JOURNAL OF ENVIRONMENTAL DESIGN (JED)

A Journal of the Faculty of Environmental Studies, University of Uyo, Uyo, Nigeria
Vol. 6, No. 2, Nov. 2011

TABLE OF CONTENT

EDITORIAL STYLE GUIDE FOR AUTHORS

Please submit your manuscript double spaced in MS Word 2003 - 2007 version. Provide tables and figures in a separate file (or files) in MS Excel MS Word. or EPS format. Maps should be supplied in EPS format.

Because manuscripts will undergo a blind review, submit two title pages. the first showing the title of the manuscript. author name, title. affiliation, telephone number. e-mail address. and the date of the manuscript. The second title page should contain only the title of the paper.

Third-person style is always preferred. If appropriate, authors may make limited use of first-person singular, but a single author should not refer to himself or herself as "we."

Biography: The manuscript should include, on a separate page or the "first" title page described above. a sentence listing each author's name, and affiliation.

Abstract. Include a one-paragraph abstract *not exceeding* 150 *words* and place it on the first page of the text. The abstract, describe the issue(s) or question(s) the paper addresses and state the major findings, conclusions and recommendations.

Keywords: To help users reference the JED published research, keywords are included with journal articles. Please suggest two keywords for your manuscript.

Abbreviations: The definition of an abbreviation or acronym is given the first time it appears; afterward, only the abbreviation is used. However, an abbreviation that is defined in the abstract should also be defined in the article. An abbreviation that appears only once in an article should be deleted and the full wording used.

If an abbreviation is first defined in the text, the abbreviation alone can then be used in subsequent footnotes or tables; however, if the abbreviation is first defined in a footnote or table, the abbreviation should be defined again when it first appears in the following text.

Text Headings: Headings are not numbered and are placed to the left. First-level headings are bold; second-level headings are italic; and third-level headings are italic with a period that leads directly into text.
Example:

First- levelheading
Second-levelheading
Third-levelheading. The text continues …

Tables and Figures: Use arabic numerals to number tables and figures consecutively in separate series in order of appearance. Include a brief descriptive title at the top of each. Tables and figures should be in separate page not integrated into the text. The text must contain a reference to each table not integrated into the text. The text must contain a reference to each table or figure.

Equations: Make sure that all symbols in equations are clear and that all equations (except those in footnotes) are numbered. Single-letter variables should be italicised. Multiple-letter variables and abbreviations (e.g., AGE) and functions (e.g., expo min. In) should not be italicised: neither should numbers. Parentheses, or mathematical operations. Vectors and matrices should be in bold (not italicized).

References: The manuscript must include complete and accurate citations of all materials referenced in the manuscript that are not of your original authorship. Please double-check your references to ensure that names and date are accurate, that Web pages are still active, and that there are no discrepancies between the text and the reference list. The APA style is strongly recommended.

SUPERVISED MAXIMUM LIKELIHOOD CLASSIFIER FOR MULTI-SPECTRAL SATELLITE REMOTE SENSING IMAGE CLASSIFICATION

Onuwa Okwuashi[1] Mfon Isong[1], Aniekan Eyoh[1], Etim Eyo[2], and Aniekan D. Ekpo[3]
onuwaokwuashi@yahoo.com.

[1] Department of Geoinformatics & Surveying, University of Uyo. [2]School of Civil Engineering & Geosciences, Newcastle University, Uk. [3]Amana Consortium Engineers Limited, Uyo.

ABSTRACT

This experiment is based on a multi-spectral Landsat 7 ETM imagery of Porirua, New Zealand. MATLAB is used for implementing the maximum likelihood classifier based computer program, while geographic information systems are used for data preparation and visualisation. The satellite image consists of three land use classes (water, undeveloped, and developed). The entire study has a dimension of 250 x 250 pixels that equal 62,500 pixels. Training of the maximum likelihood classifier is required since maximum likelihood classifier is a supervised classifier. Three hundred points that consist hundred points each from the three land use classes are selected as training set using the stratified random sampling technique. The training of the classifier requires that the standard deviation of the training set is first computed; second the means of each of the three land use classes to be classified are computed; and third, is the computation of the maximum likelihood probabilities of the three land use classes. The land use class with the highest maximum likelihood probability wins the classification for that pixel. All the training points are correctly classified by the maximum likelihood classifier. The test data constitute the remaining points not included in the training set. Some of the pixels in the test data are wrongly classified. The experiment is validated using Kappa statistic, based on the information from a confusion matrix derived by comparing the predicted with the reference data. The calculated Kappa statistic is 0.7161 which indicates substantial agreement between the predicted and the reference data.

