LANDUSE AND LAND COVER MAPPING OF THE SIMIYU CATCHMENT (TANZANIA) USING REMOTE SENSING TECHNIQUES

Justus RWETABULA, Florimond DE SMEDT (Vrije Universiteit Brussel, Belgium)

Key words: Simiyu catchment, Satellite images, Landuse/land cover, GIS

Abstract

A process of integrating remote sensing techniques and field data to accurately map landuse and land cover of the Simiyu catchment is described. Satellite images of LANDSAT 7 ETM+ of 3/4/2001 and 12/5/2001 were processed and registered using Idrisi32 release 2 image processing software and topographical maps, and enhanced for better interpretation. Images were interpreted using frequency histograms of each band, colour composite images, principal component images, and prior knowledge of land cover. Additionally, training data for supervised classification were collected in the study area in the year 2002 and 2003. Signature development was carried out and evaluated. Training sites were re-defined such that significant separability was obtained for all six bands of LANDSAT 7 EMT+. Finally, a maximum likelihood classifier was applied to classify the satellite images. Six major landuse classes were identified and mapped for the Simiyu catchment. These are: mixed short grasses with/or bare land, dense tall grassland, bush land, cultivated land, medium size grassland, and surface water.

1. INTRODUCTION

Trends of increasing contaminant (phosphorous and pesticides) concentrations in rivers such as Simiyu river have been observed in the recent past (Rwetabula et al., 2004). These contaminants possibly originate from agricultural fields where, they are applied and transported to surface water in either dissolved form or associated with sediments. Better estimations (predictions) of surface water quantity and quality in the Simiyu catchment cannot be performed without clear understanding of the landuse/land cover of the whole catchment.

Many studies which have been conducted on mapping landuse and land cover of Africa including Tanzania (Hunting Technical Services, 1997 and Mayaux et al., 2003) do not focus on pollution issues, and cover big areas such that one cannot extract useful information on particular smaller areas of interest. Additionally, most studies are old and do not give reliable information of the Simiyu catchment at present times.

A process of integrating remote sensing techniques and field data to accurately map landuse and land cover of the Simiyu catchment is described. Satellite images of 2001 were processed and classified using Idrisi32 release 2 image processing software and supervised classification techniques.

2. STUDY AREA AND DATA ACQUISITION

2.1 Study area

The Simiyu catchment is located in Tanzania southeast of Lake Victoria. (Fig. 1). The catchment has a drainage area of about 10,800 km^2 and is mainly covered by agricultural landuse for farming, grassland for grazing and bushland. The southeastern part is covered by the Serengeti game reserve. Short rainfall appears mainly in November and December and long rains in March to May resulting in a total annual rainfall of 700 to 1000 mm.

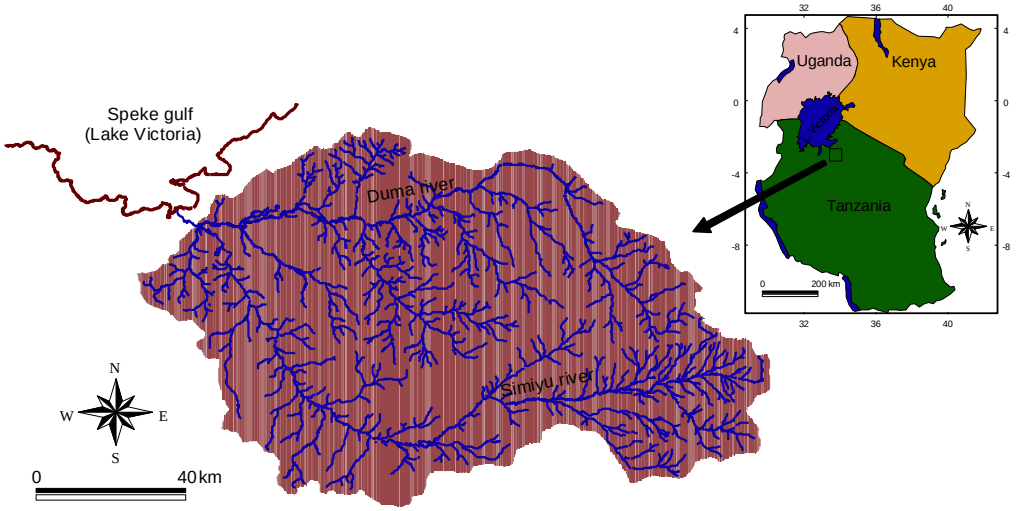

Figure 1. Simiyu catchment, Tanzania, East Africa

The catchment has a warm tropical savannah climate with an average temperature of about 23^{0}C. Sandy loam soil covers about 60% of the total catchment area (Rwetabula et al., 2004). The headwaters of the Simiyu catchment start at elevation of 1640 m and descend to 1140 m at Speke gulf of Lake Victoria. The most frequently grown crops are cotton and maize. Due to rainfall uncertainties most cropping systems practiced in this area are staggered planting (planting crops on different dates), intercropping and flat cropping system (Meertens and Lupeja, 1996). Communal grazing is the most practiced livestock management system in the Simiyu catchment.

