Accuracy Assessment of Land Use & Land Cover Classification (LU/LC) "Case study of Shomadi area- Renk County-Upper Nile State, South Sudan"

Abubaker Haroun Mohamed Adam [*], Elhag A.M.H [**], Salih, Abdelrahim . M [***]

[*] Faculty of Agriculture (Al- Kadaro- Khartoum North), University of Bahri- Sudan
[**] Department of Basic Science, College of Natural Resources (Al- Kadaro-Khartoum North), University of Bahri-Sudan.
[***] National Center for Research, Reomte sensing Authority, Natural Resource mapping , Department., Ministry of Sciences' and technology, Khartoum, Sudan

Abstract- This study was carried out in White Nile state to evaluate and tested the accuracy assessment (data quality) to monitor and assess the land cover land use mapping: The accuracy assessment model was used to measure of how many ground truth pixels are correctly classified. To achieve this objective sample points are located by Geographic Positioning System (GPS). Where, Landsat satellite imagery (ETM) year 2010 represented same results with the GPS and Geographic Information System (GIS). The result displayed an accuracy of 94.04 % with over all Kappa statistics (ka) of 91.26 %, however Remote sensing data, GPS data (ground truth), positional accuracy, logical consistency matching with the information on the map.

Index Terms- Accuracy Assessment. Ground-truth. Kappa. Pixels, land use/land cover

I. INTRODUCTION

The accuracy of spatial data has been defined by the United States Geological Survey USGS, 1990 as: "Accuracy assessment or validation is an important step in the processing of remote sensing data. It determines the information value of the resulting data to a user. Productive utilization of geo-data is only possible if the quality of the data is known. Furthermore, integrated processing of different types of geo-data cannot be effective if the data quality is not known. During the last two decades, numerous studies have been published concerning accuracy assessment of (LC) classifications (Congalton, 1996, Rosenfield, et al., 1986, and Foody, 1992).

The accuracy of any map may be tested by comparing the positions of points whose locations or with corresponding positions as determined by surveys of a higher accuracy. Tests shall be made by the producing Researcher or agency, which shall also determine which of its maps are to be tested, and the extent of such testing, (The National Map Accuracy Standards (1998).

The closeness of results of observations, computations to the true values or the values accepted as being true This research is conducted at Shomadi area- Renk County- Upper Nile State- South Sudan in 2010 (Map 1). The area has great potentiality of natural resources which favoured various land use types. Among these, the Mechanized rain-fed farming system, irrigation agriculture, grazing, settlement sprawling, forest related activities, petroleum exploitation etc. The anthropogenic activity has exerted excessive pressure on the existing resources. As consequence, environment and socio-economic set up have been negatively influenced. But, no previous study is carried to assess this situation. Therefore, it became important study the area (Haroun, 2012)

Research Objectives

The objective of this study is to compare the predicted classification of each site with the actual classification as discovered by ground truth [McCoy, (2005), Congalton, (1996), Burnett, et al., (2003) and Habtamu, (2006)].

Having the Classified map following segmentation, accuracy assessment is carried out with two goals in mind, namely:

- To assess how well a classification is worked,
- To understand how to interpret the usefulness the classification.

Research questions

Based on the above mentioned objectives, the accuracy assessment seeks to answer the following questions:

1. The errors' frequency and how often do they occur?
2. The nature of the errors: what kinds of information are confused?
3. The magnitude of errors: how bad are they? e.g., confusing Mechanized rain-fed farming with traditional rain-fed farming is not as' bad' an error as confusing forest with range; confusing similar soils is not as' bad' an error as confusing very different soils.
4. The source of errors: why did the error occur? This allows us to refine the survey methodology.

Statement of the problem

The error matrix and kappa coefficient have become a standard means of assessment of image classification accuracy. This method of determining image classification accuracy resample classified imagery against ground truth field samples often obtained with a Global Positioning System (GPS). Accuracy assessment is further discussed in Chapter four, Results and Analysis. Thus, accuracy assessment is important for the judgment if they performed classification coincides with the nature attributes. Assess accuracy of a remote sensing output is one of the most important steps in any classification exercise!!

