ISSN : 2248-9622, Vol. 4, Issue 11(Version - 4), November 2014, pp.01-12

RESEARCH ARTICLE OPEN ACCESS

Classification accuracy analyses using Shannon's Entropy

Shashi Poonam Indwar*, Dr.Nilanchal Patel**
*(Scientist B, Groundwater Hydrology Division, NIH Roorkee, Uttrakhand, India;)
** (HOD Department of Remote Sensing, BIT Mesra, Ranchi, Jharkhand, India;)

Abstract.
There are many methods for determining the Classification Accuracy. In this paper significance of Entropy of training signatures in Classification has been shown. Entropy of training signatures of the raw digital image represents the heterogeneity of the brightness values of the pixels in different bands. This implies that an image comprising a homogeneous lu/lc category will be associated with nearly the same reflectance values that would result in the occurrence of a very low entropy value. On the other hand an image characterized by the occurrence of diverse lu/lc categories will consist of largely differing reflectance values due to which the entropy of such image would be relatively high. This concept leads to analyses of classification accuracy. Although Entropy has been used many times in RS and GIS but its use in determination of classification accuracy is new approach.
Keywords: Classification, Entropy, Training signatures, Homogeneous, Heterogeneity

I. Introduction

Accuracy is considered to be the degree of closeness of results to the values accepted as true. Some of the accuracy assessment methods are: the variance analysis, minimum accuracy value used as an index of classification accuracy, spatial error and class attribute errors, a probabilistic approach for change detection and land cover classes are abstraction and generalizations of the real world in order to provide discrete values for continues. Techniques developed for accuracy assessment must take into consideration the factors that are sources of error in image and the methods used for assessing accuracy in a single image and for a pair of images. Assessing the accuracy of change detection products is an important step for the integration of remote sensed data to environmental management system as a decision making support tool. The influence of accuracy and classification performance based on the confusion matrix and derived; overall classification accuracy, producer's accuracy and kappa coefficient in change detection studies, the factors that are influencing the accuracy assessment and the accuracy assessment aspects for change detection and classification, with and without test data and cross-validation methods (**Fl.Zavoianu, A. Caramizoiu , D.Badea).** There are many sources of both conservative and optimistic bias in classification accuracy assessment. The three sources of optimistic bias: use of training data for accuracy assessment, restriction of reference data sampling to homogeneous areas, and sampling of reference data not independent of training data. The magnitude and direction of bias in classification accuracy estimates depends on the methods used for classification and reference data sampling **(T.O. Hammond and D. L. Verbyla).** The main objective of the paper was to assess classification accuracy of classified forest map on Landsat TM data from different number of reference data (200 and 388 reference data). This comparison was made through observation (200 reference data) and interpretation and observation approaches (388 reference data). Five land cover classes namely primary forest, logged over forest, water bodies, bare land and agricultural crop/mixed horticultural can be identified by the differences in spectral wavelength. The result showed that an overall accuracy from 200 reference data was 83.5% with (kappa value 0.7502459, kappa variance 0.002871). However, when 200 references was increased to 388 in the confusion matrix, the accuracy slightly improved from 83.5% to 89.17% with Kappa statistic increased from 0.75022459 to 0.8026135, respectively**(Mohd Hasmadi Ismail and Kamaruzaman Jusoff).** A geostatistical (model-based) framework for spatial accuracy assessment of land-cover classifications was developed. The key component of the proposed framework was its ability to account for spatial or spatiotemporal correlation in observed classification errors, as well as to accommodate different data supports, without relying on probability-based sampling designs. Under this geostatistical framework, confidence intervals were derived for classification accuracy in each class, overall accuracy among all classes and the kappa coefficient **(Phaedon C. Kyriakidis and Jingxiong Zhang).** 48 surface soil samples representing Yazd-Ardakan plain were collected and surface soil salinity was measured. Ten soil samples for investigation of map accuracy were applied. The obtained soil samples and other more ten soil samples which basically had high similarity in spectral reflectance and geomorphological characteristics were used to examine the produced soil salinity map and to assess its accuracy. According to results the produced soil salinity map had an overall accuracy equal to 87% and Kappa index equal to 47% indicating an acceptable accuracy for this classification **(R. Taghizadeh Mehrjardi, Sh. Mahmoodi, M. Taze and E. Sahebjalal)**. Before implementing a classification accuracy assessment, one needs to know the sources of errors **(Congalton and Green 1993, Powell et al. 2004).** 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. In the process of accuracy assessment, it is commonly assumed that the difference between an image classification result and the reference data is due to the classification error. However, in order to provide a reliable report on classification accuracy, non-image classification errors should also be examined, especially when reference data are not obtained from a field survey. A classification accuracy assessment generally includes three basic components: sampling design, response design, and estimation and analysis procedures **(Stehman and Czaplewski 1998).** Selection of a suitable sampling strategy is a critical step **(Congalton 1991).** The major components of a sampling strategy include sampling unit (pixels or polygons), sampling design, and sample size **(Muller et al. 1998).**Possible sampling designs include random, stratified random, systematic, double, and cluster sampling. A detailed description of sampling techniques can be found in previous literature such as **Stehman and Czaplewski (1998)** and **Congalton and Green (1999).**The error matrix approach is the one most widely used in accuracy assessment **(Foody 2002b).** In order to properly generate an error matrix, one must consider the following factors: (1) reference data collection, (2) classification scheme, (3) sampling scheme, (4) spatial autocorrelation, and (5) sample size and sample unit **(Congalton and Plourde 2002).** After generation of an error matrix, other important accuracy assessment elements, such as overall accuracy, omission error, commission error, and kappa coefficient, can be derived. Previous literature has defined the meanings and provided computation methods for these elements (Congalton and Mead 1983, Hudson and Ramm 1987, Congalton 1991, Janssen and van der Wel 1994, Kalkhan et al. 1997, Stehman 1996, 1997, Congalton and Green 1999, Smits et al. 1999, Congalton and Plourde 2002, Foody 2002b, 2004a). Meanwhile, many authors, such as Congalton (1991), Janssen and van der Wel (1994), Smits et al. (1999), and Foody (2002b), have conducted reviews on classification accuracy assessment. They have assessed the status of accuracy assessment of image classification, and discussed relevant issues. **Congalton and Green (1999)** systematically reviewed the concept of basic accuracy assessment and some advanced topics involved in fuzzy-logic and multilayer assessments, and explained principles and practical considerations in designing and conducting accuracy assessment of remote-sensing data. The Kappa coefficient is a measure of overall statistical agreement of an error matrix, which takes non-diagonal elements into account. Kappa analysis is recognized as a powerful method for analysing a single error matrix and for comparing the differences between various error matrices **(Congalton1991, Smits et al. 1999, Foody 2004a).** Modified kappa coefficient and tau coefficient have been developed as improved measures of classification accuracy **(Foody 1992, Ma and Redmond 1995).** Moreover, accuracy assessment based on a normalized error matrix has been conducted, which is regarded as a better presentation than the conventional error **matrix (Congalton 1991, Hardin and Shumway 1997, Stehman 2004).**The error matrix approach is only suitable for 'hard' classification, assuming that the map categories are mutually exclusive and exhaustive and that each location belongs to a single category. This assumption is often violated, especially for classifications with coarse spatial resolution imagery. 'Soft' classifications have been performed to minimize the mixed pixel problem using a fuzzy logic. The traditional error matrix approach is not appropriate for evaluating these soft classification results. Accordingly, many new measures, such as conditional entropy and mutual information **(Finn 1993, Maselli et al. 1994),** fuzzy-set approaches **(Gopal and Woodcock 1994, Binaghi et al. 1999, Woodcock and Gopal 2000),** symmetric index of information closeness **(Foody 1996),** Renyi generalized entropy function **(Ricotta and Avena 2002),** and parametric generalization of Morisita's index **(Ricotta 2004)** have been developed. However, one critical issue in assessing fuzzy classifications is the difficulty of collecting reference data. More research is thus needed to find a suitable approach for evaluating fuzzy classification results. In summary, the error matrix approach is the most common accuracy assessment approach for categorical classes. Uncertainty and confidence analysis of classification results has gained some attention recently **(McIver and Friedl 2001, Liu et al.2004),** and spatially explicit data on mapping confidence are regarded as an important aspect in effectively employing classification results for decision making **(McIver and Friedl 2001, Liu et al. 2004).** In this paper a new concept has been introduced that attempts to establish relationship between homogeneity of training Brightness Values and Classification Accuracy determined. Three sets of training signatures viz. pure, impure and moderately pure for individual lu/lc categories were extracted from the heart or centre, periphery and somewhere between centre and periphery of the respective classes. Subsequently, classification was performed using each of the three sets of training signatures separately. Classification Accuracy was determined from the classified images generated by using the respective training samples. The classification accuracy was examined in relation to the training samples of three types.This is based on the logic that pure training pixels extracted from the centre would comprise more homogeneous Brightness Values resulting in low entropy value while training signatures (impure) extracted from the periphery of a class would comprise relatively heterogeneous Brightness Values with higher entropy value.

