Classification of Heterogeneous Sites using IRS-P5 Panchromatic Satellite Imagery

Sinchana G S

Government Engineering College, Kushalnagar, Karnataka, INDIA

Choodarathnakara A L and Dr. Shivaprakash Koliwad

Government Engineering College, Kushalnagar, Karnataka, INDIA

Malnad College of Engineering, Hassan

ABSTRACT

Remotely Sensed data is an important component of Land Use/Land Cover (LU/LC) studies. This paper compares the performance of ISODATA classification with Mahalanobis Distance classification for Arsikere semi-urban study area of Karnataka State in INDIA using IRS-P5 PANF satellite imagery. The Arsikere study area is an heterogeneous region comprised primarily of water features mixed with impervious features. During the beginning of experiment, ISODATA unsupervised classification was applied on IRS data. Later, Mahalanobis Distance classification was applied on IRS data with 6000 training sites and 100 validation points for water, vegetation, soil and impervious surface features which were randomly generated using a stratified sampling approach. The LU/LC data associated with these points were then compared with Topographic Maps (Survey of India, No. D43Q3, D43Q7) and Ground Truth Data for performance analysis. Based on the confusion matrices obtained for the sample set, the OCA, Kappa statistics were compared with ISODATA. The experimental analysis shows that unsupervised ISODATA classification provides accuracy of 88% in Arsikere, semi-urban area, however Mahalanobis distance classification give up 85% OCA with TS = 6000 and VS = 100.

Keywords

Remote Sensing, Heterogeneous-sites, ISODATA, Mahalanobis Distance.

1. INTRODUCTION

Remote sensing is the sensing of an object or a phenomenon from a remote distance. In remote sensing, information transfer is accomplished by use of Electromagnetic Radiation (EMR). EMR is a form of energy that reveals the matter presence by producing reflected energy when it strikes the matter. In remote sensing, the important task is classifying the spectral measurements which are obtained from remote sensors. Classification of image is a difficult task that affects many factors. For improving the classification accuracy through essential use of features present in remote sensed data expect suitable classification techniques.

D. Lu and Q. Weng [1] mentioned that the accuracy of classification accuracy improves the type of satellite image spatially and spectrally. The spatial variability and attributes are determined by other factors as well. However, most of the classification results are affected by data itself than preprocessing, enhancement techniques and classification schemes adopted.

Tibebu Kassawmar et al. [2] proposed ISODATA clustering method using Landsat TM images of Ethiopian highlands to classify forest, woodland, shrub land/ bush land, crop land, grassland, barren land with most of the classes showing a user's and producer's accuracy of more than 80%. M. Mohammady et al. [3] focused to use unsupervised ISODATA and supervised Maximum Likelihood Classifier techniques for land mapping with the case study of the Baghsalian watershed in Iran. Broad leaf forest, conifer forest, rangeland, agricultural land, water bodies and residential land are the main land use types in study area. The synthetic approach involves supervised and unsupervised techniques along with the decision rules and yields an accuracy of 98.2%. For planning and management purposes 85% level of accuracy is considered to be suitable.

Shrinivas Khandare and Urmila Shrawankar [4] proposed an integrated method consisting of maximum likelihood and fuzzy classification. In this method, a suitable partitioning is detected in an iterated manner. The resultant classes of both the procedures are compared with k-means, ISODATA, ML classification algorithms. The fuzzy based classification method found to give better results. In Sahar A. El_Rahman's [5] paper, unsupervised hyper spectral image classification such as Iterative Self-Organizing Data Analysis method was used to extract agricultural information using ENVI software. After applying ISODATA algorithm and Principal Component Analysis (PCA) the performances were evaluated using OCA. In this experiment, the ISODATA image classification gives 75.6187% OCA.

Giles M. Foody [6] has elucidated the techniques used for assessing the classification accuracy that are commonly suggested. In image classification various types of errors will encounter frequently which affect the classification of image. Ashok Kumar T [8] presented the performance and employability of the decision tree classification algorithm in respect of varying training dataset size for class hierarchy levels I and II along with effects of ancillary data on tree complexity with number of rules induced.

