International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

Approach to Accuracy Assessment tor RS Image Classification Techniques

Pravada S. Bharatkar[1] and Rahila Patel[1]

ABSTRACT - Several techniques exist for remote sensing (RS) image classification, which includes supervised and unsupervised approaches. Classified maps are the main product of remote sensing image classification. Accuracy assessment of these classified maps is one of the foremost and important tasks of RS image classification technique. Without accuracy assessment the quality of map or output produced would be of lesser value to the end user. However, supervised and unsupervised techniques show different levels of accuracy after accuracy assessment was conducted. This paper describes a study that was carried out to perform supervised and unsupervised techniques on remote sensing data for land use/cover classification and to evaluate the accuracy result of both classification techniques. The study used IRS 1C LISS III satellite image consists of 26718 pixels, which covers Ralegaon Siddhi watershed in Ahmednagar district of Maharashtra state, India as a primary data and topographical map of SOI, cadastral map, and district statistical handbook containing land use/cover information as ancillary data. The land use/cover classes for the study area were classified into 5 themes namely, agricultural land, shrubs, water body, wasteland and barren land. Ground verification was carried out to verify and assess the accuracy of classification. A several sample points with sufficient numbers of samples were collected based on Systematic Random sampling criteria. The comparative analysis based on the overall accuracy and Kappa statistics for the various classiers reflects the better performance of maximum likelihood classification technique.

Keywords: Accuracy assessment, image classification, land use/cover, remote sensing, satellite image

1.0 Introduction:

With the advent of more advanced digital satellite remote sensing techniques, the necessity of performing an accuracy assessment has received renewed interest. This is not to say that accuracy assessment is unimportant for the more traditional remote sensing techniques. However, given the complexity of digital classification, there is more of a need to assess the reliability of the results. Accuracy assessments determine the quality of the information derived from remotely sensed data [1]. It can be either qualitative or quantitative. In qualitative assessments, we determine if a map "looks right" by comparing what we see in the imagery with what we see on the ground. This is not usually done in a rigorous way, but rather in a "quick and dirty" way. General accuracy is the goal in this case, and error and its sources are not as important. This is usually a first cut assessment. However quantitative assessments attempt to identify and measure remote sensing-based map error. In such assessments, we compare map data with reference or ground truth data.

[1]*Department of Computer Science, RCERT, Nagpur University, Chandrapur, India*
Corresponding Author: [1]*pravadadilip@gmail.com*

As we know, Remote Sensing (RS) image classification is a complex process, which includes determination of a suitable classification system, selection of training samples, image preprocessing, feature extraction, selection of suitable classification approaches, post-classification processing, and accuracy assessment as a major steps. Evaluation of classification performance is an important step of the classification process. Different approaches ranging from a qualitative evaluation based on expert knowledge to a quantitative accuracy assessment based on sampling strategies may be employed for this purpose. [2] Proposed six criteria: accuracy, reproducibility, robustness, ability to fully use the information content of the data, uniform applicability, and objectiveness to evaluate the performance of classification algorithms. Accuracy assessment is the last and essential step of classification to evaluate the performance of classification.

1.1 Accuracy Assessment

As mentioned, accuracy assessment is an important and essential step in the classification process. The goal is to quantitatively determine how effectively pixels were grouped into the correct feature classes in the area under investigation. The accuracy assessment assesses how well a classification worked and understands how to interpret

International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

the usefulness of someone else's classification. It generally includes three basic components: sampling design, response design, and estimation and analysis procedures [3]. Selection of a suitable sampling strategy is a critical step. The major components of a sampling strategy include sampling unit (pixels or polygons), sampling design, and sample size. Depending on the goal of the accuracy assessment the number of sample plots can be calculated with different methods. As a rule of thumb, [1] recommends at least 50 training pixels per class. If the area exceeds 500 km^2 or the number of categories is more than 12, then at least 75-100 training pixels should be taken per class. Any sampling scheme should satisfy three criteria: It should have a low probability of accepting a map of low accuracy, it should have a high probability of accepting a map of high accuracy, as well as it should require a minimum number, N, of ground truth samples.