Thus entropy of the training signatures could be used as a potential indicator of the purity of the signatures.

1.1 Significance of Entropy of Training Signatures in Classification:

Entropy of the raw digital image represents the heterogeneity of the brightness values of the pixels in different bands. This implies that an image comprising a homogeneous lu/lc category will be associated with nearly the same reflectance values that would result in the occurrence of a very low entropy value. On the other hand an image characterized by the occurrence of diverse lu/lc categories will consist of largely differing reflectance values due to which the entropy of such image would be relatively high. On the classified image, entropy would be determined by the number of the land use and land cover categories present. If an image consists of only one category, its entropy would be zero. Mathematically, entropy expresses the disorder of a system (here, spectral band) which is given by the following formula (O'Neill et al, 1988).

$$H = -\sum pi \log pi$$

Where

H = Entropy of the spectral band of the image
pi = fi / N
fi = frequency
N = total number of pixels.

Entropy of training samples can be used to determine its purity i.e. how homogeneous the training pixels are. Entropy helps in determining the purity of the training samples. Entropy of training data set is related to the accuracy of classification. As classification is one of the major techniques used for mapping of impervious and pervious layers in this thesis, it becomes essential to determine the purity of the training samples. If impurity is high for training data set then it will signify the presence of heterogeneous signatures which will lead to less classification accuracy whereas if impurity is less for training samples it will signify homogeneous signatures resulting in more classification accuracy.

The entropy of the pure training samples will be lower while the impure training samples will exhibit higher entropy values. It is normally expected that for purer training samples a particular lu/lc category will result in high classification accuracy. Therefore, the classification accuracy of any category could be related to the entropy of the training samples of that category. Higher classification accuracy of a particular category will be achieved if the entropy of the training samples for that category is smaller and vice versa. In other words, entropy of the training samples can be considered as a significant indicator of how pure they are, which in turn could direct the analyst to choose more accurate training samples and /or change sampling strategy. In addition, the classification accuracy could also be tested vis-à-vis the entropy of the training samples of different categories although there might be other factors playing vital role in the classification accuracy such as the classification technique involved, type of data used i.e. whether it is raw data or atmospherically corrected data, resolution of the data etc. **(Congalton & Green, 1999)**

1.2 Objectives: The present study has been carried out with the following objectives in mind.

(i) To compute the entropy values of the three different types of the training signatures (viz. pure, moderately pure and impure) based on the purity or homogeneity of the pixels values for the respective four lu/lc categories considered for carrying out the investigation such as Standing water bodies, Forest, Agriculture land and Dense Built-up area.