A. L. Choodarathnakara [9] was taken up research work with the objective of designing an efficient and reliable classification strategy in an attempt to find answers to some of the conflicting issues dealt within the existing literature pertaining to classification of fine resolution RS data. The authors have conducted experiment on the MS data of IRS LISS-IV sensor of 5m spatial resolution and PAN data of 2.5m spatial resolution. Authors have concluded that the hard classification procedure fails to classify mixed pixel problem in Arasikere semi-urban area of Karnataka State, INDIA. To overcome this problem, authors proposed Decision Tree technique along with Mamdani_Fuzzy Inference System (M_FIS) as a hybrid classifier and concluded that Mamdani_FIS was a powerful candidate to classify mixed pixels present in semi-urban areas. Ajay D. Nagne et al. [10] proposed Mahalanobis classifier to assess a Land Use Land Cover of Aurangabad Region on Hyperspectral EO-l Hyperion imaging Data. The image was classified into water, vegetation, hill without vegetation, settlement area and bare soil, resulting in 88.46% accuracy of classifier with Kappa Coefficient 0.84.

B. R. Shivakumar and S. V. Rajashekararadhya [11] analyzed the heterogeneous multispectral LANDSAT8 data with three traditional classifiers namely Maximum Likelihood classifier, Minimum-distance-to means and Mahalanobis distance by defining separability of classes using Normalized Euclidean distance. The performance of all classifier is carried out statistically by evaluating the accuracy assessment. A relative analysis was conceded to highlight the dissimilarity in classification outcome. A. L. Choodarathnakara et al. [12] proposed PCA method to detect built-up features using LANDSAT 7 ETM+ Satellite Imagery. In this method, first three components PCA1, PCA2 and PCA3 were fused to get PCA1+PCA2+PCA3 with 98% of six dimensions B1, B2, B3, B4, B5, B7 and PCA model was successful with 98% accuracy.

The objective of this research work was to assess and compare the accuracy of unsupervised ISODATA classification with Mahalanobis Distance. This research analyzed the study areas with heterogeneous land cover / land use compositions. The aim of this paper is to study semi-urban area for urban planning purpose using supervised Mahalanobis Distance Classification and unsupervised ISODATA Classification with IRS Satellite Imagery. The rest of the paper is organized as follows, Section I contains the introduction about remote sensing, Section II contain the remote sensing image classification, Section III contain the study area and the methodology proposed, Section IV contain results and discussion, section V concludes the final comments on the research work.

2. REMOTE SENSING IMAGE CLASSIFICATION

Image classification has formed a significant part in the field of Remote Sensing, Pattern Recognition and Image Analysis. In image classification, all the pixels can be sorted in an image into a limited number of individual classes. Classification of image involves supervised and unsupervised, non parametric and parametric, soft and hard classification, per-pixel, subpixel and perfield techniques. The classifiers are categorized into hard and soft classifiers depending on whether output is a definitive decision about land cover class or not a class. The classification paradigm can be expressed in the form of a tree as shown in Fig. 1.