Possible sampling designs include random, stratified random, systematic, double, and cluster sampling. A detailed description of sampling techniques can be found in [1]. In Simple_Random Sampling, observations are randomly placed, while in Stratified Random Sampling, minimum numbers of observations are randomly placed in each category. With systematic Sampling, the observations are placed at equal intervals according to a strategy and systematic Non-Aligned Sampling, a grid provides even distribution of randomly placed observations. Finally, in the cluster sampling, placed "centroids" used as a base of several nearby observations.

The purpose of quantitative accuracy assessment is the identification of the sources of errors. In addition to errors from the classification itself, other sources of errors, such as position errors resulting from the registration, interpretation errors, and poor quality of training or test samples, all affect classification accuracy. Errors exist in any classification due to misidentification, excessive generalization, misregistration, etc. One of the most common means of expressing classification accuracy compare on a category by category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. The most common error estimate is the overall accuracy, while Kappa coefficient (K) is the measure of agreement of accuracy. It provides a difference measurement between the observed agreement of two maps and agreement that is contributed by chance alone. Kappa is usually attributed to [4], but Kappa has been derived independently by others and citations go back many years [5]. It became popularized in the field of remote sensing and map comparison by [1, 6, and 7] to name a few. In particular, it is state that "Kappa analysis has become a standard component of most every accuracy assessment and is considered a required component of most image analysis software packages that include accuracy assessment procedures" [1]. Kappa analysis is recognized as a powerful method for analyzing a single error matrix and for comparing the differences between various error matrices [7, 8, and 9]. Smits et al. [7] studied number of evaluation methods of accuracy assessment and concluded that the methods based on confusion matrices and the K_{hat} statistical analysis are the most suited.

Hence, there is scope to evaluate the performance of various RS image classification methods based on some statistical parameters via; confusion matrix and its kappa co-efficient to suggest the most efficient and accurate RS image classification method for effective land use mapping.

2.0 Materials and Method

2.1 Study Area

The present study has been carried out using the Indian Remote Sensing Satellite IRS – 1D LISS-III (Indian Remote Sensing-Linear Imaging Self-Scanning Sensor III) data of Ralegaon Sidhi watershed, Maharashtra. This data has been procured from National Remote Sensing Agency (NRSA), Hyderabad, INDIA and contains the information in bands R, G, B, NIR and SWIR with a swath of 141km in the format of LGSOWG ((Landsat Ground Station Operators Working Group) of the electromagnetic spectrum with a spatial resolution. The dataset consists of 2282 x 2507 pixels. The advantage of using this dataset is the availability of the referenced date set (Table 1) produced from field survey, which is used for the classification accuracy purpose. The study watershed is lying between 18^{0}54' N to 18^{0}57' N and longitudes of 74^{0}23' E to 74^{0}27' E as shown in Fig. 1. The total geographical area of the watershed is 1070.52 ha. Physiographically, the area is surrounded by small hillocks with fractured rocks. The average temperature of the area varies between 12^{0}C to 44^{0}C. The average annual rainfall of the region is 601 mm.

3.0 Methodology

The IRS 1D LISS III satellite image and scanned survey of India (SOI) toposheet were digitally processed using the GIS software namely, Integrated Land and Water Information System (ILWIS) to prepare the base map of the watershed boundary, contour drainage and road maps of study area for ground verification. There are basically two types of data collected in support of remote sensing accuracy assessments: other remote sensing data and ground-based data. Reference data was taken using the same schemes used in the classification efforts. Since, ground based data is assumed to be 100% correct in accuracy assessments, due care was taken during the data

International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

collection. The accuracy data is independent of reference or groudtruth data.

3.1 Accuracy

One basic accuracy measure is the overall accuracy, which is calculated by dividing the correctly classified pixels (sum of the values in the main diagonal) by the total number of pixels checked.