2.2 Satellite images data acquisition

Satellite images of LANDSAT 7 ETM+ (Enhanced thematic mapper plus) were ordered from US Geological Survey in consideration of cost, coverage, cloud cover, resolution and product format. The images were acquired in two scenes (each with eight bands) 170/062 (path/row) of 3/4/2001 and 169/062 of 12/5/2001 covering the whole area of the Simiyu catchment. The pixel size are 28.5 m x 28.5 m for bands 1, 2, 3, 4, 5 and 7, 57 m x 57 m for thermal band 6 and 14.25 m x 14.25 m for panchromatic band 8. All are in GeoTIFF format. The overall image quality of both scenes is 99%.

3. DIGITAL IMAGE PROCESSING

3.1 Geometric correction and accuracy assessment

Digital image processing followed procedures described in Idrisi32 release 2 software manual (CLARK LABS, 2001) and other text books (Campell, 1996; Lillesand and Kiefer, 2000). It involves image restoration (geometric correction), image enhancement and image classification.

Geometric correction of LANDSAT 7 ETM+ scenes were carried out using ground control points (GCPs) extracted from 42 topographical maps of scale of 1:50,000 and the corresponding UTM system**.** The selection of GCPs was done considering points like road junctions, bridges and other well defined man-made features (Roy et al., 1991; Lillesand and Kiefer, 2000). Sixty GCPs were used for the first scene and fourty four for the second scene. Registration was done separately for each scene but ten common control points were incorporated in both scenes.

A linear transformation and elimination of undesirable control points were carried out systematically by keeping a close watch on the residual errors, which were limited to 0.5 pixel scene (CLARK LABS, 2001). The first order transformation model (Deppe, 1998) obtained from ten control points for scene 1 and six for scene 2 were finally considered for image rectification. The obtained RMS values were 0.44 pixel (12.5 m) for scene 1 and 1.3 pixel (37 m) for scene 2. The slightly larger RMS with second scene can possibly be attributed to the absence of definite features (Gao, 1997) as most of ground control points in the national park of Serengeti were taken from old topographical maps of 1975.

The validation of the registered images was assessed by selecting a few but well defined points (Gao, 1997) from the rectified images and by calculating the corresponding map points (excluding the previous GCPs) with the transformation model and vice versa. The estimated points in both cases were within the obtained RMS values of 0.44 pixel for scene 1 and 1.3 pixel for 169/062 scene 2. To increase the processing speed and minimize the storage capacity, the areas that did not cover the catchment under study were windowed (cut) from both scenes.

3.2 Image enhancement

An attempt was made to look at the differences among the scenes acquired on different dates. The most used and proved to be most useful in vegetation interpretation, a false colour composite image of bands 3, 4, and 5 was used. This combination contains one band from each of the three spectral zones: the visible (bands 1, 2, and 3), near-infrared (band 4) and mid-infrared (bands 5 and 7). Insignificant differences were observed among them, which is probably due to the small time span between the two scenes.

The image bands of both scenes were joined (mosaicked) together and processed as one single images. Table 1 and Figure 2 show statistical data and frequency histograms of the mosaicked images. The thermal band and panchromatic band were not used as suggested by Batelaan (1994), because they possess different spatial resolutions. The mid-infrared bands (band 5 and 7) have a relatively very high standard deviation, followed by the near-infrared band (band 4), which is probably due to highly vigorous reflecting bare soils and green vegetations (Haack and Jampoler, 1995).

Table 1. Statistics of the LANDSAT 7 ETM+ mosaicked bands

Band	Wavelength (μm)	Energy type	Range	Mean value	Standard deviation
1	0.45-0.515	Reflected blue	46-255	68	10.6
2	0.525-0.605	Reflected green	28-255	59.6	12.1
3	0.630-0.690	Reflected red	19-255	59.5	18.5
4	0.750-0.900	Reflected near -IR	7-227	70.4	19.4
5	1.55-1.75	Reflected mid-IR	5-255	100.1	31.9
7	2.09-2.35	Reflected mid-IR	2-255	63.2	26