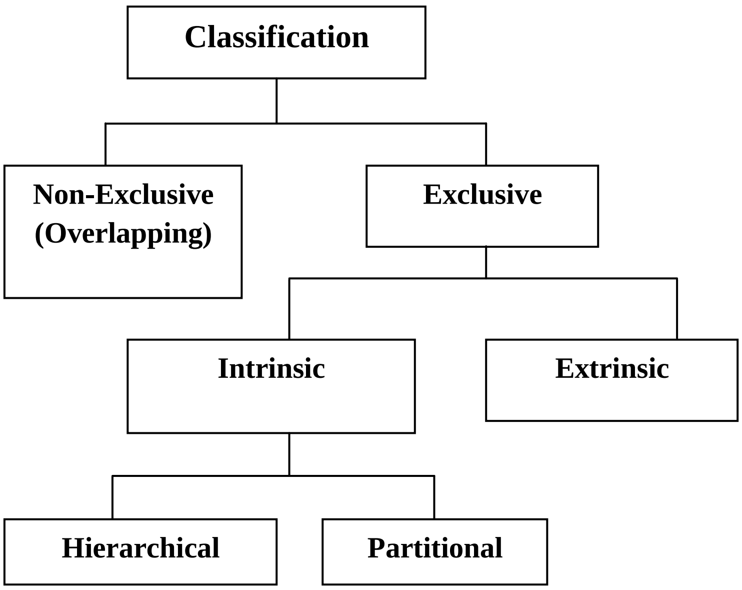

Fig 1: Classification Tree

2.1 Unsupervised ISODATA Classification

In unsupervised classification, most commonly used algorithm is ISODATA. Every class requires a class mean and covariance matrix to follow a multivariate normal distribution. It performs an iterative method and hence it is referred as an extension of the K-means algorithm. To find the data cluster centre, the author proposed K-means clustering procedure. Use of K-means clustering results in K number of clusters. Initially, cluster means and covariances are calculated followed by classification of each pixel by assigning to the nearest cluster. This process repeats until the iterations are "small enough". One of the main differences between K-means algorithm and ISODATA is with different number of clusters but the K-means assumes the number of clusters is known a priori [7].

2.2 Supervised Mahalanobis Distance Classification

Mahalanobis Distance is same as minimum distance except the covariance matrix. The covariance and variance parameter form clusters which are highly varied lead to similarly varied classes and vice versa. The Accuracy is a measure of how well the model correlates an outcome with attributes in the data that has been provided. There are various measures of accuracy, but all measures of accuracy are dependent on the data that are used. But in this work, UA, PA, kappa statistics are used as a measure of accuracy.

3. STUDY AREA AND METHODOLOGY

3.1 Satellite Data Products and Study Area

The data product utilized for the study is the Panchromatic RS image of IRS-P5 Cartosat-I satellite which has been launched and further supervised by ISRO. This satellite data product was procured from the NRSC, Hyderabad, India. Table I provides the specifications of satellite data utilized for the purpose of semi-urban study. The study area considered for this research work is semi-urban area of Arsikere taluk situated in Hassan district of Karnataka State, India with geographical coordinates of 13° 18' 50" North, 76° 15' 22" East and with original name ARASIYA KERE. Fig. 2 shows the satellite image of Arsikere semi-urban study area of Hassan district, Karnataka State.

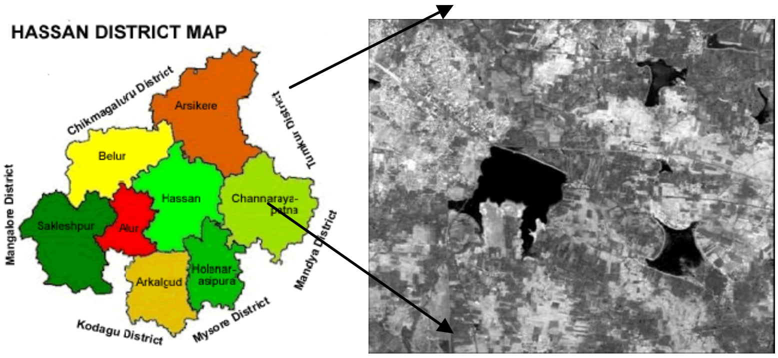

Fig 2: Arsikere semi-urban study area of Hassan District, Karnataka state, INDIA

TABLE I. Details of the satellite data products used in the study

Satellite and Data type	Date of Acquisition	Spectral Resolution	Spatial Resolution
IRS-P5 PANF	04/04/2011	0.55-0.85 µm	2.5 m

