Land Cover Classification: Comparison between Fuzzy and Boolean Classifier

Abdullahi Ahmed KUTA, Oluibukun Gbenga AJAYI, Ekundayo Abayomi ADESINA, Zitta NANPON, Hassan .A. SAMAILA-IJA., NIGERIA

Key words: Remote Sensing, Image Classification, Earth monitoring, Image Resampling, Spatial resolution, Land Spectral Classes.

SUMMARY

Production of Land Use/Land cover maps is very important for environmental monitoring and development. Image classification using either hard and/or soft classifiers is crucial in the production of these maps. While fuzzy classification is suitable for modelling vagueness due to mixed pixels in the land cover, Boolean on the other hand is suitable for modelling land cover with well-defined boundary. The analyst's choice of image classifier is a very important decision in image classification as this determines the classification output. Using Landsat5 TM of 1984, Landsat 4TM of 1992 and Landsat7 ETM+ of 2000 satellite images, this research looks at the comparison between soft (Fuzzy) and hard (Boolean) classifiers. The Landsat ETM+2000 of a 15m spatial resolution was resampled to a 30m pixel size so that the three images would be of the same pixel size to effectively carry out pixel-to-pixel analysis. Due to the nature of the landscape and bearing in mind that land cover responds differently to various Landsat spectral bands, three band combinations (image bands 2, 3, and 4) were considered for the classification. The images were classified into four (4) different land spectral classes by employing the fuzzy membership function and maximum likelihood classification tools in Idrisi Taiga 16 software. The results obtained shows that the spatial distribution of the modelled land cover classes for both Fuzzy and Boolean is basically the same which buttresses the performance level of both models. The major difference of the two models lies in the output; while fuzzy shows a subtle representation according to degree of membership function of each land cover class, the Boolean on the other hand represented the land cover types with a well-defined boundary. Also, summation of the fuzzy land cover areas is not equal to the size of the study; 108% in 1984, 107% in both 1992 and year 2000 are unlike the Boolean with 100%.

Land Cover Classification: Comparison between Fuzzy and Boolean Classifier

Abdullahi Ahmed KUTA, Oluibukun Gbenga AJAYI, Ekundayo Abayomi ADESINA, Zitta NANPON, Hassan .A. SAMAILA-IJA., NIGERIA

1. INTRODUCTION

The need for timely and accurate land cover modelling is instrumental to the effective management of our physical and natural resources. Land cover mapping for large regions often employs satellite images of medium to coarse spatial resolution (Colditz, 2015) which have become a very important source of information for adequate management of resources. The major advantage of remote sensing is that it provides spatio-temporal information about the earth's surface over a large extent, which makes it an ideal tool for land cover modelling and therefore provides ideal data for extracting land cover information (Kuta and Comber, 2015). Classification is crucial for land cover change modelling which is usually achieved using satellite images. Various classification approaches are used depending on the nature of feature or objects to be classified. Basically, Hard and soft classification techniques are the conventional methods of image classification for satellite data, but they have their own advantages and drawbacks (Hu et al., 2013). Liu et al. (2011) stated that among the two types of classifiers, soft classification provides more information than hard classification and consequently, it is required in certain situations where the probability information is useful. Nevertheless, if the class probability function is hard to estimate in some complicated problems, hard classification may produce more accurate classifiers by targeting on the classification boundary only (Wang et al., 2008). The choice of a Soft (fuzzy) or hard (Boolean) approach to model land cover has to do with the nature of the land cover to be modelled.

The Fuzzy approach is able to solve the major problems in image classification (Hegde, 2003), which makes it suitable for modelling vagueness resulting from a mixed pixel - a pixel not completely occupied by a single, homogeneous category (Zhang, 1998). While Boolean is suitable for modelling land cover with well-defined boundary but they all make a definitive decision about the land cover class to which any pixel belongs (Eastman, 2009). A fuzzy and Boolean approach for modelling land cover types will be compared with the aim of analyzing the qualitative (graphical presentation) and quantitative (computed land area) output produced by each approach.

2. LITERATURE REVIEW

2.1 Land Cover Classification

One of the primary fields in remote sensing is Land cover classification from satellite images (Coddity, 2015). Classification is an automated computer assisted grouping of pixels of remotely sensed images into information (Kata and Comber, 2015) into land cover classes in other to convert the data (images) into distinct land spectral classes. The main objective of classification is to classify features of interest into distinct land spectral classes and the approach chosen depends on the need for the classification, purpose of the classification, how conversant the analyst is with the scene or features to be classified, the accuracy required etc. There are basically two types of classification namely, supervised and unsupervised classification, usually performed to produce land cover maps from remote sensing data, mostly for large areas (Saha et al., 2005).