Overall accuracy (%) = (Correctly classified pixels/Total number of pixels) ---------------------(1)

Besides the overall accuracy, classification accuracy of individual classes can be calculated in a similar manner. Two approaches are possible. More specific measures are needed because the overall accuracy does not indicate how the accuracy is distributed across the individual categories. The categories could, and frequently do, exhibit drastically differing accuracies but overall accuracy method considers these categories as having equivalent or similar accuracies. There are at least two methods can be used to determine individual category accuracies.

(1) The ratio between the number of correctly classified pixels and the classified totals pixels of particular LULC class is the user's accuracy - because users are concerned about what percentage of the classes has been correctly classified.

User's accuracy (%) = (Correctly classified pixels /Classified total pixels) --------------------(2)

(2) The ratio between the number of correctly classified pixels and the reference total pixels for particular LULC class is called the producer's accuracy.

Producer's accuracy (%) = (Correctly classified pixels/Reference total pixels) ------------------(3)

A more appropriate way of presenting the individual classification accuracies are as follows;

Commission error = 1 - user's accuracy---------------------(4)

Omission error = 1 - producer's accuracy ----------------(5)

The kappa coefficient (K) can be computed as follows,

$$K = \frac{P_0 - P_c}{1 - P_c} \qquad \text{--------------------------(6)}$$

Where,

P_0 = proportion of units which agree, = overall accuracy

P_c = proportion of units for expected chance agreement

A Kappa coefficient of 90% may be interpreted as 90% better classification than would be expected by random assignment of classes. The general range for Kappa values are if K < 0.4, a poor kappa value; while, if 0.4 < K < 0.75, is a good kappa value and if K > 0.75, it is an excellent kappa value.

4.0 Results and Discussion:

In the present study, various classification methods have been applied on the procured RS image and the land use/cover map are prepared based on the training data statistics (Table 2). The five LULC classes are considered namely, agricultural land, shrubs, water body, wasteland and barren land. The total pixels classified for particular LULC class and the number of pixles found corrected through ground truth along with the total number of reference pixels from reference data set were computed and tabulated as presented in Table 3 for each classifier. By using the formulae (Eqs. 2 and 3) described in methodology for computing the LULC class wise users and producers accuracy for all classifier considered in the present study were computed and tabulated as presented in Table 4.

The user's accuracy is found highest in case of shrubs for all classifier except box classifier due to its distinct features observed in the image. The user's accuracy which is the ratio between the number of correctly classified and the classified total pixels reflects the accurate classification of individual land use/cover class. The shrub is better classified in all classification method, while waste land with scrub giving very poor users accuracy as compared to other LULC classes. This may be due to the reasons that the misclassification of some of training pixels of waste land as agricultural land. Since, users are concerned about what percentage of the classes has been correctly classified, this accuracy should be better. Similarly, producer's accuracy is also reflects the exact classification of particular land use/cover class and the matching of correctly classified pixels by classifier in comparison to ground truth data (reference total). As can be seen from Table 4, this accuracy also gives better results for shrubs except box classifier.

As an established fact that any image classification process will have an error due to various reasons. Classification error in an image is not randomly distributed, but show certain systematic and regularity [10]. The classification accuracy is usually used to denote the precision of the classification result in the RS image classification. The misclassification error is called commission error, which is a measure of overestimation. The off-diagonal elements in each column of confusion matrix are those samples being omitted by the classifier, which committed the misclassification error also called omission error (measure of under estimation). The commission and omission errors are computed by using Eq. 4 and 5 and graphical shown as Fig. 2 and 3. As can be seen that the commission error for mapping agriculture land use is comparatively more for all image classification methods due the reasons that the misclassification of some of training pixels of waste land as agricultural land. From the graphs, it can also be observed that, the omission error is more for waste land with scrubs due to its underestimation of this land use class.