3.2 Proposed Methodology

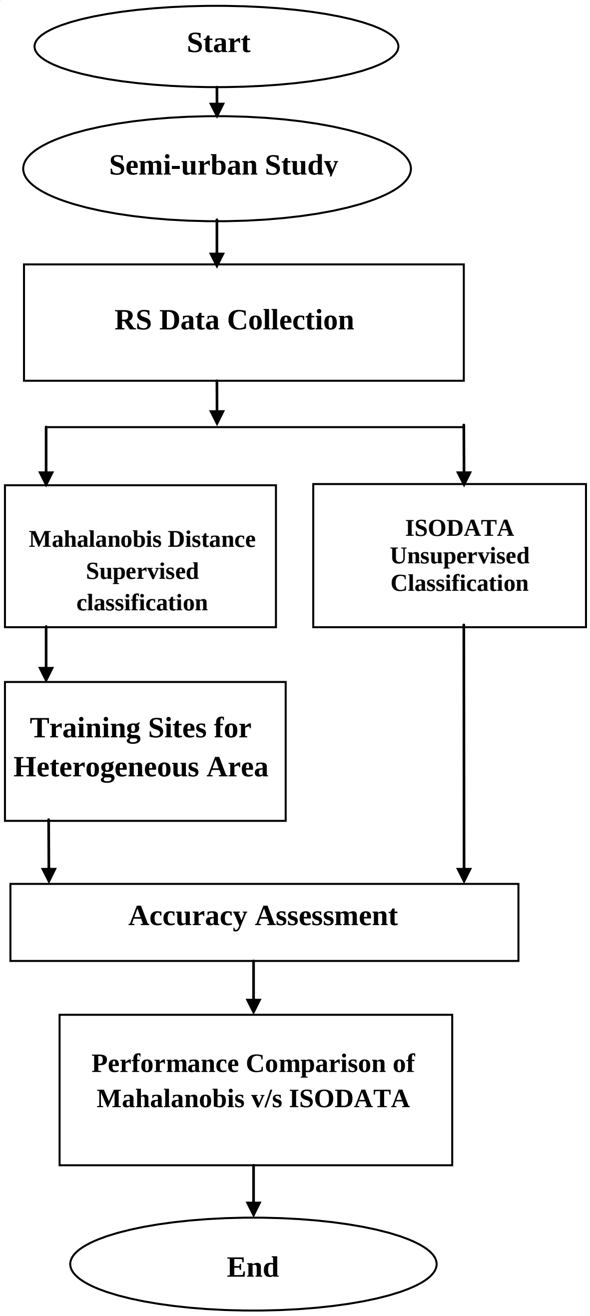

Fig 3: Proposed Methodology for classification of LU/LC features in Arsikere Semi-urban

The township of Arsikere study area is undergoing lot of changes being a semi-urban area between city and village land. Arsikere is connected to various cities in the karnataka via bus and rail transport. In order to plan such a semi-urban land, the accurate classification of land use/ land cover features is necessary. In this context, the methodology is proposed which has been depicted in Fig 3. During the first phase of the work the RS data was procured from NRSC Hyderabad. In the second phase, heterogeneous sites are identified for generating training site samples of water, vegetation, soil and impervious features of about 6000 TS. By employing Mahalanobis distance supervised classification confusion matrix was analyzed for four classes with 100 validation points for heterogeneous sites. The performance comparison of ISODATA versus Mahalanobis distance classification was dealt with the help of OCA and cross table as well.

4. EXPERIMENTAL RESULT ANALYSIS

The classification of geo-coded 2.5m spatial resolution data has been made using unsupervised ISODATA and supervised Mahalanobis distance algorithms. Evaluation of these classifiers is performed using accuracy assessment which was carried out using ERDAS IMAGINE V 9.2 RS image processing software.

4.1 Mahalanobis Distance Supervised classification

Once the heterogeneous areas were identified, supervised Mahalanobis distance classification was performed on the study area. Fig. 4 depicts Heterogeneous training sites of the Arasikere semi-urban study area. Fig. 5 shows Supervised Mahalanobis Distance Classified Image for Heterogeneous sites.