INTRODUCTION

The main product of remote sensing image processing is classification map. Earlier classification applications involve the use of hard classifiers such as minimum distance to means and the box classifiers (Peddle, Foody, Zhang, Franklin, & LeDrew, 1994; Rogan, Franklin, & Roberts, 2002; Li, Chen, & Su, 2003; Mahesh & Mather, 2003). Recently soft classification algorithms such as Artificial Neural Network (ANN) (Benediktsson, Palmason, & Sveinsson, 2005; Del Frate, Pacifici, Schiavon, & Solimini, 2007), K Nearest Neighbour (KNN), and Maximum Likelihood Classifier (MLC) have become part of the mainstream classification algorithms. The objective of this research therefore is to illustrate how the MLC algorithm can be applied to multi-class problems in satellite remote sensing image classification.

MAXIMUM LIKELIHOOD CLASSIFIER

The maximum likelihood principle is illustrated in an example with a one-dimensional data distribution $\{x_i\}$, $i = 1,\ldots,n$. We assume that the data originate from a Gaussian distribution $p(x)$ with parameters σ and μ (Hoffmann, 2005),

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right). \quad (1)$$

According to the maximum likelihood principle, the unknown parameters are chosen such that the given data are most likely under the obtained distribution. The probability L of the given data set is (Hoffmann, 2005),

$$L(\sigma,\mu) = \prod_{i=1}^{n} p(x_i) = \left(\frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp\left(-\frac{\sum_{i=1}^{n} (x_i-\mu)^2}{2\sigma^2} \right) \quad (2)$$

We want to find $\hat{\sigma}$ and $\hat{\mu}$ that maximise L. Maximising L is equivalent to maximising log L, which is also called the log-likelihood ℓ,

$$\ell(\sigma,\mu) = \log L(\sigma,\mu) = -n\log\sigma - \frac{\sum (x_i-\mu)^2}{2\sigma^2} + const \quad (3)$$

To find the maximum we compute the derivatives of the log-likelihood ℓ and set them to zero:

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_i (x_i - \mu)^2}{\sigma^3} = 0, \tag{4}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{\sum_i (x_i - \mu)^2}{\sigma^2} = 0. \tag{5}$$

Thus, we obtain the values of the parameters $\hat{\sigma}$ and $\hat{\mu}$:

$$\hat{\sigma}^2 = \frac{\sum_i (x_i - \hat{\mu})^2}{n}, \tag{6}$$

$$\hat{\mu} = \frac{\sum_i x_i}{n}. \tag{7}$$

The resulting $\hat{\sigma}^2$ is the variance of the distribution and $\hat{\mu}$ is its centre. The extremum of ℓ is indeed a local maximum, as can be seen by computing the Hesse matrix of ℓ and evaluating it at the extreme point ($\hat{\sigma}$, $\hat{\mu}$) (Hoffmann, 2005):

$$H_\ell = \begin{vmatrix} \frac{\partial^2 \ell}{\partial \sigma^2} & \frac{\partial^2 \ell}{\partial \sigma \partial \mu} \\ \frac{\partial^2 \ell}{\partial \mu \partial \sigma} & \frac{\partial^2 \ell}{\partial \mu^2} \end{vmatrix}, \tag{8}$$

$$\left.\frac{\partial^2 \ell}{\partial \sigma^2}\right|_{\sigma=\hat{\sigma},\mu=\hat{\mu}} = \frac{n}{\hat{\sigma}^2} - \frac{3\sum (x_i - \hat{\mu})^2}{\hat{\sigma}^4} = \frac{n}{\hat{\sigma}^2} - \frac{3n\hat{\sigma}^2}{\hat{\sigma}^4} = -\frac{2n}{\hat{\sigma}^2}, \tag{9}$$

