IOSR Journal of Humanities and Social Science (IOSR-JHSS)
Volume 24, Issue 3, Ser. 3 (March. 2019) 77-84
e-ISSN: 2279-0837, p-ISSN: 2279-0845.
www.iosrjournals.org

Accuracy Assessment of Satellite Image Classification for Land Cover and Land Use Using MultiSpec as a Tool

*Asikunnaby[1] and Irtifa Alam Nabila[2]
[1,2] (Lecturer, Department of Disaster and Human Security Management, Bangladesh University of Professionals, Bangladesh)
**Corresponding Author: Asikunnaby*

Abstract : Image classification works as a significant tool that revealsLand Use andLand Cover (LULC)and helps in policy formulationsthat are related to human-environment relationship. The broad objective of this study is to clarify an image classification system of LULC which can be explained easily with scientific justification. This study used MultiSpec software as a tool to find out a more accurate image analysis method than traditional processes used such as ArcGIS and the data source for the task was taken from free satellite images database of the United States Geological Survey (USGS). In the image analysis of the study, Maximum likelihood classification method was adopted for a robust statistical basis and simplistic interpretation. Confusion Matrix is a key focus of this research as the table shows accuracy level, Kappa Statistic, actual and predicted classifications. In this study the confusion matrix gives 95.9% overall accuracy, Kappa Statistic of 94% and Kappa Variance of 0.000043 which means a good statistical and scientifically reliable result for image classification. Kappa coefficient for this study is 0.94 which validates a good agreement as in the scale for Kappa a highly reliable range is in between 0.80 to 1.00. The results of this study show a possibility of a simple but well-built image analysis procedure that can be followed in related research areas.

Keywords: Image Analysis,Confusion Matrix, Maximum Likelihood, Supervised Classification

---

Date of Submission: 25-02-2019 Date of acceptance: 11-03-2019

I. "INTRODUCTION"

Human beings are influencing land use since an early age.Among many concerns of land use alteration,urban land use has received a lot of attention lately.Land use and land cover change have started as a worldwide phenomenon of academic study and it is probably the most important environmental phenomenon in recent decades.Rapid urbanization and industrialization, agricultural practices and significant changes in human actions have been recognized as the main drivers of change for land cover and land use patterns worldwide.Drastic changes in land cover and land use that once tookhundreds of years now occur within two or three decades [1].In recent times, the growing apprehension for managing natural resources has been caused by growing population pressure and related human actionswhich have resulted in severe environmental disturbance and loss of ecological diversity.The effects of major changes in land use have increased to threatening extentsOver the period of last 300 years [2].

Due to its great coverage and repetitive observations, space-based remote sensing is an acceptable and powerfulapproach for land cover investigation [3].Remote Sensing (RS) techniques are now used to produce base maps for the analysis of the urban area;RS is primarily based on advanced sensor data from space-borne frameworks.This is in fact due to reliable sources of high spatial data, a wide range of multi-spectral and multi-temporal sources, advanced statistical and geo-spatial approaches and compatibility with multi-spectral and GIS data sources and approaches. However, researchers realize that a pixel having the spectral information is ablend of several urban surfaces, even with the smallest pixel size.Therefore, the objectives of sub-pixel classification methods are to measure the surface mixture statistically to enhance the total accuracy of the classification [4].Although there are pixel differences, there is also noteworthy evidence that adjacent pixel groups have alike spectral information and therefore belong to the similar classification category.Before classification, object-oriented techniques are required to develop group pixels based on spectral similarity and spatial proximity. The accurateness of categorization using object-based methods shows considerable success and potential for various applications related to urban[5].In the classification process, geo-spatial approaches for built-up area mapping also uses adjacent pixels, such as object-oriented methods that spot the significance of spatial proximity.

In recent decades, Bangladesh has experienced rapid urban growth that has caused arable land loss, habitat destruction, decline of wetlands and natural vegetation cover [7].Typically, urban areas are a mixed patchwork of land cover and land use that is difficult to classify in certain classes using remote sensing data.The derivation of classification methods for urban landscape features has been adapted in conjunction with increasing spatial, spectral and temporal resolutions of digital classification systems for remote sensing instruments. Moreover,this development of classification techniques does not showthat one method is useful than another. The use of a clearly defined algorithm for the categorization of urban land cover and land use depends on the user's objectives and the level of detail, frequency and sensors essential for the expected or resulting outputs; Like the data from satellite remote sensing which are used for analysis. In order to analyze the spatial and temporal dynamics of land use and land cover, new GIS and remote sensing tools and techniques are powerful and cost-effective [8][9][10].Remote sensing offers useful multi-temporal data on land use methods and changes in land cover where GIS is a suitable tool to map and analyze these forms [11].

