MAPPING THE FOREST COVER OF UGANDA WITH SPOT (XS) AND LANDSAT (ETM+) IMAGES:

(A case study of Bwindi Impenetrable National Park, Uganda)

BY

JOHN RICHARD OTUKEI

Thesis Submitted to the University of Cape Town in fulfillment of the requirements for the a ward of the Degree of Master of Science in Engineering

Prof. Heinz Ruther
(First Supervisor)

Dr. Yafesi Okia
(Co-Supervisor)

SCHOOL OF ARCHITECTURE, PLANNING AND GEOMATICS,
UNIVERSITY OF CAPE TOWN,
CAPE TOWN-SOUTH AFRICA.

August 2004

DECLARATION

I hereby declare that this is my original work, and where appropriate I have acknowledged the work of others to the best of my knowledge. The thesis describes the work undertaken for a study leading to the award of Master of Science in Engineering at the University of Cape Town (UCT) and has not been submitted in any form to another University.

Otukei John Richard
Cape Town
August, 2004

ABSTRACT

It is well established that land cover information is an essential component in the creation of spatial information systems. Lack of current land cover information constitutes a weakness in land resource management especially in developing countries like Uganda. In response to this need, the thesis reports on a case study on tropical forest mapping in Uganda. The geographic area of study is the Bwindi Impenetrable National Park located in the southwest of the country.

Digital image processing techniques were applied to SPOT and Landsat Imagery using Erdas Imagine (8.4) and Arc View GIS software. A combination of original and derived bands (Principal Components, Tasseled Cap and Texture Analysis) was used for the image analysis. Both supervised and unsupervised classification approaches were used. The optimal combination of bands was selected on the basis of secondary correlation analysis of the derived as well as original bands. Further identification of the best bands was based on separability indices.

With the band combination selected, four main land cover classes were identified in the forest i.e. dense evergreen forest, evergreen forest, mixed rangeland and brush land. In addition to this, three other land use types were extracted from the imagery within the neighborhood of the forest and these were subsistence farmlands, plantation farmlands (tea plantation) and woodland. The results were confirmed by post processing field inspection.

Supervised classification was performed with two classification algorithms: A maximum Likelihood Classifier (MLC) and a combination of MLC and Parallelepiped classifier. The results of accuracy assessment showed that MLC generally gave better accuracies and kappa statistics than the combination of MLC and Parallelepiped classifier

ACKNOWLEDGEMENTS

I wish to register my sincere gratitude to my principal supervisor, Prof. Dr. Heinz Ruther for guidance, valuable suggestions and constructive criticism throughout this research. Further thanks go to Dr. Yafesi Okia, co-supervisor, who gave useful ideas during the course of this work.

I am indebted to the USHEPiA funding organization for providing me with funds that enabled me to carry out this research. In particular, I would like to thank Lesley, Nan, Carol and Zubaida for all the support they rendered to me while in Cape Town.

Further thanks go to my colleagues, friends and in the Department of Architecture, Planning and Geomatics at UCT; Anthony, Eric, Ambogo, Ivan, Nasani, Themba, Chedza, Kahilu, Kolade, Annette, Justine. I would also like to thank the staff of department of Geomatics especially Prof. Charles Merry, Solomon, Terry, Jenny, Nick, Sydney, Sandra, Andrea, Peter and all those whose names have not appeared here.

To all USHEPiA colleagues; Joe, Jim, Bagai, Siphila, Fredy, Wayona, Karen, Karanja, May, Purity, Vivian and Goretti, it was nice being with you and I hope to meet you again.

I would also like to recognize the support I got from the Department of Surveys and Mapping for allowing me access to remote sensed images and ancillary data. In this respect, my thanks also go to the management of AFRICOVER mapping organization in Nairobi for the effort they put in to send remote sensed images free of charge.

I would also like to pay tribute to Makerere University for identifying me as a potential candidate for postgraduate studies especially Dr. Badru Kigundu, the former Dean of Faculty of Technology.

Great thanks also go to the management of Uganda Wildlife Authority (UWA), the National Council of Science and Technology (NCST) and the management of Bwindi

Impenetrable Forest for facilitating the research activities in Uganda. Special thanks go to Julius Ecuru from NCST for his support towards the realization of this thesis.

Finally, I would like to thank the entire family of **Amoding Mary Teresa**. Special thanks go to my Mum, without you this thesis would be a myth. To all my brothers and sisters: Betty, Imelda, Mary, Francis, Robert and Thomas, thanks for your support

DEDICATION

I dedicate this thesis to the entire family of Amoding Mary Teresa. Sincerely, you have always been there whenever I call for your support.

"IDARE IYESI LOKASUBAN PAPA OPEDORI KERE"

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 CONTEXT

Land use and land cover (LULC) information has become a central component in current strategies for managing natural resources and monitoring environmental changes. Accurate and up-to-date information on LULC is critical for the effective management of natural resources. This study addresses the problem of land use and land cover mapping of forested areas in Uganda by means of remote sensing technology. The case study area is the Bwindi Impenetrable National Park (BINP), South West Uganda.

According to Sabins (1997) and Lillesand and Kiefer (1999) *land use* describes how the parcel of land is used (such as for agriculture, residences, or industry) and *land cover* describes the materials (such as rocks, forest, buildings) that are present on the surface.

Howard (1991) argues that BINP is the most important forest for conservation of biodiversity in Uganda. Like any other tropical forest, BINP has multiple social values such as soil and watershed protection, climatic modification and carbon storage as well as economic values of numerous products which can be extracted from the forest. Bwindi forest provides the indigenous and other forest dependent people with medicinal plants, bush meat, fruits, honey, shelter, timber and firewood. The BINP is a home to more than half of one of the world's endangered species, the Mountain Gorilla (Figure 1-1)

Figure 1-1 The Mountain Gorilla

Additionally, it is the only place in the world where Mountain Gorillas and the Chimpanzees co-exist. The forest has a variety of species of birds, snakes, butterflies, chameleons and frogs which have promoted the development of the tourism industry in Uganda (Howard, 1991).

In 1932 BINP was designated a forest reserve. At that time much of the forest remained outside the gazetted forest reserve (Hamilton, 1994). Today, there is virtually no forest left outside the boundaries of the gazetted forest reserve. The ever-increasing population around BINP has exerted a lot of pressure on the park in search of the aforementioned forest resources as well need for more land for cultivation.

The population of Kabale, Kanungu and the Rukungiri Districts, where BINP is located, increased by 90% between 1948 and 1980 (Butynski, 1984). Giri et al. (1996) suggest that, whether by overexploitation or mismanagement, changes in land use and land cover will continue taking place at an unprecedented rate.

The conversion of forest cover in general and tropical forests in particular has severe long term environmental and social economic consequences locally and globally. The major environmental problems caused by forest cover change are changes in global climate, habitat degradation and unprecedented species extinction (Goldsmith, 1988).

The increased awareness of environmental issues and the need to strive for the sustainable use of natural resources has focused attention on the need to map and monitor changes in land cover at different temporal and spatial scales. Timely and accurate information on the forest types and areas at regional scale is needed for natural resource management, carbon cycle studies and modeling of biochemistry, hydrology and climate (Xiao et al., 2002). A land use and land cover map is therefore essential for the effective management of natural resources, especially the conservation of biodiversity in BINP, Uganda. Satellite based remote sensing products provide one option to meet these demands.

1.2 Remote sensing and land cover mapping.

Since the late 1960s, the rapid development of the concept of vegetation mapping has led to increased studies of land use and land cover changes worldwide. Aerial photographs have been employed for mapping forest types since 1930 (Boontawee *et al.* 1995). Generally, aerial photographs, satellite images, field surveys and Geographical-Positioning Systems (GPS) are regarded as the main sources of information for the preparation of land cover and land use maps.

Aerial photographs and field surveys other than GPS are considered conventional methods of obtaining information necessary for mapping, with aerial photographs being more popular because of the high spatial resolution, considerable accuracy and speed in comparison to ground survey methods (Lillesand and Kiefer, 1994).These conventional ground methods of land use mapping are labour intensive, time consuming and therefore done relatively infrequently.

Since the early 1970s, satellite remote sensing techniques have been developed which have proved to be of significant value for preparing accurate land cover maps and monitoring changes at regular intervals. Remotely sensed data have been used in forest mapping for inventory or monitoring purposes in many parts of the world (Trisurat et al., 2000).

The use of remote sensing techniques in forest mapping is possible because of the significant differences in responses of forest cover in the red, near-infrared, middle-infrared spectral regions compared with non-forested areas (Trisurat et al., 2000).

Although satellite imagery has been demonstrated to be a cost- effective method for classification of forests all over the world (Schriever and Congalton, 1995), it has not been adopted at an extensive scale in Uganda. Most of the existing forest maps in Uganda were derived from black and white aerial photographs acquired in 1954 by Air Survey Company and after 40 years these maps must be expected to be out of date. The National

Base Mapping Project of Kampala and Jinja Blocks (1994-1998), North of Lake Victoria, was an attempt to update the existing maps but was restricted to large-scale mapping (http://www.worldbank.org/afr/findings/infobeng/infob73.pdf [Accessed: 3rd January 2004].

The Food and Agricultural Organisation (FAO) under the AFRICOVER project have derive regional land use and land cover maps for Uganda using Landsat TM data at the scales of 1:250000 and 1:20000 but no detail mapping was undertaken. The Systeme Probatoire d'Observation de la Terre (SPOT) satellite images were used under the National Biomass study program for mapping selected urban areas in Uganda (National Biomass study Report, 1991) but no attempt was made to carry out detail mapping of the forested areas.

1.3 Problem statement

BINP has fallen under various protected area categories for over half century (Uganda National Parks, 1995). The forest now known as Bwindi (meaning "dark" in the indigenous language) was managed by the British during colonial rule and, later, by the Ugandan authorities. It was designated a forest reserve in 1932 and as both animal sanctuary and forest reserve in 1961. In 1991, the reserve was formally gazetted as the Bwindi Impenetrable National Park (BINP). The change to national status reflected the growing recognition of Bwindi Impenetrable Forest as a vital refuge for some of Uganda's rarest and most unique fauna and flora. In 1994, BINP was declared a natural World Heritage Site (WHS) under UNESCO's World Convention (WC) concerning the protection of the World Cultural and Natural Heritage (WCNH).

The increased legal protection afforded to the forest, however, cannot by itself, ensure the long term security of the park as an enclave for the flora and fauna that it is designated to preserve. The conservation of BINP requires active management that addresses both its biological and physical characteristics, and, perhaps more importantly, the social and cultural context in which it is located.

Like many other national parks in Uganda, BINP has problems associated with its conservation for example conflicts of interest over land usage especially the desire to utilize the park's resources by the local community as they have done historically. There have been management problems including among others the following: poaching, pit sawing (for timber), mining of gold and other minerals, wildfires, livestock grazing, crop raiding by animals, lack of trained personnel, lack of infrastructure and lack of sufficient ecological knowledge to facilitate conservation management. These problems are considered serious constraints in the conservation of the park and many can be addressed fully through analysis of the relationships that exist between the park and its neighboring communities.

Mindful of the above problems existing in BINP, there is a need to obtain and link the data about social-economic as well as biophysical conditions with the human pressure on the forest resources so as to understand why, where and when illegal activities take place, and to predict where they can happen in the future. Remote sensing and Geographical Information System (GIS) technology can be used to supply information as well as analyse the spatial relationships necessary to support the decision-making process. The European Space Agency and European Commission Joint Research Center under the umbrella of Tropical Ecosystem Environment Observation by Satellites (TREES) proposed the application of SPOT and Landsat TM for land forest change assessment (Achard, 2002)

A land cover map is considered to be first on the list of the data sets required for the establishment of GIS source of such information (Apan, 1997). Availability of a land cover map of the BINP is an important element in making informed decisions concerning Earth's resources and environment.

Land use and Land Cover information can be integrated with other data sets, proposed by UNESCO and ESA in the context of Geographical Information Systems. The GIS database can then be used by planners, resource managers, scientists, and decision

makers from the Local Government, State, regions and private co-operations for a variety of purposes. UNESCO and ESA, under the umbrella of, Build Environment for the Gorilla Project (BEGO), have proposed the following data sets required for the establishment of GIS database for BINP and the Virunga National Park in the eastern part of the Democratic Republic of Congo:

- Land use and land cover classification map (containing information like forests, human settlements, water bodies, grassland, agriculture and bare soil).
- Infrastructure (containing information such as roads, large footpaths, park infrastructure, trail networks in the protected areas, airstrips).
- Human activities (such as forest exploitation, mining and agriculture).
- Administrative boundaries (including: National and International boundaries, park limits, hunting areas, local administrative units and adjacent trans-boundary protected areas).
- Toponyms such as main cities (if applicable), towns and names of local places).
- Hydrography (which includes rivers, lakes and ponds).
- Topography (contour lines).

Moreover, the BEGO project will require that the land cover map be updated after every five years. The existing set of land cover maps in BINP is a product of 1954 photography carried out during the era of the colonial Government. These maps were produced in 1965 and reprinted in 1978 and 2000. No detailed attempts have been made to up-date the land cover maps either using aerial photographs or remotely sensed images.

This study explores the utility of SPOT (XS) and Landsat (ETM) remote sensed images for land cover mapping of Bwindi Impenetrable National Park (BINP). The results of this study serve as a guide to implementation of digital image processing for land cover mapping as well as map revision.

1.4 Objectives of the study and research questions

Specific objective: The Production of a land use and land cover (LULC) map for the Bwindi Impenetrable National Park (BINP) and its immediate surrounding areas using remote sensing technologies.

General objectives:

1. Identify and classify land use and land cover features by digital image processing of original SPOT (XS) imagery, Principal Component Analysis (PCA), Vegetation Indices (VI) and Texture Analysis (TA).
2. Identify and classify land use and land cover features by digital image processing of original Landsat TM imagery through PCA, Tasseled cap transform (TCT), vegetation indices and texture analysis.
3. Identify an appropriate method based on either secondary correlation or vegetation indices suitable for band selection for supervised classification.
4. Test the classification accuracies of classified SPOT and Landsat images and compare the classification results
5. Integrate the land use and land cover map of the Bwindi with other data layers into a prototype GIS.

1.5 Research questions

1. What information is obtained from land use and land cover mapping of the Bwindi impenetrable forest using SPOT and Landsat data?
2. What bands (original and derived) are suitable for classification of (SPOT) as well as Landsat data in tropical forest mapping?
3. Which method is suitable for appropriate band selection?
4. What accuracies are obtained through the classification of SPOT (XS) and Landsat TM images?

1.6 Potential users of the research results

The findings of the study can serve as a guide in implementing digital image processing for forest mapping in Uganda and the establishment of a Geographical Information System (GIS) database for the management of forest resources. A GIS forest database is in line with United Nations Educational Scientific and Cultural Organisation (UNESCO) and European Space Agency (ESA) joint initiative demonstrating the application of Earth Observation and other Technologies (e.g. Navigational and positioning) in support of State signatories of the Convention to assist Developing Countries in fulfilling the World Heritage Conservation (WHC).

The results of this research can be useful to the managers of the Bwindi Impenetrable National Park as well as local government for planning activities that reduce the pressures on the forest resources. The GIS database provides information on land cover, water sources, trading centers, park trails, topography, park boundaries as well as road network which are all necessary for planning purpose. Uganda Wildlife Authority (UWA), Uganda National Council of Science and Technology (UNCST) and educational institutions may find the results of the study useful.

1.7 Research methodology

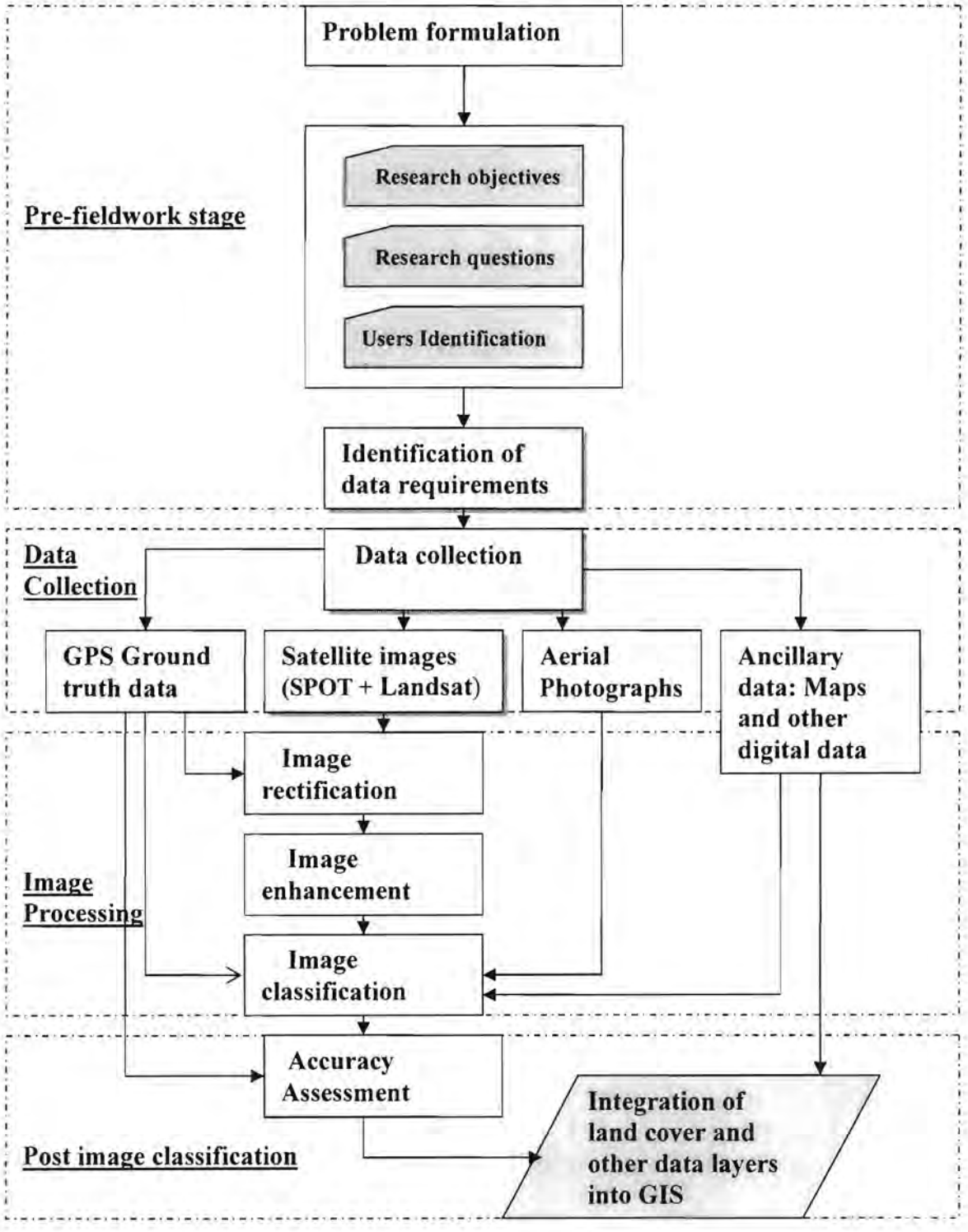


Figure 1-2 Research conceptual framework

1.8 Thesis outline

The thesis is organized into six chapters. The summaries of the subsequent chapters are as follows:

Chapter 1 presents a general introduction to land use and land cover (LULC) and discusses the importance of land use and land cover information. The chapter highlights the different methods used for LULC mapping and further gives an insight into the importance of the BINP, objectives of the study, justification of the study, problem statement, methodology and definition of study area.

Chapter 2 briefly reviews the theoretical background of some methods and technologies used in digital image processing. The focus is on digital image processing and feature extraction methods (Principal components, tasseled cap, image ratios and texture transform) as well as the developments in satellite technology with a focus on Landsat TM and SPOT satellites is presented.

Chapter 3 reviews the main land use and land cover mapping methods that have been developed to date. The chapter highlights the advantages and disadvantages of each method, the aim of which will be to identify the remaining problem areas, which forms the objectives of the study.

Chapter 4 describes the study area in much detail. Aspects of geographical location, physical features, geology, vegetation, fauna, and conservative value as well as limitations to the study are presented here.

Chapter 5 is a detailed description of the methodology used in this project for digital image processing. The various approaches used for digital image processing in this research are presented here. It focuses on the main spectral enhancements techniques used in this research (principal component analysis, tasseled cap and vegetation indices).

Chapter 6 presents the results of accuracy assessment. The results obtained from supervised classification of SPOT and Landsat images are tested using the kappa statistic and overall accuracy.

Chapter 7 presents the analysis of the results. The chapter analyses the results of different land cover types derived from the classification, the method of band selection, classification algorithms and comparison of accuracies obtained from accuracy assessment.

Chapter 8 presents the conclusions and recommendations based on the findings.

Chapter 9 presents the references and bibliography.

Chapter 10 presents the appendices.

2 THEORETICAL BACKGROUND

2.1 Introduction

This chapter provides a theoretical background of the principal concepts of remote sensing. The chapter gives the basic steps in digital image processing and Emphasis is placed on principal component analysis (PCA), Tasseled Cap Transforms (TCT), Band Ratios (BRs) and Texture Analysis (TA) since these form the basis of the methodology applied in this research.

2.2 Definition of remote sensing

Remote sensing is defined as " the science and the art of obtaining information about an object, area, or phenomenon through analysis of data acquired by a device that is not in contact with the object, area or phenomenon under investigation" (Lillesand and kiefer, 1999; Gibson and Power, 2000). It specifically refers to the methods that employ electromagnetic energy such as light, heat, and radio waves as a means of detecting and measuring target characteristics (Sabins, 1997). Remote sensing excludes geophysical methods such as electrical, magnetic and gravity surveys that measure force fields rather than electromagnetic radiation (Sabins, 1997).

2.3 Characteristics of imaging remote sensing systems

The major characteristics of an imaging remote sensing system instrument operating in the visible and infrared spectral region are described in terms of spatial, spectral, temporal and radiometric resolution (Mather, 1999).

2.3.1 Spatial Resolution

Spatial resolution is the ability to distinguish between closely spaced objects on an image. Spatial resolution is also defined as the fineness of spatial detail visible in an image. It is normally described by the terms fine or coarse spatial resolution. Coarse spatial resolution is where only large features are visible in an image as opposed to fine spatial resolution where small features can be detected.

2.3.2 Spectral resolution

Spectral resolution on the other hand refers to the specific wavelength intervals in an electromagnetic spectrum recorded by a sensor. Wide intervals in electromagnetic spectrum are referred to as coarse spectral resolution and narrow intervals are referred to as fine spectral resolution. For example, SPOT panchromatic sensor has coarse spectral resolution because it records Electromagnetic Radiations (EMR) between 0.51 and 0.73 μm while band 3 of Landsat TM sensor has fine resolution because it records EMR between 0.63 and 0.69μm (Jensen, 1996).

2.3.3 Radiometric resolution

Radiometric resolution is defined as the number of brightness levels, which a sensor can record. It is expressed in terms of bits. An 8-bit radiometric resolution for example has 256 levels (grey values) i.e. 2^8 while a 3-bit has 8 levels i.e. 2^3.

2.3.4 Temporal resolution

Temporal resolution refers to the frequency a sensor obtains imagery of a particular area. For example, the Landsat satellite can view the same area of the globe once every 16 days while on the other hand SPOT (XS) can view the same area every 26-days. Temporal resolution is an important factor to consider in change detection studies.

2.4 Radiometric characteristics and features of Landsat (TM) and SPOT (XS)

2.4.1 Landsat TM/ETM+

The Landsat Thematic Mapper (TM) is a sensor onboard the Landsat 4 and Landsat 5 spacecrafts. It provides information on the earth's surface in the visible, near, middle and thermal infrared regions of the electromagnetic spectrum. The newest in this series of remote sensing satellites is Landsat 7. Launched on 15 April 1999, Landsat 7 has the new Enhanced Thematic Mapper Plus (ETM+) sensor. This sensor has the same 7 spectral bands as its predecessor, TM, but has an added panchromatic band with 15 meter resolution and a higher resolution thermal band of 60 meters. The ETM+ sensor also has a five percent absolute radiometric calibration and Table 2.2 show the radiometric characteristics of the Landsat image and the sensor system.

2.4.2 SPOT (XS) Satellite

SPOT (XS) satellite provides information on the earth's surface in three spectral bands. The SPOT satellites orbit at an altitude of 822 km. SPOT scene sizes are typically 60 km by 60 km (vertical viewing) or 60 km by 80 km (oblique viewing). The optical imaging instruments (HRVs) are steerable to either side of the ground track – east to west by up to 30 degrees. Table 2.3 is a summary of the characteristics of SPOT XS imaging system.

Table 2.1 Radiometric characteristics of Landsat TM and ETM+ sensors

Band Number	Spectral Range (Microns)	EM Region	Generalized Application Details
1	0.45 - 0.52	Visible Blue	Coastal water mapping, differentiation of vegetation from soils
2	0.52 - 0.60	Visible Green	Assessment of vegetation vigor
3	0.63 - 0.69	Visible Red	Chlorophyll absorption for vegetation differentiation
4	0.76 - 0.90	Near Infrared	Biomass surveys and delineation of water bodies
5	1.55 - 1.75	Middle Infrared	Vegetation and soil moisture measurements; differentiation between snow and cloud
6	10.40- 12.50	Thermal Infrared	Thermal mapping, soil moisture studies and plant heat stress measurement
7	2.08 - 2.35	Middle Infrared	Hydrothermal mapping
8	0.52 - 0.90 (panchromatic)	Green, Visible Red, Near Infrared	Large area mapping, urban change studies

Table 2.2 Other Landsat satellite characteristics

Property		**Landsat 7 ETM+**	**Landsat 5 TM**
Ground Sampling Interval (GSI) (pixel size)	Bands 1-5 & 7 Band 6 Band 8	30 x 30 m 60 x 60 m 15 x 15 m pixel size (18 x 18m GSI)*	30 x 30 m 120 X 120 m N/A
Swath width		185 km	185 km
Repeat coverage interval		16 days (233 orbits)	16 days (233 orbits)
Altitude		705 km	705 km
Quantization		Best 8 of 9 bits	8 bits (256 levels)
On-board data storage		375 Gb (solid state)	Magnetic tape failed
Orbit type		Sun-synchronous	Sun-synchronous
Inclination		98.2°	98.2°
Equatorial Crossing		Descending node: 10:00am	Descending node: 10:10am

**ETM+ band 8 (panchromatic) was designed to be acquired at 15m resolution, but post-launch testing shows a ground sampling interval closer to 18m.*

Source: http://www.ga.gov.au/acres/pro_ser/landdata.htm

Table 2.3 Characteristics of SPOT (XS) sensors

Mode	Band	Micrometers	Uses	Resolution
Multispectral (XS)	1 (Green)	0.50-0.59 micrometers	Recommended for use in combination with other bands because of low contrast and sensitivity to haze	20 meters
	2 (Red)	0.61-0.68 micrometers	Best for showing roads and bare soils. This band heightens the contrast between vegetated and non-vegetated areas	20 meters
	3 (Near Infrared)	0.79-0.89 micrometers	This band is used to evaluate vegetation biomass and separates water from vegetation.	20 meters

In addition to the multispectral bands (XS) of SPOT presented in Table 2.3, the early SPOT satellites (SPOT 1, SPOT 2 and SPOT 3) carried a panchromatic (P) sensor on board. The P sensor acquired images in the range of (0.51- 0.73) μm with a resolution of 10 m x 10 m and is intended for applications requiring fine geometrical details. The P sensor (or P band) is not emphasized here since it was not used for analysis. A detailed analysis of the SPOT program can be found in Jensen (1996) and Sabins (1996). A new Vegetation instrument was launched on-board SPOT-4 in March 1998 and has been used to acquire images for vegetation studies.

2.5 Digital image processing

Digital Image Processing (DIP) involves the manipulation and interpretation of digital images with the aid of a computer (Lillesand and Kiefer, 1999). Numerous authors have dealt with the subject of DIP (Sabins, 1997; Gibson and Power, 2000; Jensen, 1996; Erdas, 1999; Mather, 1999; and Lillesand and Kiefer; 1999). In this section, an illustration of some concepts of DIP is presented with specific attention to Principal

Components Analysis (PCA), Tasseled Cap Transform (TCT), Image Ratios (IRs) and Texture Analysis (TA).

2.5.1 Image restoration

Image restoration methods compensate for data errors, noise and geometric distortions introduced during the scanning, recording and playback processes. Examples of such operations are: restoring of; line dropouts, periodic line striping and line offsets, filtering random noise; correcting for atmospheric scattering and geometric distortions (Sabins, 1997, Gibson and Power, 2000, Jensen, 1996, Erdas, 1999 and Lillesand and Kiefer, 1999).

2.5.2 Image enhancement

According to Mather (1999) image enhancement is the alteration of the appearance of an image in such a way that the information contained in the image is more readily interpreted visually in terms of a particular need. Some examples of image enhancement processes include: contrast enhancement, density slicing, edge enhancement, digital mosaics, intensity hue and saturation transformations, merging data sets and synthetic stereo images (Sabins, 1997).

2.5.2.1 Image transforms

Image transforms are the arithmetic operations of addition, subtraction, multiplication and division that are used to generate a new, composite, image from two or more bands of a multispectral or multitemporal image (Mather, 1999). The arithmetic operations can be applied in the following areas:

- Principal Component Analysis (PCA).
- Vegetation indices (VIs)
- Tasseled Cap Transforms (TCT)

- Texture Analysis

2.5.2.2 Principal Component Analysis

Extensive interband correlation is a problem frequently encountered in the analysis of multispectral image data. That is, images generated by digital data from various wavelength bands often appear similar and convey essentially the same information (Lillesand and Kiefer, 1999: Sabins, 1997: Erdas, 1999; Jensen, 1996). When the correlated bands are plotted, for example, TM band 1 against TM band 2 of the Landsat image, data points are distributed in an elongated manner which indicate that as DNs increase for one band they increase for the other band as well (Sabins, 1997). If this redundancy were reduced, the amount of data required to describe a multispectral image could be compressed (Sabins, 1997).

The Principal-Component Transformation (PCT), originally known as the eigenvector, principal component, or discrete Karhunen-Loéve transform (Gonzalez and Witz, 1977), involves a mathematical procedure that transforms a number of correlated variables into uncorrelated variables called *principal components* (Sabins, 1997; Erdas, 1999; Richards, 1999; Gibson and Power, 2000; Jensen, 1996; Lillesand and Kiefer, 1999; Mather, 1999; Swan, 1995). The first principal component accounts for most of the variance for the input data and each succeeding component accounts for less variance. The PCA may therefore be used to compress the information content of a number of bands of the imagery (e.g., seven Thematic Mapper bands) into just two or three transformed principal components.

Figure 2-1 shows a hypothetical plot of DN values of TM band 1 (axis X_1) and TM band 2 (axis X_2) axis of Landsat satellite image. For the two bands of data, the transformation defines a new axis (principal component axis) Y1 oriented in the direction of maximum variance while axis Y_2 is perpendicular to Y_1.