Unsupervised classification has to do with the classification of all the pixels with unknown identities, blindly grouped into a certain number of clusters according to the similarities in their digital numbers (Gao, 2009). This type of classification is more machine dependent; the analyst does not have any control over assigning the classes but only indicates the number of the proposed land cover clusters. Prior knowledge of the area is not required by the analyst. However, in a supervised classification, the analyst has more control over the classification by assigning to the machine the pixels that belong to a cluster. The analyst "supervises" the pixel categorization process by specifying to the computer algorithm numerical descriptors of the various land cover types present in a scene (Lillesand et al., 2008). The knowledge of the study area is essential and could also produce more accurate results than unsupervised classification. Supervised classification is divided into two stages: training and classification stages. The training stage involves developing spectral signatures for various predefined land cover classes in the scene, which the machine uses as the basis for its classification. The training stage is both an art and a science and requires a close interaction between the image analyst and the image data (Lillesand et al., 2008). The success of supervised classification depends on the quality of the training data.

2.2 Classifiers

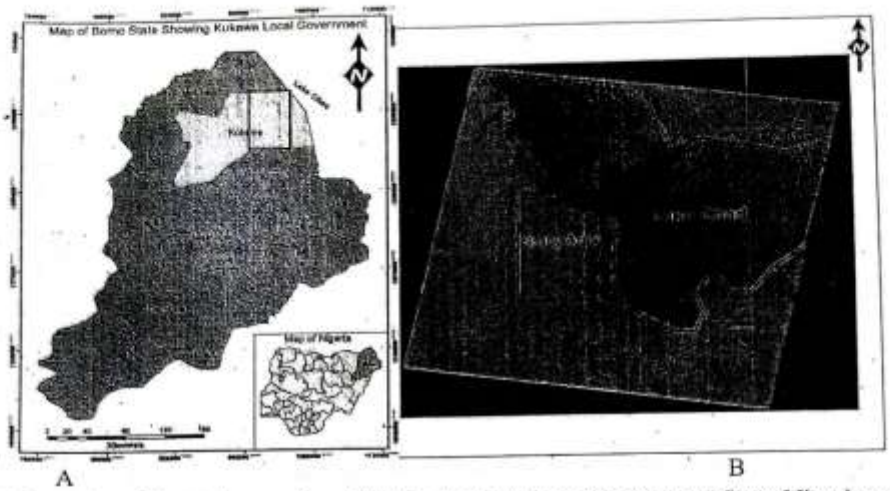
Among many classifiers, some are hard classifiers while some are soft ones which are based on different philosophies with each having its merits and demerits, depending on the landscape on which they are used (Liu, et al., 2011). The hard traditional (crips) classification is use when the objects/features have well-defined boundaries (homogeneous) with assumption that a pixel can only belong to a particular class hence it is not suitable for modelling heterogeneous landscapes. While the soft classifier is used when the land cover classes have no well-defined boundaries. Its ability to represent vaguely defined geographical phenomena makes fuzzy classification the preferred approach over Boolean while Boolean is preferred in a land scape with a well-defined boundary. To solve this problem, a notion of partial truth was formulated mathematically by Zadeh in 1965 and processed by

computers, in order to apply a more human-like way of thinking in the programming of computers, which is an alternative to the traditional notion of set membership and logic, which originated from the ancient Greek philosophy called fuzzy logic (Hellmann, 2002). Basically, fuzzy logic is a multivalued logic that allows intermediate values to be defined between conventional evaluations like true/false, yes/no, high/low, etc. which can be formulated mathematically and processed by computers to handle the concept of partial truth relating to classes of objects whose boundaries are continuous; the boundaries of perceived classes are fuzzy. The true value may range between completely true and completely false; that is, there is a degree of belonging that has a fuzzy set membership unlike the traditional set membership logic theory (Boolean logic), where binary sets have two values, true or false (Hellmann, 2002). In fuzzy logic, each input-point is mapped based on the membership value as defined by a fuzzy membership function between 0 and 1. Probability attempts to model clearly defined and randomly occurring events, where the highest probability class is interpreted as the actual class, while fuzzy logic is concerned with the vagueness or ambiguity occurring when describing the event itself.

3. METHODOLOGY

3.1 Study Area

The study area is located in Kukawa; a town and local government area in Borno, a north-eastern state of Nigeria, bordering Lake Chad on the north- east side. It shares boundaries with both Cameroun and the Republic of Chad in the east, Mobba local government area in the north-west, the local government areas of Kaga and Madukur in the south-west and Munguno in the south, which are all in Nigeria (Kuta and Comber, 2015). Kukawa is located between latitudes 12° 20' 00"N and 13° 20' 36"N and longitudes 12° 20' 00"E and 13° 15' 34"E; it has an estimated population of 3,576 people and it is 111.5 miles away from Maiduguri, Borno's state capital (http://en.wikipedia.org/wiki/Kukawa). Figure 1 shows a diagrammatic representation of the study area.

