used to assess classification accuracy is with the use of an error matrix [1]. The rows in the matrix represent the remote sensing derived land use map (i.e., Landsat data), while the columns represent the reference data (i.e., ground truth, in-situ samples). These tables produce many statistical measures of thematic accuracy including overall classification accuracy, percentage of omission and commission error by category, and the KHAT coefficient.

To assess classification accuracy we need to compare two different maps first the classified map derived from remotely sensed data and second the existing sources of reference information. Usually, the "assumed-true" reference data are derived from ground truth data. These reference pixels are randomly selected from the available satellite imagery of land cover reference information. An important factor in determining the accuracy of a classification is the number of reference pixels used. Congalton states that more than 250 reference pixels are needed to estimate the mean accuracy of a class to within plus or minus five percent [1].

Remote sensing technology is used for collecting the requisite data to solve problems relate to land use planning. The advantages of remote sensing technology have attracted great interest in the scientific and engineering community [7]. Modeling environmental phenomena usually needs some spatial information about the distribution and the types of land cover and land use (LCLU) as well as soil types [2]. Ragan and Jackson investigated the use of computer analysis of Landsat Satellite Multispectral Scanner data for estimating the land cover distributions needed in operating the Soil Conservation Service (SCS) models [3].

This paper evaluates four remote sensing classification methods for automatically obtaining LCLU in Bhopal city (in Madhya Pradesh State of India) from IRS images. The study was made to find out various aspects of classification and

1.. Paliwal M.C ,Research Scholar, Department of Civil Engineering,; Maulana Azad National Institute of Technology, Bhopal ,India
e-mail – mcpaliwal@rediffmail.com
2. Katiyar S.K , Professor ,Department of Civil Engineering, Maulana Azad National Institute of Technology (Under Govt. of India, Ministry of Human Resource Development),Bhopal – 462 051, India,
e-mail – skkatyar7@rediffmail.com)

accuracy assessment of satellite imagery using advanced methods.

II. Objectives

Following were the objectives of the study:-

1) To classify satellite imagery using different methods;
2) To find out accuracy of different classes;
3) To verify the accuracy by ground truth using GPS;
4) To suggest best method to find out accuracy

III. Area Under Study

In this study Bhopal city area has been taken for research study. The geographical location of the area lies in UTM zone 43N. The extant of study area is between latitudes 220 11’-30.44” N and 230 27’ 39.96” N, and between longitudes 770 26’ 32.80’’ E and 770 27’ 59” E. It falls in the Malwa Plateau region. It has a good fertile soil cover, which encourages agricultural practices in the region. Agriculture has been a major occupation for the residents. Bhopal has an average elevation of 523 m (MSL) with an annual rainfall of 1146 mm and average temperature of 250C. Bhopal has slightly deep, well-drained, calcareous clayey soils on gently sloping plain with narrow valleys with moderate erosion and has an irrigated agricultural practices. The main agricultural crops are wheat with gram, pea (Pisum sativum), and other crops. Well-distributed sample plots (14 in the Bhopal city) were selected as ground truth and validating the results. Location of study area and LISS IV sensor image of Bhopal are shown in Fig. 1 and Fig. 2, respectively.