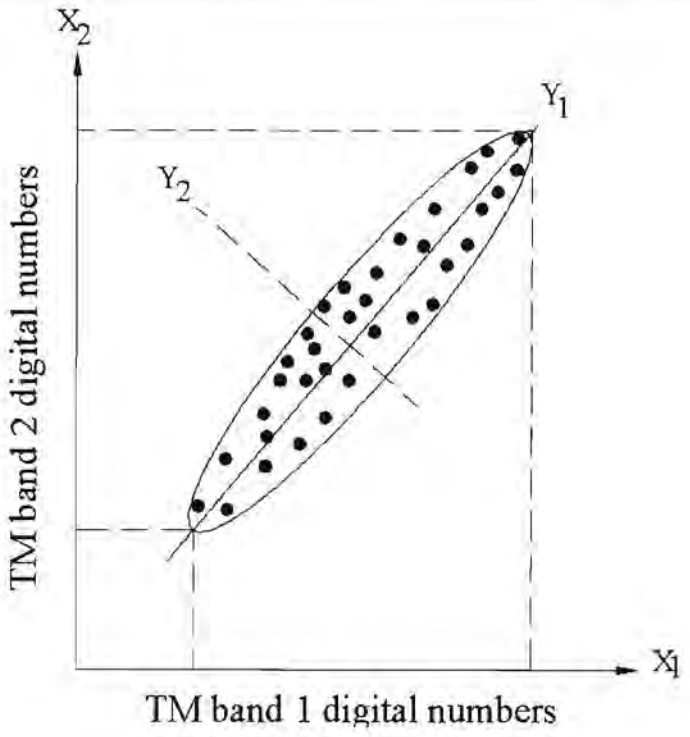


Figure 2-1 Principal Component Analysis

(After Sabins, 1997: Figure 8-26)

In order to perform PC transform, the following steps are necessary (Richards, 1999):

- Assembling the covariance matrix of the image to be transformed,
- Determining the eigenvalues and eigenvectors of the covariance matrix
- Obtaining the transformation components (matrix) using eigenvectors of the covariance matrix as the weighting coefficients.

The covariance matrix is one of the most important mathematical concepts in the analysis of multispectral remote sensing data. It describes the spread or scatter of the data in a multispectral space as well as interband correlation.

An unbiased estimate of the covariance matrix is given by the following equation:

$$\sum_x = \frac{1}{k-1}\sum_{k=1}^{k}(x_k - m)(x_k - m)' \quad \text{..... (2-1)}$$

Where $\sum_x$ is the covariance matrix of the data to be transformed, K is the total number of pixel vectors, x is the pixel vector, m is the mean of the pixel vectors and t is the transpose.

An ***eigen vector*** of a matrix A is a vector **X** such that A sends **X** to a multiple of itself. The multiple is called the ***eigen value*** and it corresponds to the vector X. If the ***eigen value*** is denote as λ, the equation (2-2) holds

$$AX = \lambda X \quad \text{(2-2)}$$

For data transformation, a matrix A, defines covariance of the data to be transformed i.e. $\sum_x$

Equation (2-2) can be reduced to

$$(\sum_x - \lambda I)X = 0 \quad \text{(2-3)}$$

Where I is an identity matrix

Equation (2-3) can be solved by first determining the eigen values λ and then using the values of λ to find the eigen vectors X.

The eigen values are determined by solving the characteristic equation

$$\left|\sum_x - \lambda I\right| = 0 \quad \text{(2-4)}$$

The values for λ and $\sum_x$ are substituted in equation (2-3) in order to determine the eigen vectors X.

The transposed matrix of eigen vectors of $\sum_x$ is the transformation matrix G required to transform the original data to the new axis (principal components) based on the following equation.

$$Y = GX \quad \text{(2-5)}$$

Where Y represents the positions (brightness values) of the new axes (principal components), G is the transformation matrix (transpose of a matrix of eigen vectors) and X represents the positions (brightness values) in the original axes (correlated axes).

The Eigen values provide a measure of the percentage of the variance contained in each component while the eigenvectors show the direction of the new principal axis or principal components. The first principal component (PC1) corresponds to the major axis of the ellipse of the uncorrelated data (Y_1) while the second principal component (PC2) is perpendicular to PC1

Principal components can be analyzed as separate black and white images or any three of the components can be combined and processed as a colour composite.

2.5.2.3 Vegetation Indices (VIs)

Vegetation Indices (VIs) are spectral measures of the amount of vegetation present on the ground in a particular pixel. VIs are used in remote sensing to estimate the likelihood that vegetation was actively growing at a particular location whenever it was observed (http://www.gvm.jrc.it/stars/vis.htm: accessed on 9th /8/2004). The VIs are therefore based on the typical spectral characteristics of healthy green vegetation, dead or senescent vegetation and bare soil.

According to Verstraete (2001) the following are advantages of using VIs

- Simplicity of implementation.
- Low computational cost.
- Reduction in data volume
- Fast estimate

VIs can reduce the confounding effects of the topography, soil background, atmosphere and view/sun angle (http://www.ntsg.umt.edu/Courses/dip/lectures/vi/sld002.htm, accessed on 9th/8/2004).

However, there are some disadvantages associated with using VIs and among these are the following:

- VIs are sensitive to many geophysical variables not of direct interest or relevance (e.g., atmospheric effects, soil brightness changes with humidity).
- VIs are sensitive to geometrical conditions of illumination and observation, as well as to the particular anisotropy (i.e. having different values when measured in different direction) of the observed system.
- VIs are sensitive to the particular sensor used
- VIs are not a measurable geophysical product
- VIs have no clear, unique, definite meaning or interpretation.
- Values are difficult or impossible to compare with great accuracy if they describe locations separated by large distances, locations observed on different days, or locations observed with different sensors.

Despite the disadvantages of using VIs, researchers have continued to use VIs especially in applications such as:

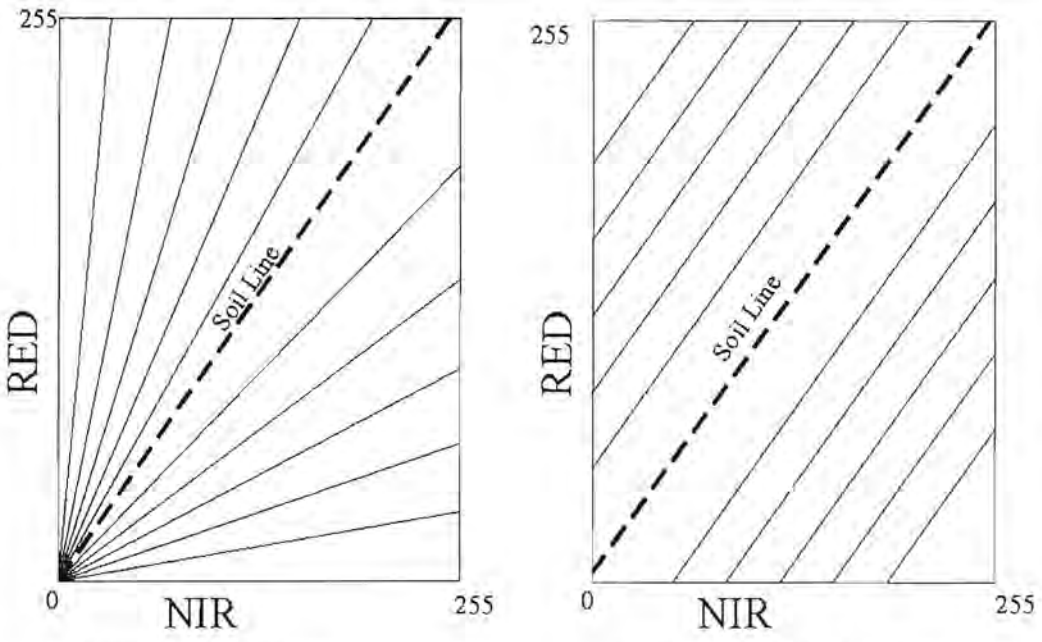
- Mineral exploitation
- Determination of the quantity and condition of vegetation present on the landscape.
- Change detection and monitoring
- Estimation of biophysical parameters (e.g. biomass, leaf area index and productivity).

In all vegetation indices, two assumptions are made i.e.:

- An algebraic combination of remotely sensed spectral bands can provide useful information about vegetation
- All bare soil in an image will form a line in the spectral space i.e. the soil brightness line of zero vegetation.

The second assumption leads to two schools of thought about the orientation of the lines of equal vegetation (isovegetation lines) in relation to the soil brightness line i.e.

- The isovegetation lines converge at a single point. Examples of the ratios employing this assumption are ratio based indices which measure the slope of the line between the point of convergence and the red/near-infrared point of the pixel. These ratios are, Normalised Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Transformed Vegetation Index (TVI) and Ratio Vegetation Index (RVI). Figure 2-2 (a) shows how the isovegetation lines would appear for NDVI/RVI scatter plot (i.e. the isovegetaion lines intersect at the origin of the RED/NIR). In other cases the intersection of the isovegetation lines may not occur at the origin of the RED/NIR. For example the SAVI isovegetation lines intersect at the negative RED and negative NIR quadrant. For slope based indices shown in Figure 2-2 (a), the maximum positive slope (as measured from the soil line) would be +1 while the maximum negative slope (as measured from the soil line) would be -1. The positive slopes are associated with the presence of vegetation with +1 indicating total vegetation cover. The slope of 0 indicates bare soil while negative slopes indicate the presence of water. It should be noted that slope refers to the ratio of the respective index.
- The isovegetation lines remain parallel to the soil line. These indices measure the perpendicular distance from the soil line to the Red/near-infrared point of the pixel. Examples of these indices are Perpendicular Vegetation Index (PVI), Weighted Difference Vegetation Index (WDVI) and Difference Vegetation Index (DVI). Figure 2-2 (b) shows how the plot of the isovegetation lines for the distance based indices would appear in the RED/NIR scatter plot.



a) Slope based indices b) Distance based indices

Figure 2-2 Slope based indices and distance based indices

Table 2.4 and Table 2.5 show some of the vegetation indices and their properties. The following should be noted for all the cases:

- RED refers to reflectance in the red band of electromagnetic spectrum
- NIR is the near infrared band reflectance
- g is the gradient of the soil line
- a is the y intercept of the soil line.
- L is the correction factor designed to correct for the soil reflectance component of the energy detected by the sensor. According to Huete (1988), L is equal to 0.5
- For the cased of the Perpendicular Vegetation Index, S_{RED} and S_{NIR} are the RED and Near-Infrared (NIR) reflectance for soil at the point where the perpendicular from a pixel with Digital Number (DN) values V_{RED} and V_{NIR} meets the soil line. V_{RED} and V_{NIR} are DN values of the pixel in the RED/NIR scatter plot. The procedure for determining the values (co-ordinates) of the intersection of the perpendicular from the pixel to the soil line (i.e. S_{RED} and S_{NIR}) is described by Richardson and Wiegand (1977).

Table 2.4 Vegetation indices and their properties

Index and source	Name	Formula
RVI Jordan (1969)	Ratio Vegetation Index	$RVI = \frac{NIR}{RED}$
NDVI Rouse et al. (1973)	Normalised Difference Vegetation Index	$NDVI = \frac{NIR - RED}{NIR + RED}$
PVI Richardson and Wiegard (1977)	Perpendicular Vegetation Index	$PVI = \sqrt{(S_{RED} - V_{RED}) + (S_{NIR} - V_{NIR})^2}$
DVI Tucker (1979)	Difference Vegetation Index	$DVI = NIR - RED$
IPVI Crippen (1990)	Infra-Red Perpendicular Vegetation Index	$IPVI = \frac{NIR}{NIR + RED}$
WDVI Clevers(1988)	Weighted Difference Vegetation Index	$WDVI = NIR - g * RED$
SAVI Huete (1988)	Soil Adjusted Vegetation Index	$SAVI = \frac{(NIR - RED}{(NIR + RED + L) * (1 + L)}$
TSAVI Baret et al. (1991)	Transformed Soil Adjusted Vegetation Index	$TSAVI = \frac{a(NIR - a * RED - g)}{g * NIR + g * a + X * (1 + a^2)}$
MSAVI Qi et al. (1994)	Modified Soil Adjusted Vegetation Index	$MSAVI = \frac{NIR - RED}{(NIR + RED + J) * (1 + J)}$ $J = 2 * g * NDVI * WDVI$
GEMI Pinty and Verstraete (1991)	Global Environment Monitoring Vegetaion Index	$GEMI = \eta * (1 - 0.25 * \eta) - \frac{RED - 0.125}{1 - RED}$ $\eta = \frac{2(NIR^2 - RED^2) + 1.5 * NIR + 0.5 * RED}{NIR + RED + 0.5}$

Table 2.5 Vegetation indices and properties (continued)

Index	Other Characteristics
RVI	Slope based, range 0 to ∞, the isovegetation lines converge at the origin. The soil line has a gradient of 1.
NDVI	Slope based and has a range of between -1 to +1. The soil line has a gradient of 1 and the isovegetation lines converge at the origin.
PVI	Distance based, varies between -1 to +1 and isovegetation lines parallel to the soil line. The soil line passes through the origin.
DVI	Distance based, the range is between -255 and +255 for data recorded in 8-bit scale. The isovegetation lines are parallel to the soil line. The soil line has passes through the origin and has an arbitrary* gradient. The range is infinite
IPVI	Slope based, has a range of 0 to +1, soil line has a gradient of 1 and passes through the origin. The isovegetation lines converge at the origin.
WDVI	Distance based, isovegetation lines are parallel to the soil line, the soil line passes through the origin and has an arbitrary* gradient. The range is infinite.
SAVI	Slope based, range -1 to +1, the isovegetation lines converge in the negative RED and negative NIR.
TSAVI	Slope based, range -1 to +1, the isovegetation lines converge in the negative RED and negative NIR.
MSAVI	Slope based, range -1 to +1, the isovegetation lines converge in the negative RED and negative NIR.
GEMI	Varies between 0 and +1.

**Arbitrary gradient: In this case the index does not assume bare soil to be represented by a single line. The index allows for bare soils with varying gradients and an arbitrary gradient is chosen to compute the index. The properties of these indices are extracted from Gibson and Clare (2000) as well as the following website http://www.shef.ac.uk/~bryant/6370/veg/vegbasic.htm [accessed on 10th/8/2004]*

The vegetation indices can be used to study vegetation in three ways namely:

- Using any one vegetation index to analyze the presence or absence of vegetation
- Combine any three different vegetation indices to form colour composite

- Combine vegetation indices with other bands (original or derived). The derived bands could be principal components, tasseled cap or texture band while the original bands could be the TM bands of Landsat or XS bands of SPOT.

2.5.2.4 Tasseled Cap Transforms

Tasseled Cap Transform (TCT) is similar to PCA in that it attempts to reduce the amount of data layers (dimensionality) needed for image viewing, interpretation and/or analysis. It was originally developed for Landsat MSS data (Kauth and Thomas, 1976) but has been further developed to accommodate all six non-thermal bands of the thematic data (Gibson and Power, 2000).

The tasseled cap technique has also been extended to other satellite images for example IKONOS and SPOT. The output values for the data structures are calculated through a linear equation with the coefficients specific to the sensor that produced the image. Three data structure axes are in use for defining the vegetation information content (Crist et al, 1986). These data structure axes are brightness (defining soils), greenness (vegetation) and wetness (vegetation vigor).

The name "tasseled cap" comes from the fact that when the greenness and brightness of a typical scene are plotted perpendicular to each other, the resulting graph has the appearance of a cap (Jensen, 1996). Figure 2-3 (a) and (b) which was extracted from, *http://rangeview.arizona.edu/images/red_reflectance.jpg* shows the tasseled cap appearance of the DN values of NIR reflectance plotted against the RED reflectance. Figure 2-3 (a) shows the distribution of pixels in RED and NIR spectral space and the directions of the brightness axis (shown along the soil line) and greenness axis (perpendicular to the soil line) are also shown. It is important to note that the tasseled cap transform was developed for use in agricultural research (Jensen, 1996). Figure 2-3 (b) shows the migration of the pixels in the RED and NIR spectral space during a growing season.

According to Gibson *et al.* (2000), if the RED and NIR responses of many different bare soil surfaces are plotted on a scatter diagram, the soil response would form a more or less straight line because as the response increases in the visible RED, it also increases in the NIR. This line is called the soil brightness line (Figure 2-3 (a) and (b)) and the direction along this axis is the soil brightness axis. The wet soils are at the bottom of the line as they appear to be darker than the dry soils. The soil line forms the basis for the derivation of vegetation indices such as perpendicular indices as well as Tasseled Cap Indices (TCI). This is because of the assumption that spectral response of living vegetation always has a consistent relation to this line. The soil line is considered to be the zero vegetation line. The soils have a fairly high response in visible RED as well as NIR while healthy vegetation has a high response in NIR and a low response in the visible RED. The vegetation response therefore appears above the soil line. The line perpendicular to the soil line is therefore the vegetation line (axis). The amount of vegetation along this line varies according to the amount of vegetation cover present. Areas with less cover are near the soil line while the vegetated areas are at the top of the cap.

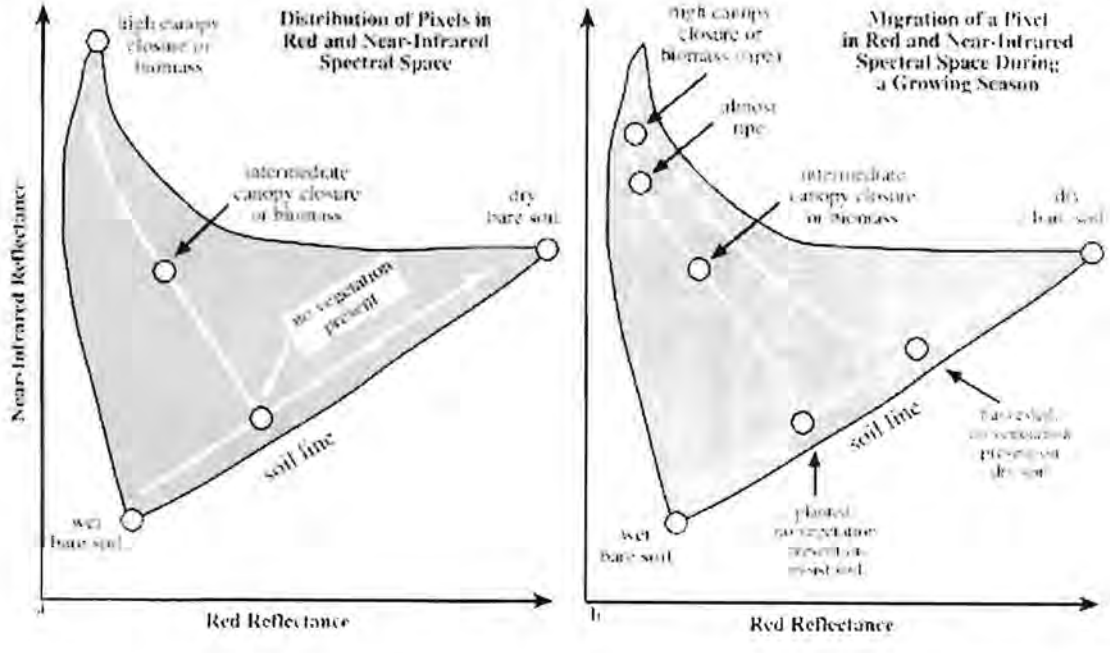


Figure 2-3 Tasseled cap and the soil line

The TCT algorithm implemented in Erdas Imagine provides the correct coefficients for Landsat MSS, TM4 and TM5 imagery. For TM4, the brightness, greenness and wetness indices are calculated as follows:

$$Brightness = 0.3037(TM1) + 0.2793(TM2) + 0.4743(TM3) + 0.5585(TM4) + 0.5082(TM5) + 0.1863(TM7) \quad (2-6)$$

$$Greenness = -0.2848(TM1) - 0.2435(TM2) - 0.5436(TM3) + 0.7243(TM4) + 0.0840(TM5) - 0.1800(TM7) \quad (2-7)$$

$$Wetness = 0.1509(TM1) + 0.1973(TM2) + 0.3279(TM3) + 0.3406(TM4) - 0.7112(TM5) - 0.4572(TM7) \quad (2-8)$$

Source: Erdas, 1999 and Jensen, 1996

The TCT produces components that are related to the geophysical properties of the scene i.e. the first component (brightness index) relates to the soil properties while the second component (greenness index) relates to the vegetation. The TCT derived bands can be used for image analysis as either single gray bands, colour composite bands (involving three TCT derived bands) or as a combination with other bands (e.g. original TM bands, ratio images and PCs)

2.5.2.5 Texture Transforms

According to Erdas (1999) many portions of images of natural scenes are devoid of sharp edges over large areas. In these areas, the scene can often be characterized as exhibiting a consistent structure analogous to the texture of a cloth. Image texture measurement can be used to segment an image and classify its segments.

A texture may be defined as the frequency of change and arrangement of tones on an image (Sabins, 1997) and could be described using the terms fine, medium and coarse or rough and smooth. Texture Transforms (TT) has been particularly applied in the enhancement of radar imagery although they can also be applied to any other raster data

(Erdas, 1999). Three approaches can be used for applications involving the use of image texture i.e.:

- Using three different sizes of a moving window to obtain a three layered variance image. The new layers can be combined into a three colour RGB (red, green, blue) image that is useful for vegetation studies or geological discrimination. Examples of window sizes used for geological studies with Radar imagery are: 15 x 15, 31 x 31 and 61 x 61.
- Preparing a three colour RGB image using three different functions (e.g. Euclidian distance, variance, kurtosis) operating through the same moving window.
- Three colour RGB image can be formed by combining texture derived layer (s) with other layers derived from different techniques (ratio indices, principal components e.t.c.) or with original bands for example TM bands for the case of Landsat imagery. This approach has been particularly applied by Ayanz (1997).

The image texture in a scene can be measured using the following algorithms: mean Euclidian distance, variance, skewness, Kurtosis, entropy, angular second moment and inverse difference moment. The details of these algorithms can be found in (Erdas, 1999; Jensen, 1996 and Peddle et. al., 1991). Below are the definitions of the texture algorithms.

$$Variance(V) = \frac{\sum(x_{ij} - M)^2}{n-1} \quad \text{.....(2-9)}$$

Where x_{ij} is DN value of the pixels *(i,j)*

n = the number of pixels in a moving window

M = the mean of the moving window

$$\text{Where } M = \frac{\sum x_{ij}}{n} \quad \text{.....(2-10)}$$

$$Mean\ Euclidian\ Distance = \frac{\sum[(\sum_{\lambda}(x_{c\lambda} - x_{ij\lambda})^2]^{\frac{1}{2}}}{n-1} \quad \text{.....(2-11)}$$

$$Skew = \frac{\sum\left|(xij - M)^3\right|}{(n-1)(V)^{\frac{3}{2}}} \quad \text{.....(2-12)}$$

$$Kurtosis = \frac{\Sigma(x_{ij} - M)^4}{(n-1)(V)^2} \quad \text{(2-13)}$$

Where

V = the variance and the rest of the variables are as in equation for the variance,

$x_{ij\lambda}$ =DN value for spectral band λ and pixel (i,j) of a multispectral image

$x_{c\lambda}$ = DN value for spectral band λ of the windows center pixel

2.6 *Image classification*

Image classification is a process of assigning individual pixels of an image to categories generally on the basis of spectral reflectance characteristics (Sabins, 1997). Figure 2-4 shows the spectral reflectance curves of three earth features i.e. Kaolinite, vegetation and soils derived from Landsat imagery.

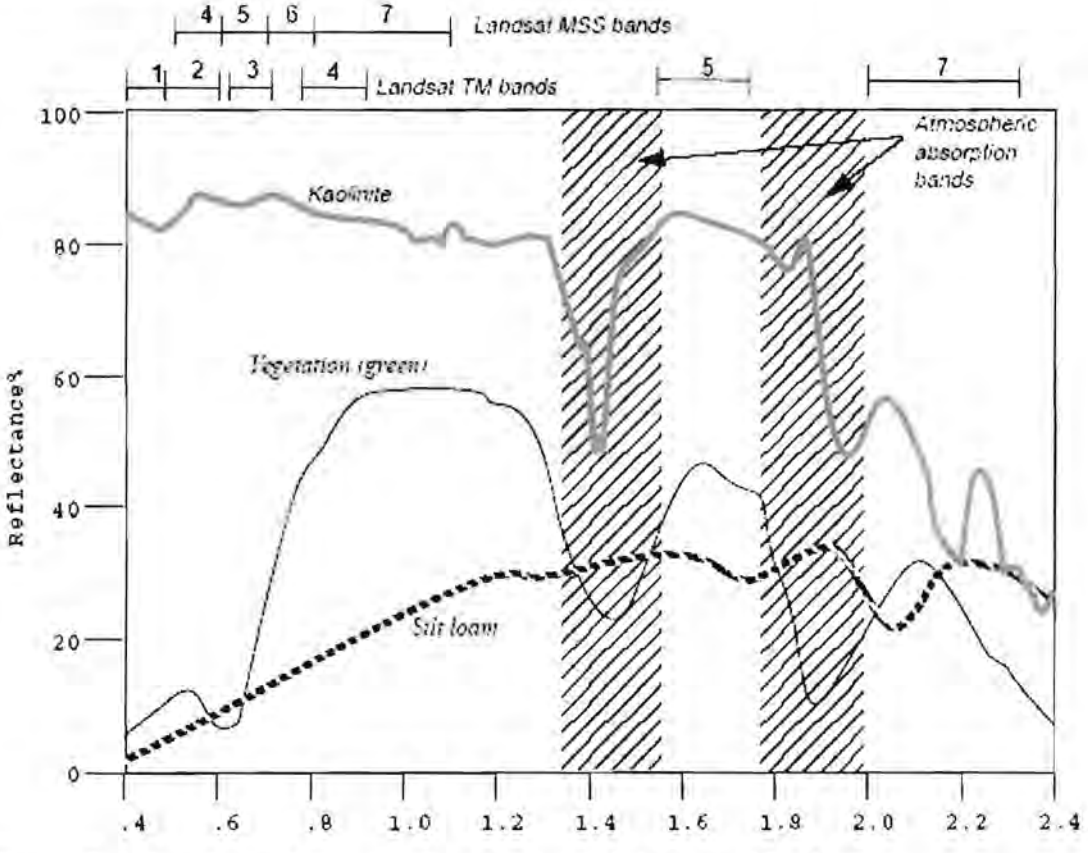


Figure 2-4 Spectral reflectance of vegetation, silt loam and kaolinite

(Extracted from Erdas field guide, 1999)

Image classification normally involves the analysis of multispectral image data (spectra) and application of statistical-based decision rules for determining the land cover identity of each pixel in an image (Lillesand and Kiefer, 1999). The curves in the figure above show how the three earth features will reflect the sun's energy in the Visible (VIs) and Infrared regions (IR) of the electromagnetic spectrum. The reflectance spectra can be used in interpreting the Landsat data. For example, the Normalised Vegetation Index (NDVI) ratio of the three earth features would be very different and therefore could be used to discriminate them.

A variety of algorithms can be used to perform multispectral classification including hard classification, classification using fuzzy logic and hybrid approaches often involving the used of ancillary (collateral) information (Sabins, 1997: Erdas, 1999: Lillesand and Kiefer, 1999). The scope of this study involves the use of hard classification approaches.

2.6.1 Supervised classification

Supervised classification is the procedure most often used for quantitative analysis of remote sensing image data. It rests upon using suitable algorithms to label the pixels in an image as representing particular ground cover types, or classes (Richards, 1999).

According to (Jensen, 1996; Erdas, 1999 and Richards, 1999) the essential practical steps necessary to perform supervised classification are:

- Decide the set of ground cover types into which the image is to be segmented. These are the information classes and could, for example be water, urban regions, croplands and rangelands, e.t.c.
- Choose representative pixels from each of the desired set of classes. These pixels form the training data. Training sets for each class can be established using site visits, maps, air photographs or even photo interpretation of a colour composite product formed from the image data (either on hard copy form or on colour display).

- Use the training data to estimate the parameters of a particular classifier algorithm to be used. These parameters will be the properties of the probability model used to define partitions in the multispectral space. The set of parameters for a given class is also referred to as the signature of that class. At this stage, the training data is evaluated. Evaluation of training data involves the determination of class separability using separability indices (e.g. Euclidian distance, divergence, transformed divergence or Jeffries-Matusita distance), contingency matrices, mean signature plot and feature space plots.
- Using the trained classifier to label and classify the pixels in the image into one of the desired information classes.
- Produce tabular summaries or thematic (class) maps which summarise the results of the classification.

2.6.2 Unsupervised classification

Unsupervised classification is a technique by which pixels in an image are assigned to spectral classes without the user having *a priori* knowledge of the existence or types of those classes. The unsupervised classification uses statistical clustering techniques to combine pixels into groups (classes) according to the degree of similarity of their brightness value in each spectral band. The analyst then combines and re-labels classes *a posteriori* into real land cover type as unambiguously as possible using maps and field knowledge (Jensen, 1996)

2.6.3 Separability indices

Separability Indices (SIs) are statistical measures of the distance between two signatures. These indices are used to evaluate training data as well as determining appropriate sets of bands for supervised classification. Separability indices include: Euclidian distance (ED), Divergence (D), Transformed Divergence (TD), and Jeffries-Matusita distance

(JM). The D, TD, and JM take into account the covariance of the signatures in the bands being compared as well as the mean vectors whereas the ED does not.

2.6.4 Euclidian distance (ED)

Euclidian distance separability index calculates the spectral distance between two signatures means. The following algorithm can be used to calculate the Euclidian distances of between any two mean vectors in the two band scatter plot.

$$Dist = \sqrt{((\mu_{ak} - \mu_{bk})^2 + (\mu_{al} - \mu_{bl})^2} \quad \text{(2-14)}$$

Where μ_{ak} and μ_{al} are the mean vectors of class *a* measured in bands *k* and *l* while and μ_{bk} and μ_{bl} are the mean vectors of class b in bands *k* and *l*.

2.6.5 Divergence and Average divergence

Divergence measures the separability of any given two classes using the mean and covariance matrices of the class statistics collected in the training phase of supervised classification.). The formula below which is an extract from Jensen (1996) computes the divergence between two classes *c* and *d*:

$$Diver_{cd} = \frac{1}{2} tr\left[(V_c - V_d)(V_d^{-1} - V_c^{-1}\right] + \frac{1}{2} tr\left[(V_c^{-1} + V_d^{-1})(M_c - M_d)(M_c - M_d)^T\right] \quad \text{(2-15)}$$

Where

$Diver_{cd}$ = the divergence between signatures c and d

tr [.] = the trace of a matrix in the brackets

V_d and Vc = are the covariance matrices for the two classes *c* and *d*

M_c and M_d = are the mean vectors for classes *c* and *d*.

The order of the matrices used in the above equation is determined by the number of bands used in the training process. If six bands were trained upon then the order of the

matrices would be 6 x 6. In this case therefore, the formula for divergence would be used to compute the statistical separability between the two classes using six bands of the training data. However, the usual goal of applying divergence is not just compute separability but to know the optimum sub set of bands for supervised classification. If a subset of three bands is required, then for six band combinations, there are 20 possible three band combinations. The divergence is computed for the two classes of interest in each of the three band subsets. The subset with the largest divergence value is then chosen as the best subset for classification.

In case there are more than two classes, the common solution is to calculate the average divergence. This involves computing the average of all pairs of classes while keeping the subset of the bands constant. The subset of the bands with the maximum average separability may be the optimal set of bands to use in the classification algorithm. The following equation is used to compute average divergence ($Diver_{avg}$) of more than two classes (Jensen, 1996):

$$Diver_{avg} = \frac{\sum_{c=1}^{m-1}\sum_{d=c+1}^{m} Diver_{cd}}{C} \quad \text{(2-16)}$$

Where, m is the number of classes and C is the number of bands in a subset.

2.6.6 Transformed Divergence (TD)

Transformed divergence is a modification of divergence separability index. The formula for computing transformed divergence is as follows:

$$TD_{cd} = 2000\left[1 - \exp\left(\frac{-Diver_{cd}}{8}\right)\right] \quad \text{(2-17)}$$

According to Jensen, this statistic gives an exponentially decreasing weight to increasing distances between the classes. It also scales the divergence values to lie between 0 and 2000. A transformed divergence value of 2000 suggests excellent separability between classes, above 1900 provides good separation while below 1700 is poor. Other methods

used to determine separability are the Bhattacharyya distance and Jeffreys-Matusita (JM) distance. The details of these methods can be found in Jensen, 1996).