This study will attempt to identify land use and land cover of the Dhaka Metropolitan Police Area (DMPA) with MultiSpec and ArcGIS software as tools and analyze their accuracy along withsignificance for statistical appropriations from Landsat 8 imagery of 2018.

II. “IMAGE CLASSIFICATION”

The main purpose of the classification method is to classify eachpixel in a digital image into multiple land cover “classes” or “themes”. These categorized data can be used to map the land cover of the image for a specific study area. Multi-spectral data are normally used for classification purposes and the spectral pattern in the data for each pixel is used as a categorization mathematical basis [12].The purpose of the image classification is to identify and depict the characteristics of an image as a unique gray color in terms of the object or type of land cover by these characteristics on the ground.

Possiblythe image classification is the utmost significant part of digital image analysis. It is good to have a “pretty picture” or an image that shows a degree of colors that illustrate the different characteristics of the basic landscape, but if one does not know what these colors mean, it is useless[13]. Supervised classification and Unsupervised classification are the two key classification methods of image analysis.

In the supervised classification, researcherslabeled the examples of classes (i.e. type of land cover) of interest in the image as “training sites”. The image processing software classification is then applied to develop a characterization of statistical reflection for individual information group. This phase is sometimes referred as “signature analysis” and can involve the development of a characterization that is as simple as the mean or ranges of reflection on each band or as difficult as detailed analyzes of the mean, variance and covariance across every bands. When a statistical characterization has been attained for each information class, the image is categorized by analyzing the reflection for individual pixel [14].

Unsupervised classification is a technique that analyzes many unknown pixels and divides them into several classifications based on natural subclassesof image value. In contrast to the supervised classification, the unsupervised classification does not need the analyst’s training data.The elementary principle is that values in the measuring space in a given type of land cover should be close to each other (e.g. similar gray levels), while data should be relatively segregated (e.g. very different gray levels)in different classes [12] [13] [14].

Maximum likelihood Classification is a principle of statistical approach to help classify overlapping signatures where pixels are assigned to the maximum probability class. The maximum probability classifier is assumed to provide more preciseoutcomes than the parallelepiped classification, but additional calculations makes it slow.The authorsplaced the term“accurate” in quotes as it perceives that the input data classes have a Gaussian distribution and the signatures have been selected in a better way; this is not a standard assumption for all time.

III. “STUDY AREA”

Dhaka Metropolitan Police Area (DMPA) has been selected as the defining boundary for this research. The study area is located geographically between $23^{0}40^{/}00^{//}$ North Latitude to $23^{0}55^{/}00^{//}$ North Latitude and $90^{0}20^{/}00^{//}$ East Longitude to $90^{0}30^{/}00^{//}$ East Longitude (Fig.1). Dhaka City is on the bank of the river Buriganga.

Dhaka municipality was established in 1864 and in 1960 it became a town committee. The Town Committee's name changed in 1972 to Dhaka Municipality. In 1983, it became a municipal corporation. Finally, it was elevated to City Corporation in 1991. Dhaka Metropolitan Police Area has a region of 339.18 square Kilometer with a population of 8906039 inhabitants and density of 62844 people per sq. km [15]. The city was separated into two City Corporation areas in 2011: Dhaka North consists of 36 wards and Dhaka South consists of 56 wards. This is the country’s most developed city and its pattern of land use is completely urbanized and heterogeneous.