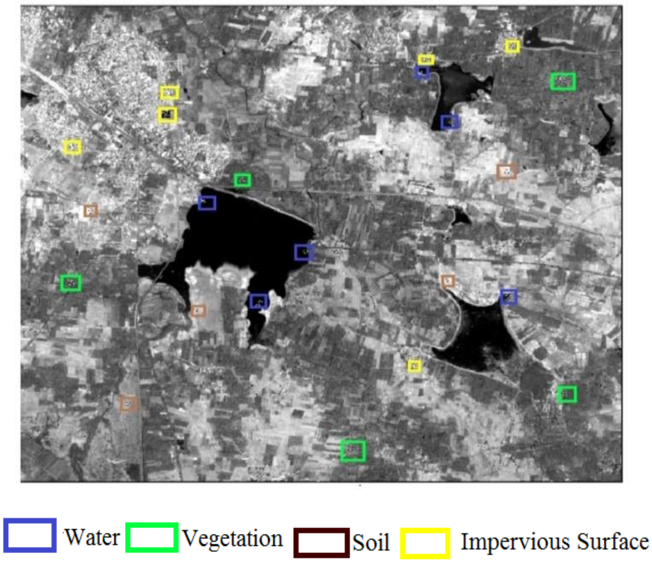

Fig 4: Heterogeneous training sites of the Arasikere semi-urban study area.

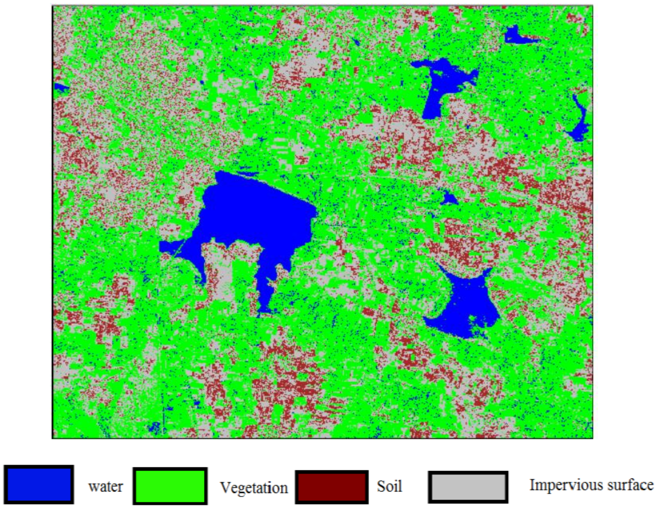

Fig 5: Supervised Mahalanobis Distance Classified Image of Heterogeneous Study Area.

TABLE II. Confusion Matrix & Kappa values obtained for Mahalanobis Distance supervised classification with TS = 6000 with VS = 100

Classes	1	2	3	4	Row Total	UA%
1	8				8	100
2	2	44	1	3	50	88
3			6	4	10	60
4		3	2	27	32	84.38
Column Total	10	47	9	34	85	
PA%	80	93.62	66.67	79.41		OCA: 85.00%
Kappa	1.0000	0.7736	0.5604	0.7633		OKS: 0.7653

Class Legends: 1: Water; 2: Vegetation; 3: Soil; 4: Impervious surface

In TABLE II, out of ten reference pixels of water, eight are correctly classified as water and the rest 20% are misclassified to vegetation class producing a PA of 80% for water. In other words eight pixels are correctly classified as water on the image and produce an UA of 100%. Out of 47 reference pixels of vegetation 44 are correctly classified as vegetation and the rest 6.38% are misclassified to soil class producing a PA of 93.62%. In other words out of the total 50 pixels which are classified as vegetation on the image, only 44 pixels represent vegetation and produce an UA of 88%. The remaining 12% of the pixels which are classified as vegetation are the misclassified pixels from water, soil and impervious surface.

Out of nine reference pixels of soil six are correctly classified as soil and the rest 33.33% are misclassified to vegetation and impervious surface classes producing a PA of 66.67%. In other words out of the total ten pixels which are classified as soil on the image, only six pixels represent soil and produce an UA of 60%. The remaining 40% of the pixels which are classified as soil are the misclassified pixels from impervious surface.