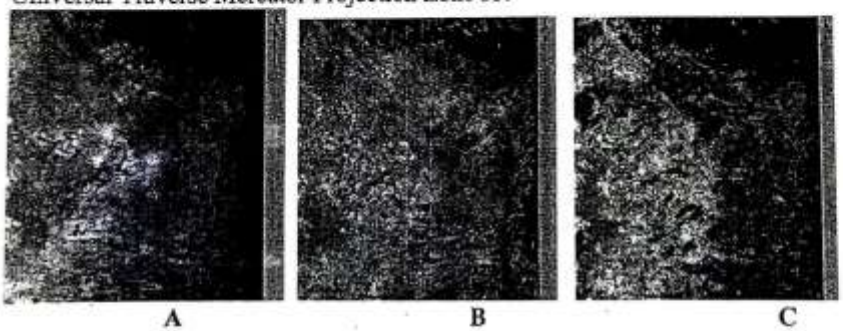


A B

Figure 1: Location of the study area in red: (A) in relation to the map of Borno State, Nigeria and (B) in relation to a Landsat satellite image of 2000 (source: Kuta and Comber, 2015).

3.2 Data

Three Landsat images acquired of three different epochs were used: Landsat5 TM of 1984(30m pixel), Landsat 4TM of 1992 (30m pixel) and Landsat7 ETM+ of 2000(15m pixel). The images were downloaded from the USGS website and geometrically corrected and projected to WGS 1984 Universal Traverse Mercator Projection Zone 33.



A B C

Figure 2. Composite images of the study area: **(A)** 30m pixel size, 1984 Landsat5 TM 234 band composite of the study area: **(B)** 30m pixel size, 1992 Landsat4 TM 234 band composite of the study area, and **(C)** 15m pixel size, 2000 Landsat5 TM 234 band composite of the study area.

3.3 Methods

The three satellite images were sub-setted to carve out the area of interest before combining the three image bands. After the generation of colour composite images of Landsat TM5, TM4 and ETM+ images shown as images A, B and C in Figure 2. The Landsat ETM+2000 of a 15m ground sampling distance was resampled to a 30m pixel size to make the three images have the same pixel size for effectively pixel-to-pixel analysis and change detection.

Before the classification, the signature files for the four land cover classes were created after digitizing the training sites to enable the software to recognise the land cover classes since it could only recognise the identifier integer 1- 4 allotted during training sites development (Ojigi, 2006). The developed signature file contains the statistical information about the reflectance values of each pixel within the training sites representing the four land cover classes (Kuta and Comber, 2015). The integers (1-4) were replaced by their respective land cover classes after signatures were developed for supervised classification for both soft and hard.

3.4 Fuzzy and Boolean Supervised Classification

The last stage was fuzzy and Boolean supervised classification after the signature development. The normalized supervised sigmoidal fuzzy membership function was used for fuzzy classification in Idrisi Taiga 16.0 because it is believably the most commonly used function in fuzzy set theory, and it is produced using the cosine function, better suitable for continuous surfaces. The membership function was normalized and the z-score set at 2.58. These two parameters are important in fuzzy classification. The normalization of the membership value is based on the assumption that the classes are exhaustive; that is, the membership value for all the classes for a single pixel must add up to 1.0. (Eastman, 2010). For the hard classification, the maximum likelihood was used which is more straight forward. The Maximum Likelihood procedure, provided by the MAXLIKE module in IDRISI, is based on Bayesian probability theory. The information from a set of training sites is used as the mean and variance/covariance data of the signatures to estimate the posterior probability that a pixel belongs to each class (Eastman, 2009).

3.5 Fuzzy and Boolean Area Computation.

The total area of any fuzzy land cover class was computed by taking the sum of all the pixel membership values for that land cover class and multiplying it by the size of the pixels on the ground (Fisher et al., 2006). Equation 1 shows the formula for computing fuzzy areas. For Boolean area, the total area is computed by multiplying the pixel count for each land cover class with the size of the ground pixel (30m x 30m = 900 square metres)