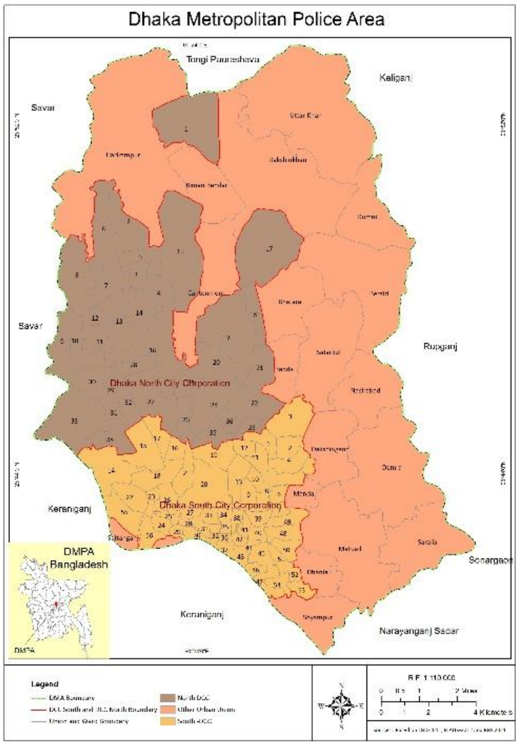


Figure 1: Study area map (Source: Authors)

IV. "DATA AND METHODS"

4.1 Selection of RS data and Image Pre-processing

Geo-corrected Landsat 8 images (year 2018, 25 November) collected from the United States Geological Survey website (USGS) has been used to identify urban land use and land cover classification. Landsat 8 satellite images are available in multi-spectral bands that allow satellite images to be better identified by processing them in a few steps.Landsat 8 images comprise 11 spectral bands having a 30m spatial resolution.

The satellite image metadata contains wrs path 137, wrs row 44, false easting 500000.0, false northing 0.0, central meridian 93.0, scale factor 0.9996, origin latitude 0.0, linear unit meter (1.0), etc.

The falsecolor composition was Layer 5, Layer 4 and Layer 2 after ArcGIS 10.3 converted individual bands into raster composite. The composite image was projected in wgs_1984_utm_zone_46. For further analysis in MultiSpec software, the processed image was clipped with mask tool in GIS for DMPA area.

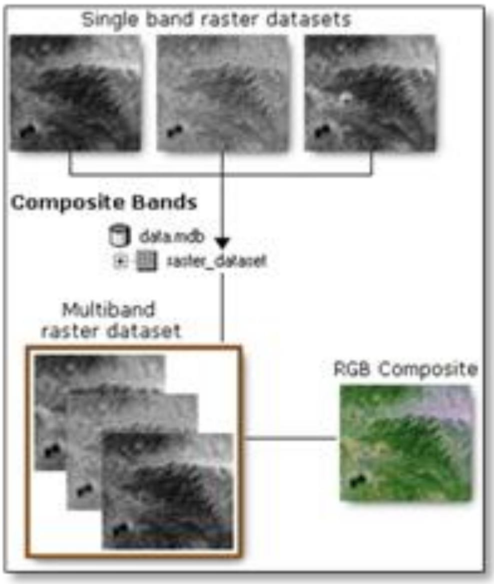


Figure 2: Image pre-processing (Source: ESRI 2015)

4.2 Image classification

The study used MultiSpec software for supervised Landsat imagery classification developed by Purdue University to analyze multispectral and hyperspectral imagery for land use and land cover (LULC) analysis.It is a processing scheme for interactive investigation of multispectral image data from Earth observations.The next phase of the classification was the selection of six training fields and four test fields.

Training field means the selection of training sites from known areas for specified classes. This study followed the maximum likelihood procedure for image classification.The maximum probability classification assumes that each band's statistics are ordinarily distributed and the probability that a particular pixelgoes to a particular calculated class.All pixels are classified if one does not select a probability threshold and each pixel is assigned to the highest probability class. If the maximum probability is lower than the specified threshold, the pixel is not classified [16].Instead, the maximum probability decision rule is based on training class multispectral distance measurements.For image classification, the classifier uses the following rules:

The likelihood Lk is well-defined as the subsequent probability of a class k pixel.

$$\mathbf{Lk = P(k/X) = P(k)*P(X/k) / P(i)*P(X/i)}$$

Where, P(k) = prior probability of class k

P(X/k) = Conditional probability to detect X from class k, or probability density function. Usually **P(k)** are expected to be equal and **P(i)*P(X/i)** is also common to all classes. Therefore, **Lk** depends on **P(X/k)** or the probability density function.