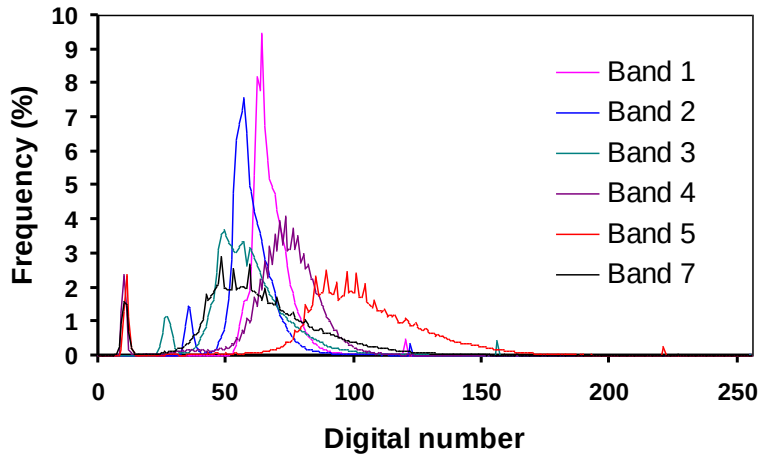

Figure 2. Frequency histograms of the mosaicked subscene images

Again, images enhancements and colour composite of bands 3, 4, and 5 were performed for better interpretation of landuse and land cover of the Simiyu catchment. First the bands were stretched before being combined in colour. Also a true colour composite of bands 1, 2 and 3, was made although this contains correlated images (Beaubien, 1994) but gives the colors that the human eye would see. However, using those combinations we could not identify all landuse and land cover types of the Simiyu catchment.

Therefore, a principal component analysis (PCA) was carried out to identify band images which carry the most genuine information about the full band set for digital analysis and/or color compositing (Chaves and Kwarteng, 1989; Beaubien, 1994; CLARK LABS, 2001). PCA is related to factor analysis and can be used to transform a set of image bands such that the new bands (called components) are uncorrelated with one other and are ordered in terms of the amount of image variation they can explain (CLARK LABS, 2001). The degree of correlation between two images is related to the amount of spectral contrast. The higher the correlation the, less the contrast and vice versa (Chaves and Kwarteng, 1989). In PCA an orthogonal transformation of the coordinate axes of the multivariate system to new orientations is performed in such a way that small number of components (smaller than the original dimension) explains as much as possible the variance in the original data. The eigenvalues express the amount of variance explained by each component and the eigenvectors are the transformation equations. The factor loadings refer to the degree of correlation between components and original bands. Also, the factor loadings give the contribution of the original bands to the newly transformed axes**.** In general the first three

components explain more than 90% of the variance in the original set of bands (Batelaan, 1994; Haack and Jampoler, 1995; CLARK LABS, 2001).

Principle components analysis has been applied by many researchers in classification of satellite images to identify water, vegetations and soil features (Chaves and Kwarteng, 1989; Batelaan, 1994; Beaubien, 1994; Haack and Jampoler, 1995), because it reduces highly correlated data with a minimum loss of information and reduces atmospheric noise in the images (Sabins, 1986; Batelaan, 1994; Haack, 1995). Caution should be taken for the use of selective images as the first principal component image because information that is not mapped to the selected images can be of significant interest (Chaves and Kwarteng, 1989).

Table 2a. Correlation matrix of the standardized LANDSAT 7 ETM+ bands of the mosaicked subscene images

	Spectral band					
	1	**2**	**3**	**4**	**5**	**7**
1	1.00					
2	0.94	1.00				
3	0.90	0.95	1.00			
4	0.41	0.62	0.50	1.00		
5	0.67	0.82	0.85	0.71	1.00	
7	0.74	0.83	0.90	0.54	0.95	1.00

Table 2b. Principle components with explained variances eigenvalues and eigenvectors

	Components					
	1	**2**	**3**	**4**	**5**	**6**
%var	**80.57**	**11.98**	**6.08**	**0.75**	**0.39**	**0.23**
eigval	**4.83**	**0.72**	**0.36**	**0.05**	**0.02**	**0.01**
eigv 1	0.40	-0.44	-0.44	-0.54	0.01	-0.39
eigv 2	0.44	-0.14	-0.34	0.17	-0.26	0.76
eigv 3	0.44	-0.25	0.05	0.75	0.20	-0.37
eigv 4	0.31	0.81	-0.40	-0.01	0.26	-0.11
eigv 5	0.42	0.25	0.44	-0.11	-0.71	-0.21
eigv 6	0.42	-0.03	0.57	-0.31	0.57	0.27

Table 2c. Principle component factor loadings

	Components					
	1	**2**	**3**	**4**	**5**	**6**
Band 1	0.88	-0.37	-0.27	-0.12	-0.001	-0.05
Band 2	0.97	-0.12	-0.21	0.04	-0.04	0.09
Band 3	0.96	-0.21	0.03	0.16	0.03	-0.04
Band 4	0.68	0.69	-0.24	-0.002	0.04	-0.01
Band 5	0.93	0.21	0.27	-0.02	-0.11	-0.03
Band 7	0.93	-0.03	0.35	-0.07	0.09	0.03

Two types of PCA are offered in Idrisi32 release 2 image processing software: standardized and unstandardized. In standardized PCA all bands have equal weight while with in unstandardized PCA, those bands with a greater variance, usually the infrared bands will have a greater weight. In this study standardized PCA was opted since gives significant improvement in image enhancement (Singh and Harrison, 1985).