$F_{AC}=\sum \mu c x$ for all values of x . eqn 1

F_{AC}=fuzzy area of class C

μc=membership function of class C

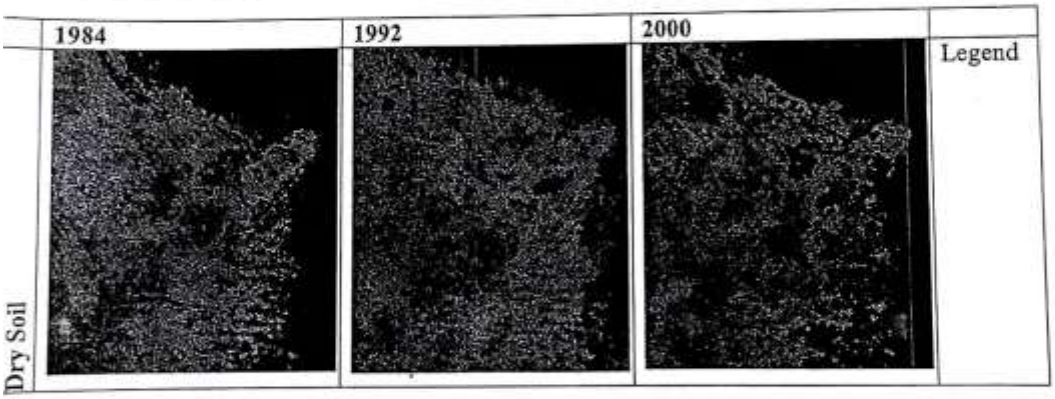
x = pixel size on the ground. Since the image pixel size is 30m, therefore, the pixel size on the ground is 30m x 30m = 900 square metres.

4. RESULTS

4.1 Fuzzy Maps of Various Types of Land Cover

Figure 3 shows the fuzzy maps of various land cover classes in the study area in 1984, 1992 and 2000. The columns represent the years, whilst the rows represent the fuzzy land covers maps. For example, row 1 is for dry soil for the periods 1984, 1992 and 2000, whilst row 2 is moist soil, row 3 is vegetation row 1 is for dry soil for the periods 1984, 1992 and 2000, whilst row 2 is moist soil, row 3 is vegetation and, lastly, row 4 is for water. Column 1 is for the classes of land cover in 1984, column 2 is for 1992 and, lastly, column 3 is for 2000. The legend indicates the degree of fuzzy membership function of each class; that is, black (0) indicates areas without fuzzy membership, whilst red (1) shows areas with full fuzzy membership rising from above 0.

Looking at row 1, for dry soil, the maps reveal that the north-eastern and south-eastern regions are dark. This is because the study area is bordered by Lake Chad and the areas are mostly water and vegetation. Rows 3 and 4, on the other hand, shows the occurrence of vegetation and water on that part of the map (north-eastern and south-eastern). It is possible to determine, in the three fuzzy dry soil maps, that the spatial extent of dry soil was the largest in 1992, having increased from 1984, and it reduced in 2000. Moreover, the second row shows that the largest extent of moist soil occurs in 2000. In addition, vegetation is less extensive during this year. Water, however, may be more extensive in this year, although it is hard to be ascertain this from the map.