(ii) To determine the difference in the entropy values of the three different types of signatures for the respective lu/lc category.

(iii) To correlate between the entropy computed for each type of the training signatures (viz. pure, impure and moderately pure) with the classification accuracy obtained by using the corresponding training signature for the respective lu/lc category.

These tasks have been performed for the satellite data of both the years, i.e. 1996 and 2004 in order to revalidate the analysis.

1.3 Data used: IRS 1C LISS-III (105/055) of 22nd December, 1996
IRS P6 LISS-III (105/055) of 20th February, 2004

1.4 Software used: (i) Erdas Imagine 8.5
(ii) Arc view
(iii) Arc-GIS (Arc Map)
(iv) MS- Excel

1.5 Detailed Methodology:

(i) Three sets of training signatures viz. pure, moderately pure, and impure were extracted for four classes i.e. standing water bodies, forest, agricultural lands and built-up using ERDAS 8.5 Software.

(ii) Training samples (aoi) of each class i.e. standing water bodies, forest, agricultural lands and built-up were subset from the main image through Data Preparation-Subset Image, using group icon of aoi.

(iii) Frequency and total number of training pixels band wise were noted from the subset images of each class via the raster attribute table.

(iv) From raster attribute editor, the following command was executed: "select all and copy frequency and total number of pixels in Microsoft excel to calculate pi and log pi".

(v) Put the values of pi and log pi in the formula of H = -∑ pi*logpi.

(vi) Represent H values in tables and graphs to show the relation between impurity and classification accuracy.

(vii) Classification of images of each year based on the three types of training signatures i.e. pure, moderately pure and impure samples.

(viii) Determination of Producer's accuracy based on the different types of training signatures.

(ix) Generation of the (a) Plot between purity of training signatures and entropy (b) Plot between entropy and Producer's accuracy (c) Plot between purity of training signatures and Producer's accuracy and comparative analysis between the two years.

(x) Comparative analysis and discussion of results.

1.5.1 Methodology Flowchart:

Selection of pure, moderately pure and impure data set for the four lu/lc classes respectively from the images of both years i.e. 1996 & 2004 respectively

Subset the training data sets selected from the main image (FCC)

Note frequency and total number of pixels band wise from attribute table of training data sets for each class

Calculate pi and log pi in Microsoft excel

Put the values of pi & log pi in formula of entropy (H) = -∑pi * log pi and compute H

Classification of images of each year based on the three types of training signatures i.e. pure, moderately pure and impure samples

Determination of Producer's Accuracy (PA) based on the different types of training signatures

Generation of the (i) plot between purity of training signatures and entropy (ii) plot between entropy and PA (iii) plot between purity of training signatures and PA and comparative analysis between the two years

Comparative Analysis and Discussion of Results

1.6 Procedure/ Strategy of Sampling of the Training Signatures:

In the present study four lu/lc categories were selected for performing the task of determining the relation between entropy and classification accuracy based on the purity of the signatures viz. standing water bodies, forest, agricultural land and built-up area. The standing water bodies comprise a large reservoir and a lake. The forest is characterized by three crown densities such as dense, moderate and open; the built-up area mainly comprises the highly dense congested conglomerate of the residential and shopping areas located at the centre of the city; and the agricultural land comprises the vast stretch of fallow land without standing crop.

For each of these categories, training samples corresponding to the pure, impure and moderately pure type were extracted from the centre, periphery and somewhere between the centre and periphery respectively taking into consideration the homogeneity of the brightness values at these three locations. For example, the training pixels are expected to be more homogeneous towards the centre of any lu/lc and become more heterogeneous towards the periphery. Similar criteria for the selection of training samples have been employed on the images of both the years i.e. 1996 and 2004. The purpose of performing the entropy-training samples homogeneity-classification accuracy relationship analyses in two different years is primarily to revalidate the findings from this study keeping in view the lu/lc dynamics and its impact on the pixel homogeneity within the different categories. The standard Maximum Likelihood classifier has been employed to classify the satellite images of both the years' i.e. 1996 and 2004 using the three types of training signatures viz. pure, impure and moderately pure selected for the four different lu/lc categories in the present study.

Table 1.1 ENTROPY (H) values of the three types of signatures for different lu/lc categories

(a) Impure Training Signature:

Lu/Lc	Standing water bodies		Forest		Agricultural Lands		Bt-up	
Year	1996	2004	1996	2004	1996	2004	1996	2004
Band 1(G)	0.27643	0.88737	0.53009	1.06534	1.01555	1.14234	1.09333	1.28063
Band 2(R)	0.27643	0.82047	1.06758	1.13094	1.08071	1.15497	0.97657	1.24588
Band 3(NIR)	0.47712	0.72832	0.98863	1.15132	1.00041	1.04199	1.11772	1.23226
Band 4(MIR)	0.47712	0.82047	1.2408	1.18921	1.00041	1.10889	1.21037	1.27702

(b) Moderately Pure Training Signature:

Lu/Lc	Standing water bodies		Forest		Agricultural Lands		Bt-up	
Year	1996	2004	1996	2004	1996	2004	1996	2004
Band 1 (G)	0.30102	0.73644	1.05843	1.09636	0.97882	0.93191	1.09333	1.2455
Band 2 (R)	0.30102	0.87958	1.09252	1.09636	0.92865	0.87718	0.97657	1.22924
Band 3 (NIR)	0.30102	0.93979	1.07547	1.10833	0.7005	0.89197	1.11772	1.17287
Band 4 (MIR)	0.30102	0.81937	1.1679	1.10833	0.85954	1.04137	1.21037	1.17287

(c) Pure Training Signature:

Lu/Lc	Standing water bodies		Forest		Agricultural Lands		Bt-up	
Year	1996	2004	1996	2004	1996	2004	1996	2004
Band 1 (G)	0.30102	0.30102	0.74707	1.04199	0.71496	0.77814	1.00092	36
Band 2 (R)	0.30102	0.30102	0.82916	1.02117	0.82785	0.77814	0.9149	1.0557
Band 3 (NIR)	0.30102	0.30102	0.91662	0.81603	0.82785	0.6778	1.10317	0.85499
Band 4 (MIR)	0.30102	0.30102	0.9678	1.10889	0.90312	0.6778	1.10317	1.04056

1.7 Results and Discussion:

1.7.1 Entropy values of the three types of training signatures

1.7.1.1 Standing water bodies

Examination of the figures 1.1 a (i) and (ii) reveal the following observations.