Table 2 gives detailed results of the PCA parameters: (a) the correlation matrix, (b) the principle components with explained variances eigenvalues and eigenvectors, and (c) the principle components factor loadings. Table 2a shows a very high positive correlation between the three images in the visible spectrum (bands 1, 2, and 3) ranging from 90 to 95%, between the two images in the middle infrared spectrum (band 5 and 7) of 95%, but also high correlation is observed between visible spectrum (i.e. band 3) and middle infrared spectrum (i.e. band 5 and band 7) with correlation values of 85% and 90%. Generally, a high correlation is observed between bands 1, 2, 3, 5 and 7, while band 4 is less correlated. In the present case we see that band 4 is still reasonably correlated to other bands, could be due to scattered reflectance of green vegetation or less reflectance of green vegetation (Hung, 2001). Also, the relatively high correlation between all image bands indicates that there is little contrast between them (Chaves and Kwarteng, 1989).

Table 2b gives the results of PCA. The first component explains 80.6%, the second component 12% and the third component 6.1%. Components 4, 5 and 6 in total explain only 1.4%. Components 1, 2 and 3 together explain almost 98.7% of the total information in the data set generated by the LANDSAT 7 ETM+ images. This reveals that all information is explained by the first three principal components, the highest component 1 followed by components 2 and 3. Also in the same table it can be concluded that the variance of principal component 1 is mainly explained by the eigenvectors 2 and 3 (band 2 and band 3) followed by eigenvector 5 and 6 (band 5 and band 7). The variance in principal component 2 is strongly related to eigenvector 4 (band 4). The variance in principal component 3 has no clear relation with a particular band but the most influential are bands 5 and 7 (eigenvector 5 and 6).

The loadings of the original bands on the newly derived components are explained by Table 2c. It is clear that principal component number one is highly contributed by the visible bands 2 and 3 followed by the mid infrared bands 5 and 7, and lastly by visible band 1 and the near infrared band 4. Principal component number two is highly contributed by band 4, and principal component number three is highly contributed by bands 5 and 7. This suggests that an analyst wishing to reduce the number of bands while maintaining good information, could use one band from visible spectrum, one from the near infrared, and one from the mid infrared, since all contribute their information in the first three principal components. Three bands 3, 4, and 5 would be a good combination to discriminate and interpret digital images containing water, vegetation and soil (Janssen et al., 1990; Roy et al., 1991; Hass, 1992; Beaubien, 1994). Also, a false colour composite of bands 3, 4, and 7, the colour composite of the first three components (Chaves and Kwarteng, 1989; Haack and Jampoler, 1995), or the use of the first principal component number one (Batelaan, 1994) could be an option.

Interpretation of colour composite images of bands 3, 4 and 5, and especially the first principal component image were performed, but not all landuse categories could be identified. This is not uncommon for catchments containing small fields, intermixed with savanna and fallows (Mayaux et al., 2003). Hence, after this failure to identity all features of the study area by colour composite images and PCA it was decided that, the only solution was to use the field data to classify the satellite images to map the landuse of the Simiyu catchment.

Table 3. Predefined landuse/landcover categories of the Simiyu catchment

Land use/ land cover category	Description (general observation, possible use of fertilizer and pesticides)	Identified landuse/ land cover type features in the field
1	Mainly observed in the game reserve and national park. These are mixed dense shrubs and some dense grasses with little woods. No application of fertilizers and pesticides. Generally no cultivation.	Bushland
2	Mainly observed in the game reserve followed by agricultural fields but not in extensive cultivated land. Most present in Duma subcatchment. These are matured, dense and tall grasses. No fertilizers and pesticides applications. Also no cultivation. *This category is very close to category 9 and 10.*	Grassland (type 1) matured and tall
3	Mainly observed in agricultural fields with mixed cropping. These are cultivated and harvested areas. Application of pesticides and fertilizers is possible. Soil disturbed for growing non rain season crops like peas.	Cultivated land (type 1)
4	Mainly observed in the national park. They are grasses with medium height, somehow open (less dense) and look like dry grasses. Application of fertilizers and pesticides is negligible.	Grassland (type 2) (medium size)
5	Surface water (lakes and reservoirs).	Surface water
6	Mainly observed in agricultural fields. These are cultivated areas, sometimes harvested and disturbed for the next plantings. Soil disturbed for growing non rain season crops like peas and some are very young crops. Application of pesticides and fertilizers is possible. *This category is very close to category number 3.*	Cultivated land (type 2)
7	Mostly observed in agricultural fields, but in non extensive cultivated land. Fertilizers and pesticides application is less significant unless mixed with crop area. *This category is very close to category number 8.*	Grassland (type 3) (Short grasses mixed with bare soil)
8	Mostly observed in agricultural fields in non extensive cultivated land. These are white, thin and short grasses. Appear during rain season and disappear in dry season leaving land bare. Fertilizers and pesticides application is less significant unless mixed with crop area. *This category is very close to category number 7.*	Grassland (type 4) (White and somehow short)
9	Mainly observed in non extensive cultivated land and in the game reserve, and very close to river channels (i.e. Duma river). These are tall, dense, and greenish grasses, also observed in seasonal wetlands. Fertilizers and pesticides application is negligible. *This category is very close to category number 2 and 10.*	Grassland (type 5) (Dense and tall, in seasonal wet areas)
10	Mainly observed in non extensive cultivated land and in the game reserve. These are, tall and dense grasses, but more greenish, ere also observed in areas very near to seasonal wetlands. *It is very close to category number 2 and 9*	Grassland (type 6) (Dense and tall, near seasonal wet areas)
11	Mainly observed just outside the study area (Simiyu catchments) and not in the game reserve. These were in degraded lands, sloppy and sometimes with very short grasses. Fertilizers and pesticides application is less significant unless mixed with crop area. They are mainly bare soils. *It is very close to category number 7 and 8.*	Bare land