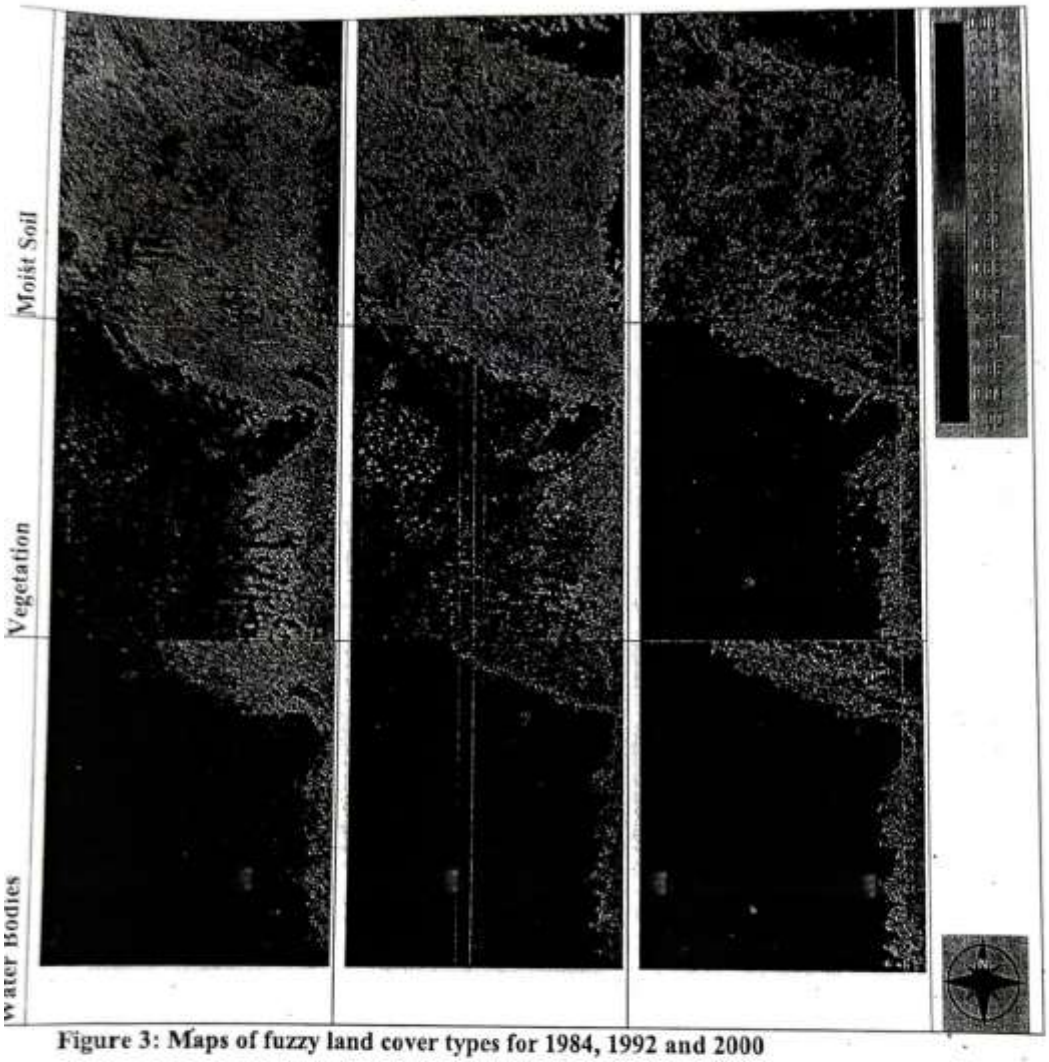
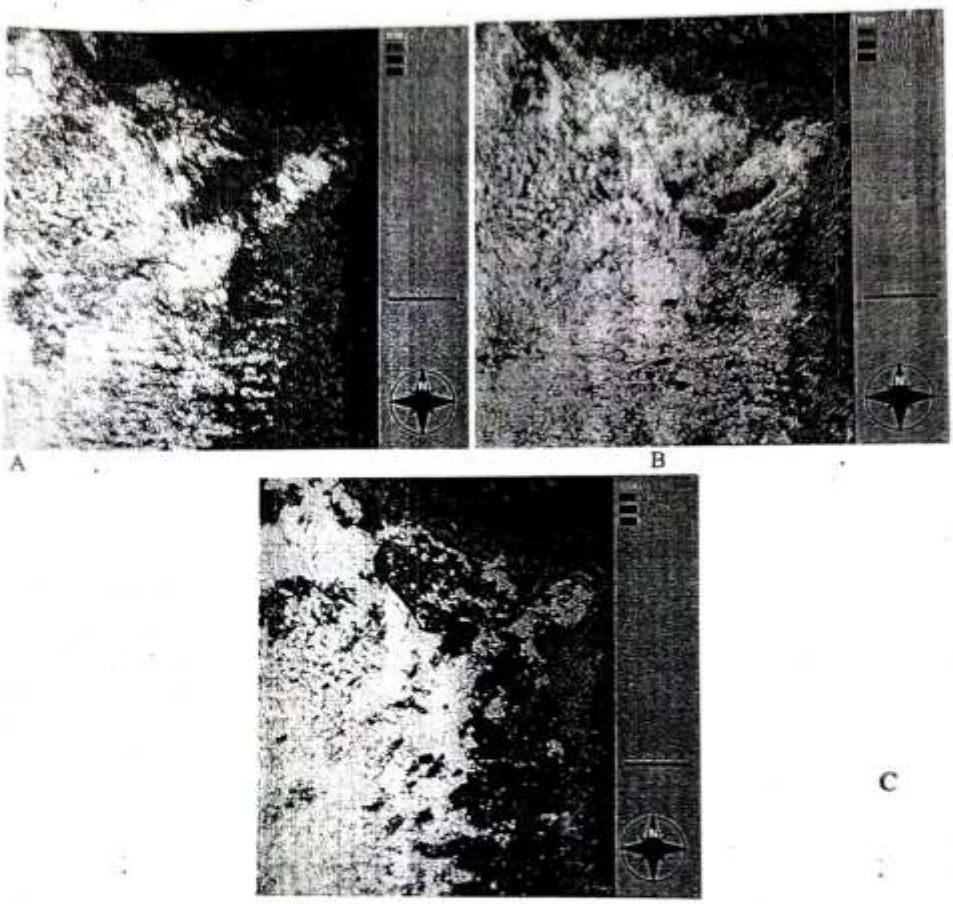


Figure 3: Maps of fuzzy land cover types for 1984, 1992 and 2000

4.2 Boolean Maps of Various Types of Land Cover

The maps in Figures 4 shows Boolean land cover maps of 1984, 1992 and 2000. Figure 4B shows that the dry soil and water has the largest extent in 1992 increasing from 1984 and decreasing in 2000.