(i) The entropy values decrease as the purity of the training signature increases.

(ii) Comparative analysis of the entropy values between the years 1996 and 2004 reveals that the entropy values are considerably lower in the year 1996 in the respective categories of training signatures as compared to 2004.

(iii) The entropy values for the three different categories of training signatures are nearly same in 1996 while in the year 2004 there occur considerable difference in the entropy between impure and pure training signatures.

(iv) The pure training signatures are associated with very low entropy values. This signifies that there occurs maximum homogeneity in the brightness values at the centre of water bodies from where pure training samples are selected while the pixels homogeneity drastically decreases as one move away from the centre towards the periphery.

(v) There occurs no systematic variation of the entropy values among the different spectral bands.

(vi) The occurrence of the first observation could be attributed to the prevalence of homogeneous condition within the water body in the year 1996 as compared to 2004.

1.7.1.2 Forest

Comparative analyses of the figures 1.1 b (i) and (ii) indicate the following.

(i) There occurs little variation in the entropy values between the two years; however, the entropy values decrease as the purity of the training signatures increase in both the years which could be attributed to the occurrence of pixel homogeneity towards the central portion of the forest class as it is observed in the field while the pixel homogeneity decreases as one moves towards the periphery where the canopy density decreases giving rise to the prevalence of mixed or impure training signatures.

(ii) Another significant observation that is apparent from the analysis of the figures is the occurrence of highest entropy in the MIR band which could be attributed to the high spectral reflectance characteristics of vegetation.

1.7.1.3 Agriculture

Agricultural lands exhibit nearly the similar observation as that of the forest. Entropy increases as the impurity of the training signature increases (Figures 1.1 c (i) and (ii).

1.7.1.4 Built-up areas

In both the years i.e. 1996 and 2004, the entropy values for the impure and moderately pure training signatures are nearly same with the former category possessing slightly more entropy value than the latter one. However, the pure training signatures are associated with significantly lower entropy values (Figure 1.1 d (i) and (ii)).

Comparison among the entropy values of the respective training signatures viz. pure, impure and moderately pure among the four lu/lc categories in the individual years indicates that the water bodies possess the smallest values while the remaining three lu/lc categories viz. forest, agricultural land and built-up are associated with nearly the same entropy values. This observation is attributed to the fact that water bodies in the study area are characterized by significantly large amount of homogeneity as compared to the other classes. As such the agricultural land in the study area comprises the fallow land with no standing crop thereby resulting in larger variation in the brightness values.

1.7.2 Correlation between Entropy and Producer's Accuracy

Figures1.2 (a-d) show the relationship between the Producer's accuracy and combined entropy of the three different types of training signatures viz. pure, impure and moderate for the four lu/lc categories in the respective years, i.e. 1996 and 2004. Analyses of the figures reveal that the classification accuracy decreases as the entropy or the impurity increases. There occurs little or no variation in the classification accuracy obtained by considering the pure and moderately pure training signatures while the classification accuracy drastically decreases for the impure samples.

1.7.3 Comparison among the Producer's Accuracy of Different Types of Training Signatures

Figures1.3 (a-d) shows together the bar chart of the producer's accuracy determined for the lu/lc categories by considering the three types of training signatures for 1996 and 2004. Comparative analysis of the classification accuracy bar charts of the lu/lc categories reveals the following.

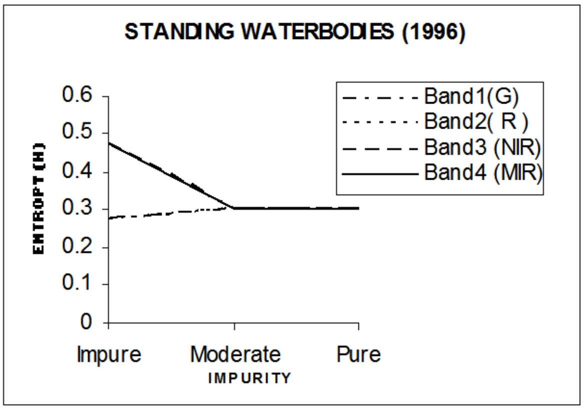
(i) First, there occurs highest classification accuracy for the pure samples with a decreasing trend towards the impure samples in both the years.

(ii) Second, the classification accuracy associated with the impure samples is appreciably lower in the year 2004 as compared to 1996 that may be attributed to the occurrence of a large amount of heterogeneity in the spectral characteristics of different lu/lc classes (towards their periphery) in the year 2004.

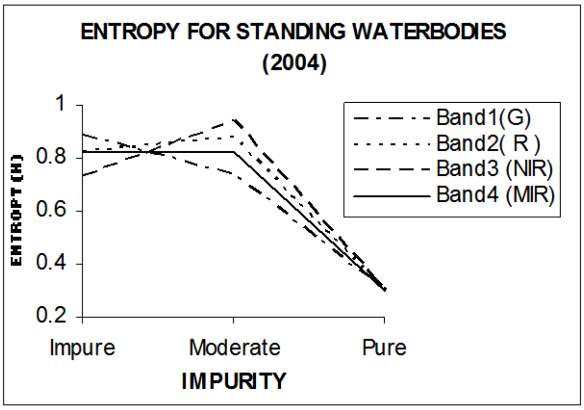
1.8 Conclusion:

The following conclusion can be drawn from the analyses carried out in this chapter.