3 LITERATURE REVIEW

3.1 Introduction:

This chapter reviews the work done by previous researchers using remote sensing data for land use and land cover (LULC) mapping. The literature review focuses on the sensor characteristics of Landsat (TM/ETM+) and SPOT (XS) satellites and the use of Landsat TM and SPOT XS imagery, and digital camera imagery for LULC mapping. This chapter is divided into three sections. The first section summarizes the radiometric characteristic or features of Landsat TM as well as Landsat (ETM+) and SPOT (XS) sensors. The second section reviews LULC techniques using Landsat TM imagery and the third part reviews LULC techniques using SPOT data.

3.2 Land use and Land Cover mapping with Landsat TM imagery

The definitions and the need for LULC were presented in *section 1.1*. This section focuses on the methodologies that have been used for LULC classification using remote sensing technologies.

The increasing use of remote sensing techniques in forest mapping is possible because of the high reflectance values from forested areas in the near infrared, moderate reflectance in the middle-infrared and low reflectance in the red spectral regions compared with non-forested areas (Trisurat *et al.*, 2000). The use of Landsat TM data, at 30m resolution has become the standard for landscape studies (Diamond et al., 2002). In addition to spatial information, Landsat TM has spectral advantages over other satellite data such as Advanced Very High Resolution Radiometer (AVHRR) data in that it captures mid-infrared vegetation response (Jensen, 1996).

Mapping of tropical forest in mountainous areas has however, been experienced as a difficult task (Trisurat *et al.* (2000). According to Trisurat *et al.* (2000), two main reasons

explain the difficulty in mapping tropical forests in mountainous areas: (1) large number of vegetation species, which leads to complexities in spectral classification and (2) the shadow effects in the image as a result of topographical variation, which becomes problematic in classification.

In earlier work reported by Sader et al. (1990), the primary forest (relatively undisturbed) could be identified with high accuracy using Landsat Thematic Mapper (TM) data and the secondary (disturbed forest) could not be distinguished. Mausel *et al* (1993) reported that different successive stages of growth of secondary forest could be mapped using TM data. His results show that three classes of successive forests were mapped with 82-88% accuracy.

Wolter *et al.* (1995) used TM data for mapping forests. He reported that, bands 3, 4 and 5 were particularly useful for separating forest from non- forest areas as well as the stratification of forested regions into conifer, hardwood and mixed conifer-hardwood classes. TM data were also found to be useful in detecting the presence or absence of under storey vegetation for varying canopy closures.

Stenback and Congalton (1990) carried out research on the use of Landsat TM to examine forest under storey. They reported that the spectral information provided by band 5 appeared as a key band for the classification of forests.

Apan (1997) argues that detailed mapping at the level of species and forest type is unreliable using TM data. This is due to mainly the low spatial resolution i.e., 30m x 30m.There are, however, several techniques that have been suggested to improve land use and land cover mapping using Satellite data. One of the common techniques in use today is band ratio transformation, which is used for topographic normalization and shadow reduction (Holben and Justice, 1981; Jensen, 1996, Sabins, 1997).

Colby (1991) carried out research on topographic normalization in rugged terrain. He reported that the ratio of bands 5 and 4 of Landsat TM was found to give satisfactory

results. Ratio transformations minimize the differences in the brightness values from similar surface materials due to topographical conditions, shadows, or seasonal changes in sunlight illuminations factors. Unfortunately, variations in albedo of materials may be obscured in ratio images if the materials have the same spectral curve. The information lost in the ratioing process can be restored by color combining at least one original band with ratio images (Mohammed, 1997).

Trisurat et al. (2000) carried out research on the improvement of tropical vegetation mapping using a combination of original and derived band. The objective of the research was to explore the potential of digital image processing over the existing technique of visual image interpretation. Six original bands and seven derived bands were used for the study. A total of thirteen bands (original and derived) were processed using supervised and unsupervised approaches. The seven additional bands included the first three principal components of the original six bands and four ratio bands, namely R4/3 (ration of band 4 and band 3), R5/2, R5/4 and R5/7. The results indicated that thematic classes derived from supervised classification produced higher overall accuracy than unsupervised classification. The combination of ratio bands R4/3, R5/2, R5/4 and R5/7 ranked the highest in terms of accuracy (65% for unsupervised and 79% for supervised. The combination of bands 2, 3 and 4 gave the least accuracy (56% for both methods). The results also indicated that the separation between dry evergreen forest and old secondary growth was not possible with unsupervised classification. He also reports the presence of class confusion during digital image processing especially among evergreen forest types. His recommendation was that rather than attempting to create very many forest classes, which lead to class confusion, it is necessary to define generalized land cover classes.

Mohammed (1997) carried out research based on visual interpretation of band ratios. The objectives of the research were: -

- To identify LULC using band ratio techniques
- To identify different band ratios as well as different band ratio combinations suitable for delineating different LULC features

In his study, Mohammed investigated eleven band ratios. He reported that each band ratio enhanced two or three LULC features but it could not enhance all the features. He further reported that in ratio combinations, all the enhanced features were revealed in different colours. His study concluded that although band ratioing is a relatively rapid means of identifying LULC classes, it is not the best and suitable method for classifying LULC features in a region. Band ratios can be used in studies such as detecting altered rocks, soil-moisture, drought study, canopy change and vegetation vigor.

Frankovich (1999) carried out research on LULC classification using Landsat TM data in Midland, Michigan. In his study, he used a combination of Principal Component Analysis, some original bands and Tasseled Cap Transform (TCT) for LULC classification. Frankovich revealed that it is possible to delineate LULC features using this approach. Frankovich reported that residential areas were classified mostly as trees. The business and industrial districts were identified and classified as open urban areas although they were speckled with pixels classified as bare fields, which were misclassified. This occurred because some urban features have the same spectral signatures as bare fields. The study reports that the biggest problem with the classification was the issue of crops. The image was largely characterized by large bare field areas, which could not easily be separated from croplands. It was therefore difficult to create a layer called crops without a lot of misclassified pixels in other areas. The few crops in the areas were grouped with either trees/shrubs class or grassy areas class.

3.3 Land use and Land cover mapping from SPOT imagery

Different scientists have used different products of SPOT satellite imagery in a number of ways. The following paragraphs present a review of the studies involving the use of the SPOT data for mapping of land cover.

Nezryt (1998) carried out research on the use of SPOT and Radar data to inventory forests in Sarawak- South East Malaysia. SPOT panchromatic and multispectral imagery, RADARSAT SAR stereo pair as well as ERS data were used in the study. The research

was aimed at establishing GIS database for forest management. The radar data complemented SPOT imagery in zones where cloud cover made acquisition of optical imagery impossible and was merged with SPOT data to yield thematic information that helped discriminate different forests types.

Two classification approaches were used: for cloud free part of SPOT multispectral images (XS); the forests were classified using indices derived from SPOT and RADARSAT data meanwhile for the cloudy part of the SPOT multispectral image (XS); the classification was performed using SAR images, and SPOT panchromatic image (Nezryt, 1998). The research does not explain why it was not possible to use SAR in both instances. The research reports however, that combining optical and Radar data effectively highlights features that might otherwise prove difficult to detect. It was further reported that the two partial classifications matched closely enough to be merged into a single final image containing six land cover classes. The details on the land cover classes can be found in Nezryt (1998). According to Nezryt (1998) the preliminary checks on the land cover classes produced satisfactory results

Achard et al. (2002) report on the use of SPOT and Landsat data sets for the determination of deforestation rates of the world's humid tropical forests. This was a study set up by the European Commission Joint Research Center (JRC) and the European Space Agency in the early 90s under the umbrella of Tropical Ecosystem Environment Observation by satellite (TREES). The project was designed for the production of a global reference map of the world's tropical forests from Satellite images. It was also intended to develop a reliable method for forest change assessment in the humid tropics and make information available in the Tropical Forest Information System (TFIS).

The project was divided into two phases. The first phase of TREES was carried out using imagery provided by United States of America (USA) National Oceanic and atmospheric Administration (NOAA) satellites for the earlier 1990's and was later completed using more recent data from European Space Agency ERS satellites and from the VEGETATION sensor of the French SPOT satellite. The first phase resulted in the

production of the first world map of tropical forests. However, the images used were lacking in definition characterized by the coarse resolution of the satellite images (http://www.gvm.sai.jrc.it). The first phase of TREES recommended that better quality data such as delivered by LANDSAT (TM sensor) and the SPOT (HRV sensor) was necessary for accurate interpretation and change assessment (Achard et al., 2002).

In the second phase of the TREES project therefore, SPOT (HRV) and Landsat (TM) images were used. The aim was to develop a reliable method for forest change assessment in the humid tropics exploiting the global imaging capabilities of the Earth Observing satellites. This second phase was completed with the production of up-to-date information on the status of the world's humid tropical forests cover (Achard, 2002). The study discussed the potential of remote sensing in the determination of forest cover for 1990 and 1997. Issues regarding annual deforested area and rate of deforestation, annual re-growth and rate of re-growth, annual net cover change and rate, and annual degraded area and rate were discussed. The maps were produced at a global scale of 1: 5,000,000. The article does not, however, describe the method used for image processing.

Ayanz et al (1997) carried out a study to compare a Single-stage and Multi-stage classification approaches for LULC mapping with TM and SPOT data. Original as well as derived bands of SPOT and Landsat TM were used for image processing. The derived bands used were Principal components, band ratios and texture. The details of the derived bands can be found in Ayanz et al (1997). In a single stage classification approach the land cover categories were classified at once while under multi-stage classification one or two bands are classified at a time. The single stage classification differed from one another in the band selection process, the use or not of prior probabilities, and /or the use supervised or unsupervised classification. Ayanz reported that results from multi-stage classification were generally superior to any other classification method. He further noted that TM data proved superior to SPOT when the iterative classification approach was used. The higher spectral content of TM imagery allowed a better separability of the classes and therefore produced results of higher accuracy.

Other studies involving the use of SPOT (XS) or SPOT (PAN) for Land use and land cover classification can be found in Sabins (1996), Yousfi (2000) and Goksel (2000).

Recently, there has been research on the use of SPOT-4 VEGETATION data for vegetation mapping. The SPOT-4 VEGETATION sensor records four spectral bands covering the green, red the near infrared and short wave infrared portion of the spectrum. Mayaux (2000), Xiao (2002) and Wolf (1998) for example have used SPOT-4 VEGETATION data for LULC classification. These researchers report that it is possible to derive forest maps with the SPOT-4 VEGETATION data. However, the SPOT-4 VEGETATION data has a coarse spatial resolution of 1km and this makes it more suitable for regional and global mapping (Mayaux, 2000). Better quality data, as reported by Achard et al. (2002) such as SPOT (XS) and Landsat (TM) should still be used for more detailed forest inventory.

3.4 Conclusions

From the available literature, it is clear that efforts have been made to use remote sensing technology for LULC mapping in most parts of the world. However, Uganda is yet to embrace this technology in an extensive way. This is due to the following reasons among others:

- Lack of manpower.
- Inadequate financial support.
- Insufficient recognition of the Technology by policy makers.
- Lack of images and Equipment.
- Brain drain.

These problems are not unique to Uganda but are common to most developing countries

It also follows from the literature review that there seems to be no universally accepted method for land use and land cover mapping. Researchers have experimented with different methods to improve on the results of the classification. The most common

methods used are either the original bands or derived bands (ratios, PCA or TCT). This is due to the complexities of land cover types in different parts of the world. Until such a time as when a universal method is developed, research will have to continue on local areas/zones to determine which methods give acceptable results.

In this study, a modification of the methods used by Trisurat et al. (2000) and Ayanz et al. (1997) is adopted. This is in reference to Landsat data. Trisurat et al. (2000) used the original 6 bands of Landsat (excluding the thermal band), first three principal components (based on all the six bands) and four derived image ratios. This study modifies or adopts the methodology used by Trisurat et al. (2000) and or Ayanz et al. (1997) in the following way:

- It involves the use of Tasseled Cap Indices (TCI). Tasseled cap indices were developed specifically for vegetation studies (Jensen, 1996) and this study incorporated TCI, a method adopted by Frankonovic (1999) but was not used by either Trisurat et al or Ayanz et al.
- It involves the use of 6 original bands as did Trisurat et al and Ayanz et al.
- It involves the use of Texture bands as Ayanz et. al. did but Trisurat et al did not.
- Principal components are also used but based on the method used by Trisurat et al other than Ayanz et al. This involves the use of all the six bands for the derivation of principal components.
- Lastly, band ratios are also used in a way similar to that used by Trisurat et al (2000), Ayanz et al. (1997) as well as Mohammed (1997).

The use of both original bands and derived bands should provide a number of bands from which the best possible band combination (s) is selected for image classification. The method used for the classification of SPOT image was similar to the method used for Landsat image classification. However, a lesser number of band ratios were derived from SPOT image owing to its low spectral resolution. The Principal component analysis was performed using the three original SPOT bands while no TCI were derived. This was due to the limitation of the software used which does not provide the transformation parameters for the tasseled cap transform of the SPOT bands.

4 STUDY AREA

4.1 Introduction

This chapter presents a description of the study area. Factors such as the geographical location, altitude, physical features, vegetation types and climate are discussed here.

4.2 Geographical location

Bwindi Impenetrable National Park (BINP) is located in South Western Uganda and covers an area of 331km^2. It is within the latitudes of $0^0$53'S-$1^0$08'S and longitude of $29^0$35'E-$29^0$50' E. (Figure 4-1). It lies at the edge of the great Western Rift Valley, and is part of the highest blocks of the Kigezi and Rukiga Highlands. BINP lies next to the Democratic Republic of Congo (DRC) border in the east, about 29 km by road North West of Kabale, 20 km north of Kisoro, and 40km south east of Lake Edward to the closet boundary point. It is located in four districts i.e. namely Kabale, Rukungiri, Kisoro and Kanungu

4.3 Physical features/Landscape

BINP has a wide altitude range, from 1160m (3805ft) at the tip of the northern sector to 2607m (8551ft) at the Rwamunyonyi hill on the southeastern edge of the Park (UWA, 1998). The landscape is generally rugged and is characterized by deep narrow valleys and steep sided hills. Flat areas can only be found in the Mubwindi and Ngoto swamps. The park constitutes an important water catchment area, serving surrounding agricultural land. Three major tributaries of the Ishasha River drain into Lake Edward to the north, and the Ndengo, Kanyamwabo and Shongi rivers flow southwards towards Lake Mutanda and Lake Bunyonyi (UWA, 1998; WCMC, 2001).

4.4 Geology

The area is associated with upwarping of the western Rift Valley and its underlying rocks are phyllites and shales, with some quartz, quartzite and granite (Howard, 1991). Significant deposits of iron, wolfram (tungsten ore) and gold have been found in and a round the park. The soils are iron rich with varying levels of humus. They are moderately poor in structure and their acidity ranges from PH 2.9 to PH 5.2 (Howard, 1991, Uganda National Parks, 1995).

4.5 Climate

The climate in the area is tropical with two rainfall peaks from March to May and September to November. The area also has two dry seasons from December to January and June August. The annual mean temperature range is 7-15^{0}C minimum to 20-27^{o}C maximum. The annual precipitation lies in the range 1,130-2,390mm (Howard, 1991; UWA, 1998; UNDP, 1993).

4.6 Vegetation

Bwindi is an afromontane forest which is considered to be the rarest vegetation type on the continent and while BINP does not feature as one of the largest of protected afromontane forests in the region, its contribution to the total representation of this forest type is significant. Current evidence indicates that Bwindi forest is the most diverse forest in East Africa for tree species (more than 200 species) and ferns (more than 104 species). This could have led to the recognition of BINP, by the International Union for the Conservation of Nature (IUCN) plant program, as one of the 29 forests in Africa most important for conservation. The forest gets the name "***impenetrable***" from the dense cover of herbs, vines, and shrubs inhabiting the valley bottoms. According to Longdale-Brown (1964), the area is broadly classified as having two types of vegetation, Medium altitude moist evergreen forest and high altitude forest.

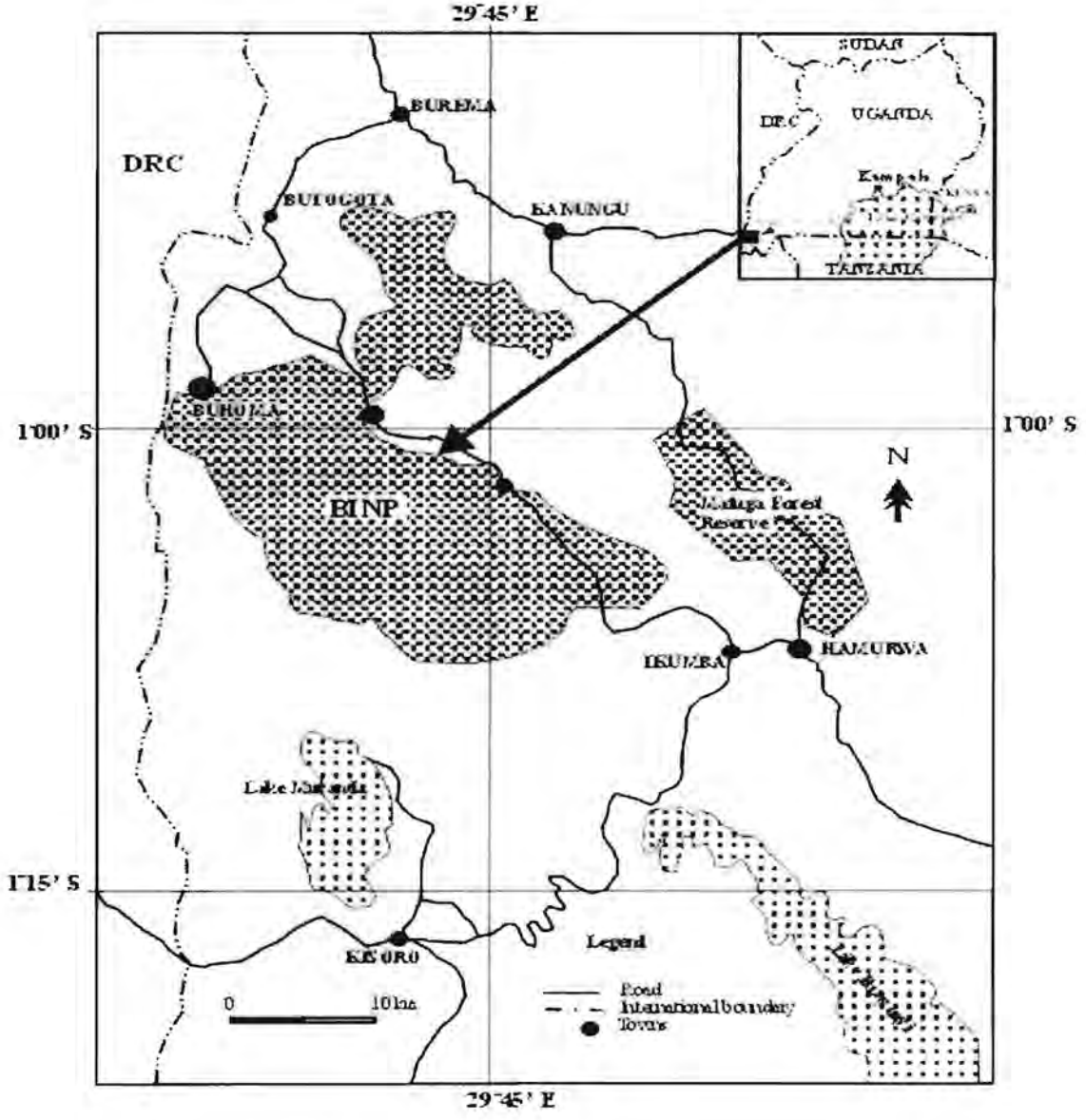


Figure 4-1 Study area

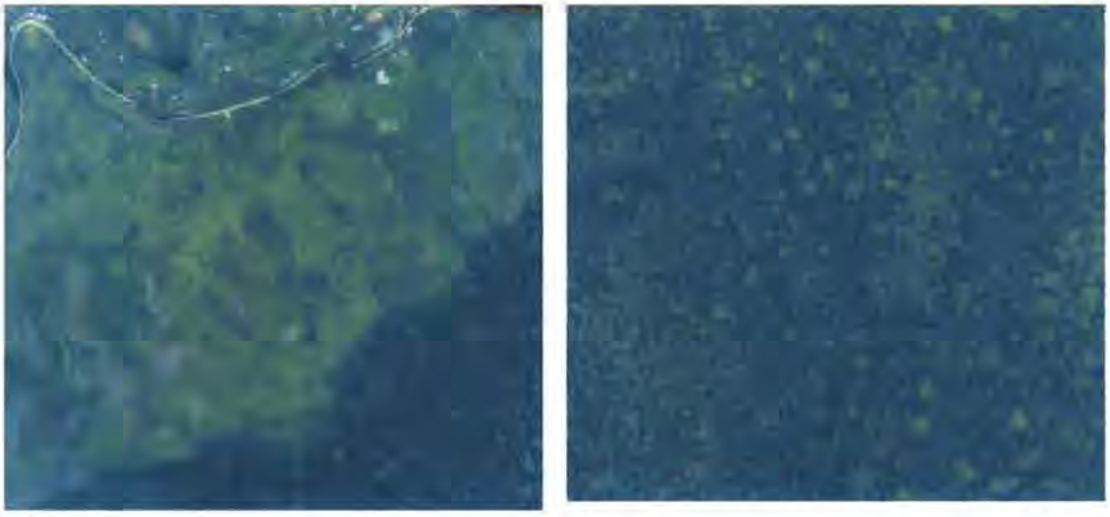
Approximately 40% of the forest is rich to medium-rich mixed forest, including key species such as *prunus africana*, nationally threatened *Newtonia buchananii* (V), *symphonia globulifera*, *chrysophyllum* spp., *podocarpus* ssp., and *strombosia scheffleri* (Howard, 1991).

Howard (1991) reports that three plant communities tend to dominate depending on altitude. In the low-lying areas around 1500m, *parinari exelsa* (Figure 4-2) is the dominant species (about 10% of the park); at altitudes of 2000m, it is the *Newtonia buchananii* (about 11% of the park) and at altitudes of 2200m, it is *chrysophyllum gorungosanum* (about 8% of the park) that dominates. Low stature communities, classified as poor hill and colonizing types, occupy almost 30% of the park.

Figure 4-2 Parinari excelsa

(Source: http://www.mckono.org/tfrrung.html)

There are also small areas of swamp and grassland. Bamboo forest is restricted to less than 100ha. The total number of tree species in Bwindi is not known. However, some 200 species (47% of the countries total) have been discovered. These include 10 species not found elsewhere: *croton bukobensis, strombosiosis tetrandra, Brazzeia longipedicellata, Grewia milbraedii, Maesobotrya purseglovei, Balthasaria schliebenii, Xylopia staudtii, Allanblanckia kimbiliensis, Memecylon* spp., and *Guarea mayombensis* (Howard, 1991). The aerial photographs in the Figure 4-3 (a) and (b) acquired from a light aircraft and DCS Proback Kodak digital camera demonstrates the complexity of the vegetation of Bwindi.



(a) (b)

Figure 4-3 Bwindi Land cover classes (aerial Photographs)

The Photograph in Figure 4-3(a) shows the following characteristic land cover types: The dense forest at the bottom right corner and subsistence farmlands on the rest of the photography. Within the subsistence farmlands, some settlements can clearly be seen as well as communication infrastructure. Photograph Figure 4-3(b) shows an area completely covered by dense forest. The vegetation is dense and dark especially when observed from a distance. This is the very reason the forest was given the name Bwindi which in the indigenous language means dark.

4.7 Fauna

According to the United Nations Environmental Program (UNEP) report (2001), Bwindi is believed to hold the richest faunal community in East Africa. BINP contains more than half of the world's population of mountain gorillas *Gorilla gorilla berengei* (about 300 out of 650), which live in some 23 family units (von Zeipel, 1996). A total of 12 species of bird, one species primate and 3 species of butterflies occur only in Bwindi. The BINP also has Chimpanzees and 8 other species of primates. Elephant, bush pig, giant forest hog, many species of bats and rodents, at least 346 species of birds including rare forest

birds, at least 14 species of snakes, 27 species of frogs and toads, 6 species of chameleons and14 species of lizards.
(http://www.kibokotravels.org.rw/english/uganda/bwindiforest.htm)

4.8 Conservation value

Bwindi is certainly the most important area in Uganda for species conservation due to its exceptional species diversity, including many Albertine Rift endemics and 9 globally endangered species (http://www.unep-wcmc.org/protected_areas/data/wh/bwindi.html.)

It is also believed to hold the richest faunal community in East Africa. In addition, Bwindi is internationally important as the habitat of more than half of the world's population of mountain gorillas.

4.9 Limitations to the study

Major limitations to the study area are high cloud cover, cloud and vegetation shadows, steep hills and ragged terrain, insecurity (due to political instability in Congo) and lack of reference data (old maps, photographs). These factors will therefore have a bearing on the final results of this research.

4.10 Conclusion

Following the above discussion, it can be concluded that BINP is one of the most important areas for conservation as has been recognized by UNESCO as well as IUCN. It is rich in both flora and fauna. The forest resources if used in a sustainable way can provide economic benefits to not only the local people in the area but Uganda as a whole. Some of the benefits that are provided by the BINP are:

- It is a tourist attraction site - it therefore generates some revenue for the Government as well as the local community.

- It provides raw materials for building especially timber and wood. Some logging is allowed in Bwindi but it is known that some timber is obtained illegally.
- There are mineral deposits especially Gold which are mined illegally.
- It also provides water to the local community as well as near by lakes.

The BINP, therefore, plays an important role in the region and all attempts are being made by the Government of Uganda as well as the international community to protect existence of BINP.

5 Digital Image Processing

5.1 Introduction

This chapter presents the materials and the methods used for this study. The chapter presents a discussion on the data used, ground truth data collection methods, classification scheme used as well as methods used for processing SPOT (XS) and Landsat (ETM+).

5.2 Satellite data

A geo-referenced, 20 x 20 meter resolution SPOT XS data acquired in 1993 as well as a geo-referenced 30 x 30 meter resolution Landsat ETM scene acquired in June, 2000 were obtained from Department of Surveys and Mapping –Entebbe, Uganda. Ground truth data collected with Gamin GPS in 2003 and 2004 was used for geo-referencing as well as performing accuracy assessment for the classified data.

5.3 Aerial Photographs

Aerial photographs acquired in October, 2003 with KODAK professional proback digital camera systems. A four-seater 5X-CAT single engine aircraft was used to acquire aerial photographs which were used as ground truth data for supervised classification.

5.4 *Ancillary data*

The ancillary data included the following:

- Shape files obtained from Makerere University Institute of Environment and Natural Resources (MUIENR). The shape files comprised of road network, rivers, park boundaries, trading centers and contours.
- Four sheets of maps for BINP area.

5.5 *Software*

The software used is presented in Table 5.1

Table 5.1 Software

Software	Use(s)
ERDAS 8.4	Image processing
ARCVIEW 3.2	Image analysis , map processing and shape file projection
Adobe photoshop 7.0	Visualisation of aerial photographs
Photodesk DCS	Reading Aerial photographs from DCS format
AutoCAD	Drawing
Corel draw 11	Drawing
Microsoft excel	Data entry and processing
Microsoft word	Word processing

5.6 Equipment

The equipments used for the study are presented in Table 5.2.

Table 5.2 Equipment

Equipment	**Use(s)**
Garmin GPS	Data collection (Ground truth data)
Computer Pentium (IV), 512Mb RAM and 40GB hard drive	Data processing
5X-CAT single engine aircraft	Aerial photography
KODAK professional proback digital camera	Aerial photography
Magnetic Compass	Showing direction
Printer	Printing

5.7 Ground truth data (Field sampling)

Fieldwork was carried out during the months of October, 2003 and March, 2004 to collect training data as well as data for testing the classification accuracies of the land cover categories. The data collection exercise was done using Garmin Global Positioning System (GPS). Simple random sampling technique was used for data collection. However, systematic random sampling was not possible due to the complex nature of the study area i.e. very steep slopes as well as dense vegetation cover. Therefore, a modified random sampling technique, where the samples were selected in easily accessible areas was used. An attempt was made to obtain samples representing the desired land cover categories. Figure 5-1 shows the field observed GPS points superimposed on to a section of Landsat (ETM+) image.

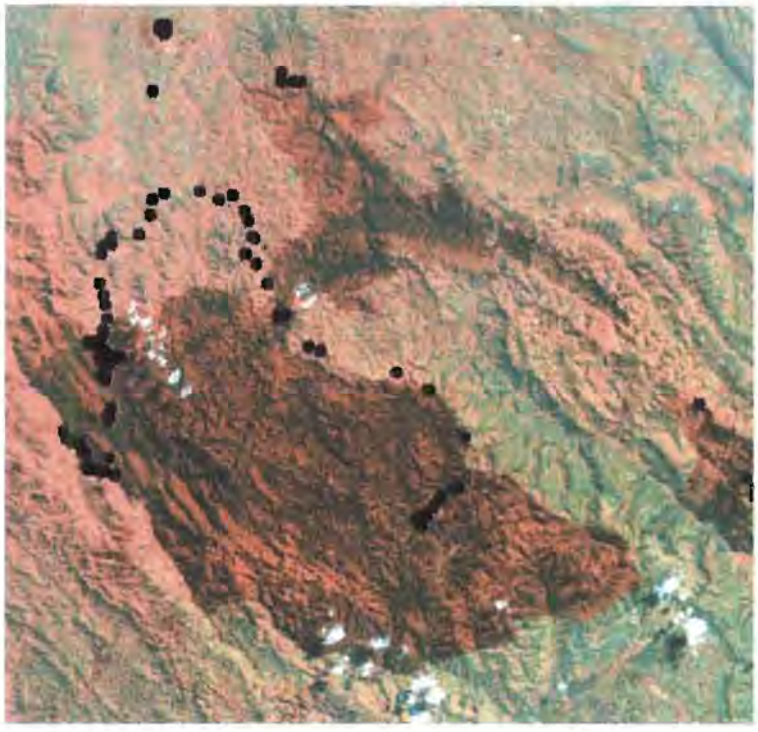


Figure 5-1 Landsat (ETM+) and Ground truth data

5.8 Land cover classification scheme

For this study, land cover classification scheme is defined as a standard base for classification of land cover activities. In a land cover classification scheme, all classes of interest must be carefully selected and defined to successfully classify remote sensed data into land cover (or land use) categories. This section presents the description of the land cover categories used for the classification of Landsat (ETM+) as well as SPOT (XS) satellite images. The classification scheme is based on the modified level II classification scheme of Anderson *et al.* (1976)

- **Dense Evergreen Forest (DEGF):** This class consists of forest areas that are undisturbed, contain a dense canopy, vibrant trees and evergreen forest.
- **Evergreen Forest (EGF)**: Mixed forest with, less dense vegetation, which remains green through out the year. Disturbed forest falls in this category. It also includes Eucalyptus forests.

- **Brush land**: Open spaces in the forest. This class includes areas covered by herbaceous as well as creeping plants.
- **Mixed rangeland**: This includes all grassland areas as well as areas with the intermixture of grassland and shrubs.
- **Subsistence farmlands (S):** areas that support subsistence farming with banana plantations being the dominant food crop.
- **Tea plantation**: These are small farmlands used for tea growing.
- **Bush land (B)**: areas with low dense vegetation.

Figure 5-1 and Figure *5-2* are extracts from the aerial photographs taken with KODAK Professional proback Digital Camera System (DCS) showing some of the land cover categories. The identification of the land cover types was based on visual photo interpretation as well as field work.