$$\left.\frac{\partial^2 \ell}{\partial \sigma \partial \mu}\right|_{\sigma=\hat{\sigma},\mu=\hat{\mu}} = \left.\frac{\partial^2 \ell}{\partial \mu \partial \sigma}\right|_{\sigma=\hat{\sigma},\mu=\hat{\mu}} = \frac{2\sum_i (x_i - \hat{\mu})}{\sigma^3} = 0, \tag{10}$$

$$\left.\frac{\partial^2 \ell}{\partial \mu^2}\right|_{\sigma=\hat{\sigma},\mu=\hat{\mu}} = -\frac{n}{\hat{\sigma}^2}. \tag{11}$$

It follows that the Hesse matrix at the extremum is negative definite,

$$H_\ell\big|_{\sigma=\hat{\sigma},\mu=\hat{\mu}} = \begin{vmatrix} -\frac{2n}{\hat{\sigma}^2} & 0 \\ 0 & -\frac{n}{\hat{\sigma}^2} \end{vmatrix}. \tag{12}$$

Therefore, the extremum is a local maximum. Moreover, it is also a global maximum. First, for finite parameters, no other extrema exist because ℓ is a smooth function. Second, ℓ is positive for finite parameters, but approaches zero for infinite values. Thus, any maximum must be in the finite range (Hoffmann, 2005).

APPLICATION

The experiment was based on a multi-spectral Landsat 7 ETM imagery of Porirua, New Zealand, acquired in 2006 (see Figure 1). The Landsat image consists of seven spectral bands, and has a cell size of 25m x 25m. The original satellite data were first reviewed in GIS (ArcGIS software), and all seven spectral bands were extracted using the *layer properties tool* and visualised in MATLAB (see Figure 1). Before importing the data into MATLAB, they were first converted from raster to ASCII data using the ArcGIS *conversion tool*. MATLAB cannot read raster files; hence the data must be in ASCII format for onward processing in MATLAB. In MATLAB the final study area was extracted from the original satellite image. The final image used for the classification was 250 x 250 pixels, which amounts to 625,000 pixels. All the seven spectral bands were used for the classification. The satellite image consists of three distinct land use classes: water, undeveloped, developed cells. The aim of this experiment therefore is to classify the satellite image into these three land use classes.

The first step in supervised classification is the selection of training sample. Good results are obtained when the training sample is non-redundant, relatively concise, and randomly selected (Lo & Yeung, 2007). Selection of the training sample was based on the stratified random sampling. Three hundred points were selected for training, while the remaining pixels were used to test the accuracy of the classification. The 300 points consisted 100 points each from the three land use classes (water, undeveloped, and developed). The classification was based on the MATLAB function *[IDX,prob,STD,mean_WAT,mean_UNDEV,mean_DEV] MLC(train,test,m1,m2,m3,m)*. The right hand side of the equation represents the known input parameters, while the left hand side of the equation represents the unknown output parameters. *Train* is the training set; *test* is the test data; *m1, m2, m3* are the number of training points that belong to water, undeveloped, developed cells respectively; *m* is the size of the test data; *mean_WAT,mean_UNDEV,mean_DEV* are the calculated means of water, undeveloped, and developed cells from the training set; *STD* is the calculated standard deviation of the training sample; *prob* is the calculated maximum likelihood probability of the test data; and *IDX* represents the classification indices for the three land use classes (water=1; undeveloped=2; developed=3). The computed standard deviation and means for the three land use classes are given in Tables 1-4; while some of the computed maximum likelihood probabilities for the training and test data are given in Tables 5 and 6. The land use class with the highest maximum likelihood probability wins the classification for that pixel (see Tables 5-6).The training accuracy was 100%, since all the training points were correctly classified. Not all the test data were correctly classified. The classification results were visualised in ArcGIS (see Figure 2).