International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

The overall accuracy and kappa values were also computed as presented in Table 5 for each RSI classification methods. The overall accuracy represents the sum of all correctly classified pixels divided by the total number of test pixels, while kappa coefficient provides a difference measurement between the observed agreement of two maps and agreement that is contributed by chance alone. As stated earlier, kappa coefficient above 75% may be interpreted as better classification than would be expected by random assignment of classes [10]. As can be seen from the table that the overall accuracy and kappa coefficient are comparatively more for maximum likelihood classifier. The kappa value of 0.842 for maximum likelihood classifier indicates better classification of land use map. The LULC map using maximum likelihood is shown in Fig. 4. As can been seen from Fig. 4, agriculture land is predominate land use class covering more than 50% of total area depicting the agriculture watershed. The derived LULC map of the study watershed may be used for effective land use planning Also, the LULC coverage of agricuture and waste land could be used for waste land management as well as enhancing agricultral productity to make the Ralegoan Siddhi villages as self-sufficient in all respect and to meet the needs of the growing population.

5.0 Conclusion:

In the present study, the classification accuracy assessment has been performed to evaluate an efficient classification method for deriving the land use land cover map of Ralegaon Sidhi watershed. As can been seen from the results, the maximum likelihood classifier gives better results in terms of overall accuracy of 88.52% and excellent kappa (K=0.842) value. For the individual LULC class, it is observed that the classification accuracy of shrubs is found better for all classifiers except box classifier. The results depicts that the overall accuracy and kappa value are reasonably better for maximum likelihood classifier as compare to box, minimum distance to mean and Mahalanobis classifiers. Hence, it is concluded that maximum likelihood classifier is better classifier for effective LULC mapping. The resulted LULC map and quantitative assessment of agriculture and waste land use could be utilizes for better planning of waste land management and agriculture development schemes.

References

Congalton, R. G. and Green, K., Assessing the Accuracy of Remotely Sensed Data, Principles and practices (Boca Raton, London, New York: Lewis Publishers (1999).

Cihlar, J., Xiao, Q., Chen, J., Beaubien, J., Fung, K. and Latifovic, R., Classification by progressive generalization: a new automated methodology for remote sensing multispectral data. International Journal of Remote Sensing, 19:2685–2704 (1998).

Stehman, S.V. and Czaplewski, R.L., Design and analysis for thematic map accuracy assessment: fundamental principles. Rem. Sens. Environ. 64:331–344 (1998).

Cohen, J.: A coefficient of agreement for nominal scales, Educ. Psychol. Measurement 20(1):37-46 (1960).

Goodman, L. A. and Kruskal, W. H., Measures of Association for Cross Classification. Jour. Amer. Stat. Asso.49: 733-764(1954).

Wilkinson, G.G., A review of current issues in the integration of GIS and remote sensing data. International Journal of Geographical Information Systems, 10:85–101(1996).

Smits, P.C., Dellepiane, S.G. and Schowengerdt, R.A., Quality assessment of image classification algorithms for land-cover mapping: a review and a proposal for a cost based approach. Intern. Jour. Rem. Sens. 20:1461–1486 (1999).

Foody, G.M. and Mathur, A., A relative evaluation of multiclass image classification by support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 42:1336–1343 (2004).

Zheng, Xiaoshen, Liu, Wenling, Qian, Li, and Liang, Xiao, Research on Pattern Recognition Methods of TM Remote Sensing Images, Proc. IEEE, Seventh International Conference on Fuzzy Systems and Knowledge Discovery, doi:978-1-4244-5934-6/10 (2010).

Vieira, Susana M., Kaymak, Uzay and Sousa, Joao M. C., Cohen Kappa Coefficient as a performance Measure for Feature Selection, Pro. IEEE International Conference, doi: 978-1-4244-8126-2/10 (2010).

International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

Table 1. Reference data set

Sr. No.	LULC classes	Total nos. of ground truth pixels
1	Agriculture Land	214
2	shrubs	82
3	Waste land	111
4	Barren land	59
5	Water body	74
	Total	540

Table 2. Sample pixels statistics for various land use classes selected for training

Land use/cover class	Sample (Training) pixels statistics			
	Band	Mean	St. Dev	Total
Agriculture Land	1:	80.5	6.2	522
	2:	76.5	7.1	522
	3:	67.3	7.6	522
	4:	114.7	13.4	522
Shrubs	1:	67.3	3.8	64
	2:	50.1	5.3	64
	3:	105.6	8.8	64
	4:	102.5	6.9	64
Waste land	1:	91.3	2.6	103
	2:	89.8	3.4	103
	3:	79.2	4.1	103
	4:	143.5	8.6	103
Barren land	1:	103.9	4.7	76
	2:	103.9	6.4	76
	3:	92.3	8.5	76
	4:	153.6	8.1	76
Water body	1:	63.6	1.7	71
	2:	53.4	2.1	71
	3:	44.9	2.4	71
	4:	96.2	5.8	71