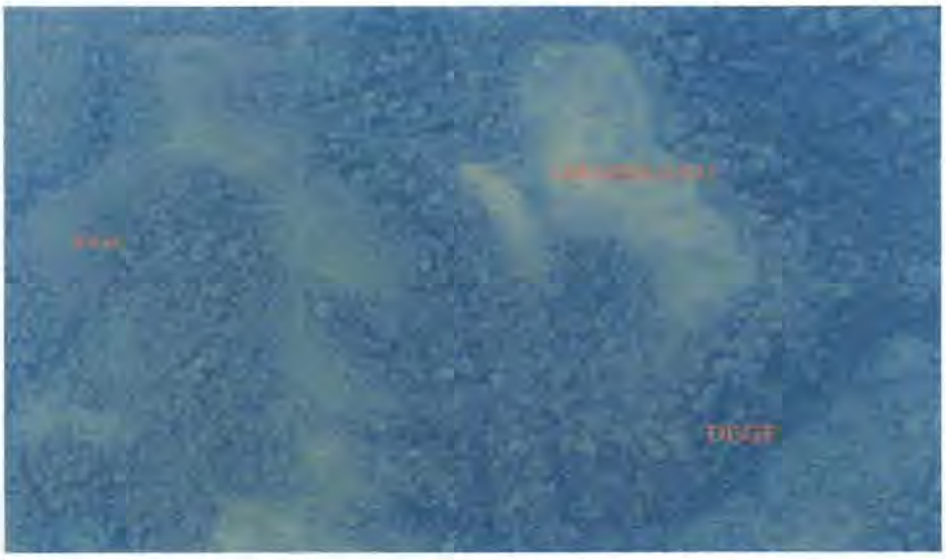


Figure 5-2 Location of EGF, DEGF and Mixed rangeland classes

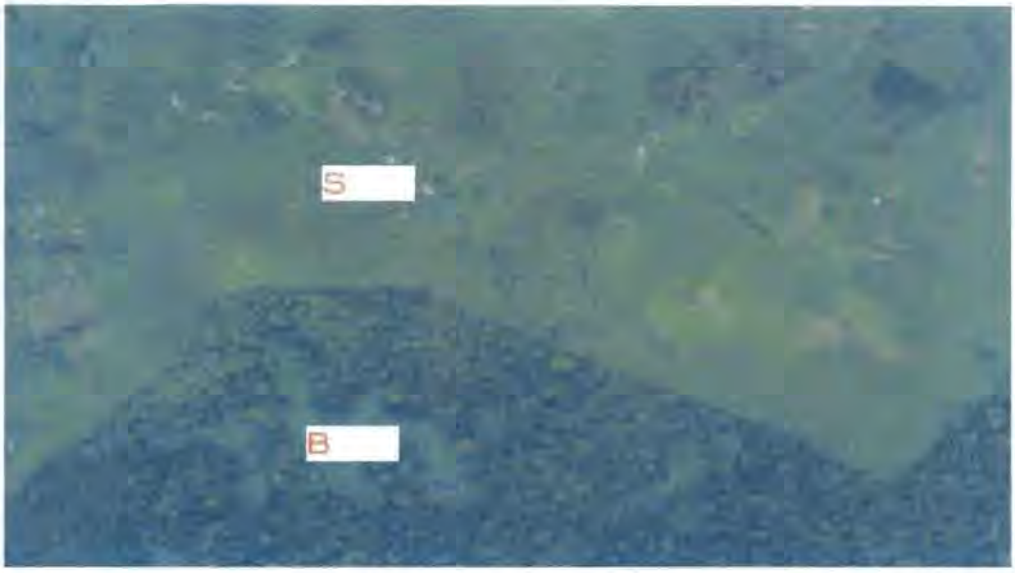


Figure 5-3 Location of subsistence farmlands and Brush land areas

5.9 Digital processing of SPOT image

5.9.1 Layer stacking, segmentation and cloud masking of SPOT image

The three SPOT bands were layer-stacked using the layer-stacking algorithm. A sub set of the image covering the study area was then extracted. The clouds and cloud shadows were masked out using Erdas 8.4. The procedure used for cloud masking included the following steps:

- Creation of new image based on Areas Of Interest (AOI) objects. The AOI were obtained by digitising areas covered by clouds and cloud shadows.
- The new AOI image was classified using unsupervised classification. Only two classes were desired in this case. One class representing AOI while the other class representing the rest of the image.
- The unsupervised classified image was used as an input mask for the cloud covered satellite image. The AOI were assigned value 0 (since this was unwanted area to be used as a mask while the rest of the image was assigned value 1. This resulted into an image with clouds and cloud shadows masked out.

For a detailed procedure on cloud masking, reference is made to Erdas field guide (1999). After clouds as well as cloud shadows were masked, the resulting image was easier to interpret than the source image.

5.9.2 Rectification of SPOT satellite image

Rectification is a process of making the image data conform to a map projection system (Erdas, 1999). Image rectification involves the conversion of image or grid co-ordinates to real-world coordinates and typically involves rotation and scaling of grid cells, and thus requires re-sampling of values. Rectification is necessary because of the following reasons:

- Comparing pixels scene by scene in applications such as change detection.
- Developing GIS data bases for GIS modeling.
- Identifying training samples according to map coordinates prior to classification.
- Overlaying an image with vector data
- Comparing images that are originally at different scales
- Extracting accurate distance and area measurements
- Mosaicking images
- Performing any other analysis requiring precise geographic location.

The re-rectification of the SPOT (XS) image was necessary because it was not possible to establish the ellipsoid against which it was originally geo-referenced. Performing an accuracy assessment of the classified SPOT image using GPS data would therefore be problematic. The initial expectation was that the image was geo-referenced to the Clark 1880 modified ellipsoid, used for most topographical maps in Uganda. Attempts to confirm this were fruitless. However, before re-rectification of the SPOT image was performed, a comparative co-ordinate analysis was done to confirm whether the SPOT co-ordinates were based on Clark 1880 ellipsoid or not. Table 5.3 illustrates the co-ordinate types (derived from different sources i.e. map and images) while tables Table 5.4

and Table 5.5 show the co-ordinates comparison results. Table 5.4 has fewer points than Table 5.5 because of the difficulty encountered in identifying common points between CLARK_1 and CLARK_2 data sources.

Figure 5-4 shows a sketch of points whose co-ordinates were obtained either with GPS or scaled from the maps. From this sketch, it clearly shown that some co-ordinates had either reliable scaled co-ordinates in the easting or northing or both. For example point 2 had reliable easting than northing while point 10 had reliable northing than easting. The word reliable is somehow used in a loose sense and it refers to how easy or difficult it was to estimate the co-ordinates. In some cases, it was either possible to estimate the eastings or northings while in other cases it was easy to estimate both the eastings and northings.

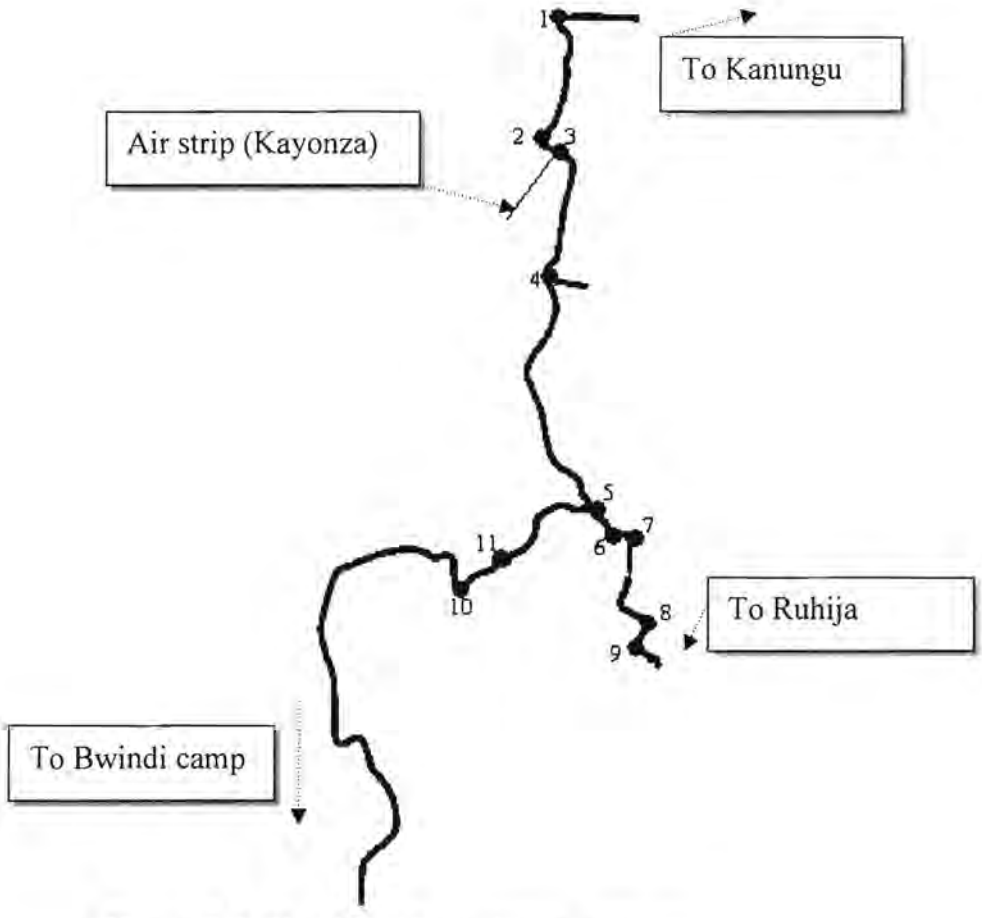


Figure 5-4 Sketch for the scaled co-ordinates

Table 5.3 Co-ordinate sources (ellipsoids)

Co-ordinate source (Ellipsoid)	Comments
CLARK_1	Co-ordinates of selected features (road junctions and bends) scaled from a topographical map of the study area. The map was at a scale of 1:50000 and the cordinates were based on Clark (1880) modified ellipsoid
CLARK_2	Co-ordinates of selected features measured from shapefiles (roads and bends). The shapefiles were derived by on screen digitising of the topographical GIS maps whose co-ordinates were based on Clark (1880) ellipsoid. It was expected that CLARK_1 and CLARK_2 co-ordinates would be in agreement since both are based on Clark (1880) ellipsoid.
SPOT	Co-ordinates of selected features obtained from the SPOT image refered here as SPOT co-ordinates. If the SPOT image was geo-referenced to Clark (1880) ellipsoid then SPOT co-ordinates would be in agreement with CLARK_1 as well

Table 5.4 Co-ordinate comparison (1)

	CLARK_1		CLARK_2		Cordinate difference	
ID	**Easting (m)**	**Northing (m)**	**Easting (m)**	**Northing(m)**	**ΔE (m)**	**ΔN (m)**
1	793675	9907300	793694	9907329	-19	-29
4	793600	9903425	793608	9903413	-8	12
5	794325	9900050	794264	9900024	61	26
8	795075	9898350	795006	9898426	69	-76
Average co-ordinate differences					**26**	**-17**

(ΔE and ΔN are the changes in the eastings and northings respectively).

Table 5.5 Co-ordinate comparison (2)

	CLARK_1		SPOT		Co-odinate Differences	
ID	**Easting (m)**	**Northing (m)**	**Easting (m)**	**Northing (m)**	**ΔE (m)**	**ΔN (m)**
1	793675	9907300	794164	9907333	-489	-33
2	793550	9905500	793926	9905521	-376	-21
4	793600	9903425	793972	9903414	-372	11
5	794325	9900050	794647	9900008	-322	42
6	794450	9899725	794833	9899766	-383	-41
7	794675	9899800	794965	9899766	-290	34
8	795075	9898350	795377	9898304	-302	46
9	794925	9898050	795211	9898011	-286	39
10	792725	9899350	793024	9899312	-299	38
11	792325	9898850	792640	9898853	-315	-3
Average co-ordinate differences					**-343**	**11.2**

The following observations were made from Table 5.4 and Table 5.5

- CLARK_1 and CLARK_2 co-ordinates were in general agreement to within ±1mm at map scale, in both the north and east directions. Taking into account the discrepancies between CLARK_1 and CLARK_2 caused by either the difficulty in scaling co-ordinates from the topographic map or measuring co-ordinates from the digital vector data (shape files), it can be concluded that no significant difference exists between CLARK_1 and CLARK_2.
- CLARK_1 and SPOT co-ordinates were in agreement to within ±7mm and ± 1mm at the map scale in the easting and northing directions respectively. This indicated a significant difference in the eastings between CLARK_1 and SPOT.

Following the above observations, it was therefore necessary to correct the SPOT image co-ordinates to conform to either CLARK (1880) or World Geodetic System (WGS 84) co-ordinates. The procedure involved the determination of the shifts in the co-ordinates in the easting and northing and the average shifts were used to correct the SPOT image. In Erdas Imagine software, this can be done by applying the shifts (add or subtract) to the origin of the image co-ordinates:

Three different options could have been used to determine the shifts and hence correct the SPOT co-ordinates i.e.

- Determine the shifts by getting the differences between Clark 1880 (either map or digital vector data) co-ordinates and SPOT coordinates. The shifts would then be used to correct the SPOT coordinates and the resulting co-ordinates would be in Clark (1880).
- Determine the shifts by computing the differences between SPOT and GPS (WGS84) co-ordinates and applying to shifts to correct SPOT images. In this case, the resulting SPOT co-ordinates would be in WGS84 (ellipsoid).
- Lastly computing the shifts by considering the differences between SPOT and Landsat co-ordinates. The Landsat co-ordinates were in WGS84 ellipsoid and hence the resulting SPOT co-ordinates would be in WGS84.

The third option was preferred because of the following:

- The GPS ground truth data used for accuracy assessment was based on WGS84 ellipsoid.
- The GPS co-ordinates that could have been used in case two above were not well distributed in the image.

It was however, to first confirm if the GPS and Landsat co-ordinates were in agreement. Table 5.6 is a comparison of Landsat and GPS co-ordinates. Column 6 and 7 show the differences in the eastings and northings of Landsat and GPS co-ordinates. Again considering the errors involved in estimating co-ordinates from the Landsat image as well as measuring GPS co-ordinates on the ground, it can be concluded that GPS and Landsat co-ordinates are similar.

Table **5.6** *Comparison between Landsat and GPS co-ordinates*

LANDSAT			GPS		DIFFERENCE	
ID	**Easting**	**Northing**	**Easting**	**Northing**	**ΔE (M)**	**ΔN (M)**
3	794060	9904957	794035	9904979	25	-22
5	794440	9899711	794433	9899732	7	-21
9	795067	9897739	795041	9897726	26	13
10	792887	9899045	792900	9899058	-13	-13
11	792452	9898552	792410	9898584	42	-32

Table 5.7 shows the shifts between SPOT co-ordinates and WGS84 co-ordinates while Table 5.8 shows the shifts that would have been obtained if the SPOT image was in CLARK (1880) ellipsoidal co-ordinates. The shifts in both cases are expressed as differences in the eastings (ΔE) or northings (ΔN).

Table 5.9 is a summary of the average shifts, ranges as well as maximum and minimum deviations from the means, for the results Table 5.6 and Table 5.7.

Table 5.7 Comparison between SPOT and WGS co-ordinates

	SPOT_CORDS		WGS_CORDS		DIFFERENCES	
ID	**Easting(m)**	**Northing(m)**	**Easting (m)**	**Northing (m)**	**ΔE (m)**	**ΔN(m)**
1	794164	9907333	793852	9907074	312	259
2	793926	9905521	793690	9905240	236	281
3	794295	9905239	794060	9904957	235	282
4	793972	9903414	793739	9903129	233	285
5	794647	9900008	794440	9899711	207	297
6	794833	9899766	794602	9899474	231	292
7	794965	9899766	794818	9899476	147	290
8	795377	9898304	795218	9897986	159	318
9	795211	9898011	795067	9897739	144	272
10	793024	9899312	792887	9899045	137	267
11	792640	9898853	792452	9898552	188	301
Average of co-ordinate differences					**211**	**286**

(The highlighted differences are the reliable differences based on the reliability (ease of scaling the co-ordinates).

Table 5.8 Comparison between Clark (1880) and WGS 84

	CLARK_CORDS		WGS_CORDS		DIFFERENCES	
ID	**Eastings (m)**	**Northings (m)**	**Eastings (m)**	**Northings (m)**	**ΔE (m)**	**ΔN (m)**
1	793675	9907300	793852	9907074	-177	226
2	793550	9905500	793690	9905240	-140	260
4	793600	9903425	793739	9903129	-139	296
5	794325	9900050	794440	9899711	-115	339
6	794450	9899725	794602	9899474	-152	251
7	794675	9899800	794818	9899476	-143	324
8	795075	9898350	795218	9897986	-143	364
9	794925	9898050	795067	9897739	-142	311
10	792725	9899350	792887	9899045	-162	305
11	792325	9898850	792452	9898552	-127	298
Average of co-ordinate differences					**-144**	**303**

Table 5.9 Shifts between (SPOT and WGS84) and (Clark and WGS84)

Variables	**SPOT_WGS**		**CLARK_WGS**	
	Easting (m)	**Northings (m)**	**Easting (m)**	**Northing (m)**
Mean shifts	211	286	-144	303
Maximum deviation from the mean	101	15	2	36
Minimum deviation from the mean	-67	-19	-33	-52
Range	168	34	35	88

The results in Table 5.7 and Table 5.8 were used for comparative analysis i.e. compare the shifts between SPOT and WGS84 as well as Clark 1880 and WGS84. The following observations were made:

- Shifts in eastings between Clark 1880 and WGS84 co-ordinates were negative for all the points used. This was in contrast to the shifts in easting co-ordinates between SPOT and WGS84.
- Shifts in the northings in both cases were positive and approximately having the same average shift.
- The range for the shifts in the easting co-ordinates between SPOT and WGS84 was relatively high.

The average shifts in Table 5.6 were used to correct SPOT co-ordinates. The method used involved the subtraction of the shifts from the co-ordinates of the origin. Unfortunately, when some points from the resulting image were checked with the GPS co-ordinates, the co-ordinates did not correspond well. Some GPS co-ordinates matched with their corresponding points from the corrected image while some points did not. The following conclusions were therefore made:-

- Scale distortions existed in the SPOT image
- The SPOT image did not have reliable co-ordinates

Another approach for rectifying SPOT image was therefore used. This approach was based on image-to- image rectification. Image-to-image rectification is possible if the reference image is already geo-referenced. In this case, the geo-referenced Landsat image was used as the reference image. Erdas Imagine 8.4 software was used for geometric rectification of SPOT image using Landsat image.

A detailed procedure for image-to-image rectification can be found in Erdas (1999). A first order polynomial was used for image rectification. Control point errors in X and Y were 0.13 and 0.84 respectively equivalent to a total error of 0.86, measured in pixels. The check point errors in X and Y were 1.8 and 0.74 with a total error of 2.02. The

rectified image was resampled using Nearest Neighbour algorithm. The rectified image was checked using the reference image (Landsat). This was done by displaying the two images in two different views. The two views were linked and compared using *inquire cursor*. Some features common to both images were used for comparison and there was a general agreement in the resulting coordinates. The resulting corrected SPOT image as well as Landsat image was used for further processing.

5.9.3 Correlation between original SPOT bands

Application of Principle Component Analysis (PCA) as data compression technique is suitable when there is high correlation in the original data. It was therefore necessary to examine the correlations between the original SPOT bands. The correlations between the SPOT bands were computed based on the following equation.

$$r_{kl} = \frac{\sum_{i=1}^{n}(k_i - \bar{k})(l_i - \bar{l})}{(n-1)\delta_k \delta_l}$$

Where r_{kl} is the correlation coefficient, $\delta_k \delta_l$ are the standard deviations of elements (DN values) in band k and l, $\bar{k}, \bar{l}$ are the mean values of elements in bands k and l while n is the count (number DN values). The formula for computing standard deviations as well as correlation coefficients can be found in most statistical books as well as some remote sensing text books for example Jenson, 1996), Swan (1995) and Erdas (1999). The inter-band correlations were computed with the help of spatial modeler functions in Erdas Imagine software. The results are summarized in Table 5.10

Table 5.10 Correlation between the original SPOT bands

SPOT BAND	XS1	XS2	XS3
XS1	1.000	0.884	0.676
XS2	0.884	1.000	0.493
XS3	0.676	0.493	1.000

The correlation matrix shows an expected value 1 along the leading diagonal (Correlation between one band and itself). The off diagonal elements show the inter-band correlations. The table shows a high correlation between band1 and band 2 i.e. 88% while the correlation between band 2 and band 3 is the lowest i.e. 49%. It was therefore suitable to apply PCA in order to remove the correlation between the original SPOT bands, hence, offering a way to optimize data viewing as well as analysis.

5.9.4 Principal Component Analysis (SPOT image)

PCA was applied to the original SPOT bands. The application of PCA resulted in an uncorrelated data set. Table 5.11 shows the correlation between the principal components (PCs) of the SPOT bands (rounded to 3dp). It can be clearly seen that no correlation exists between the derived components.

Table 5.11 Correlation between PCs of SPOT bands

	PC1	PC2	PC3
PC1	1.000	0.000	0.000
PC2	0.000	1.000	0.000
PC3	0.000	0.000	1.000

Table 5.12 shows the Eigen values for the three derived PCs. Eigen values express the amount of information content contained by each principal component. It can be observed that the first principal component (PC1) contains the highest information content of the image with the least information contained by the third principal component (PC3). Table 5.13 presents the results for the correlation between original SPOT bands (XS) and the derived Principal Components (PCs). The correlation between XS and PCs can be used to examine how each band "loads" or is associated with each principal component (Jensen, 1996).

Table 5.12 Eigen values from SPOT image

P-Component	Eigen Value	Percent (%)
PC1	892.330	94.934
PC2	44.214	4.704
PC3	3.401	0.362
TOTAL	**939.946**	**100.000**

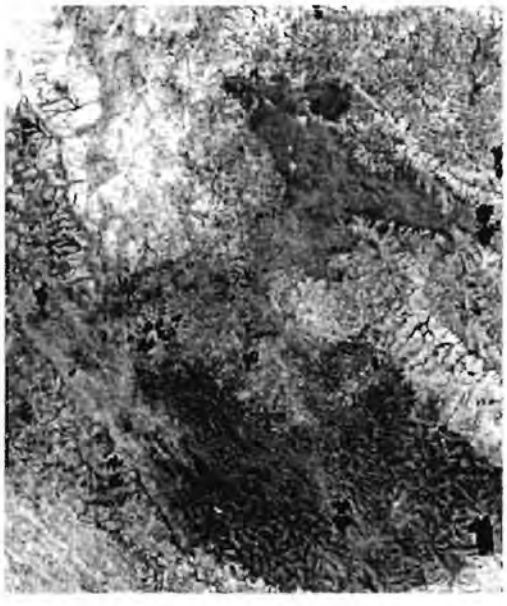
Table 5.13 Correlation between PCs and SPOT (XS) bands

	PC1	PC2	PC3
XS1	0.706	0.685	-0.174
XS2	0.527	0.810	0.254
XS3	0.999	-0.042	0.003

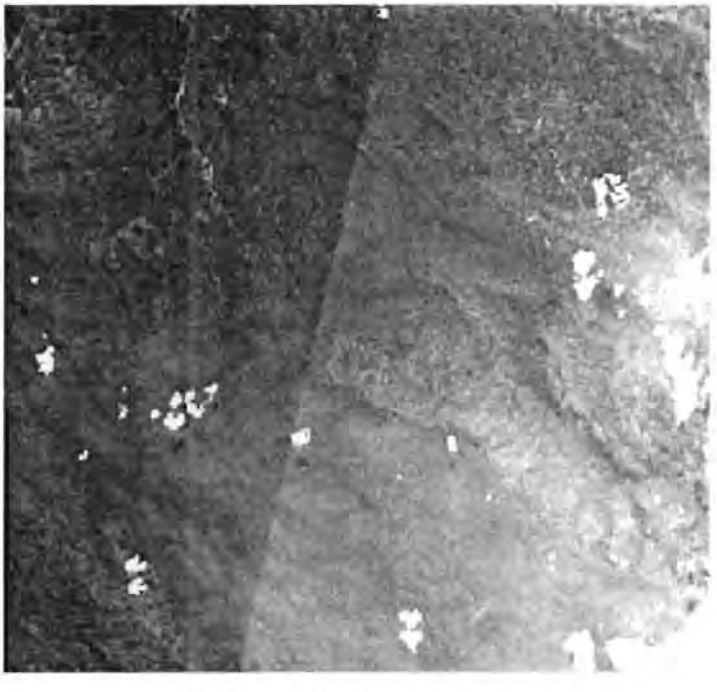
The results in Table 5.13 show that PC1 has a high association with XS3 as well as XS1. In this component, dark pixels are associated with the presence of water. Most of the area within the forest appears dark indicating the presence of water moisture in the forest canopy. The bright pixels are associated with high reflectance vegetation. The roads are not clearly delineated in this band due to the fact that XS2 has the lowest loading for PC1. XS2 is suitable for delineating bare ground areas from vegetated areas.

PC2 has a high association with XS2 and XS1. The high association of this component to XS2 enabled the distinction between vegetated areas and unvegetated areas. Roads as well as grassland areas appear bright in this component. PC3 is mainly noise. It has the highest loading with XS2 and this is why the roads are clearly seen. Figure 5-5 (a, b and c) show PC1, PC2 as well as PC3.

a) First principal component



b) Second principal component



c) Third principal component

Figure 5-5 Principal Component images derived from SPOT

5.9.5 Image ratios

Image ratios were computed as input data for the subsequent image classification. The derived ratios are summarized in Table 5.14. The first three ratios are the three possible image ratios for a three band SPOT image. The fourth ratio is the Normalized Difference Vegetation Index (NDVI). The ratios enhanced some features but none could distinguish all the features of interest uniquely. The following are some examples: -

- *Ratio band XS2/XS1* distinguished forest and grassland areas uniquely (which appear as bright pixels) but makes subsistence farmlands and forest area appear similar (relatively dark pixels).
- *Ratio band (XS3-XS2)/ (XS2+XS3)* enhanced brush land areas (very bright pixels), dense forest (dark pixels), less dense forest as well as subsistence farmlands. However this ratio band makes dense forest, grassland, bush land and built up area to appear similar (dark pixels).

Table 5.14 Image ratios of the SPOT bands

Ratio definition	Description
XS2/XS1	RED/GREEN
XS3/XS1	NIR/GREEN
XS3/XS2	NIR/RED
(XS3-XS2)/(XS2+XS3)	NIR-RED/(NIR+RED)

5.9.6 Texture Transform for the SPOT image

Texture transform was used for the derivation of a texture image used as one of the input bands prior to image classification. The texture band was derived using SPOT NIR band. The SPOT NIR band was chosen because of the high reflectivity of vegetation in this band. A variance algorithm with 3 x 3 window size was used to derive the texture band (TB). The resulting image was rescaled to a range of 0-255. For the texture band, areas with high heterogeneity (more coarse texture) appear as bright pixels. Forest edges for

example appear bright which suggest high heterogeneity while areas with less heterogeneity (homogenous) appear dark.

5.9.7 Layer Stacking of original and derived bands SPOT bands

Original SPOT XS bands were layer stacked with the derived bands. This resulted in an image having a total of ten bands (original as well as derived), also referred to as layers. Table 5.15 presents a summary of bands as well as the number used to identify each band. These numbers are used for easy analysis.

Table 5.15 Original and derived SPOT bands

ID	Band	Frequency ranges (μm)
1	XS1 (SPOT original band 1)	0.50-0.59
2	XS2 (SPOT original band2)	0.61-0.68
3	XS3 (SPOT original band3)	0.79-0.89
4	R2/1 (Band 2 divided by band1)	-
5	R3/1 (Band 3 divided by band1)	-
6	R3/2 (Band 3 divided by band2)	-
7	NDVI $(\frac{band3 - band2}{band3 + band2})$	-
8	PC1 (Principal component 1)	-
9	PC2 (principal component 2)	-
10	T (texture band)	-

5.9.8 Classification of the SPOT image

Image classification was carried out using Erdas Imagine 8.4 software. Supervised as well as unsupervised classifications of SPOT image were performed. The unsupervised and supervised procedures are explained below.

5.9.8.1 Unsupervised classification

Unsupervised classification of the SPOT image was performed using the ISODATA algorithm available in Erdas Imagine. The input for the unsupervised classification involved the specification of the following:

- *The number of spectral classes*: The classes that result from unsupervised classification are spectral classes which are based on natural groupings of the image values. The identity of the spectral class will not be initially known. Once the data are classified, the analyst attempts *a posteriori* to assign the spectral classes to the information classes. This is done by comparing classified data to some form of reference data (such as larger scale imagery, maps, or site visits) to determine the identity the spectral values. Some spectral classes may be difficult to assign to the information classes since some classes may represent mixed classes of the surface cover. For this study, the number of initial spectral clusters was set to 15. Theoretically, the number of spectral classes should be higher than the number of information classes required but should not be very high as this would cause the resulting image to be too cluttered and more difficult to interpret.
- *Convergence threshold*: this is the maximum percentage of pixels whose class values are allowed to be unchanged between iterations. When this number is reached, the ISODATA algorithm terminates. Some datasets may never reach the desired percentage of unchanged pixels between iterations and when this happens, the processing is terminated after a preset number of iterations and the parameters (convergence threshold, number of classes and maximum number of iteration are edited. A 95% convergence threshold was used in this study.
- *Maximum number of iterations*: the number of times the ISODATA is to classify pixels and re-calculate the cluster mean vectors. The ISODATA terminates when this number is reached. This happens even if the maximum threshold has not been reached. In this study the maximum number of iterations was set to 24.

The Figure 5-6 shows the results of unsupervised classification of SPOT data with 15 spectral classes. Each of the initial fifteen spectral classes was evaluated (merged,

deleted, renamed) and assigned to one of the predetermined information (LULC) classes. The resulting image as well as the information classes is shown in Figure 5-7

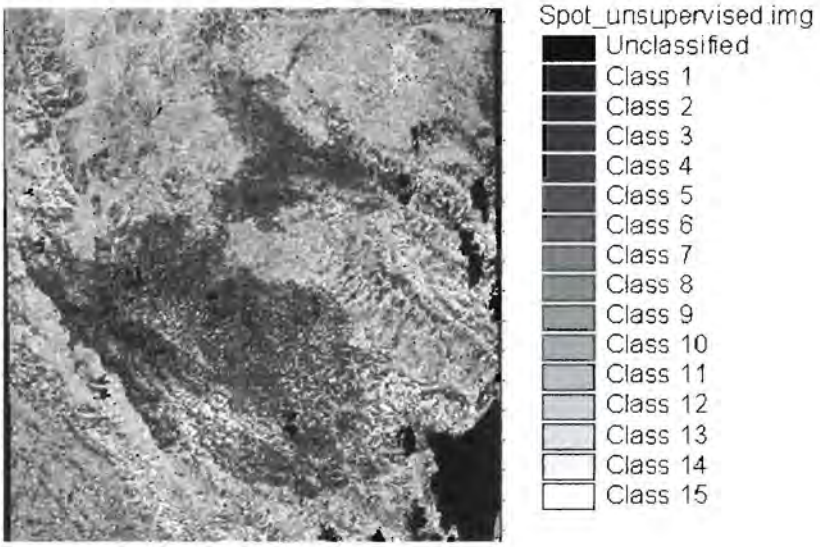


Figure 5-6 Spectral Classes from unsupervised SPOT image

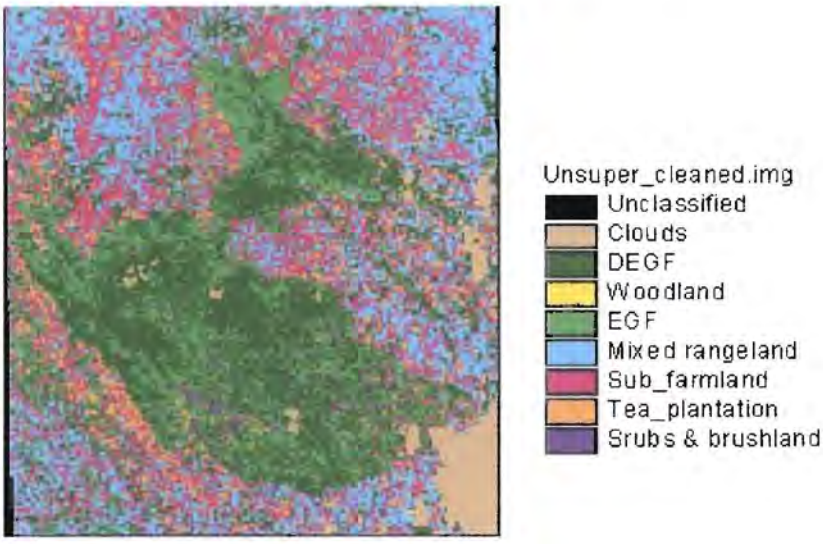


Figure 5-7 Information classes derived from unsupervised classified SPOT image

The following observations were made from the results in Figure 5-7

- Most of the pre-determined classes in the classification scheme were identified with the exception of the woodland class. The woodland class was classified as dense evergreen forest as it was not possible to separate it from the dense forest.