Without an accuracy assessment the output or results is of little value. In many studies, the created thematic maps do not coincide with the natural location maps. This happens due to error from the expressed number of sample units (pixels) assigned to a particular category relative to actual category as indicated by reference data.

II. STUDY AREA

The study area situated in Renk County-Upper Nile State-South Sudan. It lies at latitudes 11° 29' 11.76" N - 32° 43' 39.45"E and longitudes 11° 43' 40.62" N - 32° 48' 16.59"E with total area of 169,283.79 ha. (Fig.1) (Haroun, 2012). The study area has huge potentiality of natural resources such as forest, range, animal resources, game, water resources, fertile soil and arable cultivable land, fisheries, etc. Therefore, it became attractive to investors.

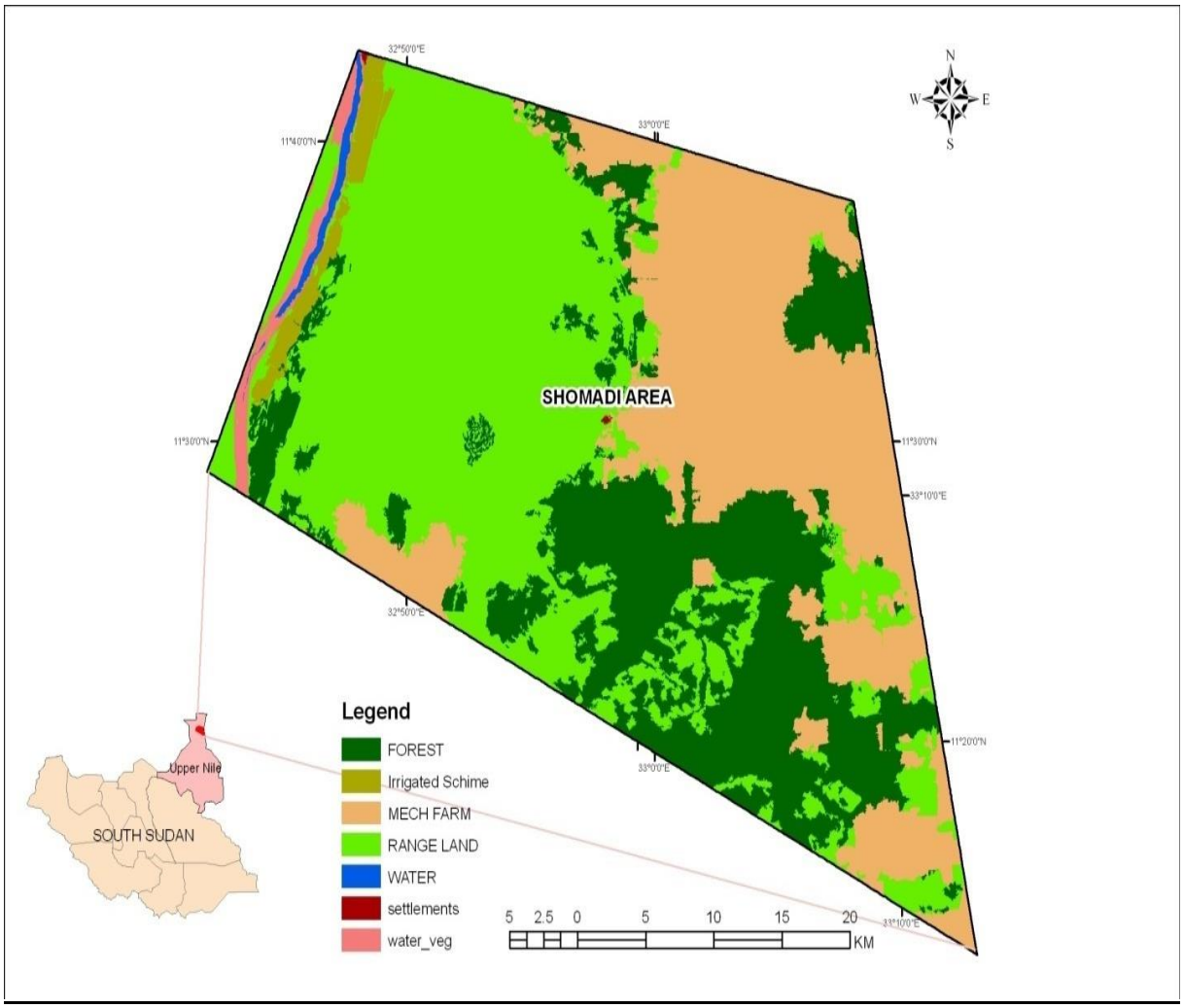

Fig (1) Location map of the study area

III. METHODOLOGY

Ground Control Point Establishment:

In parallel to the remote sensing work; field work is carried to collect data. The field data is collected by measurements approach. Before commencing the field measurement as a pre-requisite; reconnaissance in the form of three field trips. During these field trips, points are selected. Measurements to soil, vegetation and fixing corner points for the boundaries of the study area. The details of all above are illustrated below:

During the second trip, both Ground Control Points (GCPs) and the Area of Interest (AOI) of study area are selected, demarcated & measured by Global Positioning System (GPS). This is done to get an accurate reading for the specific locations for samples positions. In each position 3-4 readings of the GPS are taken. Each Positional measurement is averaged. Alleviations are obtained in field in feet; then, converted to meters (1 Foot =

0.3048 meters). Alleviations are obtained in field in feet; then, converted to meters (1 Foot = 0.3048 meters).

Systematic sample is taken to evaluate the accuracy of the LULC classification as well as soil type of the study area. SPOT-5 map of 2010 and ERDAS imagery version 9.1 are used (Fig.2) The procedure for accuracy assessment: The accuracy assessment is carried out in the following manner:

- Reference data is collected by GPS. Class types are then determined at specific locations which are known as ground truth (Table 1).