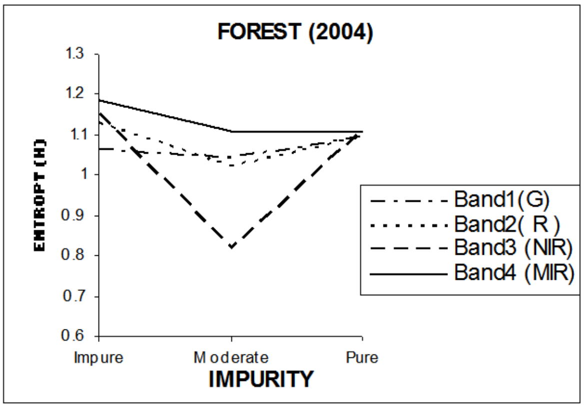
1. The entropy of the training signatures is strongly related to their purity. Pure training signatures give rise to low entropy that is characteristic of the homogeneity of brightness values whereas impure training samples are associated with large entropy resulted due to heterogeneous pattern of brightness values. In other words, entropy of the training samples could be used as a potential indicator of the purity of the training samples.
2. There also occurs significant relationship between producer's accuracy and the entropy of the training samples. Producer's accuracy is found to be higher for pure training samples characterized by the low entropy values whereas moderately pure and impure training samples associated with larger entropy values lead to lower producer's accuracy.
3. Water bodies as expected are associated with larger homogeneity of brightness values with lower entropy values as compared to the other lu/lc categories.



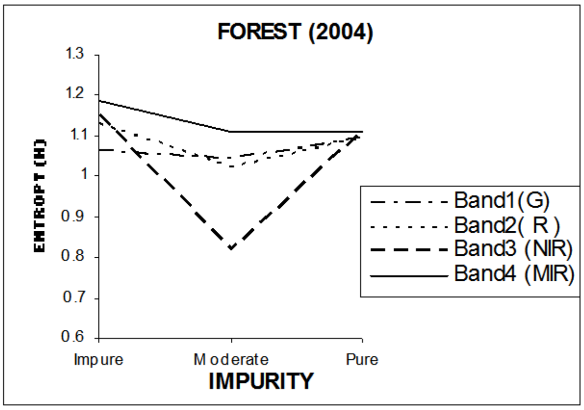
a(i)



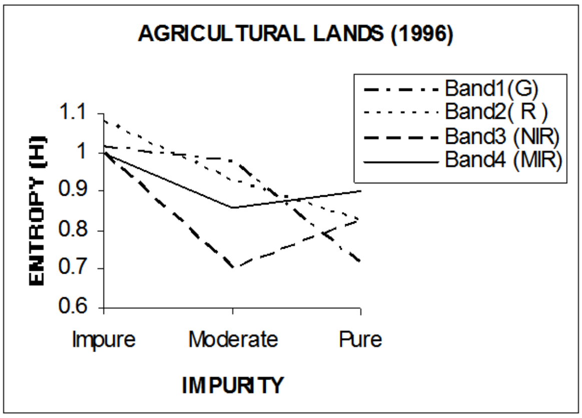
a(ii)



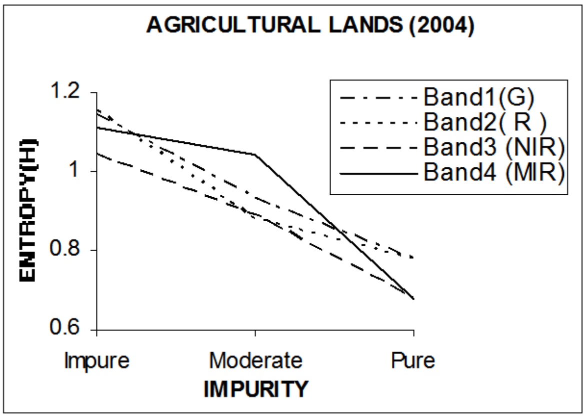
b (i)



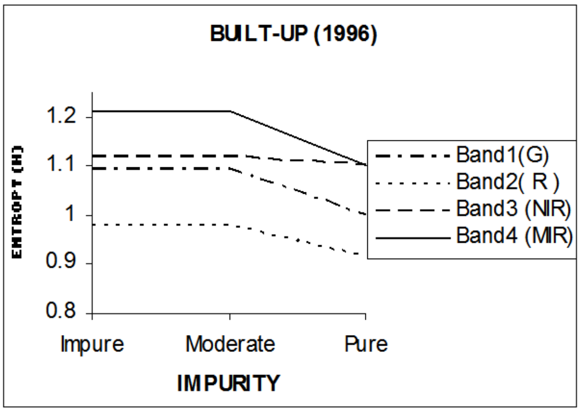
b(ii)



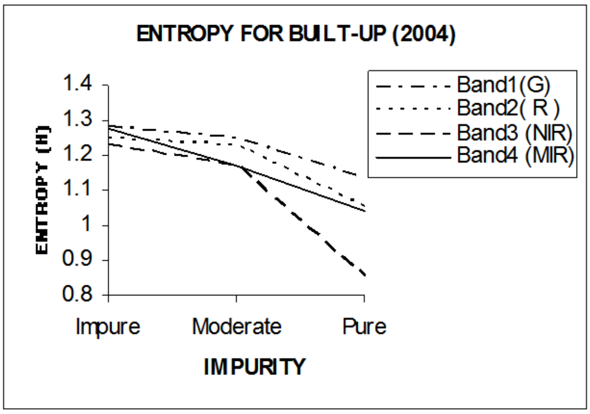
c (i)



c(ii)