- There was also an overlap between tea plantation and brush land classes. Some areas thought to be brush land were classified as tea plantation. Since most of the brush land class is located in the forest, and in view of the fact that no tea farming exists in the forest, it was concluded that the brush land areas were erroneously classified as tea.

5.9.8.2 Supervised classification

In supervised classification, the analyst more closely controls the classification process than in unsupervised classification. The methodology used to perform supervised classification in this study was divided into three main areas i.e.:

1. Generation of signature sets: a signature is a set of data (or attributes) that defines a training sample. The attributes may include the name of the signature (class name), colour, output value (DN value), and order of processing the signature and parallelepiped limits. In this study, the signatures were generated with the help of ERDAS IMAGINE signature editor. The signature editor allows the creation, management, evaluation and editing of the signatures (.sig extension). The training areas were identified based with the help of ground truth data (maps, GPS data and aerial photographs). The Erdas Imagine software allowed on-screen digitizing and editing of training samples.

2. Evaluation of signature sets: This involves the tests carried out to help in determining whether the signature is a true representation of the pixels to be classified for each class. In ERDAS IMAGINE, the evaluation methods include: signature alarms, signature ellipse, contingency matrices, divergence measure, statistics and histograms.

   Signature alarm: the signature alarm allows one to compare an estimate classification of one or more signatures against the original data as it appears in the viewer (image processing window). The alarm utility creates a functional layer using parallelepiped

decision rule and the viewer allows the analyst to toggle between the image layer and the functional layer.

Signature ellipse: in this method, ellipses are generated using the mean and standard deviation of each signature file. It is also possible to generate parallelepiped rectangles and labels of each signature.

Contingency matrix: used to perform a quick classification of the sample pixels (i.e. pixels in the selected Areas Of Interests (AOIs) also referred to as training pixels) to determine what percentage of the sample pixels are actually classified as expected. These percentages are expressed in contingency matrix.

Divergence: measures the divergence (statistical distance) between two signatures and determine suitable band subsets for classification.

Statistics and histograms: in this method, the statistics and the histograms of the signatures are analyzed to make evaluations and comparisons.

The signatures were evaluated using separability indices, mean signature plots and feature space plots. Table 5.16 shows the results derived from the Transformed Divergence (TD) separability index algorithm in Erdas software. According to Jensen (1996), if the TD index is between 1900 and 2000, the classes are considered separable. Between 1700 and 1900, the separation is fairly good and below 1700, the separation is poor.

In Table 5.16 the separability between classes (except classes 6 and 7) is 2000, which indicates that the classes are separable. The separability of 1995 between classes 6 and 7 is within the range of 1700 and 1900 which also indicates a good separation of the classes. The TD algorithm did not give the results for the separability listings between the masked areas (clouds) and other classes. This could have been due to two reasons:

- The separability between class 1(clouds) and other classes was not significant and the program ignored this value.

- The masked areas did not have proper parameters (mean and variance) required for the computation of separability indices.

Table 5.16 Signature separability (transformed divergence)

Signature Separability Listing

File: c:/spot_2004/last_super.sig

Distance measure: Transformed Divergence

Using bands: 1 2 3 4 5 6 7 8 9 10

Taken 10 at a time

Class

2 DEGF
3 Woodland/bushland
4 EGF
5 Mixed_rangeland
6 Subsistence Farmalnds
7 TP
8 Shrubs/brushland

Best Minimum Separability

Bands	AVE	MIN	Class Pairs:						
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8
1 2 3 4	2000	1995	2000	2000	2000	2000	2000	2000	2000
5 6 7			2000	2000	2000	2000	2000	2000	2000
8 9 10			2000	2000	2000	2000	1995	2000	2000

Figure 5-8 shows the mean signature plots for the data set in the training areas for the different land cover types. The mean DN values of the data in the training area are plotted against the respective bands. The results in Figure 5-8 suggest that bands 2, 4, and 9 do not provide good separability between classes. If one single band was required for classifying an image either band 2, 4 or 9 would not be a good choice. It is, however, possible to obtain good separability when these bands (2, 4, and 9) are analyzed with other bands. This is because some classes may not be separable in one band but may be separable in other bands.

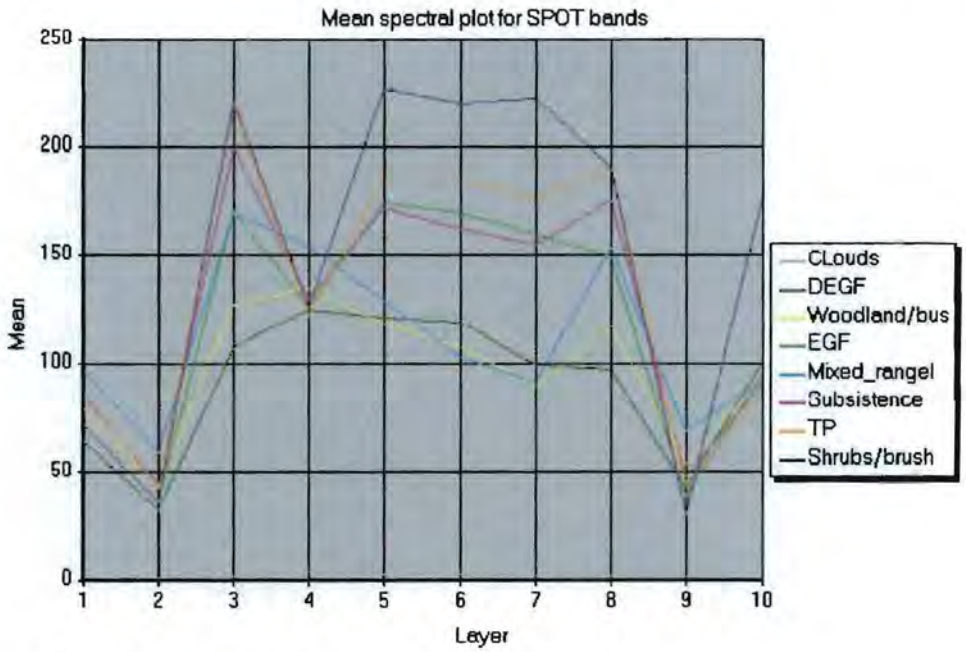


Figure 5-8 Spectral signatures (SPOT data)

The last signature evaluation method involved the generation of signature ellipses. Figure 5-9 shows signature ellipses for the different land cover types (see embedded text). The signature ellipses were obtained using the signature means and standard deviation of value two. It can be seen that good separation exits between the different signatures (land cover types). The labels for each land cover type are also shown.

Figure 5-9 Feature space objects (SPOT data)

3. Processing a supervised classification for SPOT image

This was performed after evaluating the signatures representing different land cover classes. One of the objectives of this study was to establish which bands (original or derived) were suitable for classification. Two approaches were used for selection of the best band combination (s) suitable for classification i.e.:

- Correlation approach.
- Signature separability indices

In this section, the correlation approach is discussed while the signature separability indices approach is presented in the later sections. By analyzing the correlation between bands, one is able to get an insight in the relationship between spectral information of the bands used. Separability indices such as divergence, transformed divergence, the Euclidian distance measure and Jeffreys Matusita are commonly used but one can not easily get a fundamental understanding of the relationship between bands being analyzed

Table 5.17 shows the correlations of the ten original and derived SPOT bands used for analysis of SPOT data. Previous research by Ayanz (1996) showed that maximum accuracy was obtained with a six band combination, thereafter the accuracy starts decreasing. Based on this study, six bands were derived for use in supervised image classification.

Table 5.17 Correlation between original and derived SPOT bands

Band	S1	S2	S3	R2/1	R3/1	R3/2	NDVI	PC1	PC2	Text (4)
S1	1	0.884	0.676	0.359	0.237	0.081	0.149	0.706	0.685	-0.083
S2	0.884	1	0.493	0.745	0.08	-0.2	-0.081	0.527	0.81	-0.035
S3	0.676	0.493	1	0.05	0.872	0.739	0.812	0.999	-0.043	0.087
R2/1	0.359	0.745	0.05	1	-0.158	-0.511	-0.364	0.075	0.636	0.039
R3/1	0.237	0.08	0.872	-0.158	1	0.926	0.977	0.851	-0.496	0.179
R3/2	0.081	-0.2	0.739	-0.511	0.926	1	0.985	0.712	-0.657	0.142
NDVI	0.149	-0.081	0.812	-0.364	0.977	0.985	1	0.787	-0.601	0.161
PC1	0.706	0.527	0.999	0.075	0.851	0.712	0.787	1	0	0.08
PC2	0.685	0.81	-0.043	0.636	-0.496	-0.657	-0.601	0	1	-0.154
Text (4)	-0.083	-0.035	0.087	0.039	0.179	0.142	0.161	0.08	-0.154	1

The best six band combination was obtained by ranking the correlations of the band pairs in Table 5.17 . The correlations (entries) along the leading diagonal were not ranked since it involves the correlation of one band and itself. The results of Table 5.18 show the ten band pairs with the least correlation. The entries of Table 5.17 were sorted out in ascending order and the pairs with the least correlation were extracted. To get the best six bands the following methodology was adopted:

- The first two band pairs in the first row, with the least correlation (i.e. 0) were selected.
- The second band pair in the second row (i.e. bands with the second least correlation) was also selected.
- In the third row, only band 4 was considered since band 10 had already been selected in row 2.
- In the fourth row, only band 3 was considered because band 9 had already been considered in row 1. This approach resulted in six bands with the least correlation that were used for supervised classification (i.e. band 2, 3, 4, 8, 9, 10).

Table 5.18 Band pairs with least correlation

ID	Band A	Band B	Correlation
1	8	9	0.000
2	2	10	0.035
3	4	10	0.039
4	3	9	0.043
5	3	4	0.050
6	4	8	0.075
7	2	5	0.080
8	8	10	0.080
9	1	6	0.081
10	2	7	0.081

The resulting six-band combination was used for performing supervised classification with Maximum Likelihood Classifier (also called the Bayesian Classifier) and Parallelepiped Classifier.

The classification algorithms were used in a two step process:

- The first step involved the application of Maximum Likelihood Classifier (MLC) as the classification algorithm.

The Maximum Likelihood Classifier (MLC) was chosen because of the following reasons:

- It is the most accurate classification algorithm in Erdas Imagine system (Erdas, 1999)
- It takes the variability of the classes into account by using the covariance matrix

The equation for the MLC is as follows:

$$D = \ln(a_c) - [0.5\ln(|Cov_c|)] - [0.5(X - M_c)^T (Cov_c - 1)(X - M_c)]$$

Where:

D = weighted distance (likelihood)

c = a particular class

X = the measurement vector of the candidate pixel

M_c = the mean vector of the sample of class c

a_c = percentage probability that any candidate pixel is a member of class c (defaults to 1.0, or is entered from *a priori* knowledge)

Cov_c =the covariance matrix of the pixel in the sample of class c

$|Cov_c|$ = determinant of Cov_c

Cov_c -1 = inverse of Cov_c (matrix algebra)

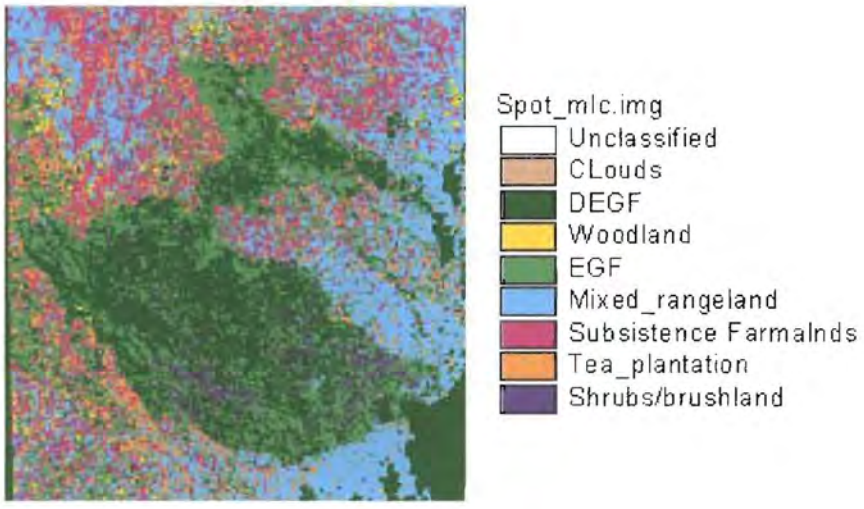
ln = natural logarithm function

T = transposition function

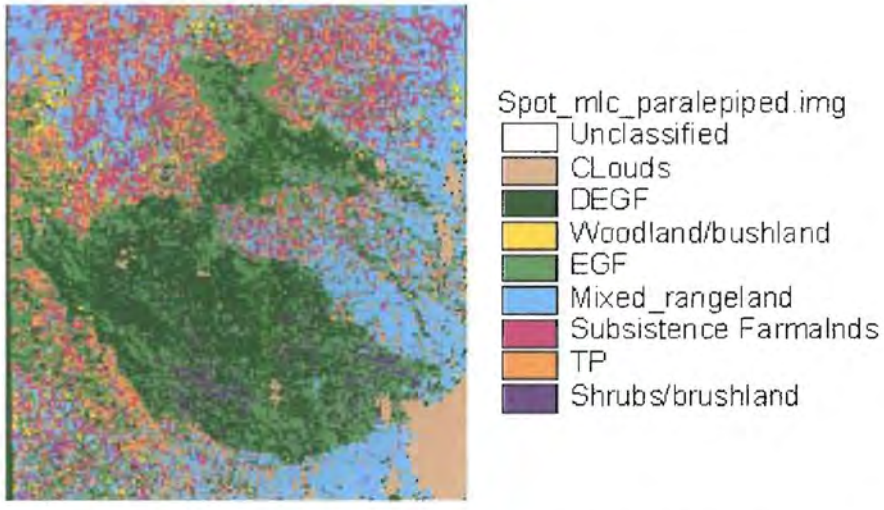
- In the second step, a modification of the first step was employed and this involved using both Parallelepiped (box) and Maximum likelihood classifies. In this approach, the Parallelepiped classifier classifies the data within the training areas based on the size of the Parallelepiped (length and width). The length and width of the Parallelepiped may be defined by the maximum and minimum DN values in the respective bands. The Parallelepiped classifier then uses two decisions based on inclusion or exclusion i.e. when the pixel value falls within the box it is defined as belonging to that class while if it is not within the box then it is outside the class boundaries. The DN values that will be unclassified by the parallelepiped method are then classified using MLC. This method is recommended especially when the training data is multi- model (having more than one mode). A comparison of the results

(specifically the accuracies) of these two methods was made. The results are presented in chapter six. Figure 5-10 shows the supervised classified images that were obtained using the two sets of the classification algorithms described above.

Seven land cover classes (*see legend*) were delineated including woodland class which was not possible to delineate using unsupervised classification. In this method, the masked areas (clouds) were erroneously classified as dense evergreen forest. The tea plantation (TP) and the brush land classes were separated properly i.e. there was no overlap between these two classes unlike the case for the unsupervised classification. This was because the classification was controlled by the analyst unlike the case for unsupervised classification.



a) *Classification with Maximum likelihood classifier*



b) *Classification with parallelepiped and maximum likelihood classifier*

Figure 5-10 Results for supervised SPOT image

5.10 Processing of Landsat Image

5.10.1 Layer stacking of the Landsat (ETM+) and segmentation

Six of the original Landsat TM bands (excluding band six) were layer stacked and a subset of the image covering the study area was extracted. Landsat TM band 6 (thermal band) was not included in the analysis because it has a coarse resolution and it is not normally used for vegetation studies. The clouds and cloud shadows were masked out and the resulting image was used for further analysis.

5.10.2 Correlation between Original Landsat Bands

The correlations between the original Landsat bands were computed using the same approach as used for the SPOT data. Table 5.19 shows the correlations of the original Landsat bands. The table clearly shows high correlation between most of the bands. This means that redundant information in the data set exists. It was therefore suitable to apply PCA to remove the correlation between bands. The results of the correlation analysis are shown in Table 5.20

Table 5.19 Correlation between original Landsat ETM bands

Band	TM1	TM2	TM3	TM4	TM5	TM7
TM1	1.00	0.90	0.83	0.40	0.73	0.70
TM2	0.90	1.00	0.93	0.57	0.89	0.85
TM3	0.83	0.93	1.00	0.37	0.88	0.93
TM4	0.40	0.57	0.37	1.00	0.64	0.41
TM5	0.73	0.89	0.88	0.64	1.00	0.94
TM7	0.70	0.85	0.93	0.41	0.94	1.00

5.10.3 Principal Component Analysis (PCA)

The results of the PCA are shown in Table 5.20.

Table 5.20 Principal components for Landsat image

Component	Eigen Value (EV)	EV%
PC1	1266.13	80.52
PC2	254.40	16.18
PC3	35.14	2.23
PC4	10.13	0.64
PC5	4.54	0.29
PC6	2.13	0.14
SUM	**1572.47**	**100.00**

According to the results in, PC1 contains the highest information content of the original image i.e. 80.52% while PC2 contributes 16.18%. Thus PC1 and PC2 together contribute 96.8% of the total information in the scene. These two principal components were used as one of the input bands prior to image classification. The remaining four components contributing 3.2% of the information content of the original Landsat data were not used as input bands for supervised classification. Table 5.21 shows the correlations of the resulting principal components of the Landsat image. It is clear that no correlation exists between the resulting components

Table 5.21 Correlation of the resulting components

COMPONENT	PC1	PC2	PC3	PC4	PC5	PC6
PC1	1.000	0.000	0.000	0.000	0.000	0.000
PC2	0.000	1.000	0.000	0.000	0.000	0.000
PC3	0.000	0.000	1.000	0.000	0.000	0.000
PC4	0.000	0.000	0.000	1.000	0.000	0.000
PC5	0.000	0.000	0.000	0.000	1.000	0.000
PC6	0.000	0.000	0.000	0.000	0.000	1.000

5.10.4 Ratios

Seven different image ratios were derived. These ratios were derived based on the studies carried out by Mohammed (1997), Trisurat et al (2000) and Ayanz (1997). Mohammed (1997) gave a detailed analysis of features delineated by some of the ratios given in Table 5.22.

Table 5.22 Derived image ratios for Landsat ETM image

Number	Ratio
1	TM4/TM2
2	TM4/TM3
3	TM5/TM4
4	TM7/TM5
5	TM7/TM3
6	TM5/TM3
7	NDVI

5.10.5 Derivations of Tasseled Cap Transform bands

Tasseled Cap Transform (TCT) was performed on the six TM bands resulting in six new layers (in this study also referred to as derived bands). Of the six derived bands, the first two were used for image analysis since they are reported to contain most of the information present in the original scene (Jensen, 1996). These derived bands are referred to as the Soil Brightness Index (SBI) and Greenness Index (GI).

5.10.6 Derivation of Texture band

The derivation of the Texture band was performed using TM band 4 since it is sensitive to changes in vegetation. It should be noted that the near infrared band of SPOT (XS3) and the near infrared band of Landsat (TM4) operate in the same region of electromagnetic spectral (i.e. 0.79-0.89,0.76 - 0.90)μm respectively. A variance algorithm

with 3x3 size moving window was used to derive the texture band. The resulting band was incorporated for image classification.

5.10.7 Layer Stacking of original and derived bands

The original Landsat TM bands were layer stacked with the derived bands (Principal Components, Image Ratios, Tasseled Cap Indices and Texture Transforms). A total of 18 layers (representing original and derived bands) was obtained and used for further image analysis. Table 5.23 shows the different bands and the corresponding numbers (ID) used in the subsequent sections.

Table 5.23 Original and derived Landsat ETM+ bands

ID	BAND
1	TM1
2	TM2
3	TM3
4	TM4
5	TM5
6	TM7
7	R4/2
8	R4/3
9	R5/3
10	R5/4
11	R7/3
12	R7/5
13	NDVI
14	PC1
15	PC2
16	SBI
17	GI
18	TEXT4

5.10.8 Image classification

Image classification was carried out using both unsupervised and supervised approaches. The process was similar to the one adopted in the classification of the SPOT image. The next sections give an illustration of the main steps followed during classification.

5.10.8.1 Unsupervised classification

Unsupervised classification was performed using all the bands (original and derived). Fifteen initial clusters were chosen which were analyzed and reduced to 8 information classes after evaluation. The resulting classified map was recoded into 8 classes and the results are shown in Figure 5-11.

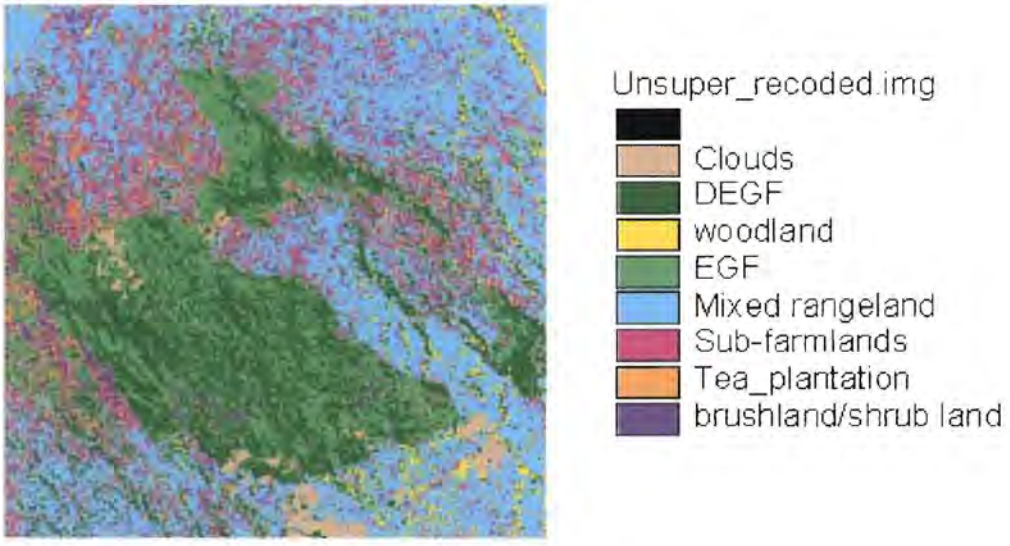


Figure 5-11 Unsupervised Landsat (ETM+) image

The following observations were made from the results for supervised classification:

- The woodland class was identified unlike the case for unsupervised classification of SPOT image where the woodland class was classified as DEGF.
- The brush land and tea plantation were identified unlike the case for unsupervised classification of SPOT where the two classes overlapped.

- The areas that were masked (clouds and cloud shadows) and referred in here as clouds were classified as well and are shown in the legend. These areas essentially have DN values of 0.

5.10.8.2 Supervised Classification of Landsat image

A similar approach used for supervised classification of the SPOT image was used for supervised classification of Landsat image i.e.

- Training site development and signature generation.
- Evaluation of the signatures.
- Choosing the desired bands for classification.
- Performing supervised classification.

Figure 5-12 shows a graph of the mean signatures derived from Landsat image (i.e. a plot of the means of the pixels within trainings of each land cover class against the respective band). Analysis of the signature plot suggests that separability exists for at least more than one signature in all the bands. A general examination of the graph suggests that band 4 separates all the signatures, while bands 10 and 18 seem unsuitable for classification. The signature plot method can be used to examine the signatures but may not be suitable for identification of the best band combination for classification. This assertion is supported by the results reported by Ayanz (1996) who compared the performance of signature plot and spectral indices as methods of selecting best band combination for classification. Signature plot method generally gave results of low accuracies in comparison to the results obtained when band selection is performed with spectral indices.

It should also be noted that signature plot helps to analyze signatures on band to band basis. It is therefore possible to obtain the signatures that are not separable in one band but separable if two bands were used.

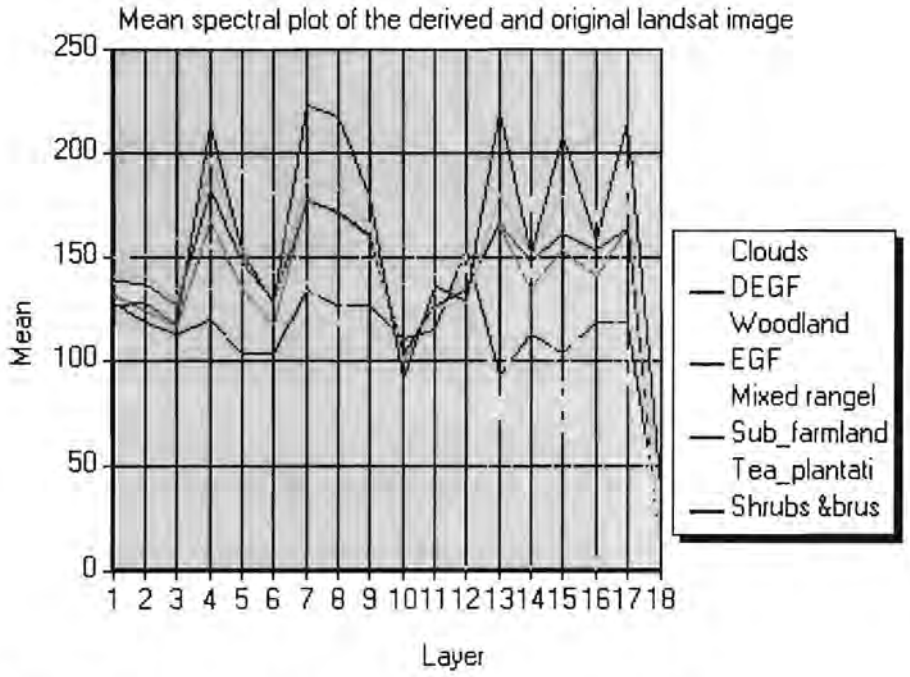


Figure 5-12 Mean spectral signatures (Landsat Data)

Figure 5-13 shows the signature ellipse plot as well as the respective signature means. A standard deviation of two was used to generate the signature ellipses. It can be observed from these ellipses that small overlap exists between the classes especially between woodland and mixed rangeland as well as tea plantation and subsistence farmland. Since it is difficult to obtain hard boundaries defining land cover features, these overlaps are expected. The signatures were however, considered suitable for classification.

Figure 5-13 Feature space objects (Landsat)

Processing a supervised classification of Landsat data

The methodology adopted for processing supervised classification of Landsat was similar the one used for supervised classification of SPOT data. This involved selection of the best six band combination as well as selection of classification algorithm. The best band combination (s) for classification was selected using the following methods:

- Correlation approach
- Signature separability indices

(See related discussion under supervised classification of SPOT data)

Correlation approach

In this section, the correlation approach is discussed while the signature separability indices approach is presented in the later sections. The correlation approach was based on the analysis of the correlations between the original and derived Landsat bands shown in Table 5.24. The correlations were sorted out in ascending order and the bands with the least correlation were selected. Table 5.25 shows the ten least correlated pairs of bands.