- Reference data is compared to map to find out if the class type on classified map equals to class type determined from reference data (Table1).
- Sample size, sample number: In this study sample size of 288 is taken. Sampling method – the systematic-non aligned sampling (a grid provides even distribution of randomly placed GPS ground truth) method is used.
- The data is summarized and quantified by using error matrix (Table 1). The result is presented in a table which displays the followings:

Total Accuracy: Number of correct plots / total number of plots:

- Diagonals represent sites classified correctly according to reference data.
- Off-diagonals are mis-classified.
- The problem with total accuracy is that the summary value is an average which does not reveal if error is evenly distributed between classes or if some classes are really bad and some really good, therefore, it includes other forms, and these are: User's accuracy. It corresponds to error of commission (inclusion): example 1 mechanized farm and 3 forest sites included erroneously in rangeland category. Producer's accuracy this corresponds to error of omission (exclusion), example 7 swamp area (vegetation in wetland) and 1 water body sites omitted from irrigated agriculture category.

⇨ User's Accuracy: From the perspective of the user of the classified map, how accurate is the map?

- For a given class, how many of the pixels on the map are actually what they say they are?
- Calculated as: Number correctly identified in a given map class / Number claimed to be in that map class.

⇨ Producer's Accuracy: From the perspective of the maker of the classified map, how accurate is the map?

- For a given class in reference plots, how many of the pixels on the map are labeled correctly?
- Calculated as: Number correctly identified in ref. plots of a given class / Number actually in that reference class.

⇨ Kappa: Estimated as ($\hat{K}$). It reflects the difference between actual agreement and the agreement expected by chance.

Kappa of 0.75 means there is 75% better agreement than by chance alone

Kappa = observed accuracy – chance agreement
1- Chance agreement

- Observed accuracy determined by diagonal in error matrix.
- Chance agreement incorporates off-diagonal.