The "Training Class Performance (Re-substitution Method)" table arranges the pixels of each field and class that were classified. This table is known as the Confusion Matrix, where classes with accuracy level and Kappa statistics are represented.

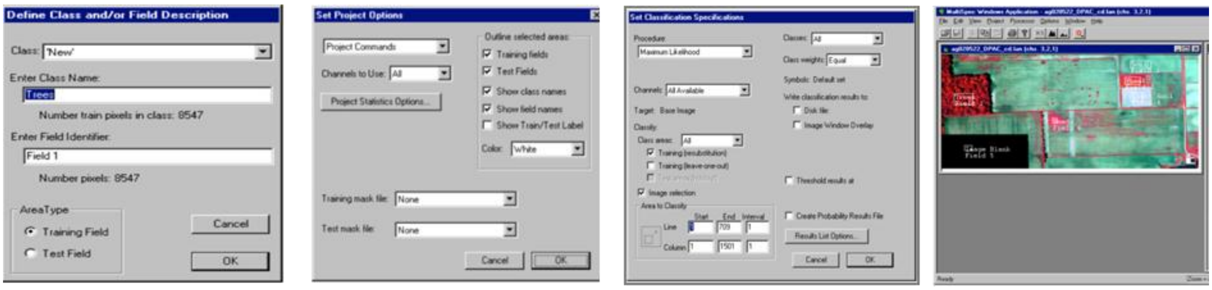


Figure 3: Selection of training fields and classification procedure (Source: MultiSpec)

The results of the image classification are compared to additional ground information in this confusion matrix. It identifies the nature and quantity of the classification errors and is the strength of a confusion matrix.
The validation classes, or "actual" values are sited along the x axis and the "predicted" classes, or classified land cover and land use are positioned along the y axis (Fig. 4). The matrix shows "the accuracy of the user", "the accuracy of the manufacturer" and "the overall accuracy".A comprehensive image analysis scenario is found on the basis of the change between the real agreement in the error matrix and the chance agreement showed in the overall row and column [17].

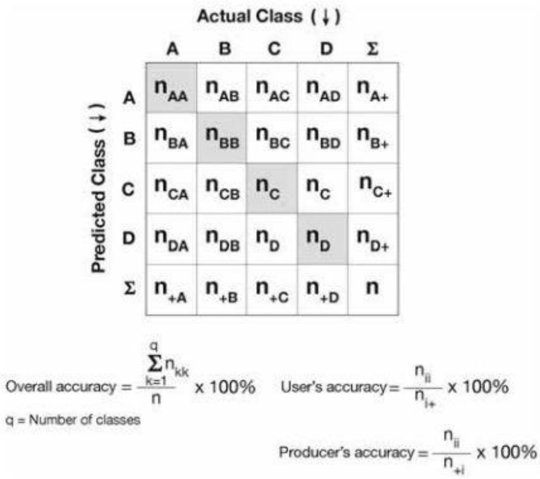


Figure 4: Layout of a confusion matrix (Source: Congalton, 2008)

V. "RESULTS AND DISCUSSION"

For this research, the authors followed Anderson's land use and land cover categorizationmethod and adapted level 1 and level 2 in a further generalized manner for understanding the case of the metropolitan area of Dhaka [18].The outputs of MultiSpec software from supervised classification using the maximum likelihoodclassification provided the following classes.The five classes of the supervised image include Open Space, Vegetation, Water Body, Building and Road.The Confusion Matrix table or popularly referred to as the Error Matrix table provides key image classification information and enables the researchers to verify its accuracy.

Table 1: Confusion Matrix of Image Classification

Class Name	Producer's Accuracy (%)	Number of Samples	Number of Samples in each class				
			Open Space	Vegetation	Water Body	Building	Road
Open Space	99.1	337	334	0	0	0	3
Vegetation	98.9	275	0	272	0	0	3
Water Body	99.2	377	0	2	374	1	0
Building	91.0	612	2	0	0	557	53
Road	88.5	148	0	0	0	7	141
Total		1749	336	274	374	575	190
User's Accuracy (%)			99.4	99.3	100	96.9	95.3
Overall class performance (1678 / 1749) = 95.9%							
Kappa Statistic (X100) = 94.0%. Kappa Variance = 0.000043.							