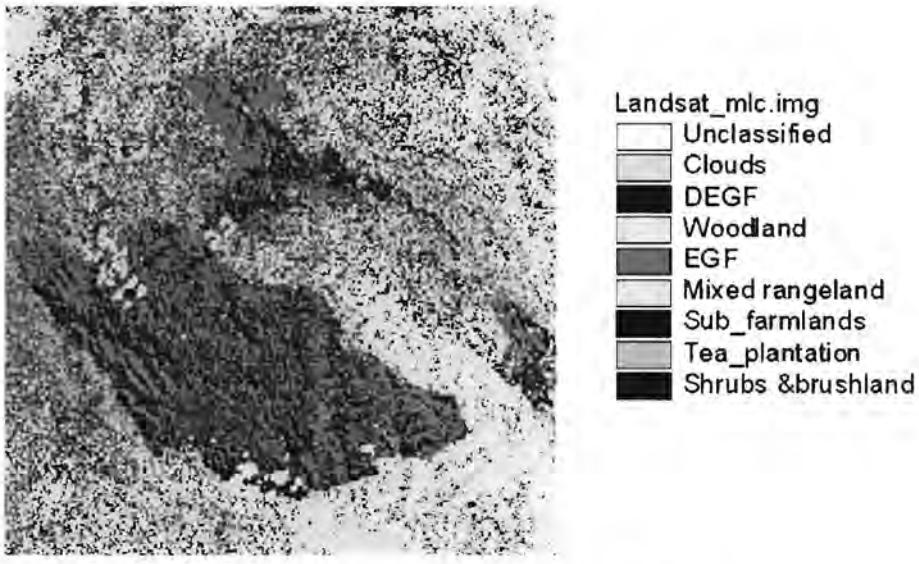
Table 5.24 Correlation between original and derived Landsat ETM+ bands

BAND	TM1	TM2	TM3	TM4	TM5	TM7	R4/2	R4/3	R5/3	R5/4	R7/3	R7/5	NDVI	PC1	PC2	TC1	TC2	TEXT4
TM1	1	0.9	0.83	0.4	0.73	0.7	-0.08	-0.17	0.49	0.57	0.46	-0.25	-0.17	0.75	-0.15	0.76	-0.28	-0.1
TM2	0.9	1	0.93	0.57	0.89	0.85	0.06	-0.07	0.67	0.62	0.64	-0.33	-0.07	0.91	-0.07	0.92	-0.17	-0.1
TM3	0.83	0.93	1	0.37	0.88	0.93	-0.13	-0.31	0.59	0.8	0.7	-0.1	-0.32	0.9	-0.33	0.86	-0.41	-0.1
TM4	0.4	0.57	0.37	1	0.64	0.42	0.86	0.76	0.78	-0.11	0.44	-0.8	0.76	0.67	0.74	0.77	0.69	-0
TM5	0.73	0.89	0.88	0.64	1	0.94	0.23	0.04	0.9	0.69	0.87	-0.41	0.04	0.99	-0.03	0.97	-0.06	-0.1
TM7	0.7	0.85	0.93	0.42	0.94	1	-0.02	-0.21	0.75	0.83	0.91	-0.09	-0.22	0.94	-0.3	0.88	-0.31	-0.1
R4/2	-0.08	0.06	-0.13	0.86	0.23	-0.02	1	0.97	0.54	-0.5	0.15	-0.76	0.96	0.25	0.93	0.37	0.94	0.07
R4/3	-0.17	-0.07	-0.31	0.76	0.04	-0.21	0.97	1	0.4	-0.65	-0.02	-0.74	1	0.07	0.97	0.2	0.99	0.09
R5/3	0.49	0.67	0.59	0.78	0.9	0.75	0.54	0.4	1	0.43	0.83	-0.64	0.39	0.87	0.28	0.88	0.31	-0.1
R5/4	0.57	0.62	0.8	-0.11	0.69	0.83	-0.5	-0.65	0.43	1	0.72	0.22	-0.66	0.65	-0.73	0.53	-0.72	-0.1
R7/3	0.46	0.64	0.7	0.44	0.87	0.91	0.15	-0.02	0.83	0.72	1	-0.12	-0.03	0.85	-0.18	0.78	-0.12	-0.1
R7/5	-0.25	-0.33	-0.1	-0.8	-0.41	-0.09	-0.76	-0.74	-0.64	0.22	-0.12	1	-0.75	-0.39	-0.75	-0.5	-0.71	0.04
NDVI	-0.17	-0.07	-0.32	0.76	0.04	-0.22	0.96	1	0.39	-0.66	-0.03	-0.75	1	0.06	0.97	0.19	0.99	0.08
PC1	0.75	0.91	0.9	0.67	0.99	0.94	0.25	0.07	0.87	0.65	0.85	-0.39	0.06	1	0	0.99	-0.04	-0.1
PC2	-0.15	-0.07	-0.33	0.74	-0.03	-0.3	0.93	0.97	0.28	-0.73	-0.18	-0.75	0.97	0	1	0.15	0.98	0.08
SBI	0.76	0.92	0.86	0.77	0.97	0.88	0.37	0.2	0.88	0.53	0.78	-0.5	0.19	0.99	0.15	1	0.09	-0.1
GI	-0.28	-0.17	-0.41	0.69	-0.06	-0.31	0.94	0.99	0.31	-0.72	-0.12	-0.71	0.99	-0.04	0.98	0.09	1	0.09
TEXT4	-0.14	-0.12	-0.12	-0.01	-0.11	-0.1	0.07	0.09	-0.07	-0.13	-0.06	0.04	0.08	-0.1	0.08	-0.1	0.09	1

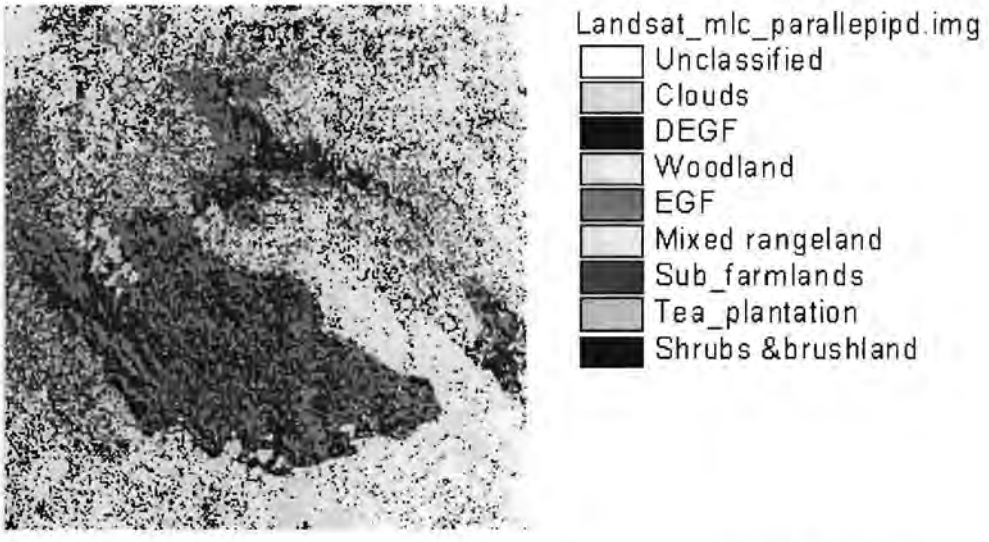
Table 5.25 Ten least Correlated band pairs of the Landsat bands

ID	Band A	Band B	Correlation
1	14	15	0
2	4	18	0.01
3	6	7	0.02
4	8	11	0.02
5	5	15	0.03
6	11	13	0.03
7	5	8	0.04
8	5	13	0.04
9	11	18	0.04
10	12	18	0.04

Analysis of the results in Table 5.25 led to the conclusion that bands 4, 6, 7, 14, 15 and 18 were suitable for classification (least correlated bands). The six-band combination selected based on correlation analysis was used for supervised classification of Landsat data. The results of the supervised classification are shown in Figure 5-14.



a) *Classification with Maximum likelihood classifier*



b) *Classification with Parallelepiped and Maximum likelihood classification*

Figure 5-14 Supervised classification of Landsat (correlation approach)

From the results in Figure 5-14, it can be observed that all the seven predetermined land cover types were classified. However, the woodland class was over classified. For example areas that were originally masked and labeled as clouds were classified as woodland. It therefore means that with the six band combination, woodland was not separated from the masked areas. Other land cover classes were, however, well delineated.

5.11 Determination of separability using separability indices

One of the objectives of the study was to compare the correlation and spectral indices approaches for selecting suitable combination of bands for supervised classification. In the previous sections, the correlation approach was explored. This section explores the application of the spectral indices for suitable band selection. Two separability indices were explored i.e.:

- Transformed Divergence (TD) separability index.
- Euclidian Distance (ED) separability index.

According to Jensen (1996), the TD gives the better results as compared to ED. This is because it takes into account the variance of the data set. It also weighs the separability between classes in such a way that as the distance between land- cover classes increases, the weight corresponding to that distance decreases exponentially. The weighting of the separability is advantageous in that it minimizes the effect of outlying separable classes which tend to weight the average divergence upwards in a misleading fashion. However, Jensen (1996) reports that TD produces better results when the training data is normally distributed. Obtaining data that is normally distributed is always difficult. This is due to the heterogeneity of the land cover classes as well as generalization of the land cover classes. It is also possible to have training data that has a normal distribution in some bands but a non-uniform distribution in other bands. Figure 5-15 shows two histograms derived from the training data for the mixed rangeland class.

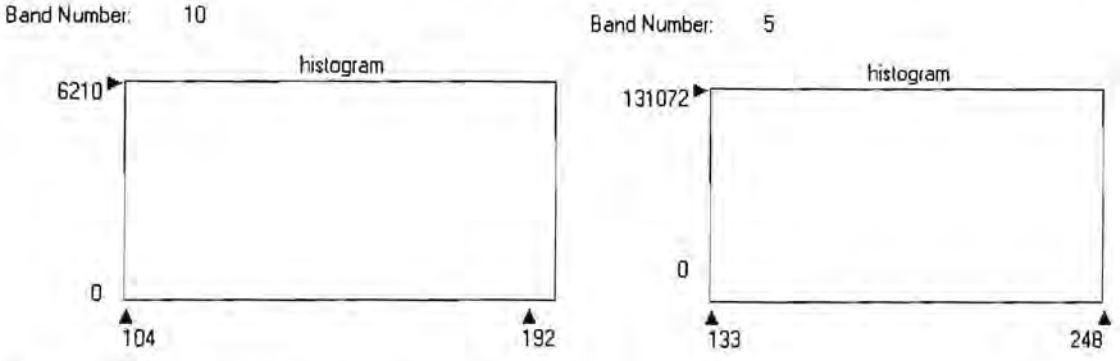


Figure 5-15 Histograms for mixed rangeland

The histogram on the left is derived from band 10 (i.e. band 5/ band 4) while the histogram on the right is derived from original ETM band 5. Two observations can be made from the above histograms:

- The histogram on the left is normally distributed which represents one class.
- The histogram on the right hand side is bi-modal which indicates that there are two classes in the class labeled mixed rangeland.

From the above observations, it is expected that the training data may not totally be uni-modal in all the bands.

As mentioned earlier on, ED does not take into account the variability of the data but according to Erdas (1999), ED may be appropriate to use when the training data is not normally distributed. This study therefore examines the two separability indices and a comparison of the results is made.

The separability index algorithms in Erdas Imagine software compute the separability between any two land cover classes based on the following:

- Best minimum separability.
- Best average separability.

For the best minimum separability, the divergence between the land cover classes is computed for given combinations. The band combination with the best minimum separability is used. Likewise for the best average separability, the average separability of the land cover classes for given combinations is computed and the combination with the best average separability is chosen. *Section 5.11.1* and *section 5.11.2* present the optimum band combinations of the SPOT and Landsat (ETM+) image respectively

5.11.1 Optimum band combinations for SPOT image

Table 5.26 and Table 5.27 show the optimum band combinations for SPOT image based on best minimum separability and best average separability of the Euclidian and Transformed Divergence Vegetation indices. From Table 5.26 and Table 5.27, the TD

separability index selected the same band combination for the best minimum and average separability. The results are shown in appendix c. It therefore follows that three optimum band combinations were derived for the SPOT image.

Table 5.26 Optimum bands using best minimum separability (SPOT)

Spectral index	Optimum band combinations
Euclidian Distance (ED)	1,3,6,7,8,9
Transformed Divergence (TD)	1,2,3,4,5,6

Table 5.27 Optimum bands obtained best average separability (SPOT)

Spectral index	Optimum band combinations
Euclidian Distance (ED)	3,4,5,6,7,8
Transformed Divergence (TD)	1,2,3,4,5,6

5.11.2 Optimum band selection for Landsat (ETM+) image

Table 5.28 and Table 5.29 show the optimum band combinations for Landsat image. Like the case for SPOT image, the TD separability index selected the same band combinations for the best minimum separability and best average separability. It therefore follows that three band combinations were derived for the SPOT image.

Table 5.28 Optimum bands obtained using best minimum separability (Landsat)

Spectral index	Optimum band combinations
Euclidian Distance (ED)	4,5,13,14,15,16
Transformed Divergence (TD)	1,2,3,4,5,6

Table 5.29 Optimum bands obtained with best average separability (Landsat)

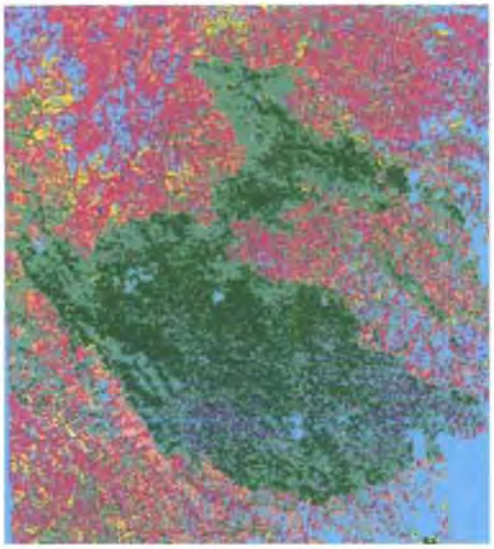
Spectral index	**Optimum band combinations**
Euclidian Distance (ED)	4,5,7,8,13,15
Transformed Divergence (TD)	1,2,3,4,5,6

5.11.3 Supervised classification -spectral index approach

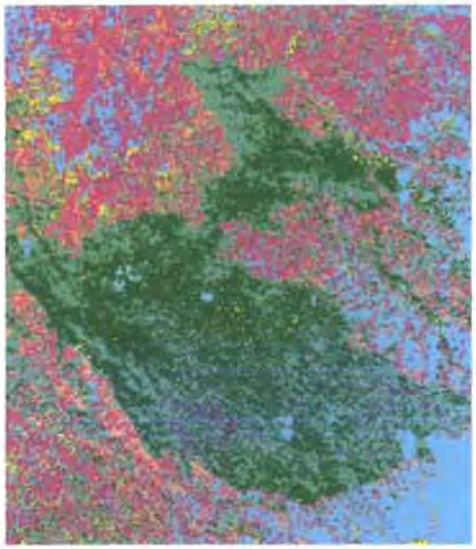
A total of six optimum band combinations were selected based on Euclidian distance and transformed divergence separability indices. These band combinations were used for performing supervised classification. The band combinations were classified using:

- Maximum Likelihood Classifier (MLC), also called the Bayesian Classifier
- Parallelepiped and Maximum Likelihood Classifier

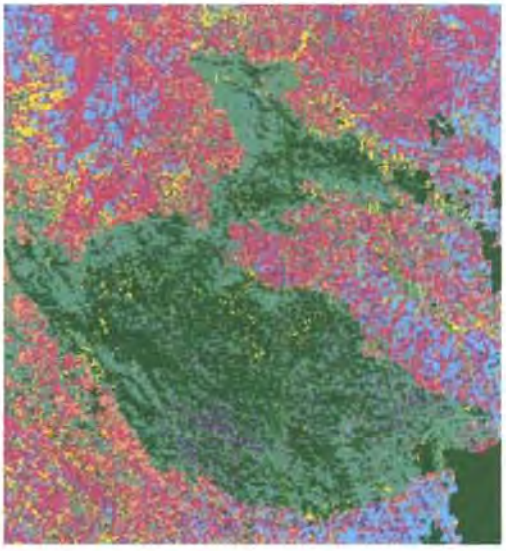
The results are shown in the Figure 5-16 (a), (b) and (c)



a) Best minimum separability (ED)



b) Best average separability (ED)

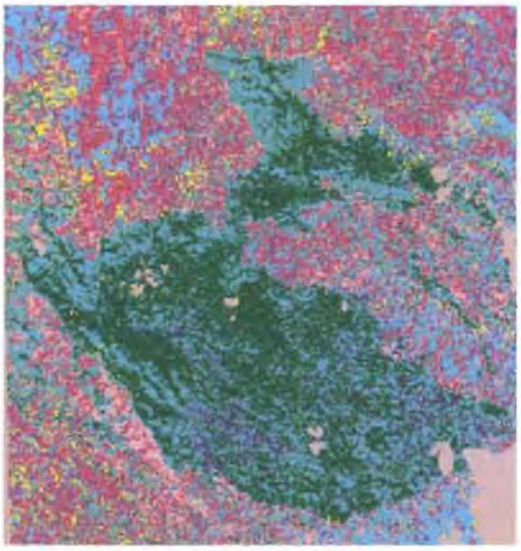


c) Transformed Divergence (TD)

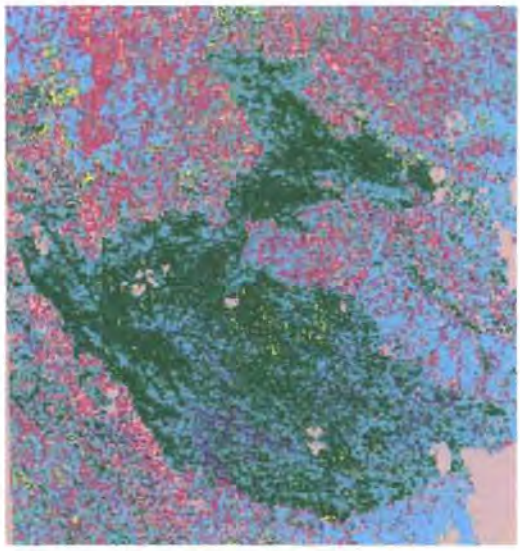
Figure 5-16 Supervised classification of SPOT using MLC

The results in Figure 5-16 show that no significant difference exists between the classified maps resulting from Euclidian distance (ED) separability indices (Best minimum and best average). In these two cases (Figure 5-16*(a) and (b)*), the masked areas were erroneously classified as mixed rangeland). The results for transformed divergence separability index Figure 5-16(c) are similar to the results from Euclidian

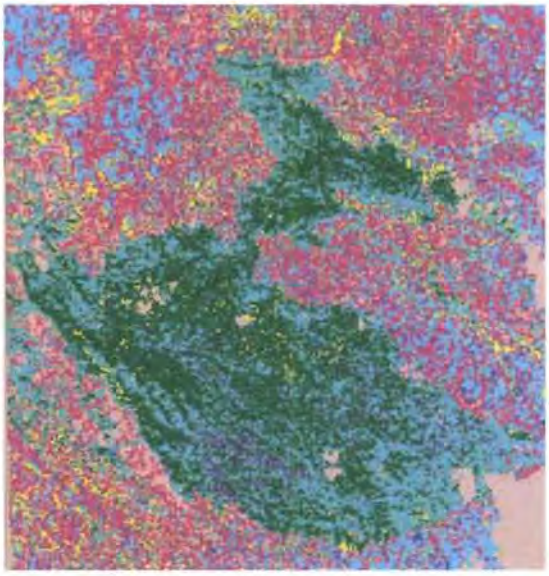
distance separability index. The main area of departure is the masked areas (clouds) which were erroneously classified as DEGF.



a) Best minimum separability (ED)



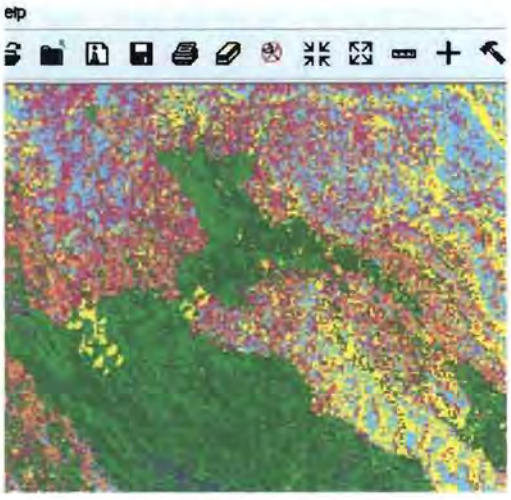
b) Best average separability (ED)



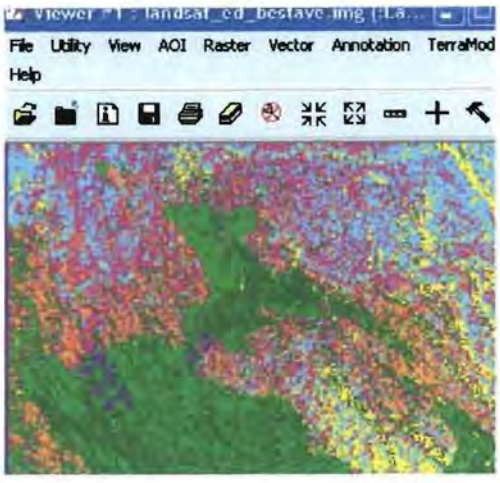
c) Transformed Divergence (TD)

Figure 5-17 Supervised classification of SPOT using MLC and Parallelepiped

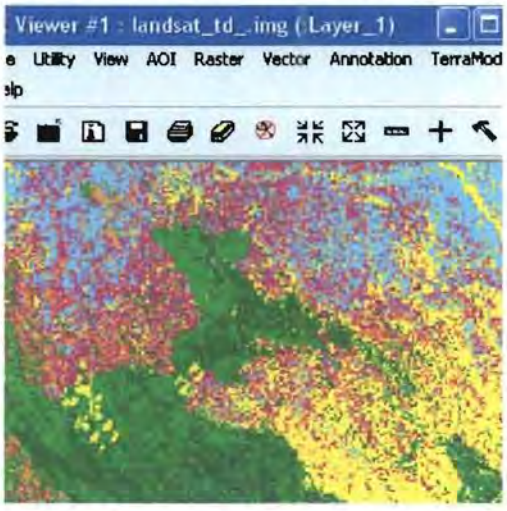
The results of the supervised classification of SPOT image with parallelepiped and MLC (Figure 5-17) show that no significant difference exists between the different approaches used (ED or TD). However, this approach identified the masked areas as a separate class unlike the case for supervised classification with only MLC.



a) Minimum separability (ED)



b) Average separability (ED)

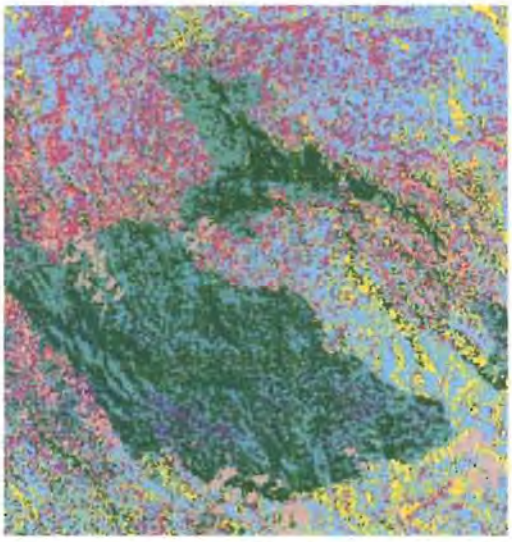


c) Transformed Divergence (TD)

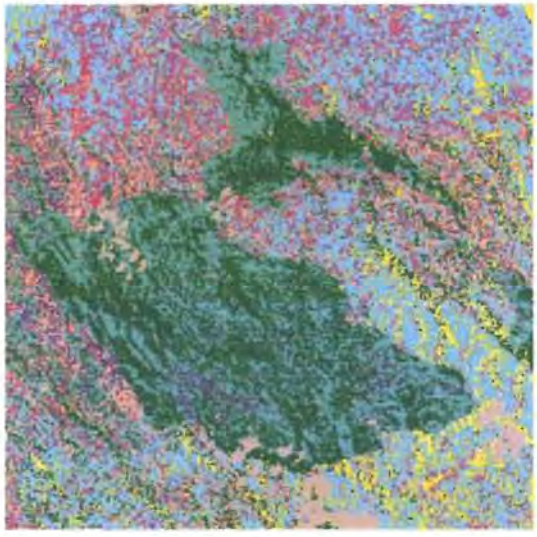
Figure 5-18 Supervised classification of Landsat using MLC

Figure 5-18 shows the results of the supervised classification of Landsat data using MLC. Similar to the supervised classification of SPOT image with MLC, this approach did not identify the masked areas as a separate class. For the case of ED (minimum separability), the masked areas were classified as woodland while for the case of ED (average separability), the masked areas where classified as brush land. The TD approach produced results similar to the results from ED (best minimum separability) i.e. the masked areas were classified as woodland. However, the woodland class seemed to be

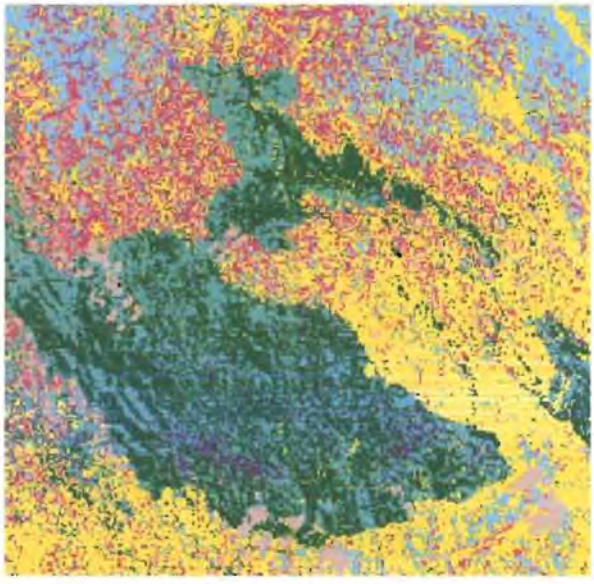
over classified as well. It should be noted that all the seven information classes desired where identified in the above cases



a) Best minimum separability



b) Best average separability



c) Transformed divergence

Figure 5-19 Supervised classification of Landsat with MLC and Parallelepiped

Figure 5-19 shows the results of the supervised classification of Landsat using a combination of MLC and parallelepiped. In this approach, the seven information classes were identified as in the previous approaches. The masked areas were identified as a

separate class. The results for supervised classification with optimum set of bands derived using TD index, over classified the woodland class.

6 ACCURACY ASSESSMENT

6.1 Introduction

This chapter presents the results of accuracy assessment for classified SPOT and Landsat images. A discussion of error matrices, overall accuracies, producer accuracy, user accuracy and Kappa static as measures of classification accuracy is presented in this chapter.

6.2 Accuracy assessment of the classified Landsat and SPOT images

The accuracies for the classified Landsat and SPOT images were assessed using error matrices (also referred to as confusion matrices). According to Jensen (1996), an error matrix, is a square array of numbers laid out in rows and columns that expresses the number of sample units (i.e. pixels, clusters of pixels or polygons) assigned to a particular category relative to the actual category as verified from the field. The columns of an error matrix represent the reference data while the rows indicate the classification obtained from the remotely sensed data.

From an error matrix, measures of accuracy such as *overall accuracy*, *producer accuracy*, *user's accuracy*, and *Kappa statistic* can be derived. *Overall accuracy* is computed by dividing the total correct (sum of major diagonal elements) by the total number of pixels in an error matrix. *Producer accuracy* measures the errors of omission in the reference data set. It measures the probability of the reference data being correctly identified. The producer accuracy is measured by dividing the total number of correctly identified reference pixels in each category by the total number of pixels of that category. User's accuracy measures the probability that a pixel classified on the map actually represents that category on the ground. It is obtained by dividing the number of correctly classified pixels by the total number of pixels in that class. The kappa statistic

symbolized as K_{hat} statistic (Jensen, 1996; Mather, 1999; and Richards, 1999), derived from KAPPA analysis also provides a measure for accuracy assessment. KAPPA analysis is a discrete multivariate technique of use in accuracy assessment. The K_{hat} statistic is computed as:

$$K_{hat} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}$$

Where r is the number of rows in the error matrix, x_{ii} are the diagonal entries in an error matrix, x_{i+} and x_{+i} are the marginal totals for row i and column j respectively.

In this study, the accuracy for the maps derived from supervised classification was evaluated using overall accuracies and Kappa statistics. The error matrix was obtained by comparing the classified pixels and the reference pixels. The comparison of pixels (classified and reference) was done with the help of the inbuilt algorithm for accuracy assessment in Erdas Imagine. A total of 105 reference points (samples) obtained from field truthing using GPS instrument were used for the creation of an error matrix from which the overall accuracies and Kappa static's were derived.

The actual number of pixels to be used as reference pixels for the testing of a classified map is often difficult to determine (Jensen, 1996; Congalton, 1984). Sample sizes are often limited by logistic and budgetary constraints that are outside the control of investigators (Ayanz, 1996). However, some methods for determining the sample size for testing the accuracy of the classified map have been suggested by Congalton (1991), Van Genderen *et al.*, (1977) and Fitzpatrick-Lins (1981).

Congalton suggests that at least 50 samples for each Land cover should be collected. He also noted that for land cover features with little variability such as water and forests, fewer points might be used. Van Genderen (1977) stated that when sampling is done for overall accuracy, the smallest sample size in each class should be 20 or 30 for maps in

which the admissible percentage error is 15 and 10 percent respectively. This approach is supported by Fitzpatrick-Lins (1981).

"Regardless of the method used, it is impossible not to lie with the error matrix" (Verbyla, 1995; 2002). This is due to the assumptions made during accuracy assessment i.e.

- The assumption that the reference data is a true representation of the entire classification.
- The assumption that the reference data and the classified image are perfectly co-registered and that there is no error in the reference data.

Because of the difficulty in assessing the accuracy of each pixel, the actual accuracy of the classification of remotely sensed data will remain unknown. It is possible, therefore, to produce misleading assessment of the classification accuracy resulting in either a conservative estimate or optimistic estimate of the accuracy results (Verbyla, 2002).

As already mentioned, a sample size of 105 pixels was used for accuracy assessment. Following the above discussion, this sample size can not be fully representative of the whole image and hence the accuracies derived out of this sample should be taken as relative accuracies rather than absolute or actual accuracies. The word 'relative' is used because different approaches (i.e. band selection methods and classification algorithms) were used for the image classification process and hence different results were obtained. The inability to access certain areas in the forest due to altitude variability, the difficulty in obtaining GPS positions due to high canopy enclosure and the lack of any other reliable ground truth data (maps and aerial photos), were some of the stumbling blocks of getting relatively a large sample size for accuracy assessment.

The GPS points and the corresponding land cover codes are presented in appendix A. The following tables give the accuracies of the classification results obtained from different classification algorithms. Table 6.1 through to Table 6.4 present the results for accuracy assessment based on the use of separability indices for optimum band selection while Table 6.5 and Table 6.6 present the results based on correlation analysis for optimum

band selection. Table 6.7 presents a summary of the classification accuracies. The accuracies have been ranked from the highest to the lowest.

Table 6.1 Maximum likelihood classification results (SPOT)

Band combination	Overall Accuracy	Kappa Statistic	Classification algorithm	Method for band selection	Satellite image
1,3,6,7,8,9	55.24	0.3788	MLC	ED: Best minimum	SPOT
3,4,5,6,7,8	50.48	0.3056	MLC	ED: Best average	SPOT
1,2,3,4,5,6	55.24	0.3755	MLC	TD	SPOT

Table 6.2 Parallelepiped and Maximum likelihood classification results (SPOT)

Band combination	Overall Accuracy	Kappa Statistic	Classification algorithm	Method for band selection	Satellite image
1,3,6,7,8,9	52.38	0.3463	P &MLC	ED: Best minimum	SPOT
3,4,5,6,7,8	42.86	0.2178	P& MLC	ED: Best average	SPOT
1,2,3,4,5,6	48.57	0.301	P&MLC	TD	SPOT

Table 6.3 Maximum Likelihood Classification (MLC) results (Landsat ETM+)

Band combination	Overall Accuracy	Kappa Statistic	Classification algorithm	Method for band selection	Satellite image
4,5,13,14,15,16	72.38	0.6149	MLC	ED: Best minimum	Landsat
4,5,7,8,13,15	73.33	0.6268	MLC	ED: Best average	Landsat
1,2,3,4,5,6	70.48	0.591	MLC	TD	Landsat

Table 6.4 Parallelepiped and MLC results (Landsat ETM+)

Band combination	Overall Accuracy	Kappa Statistic	Classification algorithm	Method for band selection	Satellite image
4,5,13,14,15,16	71.43	0.6026	P &MLC	ED: Best minimum	Landsat
4,5,7,8,13,15	70.48	0.5933	P& MLC	ED: Best average	Landsat
1,2,3,4,5,6	69.52	0.5851	P&MLC	TD	Landsat

Table 6.5 MLC results for SPOT and Landsat (ETM+)-correlation approach

Band combination	Overall Accuracy	Kappa Statistic	Classification algorithm	Method for band selection	Satellite image
2,3,4,8,9,10	54.29	0.3706	MLC	Correlation	SPOT
4,6,7,14,15,18	73.33	0.6307	MLC	Correlation	Landsat

Table 6.6 MLC and parallelepiped results –correlation approach

Band combination	Overall Accuracy	Kappa Statistic	Classification algorithm	Method for band selection	Satellite image
2,3,4,8,9,10	53.33	0.3706	P &MLC	Correlation	SPOT
4,6,7,14,15,18	71.43	0.6029	P &MLC	Correlation	Landsat

Table 6.7 Summery of the classification accuracies and kappa statistics

Band combination	Classification algorithm	Method for band selection	Satellite image	Overall Accuracy	Kappa Statistic
4,5,7,8,13,15	MLC	ED: Best average	Landsat	**73.33**	**0.6268**
4,6,7,14,15,18	MLC	Correlation	Landsat	**73.33**	**0.6307**
4,5,13,14,15,16	MLC	ED: Best minimum	Landsat	**72.38**	**0.6149**
4,5,13,14,15,16	P &MLC	ED: Best minimum	Landsat	**71.43**	**0.6026**
4,6,7,14,15,18	P &MLC	Correlation	Landsat	**71.43**	**0.6029**
1,2,3,4,5,6	MLC	TD	Landsat	**70.48**	**0.591**
4,5,7,8,13,15	P& MLC	ED: Best average	Landsat	**70.48**	**0.5933**
1,2,3,4,5,6	P&MLC	TD	Landsat	**69.52**	**0.5851**
1,3,6,7,8,9	MLC	ED: Best minimum	SPOT	**55.24**	**0.3788**
1,2,3,4,5,6	MLC	TD	SPOT	**55.24**	**0.3755**
2,3,4,8,9,10	MLC	Correlation	SPOT	**54.29**	**0.3706**
2,3,4,8,9,10	P &MLC	Correlation	SPOT	**53.33**	**0.3706**
1,3,6,7,8,9	P&MLC	ED: Best minimum	SPOT	**52.38**	**0.3463**
3,4,5,6,7,8	MLC	ED: Best average	SPOT	**50.48**	**0.3056**
1,2,3,4,5,6	P&MLC	TD	SPOT	**48.57**	**0.301**
3,4,5,6,7,8	P&MLC	ED: Best average	SPOT	**42.86**	**0.2178**

The following should be noted from the Table 6.1 up to Table 6.7:

P= parallelepiped Classifier, MLC= Maximum likelihood Classifier, ED= Euclidian Distance, TD= Transformed divergence.