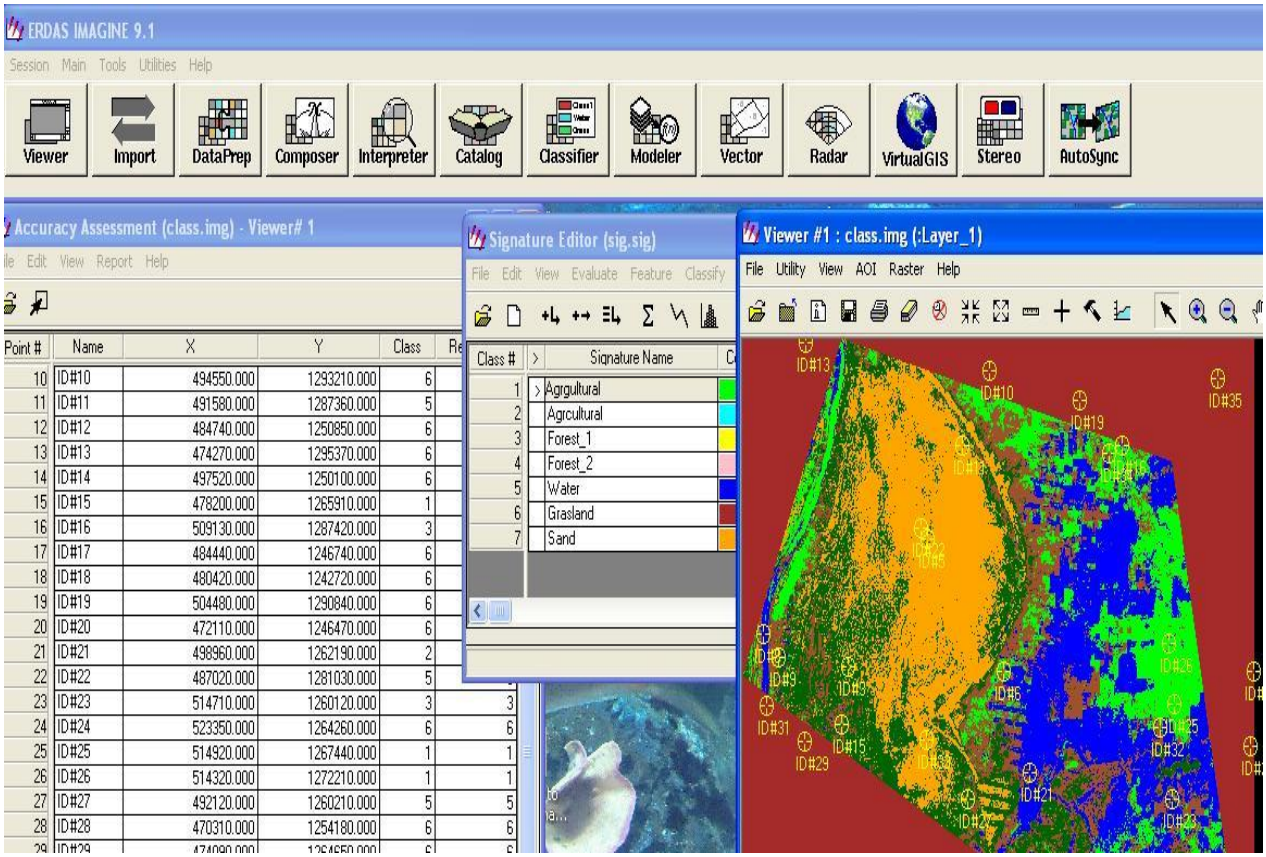

Fig.2: ERDAS accuracy assessment viewer

Reference Data: Geo-referenced Raster Image of year 2010									Percent	
									CE	**UA**
Classification	**Forest**	**Irrigate Agric.**	**M.rainfed Agric.**	**Range**	**Water bodies**	**Setttel.**	**swamps**	**Total**		
Forest	**68**	0	2	2	0	0	0	72	5.56	94.44
Irrigated agric.	0	**04**	0	0	0	0	0	4	0	100
M.rainfed Agric.	4	0	**73**	5	0	0	0	82	18.29	89.02
Range	3	0	1	**119**	0	0	0	123	21.14	96.75
Water bodies	0	0	0	0	**02**	0	0	2	0	100
Settlements	0	0	0	0	0	**02**	0	2	0	100
Swamps	0	0	0	0	0	0	**03**	3	0	100
Total	75	4	76	128	2	2	3	**288**	**OA = 94.09%**	
OE	9.34	0	3.95	8.5	0	0	0	0		
PA	90.66	100	96.05	92.74	100	100	100	100	**Kappa=0.9126**	