International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

Table 3. Classification accuracy assessment for various classifiers

Land use/Cover Class	Reference total	Name of the classifiers							
		Box		Minimum Distance		Mahanolobis		Maximum likelihood	
		Classified total	Number correct	Classified total	Number correct	Classified total	Number correct	Classified total	Number correct
Agriculture land	214	208	169	190	155	289	214	268	214
Barren land	59	27	24	57	57	65	56	61	55
Shrubs	82	44	44	82	82	82	82	82	82
Waste land with scrub	111	37	37	79	77	37	36	60	58
water body	74	48	48	132	73	67	67	69	69

Table 4. Classification accuracies for various classifiers

Land use/Cover Class	Name of the classifiers							
	Box		Minimum Distance		Mahalanobis		Maximum likelihood	
	Produce's Accuracy	User's Acuuracy	Produce's Accuracy	User's Accuracy	Produce's Accuracy	User's Accuracy	Produce's Accuracy	User's Acuuracy
Agriculture land	73.16	78.97	81.58	72.43	74.05	100.00	79.85	100.00
Barren land	38.71	40.68	100.00	96.61	86.15	94.92	90.16	93.22
Shrubs	53.66	53.66	100.00	100.00	100.00	100.00	100.00	100.00
Waste land with scrub	40.66	33.33	97.47	69.37	97.30	32.43	96.67	52.25
water body	64.86	64.86	55.30	98.65	100.00	90.54	100.00	93.24

Table 5. Overall classification accuracy and Kappa statistics for various classifiers

Sr. No.	Name of the Classifier	Overall Accuracy (%)	Kappa Coefficient
1	Box Classifier	58.15	0.443
2	Minimum distance classifier	81.85	0.762
3	Mahalanobis classifier	84.26	0.781
4	Maximum likelihood classifier	88.52	0.842