Before the start of the field campaign an identification of all possible landuse/land cover categories existing in the Simiyu catchment was made, using the frequency histograms of the six LANDSAT 7 ETM+ bands (Fig. 2), the colour composite images of bands 3, 4, and 5, the principal component images, and prior knowledge of the landuse in general. From the, frequency histograms and especially the peaks in bands 4, 5 and 7 different categories could be identified. These were superimposed on the colour composite and PCA images and based on the general known features of the area enabled to predefine eleven dominant landuse/land cover categories in the Simiyu catchment, as described in Table 3.

3.3 Training sites selection and field data collection

Training sites were selected based on pre-identified landuse/ land cover categories (Table 3) and colour composite images of bands 3, 4 and 5, the true colour composite image of bands 1, 2 and 3 and the image of principal component number one. The training site selection criteria were areas: (a) with extensive (large coverage) landuse/cover categories and (b) accessibility. Other areas which did not cover a big area but were among the predefined identified landuse categories were also included in the training sites.

A bout twenty location for each intended landuse/land cover categories were identified as "training data set" for supervised classification. These locations were distributed throughout the study area. The sizes of sites varied, ranging from meters to the order of kilometers, and provide a total of at least 100 pixels for each category as prescribed by practitioners (Campell, 1996; CLARK LABS, 2001).

Ground truthing was conducted in the year 2002 and 2003. Site visits were mostly done tallying with acquired months of the satellite images. In very rare cases, due to transport problems and accessibility, other areas were visited on different dates. Also, due to security matters (in the game reserve) some training sites were not reached. In such circumstances game officers were consulted for clarification on the features covering those areas. Training sites were identified and earmarked by taking notes, size and reference location (Campell, 1996), and digitized and stored as line vectors. Only homogenous training sites were stored for further classification. Ten to twenty training sites were visited for each landuse/land cover category in the study area. Due to varying cropping system existing in the catchment, like mixed cropping, we could not manage to define crop types individually. The last column of Table 3 describes the landuse/land cover features identified during the field work.

4. DIGITAL IMAGES CLASSIFICATION

4.1 Introduction

Digital image classification procedures were applied to the mosaicked images of the LANDSAT 7 ETM+ data using the supervised classification method (Deppe, 1998; Congalton et al. 1998). The supervised classification method requires the user to develop the spectra signatures of known categories, while the software assigns to each pixel in the image a category to which its signature is most similar. Contrary to unsupervised classification, in supervised classification the user has some control, which is artful especially if the classification is for a specific purpose, or else is tied to specific areas of known identity, known training data selected by the analyst may not be representative of all conditions encountered throughout the images. Hence, the classification may not be able to recognize

and represent special or unique categories that are not represented in the training data, due to the fact that they are not known by the analyst or occupy a very small area on the image under study (Campell, 1996).

The maximum likelihood algorithm with equal probability of occurrence was (Janssen et al., 1990; Roy et al., 1991; Deppe, 1998; Congalton et al. 1998) used in the classification exercise, because it is powerful and produces the best results if defined training sites are very good (particularly their uniformity). This method, however, needs more computations to classify each pixel than other methods like the minimum distance to means and parallelepiped classifiers (Campell, 1996; Lillesand and Kiefer, 2000; CLARK LABS, 2001). In the maximum likelihood method the distribution of reflectance values in a training site is described by a probability density function, developed on the basis of Bayesian statistics. This method uses the training data as means of estimating means and variances of the classes, which are then used to estimate the probabilities. This stage involves with training site development, signature development and correction (redefining training sites) and classification.