CONCLUSION

The result of the classification experiment displayed in Figure 2 was validated with the Kappa statistic (Cohen, 1960).). Kappa statistic can be expressed mathematically as,

$$k = \frac{P_o - P_c}{1 - P_c} \tag{13}$$

Where,

$$P_o = \sum_{i=1}^{m} P_{ii} = \frac{1}{N}\sum_{i=1}^{m} n_{ii} \tag{14}$$

and,

$$P_c = \sum_{i=1}^{m} P_{i+}P_{+i} = \frac{1}{N^2}\sum_{1=1}^{m} n_{i+}n_{+i} , \tag{15}$$

(Ma & Redmond, 1995; Lo & Yeung, 2007).

Where,

P_o = proportion agreement observed

P_c = proportion agreement expected by chance

n_{ii} = the total number of correctly classified points by class along the diagonal of the error matrix

N = the total number of points checked (sampled)

P_{ii} = the proportion of correctly classified sample points by class at the diagonal of the error matrix (i.e. n_{ii} / N)

P_{i+} = the marginal distribution of the sample data (n_{i+} / N where n_{i+} is the row sum by class)

P_{+i} = the marginal distribution of the reference data (n_{+i} / N where n_{+i} is the column sum of class)

m = the total number of classes

The Kappa statistic is more reliable than other validation techniques because it has the ability to evaluate the actual agreement and chance agreement (Fung & LeDrew, 1988). Kappa statistic is computed from an error matrix or confusion matrix resulting from the comparison of the reference with the predicted data (see Table 7). A cell-by-cell comparison between the reference and the predicted data are displayed in the confusion matrix given in Table 7. The calculated Kappa statistic using equations 13, 14, and 15 was 0.7161.

According to Landis and Koch (1977) the calculated Kappa result can be appraised based on the interpretation given in Table 8. Using Table 8 the computed Kappa statistic implied that the predicted data has a substantial agreement with the reference data.

CONCLUSION

The MLC is one of the earliest parametric soft classifiers based on simple algorithms derived through basic algebraic methods. The computed Kappa statistic from the experiment showed that the predicted data have substantial agreement with the reference data. The result of this experiment shows that the MLC remains invaluable in satellite remote sensing image classification despite the paradigm shift from parametric to nonparametric algorithms in contemporary image classification modelling.

REFERENCES

Benediktsson, J. A. Palmason, J. A., & Sveinsson, J. R. (2005). Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens., 43*, 480–490.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Del Frate, F, Pacifici, F, Schiavon, G., & Solimini, D. (2007). Use of neural networks for automatic classification from highresolution images," *IEEE Trans. Geosci. Remote Sens., 45* (4), 800–809.

Fung, T. & LeDrew, E. (1988). The Determination of optimal threshold levels for changes detection using various accuracy indices. *Photogrammetric Engineering and Remote Sensing, 54*(10), 1449-1454.

Hoffmann, H. (2005). *Unsupervised Learning of Visuomotor Associations* (unpublished doctoral dissertation). Universität Bielefeld, Technische Fakultät Logos Verlag, Berlin.

Landis, J. R. & Koch, G. G (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*,159 – 174

Li, T. S., Chen, C. Y., & Su, C. T. (2003). Comparison of neural and statistical algorithms for supervised classification of multi-dimensional data. *Int. J. Indus. Eng.—Theory Appl. Pract., 10*, 73–81.

Lo, C. P. & Yeung, A. K. W. (Eds.) (2007). *Concepts And Techniques Of Geographic Information Systems* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Ma, Z. & Redmond, R. L. (1995). Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing 61*, 435-439.

Mahesh P. & Mather, P. M. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ., 86*, 554–565.

Peddle D. R., Foody, G. M., Zhang, A., Franklin, S. E., & LeDrew, E. F. (1994). Multisource image classification II: An empirical comparison of evidential reasoning and neural network approaches. *Canadian Journal of Remote Sensing, 20*, 396–407.

Rogan, J., Franklin, J., & Roberts, D. A. (2002). A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sens. Environ., 80*, 143–156.