Out of 34 reference pixels of impervious surface 27 are correctly classified as impervious surface and the rest 20.59% are misclassified to vegetation and soil classes producing a PA of 79.41%. In other words out of the total 32 pixels which are classified as impervious surface on the image, only 27 pixels represent impervious surface and produce an UA of 84.38%. The remaining 15.62% of the pixels which are classified as impervious surface are the misclassified pixels from vegetation and soil classes.

4.2 ISODATA Unsupervised Classification

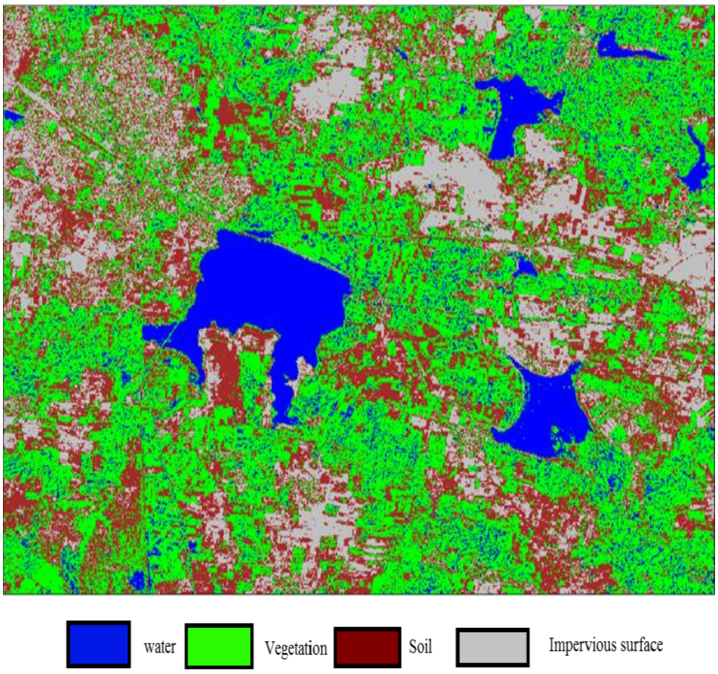

Fig 6: ISODATA classified Image for the Heterogeneous Study Area.

TABLE III. Confusion Matrix & Kappa Values obtained for ISODATA Classification for four Classes with 100 Validation Points

Classes	1	2	3	4	Row Total	UA%
1	11				11	100
2	3	36	3		42	85.71
3		4	24	1	29	82.76
4			1	17	18	94.44
Column Total	14	40	28	18	88	
PA%	78.57	90	85.71	94.44		OCA: 88.00%
Kappa	1.000	0. 761	0.760	0.932		OKS: 0.829

Class Legends: 1: Water; 2: Vegetation; 3: Soil; 4: Impervious surface

Fig. 6 shows ISODATA classified Image of the Heterogeneous Study Area. TABLE III depicts Confusion Matrix & Kappa values obtained for ISODATA Unsupervised

Classification for four classes with training sites of 6000 and 100 validation points for heterogeneous areas.

In ISODATA classification, out of 14 reference pixels of water, 11 are correctly classified as water and the rest 21.43% are misclassified to vegetation class producing a PA of 78.57% for water. But, all the 11 pixels are correctly classified as water on the image and produce an UA of 100%. Out of 40 reference pixels of vegetation, 36 are correctly classified as vegetation and the rest 10% are misclassified to soil class producing a PA of 90% for vegetation. In other words, out of the total 42 pixels which are correctly classified as vegetation on the image, only 36 pixels represents vegetation and produce an UA of 85.71%. The remaining 14.29% of the pixels which are classified as vegetation are the misclassified pixels from water and soil classes.

Out of 28 reference pixels of soil, 24 are correctly classified as soil and the rest 14.29% are misclassified to vegetation and impervious surface classes producing a PA of 85.71% for soil. In other words, out of the total 29 pixels which are correctly classified as soil on the image, only 24 pixels represents soil and produce an UA of 82.76%. The remaining 17.24% of the pixels which are classified as soil are the misclassified pixels from vegetation and impervious surface classes.