4.2 Training site and signature development

A false colour composite of bands 3, 4 and 5 was used to identify possible training sites, guided by the former predefined training sites. The training sites were digitized and saved as vector polygons while maintaining the key characteristics of the training sites, such as uniformity or homogeneity as much as possible, avoiding mixed pixels, ensuring at least 100 pixels for each landuse/land cover category, and evenly distributed over the study area (Campell, 1996; CLARK LABS, 2001).

Signature development was carried out using the MAKESIG module in Idrisi32 release 2 image processing software for each landuse/land cover category of the defined training sites. In this process signature files containing statistical information about the reflectance values of the pixels within the training sites for each class or landuse /land cover category are created, using the six bands of LANDSAT 7 ETM+ scene. Signature evaluations were made using the signature compare module (SIGCOMP) in a graphic output showing the means, minimum and maximum reflectance values for each band of the defined training sites. Using these plots similar signatures and overlapping signatures for each band were detected. A clear category separations (looking at the distance between the means), were observed in the red band 3, near infrared band 4 and middle infrared bands 5 and 7, but not in the blue band 1 and green band 2.

With maximum likelihood classification procedures, means plot that are nearly coincident are difficult to distinguish (CLARK LABS, 2001), therefore it is necessary to redefine the training sites to the extent that mean of the signatures become significantly separable. A review was made to refine the training sites such that significant separability is attained for all six bands of LANDSAT 7 EMT+. The criterion was to avoid or minimize overlapping signatures while maintaining uniformity of the training sites.

The HISTO module was used to the assess uniformity of the training sites for each band used to the create signatures. In this module frequency histograms of cell values in signature file are produced, and also statistics, such as the mean, maximum and minimum, and standard deviation. Histograms of most landuse/land cover categories showed significant uniformity. Those which exhibited bimodal histograms were either discarded from training site data set, or redefined to have an unimodal frequency distribution (Campell, 1996).

Overlapping signatures were investigated using statistical data provided by HISTO module, and box plots resulting from the SIGCOMP module. To meet the prerequisites of the maximum likelihood classifier of uniformity and non overlapping signatures, several iterations of training sites adjustments, signature development and signature evaluations were carried out until achieving satisfactory separable signatures. The training sites were redefined, by replacing bad training sites with good ones and merging overlapping training sites into one class (Congalton et al., 1998; Lillesand and Kiefer, 2000). The signature comparison of the final step of the signature development is shown in Figure 3.

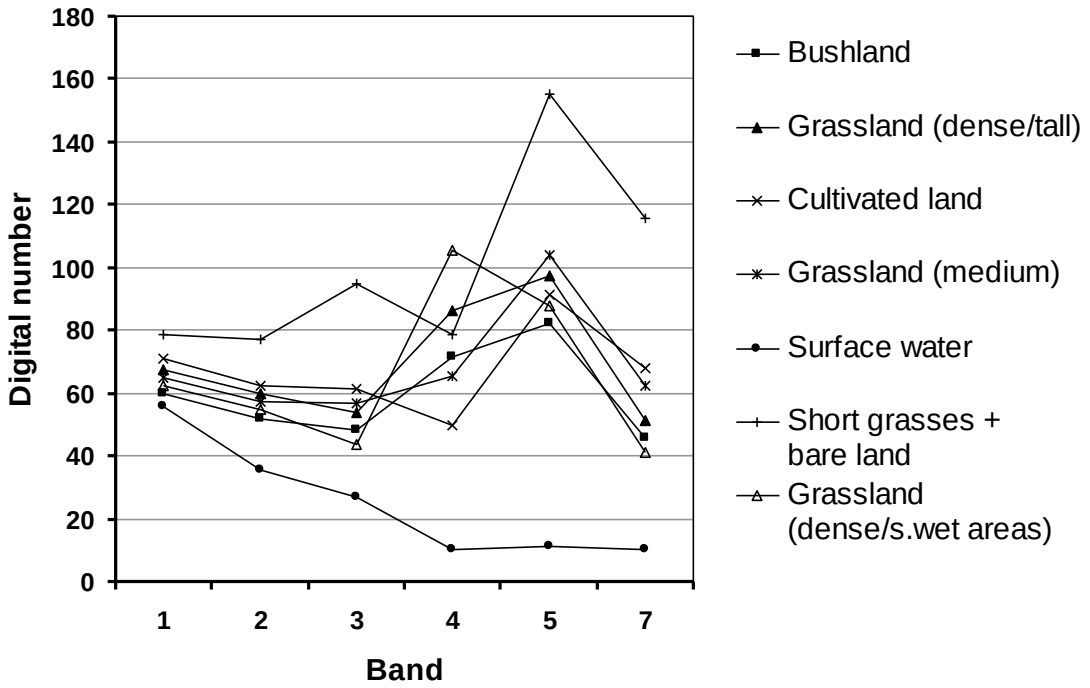

Figure 3. Signature comparison plots of the means of the selected training sites.