Table 1: *Training result: Computed standard deviation*

Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
35.2085	20.6571	34.4880	45.2743	68.6160	7.4275	43.2041

Table 2: *Training result: Computed mean for water cells*

Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
82.1000	27.2000	20.1000	8.7000	5.1000	108.3000	0.4000

Table 3: *Training result: Computed mean for undeveloped cells*

Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
69.4000	25.0000	23.9000	94.9000	68.7000	112.7000	18.0000

Table 4: *Training result: Computed mean for developed cells*

Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
135.1000	61.2000	83.2000	93.2000	150.2000	123.2000	85.4000

Table 5: *Training result: Some computed maximum likelihood probabilities for 300 points*

(Water=1; Undeveloped=2; Developed=3)

Pixels	Water	Undeveloped	Developed	MLC	Reference data	Remark
1	0.0246	0.0461	0.1082	3	3	Correct
2	0.0035	0.0171	0.0642	3	3	,,
3	0.0213	0.0339	0.1079	3	3	,,
4	0.0408	0.0585	0.0781	3	3	,,
5	0.0416	0.0589	0.1007	3	3	,,
- - -	- - -	- - -	- - -	- - -	- - -	- - - - -
101	0.1121	0.0972	0.0185	1	1	,,
102	0.1128	0.0970	0.0217	1	1	,,
103	0.1129	0.0947	0.0199	1	1	,,
104	0.1127	0.0973	0.0218	1	1	,,
105	0.1120	0.0912	0.0202	1	1	,,
- - -	- - -	- - -	- - -	- - -	- - -	- - -
296	0.0989	0.1098	0.0298	2	2	,,
297	0.096	0.1126	0.0374	2	2	,,
298	0.1022	0.1096	0.0297	2	2	,,
299	0.0896	0.1114	0.0453	2	2	,,
300	0.0631	0.0794	0.0723	2	2	,,

Table 6: *Test result: Some computed maximum likelihood probabilities for 62,500 pixels*
(Water=1; Undeveloped=2; Developed=3)

Pixels	Water	Undeveloped	Developed	MLC	Reference data	Remark
1	0.1111	0.0912	0.0147	1	1	Correct
2	0.1111	0.0910	0.0148	1	1	Correct
3	0.1111	0.0908	0.0148	1	1	Correct
4	0.1111	0.0910	0.0148	1	1	Correct
5	0.1112	0.0913	0.0148	1	1	Correct
- - -	- - -	- - -	- - -	- - -	- - -	- - -
16996	0.0551	0.0748	0.0859	3	2	Wrong
16997	0.0558	0.0760	0.0850	3	3	Correct
16998	0.0608	0.0789	0.0794	3	2	Wrong
16999	0.0638	0.0794	0.0755	2	2	Correct
17000	0.0619	0.0788	0.0775	2	2	Correct
- - -	- - -	- - -	- - -	- - -	- - -	- - -
62496	0.0890	0.1105	0.0446	2	2	Correct
62497	0.0861	0.1096	0.0470	2	2	Correct
62498	0.0870	0.1104	0.0475	2	2	Correct
62499	0.0862	0.1099	0.0481	2	2	Correct
62500	0.0840	0.1075	0.0517	2	2	Correct

Table 7: *Computed confusion matrix for MLC modelling*

	REFERENCE DATA		
	Developed	Undeveloped	Water
PREDICTED DATA			
Developed	10039	189	0
Undeveloped	7143	36629	24
Water	105	1578	6793

Table 8: *Interpretation of kappa statistic*

KAPPA	INTERPRETATION
< 0	No agreement
0.0 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