While the extent of vegetation and moist soil is directly opposite trend from the dry soil and water; least in 1992; decreasing from 1984 and then increasing in 2000.



A B C

Figure 4: Boolean land cover map of 1984 (A), 1992 (B) and 2000 (C)

4.3 Fuzzy and Boolean Area of Various Types of Land Cover

Table 1 shows the fuzzy areas covered by all the land cover classes for the periods under investigation in square kilometres and proportional percentage of the total area of the study area, which is 3,572.986

km^2. Table 2 shows the areas for the three dates both in square kilometres and as a percentage for Boolean land cover class

Table 1: Fuzzy areas of land cover classes in 1984, 1992 and 2000 (Kuta and Comber, 2015)

Land cover class	Area in square kilometres			Area as percentage (in %)		
	1984	1992	2000	1984	1992	2000
Dry Soil	1394.826	1724.60	1428.34	39.02	48.27	39.98
Moist Soil	1656.780	1325.53	1626.46	46.37	37.10	45.52
Vegetation	686.940	549.51	516.90	19.23	15.37	14.47
Water	134.660	238.25	259.69	3.77	6.67	7.27
Total	3873.206	3837.89	3831.390	108	107	107

Table 2: Boolean Areas of Land cover classes in 1984, 1992 and 2000.

Land cover class	Areas in square kilometres			Area as percentage (in %)		
	1984	1992	2000	1984	1992	2000
Dry Soil	1397.640	1756.023	1502.468	39.12	49.15	42.05
Moist Soil	1295.014	1240.668	1335.248	36.24	34.72	37.37
Vegetation	816.836	271.774	493.381	22.86	7.61	13.81
Water	63.496	304.520	241.889	1.78	8.52	6.77
Total	3,572.986	3,572.986	3,572.986	100	100	100

5. DISCUSSION OF RESULTS

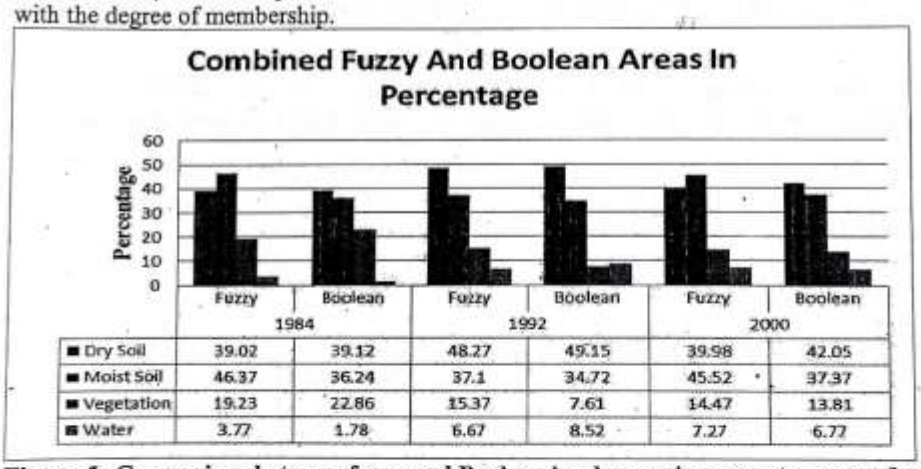
5.1 Comparison between Fuzzy and Boolean Classification

The results in Figure 3 shows the spatial extent of various fuzzy land cover classes in the study area for the periods of 1984, 1992 and 2000 while Figure 4 shows the spatial extent of various Boolean land cover classes in the study area.