+ (100 - percent omission error); also called producer's accuracy.
* (100 - percent commission error); also called user's accuracy.

For Open Space, it can be said that the accuracy level of producer is 99.1%. There are 337 samples in which 334 samples were counted as open space classes and 3 other samples were counted as roads. It makes the accuracy level 99.1%. It can also be said that the producer's level of accuracy is 98.9 percent for vegetation,where out of 275 samples 272 were counted as vegetation and 3 other samples were counted as roads.For Water Bodies, the accuracy level of producer is 99.2%. Number of samples are 377 where 374 samples calculated as Water Bodies, 2 samples were counted as Vegetation and one of the samples counted as Building. In case of Building, the accuracy level of producer is 91.0%. Number of samples are 612 where 557 samples calculated as Building, 2 samples were counted as Open Space and 53 samples were counted as Road. While for Road, the accuracy level of producer is 88.5%. The number of samples is 148, with 141 samples were counted as roads and 7 as buildings.Thus, overall number of samples were 1749, where Open Space counted as 336, total Vegetation counted as 274, total Water Bodies counted as 374, total Building counted as 575 and total Road counted as 190.

Here it is seen that User's accuracy of Open Space is 99.4 %. There is an interruption of two building samples in Open Space class. The user's level of accuracy must therefore be less than 100%. Similarly, the user's vegetation accuracy is 99.3%. Two samples of water bodies are interrupted in the vegetation class. The accuracy of the water body by the user is 100%. There is no interruption of any sample, so the user's accuracy must be 100%. User's accuracy of Building is 96.9 percent. Samples from both Water body (1sample) and Road (7 sample) have interrupted the class of building. User's accuracy of Road is 95.3%. It shows the lowest level of accuracy in this table, since all classes except the water body in this class have been interrupted. Three Open Space samples, three Vegetation samples and 53 Building samples were counted here. All these samples reduce the level of accuracy.

The overall performance of the class is 95.9 percent, almost 96 percent. The authors divided the sum of all samples in every class by the total amount of samples (ignored interrupted values). The calculated percentage is 95.9.The kappa coefficient of agreement is often used to summarize the outcomes of an accuratenesscalculation to assess the sorting of land use or land cover obtained using remote sensing technique. The kappa coefficient standard estimator together with this estimator's standard error requires a sampling model approximated by simple random sampling.

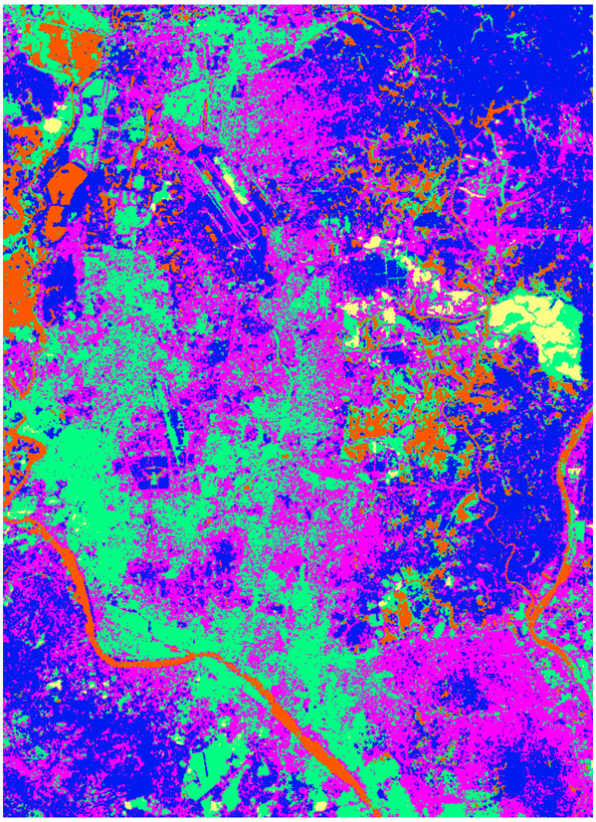
Kappa = 1, perfect agreement exists.

Kappa = 0, agreement is the equal as would be expected by chance.