Out of 18 reference pixels of impervious surface, 17 are correctly classified as impervious surface and the rest 5.56% are misclassified to soil class producing a PA of 94.44% for impervious surface. In other words, out of the total 18 pixels which are correctly classified as impervious surface on the image, only 17 pixels represents impervious surface and produce an UA of 94.44%. The remaining 5.56% of the pixels which are classified as impervious surface are the misclassified pixels from soil class.

4.3 Comparison of ISODATA and Mahalanobis Distance

Fig. 7 presents comparative graph of OCA for heterogeneous training sites for TS = 6000 with VS = 100 pixels using Supervised Mahalanobis Distance and Unsupervised ISODATA classification techniques.

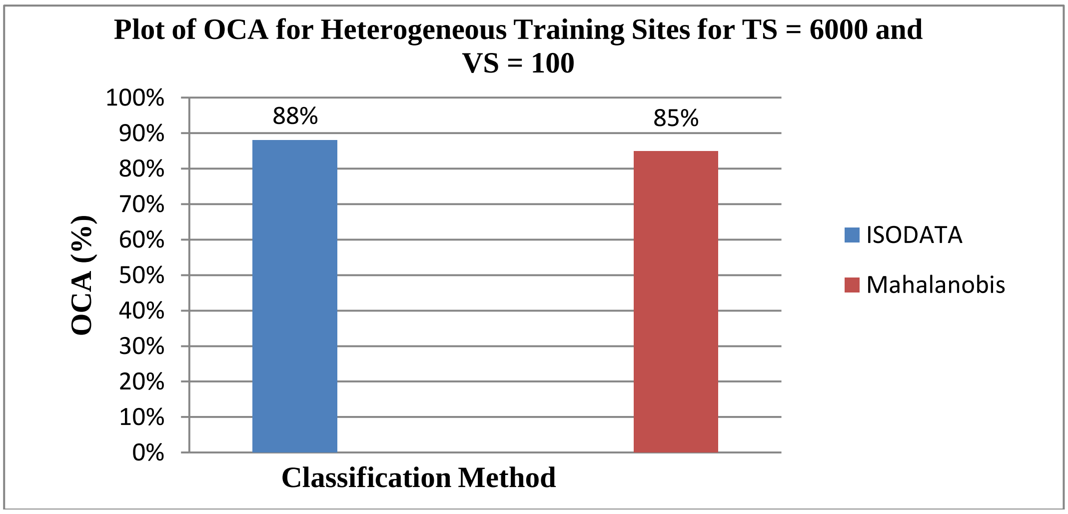

Fig 7: Plot of OCA for Heterogeneous Training Sites for TS = 6000 and VS = 100

TABLE IV. Cross Table between ISODATA and Mahalanobis Distance for Heterogeneous TS = 6000 with VS = 100

ISODATA in Pixels	Mahalanobis Distance in Pixels			
	Water	***Vegetation***	***Soil***	***Impervious surface***
Water	**0.7272**			
Vegetation		**1.2222**		
Soil			**0.25**	
Impervious surface				**1.5882**

The comparative results demonstrate that the water class estimated by Mahalanobis Distance was 0.72 times the water estimated by ISODATA. The vegetation class estimated by Mahalanobis Distance was 1.22 times the vegetation estimated by ISODATA. The soil class estimated by Mahalanobis Distance was 0.25 times the soil estimated by ISODATA. Impervious surface class estimated by Mahalanobis Distance was 1.58 times the impervious surface estimated by ISODATA.

5. CONCLUSION

In this work the study area considered is Arsikere taluk in Hassan district which is a semiconducting area with moderate rainfall. The experimental results concludes that ISODATA classification provides 88% OCA accuracy for Arasikere semi-urban area but Mahalanobis distance yields out 85% OCA with TS = 6000 and VS = 100. Moreover, the satellite data used in this study consisting of only one band and hence it is not possible to classify more land use/land cover classes. Hence the work can be continued by procuring high spatial and spectral resolution data with more number of bands. Since the Arasikere study area is a semi-urban area which consisting of mixed pixels soft classification approaches can be performed for better classification accuracy.