Table 1: Error matrix accuracy totals for the classified image (2012)

IV. RESULTS AND DISCUSSION

Table (1) which is shown above represents data of the results of accuracy assessment of SPOT-5 and field survey during 2010. It shows:

- Overall accuracy (OA), Kappa statistics (KS). Omission error (OE), Commission error (CE), Producer accuracy (PA) and User accuracy (UA).
- A pixel accurate classified: it is a comparison to a reference, determines probability that a pixel represents the class for which it has been assigned.
- The total accuracy is measured by calculating the proportion class pixel relative to total tested number of pixel (Total =Total corrected/Total tested)

Overall accuracy = (68+4+73+119+2+2+3)/(68+2+2+4 +4+73+5+119+1+3+2+2+3) = 271/2988= 94.09%

$$K = \frac{288(68+4+73+119+2+2+3) - [(72x75)+(4x4)+(82x76)+(123x128)+(2x2)+(123x128)+(2x2)n(2x2)+(2x3)]}{(288)^2 - [(72x75)+(4x4)+(82x76)+(123x128)+(2x2)+(2x2)+(2x3]}$$

= 451131/56027 =0.9126

To evaluate the accuracy of the classification system, (288) reference test pixels are identified (please refer to Appendix C). The classified image of 2010 is verified with digital (LU) and (LC) maps. The assessment showed an overall accuracy derived from the stratified random sampling method for the 2010 classified images is 94.04% with an overall kappa statistic of 0.91,26.

Considering the categories accuracy, the stratified random method provided very high accuracy assessment in irrigated agriculture, water bodies, settlements and swamps, each with an accuracy of 100%. It also provided high accuracy of 96.05%, 92.74%, 90.66 % for mechanized rain-fed farming, rangelands and forest lands respectively. This reduction in accuracy is due to the similarity between range land grasses and the follow lands in the mechanized rain-fed farming. Moreover, the open forest vegetation is similar to range vegetation, thus conversion occurred between these (LU) and (LC) categories. Similar studies displayed different accuracy assessment. This is because the probability that a reference pixel correctly classified is determined by the producer's accuracy. The probability that a classified pixel from the (LC) map accurately corresponds with the referenced data is determined by the user's accuracy (Jensen, 2005), while the Khat Coefficient (K) or Kappa statistic measures the difference between the true agreement of classified map and chance agreement of random classifier compared to reference data (Lillesand et al., 2004). It is stated that Kappa values of more than 0.80 indicate good classification performance. Kappa values between 0.40 and 0.80 indicate moderate classification performance and Kappa values of less than 0.40 indicate poor classification performance (Jensen, 2005, Lillesand et al., 2004). Based on this judgment, this study has proved high accuracy assessment for the map of 2010. However, similar accuracy assessment are computed in different studies, among these a study by Doxani, etal., (2006) to investigate urban land cover change detection of Thessaloniki, a city in the north part of Greece during 2003 - 2006. Where Object- Oriented classification approach was used. The result showed an overall accuracy and Kappa coefficient of 70-75% and 65% respectively. Lucy (2011) studied change detection in Wanjohi area, Kenya. Object – oriented image classification was carried out using ERDAS imagine objective. An overall accuracy of about 88.33% was achieved indicating the probability that a randomly selected point on the map was correctly mapped. The tree object accuracy was at 100%. 87. Another study on (LC) and (LU) change in Mbeere District, Kenya was carried out by Peter (2007). Where two (LC) maps corresponding to 1987 and 2000 Landsat images were produced. The overall accuracy of the two maps was above 85% and the overall kappa statistics was above 0.81. Different (LC) classes had differing producer's and user's accuracy levels indicating different levels of omission and commission errors. Woodland in 2000 had the lowest producer's accuracy of 74.4%, while settlement in 2000 user's accuracy of 77.8 % was the lowest.