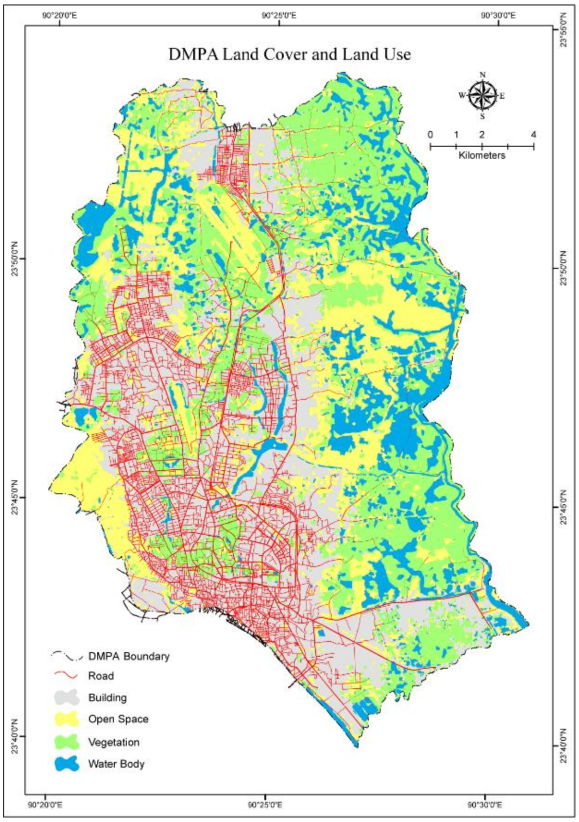
Kappa < 0, agreement is weaker than expected by chance; this seldomhappens.

Here is one possible interpretation of Kappa.

If the value is less than 0.20 then it is a poor agreement. A fair agreement is indicated by the value between 0.20 and 0.40. It will be within 0.40 to 0.60 for moderate agreement. It's going to be a good agreement if the range is from 0.60 to 0.80. The value will be 0.80 to 1.00 for a very good agreement. Here, the Kappa coefficient is 0.94, which accounts for 94.0 percent of Kappa stats. It shows very good agreement because the value ranges from 0.80 to 1.00.



5a. Raw Output from MultiSpec



5b. Processed Map from ArcGIS

Figure 5: DMPALand Cover map (Source: Authors)

Table 2: Land Use andLand Use Class Attributes

NUMBER				
	Class	Samples	Percent	Area (Hectares)
1	Open Space	157855	27.78	14206.95
2	Vegetation	108892	19.16	9800.32
3	Water Body	85675	15.07	7710.79
4	Building	196281	34.54	17665.29
5	Road	19466	3.42	1751.95
	TOTAL	568170	100	51135

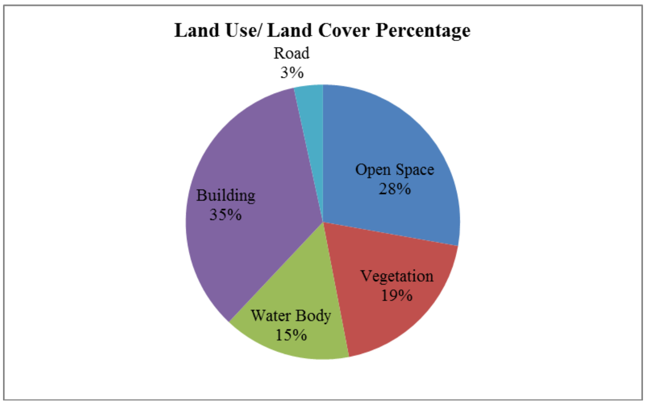


Figure 6: Land use and land cover percentage (Source: MultiSpec)

The land cover map for this study was prepared (Fig.5) with the help of Maximum Likelihood Supervised Classification. This land cover map discloses a significant scenario of Dhaka Metropolitan Police Area (DMPA). The study area covers 51,135.30 hectares, 3 percent of which is road, 19 percent vegetation, 35 percent building, 15 percent water body and 28 percent open space (Table 2). Here, the building area is 17665.29 hectares, covering nearly 35 percent of the total area and Vegetation area is 9800.32 hectares which cover 19.16% of total area. In addition, the area of the Water body is 7710.79 hectares, covering only 15.07 %. Open space covers 27.78 percent of area indicating that urbanization progressively grows and occupies open space (Fig. 6). The overall scenario supports Dhaka Metropolitan's real-life status as the city largely gives up its open space to infrastructure and increasing housing needs. The most important part is the Vegetation cover, where everyone is alarmed by the loss of green space of the city.