Generally, a significant separation of the means is observed in the all bands of LANDSAT 7 EMT+ data except for bands 1 and 2. The best separability was achieved in band 4. All bands were incorporated in the final stage of the classification even, band 1 and band 2 to avoid loosing information which is not mapped in the other bands.

4.3 Classification

Maximum likelihood classification was performed using the developed signatures of the landuse/land cover categories of bushland, dense and tall grassland, cultivated land, medium grassland, surface water, mixed short grasses with bare land and dense grassland in seasonal wet areas. Other methods like minimum distance to means and parallepiped classfier were also tested but their results were not any more convincing. Although, images were generally cloud free, some pixels were affected by shading; these were replaced by the nearest as it was observed in the field (Ryerson, 1985). Lastly classified image was exported to Arc View GIS software for better presentation. Figure 4 shows the classified landuse/land cover map of the Simiyu catchment obtained from the LANDSAT 7 ETM+ satellite images of 2001.

The final figure is shown in Figure 4. The main landuse/land cover of the Simiyu catchment consists of mixed bare land and short grasses (46.5%), dense and tall grassland (21.5%), bushland (19.7%), cultivated land (8.3 %) and medium size grassland (4%). A very small area (0.02%) is covered by surface water (Fig. 4). Generally, the results are not far from those obtained by (Hunting Technical Services, 1997), which, however were presented in a different way, possibly due to particular interests.

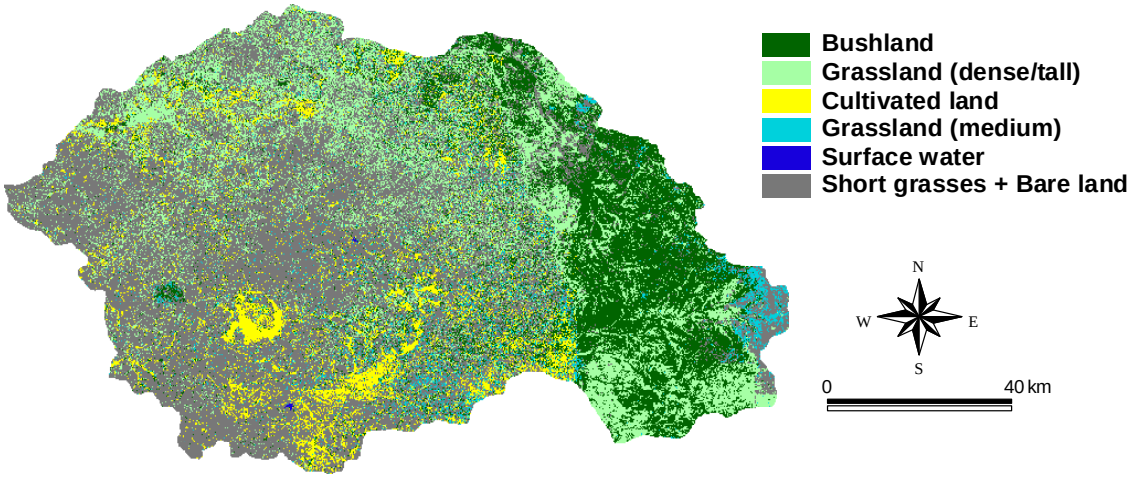

Figure 4. Landuse/Land cover map of the Simiyu catchment, Tanzania

5. CONCLUSIONS

This study combined remote sensing, Idrisi32 release 2 image processing software, GIS and extensive and detailed ground information to map landuse and land cover of the Simiyu catchment. Landuse and land cover is extremely valuable especially for water quantity and water quality predictions, and assessing the hydrological effects of landuse changes. Satellite images data acquisition, digital image processing, training sites collection, digital image classification are the processes involved to map the landuse in of the Simiyu catchment.

Image enhancement was applied to increase the contrast of the satellite images using the false colour composite image of bands 3, 4, and 5, and principal component images we could not identify all features existing in the Simiyu catchment. This is probably due to small fields intermixed and scattered in the catchment. Using the frequency histograms of the bands and prior general knowledge of the landuse together with the first principal image and false colour composite image of bands 3, 4 and 5 features existing in the catchment could be predefined and mapped. This was used to set up a field campaign to map the training sites for supervised classification.

The field measurements were conducted in the Simiyu catchments in 2002 and 2003 and training site data for supervised classification procedures were identified and mapped. Afterwards the training sites were revised based on statistical data of the histograms of the signatures, to obtain better separation between the signatures. Maximum likelihood classifier was used to obtain the final landuse map of the catchment.

The Simiyu catchment is by far dominated by mixed short grasses and open land (46.5%). This suggests that the Simiyu catchment is highly subjected to erosion.