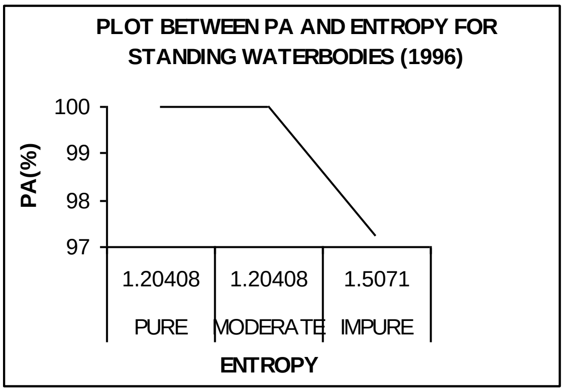


d (i)

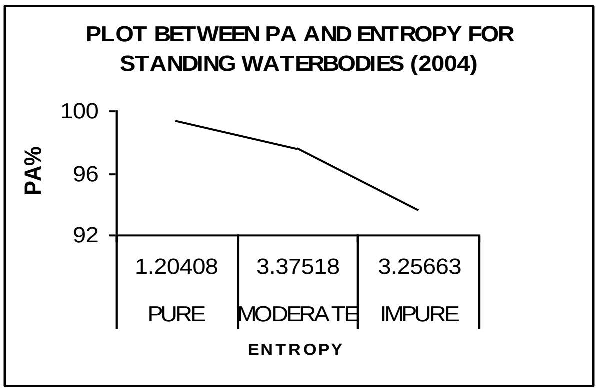


d(ii)

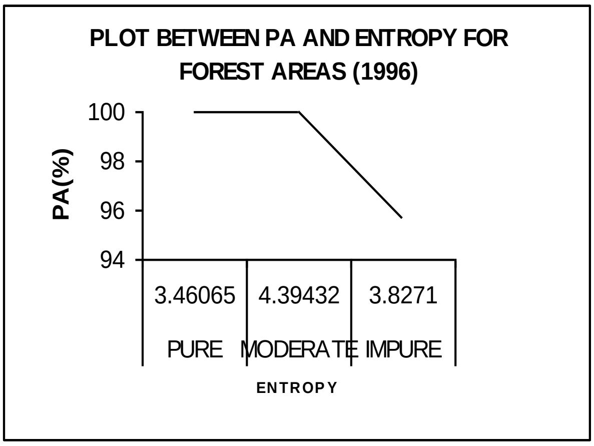
Figure 1.1 Entropy Values of the Three Signature Types for Different Lu/Lc.



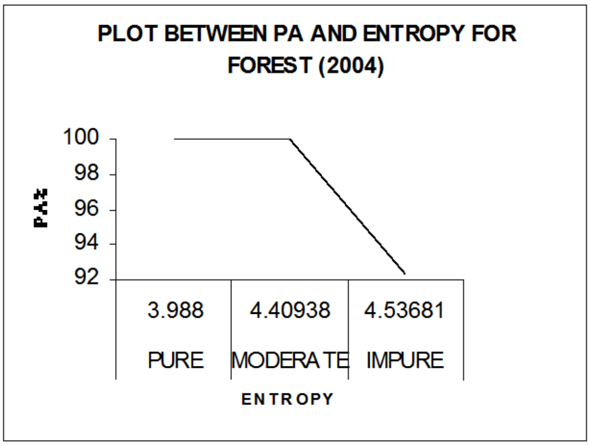
a (i)



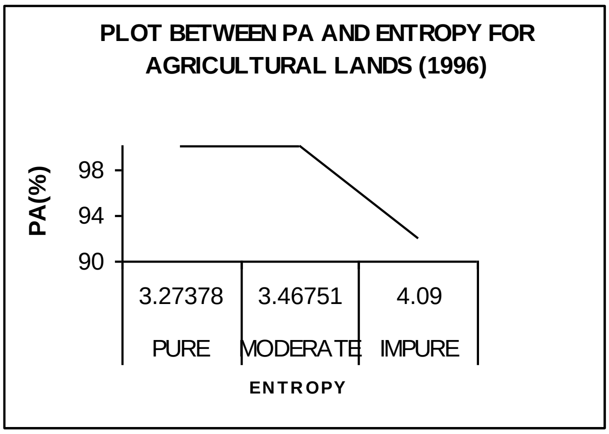
a(ii)



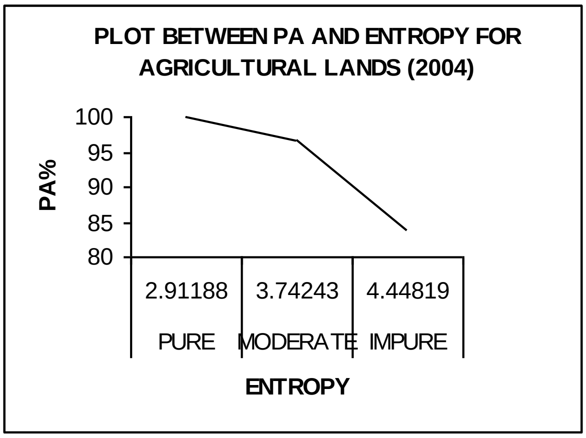
b (i)