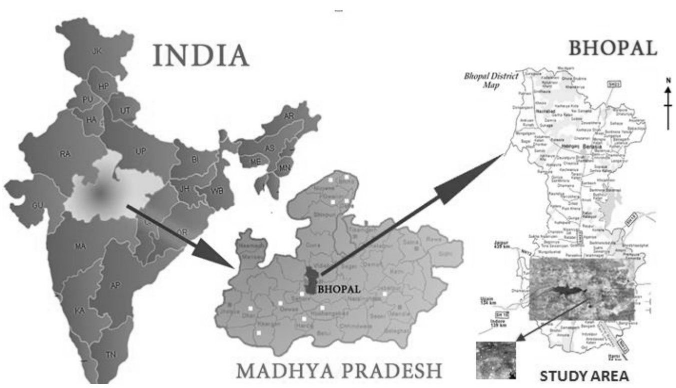

Fig. 1 Study area

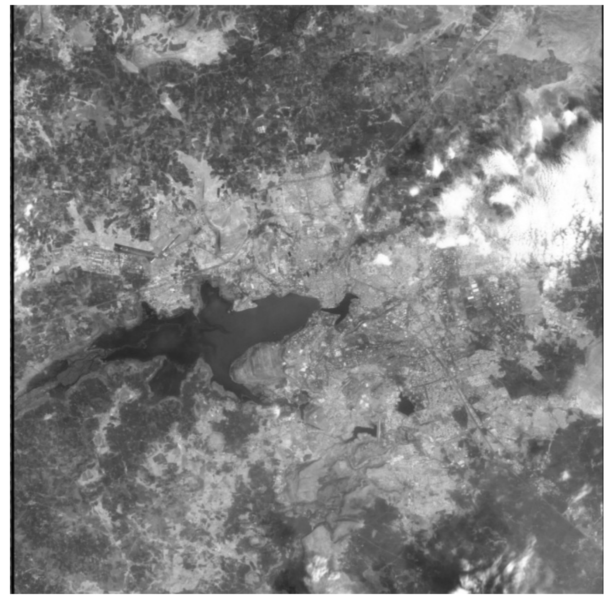

Fig. 2 LISS-IV Image of Bhopal and adjoining area

IV. Methodology

The most common types of classification techniques are unsupervised and supervised classification methods. Unsupervised classification doesn’t need a prior knowledge of the area and the supervised classification needs prior knowledge of the area [4]. The process of gaining this prior knowledge is known as ground-truthing. These ground-truths (or signatures) can be obtained from existing maps or by conducting field work in the study areas.

The classification system used in this study was the one developed by [5]. Image classification (or image information extraction) of land cover process in this study involves several steps which are given below:

1) Geo-referencing of satellite imagery
2) Geometric corrections applied to IRS-1C LISS-IV data
3) Selection of study area.
4) Ground truthing by using GPS to obtain latitudes and longitudes of reference points
5) Land use /Land Cover Classification by Supervised classification method using ERDAS Imagine software:
   a) Maximum likelihood classifier.
   b) Mahalonbis Distance Clasifier
   c) Minimum Distance Classifier.
6) Land use Land Cover Classification by Unsupervised classification by ISODATA method
7) Comparison of overall accuracies of each method with respect to Performance Evaluation / Accuracy Assessment.
8) Output generation.

After acquiring the satellite images of the study areas, classification was made by using four methods viz. unsupervised classification and supervised classification which includes maximum likelihood, Mahalanobis and minimum distance methods. After the classification accuracy was tested by different methods of accuracy assessment and comparison was done for various classes. ERDAS IMAGINE 2011 developed by Leica Geo-systems software was used for classification and accuracy assessment.

V. Unsupervised Classification

ISODATA method of unsupervised classifications has been used in this study. It needs three input parameters viz. number of classes/clusters, maximum number of iteration and maximum percentage of pixels. In the present study values selected are 14 classes, 30 iterations and 0.95 as maximum percentage of pixels respectively. After the execution of the algorithm, the assigned classes (14 classes) were grouped into a number of categories according to their spectral appearance on screen. An investigation on the Land use / Land cover mapping was done to find out accuracy aspects of satellite imagery of Bhopal [8]. Fig. 3 depicts the results of application of ISODATA algorithm for the selected area.