From the Table 6.7, it can be observed that the accuracies as well as the kappa statistics derived from the Landsat (ETM+) image are higher than the accuracies and the kappa statistics derived from the SPOT image. The highest accuracy and kappa statistic obtained from the analysis of the Landsat data are 73.33 and 0.6268 while the highest accuracy and kappa statistic are 55.24 and 0.3777 respectively.

7 DISCUSSION OF RESULTS

7.1 Introduction

This chapter discusses the results obtained from supervised and unsupervised classification of SPOT and Landsat data sets. The discussion of the results is based on the research questions. Generally, four major land cover types were identified within the forest from remotely sensed images. These include: Dense Evergreen Forest (DEGF), Evergreen Forest (EGF), brushland and grassland. Outside of the forest, three other land cover categories were identified i.e. grassland, subsistence farmlands, tea plantations and woodland.

However, one difficulty encountered in this research was ground truthing. The only main sources of data were: Old topographical maps printed in 1965, field visits as well as aerial photography acquired using a small format DCS camera in 2003. An attempt to obtain ground truth data with the GPS was generally difficult due to limitations of the terrain (hilly areas, steep valleys). The aerial photographs could have provided a better source for ground truth data but the photographs were not rectified. The identification of land cover classes were based on old maps, GPS data (not uniformly distributed through out the study area) and aerial photo interpretation (unrectified). This, therefore, affects the validation of the classification results.

In some areas visited during field work the following generally characterizes some of the land cover classes:

- DEGF were located in the valleys.
- EGF were located at medium and high altitudes.
- Brush land areas were mainly located at the higher southern altitude.

Figure 7-1 shows part of a three band composite of Landsat TM image with bands 4, 3, and 2 assigned to red, green and blue colour guns respectively. The raster contours generated from DEM of Bwindi are overlaid on the image. It can be observed that the dark areas in the forest (representing the DEGF) are associated with lower altitude (flat areas)

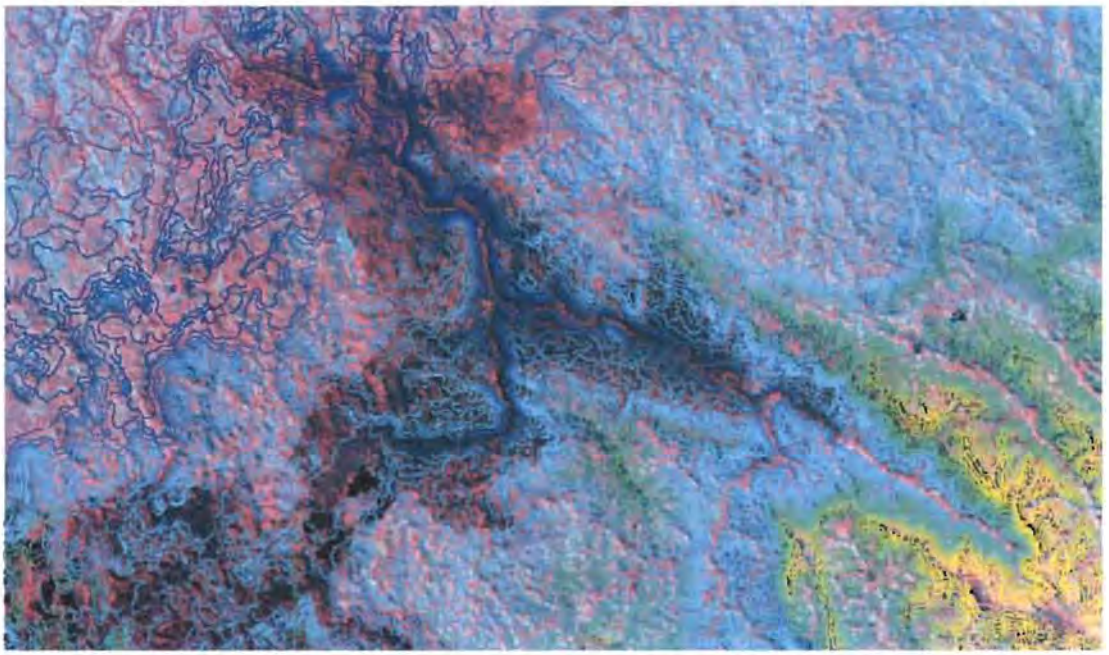


Figure 7-1 Raster contours overlaid on Landsat image (1)

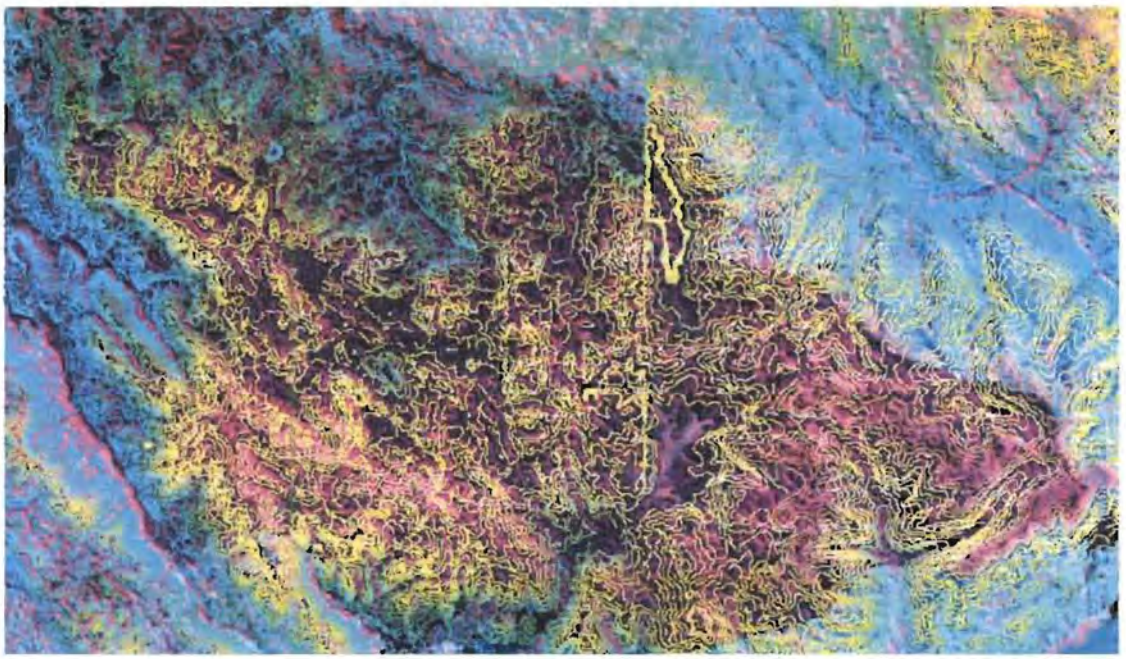


Figure 7-2 Raster Contours on Landsat image (2)

Figure 7-2 shows the southern part of Bwindi impenetrable forest. The brush land areas correspond with the higher altitudes. These areas appear as bright red from the image.

Figure 7-3 and Figure 7-4 examine the correlations between the classified data and altitude. It is observed in Figure 7-3 that a contour with value 4500m (shown with red colour) passes through both the DEGF as well as EGF

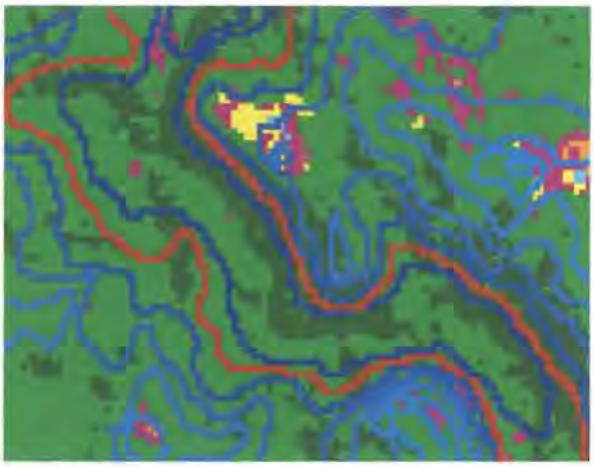


Figure 7-3 Raster contours overlaid on classified Landsat image (1)

In Figure 7-4 , a contour of value 7200m (shown with red colour) is overlaid on the classified Landsat image. The contour passes through DEGF, EGF as well as Brush land classes.

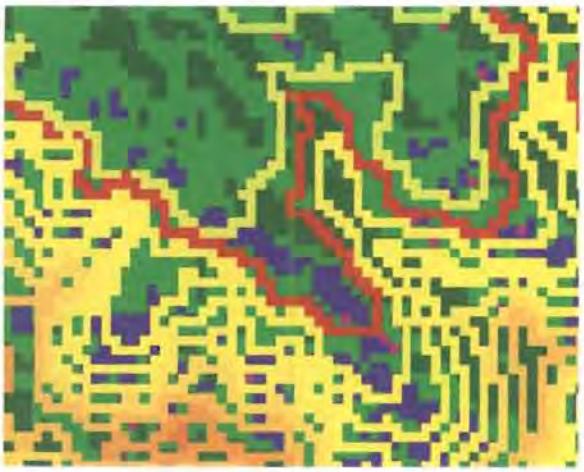


Figure 7-4 Raster contours overlaid on classified Landsat image (2)

With these observations, classification of the land cover classes based on altitude is not reliable and is inconclusive. The following sections discuss the results which provide answers for the research questions

7.2 Discussion of results for research question 1

"The information obtained from land use and land cover mapping of Bwindi impenetrable forest using SPOT and Landsat (ETM+) imagery"

7.2.1 Land cover types (information) from SPOT data

The identification of the land cover types using SPOT data was carried out in three phases:

Phase 1: Identification of land cover types using unsupervised classification

The results for the unsupervised classification of the SPOT data are shown in Figure 5-7 (chapter 5). In this approach, all the classes in the classification scheme except the woodland class were identified. The Woodland class was classified as DEGF. This is because unsupervised classification classes are dependent on the ability of the computer to identify the spectral groupings and less input is required from the analyst. As can be seen for the case of supervised classification of SPOT images, it was possible to identify the woodland class as well.

Phase 2: Supervised classification with optimum band selection using correlation analysis.

The results for this approach are shown in Figure 5-10(a) and (b) (chapter 5). Figure 5-10(a) presents the results for supervised classification using Maximum Likelihood Classifier (MLC) while Figure 5-10(b) presents the results for supervised classification using parallelepiped and MCL. The supervised classification results using MLC showed masked areas classified as Dense Evergreen forest (DEGF). This is because the DEGF class had large variance in this particular band combination and besides the masked areas

had no proper parameters (mean and standard deviation-all these had 0 values) by the MLC. The DEGF was nearest to the masked areas i.e. clouds (Figure 5-9). The classification with both parallelepiped and MLC' however, separated clouds (masked areas) from other classes. This is because the parallelepiped classifier works on the principle of the class (or DN values in the class) being either inside or outside the parallelepiped box and the unclassified pixels are then classified with the MLC.

Phase 3: Supervised classification with optimum band selection using vegetation indices

A similar approach to the one used in phase 2 above was adopted. Again all the results obtained with MLC did not identify the masked areas. The masked areas were classified as one of the other classes in the classification scheme (Figure 5-16) while a combination of parallelepiped and MLC identified the masked areas (Figure 5-17).

7.2.2 Land cover types from Landsat (ETM+) satellite image

Identification of land cover types with Landsat image followed a similar approach used for the SPOT image.

For unsupervised classification of the Landsat image all the required land cover classes were identified (Figure 5-11). For the supervised classification of the Landsat image, the trend was the same as the supervised classification of SPOT image i.e. when the MLC was used, the masked areas were classified as one of the other classes while when a combination of the parallelepiped and MLC were used, all the information classes and the clouds (masked areas) were identified.

7.3 Discussion of results based on research question 2

"The bands (original and derived) suitable for classification of the SPOT and Landsat for the Bwindi forest mapping"

7.3.1 Suitable bands for the mapping with SPOT

As mentioned in chapter five, six bands were required for the optimum set of bands for classification of either SPOT or Landsat data. The justification for six bands is presented in chapter 5.

The results for the determination of the optimum band combinations suitable for the classification of the SPOT image are presented in Table 5.18 (for correlation analysis), Table 5.26 and Table 5.27 (for separability index analysis). Table 6.7 compares the accuracies and the kappa statistics of the optimum band combinations. The results in Table 6.7 indicate that the classification accuracy does not only depend on the optimum band selection method but also on the classification algorithm used. However, the results suggest that optimum band selection based on separability indices or correlation analysis and classified with Maximum Likelihood Classification algorithm produces results of higher accuracies and kappa statistics. The best overall band combination (i.e. with highest accuracy and kappa statistic was obtained by ED separability index (best average). The second best band combination was derived from TD separability index. In these two specific band combinations it is important to note that all have original SPOT band 1 and 3 and the derived band 6 (ratio of band 3 and band 2 of the SPOT) images. This suggests the following:

- The combination of original and derived bands provide results of higher accuracy if the appropriate classification algorithm is used (in this case MLC)
- The combination of bands from different spectral regions of the electromagnetic spectrum (e.g. XS1= visible region, XS3 =infrared region) as well as the ratio between infrared/red bands provide results of better accuracies.

7.3.2 Suitable bands for mapping with Landsat

The results for optimum set of bands for Landsat data is presented in Table 5.25 (correlation analysis) as well as Tables 5.28 and 5.29 (for separability index analysis). Table 6.7 compares the accuracies and kappa statistics of the optimum band combinations. It is again evident that the accuracies and the kappa statistics depend not only on the method for band selection but also on the classification algorithm. Generally, the band selection by either separability indices or correlation analysis with maximum likelihood classification algorithm provides the results with higher accuracies and kappa statistics with exception of Transformed Divergence (TD) as band selection methods. This is because the TD band selection provided only the original bands that were highly correlated. When the accuracies obtained with TD as band selection method is compared with the accuracies obtained with either ED or correlation method, the TD gives lower accuracies. This is true whether the classification is performed using MLC or a combination of parallelepiped and MLC. This suggests that the presence of the derived bands was responsible for the higher accuracies obtained using ED or correlation as band selection methods. This is supported by the results shown by the optimum band selection using for SPOT image. The results had higher accuracies due to the presence of the derived bands.

7.4 Discussion of results for research question 3

"The method suitable for band selection"

7.4.1 SPOT image

As it has been discussed in the previous sections, band selection is dependent on the classification algorithm as well as the number of bands used. However, for this specific study, any of the band selection approaches (i.e. correlation and separability indices) can be used with MLC as the classification algorithm.

7.4.2 Landsat image

Table 6.7 suggests that either the correlation approach or separability indices (except TD) can be used for band selection with MLC as the classification algorithm.

7.5 Discussion of results for research question 4

"The accuracies obtained through the classification of the SPOT and Landsat images"

The summary of the results for accuracy assessment are presented in Table 6.7. Further details regarding accuracy assessment are presented in the appendix D. In this section, the highest accuracy derived from SPOT and Landsat are discussed. The analysis presented here applies to other results.

7.5.1 Highest accuracy from SPOT

The highest accuracy and the kappa statistic were 55.24% respectively. This accuracy was achieved with Euclidian distance as band selection method while MLC was used as the classification algorithm. From the error matrix, it is clear that the accuracy was affected by misclassification of the test data. The discussion of some entries in the error matrix is presented below.

- Out of 19 samples known to be DEGF, only 12 were classified as DEGF while 7 where classified as EGF. This is understandable as these classes overlap (i.e. it is not easy to delineate the boundary between DEGF and EGF especially at the areas where transition occurs. The resulting producer accuracy was therefore low (63.16%). The user's accuracy was also affected. Twelve sample pixels that were originally known to be EGF were classified as DEGF and the resulting producer accuracy was 50%.
- Out of 52 test samples for EGF, only 32 were identified as EGF. The rest were misclassified leading to a producer accuracy of 64%. This resulted in lower producer accuracy. Also 12 other test samples that were not known to belong to EGF were classified as EGF leading to a user's accuracy of 72.73%.

- The tea plantation had the worst producers and users accuracy. None of the 8 reference samples were classified as tea plantation while one test sample that was originally known as EGF was classified as tea-plantations. One reason that explains this could be that no tea plantation was growing at the test area when the image was acquired.

The misclassification of the pixels had an effect on the overall accuracy as well as the Kappa statistic.

7.5.2 Highest accuracy from Landsat image

The highest accuracy and kappa statistic were 73.33 and 0.6307. These results were obtained using the correlation method for band selection and MLC as classification algorithm. There was less misclassification of the pixels in validating the accuracy of the classified image (see figure 40) for related information. The lowest producer's accuracy was 46.15% resulting from Subsistence farmlands. Out of 13 test samples, only 6 were identified correctly while 7 were misclassified. It is, however, important to note that the tea plantation class was the highest. The lowest user's accuracy was 53.85% obtained from the tea plantation. This is because 6 other test samples that were not originally known as tea were in fact classified as tea.

7.5.3 Comparison of Landsat and SPOT accuracies

The comparison between the results of SPOT and Landsat analysis are presented in Table 6.7. The results from SPOT image generally had lower accuracy and kappa statistics than the results from Landsat. One of the reasons for this discrepancy could be due to the fact that the test data was collected in 2003, ten years after the SPOT image was acquired unlike Landsat TM that was relatively recent (acquired in 2000). Changes in land cover are always taking place especially in the areas used as subsistence farmlands because land use and land cover types are not properly defined.

Secondly the difference in accuracy could be due to the fact that the SPOT image was geo-referenced using Landsat image. Depending on the final fit of the resulting image, this would have a bearing on then classification accuracies. On contrary, both Landsat and GPS points acquired for testing were based on one reference system and hence less distortion are expected to arise due to rectification.

7.6 Geographical Information System (GIS) data base

GIS is a tool for managing information of any kind according to where it is located. Simply put, it combines layers of information about a place to give a better understanding of the place.

A GIS prototype was established by integrating the land cover map obtained with other data layers namely: rivers, roads, administrative boundaries, townships as well as landscape information. The ancillary data were originally in Clark 1880 modified ellipsoid. These layers were re-projected to WGS84 coordinate system before overlaying with the vegetation map. The results are shown in Figure 7-5

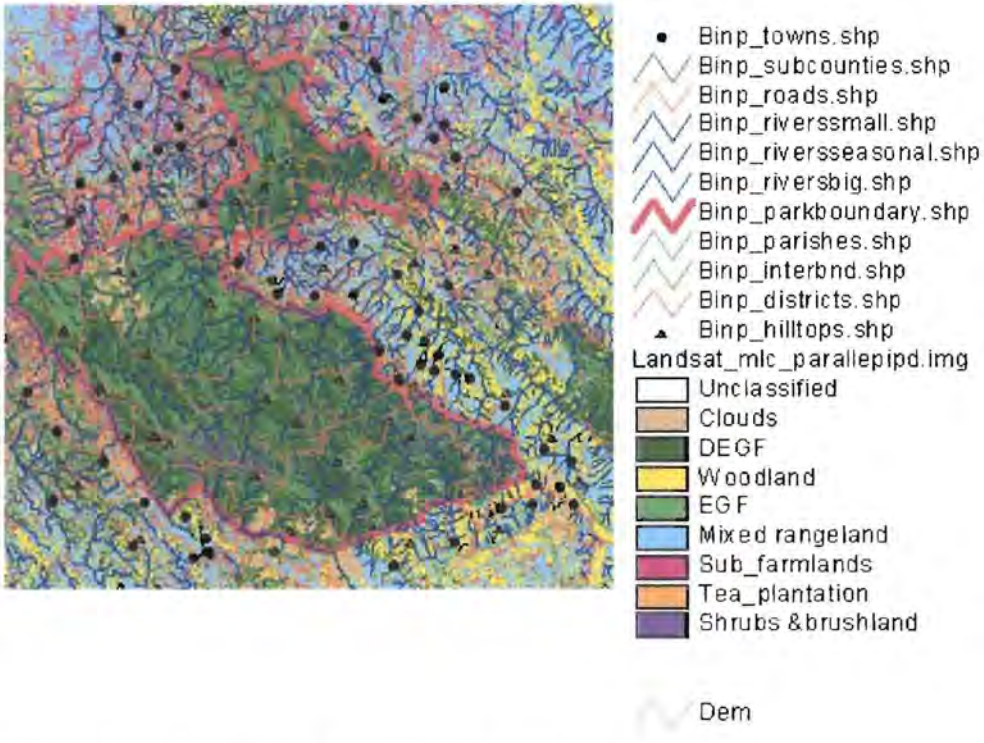


Figure 7-5 Prototype GIS data base for Bwindi National Park

The GIS database shown in Figure 7-5 can be used for forest management and/or managing activities in the forest. These may include location of existing trails and/or establishing new ones, identification of water sources, identification of areas for conservation and identification of areas for gorilla tracking.

8 CONCLUSION AND RECOMMENDATIONS

8.1 Introduction

This study addressed the application of Landsat TM and SPOT satellite imagery for land cover mapping of Bwindi Impenetrable Forest. The study focused on the suitability of SPOT and Landsat TM in identifying different land cover features. The study also identified the suitable bands for classification, the method of band selection and the classification algorithm. Of specific interest was an investigation of the potential of band selection by secondary correlation. This chapter presents the conclusions based on the findings of the study. The conclusions are considered to be specific to the data sets tested used and the study area investigated.

8.2 Conclusions

Research question 1: *Information obtained from land use and land cover mapping of Bwindi Impenetrable National Park*

- Seven land cover categories were identified i.e. woodland, grassland, tea plantation, subsistence farmland, dense evergreen, brushland and evergreen forests. The study has therefore demonstrated the potential of using SPOT and Landsat TM satellite images for land cover mapping.
- The delineation of built up areas (roads and trading centers) was not possible using either Landsat or SPOT images. Built-up areas covered small areas and were in most cases located amidst other vegetation types for example banana plantations. This resulted in spectral mixing making it difficult to separate built-up areas with the adjacent vegetation types. The built-up-areas were classified as mixed rangeland

Research questions 2&3: *Bands suitable for classification/ method for band selection*

- Two methods of band selection were used i.e. the correlation approach and separability index approach.
- The correlation approach provided one set of bands for the SPOT and one set of bands for Landsat (ETM+).
- The separability index approach provided three sets of bands for the SPOT and another three sets of bands for Landsat (ETM+) basing on either Euclidian Distance (ED) or Transformed Divergence (TD). The best minimum and average separability were obtained in both cases.
- For SPOT image, the band combination provided by ED (best minimum separability) or the TD were suitable for classification (refer to Table 6.7) with Maximum likelihood Classifier.
- The bands provided by correlation approach and ED (best average separability) were suitable for classification with MLC since they gave the highest accuracy and Kappa statistic.

Research question 4: *Accuracies of the classified maps obtained through the classification of SPOT and Landsat images.*

The accuracies obtained through the classification of the SPOT and Landsat images are summarized in Table 6.7. The accuracies vary depending on the method of band selection as well as the classification algorithm used.

8.3 Recommendations

8.3.1 Detailed classification

Detailed mapping of the study area is unrealistic using SPOT and Landsat. Further research is recommended on the use of high spatial resolution imagery, for example, IKONOS imagery and Radar. Image fusion of high resolution SPOT imagery with high spectral Landsat imagery is also recommended.

8.3.2 Land cover types

Land cover types in the forest were determined through visual analysis. The determination of land cover types on the basis of leaf area index, chlorophyll content, diameter at breast height, tree heights and or tree canopy other than visual classification and or spectral classification should be explored

8.3.3 Reference data

The reference data should be obtained as soon as the satellite image is acquired. This would minimize the errors of omission and commission obtained in the classified images which tend to lower the producer's and user's accuracies. It is also recommended that the sample size for the test data should be increased but as has been mentioned earlier, it is not possible to get absolute accuracy. A method of stratified random sampling is recommend for obtaining the test data since it allows the determination of representative samples for each class.

9 REFERENCES AND BIBLIOGRAPHY

Achard F., Hugh D. Eva, Hans-Jürgen Stibig, Philippe Mayaux, Javier Gallego, Timothy Richards and Jean-Pual Malingreau, 2002. Determination of the deforestation rates of the Worlds Humid Tropical Forests. *Science Magazine*, 297, 999-1002.

Apan, A.A., 1997. Land cover mapping for tropical forest rehabilitation planning using remotely sensed data. *International Journal of remote sensing*, 18, 1029-1049.

Armenakis, C., Leduc,.F, Cyr, I, Savopol, F. and Cavayas, F., 2003. Acomparative analysis of scanned maps and imagery for mapping applications. *ISPRS Journal of Photogrammetry & Remote Sensing*, 57, 304-314.

Baret, F. and Guyot, G. 1991. 'Potential and limits of vegetation indices for LAI and APAR assessment'. *Remote sensing of environment* 35, 53-70.

Barrett, E.C. and L.F. Curtis. (1992) *Introduction to Environmental Remote Sensing*, 3rd ed. Chapman & Hall: London.

Butynski, T.M., 1984. Ecological survey of the Impenetrable (Bwindi) forest, Uganda and recommendations for its conservation and management. Report. New York Zoological Society, New York.

Clevers, J. G. P. W. (1988) 'The derivation of a simplified reflectance model for the estimation of leaf area index', *Remote Sensing of the Environment* 34, 71-3.

Colby, J. D., 1991. Topographic normalization in rugged terrain. *Photogrametric Engineering and Remote Rensing*, 57, 531-537.

Crippen, R.E. (1990) Calculating the Vegetation Index Faster, *Remote Sensing of Environment* 34, 71-73.

Dean Christopher, Timothy A. Warner and James B. McGraw, 2000.Suitabilty of the DCS460c colour digital camera for quantitative remote sensing analysis of vegetation. *ISPRS Journal of Photogrammetry & Remote Sensing, 55, 105-118.*

Erdas, 1999. *Erdas Field Guide*, Version 8.4. Fifth edition. Erdas, Inc. Atlanta Georgia.

Fitzpatrick-Lins, K., 1981, 'The comparison of Sampling Procedures and Data Analysis for a Land-cover Map,' *Photogrammetric Engineering and Remote Sensing*, 47(4), 343-351.

Forest Department, 1991. National Biomass study. Ministry of Natural resources, Kampala, Uganda.

Forest mapping and monitoring. http://www.gvm.sai.jrc.it/forest/contentsforest.htm : 26th December, 2003]

Frankovich, J.S., 1999.Unsupervised classification of spectrally enhanced Landsat TM data of Midland, Michigan. Final project, University of Arkansas.

Gibson, Paul J. and Power, Clare H. 2000. *Introduction to Remote Sensing*: Digital Image Processing and applications. Routledge, 29 West 35th, New York.

Giri, C. and Shrestha, S., 1996. Land cover mapping and monitoring from NOAA AVHRR data in Bangladesh. *International Journal of remote sensing*, 17, 2749-2759.

Goldsmith, F. B. (1998). *Tropical rain forest*. London: Chapman $Hall.

Gonzalez, R. C. and Wintz, P. 1977. *Digital Image Processing*. Addison-Wesley Publishing Company, Inc., Massachusetts 01867, USA.

Hamilton, A. C., 1984. *Deforestation in Uganda.* Oxford University Press, Nairobi, Kenya.

Howard, P.C., 1991. Nature conservation in Uganda's tropical forest reserves. World Conservation Union, Gland, Switzerland.

Huete, A. R. (1988) 'A soil –Adjusted Vegetation index (SAVI)', *Remote Sensing of the Environment* 25, 295-309.

Jensen, John R., 1996. *Introductory digital image processing: a remote sensing Perspective.* Second edition. Prentice- Hall, New Jersey.

Jordan, C. F. (1969) 'Derivation of leaf area index from quality of light on the forest floor', *Ecology* 50, 663-6.

Kauth, R. J. and Thomas, G.S. (1976) The Tasseled Cap, a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat, in proceedings of the Symposium on Machine Processing of Remotely sensed data. West Lafayette, Purdue University.

Mason, S., Ruther, H., Smit, J., 1997. Investigation of the KODAK DCS460 digital camera for small-area mapping. *ISPRS Journal of Photogrammetry & Remote Sensing,* 52, 202-214.

Mausel, P., Wu, Y., Li, Y., Moran, E. F., and Brondizio, E.S., 1993. Spectral identification of successive stages following deforestation in Amazon. *Geocarto International,* 4, 61-71.

Mills, J.P., Newton, I., 1996. Anew approach to verification and revision of large-scale mapping. *ISPRS Journal of Photogrammetry & Remote Sensing,* 51(1), 17-27.

Mohammed M.R., 1997. Identification of Land use and land cover using band ratioing. Department of Geoscience. Murray State University.

Nezryt Edmund, 1998.Using SPOT and Radar data to inventory forestry in Sarawak, Malaysia. Report. Privateer N.V., Netherlands.

Peddle, R.D. and Franklin, S. E, 1991. Image Texture processing and data integration for surface pattern Discrimination. *Photogrammetric Engineering and Remote Sensing,* 57 (4), 413-420

Pinty, B. and Verstraete, M.M. 1991. 'Extracting information on surface properties from bi-directional reflectance measurements'. *Journal of Geophysical Research,* 96, 2865-74

Qi, J., Chenbouni, A., Huete, A.R. and Kerr, Y. H. (1994) 'Modified Soil Adjusted Vegetation index (MSAVI)', *Remote Sensing of the Environment* 101: 15-20.

Richards, J.A. and Xiuping, J. 1999. *Remote Sensing Digital Image Analysis: An Introduction.* 3rd ed., Springer, Berlin.

Richardson, A.J. and Wiegand, C.L. 1977. 'Distinguishing vegetation from soil background'. *Photogrammetric Engineering and Remote Sensing,* 43, 1541-52.

Sabins, Floyd F., 1997. Remote sensing: *Principles and interpretation.* Third edition. W.H.Freeman and Company, New York.

Sader, S.A., Stone, T. A., and Joyce, A.T., 1990. Remote sensing of tropical forest. An overview of research and applications using non-photographic sensors. *Photogrammetric Engineering and Remote Sensing,* 56, 1343-1351.

San Miguel-Ayanz and Biging, G.S., 1997. Comparison of single stage and multi stage classification approaches for cover type mapping with TM and SPOT data. *International Journal of Remote sensing and Environment,* 59, 92-104.

Stenback, J.M., and Congalton, R.G., 1990. Using Thematic Mapper imagery to examine forest under story. *Photogrammetric Engineering and Remote sensing,* 56, 1285-1290.

Thomas M.Lillesand and Ralph W. Kiefer, 1999. *Remote Sensing and Image Interpretation*. Fourth edition. John Wiley and Sons, New York.

Trisurat Y., Eiumnoh A., Murai S., Hussain M.Z., and Shrestha R.P., 2000. Improvement of tropical forest mapping using a remote sensing technique: a case study of Khao Yai National Park, Thailand. *International Journal of remote sensing,* 21, 2031-2042.