As such, it is shown that investigation of frequency histograms of the satellite images data set and prior knowledge of general catchment features are promising procedures to identify possible training sites. Using these well defined training sites, supervised classification procedures and maximum likelihood classifier prove to be useful in the classification of the satellite images to obtain a reasonable map of the landuse/land cover of the Simiyu catchment.

6. RECOMMENDATIONS

The investigation of the use of remote sensing to estimate landuse/land cover in the Simiyu area is still open. Possible improvement and subjects for future studies are: the use of various sets of satellite images covering all seasons, registration of images using recent topographical maps or GPS, more training sites that are uniformly distributed to obtain great separability in signature development procedures, to discriminate crop types from cultivated land or if exists from mixed short grasses and open land. The findings of this study demonstrate the potential use and advantages of using LANDAST 7ETM+ and establish a baseline to carry out further investigations.

REFERENCES

Batelaan, O. and F., De Smedt, 1994. Use of Landsat TM in the analysis of groundwater flow systems. *Space scientific research in Belgium, Brussels. III, Earth observation, part 2: 215-233.*

Beaubien, J., 1994. Landsat TM satellite images of forests: From enhancement to classification. *Canadian Journal of Remote Sensing. 20 (1): 12-26.*

Campell, J.B., 1996. *Introduction to remote sensing*. The Guilford Press. London. 622pp.

Chavez, Pat S. J. and A.Y., Kwarteng, 1989. Extracting spectral contrast in Landsat Thematic Mapper Image data using selective principal component analysis. *American Society for Photogrammetry and Remote Sensing. 55(3):339-348.*

CLARK LABS, 2001. *Idrisi32, Release 2, tutorial manual.* Idrisi production, USA.

Congalton, R.G., et al., 1998. Mapping and monitoring agricultural crops and other land cover in the lower Colorado river basin. *American Society for Photogrammetry and Remote Sensing. 64(11): 1107-1113.*

Deppe, F., 1998. Forest area estimation using sample surveys and Landsat MSS and TM data *Photogrammetric Engineering and Remote Sensing. 64 (4): 285-292*

Fang, H., 1998. Rice crop area estimation of an administrative division in China using remote sensing data. *Int. J. Remote Sensing. 19(17): 3411-3419.*

Gao, J. and S.M., O'leary, 1997. Estimation of suspended solids from aerial photographs in a GIS: *Int. J. Remote Sensing. 18 (10):* 2073-2086.

Haack, B. and S., Jampoler, 1995. Colour composite comparisons for agricultural assessments *Int. J. Remote Sensing. 16 (9): 1589-1598.*

Hass, R. H., 1992. Landsat Thematic Mapper products for rangeland assessment. *Geocarto International. 7: 27-33.*

Hunting Technical Service, 1997. *Forest resources management project: National reconnaissance level landuse and natural resources mapping project*. Tanzania: 275pp.

Hung, L. Q., 2001. *Remote sensing based hydrogeological analysis of the Suoimuioi catchment, Vietnam.* Msc, thesis, Interuniversity Programme in Water Resources Engineering, Vrije Universiteit Brussel: 87pp.

Janssen, L. L. F., 1990. Integrating topographical data with remote sensing for land cover classification. *American Society for Photogrammetry and Remote Sensing. 56(11):1503-1506.*

Lillesand, T.M. and R.W., Kiefer, 2000. *Remote sensing and image interpretation*. John Wiley & Sons, Inc. New York. 724pp.

Mayaux, P., 2003. A land cover map of Africa. European Commission Joint Research Centre. EUR 20665 EN. *At http://www.gvm.jrc.it/glc2000/publications.htm.*

Meertens, H.C.C. and Lupeja, P.M., 1996. *A collection of agricultural background information for Mwanza region. Kilimo/FAO plant nutrition programme in Tanzania*. Field document No. GCPF/URT/106/NET, Mwanza Tanzania: 74pp.

Roy, P.S., 1991. Tropical forest type mapping and monitoring using remote sensing. *Int. J. Remote Sensing. 12 (11): 2205-2225.*

Rwetabula, J., F. De Smedt, M. Rebhun and F. Mwanuzi, 2004. Transport of micropollutants and phosphates in the Simiyu river (tributary of Lake Victoria), Tanzania. *Submitted and presented at The 1st International Conference on Environmental Science and Technology, New Orleans, Louisiana, USA January 23-26th, 2005.*

Ryerson, R.A., et al., 1985. Timely crop area estimates from Landsat. *Photogrammetric Photogrammetric Engineering and Remote Sensing. 51(11):1735-1743.*

Sabins, J.F.F, 1986. *Remote sensing: Principles and interpretation*. Freeman and Company, New York.

Singh, A. and A., Harrison, 1985. Standardized principal components. *Int. J. Remote Sensing. 6 (6):883-896.*

ACKNOWLEDGEMENTS

The authors acknowledge Lake Victoria Environmental Management Project, A World Bank Project (LVEMP), for the support of this research