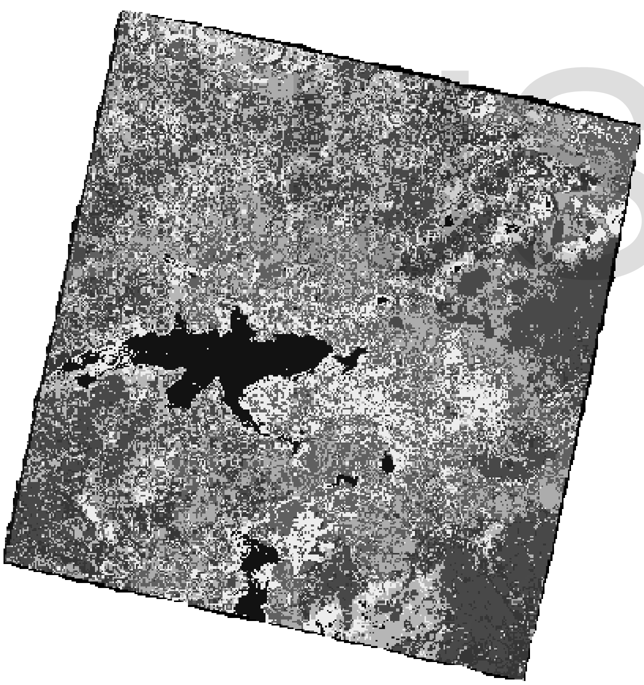

Fig. 3 Unclassified Methods- ISODATA

VI. Supervised Classification

Supervised classification algorithms need a prior knowledge of the study area (ground-truths) which may be obtained from different groups into six classes namely water bodies, agricultural fields, forest land, urban settlement, barren land and unclassified area. The ground-truth samples are introduced as sets of pixels selected to represent actual phenomena in order to train the computer system to recognize data patterns. These maps were geo-referenced and the locations as well as the distribution of feature classes of LCLU were extracted. Field visits to the study areas were undertaken during which data were collected in the area and the location of classes were recorded by GPS. According to these two sources, different ground-truths were recorded.

Using maximum likelihood method area distributed under six classes namely water bodies, agricultural fields, forest land, urban settlement, barren land and unclassified area are depicted in Fig. 4.

Using Mahalonbis method area distributed under six classes namely water bodies, agricultural fields, forest land, urban settlement, barren land and unclassified area are depicted in Fig. 5.

Using minimum distance method area distributed under six classes namely water bodies, agricultural fields, forest land, urban settlement, barren land and unclassified area are depicted in Fig. 6.

After classifications accuracy assessment was done for three supervised classifications. Producer accuracy, user accuracy and overall accuracy were calculated from the error matrices formed [9]. The results are shown in the table 1.

Identifying seed pixel was used in the supervised classification in this study for computer training. In this method the analyst defines a single pixel that is representative of the training sample and the computer system makes a comparison between the seed pixel and the contiguous pixels, based on some parameters specified by the analyst. When one or more of the contiguous pixels is accepted, the mean of the sample is calculated from the accepted pixels, and then the pixels contiguous to the sample are compared in the same way. This process repeats until no pixels that are contiguous to the sample satisfy the spectral parameters. While calculating the results the sample grows outward from the model pixels with each iteration. Three supervised classification methods were used in this study i.e. maximum likelihood, Mahalonbis and minimum distance. Producer accuracy, User accuracy and Overall accuracy were calculated by maximum likelihood method.

Maximum likelihood is one of the most popular supervised classification method used with remote sensing image data. This method is based on the probability that a pixel belongs to a particular class. The basic theory assumes that these probabilities are equal for all classes, and that the input bands have normal distributions. However, this method needs long time of computation, relies heavily on a normal distribution of the data in each input band and tends to over-classify signatures with relatively large values in the covariance matrix. The distance (spectral distance) method calculates the spectral distance between the measurement vector for the candidate pixel and the mean vector for each signature, and the equation for classifying by spectral distance is based on the equation for Euclidean distance. It requires the least computational time amongst other supervised methods, however, the pixels that should not be unclassified become classified, and it does not consider class variability.

Mahalanobis distance is similar to minimum distance, except that the covariance matrix is used instead. Unlike minimum distance, this method takes the variability of classes into account. It could be more useful than minimum distance in cases where statistical criteria must be taken into account, but the weighting factors that are available with the maximum likelihood option are not needed. However, this method tends to over-classify signatures with relatively large values in the covariance matrix. Also, it is slower to compute than minimum distance; and it relies heavily on a normal distribution of the data in each input band.