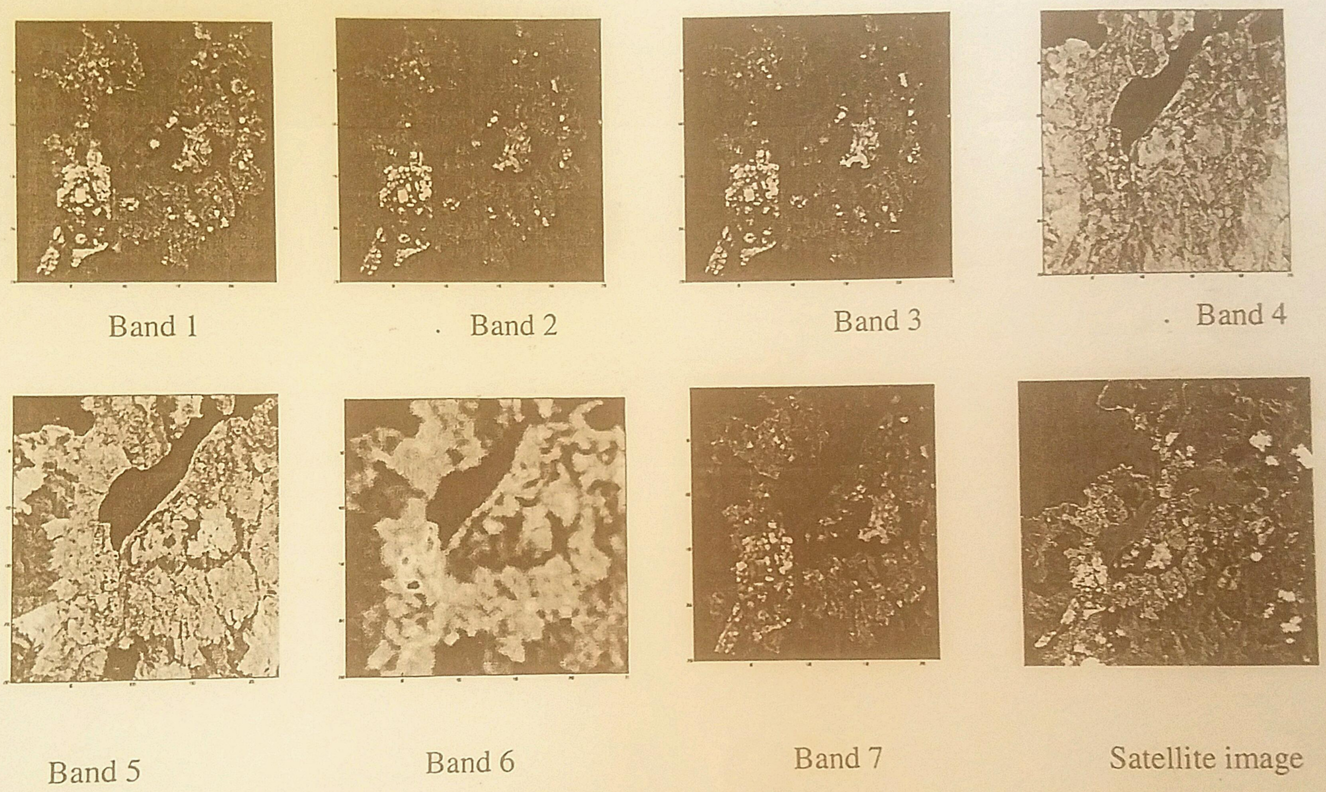

Figure 1: *Extracted spectral bands 1 - 7 and original Landsat image of Porirua, New Zealand*

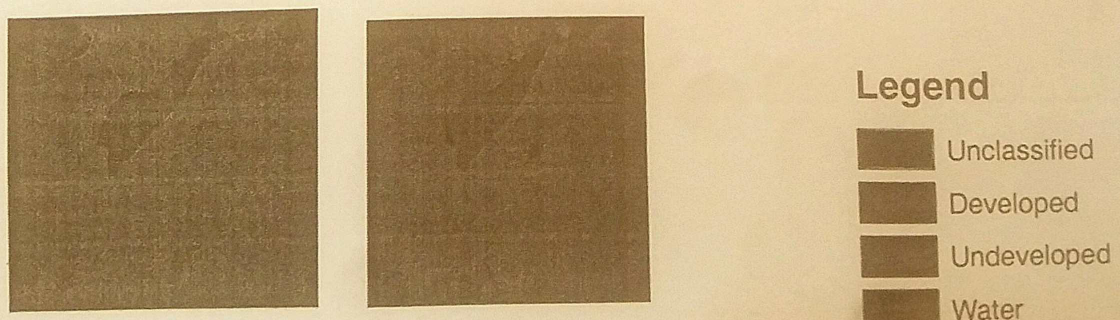

Figure 2: *MLC classification result*

Reference data MLC