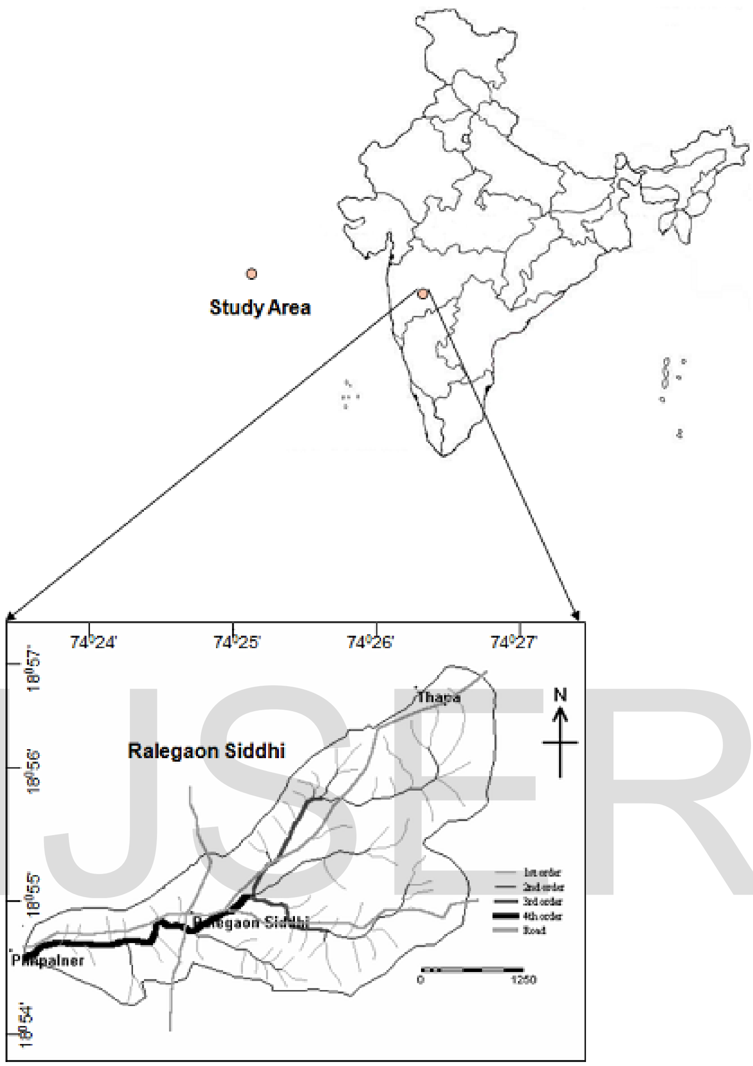

Fig. 1. Location of Study Area

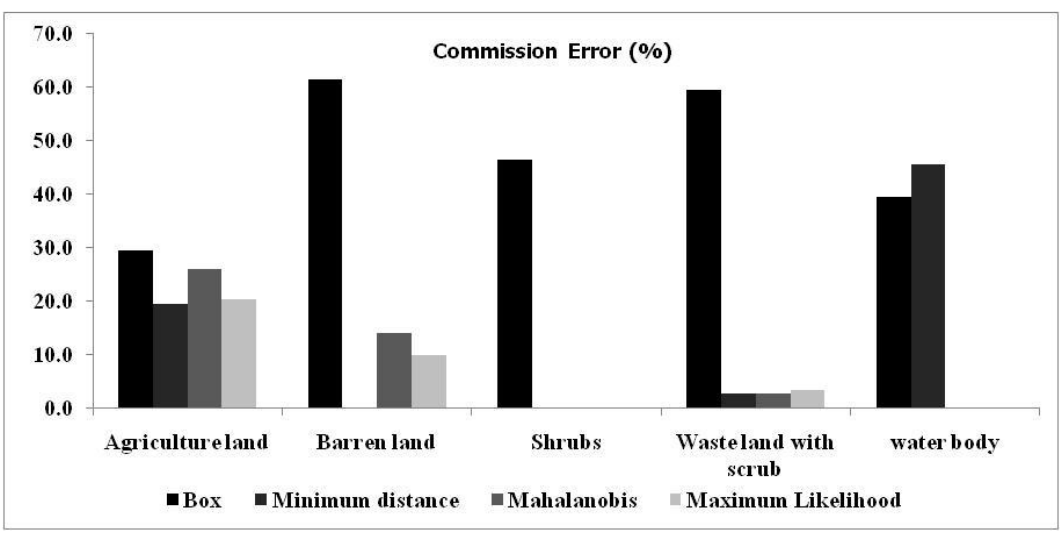

Fig. 2. Various LULC wise commission errors for classifier