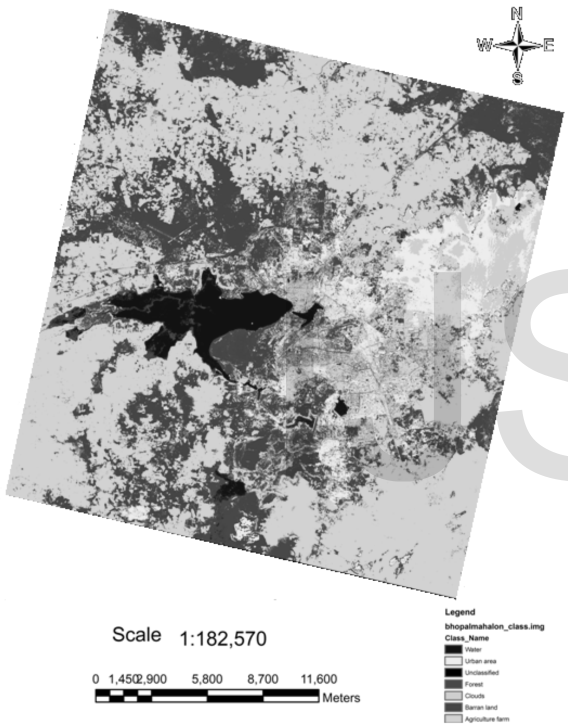

Fig. 4 Maximum Likelihood Method

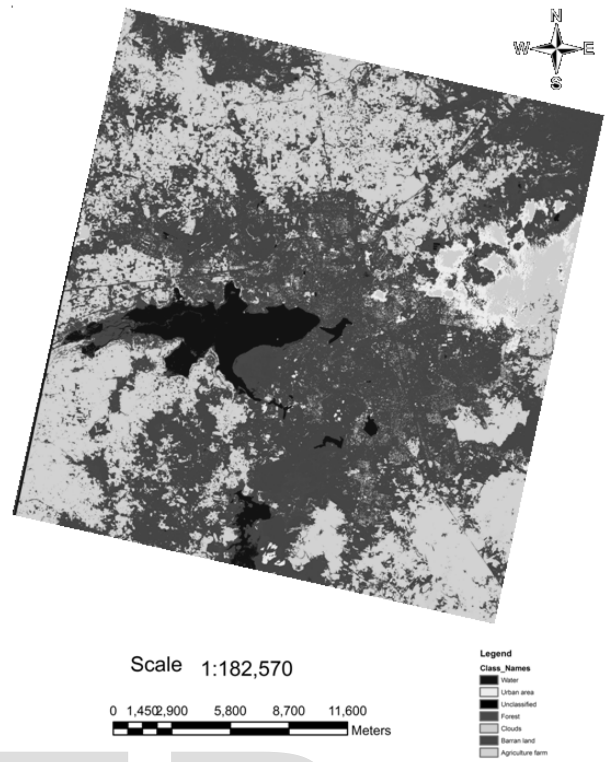

Fig. 5 Mahalanobis Method

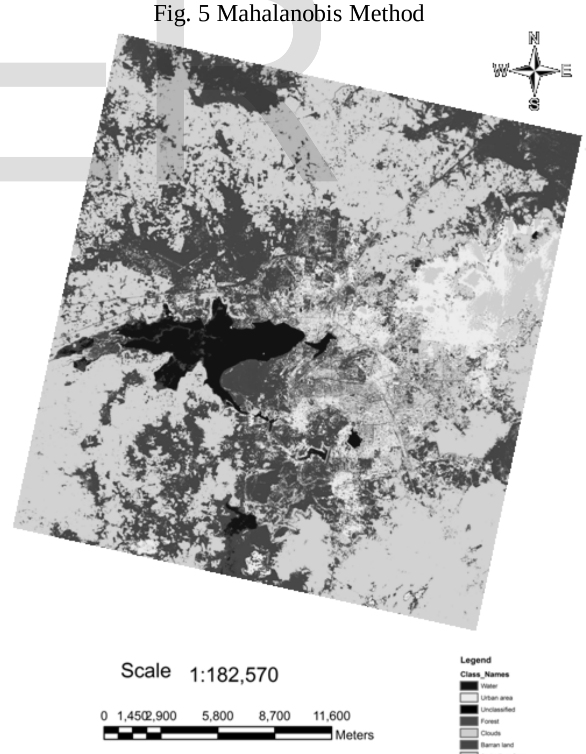

Fig. 6 Minimum Distance Method

VII. Results And Discussion

Accuracy assessment of classification can be defined as the process of comparing the classification with geographical data that are assumed to be true, in order to determine the accuracy of the classification process. Usually, the assumed-true data are derived from ground-truth data. Evaluating the accuracy of the classification was done here by applying accuracy assessment methods. In this study, the number of reference points, used for the accuracy assessment of classification were taken from the field visit, and the remaining from the topographic maps. The error matrix and the associated accuracies were computed by three methods of supervised classifications for the area under study three error matrices were produced. Table given below shows the results of error matrices of the three classification methods. The classification accuracy is most important aspect to assess the reliability of maps, especially when comparing different classification method. During this study the accuracy has been estimated on the training set samples and the results of all the classification techniques is summarized in the following table.