Coming to the effort of accuracy assessment for the Dhaka Metropolitan Police Area, the image analysis from supervised classification with the help ofMultiSpec is remarkable in defining the results of its analysis with scientific and statistical justification.Although it should be noted that the correct accuracy assessment requires a great deal of verification with ground truthing, where the authors lacked mobility, access and financial support.

VI. “CONCLUSION”

The image classification of MultiSpec provided detailed information to identify the feature classification of the Dhaka Metropolitan Police Area. The study shows that image classification with MultiSpecis easier and more accurate compared to othertraditional processes of image analysis. The Confusion Matrix and Kappa measure the level of accuracy and agreement. This research can be a way of justifying the future procedures of image analysis with a simple but powerful statistical basis.

REFERENCES

[1]. I.A. Ademiluyi, A.S. Okude, and C.O. Akanni, An appraisal of land use and landcover mapping in Nigeria, *African Journal of Agricultural Research,3(9),*2008,581-586, 1991-637X.

[2]. H.Briassoullis, Analysis of Land Use Change: Theoretical and Modelling Approaches-The Web Book of Regional Science, *Regional Research Institute*, West Virginia University, USA, 1999.

[3]. K.Y. Lee, S.C. Liew and L.K. Kwoh, Land cover Classification and Interpretation of NASA/JPL AIRSAR data based on scattering mechanisms and statistical distribution, *ACRS*,2002.

[4]. S.W. Myint, A new framework for effective urban land-use land-cover classification: A wavelet approach, *Geoscience and Remote Sensing*, *43(2)*, 2006a,155–178.

[5]. S.W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke and Q. Weng, Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery, *Remote Sensing of Environment*, *115(5)*, 2011,1145–1161.

[6]. S.W. Myint and V. Mesev, A comparative analysis of spatial indices and wavelet-based classification, *Remote Sensing Letters*, *3*, 2012,141–150.

[7]. H. Alphan, Land use change and urbanization in Adana, Turkey, *Land Degradation and Development*, 14(6), 2003, 575–586.

[8]. S. Hathout, The use of GIS for monitoring and predicting urban growth in East and West St. Paul, Winnipeg, Manitoba, Canada, *Journal of Environmental Management*, 66,2006, 229–238.

[9]. M. Herold, N.C. Goldstein, and K.C. Clarke, The spatiotemporal form of urban growth: measurement, analysis and modelling, *Remote Sensing of Environment*, *86*, 2003, 286–302.

[10]. P. Serra, X. Pons, and D. Saurı, Land-cover and land-use change in a Mediterranean landscape: a spatial analysis of driving forces integrating biophysical and human factors, *Applied Geography*, *28*,2008, 189–209.

[11]. B.P. Zhang, Y.H. Yao, W.M. Cheng, C.H. Zhou, Z.Lu, and X.D. Chen, Human-induced changes to biodiversity and alpine pastureland in the Bayanbulak Region of the East Tianshan Mountains, *Mountain Research and Development,22*,2002, 1–7.

[12]. T.M. Lillesand and R.W. Kiefer, *Remote sensing and photo interpretation* (New York, John Wiley and Sons, 1994).

[13]. PCI, *Image works version 6.2* (Richmond, Ontario, PCI,1997)

[14]. J.R. Eastman, *IDRIS for windows, user's guide* (USA, Clark University, Worcester, 1995)

[15]. Bangladesh Bureau of Statistics (BBS), *Population and housing census, national volume-3: urban area report* (Dhaka, Bangladesh, Statistics and Informatics Division: Ministry of Planning, 2011)

[16]. J.A. Richards and X. Jia, Remote Sensing Digital Image Analysis, *Springer-Verlag*,1999,194.

[17]. R.G. Congalton and K. Green, Assessing the accuracy of remotely sensed data: principles and practices, *CRC press*. 2008, 9781420055122.

[18]. J.R. Anderson, E.E. Hardy, J.T. Roach and R.E. Witmer, A land use and land cover classification system for use with remote sensor data,*US Government Printing Office ,964,*1976.