ACKNOWLEDGMENT

The authors would like to graciously thank National Remote Sensing Centre (NRSC), Hyderabad, INDIA for providing the data product for the study. Thanks to the Survey of India (SOI) for providing the Topographic Maps for this study and Karnataka State Remote Sensing Application Centre (KSRSAC), Bangalore. Acknowledgements are due towards beloved students Kiran H T USN: 4GL14EC023, Sudhan A T USN: 4GL14EC055, Indresha C A USN: 4GL15EC402, Shalini M C USN: 4GL15EC407. Special thanks to teaching and non-teaching faculty of GEC, Kushalnagar for their encouragement and guidance.

REFERENCES

[1] D. Lu and Q. Weng, 2007. "A Survey of Image Classification Methods and Techniques for Improving Classification Performance", International Journal of Remote Sensing, Vol. 28, No.5, pp. 823-870.

[2] Tibebu Kassawmar, Sandra Eckert, Kaspar Hurni, Gete Zeleke & Hans Hurni, 2016. "Reducing Landscape Heterogeneity for Improved Land Use and Land Cover (LULC) Classification across the Large and Complex Ethiopian Highlands", Geocarto International, doi: 10.1080/10106049.2016.1222637.

[3] M. Mohammady, H. R. Moradi, H. Zeinivand and A. J. A. M. Temme, 2015. "A Comparison of Supervised, Unsupervised and Synthetic Land Use Classification Methods in the North of Iran", International Journal Environment Science Technology, Springer, Vol. 12, pp. 1515-1526.

[4] Shrinivas Khandare and Urmila Shrawankar, 2014. "Classification of High Resolution Pixels into Non Heterogeneous groups Using Rule Based Fuzzy Approach", IEEE International Conference on Computational Intelligence and Computing Research.

[5] Sahar A. El_Rahman, 2015. "Hyperspectral Imaging Classification Using ISODATA Algorithm: Big Data Challenge", Fifth International Conference on e-Learning, IEEE, pp. 247-250.
[6] Giles M. Foody, 2002. "Status of Land Cover Classification Accuracy Assessment", Remote Sensing of Environment, pp. 185-201.
[7] Asmala Ahmad and Suliadi Firdaus Sufahani, 2012. "Analysis of Landsat 5 TM Data of Malaysian Land Covers Using ISODATA Clustering Technique", IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE 2012), Melaka, Malaysia, pp.92-97.
[8] Ashok Kumar T, 2010. "Advanced Image Processing Techniques and Algorithms For Classification of Semi-Urban Land Use/ Land Cover Features using High Resolution Satellite Data", Ph.D. Thesis, Visvesvaraya Technological University.
[9] A L Choodarathnakara, 2018. "A Soft Computing Approach For Classification Of Land Use / Land Cover Features In Semi-Urban Area Using High Resolution Satellite Imagery", Ph.D. Thesis, Visvesvaraya Technological University.
[10] Ajay D. Nagne, Rajesh K. Dhumal, Amol D. Vibhute, Yogesh D. Rajendra, Sandeep Gaikwad, K. V. Kale and S. C. Mehrotra Dept, 2017. "Performance Evaluation of Urban Areas Land Use Classification from Hyperspectral Data by Using Mahalanobis Classifier", 11th International Conference on Intelligent Systems and Control (ISCO), IEEE, pp. 388-392.
[11] B. R. Shivakumar and S. V. Rajashekararadhya, 2017. "Spectral Similarity for Evaluating Classification Performance of Traditional Classifiers", IEEE WISPNET Conference, pp. 1999-2004.
[12] A L Choodarathnakara, Sinchana G S, Sowkya H K and Ramya B R, April 2018. "Automated Detection of Built-up Feature in Bangalore City using LANDSAT 7 ETM+ Satellite Imagery", International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 7, Issue 4, pp. 292-301.