TABLE I
RESULTS OF ERROR MATRICES OF THE THREE CLASSIFICATION METHODS

Name of method	Producer accuracy	User Accuracy	Overall accuracy
Max. Likelihood	86.20	85.92	91.67
Min. Distance	92.36	93.84	90.83
Mahalonbis	87.04	83.88	83.33

The modern techniques of satellite image processing have been applied to extract information about land use/land cover mapping for Bhopal and adjoining areas. Land data was collected from different sources of information such as remotely sensed data and topographic maps of the area. The scope of this paper includes study of various classification methods of remote sensing satellite imagery. For classification of image, same pixel classes were considered for various classification methods.

The performance of different classifications was evaluated using same training set pixels. The analysis and result shows that maximum likelihood classifier shows highest accuracy with an average accuracy of about 91.67% followed by minimum distance classifier with 90.83% and Mahalanobis distance classifier with 83.33%. The Unsupervised classification however could not show the clear overall accuracy all the classed are intermixed with each other.

In unsupervised method there is every possibility if merging large data resulting in overlapping of classes. It is difficult to segregate the classes from one another and the errors of omission and commission are greater.

VIII. CONCLUSION

This study underlines the performance of four widely used classification techniques for land use/land cover. The performance of maximum likelihood classifier is reliable assessing high accuracy followed by Mahalanobis distance classifier. Maximum likelihood classifier also gives reliable result and can be used with neural network classifier for comparison and better assessment of the land use/land cover. Error matrices produced to evaluate the classification methods show that the best overall classification accuracy method was the maximum likelihood. The accuracy may be improved by taking more number of reference points and correct ground truthing using Dfferential GPS. Error matrices produced to evaluate the classification methods show that the best overall classification accuracy method was the maximum likelihood for all the three sub-catchments; with an average accuracy of about 91.67%. The second best overall classification accuracy method was minimum distance with an average accuracy of 90.83% and lowest accuracy was found by Mahalanobis distance; with an average accuracy of 83.33%.

REFERENCES

[1] Congalton, R., “A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment” Vol 37, 199, pp. 35-46.
[2] Engman, E. T. and Gurney, R. J., “Remote Sensing in Hydrology”, Chapman et. Al., London, 1991, 225 p.
[3] Ragan, R.M. and Jackson, T.J., “Runoff synthesis using Landsat and SCS model”, Journal of Hydraulic Division, ASCE (106), 1980, pp. 667-678.
[4] Lillesand, T.M. and Kiefer, R.W., “Remote Sensing and Digital Image Interpretation”, 2000.
[5] Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R.E., “A land use and land cover classification system for use with remote sensor data”, U.S. Geological Survey Professional Paper, No. 964, USGS, Washington, D.C., 1976.
[6] ERDAS IMGINE, “ERDAS Field Guide”, 6th Edition, 2010, pp. 686.
[7] Lyon, J. G., "Remote Sensing and Geographic In-formation Systems in Hydrology”, in Ward, A. D., and Elliot, W. J., (eds.), Environmental Hydrology, CRC Press, 1995, pp: 337-367.
[8] Paliwal M.C & Katiyar S.K., “Investigation on the accuracy aspects in land use/land cover mapping using remote sensing satellite imagery-a case study”, India Geospatial Forum, 22-24 January, Hyderabad, India. 2013, pp. 51.
[9] Claudio Conese and Fabio Maseui, (1992), “Use of Error Matrices to Improve Area Estimates with Maximum Likelihood Classification Procedures”, REMOTE SENSING AND ENVIRONMENT Vol 40, 1992, 113-124.