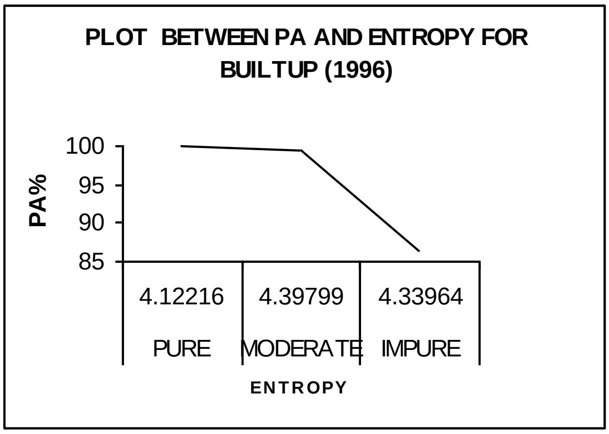
b(ii)

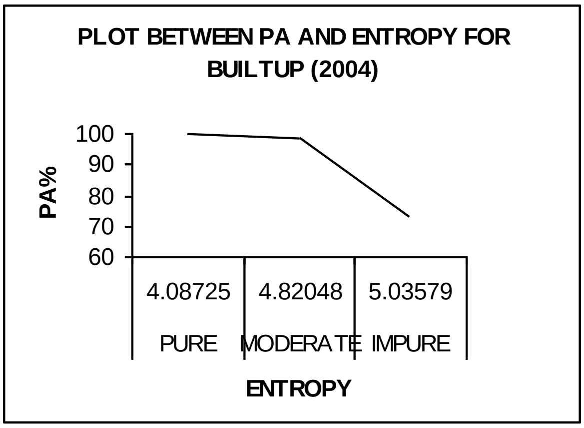


c (i)



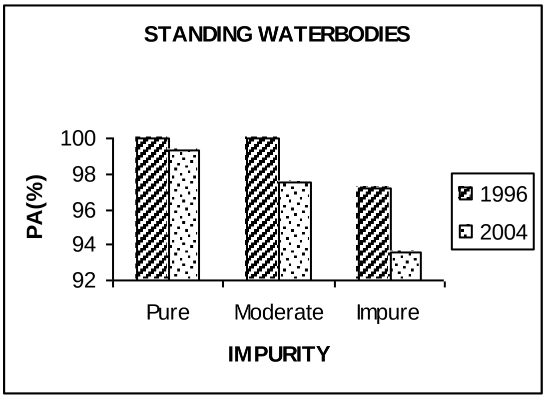
c(ii)

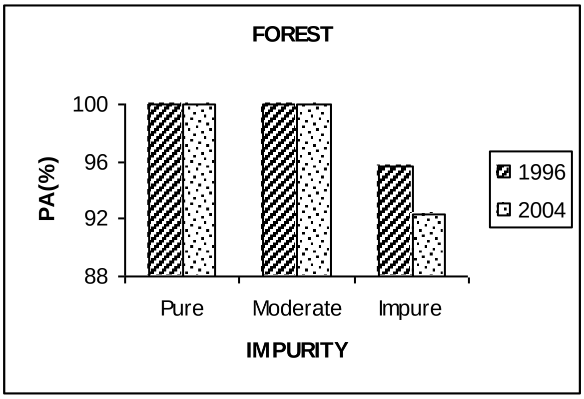




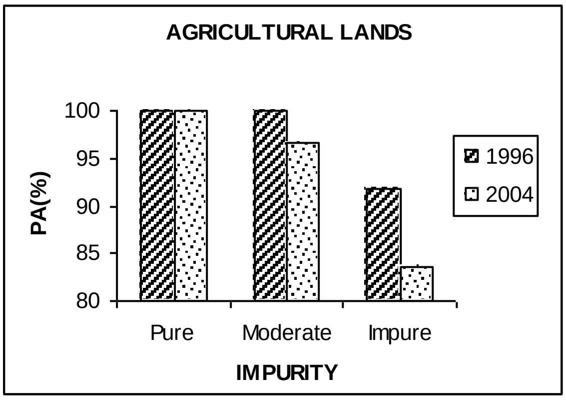
d (i) **d(ii)**

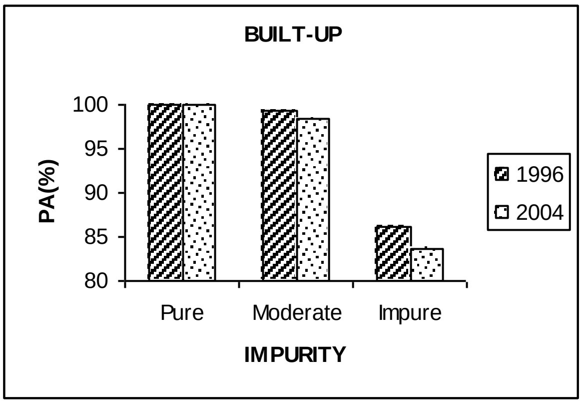
Figure 1.2 Correlation Plot between Producer's Accuracy (PA) and Entropy of Different Lu/Lc.





(a) **(b)**





(c) **(d)**

Figure 1.3 Correlation among Pure, Impure and Moderately Pure Training Signatures and Producer's Accuracy (PA) of Different Lu/Lc.