V. CONCLUSION AND RECOMMENDATION

The goals of accuracy assessment can be expanded beyond consideration of the nature and frequency of errors in maps to include assessment of the magnitude of those errors and alternative measures of their frequency.

The result in general showed that there was considerable and urgent need to compare the information on the map with what is on the ground in order to avoid repeated mistakes and misinformation when dealing with map.

The satellite remote sensing, GPS and GIS technology helped to overcome the limitations of manual system. This technique has been useful to supply temporal and synoptic data of high quality in advance of land use, land cover mapping. This has also helped to monitor natural calamities as floods and drought.

The use of accuracy assessment in thematic maps can improve lands use, land cover characterization by allowing explicit acknowledgement of heterogeneity within map units and scale.

Based on these finding the following recommendation can be stated:

1- Developing national standard and national reference for assessing maps.

2- Researchers must use modern techniques and some more global analytical modules to assess and evaluates the maps.

3- Map accuracy linked to other factors must be taken when evaluation and design map muddle.

REFERENCES

[1] Abubaker Haroun Mohamed Adam (2012): Land Use & Land Cover Classification (LU/LC) Case study of Shomadi area-Renk County-Upper Nile state, South Sudan.

[2] Burnett, C., and T. Blaschke. (2003): A multi-scale segmentation/object relationship modeling methodology for landscape analysis. Ecological Modeling 168:233-249.

[3] Congalton, R.G. (1991): A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing of Environment 1991. 37: p. 35-46.

[4] Foody, G. (1992): On the Compensation for Chance Agreement in Image Classification Accuracy Assessment. Photogrammetric Engineering & Remote Sensing. Vol. 58. No. 10. p. 1459-1460.

[5] G. Doxani *, S. Siachalou, M. Tsakiri- Strat (2006): An Object-Oriented Approach to Urban Land Cover Change.

[6] Habtamu (2006): Land covers change detection using hyperspectral imagery in the islands of schiermonnikoog, The Netherlands.

[7] Jensen, J. R. (2005): Introductory digital image processing: a remote sensing perspective. John R. Jensen. 3rd ed., Pearson Education Inc., Upper Saddle River.

[8] Lillesand TM, Kiefer RW, Chipman JW, (2004): Remote Sensing and Image Interpretation (5th_edn.), John Wiley and sons, Inc.:Hoboken, New Jersey.
Detection.http://www.isprs.org/proceedings/XXXVII/congress/7_pdf/10_ThS-18/22.pdf.

[9] Lucy Chepkosgei Chepkochei(2011): Object- Oriented Image classification of Individual trees using ERDAS Imagine objective: Case study of Wanjohi area, Lake Naivasha Basin, Kenya. Proceedings, Kenya Geothermal Conference 2011.

[10] McCoy, Roger M. (2005): Field methods in Remote Sensing. New York: The Guilford Press.

[11] Peter Masavi Maluki, (2007): Mapping land cover/land change in Mbeere District, Kenya. MSc thesis.

[12] Rosenfield, G. & Fitzpatrick-Lins, K. (1986): A Coefficient of Agreement as a Measure of Thematic Classification Accuracy. Photogrammetric Engineering & Remote Sensing. Vol. 52. No. 2. p. 223-227

[13] The National Map Accuracy Standards (October 9, 1998) Minnesota Governor's Council on Geographic Information, Spatial Data Accuracy Handbook.

AUTHORS

First Author – Abubaker Haroun Mohamed Adam, Faculty of Agriculture (Al- Kadaro- Khartoum North), University of Bahri-Sudan. , Mobile (+249) 0912892429. Email: abubakerharoun@gmail.com
Second Author – Elhag A.M.H, Department of Basic Science, College of Natural Resources (Al- Kadaro-Khartoum North), University Of Bahri-Sudan. Tel:(249) 0912815884. Email: remotesensing_9@yahoo.com or remotesensing9@gmail.com
Third Author – Salih, Abdelrahim . M, National Center for Research, Reomte sensing Authority, Natural Resource mapping , Department., Ministry of Sciences' and technology, Khartoum, Sudan, Mobile:+249-9-1242137. Email: abdelrahim_rsa95@hotmail.com