Tucker C. J. (1979) 'Red and photographic infrared linear combinations for monitoring vegetation', *Remote Sensing of the Environment* 8, 127-50.

Van Genderen, J.L. and B.F. Lock. (1977). Testing land-use map accuracy. *Photogrammetric Engineering & Remote Sensing* 43(9), 35-1137.

Verbyla, D. L. and T. O. Hammond. 1995. Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing*. 16, 581-587.

Wolter, P.T., Mladenoff, D.J., Host, G. R., and Crow, T .R., 1995. Improved forest classification in the northern lake states using multi-temporal Landsat imagery. *Photogrammetric Engineering and Remote sensing,* 61, 1129-1143.

Xiao,X., Boles, S., Liu, J., Zhunag, D., and Liu, M., 2002. Characterisation of forest types in Northern China, using multi-temporal SPOT-4 VEGETATION sensors data. *International Journal of Remote sensing and Environment,* 82, 335-348.

Yousfi, D., (2000). A methodological approach for agricultural statistics by remote, 23-29, In: Casanova, J. L., (Ed), *Remote Sensing in the 21st Century: Economic and Environmental applications*, University of Valladolid, Spain.

http://www.ntsg.umt.edu/Courses/dip/lectures/vi/sld002.htm, accessed on 9th/8/2004.

http://www.worldbank.org/afr/findings/infobeng/infob73.pdf [Accessed: 3r January 2004).

APPENDIX A: TEST DATA

ID	Easting (m)	Northing	LULC_code	Comments
1	806127	9882524	2	DEGF
2	805880	9881974	2	DEGF
3	805843	9882060	2	DEGF
4	805888	9882229	2	DEGF
5	805952	9882357	2	DEGF
6	806040	9882471	2	DEGF
7	791312	9889661	2	DEGF
8	790475	9890370	2	DEGF
9	790420	9890400	2	DEGF
10	791210	9888835	2	DEGF
11	791100	9888835	2	DEGF
12	808690	9884063	2	DEGF
13	808499	9884134	2	DEGF
14	808524	9884465	2	DEGF
15	808260	9884540	2	DEGF
16	807914	9884600	2	DEGF
17	807762	9884713	2	DEGF
18	806710	9887647	2	DEGF
19	799320	9891780	2	DEGF
20	807568	9883664	4	EGF
21	805840	9882126	4	EGF
22	805877	9882163	4	EGF
23	805902	9882268	4	EGF
24	806445	9882841	4	EGF
25	806690	9883173	4	EGF
26	806778	9883325	4	EGF
27	807362	9883578	4	EGF
28	793758	9904991	4	EGF
29	793773	9904886	4	EGF
30	790131	9885541	4	EGF
31	790166	9885511	4	EGF
32	790105	9885784	4	EGF
33	790193	9885454	4	EGF
34	791471	9885046	4	EGF
35	791472	9886993	4	EGF
36	791477	9886856	4	EGF
37	791398	9887113	4	EGF
38	791546	9887450	4	EGF
39	791288	9888666	4	EGF
40	791193	9888954	4	EGF
41	791155	9889507	4	EGF
42	791007	9889740	4	EGF

43	791410	9889605	4	EGF
44	791429	9889678	4	EGF
45	791510	9889660	4	EGF
46	791675	9889685	4	EGF
47	791670	9889790	4	EGF
48	791621	9889805	4	EGF
49	790535	9890360	4	EGF
50	791570	9889810	4	EGF
51	791475	9889813	4	EGF
52	791373	9889886	4	EGF
53	791270	9889990	4	EGF
54	790993	9890047	4	EGF
55	791130	9889950	4	EGF
56	791136	9889890	4	EGF
57	791020	9889650	4	EGF
58	791995	9889785	4	EGF
59	791000	9889737	4	EGF
60	791020	9889640	4	EGF
61	791074	9889598	4	EGF
62	791120	9889495	4	EGF
63	791445	9889388	4	EGF
64	791180	9889240	4	EGF
65	791210	9889050	4	EGF
66	791218	9888958	4	EGF
67	808500	9884250	4	EGF
68	807560	9885110	4	EGF
69	799665	9890910	4	EGF
70	794035	9904979	5	Grassland
71	793977	9904882	5	Grassland
72	793940	9904814	5	Grassland
73	793883	9904952	5	Grassland
74	793857	9904699	5	Grassland
75	791190	9890745	5	Grassland
76	789677	9885869	5	Grassland
77	789713	9885771	5	Grassland
78	791688	9884274	5	Grassland
79	793448	9901940	5	Grassland
80	799349	9902396	5	Grassland
81	800254	9902416	5	Grassland
82	799406	9902823	5	Grassland
83	793380	9896975	5	Grassland
84	801181	9890043	5	Grassland
85	789588	9885893	6	SF
86	790350	9885315	6	SF
87	790460	9884781	6	SF

88	790409	9884521	6	SF
89	790675	9884492	6	SF
90	791205	9884467	6	SF
91	791283	9884408	6	SF
92	791187	9894769	6	SF
93	791177	9894797	6	SF
94	799946	9902389	6	SF
95	793950	9897260	6	SF
96	804706	9889055	6	SF
97	798840	9892850	6	SF
98	793739	9904819	7	Tea Plantation
99	793696	9904819	7	Tea Plantation
100	793791	9904758	7	Tea Plantation
101	793831	9904734	7	Tea Plantation
102	793916	9904644	7	Tea Plantation
103	793999	9904593	7	Tea Plantation
104	794040	9904661	7	Tea Plantation
105	798140	9895178	7	Tea Plantation

APPENDIX B: DESCRIPTION OF LAND COVER TYPES

ID	Easting (m)	Northing (m)	Comments
1	807568	9883664	Forest Valley
2	806127	9882524	Forest Hill
3	805843	9881871	Swamp
4	805880	9881974	Forest Hill
5	805843	9882060	Forest Hill
6	805840	9882126	Forest Hill
7	805877	9882163	Forest Hill
8	805588	9882229	Forest Hill
9	805902	9882268	Forest Hill
10	805952	9882357	Forest Hill
11	806040	9882471	Forest Hill
12	806445	9882841	Forest Hill
13	806690	9883173	Flat Hill
14	806778	9883325	Flat Hill
15	807362	9883578	Flat Hill
16	799197	9891598	Bridge
17	794035	9904979	Road/Airfield
18	793977	9904882	Airfield/Grassland
19	793940	9904814	Airfield/Grassland
20	793883	9904952	Grassland
21	793758	9904991	Eucalyptus Trees
22	793773	9904886	Eucalyptus Trees
23	793739	9904819	Tea Plantation
24	793696	9904819	Tea Plantation
25	793791	9904758	Tea Plantation
26	793831	9904734	Tea Plantation
27	793857	9904699	Airfield/Grassland
28	793916	9904644	Tea Plantation
29	793999	9904593	Tea Plantation
30	794040	9904661	Tea Plantation
31	791190	9890745	Camp (Head Office)
32	789408	9885946	Bridge
33	789459	9886007	Bridge
34	789375	9886335	Bridge
35	789436	9886012	Bridge
36	789377	9886313	Bridge
37	789588	9885893	Grazing Land
38	789677	9885869	Grassland
39	789713	9885771	Grassland
40	790131	9885541	Forest Valley
41	790166	9885511	Forest Valley

42	790105	9885784	Forest Hill
43	790193	9885454	Forest Degraded
44	790350	9885315	Forest Degraded
45	790460	9884781	Banana Plantation
46	790409	9884521	Banana Plantation
47	790675	9884492	Farm land- Mixed
48	791205	9884467	Farm land- Banana
49	791283	9884408	Home- Banana
50	791688	9884274	Grassland
51	791471	9885046	Hill Slope
52	791472	9886993	Hill Slope
53	791477	9886856	Hill Slope
54	791398	9887113	Forest slope
55	791546	9887450	Slope of hill-Grass
56	791288	9888666	Flat area-forest
57	791193	9888954	Forest- Low altitude
58	791155	9889507	Forest- Low altitude
59	790985	9890325	Grass-Plants Road
60	790983	9890456	Road Junction
61	791039	9890605	Road near APSL
62	791155	9890734	Road-Near Camp
63	791161	9890957	Grass land tents
64	791160	9890913	Grassland
65	791184	9890872	Grassland
66	791187	9894769	Corner of photography
67	791177	9894797	Corner of photography
68	791523	9895243	Road Junction
69	798052	9895251	Road
70	798140	9895178	Tea Plantation
71	799561	9891629	Bridge
72	797696	9894456	Road
73	793448	9901940	Trading Center
74	793950	9907158	Bridge2
75	799349	9902396	Grassland ,Farm land
76	799946	9902389	Grassland ,Pineapple
77	800254	9902416	Grassland
78	799406	9902823	Grassland_ Farm
79	818735	9887561	Forest Plantation
80	820890	9884164	Forest Plantation
81	790987	9890245	path to camp
82	790790	9890302	path to camp
83	790969	9890438	path to camp
84	791060	9890533	path to camp
85	791035	9890605	path to camp
86	791085	9890665	path to camp

87	791155	9890725	path to camp
88	791195	9890810	path to camp
89	791215	9890830	path to camp
90	791220	9890845	path to camp
91	791190	9890882	End of forest
92	791200	9890975	Beginning of the village
93	791185	9891164	End of the village
94	791178	9891219	path to camp
95	791150	9891219	Tea
96	791280	9891400	path to camp
97	791320	9891400	Tea
98	791310	9891470	path to camp
99	791500	9891520	Edge of forest
100	808690	9884063	medium height vegetation/dense short trees
101	808499	9884134	tall vegetation on right
102	808500	9884250	tall trees right and short left
103	808524	9884465	medium height/not dense
104	808260	9884540	few trees on left/open right
105	807914	9884600	medium height/dense/steep
106	807762	9884713	medium height/steep
107	807560	9885110	high, small trees not dense steep
108	807840	9885650	open grassland/left and right
109	807675	9885810	open right/medium height left/valley/thick
110	807170	9886600	dense/very steep left and right with medium trees
111	807080	9887260	Farmland right/ dense steep left
112	806710	9887647	low dense vegetation left
113	805210	9888790	low lying dense bush
114	799425	9890700	surrounded by fairly dense forest/tall trees
115	799665	9890910	fairly open/eucalyptus trees/valley/dense undergrowth
116	799605	9891570	River left/open
117	799320	9891780	Dense high forest
118	799100	9892130	open, dense undergrowth (broadleaved plants)
119	798840	9892850	End of forest neck/subsistence farmland
120	798134	9893613	Grassland/steep slope
121	793968	9897262	bananas/road
122	790370	9897450	bananas/road

APPENDIX D: OPTIMUM BAND SELECTION RESULTS

Taken 6 at a time

Class

1 CLouds
2 DEGF
3 Woodland/bushland
4 EGF
5 Mixed_rangeland
6 Subsistence Farmalnds
7 TP
8 Shrubs/brushland

Best Minimum Separability

Bands	AVE	MIN	Class Pairs: 1: 2	1: 3	1: 4	1: 5	1: 6	1: 7	1: 8
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8
1 3 6 7	167	37	221	237	331	287	357	394	431
8 9			37	115	96	142	179	218	110
			61	127	166	214	104	43	69
			104	98	135	193	39	96	63

Best Average Separability

Bands	AVE	MIN	Class Pairs: 1: 2	1: 3	1: 4	1: 5	1: 6	1: 7	1: 8
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8
3 4 5 6	186	34	268	279	384	328	404	440	495
7 8			34	126	92	150	188	242	123
			61	138	179	239	110	41	68
			117	107	147	214	42	108	73

Signature Separability Listing

File: c:/spot_2004/last_super.sig

Distance measure: Transformed Divergence

Using bands: 1 2 3 4 5 6 7 8 9 10

Taken 6 at a time

	Class
2	DEGF
3	Woodland/bushland
4	EGF
5	Mixed_rangeland
6	Subsistence Farmalnds
7	TP
8	Shrubs/brushland

Best Minimum Separability

Bands	AVE	MIN	Class Pairs:						
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8
1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000
5 6			2000	2000	2000	2000	2000	2000	2000
			2000	2000	2000	2000	2000	2000	2000

Best Average Separability

Bands	AVE	MIN	Class Pairs:						
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8
1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000
5 6			2000	2000	2000	2000	2000	2000	2000
			2000	2000	2000	2000	2000	2000	2000

ERROR MATRIX _3 (Transformed Divergence *landsat*)

REFERENCE DATA

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**16**	0	12	0	0	0	0	**28**
Woodland	0	0	0	**0**	0	1	1	0	0	**2**
EGF	0	0	3	0	**37**	1	2	0	0	**43**
Mixed rangeland	0	0	0	0	0	**8**	0	1	0	**9**
Sub_farmland	0	0	0	0	0	5	**7**	1	0	**13**
Tea_plantation	0	0	0	0	1	0	3	**6**	0	**10**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	28	16	84.21%	57.14%	0.4767
Woodland	0	2	0	---	---	0
EGF	50	43	37	74.00%	86.05%	0.7336
Mixed rangeland	15	9	8	53.33%	88.89%	0.8704
Sub_farmlands	13	13	7	53.85%	53.85%	0.4732
Tea_plantation	8	10	6	75.00%	60.00%	0.567
Shrubs &brushla	0	0	0	---	---	0
Totals	**105**	**105**	**74**			

Overall Classification Accuracy = 70.48
Overall Kappa Statistics = 0.5910

Signature Separability Listing

File: c:/bwindi_cloudfreelandsat/supervised.sig

Distance measure: Euclidean Distance

Using bands: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Taken 6 at a time

Class	
1	Clouds
2	DEGF
3	Woodland
4	EGF
5	Mixed rangeland
6	Sub_farmlands
7	Tea_plantation
8	Shrubs &brushland

Best Minimum Separability

Bands	AVE	MIN	Class Pairs: 1: 2	1: 3	1: 4	1: 5	1: 6	1: 7	1: 8
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8
4 5 13 14 15 16	168	28	229	283	313	346	341	365	394
			63	110	129	131	159	204	113
			66	124	150	201	141	28	53
			95	138	158	209	28	78	52

Best Average Separability

Bands	AVE	MIN	Class Pairs: 1: 2	1: 3	1: 4	1: 5	1: 6	1: 7	1: 8
			2: 3	2: 4	2: 5	2: 6	2: 7	2: 8	3: 4
			3: 5	3: 6	3: 7	3: 8	4: 5	4: 6	4: 7
			4: 8	5: 6	5: 7	5: 8	6: 7	6: 8	7: 8

Signature Separability Listing

File: c:/bwindi_cloudfreelandsat/supervised.sig

Distance measure: Transformed Divergence

Using bands: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Taken 6 at a time

	Class
2	DEGF
3	Woodland
4	EGF
5	Mixed rangeland
8	Shrubs &brushland

Best Minimum Separability

Bands	AVE	MIN	Class Pairs: 2: 3	2: 4	2: 5	2: 8	3: 4	3: 5	3: 8	4: 5	4: 8	5: 8
1 2 3 4 5 6	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

Best Average Separability

Bands	AVE	MIN	Class Pairs: 2: 3	2: 4	2: 5	2: 8	3: 4	3: 5	3: 8	4: 5	4: 8	5: 8
1 2 3 4 5 6	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

APPENDIX D: ACCURACY ASSESSEMENT RESULTS

ERROR MATRIX _1 (*Euclidian Distance (landsat) : Best average*)

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**16**	0	10	0	1	0	0	**27**
Woodland	0	0	0	**0**	0	0	0	0	0	**0**
EGF	0	0	3	0	**39**	2	1	0	0	**45**
Mixed rangeland	0	0	0	0	0	**9**	0	1	0	**10**
Sub_farmland	0	0	0	0	0	4	**6**	0	0	**10**
Tea_plantation	0	0	0	0	1	0	5	**7**	0	**13**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	27	16	84.21%	59.26%	0.5026
Woodland	0	0	0	---	---	0
EGF	50	45	39	78.00%	86.67%	0.7455
Mixed rangeland	15	10	9	60.00%	90.00%	0.8833
Sub_farmlands	13	10	6	46.15%	60.00%	0.5435
Tea_plantation	8	13	7	87.50%	53.85%	0.5004
Shrubs &brushla	0	0	0	---	---	0
Totals	**105**	**105**	**77**			

Overall Classification Accuracy = 73.33%
Overall Kappa Statistics = 0.6268

ERROR MATRIX _2 (*Euclidian Distance (landsat) : Best minimum separability*)

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	0	0	0	0	0	0	0	0	0	**0**
Clouds	0	0	0	0	0	0	0	0	0	**0**
DEGF	0	0	16	0	10	0	0	0	0	**26**
Woodland	0	0	0	0	1	2	0	0	0	**3**
EGF	0	0	3	0	38	1	3	0	0	**45**
Mixed rangeland	0	0	0	0	0	9	0	1	0	**10**
Sub_farmland	0	0	0	0	0	3	7	1	0	**11**
Tea_plantation	0	0	0	0	1	0	3	6	0	**10**
Shrubs &brushland	0	0	0	0	0	0	0	0	0	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	26	16	84.21%	61.54%	0.5304
Woodland	0	3	0	---	---	0
EGF	50	45	38	76.00%	84.44%	0.703
Mixed rangeland	15	10	9	60.00%	90.00%	0.8833
Sub_farmlands	13	11	7	53.85%	63.64%	0.585
Tea_plantation	8	10	6	75.00%	60.00%	0.567
Shrubs &brushla	0	0	0	---	---	0
Totals	**105**	**105**	**76**			

Overall Classification Accuracy = 72.38
Overall Kappa Statistics = 0.6149

ERROR MATRIX _1 (Euclidian distance: Best average (SPOT))

REFERENCE DATA

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	1	0	0	0	0	**1**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**11**	0	17	0	0	0	0	**28**
Woodland	0	0	0	**0**	0	0	0	0	0	**0**
EGF	0	0	7	0	**29**	**4**	3	2	0	**45**
Mixed rangeland	0	0	1	0	0	**7**	0	1	0	**9**
Sub_farmland	0	0	0	0	2	**4**	**6**	5	0	**17**
Tea_plantation	0	0	0	0	1	0	2	**0**	0	**3**
Shrubs &brushland	0	0	0	0	0	0	2	0	**0**	**2**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	1	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	28	11	57.89%	39.29%	0.2587
Woodland	0	0	0	---	---	0
EGF	50	45	29	58.00%	64.44%	0.3212
Mixed rangeland	15	9	7	46.67%	77.78%	0.7407
Sub_farmlands	13	17	6	46.15%	35.29%	0.2615
Tea_plantation	8	3	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	2	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy = 50.48
Overall Kappa Statistics =0.3056

ERROR MATRIX _2 (*Best minimum: SPOT*)

REFERENCE DATA

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	2	0	0	0	0	**2**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**12**	0	12	0	0	0	0	**24**
Woodland	0	0	0	**0**	1	0	1	0	0	**2**
EGF	0	0	7	0	**32**	3	2	0	0	**44**
Mixed rangeland	0	0	0	0	0	**7**	0	1	0	**8**
Sub_farmland	0	0	0	0	2	5	**7**	7	0	**21**
Tea_plantation	0	0	0	0	1	0	3	**0**	0	**4**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	2	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	24	12	63.16%	50.00%	0.3895
Woodland	0	2	0	---	---	0
EGF	50	44	32	64.00%	72.73%	0.4793
Mixed rangeland	15	8	7	46.67%	87.50%	0.8542
Sub_farmlands	13	21	7	53.85%	33.33%	0.2391
Tea_plantation	8	4	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy = 55.24%
Overall Kappa Statistics = 0.3788

ERROR MATRIX _3 (*Transformed divergence: SPOT*)

REFERENCE DATA

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	1	0	0	0	0	1
Clouds	0	**0**	0	0	0	0	0	0	0	0
DEGF	0	0	**11**	0	10	0	0	0	0	21
Woodland	0	0	0	**0**	2	0	1	0	0	3
EGF	0	0	7	0	**33**	3	2	0	0	45
Mixed rangeland	0	0	1	0	1	**7**	0	1	0	10
Sub_farmland	0	0	0	0	2	5	**7**	7	0	21
Tea_plantation	0	0	0	0	1	0	3	**0**	0	4
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	0
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	1	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	21	11	57.89%	52.38%	0.4186
Woodland	0	3	0	---	---	0
EGF	50	45	33	66.00%	73.33%	0.4909
Mixed rangeland	15	10	7	46.67%	70.00%	0.65
Sub_farmlands	13	21	7	53.85%	33.33%	0.2391
Tea_plantation	8	4	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy = 55.24%
Overall Kappa Statistics = 0.3755

ERROR MATRIX _11 (*Euclidian Distance (landsat) : Best average*)

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	14	0	**1**	0	14	0	1	0	0	**31**
Woodland	0	0	0	**0**	0	0	0	0	0	**0**
EGF	35	2	1	0	**35**	2	1	0	0	**41**
Mixed rangeland	0	11	0	1	0	**11**	0	1	0	**12**
Sub_farmland	0	2	5	0	0	2	**5**	0	0	**7**
Tea_plantation	1	0	6	7	1	0	6	7	0	**14**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	31	16	84.21%	51.61%	0.4092
Woodland	0	0	0	---	---	0
EGF	50	41	35	70.00%	85.37%	0.7206
Mixed rangeland	15	12	11	73.33%	91.67%	0.9028
Sub_farmlands	13	7	5	38.46%	71.43%	0.6739
Tea_plantation	8	14	7	87.50%	50.00%	0.4588
Shrubs &brushland	0	0	0	---	---	0
Totals	105	105	74			

Overall Classification Accuracy =	70.48
Overall Kappa Statistics =	0.5933

ERROR MATRIX _12 (*Euclidian Distance (landsat) : Best minimum*)

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**16**	0	13	0	1	0	0	**30**
Woodland	0	0	0	**0**	0	0	0	0	0	**0**
EGF	0	0	3	0	**36**	2	2	0	0	**43**
Mixed rangeland	0	0	0	0	0	**10**	0	1	0	**11**
Sub_farmland	0	0	0	0	0	3	**6**	0	0	**9**
Tea_plantation	0	0	0	0	1	0	4	**7**	0	**12**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	30	16	84.21%	53.33%	0.4302
Woodland	0	0	0	---	---	0
EGF	50	43	36	72.00%	83.72%	0.6892
Mixed rangeland	15	11	10	66.67%	90.91%	0.8939
Sub_farmlands	13	9	6	46.15%	66.67%	0.6196
Tea_plantation	8	12	7	87.50%	58.33%	0.549
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**75**			

Overall Classification Accuracy =	71.43
Overall Kappa Statistics =	0.6026

ERROR MATRIX _13 (*Euclidian Distance (landsat) : Transformed divergence*)

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**16**	0	12	0	0	0	0	**28**
Woodland	0	0	0	**0**	1	4	2	0	0	**7**
EGF	0	0	3	0	**36**	1	1	0	0	**41**
Mixed rangeland	0	0	0	0	0	**8**	0	1	0	**9**
Sub_farmland	0	0	0	0	0	2	**7**	1	0	**10**
Tea_plantation	0	0	0	0	1	0	3	**6**	0	**10**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	-----
Clouds	0	0	0	---	---	0
DEGF	19	28	16	84.21%	57.14%	0
Woodland	0	7	0	---	---	0.4767
EGF	50	41	36	72.00%	87.80%	0
Mixed rangeland	15	9	8	53.33%	88.89%	0.7672
Sub_farmlands	13	10	7	53.85%	70.00%	0.8704
Tea_plantation	8	10	6	75.00%	60.00%	0.6576
Shrubs &brushland	0	0	0	---	---	0.567
Totals	**105**	**105**	**73**			**0**

Overall Classification Accuracy =	69.52
Overall Kappa Statistics =	0.5851

ERROR MATRIX _12_*(Euclidian Distance (SPOT) : Best Average)*

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	1	0	0	0	0	**1**
DEGF	0	0	**12**	0	24	0	1	0	0	**37**
Woodland	0	0	0	**0**	0	0	0	0	0	**0**
EGF	0	0	6	0	**21**	4	4	3	0	**38**
Mixed rangeland	0	0	1	0	1	**8**	0	2	0	**12**
Sub_farmland	0	0	0	0	2	3	**4**	3	0	**12**
Tea_plantation	0	0	0	0	1	0	3	**0**	0	**4**
Shrubs &brushland	0	0	0	0	0	0	1	0	**0**	**1**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	1	0	---	---	0
DEGF	19	37	12	63.16%	32.43%	0.175
Woodland	0	0	0	---	---	0
EGF	50	38	21	42.00%	55.26%	0.1459
Mixed rangeland	15	12	8	53.33%	66.67%	0.6111
Sub_farmlands	13	12	4	30.77%	33.33%	0.2391
Tea_plantation	8	4	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	1	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy =	**42.86**
Overall Kappa Statistics =	**0.2178**

ERROR MATRIX _12_*(Euclidian Distance (SPOT) : Best minimum)*

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	2	0	0	0	0	**2**
DEGF	0	0	**12**	0	15	0	0	0	0	**27**
Woodland	0	0	0	**0**	1	0	1	0	0	**2**
EGF	0	0	7	0	**29**	3	2	0	0	**41**
Mixed rangeland	0	0	0	0	0	**8**	1	2	0	**11**
Sub_farmland	0	0	0	0	2	**4**	**6**	6	0	**18**
Tea_plantation	0	0	0	0	1	0	3	**0**	0	**4**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	2	0	---	---	0
DEGF	19	27	12	63.16%	44.44%	0.3217
Woodland	0	2	0	---	---	0
EGF	50	41	29	58.00%	70.73%	0.4412
Mixed rangeland	15	11	8	53.33%	72.73%	0.6818
Sub_farmlands	13	18	6	46.15%	33.33%	0.2391
Tea_plantation	8	4	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy = 52.38
Overall Kappa Statistics = 0.3463

ERROR MATRIX _13_*(Transformed divergence (SPOT))*

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	1	0	0	0	0	**1**
DEGF	0	0	**11**	0	16	0	0	0	0	**27**
Woodland	0	0	0	**0**	2	0	1	0	0	**3**
EGF	0	0	7	0	**27**	3	2	0	0	**39**
Mixed rangeland	0	0	1	0	1	**7**	0	1	0	**10**
Sub_farmland	0	0	0	0	2	5	**6**	7	0	**20**
Tea_plantation	0	0	0	0	1	0	4	**0**	0	**5**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	1	0	---	---	0
DEGF	19	27	11	57.89%	40.74%	0.2765
Woodland	0	3	0	---	---	0
EGF	50	39	27	54.00%	69.23%	0.4126
Mixed rangeland	15	10	7	46.67%	70.00%	0.65
Sub_farmlands	13	20	6	46.15%	30.00%	0.2011
Tea_plantation	8	5	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy =	48.57
Overall Kappa Statistics =	0.3010

ERROR MATRIX _Correlation:landsat_maximum likelihood

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**16**	0	10	0	0	0	0	**26**
Woodland	0	0	1	**0**	0	2	0	0	0	**3**
EGF	0	0	2	0	**39**	1	2	0	0	**44**
Mixed rangeland	0	0	0	0	0	**9**	0	1	0	**10**
Sub_farmland	0	0	0	0	0	3	**6**	0	0	**9**
Tea_plantation	0	0	0	0	1	0	5	**7**	0	**13**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	26	16	84.21%	61.54%	0.5304
Woodland	0	3	0	---	---	0
EGF	50	44	39	78.00%	88.64%	0.7831
Mixed rangeland	15	10	9	60.00%	90.00%	0.8833
Sub_farmlands	13	9	6	46.15%	66.67%	0.6196
Tea_plantation	8	13	7	87.50%	53.85%	0.5004
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**77**			

Overall Classification Accuracy = 73.33
Overall Kappa Statistics = 0.6307

ERROR MATRIX _Correlation:SPOT_maximum likelihood

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**13**	0	16	0	0	0	0	**29**
Woodland	0	0	0	**0**	0	0	1	0	0	**1**
EGF	0	0	6	0	**31**	3	1	0	0	**41**
Mixed rangeland	0	0	0	0	0	**9**	1	3	0	**13**
Sub_farmland	0	0	0	0	2	3	**4**	5	0	**14**
Tea_plantation	0	0	0	0	1	0	4	**0**	0	**5**
Shrubs &brushland	0	0	0	0	0	0	2	0	**0**	**2**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	29	13	68.42%	44.83%	0.3264
Woodland	0	1	0	---	---	0
EGF	50	41	31	62.00%	75.61%	0.5344
Mixed rangeland	15	13	9	60.00%	69.23%	0.641
Sub_farmlands	13	14	4	30.77%	28.57%	0.1848
Tea_plantation	8	5	0	0.00%	0.00%	-0.0825
Shrubs &brushland	0	2	0	---	---	0
Totals	**105**	**105**	**115**			

Overall Classification Accuracy = 54.29
Overall Kappa Statistics = 0.3706

ERROR MATRIX _landsat _maximum likelihood and parallelepiped

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**16**	0	13	0	0	0	0	**29**
Woodland	0	0	0	**0**	0	0	0	0	0	**0**
EGF	0	0	3	0	**36**	1	3	0	0	**43**
Mixed rangeland	0	0	0	0	0	**12**	1	1	0	**14**
Sub_farmland	0	0	0	0	0	2	**4**	0	0	**6**
Tea_plantation	0	0	0	0	1	0	5	**7**	0	**13**
Shrubs &brushland	0	0	0	0	0	0	0	0	**0**	**0**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0
Clouds	0	0	0	---	---	0
DEGF	19	29	16	84.21%	55.17%	0.4527
Woodland	0	0	0	---	---	0
EGF	50	43	36	72.00%	83.72%	0.6892
Mixed rangeland	15	14	12	80.00%	85.71%	0.8333
Sub_farmlands	13	6	4	30.77%	66.67%	0.6196
Tea_plantation	8	13	7	87.50%	53.85%	0.5004
Shrubs &brushland	0	0	0	---	---	0
Totals	**105**	**105**	**75**			

Overall Classification Accuracy =	71.43
Overall Kappa Statistics =	0.6029

ERROR MATRIX _Correlation:SPOT_maximum likelihood and parallelepiped

Classification	Unclassified	Clouds	DEGF	Woodland	EGF	Mixed rangeland	Sub_farmland	Tea_plantation	Shrubs &brushland	Row Total
Unclassified	**0**	0	0	0	0	0	0	0	0	**0**
Clouds	0	**0**	0	0	0	0	0	0	0	**0**
DEGF	0	0	**12**	0	16	0	0	0	0	**28**
Woodland	0	0	0	**0**	0	0	1	0	0	**1**
EGF	0	0	7	0	**31**	3	1	0	0	**42**
Mixed rangeland	0	0	0	0	0	**9**	2	2	0	**13**
Sub_farmland	0	0	0	0	2	3	**3**	5	0	**13**
Tea_plantation	0	0	0	0	1	0	5	**1**	0	**7**
Shrubs &brushland	0	0	0	0	0	0	1	0	**0**	**1**
Column Total	**0**	**0**	**19**	**0**	**50**	**15**	**13**	**8**	**0**	**105**

ACCURACY TOTALS

Class Names	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy	Kappa statistic
Unclassified	0	0	0	---	---	0.0000
Clouds	0	0	0	---	---	0.0000
DEGF	19	28	12	63.16%	42.86%	0.3023
Woodland	0	1	0	---	---	0.0000
EGF	50	42	31	62.00%	73.81%	0.5000
Mixed rangeland	15	13	9	60.00%	69.23%	0.6410
Sub_farmlands	13	13	3	23.08%	23.08%	0.1221
Tea_plantation	8	7	1	12.50%	14.29%	0.0722
Shrubs &brushland	0	1	0	---	---	0.0000
Totals	**105**	**105**	**116**			

Overall Classification Accuracy = 53.33
Overall Kappa Statistics = 0.3706