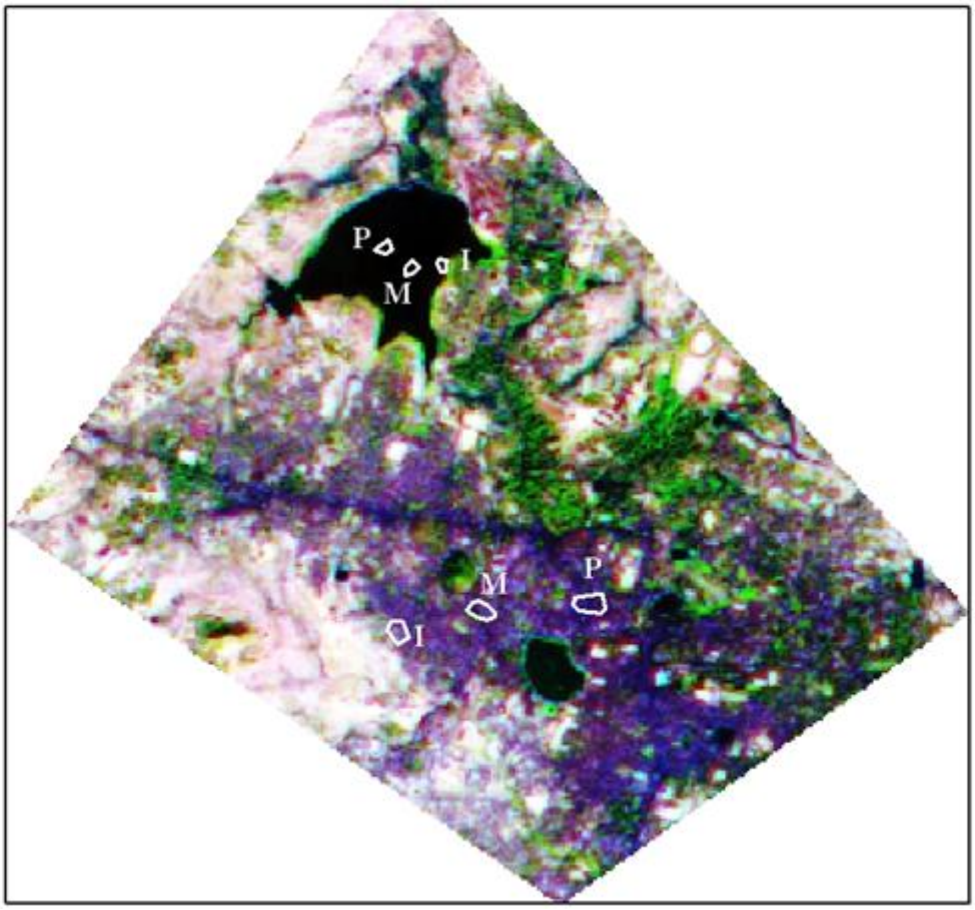


Figure 1.1 FCC of IRS 1C LISS III Data of December 1996 with Training Signatures of Different Purity Superimposed.

References

[1.] Binaghi**,** E.**,** Brivio**,** P.A.**,** Ghezzi**,** P**.** and Rampini**,** A., A fuzzy set accuracy assessment of soft classification. *Pattern Recognition Letters*, **1999***,* 20, pp. 935–948.

[2.] Congalton, R.G., A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment*, **1991**, 37, pp. 35–46.

[3.] Congalton, R.G. and Green, K., A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering and Remote Sensing*, **1993**, 59, pp.641–644.

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

[5.] Congalton, R.G. and Mead, R.A., A quantitative method to test for consistency and correctness in photo interpretation. *Photogrammetric Engineering and Remote Sensing*, **1983,** 49, pp. 69–74.

[6.] Congalton, R.G. and Plourde, L., Quality assurance and accuracy assessment of information derived from remotely sensed data. In J. Bossler (Ed.), *Manual of Geospatial Science and Technology* (London: Taylor & Francis), **2002,** pp. 349–361.

[7.] Finn, J.T., Use of the average mutual information index in evaluating classification error and consistency. *International Journal of Geographical Information Systems*, **1993,** 7, pp. 349–366.

[8.] Foody, G.M., On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing*, **1992,** 58, pp.1459–1460.

[9.] Foody, G.M., Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. *International Journal of Remote Sensing*, **1996,** 17, pp. 1317–1340.

[10.] Foody, G.M., Hard and soft classifications by a neural network with a nonexhaustively defined set of classes. *International Journal of Remote Sensing*, **2002a,** 23, pp.3853–3864.

[11.] Foody, G.M., Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, **2002b,** 80, pp. 185–201.

[12.] Foody, G.M., Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **2004a**, 70, pp. 627–633.

[13.] Gopal, S. and Woodcock, C., Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, **1994**, 60, pp. 181–188.

[14.] Hudson, W.D. and Ramm, C.W., Correct formulation of the Kappa coefficient of agreement. *Photogrammetric Engineering and Remote Sensing*, **1987**, 53, pp. 421–422.

[15.] Janssen, L.F.J. and Van Der wel, F.J.M., Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering and Remote Sensing*, **1994,** 60, pp.419–426.

[16.] Kalkhan, M.A., Reich, R.M. and Czaplewski, R.L., Variance estimates and confidence intervals for the Kappa measure of classification accuracy. *Canadian Journal of Remote Sensing*, **1997,** 23, pp. 210–216.

[17.] Ma, Z. and Redmond, R.L., Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing*, **1995**, 61, pp. 435–439.

[18.] Muller, S.V., Walker, D.A., Nelson, F.E., Auerbach, N.A.,Bockheim, J.G., Guyer, S. and Sherba, D., Accuracy assessment of a land-cover map of the Kuparuk river 862 D. Lu and Q. Weng Downloaded By: [Indiana University Libraries] At: 12:37 19 March 2007 basin, Alaska: considerations for remote regions. *Photogrammetric Engineering and Remote Sensing*, **1998**, 64, pp. 619–628.

[19.] R.Taghizadeh Mehrjardi, 2SH. Mahmoodi, 3M.Taze and 4E. Sahebjalal, Accuracy Assessment of Soil Salinity Map in Yazd-Ardakan Plain, Central Iran, Based on Landsat ETM+ Imagery. *American-Eurasian J. Agric. & Environ. Sci*., **2008**, 3 (5), pp 708-712, 2008 ISSN 1818-6769 © IDOSI Publications, 2008.

[20.] Stehman, S.V. and R.L.Czaplewski. Design and analysis of thematic map accuracy assessment: fundamental principles. *Remote Sensing of Environment*, **1998**, 64, pp 331-344.