Fuzzy output maps show each land cover class as a continuous variation having a degree of membership rather than a sharp boundary, as shown in Figure 3, twelve maps (four per date) for three epochs were produced for all the land cover classes due to the fact that each pixel may contain membership of all four land cover classes (Fisher at el., 2006). The legend shows the degree of fuzzy membership function of each class; that is, black (0) which indicates areas without fuzzy membership, whilst red (1) shows areas with full fuzzy membership rising from above 0. Fuzzy maps are able to reveal some information at sub-pixel level, which helps understand the degree of transition in land cover classes. This is one of the advantages Fuzzy has over Boolean and also makes it more suitable for modelling vague land cover classes. Whereas the Boolean output gives three maps (one per date) for three years i.e., it produces a map having all the land covers instead of producing a nap per land cover class. This type of map could be useful when simulation is to be done to project the likelihood of changes in the land cover types unlike the fuzzy. Both fuzzy (Figure 3) and Boolean (Figure 4) shows

that the dry soil is found mostly at north-west and the south western part of the study area and also that the dry soil is more extensive (1992 map). Also, the moist soil is more extensive around the centre of the study area, stretching towards the northwest and the southwest, while water and vegetation are located at north eastern part of the study area for both fuzzy and Boolean map. The spatial distributions of the modelled land cover classes for both fuzzy and Boolean is basically the same which buttresses the performance level of both models. The major difference between these two models lies in the output. While fuzzy shows a subtle representation according to degree of membership function of each land cover class, the Boolean on the other hand represented the land cover types with a well-defined boundary.

Figure 5 is the comparison between fuzzy and Boolean land cover areas. For dry soil, the difference is not much having almost the same area, e.g., 39.02% fuzzy and 39.12% Boolean in 1984 and with none having a difference of more than 3%. Moist soil has little variation; 46.37% fuzzy and 36.24% Boolean in 1984. The magnitude of underestimation is greater in 1984 and 2000 at more than 8%, whilst it is less than 3% in 1992. The area coverage of vegetation was overestimated in 1984; 19.23% fuzzy and 22.86% Boolean but underestimated in 1992 and 2000. The water body was underestimated in 1984 by 3.77% fuzzy and 1.78% Boolean. But the results in 1992 and 2000 were similar, 7.27% fuzzy and 6.77% Boolean. The over and underestimation of Boolean could be a result of the vagueness in the landscape, which it was not able to resolve. For instance, moist soil is mixed up with water and dry soil, whilst vegetation is mostly found on the water. This type of land that does not have classes with well-defined boundaries is not suitable for Boolean classification. One of the major difference between Fuzzy and Boolean area is that the summation of fuzzy area is not equal to the total size of the study area (Table 1) i.e., not equal to 100% unlike the Boolean area (Table 2). This is because fuzzy deals with the degree of membership.



	1984 Fuzzy	1984 Boolean	1992 Fuzzy	1992 Boolean	2000 Fuzzy	2000 Boolean
Dry Soil	39.02	39.12	48.27	49.15	39.98	42.05
Moist Soil	46.37	36.24	37.1	34.72	45.52	37.37
Vegetation	19.23	22.86	15.37	7.61	14.47	13.81
Water	3.77	1.78	6.67	8.52	7.27	6.77

Figure 5: Comparison between fuzzy and Boolean land cover in percentage area for 1984, 1992 and 2000 derived from Table 1 and 2.

6. CONCLUSION

Both fuzzy and Boolean classification techniques were used in this research. The output of the two classificatiops differs as fuzzy shows a subtle representation of membership functjon in each land cover, while Boolean has a null- defined boundary. A summation of fuzzy land areas did not give 100% but Boolean gave 100% total land area. An interesting research direction could be to investigate the possibility of integrating both Fuzzy and Boolean algorithms into a hybrid model and to see the behaviour of the output. Furthermore, the use of high/medium resolution satellite imageries can be used with Fuzzy and Boolean classification.

REFERENCES

Colditz, R. R. (2015). Article An Evaluation of Different Training Sample Allocation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-Based Algorithms, Remote Sensing ISSN 2072-4292 www.mdpi.com/journal/remotesensing.

Eastman, J. R. (2009). *IDRISI Taiga Guide to GIS and image Processing.*

Fisher, P., Arnot, C., Wadsworth, R. and Wellens, J. (2006). Detecting change in vague interpretations of landscapes. Ecological informatics 1, pp. 163 – 178.

Gao, J. (2009). Digital Analysis of Remotely Sensed Imagery, New York: McGraw - Hill Companies, Inc.

Hegde, S. (2003). Modelling Land Cover Change: A fuzzy Approach. Thesis submitted to the international Institute for Geo-information Science and Earth Observation, Netherlands.

Hellmann, M. (2000). Fuzzy Logic Introduction.

Hu, T. G., Xu, J: F., Zhang, D. R., Wang, J. and Zhang Y.Z. (2013). Hard and soft classification method of multi-spectral remote sensing image .based on adaptive thresholds. http://www.ncbi.nlm.nih.gov/pubmed/23841424. Accessed 28/09/2015

Kuta, A. A. and Comber, A. J. (2015). A Fuzzy approach to modelling land cover changes in north-eastern Nigeria, International Journal of Geomatics and Geosciences Volume 6, No 1, 2015, ISSN 0976 – 4380:pp1620-1637.