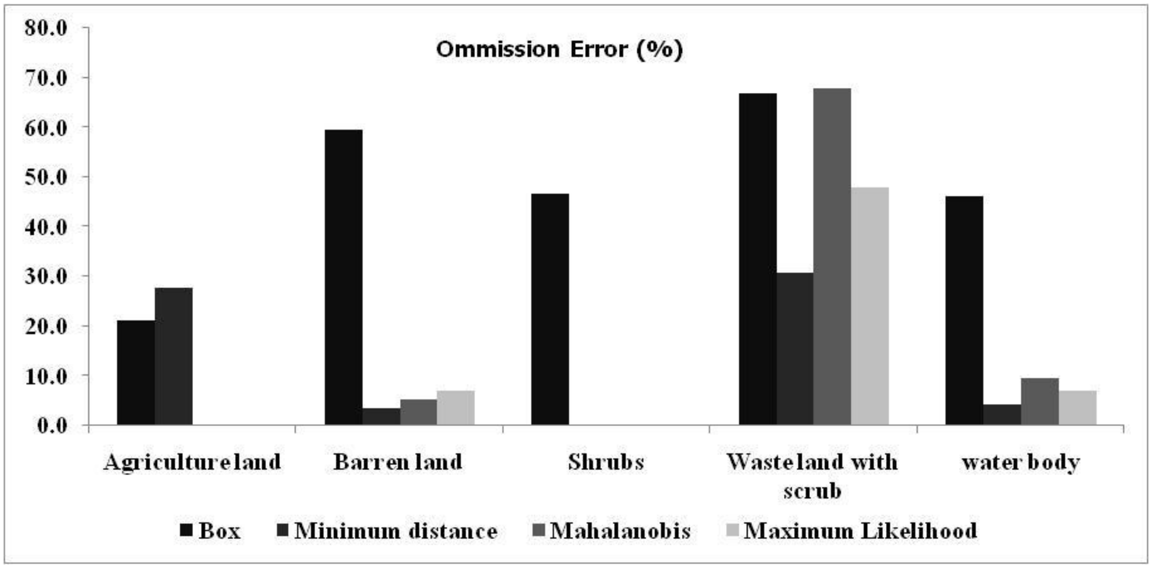

Fig. 3. Various LULC wise Omission errors for classifier

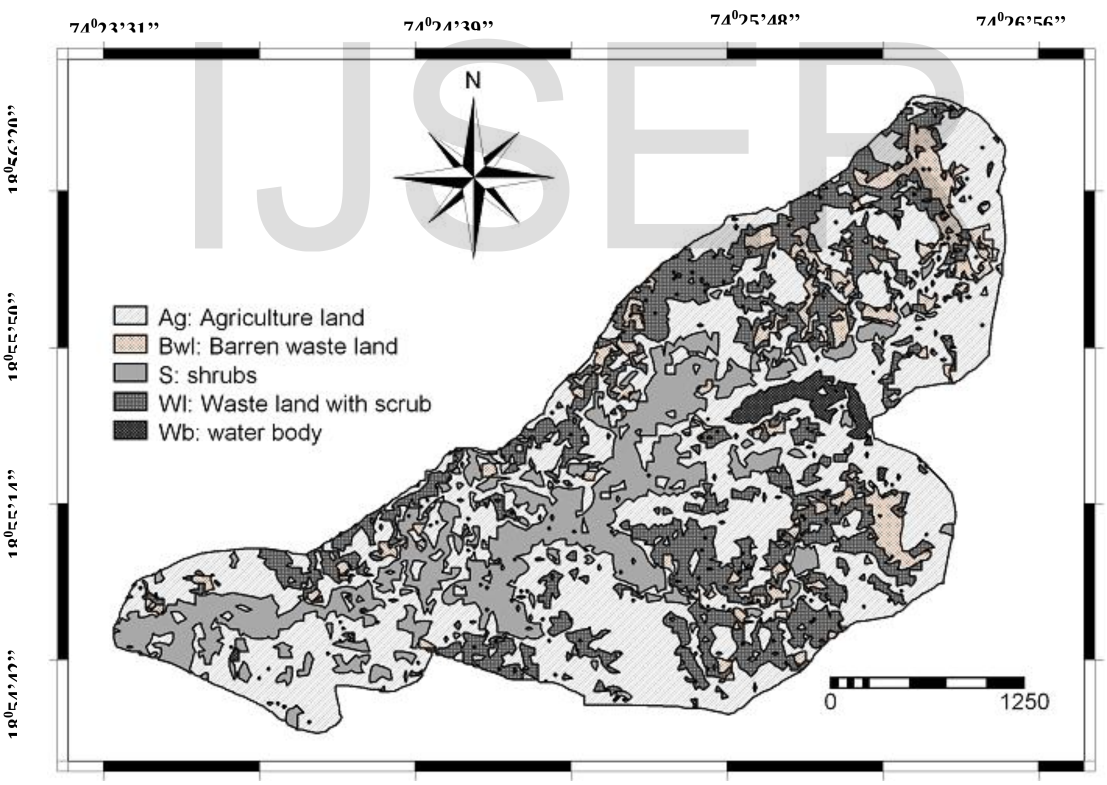

Fig. 4. LULC map using maximum likelihood

Pravada Bharatkar