Lillesand, T. M., Kiefer, R. W. and Chipman, J. W. (2008). Remote Sensing and Image Interpretation, six editions, John Wiley and Sons, Inc., Hoboken, USA.

Liu, Y., Zhang, H. H. and Wu, Y. (2011). Hard or Soft Classification? Large-margin Unified Machines, Journal of the American Statistical Association, 2011 Mar 1;106 (493):166-177.

Ojigi, L. M. (2006). Analysis of Spatial Variability of Abuja Land Use and land Cover from image classification algorithms, ISPRS Commission VII Mid-Term Symposium. Theme: Remote Sensing: From Pixel to Processes, Enschede, The Netherlands.

Saha, A. K., Arora, M. K., Csaplovics, E. and Gupta, R. P. (2005). Land Cover Classification Using IRS LISS III Image and DEM in a Rugged Terrain: A Case Study in Himalayas, Geocarto International, Vol. 20, No. 2, 33-40

Wang, J., Shen, X., and Liu, Y. (2008), "Probability estimation for large-margin classifiers," Biometrika, 95, 149–167

Zadeh, L. A. (1965). Fuzzy sets. Information and Control 8, 338–353.

Zhang, J. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery, International Journal Remote Sensing, vol. 19, no. 14, 2721 -2738.

BIOGRAPHICAL NOTES

Abdullahi. Abdullahi KUTA is a graduate from the University of Leicester, UK with M.Sc. Geographical Information Systems (2013) and a Bachelor of Surveying & Geoinformatics, FUT Minna, Nigeria. He lectures at the Surveying and Geoinformatics Department, Federal University of Technology, Minna, Nigeria. He is an expert in GIS and Remote Sensing applications.

Oluibukun Gbenga AJAYI is a First Class Graduate of The Department of Surveying and Geoinformatics, Federal University of Technology, Minna. He holds a M.Sc. degree (Distinction) also in Surveying and Geoinformatics from the University of Lagos, Nigeria. He currently lectures in the Surveying and Geoinformatics Department, Federal University of Technology, Minna. His research interest finds expression in GIS, Remote Sensing, and Digital Photogrammetry

Ekundayo. A. ADESINA is currently a lecturer in Department of Surveying and Geoinfromatics, Federal University of Technology, Minna, Nigeria. He obtained his Master of Science (M.Sc.) degree in Surveying and Geoinformatics (2015) from University of Lagos, Nigeria. His research interest is GIS, Remote Sensing and Hydrographic Surveying.

Nanpon ZITTA: A graduate of Surveying and Geoinformatics with first class honours from F.U.T. Yola (2007) and M.Tech surveying and Geoinformatics from MAUTECH Yola (2012). Presently a

lecturer with the Department of Surveying and Geoinformatics F.U.T. Minna Niger State. His research interests are GIS and Remote Sensing applications and field astronomy.

Hassan A. SAMAIL-IJA is currently a lecturer in Department of Surveying and Geoinformatics, Federal University of Technology, Minna, Nigeria. His area of specializations are Remote sensing, GIS, hydrographic and Cadastral Surveying.

CONTACTS

Name: Abdullahi. Ahmed KUTA
Institution: Federal University of Technology, Minna
Address: Surveying and Geoinformatics Department, Federal University of Technology, Minna.
City: Minna, Niger-State
Country: Nigeria
Tel. +2348055613228
Email: abdullahi_kuta@yahoo.com

Name: Oluibukun Gbenga, AJAYI
Institution: Federal University of Technology, Minna
Address: Surveying and Geoinformatics Department, Federal University of Technology, Minna.
City: Minna, Niger-State
Country: Nigeria
Tel. +2348036832797
Email: ogbajayi@gmail.com

Name: Ekundayo A. ADESINA
Institution: Federal University of Technology, Minna.
Address: Surveying and Geoinformatics Department, Federal University of Technology, Minna.
City: Minna, Niger-State
Country: Nigeria
Tel. +2348038095866
Email: adesinadon2002@yahoo.co.uk

Name: Nanpon, ZITTA
Institution: Federal University of Technology, Minna
Address: Surveying and Geoinformatics Department, Federal University of Technology, Minna.
City: Minna, Niger-State
Country: Nigeria
Tel. +2347030061896
Email: bawazitta@gmail.com

Name: Hassan .A. SAMAILA-IJA
Institution: Federal University of Technology, Minna
Address: Surveying and Geoinformatics Department, Federal University of Technology, Minna.
City: Minna, Niger-State
Country: Nigeria
Tel. +2348035996014
